CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

PHD THESIS

Optical fiber networks are the backbone of modern digital infrastructure. They carry vast volumes of sensitive information, including financial transactions, medical data, cloud services, and personal communications, across regions and continents at high reliability and extremely low latency. Yet, despite their sophistication, these fibers are physically vulnerable. A careless dig during construction can cut the fiber and result in service outages, while a deliberate bending of the fiber can lead to unauthorized access to carried data. Traditional monitoring tools rarely detect the related subtle physical-layer disturbances and typically react once the damage is already done.

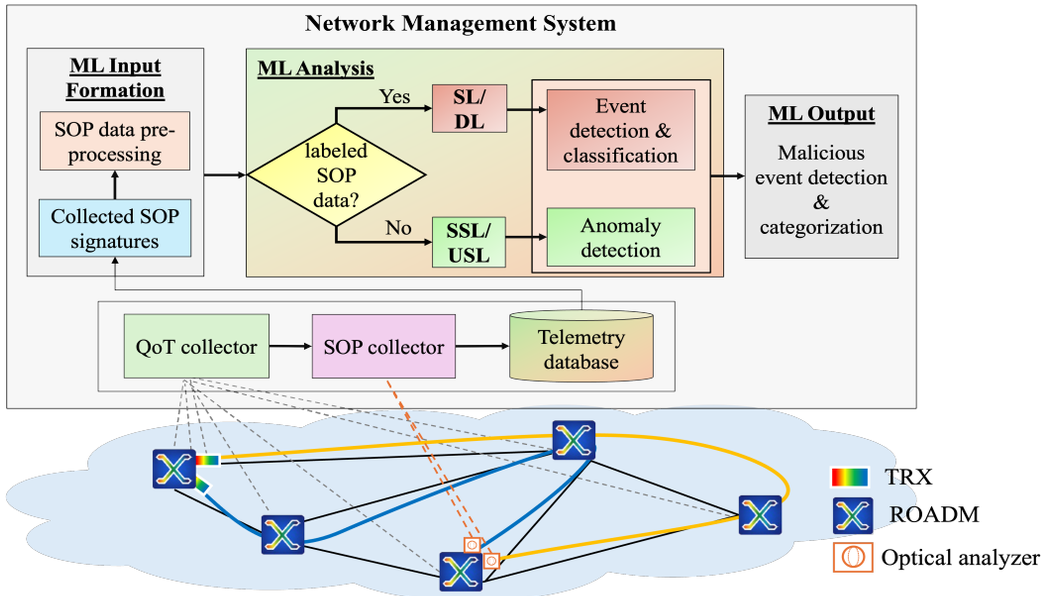
This thesis proposes a new approach: using the fiber itself as a security sensor. As light propagates along the fiber, its State of Polarization (SOP) continually changes in response to environmental conditions. Even low vibrations, temperature variations, or intentional tampering imprint distinctive signatures on the SOP. By continuously monitoring and analyzing these signatures, the network can effectively “sense” physical activity and perturbations along the fiber.

However, interpreting this complex, noisy, and high-dimensional data requires sophisticated analysis. This research applies advanced Machine Learning (ML) and Deep Learning (DL) models to distinguish the characteristic “SOP signatures" of different events—separating benign environmental noise, such as traffic-induced vibrations, from malicious actions like fiber-bending threats. The framework is validated in controlled laboratory environments and on real metropolitan network infrastructure, demonstrating that threats can be detected without disrupting live traffic or deploying specialized sensing hardware.

The innovations in this thesis enable the transformation of passive optical fibers into active, self-monitoring assets. They point toward a future where optical networks do more than carry information; they protect themselves, enabling resilient, autonomous, and secure infrastructure for the next generation of communication networks.

CHALMERS
UNIVERSITY OF TECHNOLOGY

LEYLA SADIGHI • Detection of Optical Network Breaches through ML-Based State of Polarization Analysis • 2026



Detection of Optical Network Breaches through ML-Based State of Polarization Analysis

LEYLA SADIGHI

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, 2026
www.chalmers.se

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Detection of Optical Network Breaches through ML-Based State of Polarization Analysis

Electrical Engineering department, Chalmers University of Technology

Leyla Sadighi

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden, 2026

Detection of Optical Network Breaches through ML-Based State of Polarization Analysis

Electrical Engineering department, Chalmers University of Technology

Leyla Sadighi

ISBN 978-91-8103-388-5

Acknowledgements, dedications, and similar personal statements in this thesis, reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5845
ISSN 0346-718X

Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 1000

Printed by Chalmers Digital Printing
Gothenburg, Sweden, March 2026

Abstract

Optical fiber networks form the backbone of modern communications, yet they are vulnerable to physical-layer disturbances ranging from benign environmental vibrations to malicious threats like fiber tapping. This dissertation addresses the urgent need for advanced monitoring solutions of environmental disturbances by leveraging the State of Polarization (SOP) of light as a sensitive, non-intrusive indicator of fiber events. We develop a Machine Learning (ML)–based framework that continuously analyzes SOP variations to detect and classify physical-layer anomalies. Our approach encompasses Supervised Learning (SL) for classification of known events, including Deep Learning (DL) architectures that automatically extract complex polarization features in challenging real-world conditions, as well as Semi-supervised Learning (SSL) and Unsupervised Learning (USL) techniques for detection of novel anomalies without reliance on fully labeled data.

In controlled laboratory settings, the proposed methods distinguished mechanical vibrations, eavesdropping-induced fiber bends, and other perturbations with high accuracy (exceeding 97% in multi-class classification). Field trials on live, metro-scale fibers further demonstrated robust performance, detecting intrusion attempts and accidental disturbances with minimal degradation in performance, despite real-world noise. Notably, this work provides the first validation that polarization-based sensing remains effective in standard coherent communication systems: ML models accurately detected disturbances on Dual-Polarization 16-Quadrature Amplitude Modulation (DP-16QAM) data-carrying channels with accuracy comparable to that obtained on unmodulated Continuous Wave (CW) probes.

Results confirm that ML-driven SOP-based analysis can rapidly flag fiber taps and other physical intrusions, and distinguish harmful events from harmless fluctuations with high confidence. By validating the proposed intelligent SOP-based monitoring framework over diverse real-world conditions, including different fiber types, network configurations, and signal modalities, this work demonstrates that polarization-based fiber monitoring is practically viable for deployment in real-world operational optical networks. The findings establish SOP analytics as a powerful and non-intrusive tool for enhancing the security and resilience of modern optical communication infrastructure without disrupting normal traffic.

Keywords: Optical Networks; State of Polarization (SOP); SOP Signatures; Fiber Monitoring; Polarization Sensing; Physical Layer Tampering; Machine Learning (ML); Supervised Learning (SL); Semi-Supervised Learning (SSL); Unsupervised Learning (USL); Deep Learning (DL)

To my family.

List of Publications

This thesis is based on the following publications:

[A] **Leyla Sadighi**, Stefan Karlsson, Carlos Natalino, Marija Furdek, "Machine Learning-Based Polarization Signature Analysis for Detection and Categorization of Eavesdropping and Harmful Events". Optical Fiber Communications Conference and Exhibition (OFC), May, 2024.

[B] **Leyla Sadighi**, Stefan Karlsson, Lena Wosinska, Marija Furdek, "Machine Learning Analysis of Polarization Signatures for Distinguishing Harmful from Non-harmful Fiber Events". 24th International Conference on Transparent Optical Networks (ICTON), July, 2024.

[C] **Leyla Sadighi**, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek, "Detection and Classification of Eavesdropping and Mechanical Vibrations in Fiber Optical Networks by Analyzing Polarization Signatures Over a Noisy Environment". Presented in 50th European Conference on Optical Communication (ECOC), September, 2024.

[D] **Leyla Sadighi**, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek, "Deep Learning for Detection of Harmful Events in Real-World, Noisy Optical Fiber Deployments". IEEE Journal of Lightwave Technology (JLT), February, 2025.

[E] **Leyla Sadighi**, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek, "AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges". Optical Fiber Communication Conference (OFC) 2025, Technical Digest Series.

[F] **Leyla Sadighi**, Stefan Karlsson, Carlos Natalino, Marija Furdek, "ML-based State of Polarization Analysis to Detect Emerging Threats to Optical Fiber Security". IEEE Transactions on Network and Service Management (TNSM), September, 2025.

[G] **Leyla Sadighi**, Carlos Natalino, Stefan Karlsson, Marco Ruffini, Eoin Kenny, Lena Wosinska, Marija Furdek, "Generalizability of ML-Based Classification

of State of Polarization Signatures Across Different Bands and Links". 51st European Conference on Optical Communication (ECOC), September, 2025.

[H] **Leyla Sadighi**, Stefan Karlsson, Marco Ruffini, Marija Furdek, "ML-Based Detection and Categorization of Complex Mechanical Vibrations via State of Polarization Analysis in Optical Networks". 25th International Conference on Transparent Optical Networks (ICTON), July, 2025.

[I] **Leyla Sadighi**, Carlos Natalino, Stefan Karlsson, Lena Wosinska, Eoin Kenny, Venkata Virajit Garbhapu, Marco Ruffini, Marija Furdek, "DP-16QAM Modulated vs. Unmodulated Polarization Signatures for Machine Learning-Based Fiber Sensing". IEEE Journal of Lightwave Technology (JLT), February 2026.

Other publications by the author, not included in this thesis, are:

[J] **L. Sadighi**, S. Karlsson, C. Natalino, M. Eshghie, F. Usmani, E. Kenny, L. Wosinska, P. Monti, M. Furdek, M. Ruffini, "Variational Autoencoder Domain Adaptation for Cross-System Generalization in ML-Based SOP Monitoring". Submitted to *Journal of Optical Communications and Networking (JOCN)*, March 2026.

[K] **L. Sadighi**, S. Karlsson, M. Furdek, M. Ruffini, "SOP and Phase Sensitivity for ML-Based Security and Acoustic Fiber Sensing over Bare and Patch Cables". Submitted to *European Conference on Optical Communication (ECOC)*, March 2026.

[L] **L. Sadighi**, A. Knapińska, C. Natalino, S. Karlsson, M. Ruffini, M. Furdek, "AI for Security Management in Optical Networks". *Artificial Intelligence in Optical Networks and Systems*, *Book Chapter*, eds. F. Musumeci, Q. Zhuge, D. Mello, Springer, Chapter 8, to appear (2026).

Contents

Acknowledgments

This PhD journey has been a demanding yet deeply fulfilling chapter of my life, shaped not only by academic challenges but also by the encouragement and kindness of those who supported me throughout the process.

First of all, I would like to express my deepest gratitude to my main supervisor, Associate Professor Marija Furdek Prekratić, for her unwavering support, insightful guidance, and constant encouragement throughout this journey. Her belief in my abilities consistently motivated me to push beyond my limits, while her patience and trust played a crucial role in my development as an independent researcher. Your leadership, dedication, and enthusiasm have been a continual source of inspiration and have profoundly contributed to both my academic progress and personal growth.

I would like to sincerely thank my co-supervisor, Professor Marco Ruffini at Trinity College Dublin, for his generous support, mentorship, and for providing me the opportunity to carry out one year of my doctoral research at the CONNECT Centre. Your guidance and the welcoming research environment he fostered made this experience both productive and memorable.

I am deeply grateful to Stefan Karlsson from Micropol Fiber Optics AB for the many insightful discussions, for sharing his expertise, and for his crucial contributions to the experimental aspects of this work. Your practical insights and technical knowledge have been instrumental in shaping several parts of this thesis.

I am sincerely grateful to my co-supervisor at Chalmers, Dr. Carlos Natalino da Silva, for the many insightful discussions, feedback, and new perspectives. Your dedication, clarity of thought, and willingness to engage with my ideas have had a profound impact on this work. My heartfelt thanks also go to Prof. Lena Wosinska and Prof. Paolo Monti for their continuous support and academic guidance. Your contributions have shaped my research path and strengthened my understanding of the broader field.

To all of my colleagues, especially within the Optical Networks group at Chalmers, as well as my colleagues at the CONNECT Centre, thank you for the stimulating discussions, collaborative spirit, and generous sharing of knowledge. Working alongside such talented and supportive people has been both a privilege and joy.

On a personal note, I wish to express my deepest gratitude to my late father and my late brother, whose dedication, sacrifices, and constant belief

in me laid the foundation for everything I have achieved. Though you are no longer here, your strength and values continue to guide me every day. My profound thanks also go to my mother, whose unconditional love, patience, and resilience have supported me through every step of this journey. Your encouragement has been a constant source of strength.

Finally, to my loved ones, to my brothers and sisters, and all those who have supported, encouraged, and believed in me—thank you. This accomplishment would not have been possible without your presence, kindness, and understanding.

Acronyms

1D One-Dimension

16QAM 16-Quadrature Amplitude Modulation

acc Accuracy

AI Artificial Intelligence

ANN Artificial Neural Networks

ARS Adjusted Rand Score

BER Bit Error Rate

CNN Convolutional Neural Network

CW Continuous Wave

CW-DFB Continuous Wave Distributed Feedback

CWDM Coarse Wavelength Division Multiplexing

COTDR Coherent Optical Time-Domain Reflectometry

CD Chromatic Dispersion

DAS Distributed Acoustic Sensing

DVS Distributed Vibration Sensing

DFB Distributed Feedback

DBSCAN Density-Based Spatial Clustering of Applications with Noise

DCU Dublin City University

DFOS Distributed Fiber Optic Sensing

DL Deep Learning

DNN Deep Neural Networks

DoP Degree of Polarization

DP Dual-Polarization

DP-16QAM Dual-Polarization 16-Quadrature Amplitude Modulation

DSP Digital Signal Processing

DT Decision Tree

DWDM Dense Wavelength Division Multiplexing

DP-QPSK Dual-Polarization Quadrature Phase-Shift Keying

DP-QAM Dual-Polarization Quadrature Amplitude Modulation

ECL External Cavity Laser

ET Extra Trees

FN False Negatives

FNR False Negative Rate

FOCS Fiber Optical Tactical Cable System

FP False Positives

FPR False Positive Rate

FFT Fast Fourier Transform

FTTH Fiber to the Home

GB Gradient Boosting

GAN Generative Adversarial Network

XGBoost eXtreme Gradient Boosting

HEAnet ASIERA (formerly HEAnet), Ireland's National Education and Research Network

HGB Histogram Gradient Boosting

SVM Support Vector Machine

IDS Intrusion Detection System

IF Isolation Forest

IM-DD Intensity-Modulated Direct-Detection

KNN K-Nearest Neighbors

LDA Linear Discriminant Analysis

LR Logistic Regression

LSTM Long Short-Term Memory

MLP Multi Layer Perceptron

ML Machine Learning

NMS Network Management System

NRZ Non-Return-to-Zero

NPSV Numerical Polarization State Variation

NN Neural Network

OCSVM One-Class Support Vector Machine

OPM Optical Performance Monitoring

OTDR Optical Time Domain Reflectometry

OSaaS Optical Spectrum as a Service

OSNR Optical Signal to Noise Ratio

OSI Open Systems Interconnection

PM Polarization Maintaining

PM-QAM Polarization-Multiplexed Quadrature Amplitude Modulation

PON Passive Optical Network

PSV Polarization State Variation

Φ-**OTDR** Phase-Sensitive Optical Time-Domain Reflectometry

PMD Polarization Mode Dispersion

PDL Polarization-Dependent Loss

QAM Quadrature Amplitude Modulation

QoT Quality of Transmission

Q-DAS Quantitative DAS

QPSK Quadrature Phase-Shift Keying

RBF Radial Basis Function

RF Random Forest

ROC Receiver Operating Characteristic

ROADM Reconfigurable Optical Add-Drop Multiplexer

RPM Revolutions Per Minute

SDN Software-Defined Networking

SL Supervised Learning

SM Single Mode

SNR Signal-to-Noise Ratio

SOP State of Polarization

SOP-PSDG SOP Power Spectral Density Gap

SOPAS SOP Angular Speed

SS Silhouette Score

SSL Semi-Supervised Learning

TCD Trinity College Dublin

TN True Negatives

TNR True Negative Rate

TP True Positives

TPR True Positive Rate

TRX Transceiver

USL Unsupervised Learning

WDM Wavelength Division Multiplexing

ASE Amplified Spontaneous Emission

Part I

Overview

CHAPTER 1

Introduction

1.1 Problem statement

Over the past decade, the amount of data carried by global communication networks has been increasing at an extraordinary pace. New, bandwidth-hungry applications with diverse temporal profiles and stringent performance requirements place unprecedented strain on modern optical transport networks. These networks span long-haul, regional metro, and access segments, connecting data centers, cloud platforms, enterprises, and end users, thereby forming the backbone of digital society and supporting essential services such as national defense, healthcare, financial systems, and everyday Internet connectivity [1], [2].

The critical importance of optical networks makes their resilience and security more concerning than ever. As reliance on fiber optic infrastructures deepens, the requirements on bandwidth availability and service reliability intensify, particularly for mission-critical applications. At the same time, the physical layer of optical networks remains inherently exposed to various forms of accidental and intentional disturbances. Recent reports document an increase in incidents targeting optical fiber infrastructure, ranging from acci-

dental mechanical disturbances to deliberate acts of tampering and sabotage [2], [3], [4].

The physical-layer vulnerabilities of optical networks expose them to two main types of threats, categorized according to their goal: service disruption and eavesdropping [2], [5]. Service disruption threats are aimed at disrupting the normal functionality of the system. They include accidental disturbances such as construction works which can lead to fiber cuts, and intentional actions such as jamming, polarization scrambling, or component disabling, all of which can degrade or interrupt network services [6], [7]. Fibers are also susceptible to eavesdropping, which exploits the physical properties of the medium to illegally intercept communication [8]. Advanced tapping techniques, including evanescent coupling, V-groove cut, and micro-/macro-bending, enable an attacker to tap a part of optical signal while causing negligible power loss and avoiding obvious service outages, thereby evading alarms based on power-level thresholds in conventional monitoring systems [9], [10], [11], [12]. Such stealthy intrusions pose a serious risk to the confidentiality and integrity of data transmission, as sensitive information can be intercepted while network operators remain unaware of the breach. Breaches at the physical layer can have rippling effects on higher-layer services, potentially impacting thousands of users and applications simultaneously and leading to widespread service disruption, data loss, and severe financial or societal impacts. This underscores a strong need of enhancing security and resilience of optical networks against such disturbances through advanced monitoring, detection, and protection mechanisms.

One of the key components of optical network security management is the timely detection of physical-layer disturbances or tapping attempts. Early detection of anomalies associated with external events allows network operators to respond promptly, preventing minor issues from escalating into major failures or data leaks. Techniques such as Optical Time Domain Reflectometry (OTDR) are commonly used to identify severe faults, like fiber cuts or sharp bends, by analyzing Rayleigh backscattering [13], [14], [15]. They offer relatively precise fault localization and are well established in field deployments. Nevertheless, their applicability is hindered by high deployment costs and limited scalability [16], and their sensitivity [17] is insufficient for detecting subtle disruptions such as low-amplitude vibrations or small mechanical perturbations. Other approaches, including Distributed Fiber Optic Sens-

ing (DFOS) and Distributed Acoustic Sensing (DAS) as a prominent DFOS modality for intrusion detection [18], offer enhanced sensitivity but depend on dedicated hardware and sophisticated signal processing, which substantially increases complexity and cost, thereby restricting widespread adoption.

These limitations create a strong need for a fiber sensing solution that is sensitive to subtle physical disturbances, operates non-intrusively on fibers that carry live network traffic, and is cost-effective for large-scale deployments. One promising approach is to use the SOP of light as a marker for sensing such perturbations [19]. Instead of injecting test pulses or requiring dedicated sensing equipment, SOP-based monitoring uses the signal's polarization behavior as the sensing mechanism. As light propagates along an optical fiber link, external perturbations such as mechanical stress, vibrations, bending, and temperature fluctuations alter the birefringence of the medium, thereby inducing variations in the SOP [20]. Different physical events have characteristic impact on the SOP of light, each inducing a distinct pattern of polarization variations. These variations trace unique trajectories on the Poincaré sphere, a standard geometric representation in which every possible polarization state maps to a single point on the surface of a unit sphere, forming what we refer to as an *SOP signature* that can be used to identify the underlying event. SOP monitoring offers a promising balance of sensitivity and simplicity for physical-layer security: it can detect subtle fiber perturbations that other methods might miss, using equipment already present in modern optical systems such as coherent receivers.

The task remains to analyze the rich polarization data to automatically flag intrusions and the type of disturbance. While the idea of using SOP variations for event detection is powerful, the raw collected SOP data has high dimensionality and is inherently noisy. Normal fiber operation causes constant polarization drift due to temperature changes, low vibrations, or other causes, and disturbances can imprint subtle patterns on top of this drift. Traditional monitoring based on fixed thresholds or simple rules is not suited to capture the resulting complex behavior [21], [22], [23]. Machine Learning (ML) becomes essential for this task, as ML models can analyze intricate, high-dimensional SOP trajectories and learn the subtle, non-linear relationships that distinguish benign fluctuations from meaningful disturbances. ML algorithms can be trained to recognize the SOP signatures generated by different physical events, directly learning decision boundaries from data rather than

relying on fixed decision rules or manually set thresholds. Once trained, these models can automatically detect and classify disturbances, and can improve over time as more diverse and informative training data becomes available. These capabilities motivate the development of an ML-driven framework for SOP-based fiber monitoring. By embedding ML within this SOP-based monitoring process, the system should autonomously detect security breaches and physical anomalies, providing early warning and enhancing the resilience of optical networks against both accidental and intentional threats. Moreover, because optical networks operate across different spectral bands and network topologies, with varying fiber links and cable types, and carry both modulated and unmodulated traffic, the framework must remain effective under these conditions to be practically viable. Real-world deployments also introduce continuous background noise from environmental factors such as traffic-induced vibrations and temperature fluctuations, complex and overlapping mechanical disturbances that occur simultaneously on the fiber. A practically viable framework must therefore not only achieve high classification accuracy under controlled conditions, but also maintain robust performance over these diverse and challenging operational environments.

The core problem can be stated as follows: *How can we leverage SOP variations, combined with advanced ML techniques, to detect and classify a wide range of fiber perturbations, including both accidental and malicious events?* To address it, this work employs a comprehensive set of ML techniques, including SL to identify known types of fiber perturbations, SSL and USL to detect anomalies with minimal reliance on labeled data and without prior knowledge of threats, as well as DL to handle complex real-world data representations. The novelty of this work lies in developing an SOP-based monitoring framework that integrates these different learning paradigms, SL, SSL, USL, and DL, to achieve intelligent, adaptive, and accurate detection of physical-layer disturbances across diverse optical fiber types, spectral bands, and signal modalities.

To pursue this research direction in a systematic way, the work is guided by the following key Research Questions (RQs).

1.2 Research Questions

- RQ1: How can SOP data be used to reliably detect fiber tampering or tapping events, distinguishing them from normal environmental fluctuations?
- RQ2: Upon detecting an anomaly in the SOP, how can the type of perturbation be classified to determine the origin or cause of the event using ML (e.g., differentiating a malicious fiber tap from accidental mechanical disturbances)?
- RQ3: Which ML models and algorithms are most effective and robust for analyzing SOP variation data under realistic noise conditions?
- RQ4: Does an ML model trained on event signatures collected under controlled conditions maintain its detection and classification accuracy when the same events occur dynamically during live monitoring of field-deployed fiber?
- RQ5: How can emerging or previously unseen events be detected when labeled training data is unavailable or scarce?
- RQ6: How does signal modulation affect ML-driven polarization-based sensing?
- RQ7: How do SOP-based ML classifiers trained on signals using one spectral band and fiber link generalize across different spectral bands and fiber links?
- RQ8: How can SOP-based ML models detect and distinguish complex or overlapping mechanical vibration patterns that occur simultaneously on the fiber?

By addressing these research questions, this thesis tackles some of the central scientific and engineering challenges in developing an intelligent polarization-based fiber sensing system. The questions progress from foundational issues pertinent to understanding the relationships between different SOP events, over algorithmic aspects related to evaluating ML approaches, to performance issues relevant for real-world applicability. By covering such broad spectrum of challenges, our goal is to provide a systematic path to design, implement, and evaluate a monitoring framework that strengthens optical network security.

1.3 Thesis Goals

The goal of this thesis is to design and validate an ML-driven optical network monitoring framework that relies on SOP variations to detect fiber tampering and disturbances. Achieving this broad goal involves several concrete objectives:

1. **Develop a Polarization Signature Analysis Method:** Establish a methodology for capturing and quantifying SOP variations caused by different events. This includes defining how optical signal polarization changes can be processed into meaningful SOP event signatures. The approach pursued in this work is the computation of normalized Numerical Polarization State Variation (NPSV) over time, combined with frequency-domain features (via Fast Fourier Transform (FFT)) to characterize disturbances. Such representations are essential as inputs to ML models.

2. **Apply and Compare ML Techniques and Assess Event Severity:** Investigate a range of ML approaches, including SL for event classification, SSL and USL for anomaly detection, and DL for capturing complex, real-world polarization dynamics, to interpret polarization signatures. The objective is to determine which techniques are best suited to different aspects of the problem. For instance, SL models can classify known event types (e.g., bending or vibrations at specific frequencies), while SSL and USL methods can detect novel events.

   Beyond classifying event types, an essential objective is to determine the operational relevance of detected disturbances. In practice, not every disturbance warrants an alarm. Benign perturbations (e.g., routine fiber bends) should ideally be ignored, while harmful events (e.g., tapping or excavation-induced vibrations) must be flagged. A specific goal is to enhance the system's ability to categorize event severity by defining classification schemes that group events into harmful and non-harmful categories, thus reducing false positives and focusing attention on genuine threats, while not missing events that are relevant for logging and future analysis.

   The overarching goal is to develop models that not only achieve high detection accuracy with consistently low false alarm rates but also re-

liably distinguish harmful events from benign disturbances to support practical, actionable network monitoring.

3. **Address Realistic Network Conditions:** This includes testing the framework on live fiber links to ensure robustness against practical issues such as attenuation over longer distances, environmental fluctuations, ambient noise from surrounding street traffic and infrastructure, and other sources of polarization disturbance inherent to field-deployed fibers. The evaluation also considers modulated optical signals that carry network traffic, not just unmodulated probes, thereby ensuring applicability to modern coherent optical systems. Experiments are designed with both modulated and unmodulated signals under identical disturbances to assess the impact on detection. In addition, the framework is assessed under complex and overlapping mechanical vibrations to reflect realistic multi-event vibration scenarios often encountered in operational networks.

1.4 Thesis Scope

To help guide the reader and delimit the focus of the work, the scope of this thesis is outlined below:

- This thesis focuses exclusively on physical-layer monitoring of optical fibers using polarization effects. Higher-layer security mechanisms (e.g., Intrusion Detection System (IDS) or encryption) are assumed to be in place, but they address different attack vectors and cannot detect physical-layer tampering. The proposed SOP-based approach complements them by detecting disturbances and tapping attempts that remain invisible to higher-layer mechanisms, even when traffic is fully encrypted.
- The focus is on the detection and classification of disturbances. The phenomena of interest are relatively fast polarization perturbations (at the order of milliseconds), such as those caused by bends, vibrations, or fiber taps. Slow polarization drifts (e.g., temperature-induced changes) are regarded as background noise rather than primary targets.
- Experiments cover standard Single Mode (SM) fibers commonly deployed in telecom networks: bare, indoor, Fiber Optical Tactical Cable

System (FOCS), and fibers deployed in metropolitan network segments. Multimode, hollow core, and specialty fibers are not included.

- This thesis focuses on the detection and classification of disturbances based on SOP variations and does not address the precise localization of events along the fiber. Determining the exact spatial position of a disturbance, especially on long-haul links, lies outside the scope of this work and is identified as a direction for future research.

By defining these goals and boundaries, the thesis ensures a focused investigation of the most critical aspects needed to demonstrate the viability of an SOP-based ML monitoring system. Achieving the specified objectives will help determine whether polarization monitoring can transition from a promising research concept to a practical tool for safeguarding optical networks.

1.5 Thesis Methodology and Contributions

The research journey documented in this thesis spans multiple dimensions of the problem space. It begins with fundamental questions about data acquisition and feature engineering, progresses through systematic evaluation of diverse algorithmic approaches, and advances to sophisticated DL architectures capable of handling real-world complexity. To tackle the main problem and achieve the stated research objectives, a structured experimental research approach was adopted, progressing from controlled laboratory settings to field-deployed fiber links. The methodology is built around the collection of diverse SOP datasets, the development and evaluation of ML models, and the validation of the proposed framework over increasingly realistic conditions, including the impact of signal modulation on polarization-based sensing, the generalization of trained ML models for different spectral bands and fiber links, and the detection of complex and overlapping mechanical vibration patterns. The contributions of this thesis, along with an outline of related publications, are summarized in the following subsections.

Data Collection and Processing

To support the investigation of the RQs defined in Section 1.2, a diverse collection of datasets was assembled over both controlled laboratory environments and field-deployed fiber links. Table 1.1 summarizes all datasets, the

number of event classes they contain, the research questions they address, and the ML methods applied to each. The laboratory datasets (LDS1–LDS3) were designed to capture a wide range of polarization signatures under known ground-truth conditions, encompassing individual disturbances, harmful and benign events, and complex overlapping vibration patterns. The real-world datasets (RDS1–RDS4) were obtained from an operational metropolitan fiber network with links of different lengths, enabling performance evaluation under realistic environmental noise, system-level variability, and varying signal modulation conditions. Detailed descriptions of the data collection procedures and experimental setups are provided in Chapter 4.3.

The raw data consisted of time series of SOP variations, which were then processed to extract distinctive polarization signatures for each event. As described in **Paper A**, **Paper B**, and **Paper F**, our approach was to calculate the NPSV, which captures the angular distance between successive samples of the SOP trajectories on the Poincaré sphere. We then segmented the NPSV time series into time windows and performed FFT on each window, producing a time–frequency spectrogram (visualized as the "*waterfall*" plot). Each event type yielded a unique pattern in this spectrogram.

This entire data collection and feature extraction process resulted in a first-of-its-kind dataset of polarization responses for fiber events. We gathered a wide range of distinct event SOP signatures for three different fiber cable types, including various benign and harmful events. These datasets represent a significant contribution by providing the experimental foundation for developing and benchmarking ML models in a domain where such data have been scarce. Although the datasets themselves are not publicly available, the methodology and experimental design presented in this thesis enable their reproduction and provide a basis for future work in SOP-based sensing. This contribution specifically address research question RQ1.

Table 1.1: Overview of collected datasets, associated research questions, and applied ML models. Laboratory datasets (LDS) were acquired under controlled conditions, while real-world datasets (RDS) were collected from deployed fiber links.

Data	**No. classes**	**Addressed RQs**	**Applied ML Models**	**Description and Purposes**
LDS1	13	RQ1, RQ2, RQ5	SL, SSL, USL	Lab; bare, FOCS, indoor cables; event detection; emerging threats
LDS2	5	RQ1, RQ2, RQ4	SL	Lab; indoor cable; live deployment validation
LDS3	14	RQ1, RQ2, RQ8	SL	Lab; indoor cable; broadband complex vibrations detection
RDS1	7	RQ1, RQ2, RQ3	SL	Field; indoor cable; 0.15 and 10.5 km metro links; real-world validation; OpenIreland
RDS2	14	RQ1, RQ2, RQ3	DL	Field; indoor cable; 0.15 and 10.5 km; real-world validation; OpenIreland
RDS3	4 (two sets)	RQ1, RQ2, RQ6	SL	Field; indoor cable; modulated vs. unmodulated signals; HEAnet
RDS4	3 (two sets)	RQ1, RQ2, RQ7	SL	Field; indoor cables; O-band and C-band cross-system; HEAnet and OpenIreland

SL Approaches for Event Classification

Using labeled polarization signature data for each known event scenario in LDS1, we evaluated a suite of supervised ML algorithms to automatically identify and classify different event types. We treated each event type as a class in a multi-class classification problem. To identify the most suitable classification algorithm for each scenario, we conducted a grid search for comprehensive benchmarking of SL classifiers available in the Scikit-learn library [24]. The evaluated methods span a diverse range of learning paradigms, including ensemble-based methods: Random Forest (RF), Extra Trees (ET) classifier, Histogram Gradient Boosting (HGB), eXtreme Gradient Boosting (XGBoost), and Gradient Boosting (GB); kernel-based models: Support Vector Machine (SVM); linear classifiers: Logistic Regression (LR) and Linear Discriminant Analysis (LDA); distance-based techniques: K-Nearest Neighbors (KNN); and tree-based learners: Decision Tree (DT). Through systematic experiments, we identified that ensemble methods performed particularly well on our feature space. In our initial study in **Paper A**, the XGBoost classifier achieved the highest accuracy, successfully distinguishing among the 13 event classes for three cable types with 92.3% accuracy. In **Paper B**, we achieved nearly 98% accuracy in discerning harmful disturbances from non-harmful ones in indoor cable using the HGB classifier.

These findings confirmed the viability of SOP-based event classification and highlighted the promise of ML in alleviating the need for human experts to monitor SOP data. The contribution here addressed research questions RQ1, RQ2, and RQ3 and provided a proof-of-concept supervised ML pipeline for polarization data, demonstrating effectiveness in controlled settings. Additionally, the HGB model from **Paper B** was used to address RQ4 by evaluating whether the HGB classifier trained on LDS2 can reliably detect the same disturbance categories when these events were recreated and measured live on an indoor fiber link.

Assessment on Field-Deployed Fibers Under Operational Noise

The next methodological step was to assess the approach in a real network environment outside the lab, for which we used the OpenIreland [25] testbed in Dublin and collected RDS1. Two fiber spans were used: a short 0.15 km fiber and a longer 10.5 km metro fiber link running beneath city streets. The

fibers were carrying live traffic channels with real background polarization noise from normal network operation. We emulated disturbances on these fibers within allowed experimental windows, such as deliberately bending a patch cord to simulate a tap on the short link and attaching a vibration source on the longer link. The SOP data collected from these trials was then fed into our best-performing ML models (trained on the laboratory dataset) to evaluate performance. We found that the model retained a high detection capability: for example, the HGB classifier achieved 86.5% accuracy in classifying events on the noisy 10.5 km link. This was a promising result under realistic conditions, suggesting that polarization-based intrusion detection and categorization can be applied to deployed fibers.

The results, addressing research questions RQ1 to RQ3 and presented in **Paper C**, demonstrated a slight drop in accuracy compared to lab results (from ∼97% to ∼86%), reflecting the increased complexity of real conditions, but still indicating robust performance.

DL Models for Robust Event Detection in Real-World Fiber Links

Building on the SL results, we investigated whether DL models could further improve detection accuracy under real-world conditions, particularly in handling the increased complexity and noise present in real-world RDS2 data. While traditional SL classifiers rely on manually engineered feature extraction, DL models have the capacity to learn hierarchical representations directly from data, making them inherently better suited to capture the complex and variable polarization patterns introduced by real-world noise, long fiber links, and environmental fluctuations. This motivated the development of a custom One-Dimension (1D)-Convolutional Neural Network (CNN) which takes the polarization time-series or their spectrogram representation as input and learns to classify events.

We trained and tuned this deep neural network using the RDS2 dataset. The DL approach yielded a substantial performance boost: on a field-deployed 10.5 km fiber link experiencing various disturbances, the CNN-based model achieved 92.3% accuracy in identifying harmful events, whereas the best traditional ML model had achieved 86.5% on a similar link. On a shorter 0.15 km link (representative of access networks), the CNN reached an impressive 98.6% accuracy. These results, which address research question RQ3 and are presented in **Paper D**, demonstrate state-of-the-art performance in SOP–based

event detection. Importantly, the DL models proved capable of handling the increased noise and variability associated with real, in-service conditions. This stage contributed with a critical validation of the approach's practicality and offered insights for improvement – notably, it underscored the need for techniques that could handle unlabeled data and more complex scenarios, since obtaining extensive labeled event data in a live network is difficult.

SSL and USL for Detection of Unseen Anomalies

To address the challenge of limited labeled data in large-scale deployments, we explored ML methods that do not require every event type to be labeled. We investigated one-class classification and clustering algorithms on the LDS1 polarization data. We employed a One-Class Support Vector Machine (OCSVM) as an SSL method trained only on normal operation data to learn the profile of normal SOP behavior, and detect deviations from that profile as anomalies. We also applied a USL approach called Density-Based Spatial Clustering of Applications with Noise (DBSCAN) directly on the collected signature features that groups the data into distinct clusters corresponding to different events. These techniques were tested using LDS1 collected in **Paper A**, with class labels removed to simulate a realistic deployment scenario in which the system has no prior knowledge of the event types it may encounter and must detect anomalies solely based on deviations from normal behavior. The findings were very promising: OCSVM was able to detect the majority of anomalies (harmful events) by recognizing they did not fit the "normal" pattern, and DBSCAN clustering showed ability to separate many of the event types into distinct clusters, even without labels. Notably, this approach was able to detect overlapping events (cases where two events happened concurrently) as anomalies.

A key novelty in this work is that it was the first application of SSL/USL to SOP-based fiber monitoring, filling a gap in the literature. It demonstrated that an SOP monitoring system could still function with minimal supervision, which is crucial for real deployments where new or unforeseen events might occur. The main contribution here is an adaptive anomaly detection framework that increases the robustness of the monitoring approach and has been published in **Paper F**. This research addressed research question RQ5.

Advanced Analysis: Modulation Impact, ML Generalization, and Complex Vibration Events

Impact of Modulation

All experiments described in the preceding subsections used unmodulated Continuous Wave (CW) signals. While this provides clean and controlled settings, real optical networks transmit modulated signals, that is, light waves whose properties are rapidly varied to encode and transmit information. These rapid variations introduce additional high-frequency fluctuations in the SOP that could potentially mask or interfere with disturbance signatures.

To examine whether SOP-based sensing remains reliable for signals that carry live traffic, we conducted a dedicated experiment on a 63.4 km metropolitan ASIERA (formerly HEAnet), Ireland's National Education and Research Network (HEAnet) [26] fiber ring equipped with six Reconfigurable Optical Add-Drop Multiplexers (ROADMs). Two channels were transmitted simultaneously through the same span: a 200 Gbps Dual-Polarization 16-Quadrature Amplitude Modulation (DP-16QAM) modulated signal representing a realistic coherent data channel, and an unmodulated CW signal. Both channels experienced identical physical disturbances, including a fiber-bend tap, a soft bend, and an 80 Hz vibration, resulting with dataset RDS3 with 4 classes of modulated and 4 classes of unmodulated data.

This controlled setting enabled a direct, one-to-one comparison of SOP signatures for modulated and unmodulated signals. As expected, the modulated channel exhibited additional high-frequency polarization fluctuations caused by rapid symbol transitions. Nevertheless, the lower-frequency components associated with external physical disturbances remained clearly distinghuishable. After appropriate preprocessing, the XGBoost and HGB ML classifiers trained on RDS3 successfully recognized the disturbances on both channels, with only a modest reduction in accuracy on the modulated signal.

These results, presented in **Paper I** and addressing RQ6, provide an empirical benchmark demonstrating that SOP-based monitoring is applicable not only to CW probes but also to modern coherent transmission systems.

ML Generalization

A further question concerns the extent to which models trained on one system setup remain effective when applied to different fiber links or spectral bands.

To investigate this, we evaluated generalization across datasets collected from two distinct systems (RDS4): a 21 km O-band dark fiber and a 63.4 km C-band metropolitan fiber link. We examined three scenarios using XGBoost classifier: intra-band (train and test on the data from the same system), cross-band (train on one system, test on the other), and multi-band (joint training on both).

This work, addressing RQ7 and reported in **Paper G**, showed that, while intra-band classification achieved high accuracy of up to 98.6%, cross-band accuracy was very limited, dropping to merely 8.1% in some cases. However, multi-band training mitigated this gap, improving accuracy to 91.1%. These findings indicate that polarization signatures are highly system-dependent, but that combining data from diverse links and bands enables models to learn more generalizable representations.

Complex Vibration Events

To evaluate the performance under realistic disturbance environments, we further analyzed complex vibrations. In **Paper H**, a dataset comprising 14 polarization signatures (LDS3) was constructed using pseudo-random broadband vibration patterns combined with tapping, bending, and 80 Hz malicious vibration events on both bare fiber and patch cable.

SL models were tested on these challenging scenarios, with the HGB classifier achieving an accuracy of 88.3% across all combinations. Despite the increased noise and event overlap, the models successfully distinguished harmful activities from benign disturbances and separated mixed vibration patterns. These results address RQ1, RQ2, and RQ8.

1.6 Thesis Outline

Following the introduction and problem statement in this Chapter, this thesis is structured as follows:

- Chapter 2 reviews the physical-layer threat landscape, analyzing vulnerabilities such as eavesdropping and service disruption, and evaluates current monitoring technologies including OTDR, DAS, and coherent telemetry, setting the stage for SOP-based sensing.

- Chapter 3 establishes the principles of the ML paradigms—SL, DL, SSL, and USL—utilized throughout the work.
- Chapter 4 presents the experimental framework and the data acquisition process across diverse fiber environments.
- Chapter 5 synthesizes the primary research contributions, mapping the validation of the proposed ML-driven framework across controlled, real-world, and complex operational scenarios to the specific research questions.
- Chapter 6 consolidates and highlights the author's principal contributions as presented in the included papers.
- Chapter 7 provides concluding remarks on the significance of the findings and outlines strategic directions for future research.

CHAPTER 2

Physical Layer Security in Optical Networks: Threat Landscape and Monitoring Approaches

Optical networks are increasingly attractive targets for adversaries seeking to exploit vulnerabilities at the physical layer to conduct eavesdropping or disrupt network services. The rising incidence of threats has intensified concerns about data confidentiality and network resilience, thereby establishing physical-layer security as critical in modern optical infrastructure [27]. Unlike higher-layer cryptographic measures such as encryption, which safeguard data confidentiality but cannot prevent interception or service disruption, physical-layer security focuses on securing the transmission medium itself by detecting and mitigating unauthorized access or physical intrusions [28]. Ensuring survivability in optical networks requires protection and detection mechanisms that address not only conventional failures but also intentional security breaches [27]. A robust defense can be achieved through a defense-in-depth approach, where physical-layer monitoring is integrated with higher-layer encryption schemes: the former enables timely detection of intrusion attempts, while the latter ensures that any intercepted data remains unintelligible [29].

This thesis focuses exclusively on the physical-layer monitoring aspect. This chapter provides a brief review of physical-layer security threats in optical

networks in Section 2.1, monitoring technologies available for detecting such threats in Section 2.2, and the concept and theoretical foundation behind the emerging paradigm of SOP-based fiber sensing in Section 2.3.

2.1 Physical Layer Threats in Optical Networks

Eavesdropping and service disruption represent the two primary categories of physical-layer threats. A detailed understanding of threat vectors, underlying mechanisms, and potential impact on network performance is essential, as discussed in the following subsections.

Eavesdropping

Eavesdropping threats aim at compromising data confidentiality through unauthorized interception of optical signals without necessarily causing noticeable service degradation [12]. Among the various intrusion methods, fiber tapping is the most prevalent, exploiting the physical accessibility of optical fibers to extract a portion of the transmitted light. Several tapping techniques have been reported in the literature, including micro- and macro-bending of the fiber, evanescent field coupling, V-groove cutting, optical splitting, and Bragg grating inscription on the fiber core [5], [30], [31].

Fiber Bending

Fiber bending is a well-known eavesdropping technique, which exploits the physics of light propagation in bent optical fibers [12], [30]. When an optical fiber is bent beyond a certain critical radius, the local incidence angle at the core–cladding interface is altered so that the conditions for total internal reflection are no longer fully satisfied, causing part of the guided light to leak out of the core and attenuation to increase [32]. This phenomenon, which is commonly used in fiber infrastructure monitoring [33], allows an attacker to recover a fraction of the transmitted optical signal by placing at the point of leakage either a photodetector, a device that converts the leaked light directly into an electrical signal, or a secondary fiber that collects and redirects the leaked light to a remote receiver [12], [31].

The efficiency and detectability of bending-based tapping threats depends primarily on the bend radius. Very tight bends introduce significant radi-

ation loss, whereas bends with larger radii cause only minimal attenuation and are therefore far more difficult to detect [34]. When the induced loss is sufficiently small, an attacker can extract a portion of the optical signal while remaining below the sensitivity threshold of conventional power-monitoring systems. Experiments have shown that even a small fraction of leaked light, sometimes only a few percent, can be adequate to reconstruct the transmitted data, depending on receiver sensitivity and tapping method [12], [32], [35].

The practical feasibility of this approach was reported in [36], where, by using an optical fiber clip-on coupler, an attacker in close proximity to the optical fiber could access unencrypted data traffic with minimal physical interaction. Clip-on couplers, which are commercially available devices designed to facilitate tapping, can be purchased online for as little as $200, which is quite affordable for potential adversaries [37].

Studies have shown that the G.657 bend-resistant fiber is more susceptible to certain bending-based threats than the standard G.652 fiber due to its modified refractive index profile optimized for bend insensitivity in normal operation [35]. This finding has important implications for security-aware network deployments, as fiber type selection involves trade-offs between operational flexibility (bend tolerance) and security resilience.

Beyond simple bending, numerous physical-layer tapping techniques have been demonstrated in the literature, ranging from evanescent-field coupling and fiber tapering to the exploitation of passive components such as splitters, monitoring ports, and Wavelength Division Multiplexing (WDM) elements. More invasive methods—including cladding removal, V-groove cutting, and Bragg grating inscription—have also been shown to enable partial extraction of the guided optical signal. While these approaches differ substantially in required equipment, technical complexity, and level of fiber modification, they share a common characteristic: an attacker can couple out a fraction of the transmitted light with minimal impact on link performance, often remaining below the detection threshold of conventional power- or integrity-monitoring systems [12], [30], [38], [39]. These techniques are therefore relevant primarily as evidence of the diverse and sophisticated threat landscape motivating the need for advanced monitoring methods, rather than as mechanisms requiring detailed analysis within the scope of this thesis.

Service Disruption Threats and Infrastructure Vulnerabilities

While eavesdropping threats aim to remain covert, service disruption threats deliberately degrade or completely interrupt network operation, often causing immediate and widespread impact affecting thousands to millions of users. These threats range from accidental mechanical damage during construction to deliberate sabotage motivated by terrorism, geopolitical conflict, or extortion, with fiber cuts representing the most catastrophic form of service disruption. Given the critical role that optical networks play in supporting essential services such as healthcare, financial systems, and national defense, service disruptions can have severe societal and economic consequences [40]. Understanding the mechanisms, impacts, and precursors of service disruption threats is essential for developing effective countermeasures and monitoring strategies.

Fiber Cuts and Complete Service Interruption

A *fiber cut* refers to the physical severing or damage of an active optical cable that disrupts normal network operation and necessitates immediate repair intervention. Such events typically result from accidental excavation, construction, or environmental factors near deployed fiber routes. The severity of the resulting outage depends on the number and location of affected fibers, directly influencing network performance and service availability. Fiber cuts significantly impact telecom service quality and contribute to increased operational costs and revenue loss for network operators [41].

The motivation for deliberate fiber cuts can vary widely, including geopolitical tensions leading to suspected threats on submarine and terrestrial cables connecting rival nations, activists and protesters targeting telecommunications infrastructure to disrupt government communications or draw attention to causes, organized crime cutting fibers to disable security systems during other criminal activities, and even individual vandals cutting fibers out of ignorance or malice [42]. The vulnerability of buried fiber infrastructure to construction and excavation activities represents a persistent and widespread threat, with common scenarios including construction activities where excavation for building foundations, road work, or utility installation frequently damages buried fiber cables despite legal requirements for location marking [43].

Vibration-Induced Degradation and Cut Precursors

Importantly, fiber cuts are often preceded by observable precursor events, particularly vibrations induced by excavation equipment or mechanical stress that gradually degrade fiber integrity before complete failure [42], [44]. This observation motivates the development of monitoring systems capable of detecting these precursors and triggering proactive protection measures. Mechanical vibrations from various sources can progressively damage fiber infrastructure, with construction equipment such as excavators operating near fiber routes inducing vibrations that propagate through structures. Detection of precursor vibrations enables proactive intervention, allowing operators to alert construction crews that they are working dangerously close to fiber infrastructure, reroute traffic to alternate paths before failure occurs, dispatch inspection crews to assess potential damage, and implement enhanced monitoring on affected segments [44].

Jamming and Signal Quality Degradation

Beyond complete physical damage, sophisticated attackers can employ optical-layer jamming to degrade service without causing obvious fiber cuts. Jamming signals can be, for example, Amplified Spontaneous Emission (ASE) noise sources or channels having the same wavelength as the targeted channel [28], [39]. **In-band jamming** involves injecting optical noise at the same wavelength as the targeted signal, adding unfilterable noise, degrading the Optical Optical Signal to Noise Ratio (OSNR) and increasing bit error rates. **Out-of-band jamming** exploits nonlinear effects in fiber and amplifiers, where by injecting high-power signals at wavelengths different from the target channel, an attacker can cause cross-phase modulation that adds phase noise to the target channel, four-wave mixing that generates interfering signals at the target wavelength, gain competition in optical amplifiers reducing target channel power, and stimulated Raman scattering that transfers power from shorter to longer wavelengths [28].

Threat Taxonomy and Impact Assessment

Physical-layer threats can be systematically classified to inform defense strategies:

1. **By Intent:**

    - **Malicious threats:** Deliberate actions by adversaries for eavesdropping or service disruption (sabotage)

    - **Accidental disturbances:** Unintentional damage from construction activities, environmental factors, or equipment failures

2. **By Damage:**

    - **Harmful events:** These events include deliberate or high-impact disturbances such as fiber tapping, invasive modifications (e.g., V-groove cutting or Bragg grating inscription), malicious splitter insertion, potentially harmful mechanical vibrations, or physical sabotage. Such events directly threaten network integrity and confidentiality, often resulting in measurable signal leakage, service degradation, or complete disruption.

    - **Non-harmful events:** These events represent benign perturbations arising from normal network operation or maintenance, such as soft bending during cable handling, minor mechanical vibrations from nearby activity, or slow environmental drifts. While they may induce transient polarization or power fluctuations, they pose no actual risk to data integrity or service continuity.

For critical infrastructure applications, financial services, healthcare, or national defense, even brief service interruptions can have consequences disproportionate to the duration of outage. High-frequency trading systems can lose millions in seconds. Telemedicine applications may be unable to deliver time-critical care. Military command and control may be degraded during critical operations. This comprehensive threat landscape underscores the critical importance of physical-layer monitoring and protection mechanisms that can detect both covert eavesdropping attempts and precursors to service-disrupting events. The following sections examine the technologies and methodologies that have been developed to address these challenges.

2.2 Fiber Monitoring Technologies for Threat Detection

Effective physical-layer security in optical networks requires continuous monitoring capabilities that can detect, localize, and characterize disturbances along fiber links. This section reviews standard fiber monitoring technologies, examines their principles, capabilities, limitations, and suitability for different deployment scenarios, with an emphasis on their application to security monitoring and threat detection.

Optical Time-Domain Reflectometry (OTDR)

OTDR is one of the most established and widely deployed fiber monitoring technologies, originally developed for fault detection and localization in telecommunication networks [45]. It relies on analyzing Rayleigh backscattered light that is generated when optical pulses propagate through a fiber. When a high-power optical pulse is launched into a fiber under test, a small fraction of the optical power is continuously scattered backwards due to microscopic refractive index variations inherent in the glass structure. In this way, the technique supports monitoring of signal power along the fiber, affected by different events.

Spatial resolution of OTDR depends on the pulse width. Commercial OTDR systems typically support sub-meter (for short-range, high-resolution measurements) to tens of meters resolution (for long-haul monitoring extending beyond 100 km). The trade-off between spatial resolution, sensitivity, and measurement range represents a fundamental constraint in OTDR-based monitoring systems [46].

Coherent Optical Time-Domain Reflectometry (COTDR) provides advanced sensitivity by measuring both amplitude and phase of backscattered signals. This enables detection of smaller perturbations, including subtle bending or stress that might indicate tampering attempts. Photon-counting OTDR utilizes single-photon avalanche detectors to achieve unprecedented sensitivity, enabling measurements at extremely low backscatter levels and thereby extending the achievable sensing range or improving detection of weak disturbances [47].

Recent demonstrations of OTDR applied for security monitoring include detection of fiber tapping through bend-induced loss signatures [14], [15].

Limitations and Challenges

Despite its maturity and widespread use for fault localization, OTDR faces several limitations when applied to real-time security monitoring of live networks. Conventional OTDR requires injection of high-power test pulses into the fiber, which can interfere with data transmission if performed on active links. While techniques for OTDR operation on live fibers have been developed using out-of-band wavelengths and careful power management, the need for high peak power remains a concern for network operators [46], [48].

Acquisition of an OTDR measurement typically takes a few seconds or minutes, depending on the desired spatial resolution and averaging requirements, limiting real-time monitoring capabilities [44]. This temporal resolution may be insufficient for detecting rapid, transient events such as brief vibrations or momentary fiber stress.

The sensitivity of conventional OTDR to subtle perturbations is limited. Detection of low-loss tapping methods, particularly evanescent coupling or carefully executed macro-bending with loss below the measurement noise floor ($<$0.1 dB), remains challenging [15]. Advanced techniques such as COTDR or photon-counting OTDR offer improved sensitivity but at substantially increased system complexity and cost.

Finally, deployment costs for OTDR-based monitoring at scale can be prohibitive. Each monitored fiber span requires dedicated OTDR equipment at one or both ends, and for large networks with hundreds or thousands of fiber links, the capital expenditure for comprehensive coverage becomes substantial. Additionally, OTDR equipment requires periodic calibration and maintenance, adding to operational expenses, making large-scale deployment economically challenging for many network operators.

These limitations motivate the exploration of alternative monitoring approaches that can operate continuously on live traffic, detect subtle perturbations with high sensitivity, and scale economically to large network deployments.

Distributed Acoustic Sensing (DAS)

DAS represents a significant advancement in fiber-based monitoring technology, transforming standard optical fibers into distributed arrays of acoustic sensors capable of detecting and localizing vibrations, strain, and temperature

variations along the entire fiber length. DAS relies on Phase-Sensitive Optical Time-Domain Reflectometry (Φ-OTDR), which analyzes coherent Rayleigh backscattering to detect minute phase changes induced by external perturbations. In a Φ-OTDR system, narrow-linewidth coherent optical pulses are launched into the sensing fiber. The backscattered light from different spatial locations along the fiber interferes coherently at the receiver, and any mechanical disturbance that alters the fiber length or refractive index within a resolution element causes a corresponding phase change in the backscattered signal. By continuously interrogating the fiber and analyzing temporal variations in the backscatter phase or intensity pattern, DAS systems can detect and localize vibrations with exceptional sensitivity [46], [48].

The spatial resolution of DAS systems is determined by the optical pulse width and can range from sub-meter to tens of meters, while sensing ranges extend from several kilometers in high-sensitivity configurations to over 100 km in extended-reach implementations [49]. DAS provides truly distributed sensing: every meter of fiber acts as an independent sensor, enabling detection of events occurring simultaneously at multiple locations without ambiguity. The frequency response of DAS systems typically begins at very low frequencies and extends up to several kilohertz, enabling them to detect a wide range of acoustic and vibration events relevant to security monitoring.

The high sensitivity and distributed nature of DAS make it well-suited for intrusion detection and physical-layer security applications. DAS has been successfully deployed for perimeter security monitoring, where buried fiber cables detect footsteps, digging, or vehicle movement near protected facilities [18], [49]. DAS can detect various threats to optical network security including excavation activities near buried fiber routes (providing early warning before fiber cuts occur), mechanical vibrations indicating tampering attempts, and unauthorized access to cable routes or equipment enclosures [44].

Advanced DAS variants have been developed to enhance performance for security applications. Distributed Vibration Sensing (DVS) focuses on detecting and characterizing vibration events with enhanced frequency response [49]. Quantitative DAS (Q-DAS) provides calibrated strain measurements rather than relative intensity changes, enabling more precise characterization of fiber perturbations [50]. Frequency-scanned Φ-OTDR employs wavelength tuning to enhance measurement precision and extend sensing range [51].

Limitations and Challenges

Despite its impressive sensing capabilities, DAS faces several challenges that limit its widespread adoption for network-wide security monitoring. The primary limitation is the requirement for dedicated, specialized hardware. DAS systems employ sophisticated optical interrogation units with high-performance lasers, optical pulse generators, coherent receivers, and high-speed digitizers. The capital cost of DAS equipment is substantial, typically ranging from tens of thousands to hundreds of thousands of dollars per interrogation unit, depending on the sensing range and performance specifications [46]. For monitoring large optical networks with hundreds of fiber links, deploying dedicated DAS interrogators for each link becomes economically prohibitive.

Moreover, a typical DAS system generates massive amounts of data, as the high sampling rate combined with the large number of spatial sensing points along the fiber results in continuous high-throughput data streams that require substantial computational resources for real-time processing. Analyzing this data stream in real time to detect relevant security events while minimizing false alarms requires substantial computational resources and sophisticated signal processing algorithms, adding to system complexity and operational costs.

Finally, DAS systems typically require dedicated dark fibers for sensing, as the high-power interrogation pulses can interfere with data transmission on lit fibers. The requirement for dark fiber limits DAS applicability, as many deployed fiber routes lack spare fibers available for monitoring purposes [52].

Coherent Receiver Telemetry and Digital Signal Processing (DSP)-Based Monitoring

Modern high-capacity optical networks predominantly employ Digital Signal Processing (DSP)-based coherent receivers to detect multi-level modulation formats such as Dual-Polarization Quadrature Phase-Shift Keying (DP-QPSK) and Dual-Polarization (DP)-Quadrature Amplitude Modulation (QAM) [53], [54]. These coherent receivers inherently measure amplitude and phase of the optical field in both polarization dimensions, providing access to rich telemetry data that can be leveraged for fiber monitoring purposes without requiring additional dedicated sensing hardware [55].

DSP-based coherent receivers routinely extract various performance metrics

during normal data reception, including OSNR, Chromatic Dispersion (CD), Polarization Mode Dispersion (PMD), Polarization-Dependent Loss (PDL), and frequency offset, directly from the received signal constellation [55], [56]. The diagnostic information made available through coherent-receiver telemetry provides valuable insights into link behavior and can be leveraged to identify anomalies related to threats or operational failures.

Several research efforts have investigated the use of coherent receiver telemetry for security monitoring. Changes in received power, OSNR degradation, sudden increases in bit error rate, or anomalous polarization state variations can indicate service disruption threats such as jamming or polarization scrambling [57], [58], [59], [60].

The primary advantage of coherent receiver telemetry for security monitoring is that it leverages existing network infrastructure without requiring deployment of additional sensing equipment. The telemetry data is available continuously during normal network operation, and modern coherent receivers report telemetry with sub-second update rates, enabling relatively fast detection of anomalous conditions [56]. This allows for the development of various monitoring algorithms, providing deployment flexibility and potential for adaptation to new threat types.

Limitations and Challenges

Despite these advantages, coherent receiver telemetry faces several limitations for comprehensive physical-layer security monitoring. The main limitation is related to providing only *end-to-end* visibility: link-aggregate properties are measured at the receiver and do not indicate the location of a disturbance along the multi-link path of a connection, unless additional intermediate monitoring points or correlation with other sensors is available. For long-haul links spanning hundreds of kilometers, and connections spanning multiple links and nodes, this limitation constrains operational response capabilities, as it does not allow for determining the fiber section which requires inspection or repair [55].

Secondly, telemetry-based monitoring only applies to lit fibers with coherent signals that carry traffic. Dark fibers or fibers carrying legacy modulation formats without coherent receivers remain unmonitored unless additional sensing infrastructure is deployed [55].

These considerations motivate the exploration of monitoring approaches

that build on the capabilities already available in coherent transceivers. In this context, SOP-based fiber sensing serves as a complementary technique to telemetry-based monitoring: both approaches exploit existing receiver hardware and require no additional sensing infrastructure, yet they provide different and mutually reinforcing perspectives on physical-layer behavior. SOP-based sensing captures the rapid polarization dynamics induced by mechanical and environmental perturbations and utilizes the polarization properties of optical signals as an intrinsic sensing mechanism. The combination of these two information sources enables richer and more sensitive detection of physical-layer disturbances, even though both inherit similar constraints such as the lack of inherent spatial localization.

State of Polarization (SOP): from Optical Property to Sensing Engine

The polarization state of light propagating through an optical fiber represents a fundamental optical property that is highly sensitive to mechanical and environmental perturbations [61]. Unlike intensity-based monitoring (as in conventional OTDR) or phase-based sensing (as in DAS), which measure specific aspects of the optical field, polarization monitoring captures the three-dimensional orientation of the optical electric field vector as it evolves along the fiber [62]. This rich vectorial information provides a unique signature for different types of physical disturbances affecting the fiber.

The key phenomenon enabling SOP-based sensing is that the birefringence of optical fibers, i.e., the difference in refractive index experienced by light polarized along different axes, is very sensitive to external perturbations. Mechanical stress, bending, vibration, temperature fluctuations, strain, and pressure modify the local birefringence of the fiber, which in turn causes the polarization state to evolve differently than under normal conditions [63], [64]. By continuously monitoring the SOP changes over time, one can detect and potentially characterize these external influences.

Several factors make polarization monitoring particularly attractive for security applications. First, SOP measurement can be performed using relatively simple and inexpensive equipment compared to sophisticated DAS interrogators: a polarimeter or the polarization-diverse front-end of a coherent receiver suffices [52]. Second, SOP monitoring is inherently non-invasive and can operate on fibers carrying live network traffic without requiring test sig-

nal injection or dedicated dark fibers [21], [44]. Third, polarization evolves rapidly in response to fiber perturbations, enabling near real-time detection of dynamic events [20].

The transition from viewing polarization as a passive optical property that must be managed (e.g., through PMD compensation in receivers) to leveraging it as an active sensing modality represents a paradigm shift in optical network monitoring. This paradigm enables several novel capabilities. First, it transforms the fiber into a sensing element without requiring any modifications. The standard telecom fiber becomes a sensor by virtue of monitoring a property that naturally responds to external influences [21]. Second, it exploits polarization diversity that is already present in modern coherent optical systems, where both polarization components are measured for data detection, making SOP monitoring essentially "free" in terms of receiver hardware [20], [55]. Third, it provides complementary information to traditional monitoring approaches: while OTDR detects discrete reflection and attenuation events, and DAS senses acoustic/vibrational energy, SOP monitoring detects perturbations through their impact on birefringence, including changes induced by mechanical vibration, bending, or temperature fluctuations.

The richness and complexity of polarization variation data naturally motivates the application of ML techniques for interpretation. As noted in Chapter 1, raw SOP data is high-dimensional and noisy, with normal fiber operation causing constant background polarization drift. This combination of physical sensing and computational intelligence forms the technical foundation for the research presented in this thesis. This motivates a deeper examination of the physical principles governing SOP evolution in optical fibers. Section 2.3 develops the brief foundations of SOP, which subsequently enable the ML-based analysis explored in the later chapters.

2.3 SOP-Based Fiber Sensing

This section establishes the mathematical framework and physical mechanisms underlying SOP-based fiber sensing, providing the foundation for the experimental methodologies and ML analyses presented in subsequent chapters.

Polarization States and Mathematical Representations

Light propagating through an optical fiber consists of electric and magnetic fields that move together along the fiber. The electric field oscillates perpendicular to the direction of the light path. The SOP describes how the direction of this electric field vector changes over time at a fixed point along the fiber.

For a single-color (monochromatic) light wave moving in the z direction, the electric field can be decomposed into two orthogonal components, one along the x axis and one along the y axis [65]:

$$\mathbf{E}(z,t) = \mathrm{Re}\left\{E_x e^{i(\omega t - kz + \phi_x)}\hat{\mathbf{x}} + E_y e^{i(\omega t - kz + \phi_y)}\hat{\mathbf{y}}\right\} \tag{2.1}$$

$$= E_x \cos\big(\omega t - kz + \phi_x\big)\,\hat{\mathbf{x}} + E_y \cos\big(\omega t - kz + \phi_y\big)\,\hat{\mathbf{y}}. \tag{2.2}$$

where E_x and E_y are the real amplitudes along the x and y directions, ω is the angular frequency, and ϕ_x and ϕ_y are the absolute phases of the two components. k is the wave number and is defined as $k = 2\pi/\lambda$, representing how many wavelengths fit into a unit distance along the propagation direction where λ is the wavelength of the light in the medium.

The relative phase difference $\delta = \phi_y - \phi_x$ and the amplitude ratio E_y/E_x determine the polarization state. When $\delta = 0$ or π, the light is linearly polarized; when $|E_x| = |E_y|$ and $\delta = \pm\pi/2$, the light is circularly polarized; all other combinations yield elliptical polarization.

Stokes Parameters

The Stokes parameters provide a complete description of any polarization state using four real quantities that can be measured experimentally [66], [67]. The Stokes vector is defined as:

$$\mathbf{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \tag{2.3}$$

where:

- $S_0 = I$ represents the total optical intensity (or total power) of the field.
- $S_1 = Q$ quantifies the preference for horizontal versus vertical linear

polarization. A positive value ($Q > 0$) indicates a stronger horizontal component, whereas a negative value ($Q < 0$) indicates a stronger vertical component.

- $S_2 = U$ quantifies the preference for linear polarization oriented at $+45°$ versus $-45°$. A positive value ($U > 0$) indicates dominance of the $+45°$ component, while a negative value ($U < 0$) corresponds to dominance of the $-45°$ component.

- $S_3 = V$ quantifies the preference for right-hand versus left-hand circular polarization. A positive value ($V > 0$) corresponds to right-hand circular polarization, whereas a negative value ($V < 0$) corresponds to left-hand circular polarization.

For a monochromatic plane wave with electric field components E_x and E_y, the Stokes parameters can be expressed in terms of the field amplitudes and phase difference [65]:

$$S_0 = |E_x|^2 + |E_y|^2 \tag{2.4}$$
$$S_1 = |E_x|^2 - |E_y|^2 \tag{2.5}$$
$$S_2 = 2|E_x||E_y| \cos\delta \tag{2.6}$$
$$S_3 = 2|E_x||E_y| \sin\delta \tag{2.7}$$

The Stokes parameters collectively describe all possible polarization states:

- Linear polarization occurs when $S_3 = 0$, meaning there is no phase difference between the x and y components ($\delta = 0$ or π).

- Circular polarization occurs when $S_1 = S_2 = 0$ and S_3 reaches its maximum or minimum value ($\delta = \pm\pi/2$).

- Elliptical polarization represents the most general case, occurring when both S_2 and S_3 are nonzero. In this case, the tip of the electric field vector traces an ellipse over time.

For fully polarized light, the Stokes parameters satisfy:

$$S_0^2 = S_1^2 + S_2^2 + S_3^2 \tag{2.8}$$

For partially polarized or unpolarized light:

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \tag{2.9}$$

The Degree of Polarization (DoP) quantifies the fraction of light that is polarized:

$$\text{DoP} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{2.10}$$

with DoP = 1 for fully polarized light and DoP = 0 for completely unpolarized light.

Poincaré Sphere Representation

A powerful geometric visualization of polarization states is provided by the *Poincaré sphere* which is illustrated in Figure 2.1. Each fully polarized state corresponds to a unique SOP point on the unit sphere defined by the normalized Stokes coordinates [68]:

$$(s_1, s_2, s_3) = \left(\frac{S_1}{S_0}, \frac{S_2}{S_0}, \frac{S_3}{S_0}\right) \tag{2.11}$$

The three axes of the sphere correspond to:

- The S_1 axis (H–V): Horizontal (H) vs. Vertical (V) linear polarization.
- The S_2 axis (D–A): $+45°$ Diagonal (D) vs. $-45°$ Anti-diagonal (A) linear polarization.
- The S_3 axis (R–L): Right (R)- vs. Left (L)-hand circular polarization.

The angular distance between two polarization states on the Poincaré sphere is directly related to the inner product of their Stokes vectors, providing a natural metric for quantifying polarization changes. This geometric representation is particularly valuable for SOP-based sensing, as external perturbations cause the SOP to trace specific trajectories on the Poincaré sphere, creating distinctive signatures for different types of disturbances.

Polarization Measurement Techniques

Practical implementation of SOP-based sensing requires instruments capable of measuring the Stokes parameters or equivalent polarization descriptors in

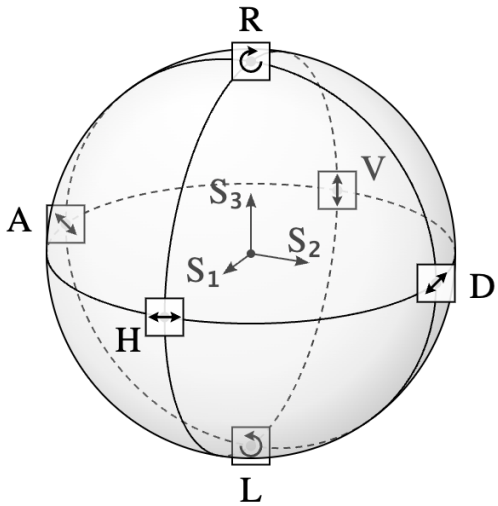


Figure 2.1: The Poincaré sphere representation of polarization states.

real-time. Two primary approaches are employed in the literature:

External Polarization Analyzer

1. **Commercial sensor modules:** Commercial sensor modules for SOP monitoring are specialized instruments that measure polarization changes in optical fibers by determining all four Stokes parameters in real-time. These modules achieve measurement rates from 100 kHz to 4 MHz, depending on the implementation. For fiber monitoring applications, these modules can be deployed either as inline sensors that continuously monitor live traffic without interruption, or as endpoint analyzers for component testing and system characterization [62], [69].

2. **Polarimeters:** A polarimeter is a device that measures the complete polarization state of light, typically by determining all four Stokes parameters. Commercial polarimeters can achieve measurement rates from kHz to MHz, enabling real-time tracking of SOP variations. For fiber monitoring applications, polarimeters can be placed at fiber endpoints to continuously monitor the output SOP [70].

Coherent Receivers

Modern coherent optical receivers used in high-speed communication systems inherently perform polarization-diverse detection, measuring both orthogonal polarization components with phase and amplitude information [53], [54]. The DSP within coherent receivers can extract Stokes parameters from the measured complex field components without requiring additional hardware.

This capability makes coherent receivers particularly attractive for SOP-based monitoring in deployed networks: every coherent transceiver can potentially serve as a polarization sensor, enabling distributed monitoring across the network infrastructure at essentially zero marginal hardware cost. The integration of SOP monitoring into existing communication equipment represents a significant practical advantage for large-scale deployment.

Birefringence and Perturbation-Induced SOP Variations

Birefringence, the dependence of refractive index on polarization state, is the fundamental physical mechanism underlying SOP-based fiber sensing. In optical fibers, birefringence arises from both intrinsic factors (such as core ellipticity and frozen-in stress from manufacturing) and extrinsic perturbations (including bending, twisting, lateral pressure, and temperature variations) [63], [64]. As light propagates through the fiber, it encounters a continuous series of birefringent sections, each causing the polarization state to rotate. This evolution can be visualized on the Poincaré sphere, where the polarization vector traces a trajectory determined by the cumulative effect of local birefringence along the entire fiber length.

Birefringence also gives rise to PMD, where the two orthogonal polarization modes travel at slightly different group velocities, causing a differential group delay between them. When an external perturbation such as fiber bending or mechanical vibration modifies the local birefringence at any point along the fiber, it alters the polarization rotation at that location, causing the output SOP to deviate from its unperturbed state. This perturbation manifests as time-varying fluctuations in the measured Stokes parameters at the receiver. The cumulative nature of polarization evolution means that disturbances anywhere along the fiber affect the final state, enabling the entire fiber to function as a sensing medium where local perturbations are encoded into observable polarization signatures [64], [71], [72].

The key principle enabling SOP-based sensing for security monitoring is that different types of perturbations produce distinctive temporal and spectral signatures in the observed polarization variations. These SOP signatures can be characterized by their temporal dynamics. The frequency content of SOP variations, revealed through spectral analysis, provides another discriminating feature, as different perturbation sources exhibit characteristic dominant frequencies and harmonics. Additionally, the amplitude of SOP variations on the Poincaré sphere, their statistical properties, and the specific trajectory patterns traced during perturbations all contribute to forming unique fingerprints for different event types. By extracting appropriate features from measured SOP time series and applying ML algorithms, it becomes possible to detect anomalous events and classify the nature of disturbances affecting the fiber. This forms the core methodology of this thesis: combining the inherent physical sensitivity of polarization to external perturbations with the capabilities of ML to achieve intelligent, automated fiber security monitoring.

2.4 Chapter Summary

This chapter provided a brief exploration of physical-layer threats in optical networks, i.e., eavesdropping and service disruption attacks. Real-world incidents and the evolving sophistication of tapping technologies underscore the critical need for effective physical-layer monitoring. Disruption vectors such as fiber cuts, mechanical vibrations, and optical jamming were examined, revealing their potential to cause severe service outages and societal impact (Section 2.1).

To address these challenges, several established monitoring technologies were reviewed in Section 2.2, including OTDR, DAS, and coherent receiver telemetry. While each offers valuable capabilities, limitations related to their deployment cost, spatial coverage, or sensitivity motivate the exploration of complementary solutions that can be integrated into existing network infrastructure without requiring specialized hardware. This motivates the introduction of polarization-based sensing, where the SOP is leveraged as a highly sensitive indicator of environmental and mechanical perturbations affecting the fiber. Section 2.3 outlined the physical and mathematical foundations of polarization, including its representation, Stokes parameters, and the Poincaré sphere. The discussion also highlighted how birefringence and external dis-

turbances induce SOP evolution, producing temporal and spectral signatures that can be exploited for monitoring.

These concepts form the physical basis for the SOP-driven sensing methodology developed throughout the thesis. Building on this foundation, the next chapter introduces the ML techniques used to detect and classify fiber perturbations in later chapters.

CHAPTER 3

Machine Learning (ML) for Detection and Classification of SOP Signatures

The SOP data carries rich information about how light polarization evolves under different physical disturbances in an optical fiber. Frequent manual analysis of these SOP variations to identify unusual patterns would require significant human effort, incurring high labor cost and scalability challenges. Traditional threshold-based monitoring approaches suffer from similar scalability issues, and cannot effectively capture the complex patterns that distinguish different types of perturbations or separate malicious events from benign environmental variations [21].

The detection of abnormal patterns in SOP data can be significantly enhanced and automated through the application of ML techniques. ML provides the computational framework for automatically extracting patterns from data without explicit programming, and making predictions over unseen data. By leveraging SL, SSL, and USL techniques, it is possible to identify and possibly classify deviations in polarization patterns that could indicate potentially harmful events or irregularities over fiber optic installations.

This chapter establishes the foundations of ML techniques used for analyzing SOP-based fiber sensing data in this thesis. It provides an overview on

the key SL algorithms and DL architectures used in this thesis in Section 3.1 and SSL/USL approaches in Section 3.2 that enable automatic detection and classification of fiber disturbances. The chapter also outlines the performance evaluation metrics and validation strategies used throughout the thesis to assess model performance and reliability in Section 3.3. Finally, Section 3.4 surveys the landscape of ML applications in optical network monitoring, identifying current gaps and contextualizing the proposed approach within the state of the art.

3.1 Supervised Learning (SL) Classification

SL classifiers aim to learn a mapping from input features to output labels using a labeled dataset. Formally, given a training set $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, where each input vector $\mathbf{x}_i \in \mathbb{R}^d$ is associated with a corresponding label $y_i \in \mathcal{Y}$ (e.g., $\mathcal{Y} = 1, \ldots, C$ for C classes), the objective is to learn a function $f : \mathbb{R}^d \to \mathcal{Y}$ that minimizes the expected loss over the data distribution:

$$f^* = \arg\min_f \mathbb{E}_{(\mathbf{x},y)\sim\mathcal{D}}[L(f(\mathbf{x}), y)] \tag{3.1}$$

where $L(f(\mathbf{x}), y)$ represents the loss function measuring the discrepancy between the predicted ($y' = f(x)$) and true (y) labels. In practice, this objective is optimized by learning from annotated (labeled) examples. The expectation is that the trained model can generalize to unseen data.

In the context of optical fiber monitoring, SL algorithms can be trained to classify and recognize different types of fiber disturbances or anomalous events based on their polarization signatures. When representative labeled data are available for diverse fiber conditions, such as normal operation, various vibration levels, or tapping attempts, these classifiers can effectively learn the distinguishing SOP patterns associated with each event type. The main strength of SL lies in its high classification accuracy (i.e., greater than 90% [73], [74]) under sufficient labeled data, enabling reliable detection of known events, including subtle changes in SOP indicative of potential security threats like eavesdropping or tampering. However, the need for labeled datasets can still be a limitation, particularly for rare or emerging disturbances. Despite this, SL remains a powerful ML technique for SOP-based monitoring, as it enables high-performing models when adequate representative labeled examples are

available. This thesis employs a diverse set of SL algorithms spanning multiple paradigms to analyze SOP data. The following subsections introduce the SL classifiers utilized in this work, outlining their principles and specific relevance to fiber perturbation detection.

Conventional Classical SL Classifiers

Decision Trees (DTs)

A DT is a hierarchical model that recursively partitions the feature space into subsets based on feature tests, forming a tree-like structure of decision rules. Each internal node represents a test on a feature, each branch represents the outcome of the test, and each leaf node represents a class label or a predicted value. DTs are among the most popular ML algorithms, producing models that are easy to interpret and visualize. However, a single DT can overfit the training data if grown too complex. An overfit DT memorizes quirks of the training set, including noise, and fails to generalize well to new data. To combat overfitting, techniques like pruning (trimming back the tree's depth) or setting minimum leaf sizes are often applied. Despite this, DTs remain a powerful baseline; they capture non-linear feature interactions and require little data preparation [75], [76].

In our SOP event classification context, a DT would sequentially evaluate conditions about polarization features to finally decide the class of disturbance. While straightforward and fast, its accuracy might be limited compared to more advanced methods, especially if the SOP patterns are complex. This motivates the use of ensemble methods that combine many DTs to improve accuracy.

Ensemble Methods: Bagging

Ensemble learning methods improve predictive performance by combining the outputs of multiple base learners. One popular ensemble technique is bootstrap aggregating (bagging), which builds many independent models on different random subsets of the training data and averages their predictions. Bagging with DTs significantly reduces variance and yields more stable predictions than a single DT [77]. In essence, bagging "de-noises" the decision by majority voting, which is very effective for noisy high-dimensional data like SOP streams. In the following, we list two main ensemble variants related to

bagging:

- **Random Forest (RF)**

  RF is an extension of bagging that further injects randomness into tree construction to decorrelate the ensemble members [78]. In an RF, each tree is trained on a bootstrap sample which is a random sample with replacement and at each split, the tree is restricted to choose the best split from a random subset of features rather than all features. This ensures that no single feature dominates all trees, and the resulting ensemble of trees, a *forest*, is more diverse. Each tree remains an accurate classifier since it selects optimal splits from the random feature subset, but by averaging many such trees the overall model achieves high accuracy. RFs are known for excellent out-of-the-box performance on classification tasks; they can handle large feature spaces, are resistant to overfitting due to averaging, and require little parameter tuning.

  In our application, an RF can capture complex nonlinear relationships between polarization features and event types, while mitigating the risk of overfitting to peculiarities of one fiber link or one environmental condition. The randomness in feature selection means even subtle or distributed polarization indicators of an attack can be picked up by at least some trees, contributing to the ensemble vote.

- **Extra Trees (ET) classifier**

  Another ensemble variant related to bagging is the ET classifier or *extremely randomized trees*. Like RF, ET builds an ensemble of many DTs and averages their results, but it increases diversity by randomizing the split thresholds (a numeric cutoff) in addition to the feature choice. In ET, when splitting a node, a split point is chosen at random for each candidate feature rather than computing the optimal cut, and then the best among these random splits is selected. Moreover, ETs typically use the full training dataset without bootstrapping. The algorithm injects randomness not through resampling but by selecting random subsets of features and randomly generating candidate split thresholds [79]. This extra randomness tends to lower the variance of the ensemble at the cost of a slight increase in bias, and can lead to even better generalization in some cases. It also makes training extremely fast, since finding random split points is computationally cheaper than searching for optimal splits.

In fiber SOP event classification, the ET algorithm can quickly generate a diverse set of decision rules capturing different aspects of polarization changes. The ensemble as a whole achieves strong accuracy, often comparable to RF, while being computationally efficient.

In summary, bagging-based ensembles like RF and ET provide powerful high-performing classifiers for our problem, leveraging the stability of many uncorrelated DTs.

Ensemble Methods: Boosting (Sequential Error Correction)

In contrast to bagging's parallel aggregation, boosting constructs an ensemble sequentially, where each model is trained to emphasize instances that were misclassified by the preceding learners [80]. Whereas bagging primarily reduces variance, boosting is designed to reduce bias by iteratively refining the decision boundary and enabling the ensemble to model increasingly complex patterns. Boosting typically employs weak learners, models that perform only marginally better than random guessing, often shallow DTs. At each iteration, the algorithm adjusts the weights of training examples or the residuals of previous models, directing subsequent learners to focus on the more challenging samples. The final decision is obtained through a weighted combination of all learners, where models with higher accuracy contribute more strongly to the ensemble output. Over time, the boosting principle has given rise to several advanced frameworks that differ in optimization strategies, regularization mechanisms, and computational efficiency. Three notable boosting approaches relevant to our work are:

- **Gradient Boosting (GB)**

  GB is a general boosting framework that typically uses shallow DTs as weak learners, adds them one by one in a stage-wise fashion, and optimizes an arbitrary differentiable loss function using gradient descent in function space. At each stage, a new tree is trained to predict the residuals, i.e., the remaining errors of the current ensemble model on the training data. By optimizing a chosen loss function, such as classification error, GB iteratively "boosts" the performance of the model, focusing on instances that previous trees misclassified. Over many rounds, the ensemble of trees forms a strong predictive model that often achieves higher accuracy than any single tree could. GB is flexible: one can

choose different differentiable loss functions and use regularization techniques to prevent overfitting [81].

In the context of classification of SOP signatures, GB can gradually capture subtle patterns in polarization behavior during training that distinguish, e.g., a harmless vibration from a malicious fiber tapping. Initially, a small tree might capture the most obvious difference; subsequent trees focus on more nuanced residual patterns, refining the decision boundary. The result is an accurate model capable of distinguishing complex abnormal polarization signatures that simpler methods might miss. Indeed, we found that gradient-boosted tree ensembles effectively differentiate between normal and anomalous SOP states, contributing to more accurate fiber disturbance monitoring.

- **eXtreme Gradient Boosting (XGBoost)**

  XGBoost is a fast and regularized version of GB that uses both gradients and Hessians, meaning the first and second derivatives of the loss function, to make more accurate split decisions. It also includes several engineering optimizations that improve speed for large SOP datasets. It also handles sparse data and missing values gracefully. These improvements allow XGBoost to build large ensembles of trees efficiently, often outperforming earlier boosting algorithms [82].

  From the perspective of fiber-perturbation classification, XGBoost provides a robust and effective learning framework for SOP datasets. Fiber monitoring can produce large amounts of polarization data, and XGBoost remains efficient to train on such datasets due to its optimized implementation and ability to handle high-dimensional feature inputs. The built-in regularization (shrinkage and tree penalization) helps prevent overfitting even as the model complexity grows, which is important because an overfit model might fail to generalize during operation. XGBoost has indeed been observed to deliver top accuracy for the classification of polarization signatures. Overall, XGBoost represents a state-of-the-art SL algorithm combining the power of boosting with practical efficiency for complex tasks like the ones investigated in this thesis.

- **Histogram Gradient Boosting (HGB)**

  HGB is an advanced variant of the GB algorithm designed for improved

efficiency on large datasets. The key idea of HGB is to discretize continuous features into a fixed number of bins (histograms), and then use these binned values to find split points, rather than considering every unique value in the data. By converting continuous inputs into uniform or quantile-based bins, the algorithm drastically reduces the number of possible split candidates it must evaluate. This histogram-based approximation leads to significant speed-ups and memory savings with only minimal loss in accuracy. Training each tree becomes faster because the data is first binned and aggregated, and splits are chosen based on these aggregated bin statistics [83].

In the fiber monitoring domain, HGB is especially advantageous when dealing with high-dimensional polarization features or very high sampling rates. The volume of SOP data can be enormous, but HGB's binned splitting can handle tens of thousands of samples and features efficiently. In summary, HGB retains the strong predictive power of boosting ensembles while optimizing performance via feature binning.

Support Vector Machine (SVM)

SVM classification from a geometric perspective is finding the optimal boundary that separates different classes while maximizing the confidence of that separation [84]. The core philosophy is elegant: find the separating boundary with the maximum margin, i.e., the widest safety zone between classes, as this generalizes best to new data. SVM strengths include powerful kernel functions, mathematical transformations that map data into higher-dimensional spaces to make non-linear patterns linearly separable, enabling classification in non-linear feature dimensions. However, it requires careful hyperparameter tuning.

Logistic Regression (LR)

Despite its name suggesting regression, LR is a linear classification algorithm that probabilistically models class membership [85]. Unlike SVM's geometric separating boundary maximization or DT's hierarchical splitting, LR directly estimates the probability that a sample belongs to each class.

LR's advantages include fast training and prediction, direct probability outputs, good interpretability since the magnitude of learned weights reflects each

feature's contribution to the prediction, and effectiveness as a baseline. that often performs surprisingly well. However, being a linear model, it struggles with complex non-linear patterns without explicit feature engineering.

In this thesis, LR is included in the comprehensive SL benchmarking as a linear baseline, providing a lower bound on the performance against which more complex ensemble methods can be compared.

Linear Discriminant Analysis (LDA)

LDA is a classical and computationally efficient technique used for both dimensionality reduction and classification. It assumes that the data from each class follows a multivariate Gaussian distribution with a shared covariance matrix and seeks to find linear combinations of features that best separate the classes by maximizing the ratio of between-class to within-class variance [86].

In SOP-based fiber monitoring, LDA can project high-dimensional polarization features into a lower-dimensional space where different event types are more distinctly separated. While it performs well when its assumptions are met, LDA is limited to linear decision boundaries and may underperform on complex, non-Gaussian data.

In this thesis, LDA is included in the benchmarking suite as a linear baseline, offering a computationally efficient reference point that contextualizes the performance gains achieved by ensemble and kernel-based approaches.

K Nearest Neighbors (KNN)

KNN is a simple yet effective non-parametric classification algorithm that assigns a class label to a new data point based on the majority class of its k closest neighbors in the feature space. It does not require an explicit training phase, as it directly uses the labeled training set for prediction by comparing distances, typically Euclidean, to determine similarity [87].

In the context of SOP-based event classification, KNN can classify polarization anomalies by referencing similar known signatures. While it performs well when class clusters are well separated, KNN can struggle with noisy boundaries, irrelevant features, and large datasets due to high computational demands at prediction time. Nonetheless, it serves as a valuable baseline method in this thesis for its simplicity and intuitive behavior.

Artificial Neural Networks (ANNs)

Although classical SL algorithms provide competitive results for many SOP-based monitoring tasks, they may struggle when the data exhibit highly noisy and nonlinear interactions or when feature engineering becomes a bottleneck. Artificial Neural Networks (ANN) address these challenges through hierarchical representation learning, allowing models to automatically extract meaningful structure from minimally processed SOP signatures. The remainder of this subsection presents the ANN models used in this work, including Multi Layer Perceptrons (MLPs) and DL architectures.

Multi-Layer Perceptron (MLP)

The MLP is a feed-forward Neural Network (NN) composed of an input layer, one or more hidden layers with nonlinear activation functions, and an output layer. An MLP with L layers consists of an input layer, $L-1$ hidden layers, and an output layer. Information flows forward through successive layers of neurons, each applying an affine transformation followed by a nonlinear activation function. An MLP is capable of learning complex, nonlinear relationships between input features and output classes through iterative training with backpropagation [88].

For SOP data classification, MLPs offer the flexibility to model subtle and high-order interactions among polarization features, enabling the classification of intricate disturbance patterns. Their probabilistic outputs are also useful for quantifying prediction confidence. However, MLPs often require more data, careful tuning, and computational resources compared to simpler models.

Deep Learning (DL) Models

DL architectures represent a distinct evolution of ANNs, defined primarily by a substantial increase in architectural depth, containing significantly more hidden layers than conventional MLPs. This crucial difference in layer count fundamentally reconfigures the network's function. The layered stack within a DL model is strategically divided: a large portion of the network is dedicated to complex representation learning, followed by fewer, specialized layers responsible for the final task execution, such as classification. This structure enables the hierarchical decomposition of the input data , where initial

layers capture low-level local features, and subsequent deep layers progressively aggregate these into abstract, non-linear, and semantically invariant representations. For SOP data, this deep feature extraction capacity allows the model to autonomously discover optimal high-order feature combinations necessary to discriminate between intricate disturbance classes, bypassing the limitations and labor associated with manual feature engineering required by shallower models [89], [90].

DL encompasses various specialized architectures, including CNNs tailored to exploit specific characteristics inherent in the input data structure. CNNs are a class of DL models particularly well suited for processing array-like data such as images, time-series data, or spectrograms [91], [92]. They operate through convolutional layers that apply learnable filters to capture local patterns in the input, followed by pooling operations that reduce dimensionality while retaining essential information. These layers enable the network to learn spatially or temporally invariant features, making CNNs highly effective for recognizing complex patterns. A typical CNN architecture consists of several convolutional and pooling blocks that extract hierarchical features, which are then passed to fully connected layers for classification through a softmax output.

3.2 Semi-Supervised Learning (SSL) and Unsupervised Learning (USL)

While SL excels when labeled examples of all important events are available, in practice, one often faces limited labeled data or entirely novel disturbances. SSL and USL techniques address these challenges by leveraging partially-labeled or unlabeled data and detecting patterns or outliers without needing a predefined class for every instance. This is highly relevant for the general task of SOP sensing, where many fiber perturbations (especially security breaches) are rare or hard to intentionally induce for labeling. Relying solely on SL could leave the monitoring blind to any event that was not in the training set. Instead, SSL/USL methods aim to detect anomalies through clustering patterns, providing a safety mechanism for unforeseen events.

In SOP sensing, a common SSL scenario is one in which data representing the normal operating state of the network is ample, examples of actual threat are rare. Semi-supervised anomaly detection algorithms can be trained on

data that represents normal operating conditions, learn its profile, and then identify any deviation from that profile as a potential anomaly without needing examples of the anomalies beforehand (i.e., during training).

USL methods go a step further by not even requiring a training phase. They directly analyze the collected SOP data to identify groups of similar polarization patterns and flag any measurements that do not fit into any group as potential anomalies, without requiring any prior knowledge of what those anomalies look like. Both SSL and USL approaches are vital for scalable fiber security monitoring, as they reduce the dependence on labor-intensive labeling of polarization events. Below, we outline some representative SSL/USL techniques and their roles in SOP-based sensing.

One-Class Support Vector Machine (OCSVM)

OCSVM is a well-established SSL algorithm for identifying anomalies. It can be trained on predominantly single-class datasets. It operates under the assumption that most data points in the training dataset represent normal behavior, while anomalies constitute deviations from this dominant pattern. OCSVM works by mapping the input data into a high-dimensional feature space through a kernel function and constructing a decision boundary that encloses the majority of the training data points. The algorithm then seeks to maximize the margin between this boundary and the origin, effectively isolating outliers. During operation, any new data point that lies within the boundary is considered normal, whereas those outside are detected as anomalies [93]. The kernel type and associated hyperparameters, most notably the kernel coefficient γ and the regularization parameter ν, are crucial in determining the model's sensitivity to outliers and its generalization capability. Proper tuning of these parameters ensures that the model captures the underlying structure of normal data while remaining robust to noise and measurement fluctuations.

In the context of SOP-based fiber monitoring, OCSVM is particularly well suited for detecting anomalous polarization behavior without requiring labeled examples of every possible threat: the model is trained exclusively on normal fiber operation data, and any observed SOP patterns that deviates significantly from this learned baseline, such as those induced by fiber tapping or harmful vibrations, are flagged as a potential security breach.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN)

DBSCAN is an unsupervised clustering algorithm that identifies dense regions in the data and separates them from sparse areas that correspond to noise or outliers. Unlike OCSVM, which requires a training phase to learn the boundary of normal behavior, DBSCAN operates directly on the dataset without prior training. It is particularly effective for analyzing high-dimensional or nonlinear data, as it can detect clusters of arbitrary shapes and isolate anomalous samples. The algorithm relies on two main parameters: the neighborhood radius ϵ and the minimum number of points required to form a dense region, $MinPts$. A data point is considered a core point if it has at least $MinPts$ neighbors within the ϵ-radius. Clusters are formed by connecting neighboring core points, while points that do not satisfy this condition are labeled as noise or anomalies [94]. Since DBSCAN operates directly on data samples without prior training, it can be flexibly applied to real-time SOP analysis without an explicit training stage. The selection of ϵ and $MinPts$ has a significant impact on clustering results. A small ϵ can cause many points to be incorrectly classified as noise, while a large value may merge distinct clusters, reducing anomaly detection sensitivity. Similarly, the choice of $MinPts$ determines the granularity of clusters and influences how well DBSCAN differentiates between normal and anomalous patterns. Through careful parameter tuning, DBSCAN can effectively identify abnormal polarization behavior, where anomalies appear as sparse or isolated points in the SOP feature space.

3.3 Performance Metrics and Model Evaluation

Evaluation of ML model performance requires both appropriate metrics that quantify different aspects of predictive accuracy and a sound validation strategy to ensure that reported results are reliable. This section covers both aspects, describing the used classification and clustering metrics as well as the cross-validation strategy used throughout this thesis.

Classification and Clustering Metrics

Confusion Matrix

Confusion matrices are a key evaluation tool for classification models. They compare a model's predicted labels with the true labels to assess classification performance. For binary classification, predicted versus actual class labels are tabulated as:

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positives (TP)	False Negatives (FN)
Actual Negative	False Positives (FP)	True Negatives (TN)

where TPs are correctly predicted positive instances; FNs are actual positives incorrectly predicted as negative; FPs are actual negatives misclassified as positive; and TNs are correctly predicted negative instances.

For multi-class classification with C classes, the confusion matrix has dimensions $C \times C$, with entry M_{ij} representing the number of samples of true class i predicted as class j. The diagonal elements M_{ii} represent correct classifications for each class, while the off-diagonal elements M_{ij} with $i \neq j$ reveal misclassifications. A well-performing classifier will have high values along the diagonal and low values elsewhere, indicating accurate and consistent predictions across all classes.

Accuracy

Accuracy measures the overall proportion of correctly classified instances. It reflects how often the classifier makes the right prediction across both positive and negative classes.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.2}$$

Recall (Sensitivity, True Positive Rate (TPR))

Recall, also known as sensitivity or True Positive Rate (TPR), measures the ability of the classifier to correctly identify all actual positive instances. It reflects how many of the true positives were successfully detected out of all actual positives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.3}$$

Precision (positive predictive value)

Precision quantifies how many of the instances predicted as positive are actually correct. It indicates the reliability of positive predictions made by the classifier.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.4}$$

F1-Score

The F1-score is the harmonic mean of precision and recall, providing a single metric that balances both false positives and false negatives. It is particularly useful in imbalanced classification settings, where relying on accuracy alone can be misleading.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{3.5}$$

False Negative Rate (FNR)

False Negative Rate (FNR), also known as the miss rate, measures the proportion of actual positive instances that the classifier fails to detect. It reflects how often the model misses true events, such as undetected anomalies.

$$\text{FNR} = \frac{FN}{TP + FN} \tag{3.6}$$

Specificity (True Negative Rate (TNR))

Specificity, also known as True Negative Rate (TNR), measures the proportion of actual negative instances that are correctly identified as negative. It reflects the classifier's ability to avoid false alarms by correctly ruling out non-events.

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3.7}$$

Metrics for Clustering and USL Anomaly Detection

In addition to traditional anomaly detection or classification metrics discussed before, SSL and USL methods, particularly those involving clustering or density estimation, require alternative performance measures that do not depend solely on labeled data.

- **Silhouette Score (SS)**

  The Silhouette Score (SS) evaluates the quality of clustering by measuring how similar an object is to its own cluster compared to other clusters. It combines notions of cohesion (intra-cluster similarity) and separation (inter-cluster dissimilarity), with values ranging from -1 to 1. A higher silhouette score indicates more distinct and well-separated clusters, which is particularly valuable when assessing the structural quality of clustering-based anomaly detectors.

- **Adjusted Rand Score (ARS)**

  The Adjusted Rand Score (ARS) quantifies the similarity between the predicted cluster assignments and ground truth class labels, adjusting for chance groupings. It accounts for both correctly clustered and misclustered pairs, providing a more nuanced performance measure. An ARS of 1 indicates perfect agreement, while an ARS near 0 suggests clustering performance no better than random.

These metrics are especially useful in evaluating models where class labels may be sparse or partially unavailable, as is common in SOP-based SSL anomaly detection scenarios.

Model Evaluation Strategy

Beyond the metrics above, robust model evaluation also requires a sound validation strategy to ensure that the reported metrics are reliable and not artifacts of a particular data split. Cross-validation serves this purpose and is described below.

Cross-Validation

Cross-validation is a robust technique for estimating how well an ML model will perform on unseen data by dividing the dataset into multiple subsets

and eliminating potential bias introduced by randomly-drawn fixed training-validation-testing splits. In K-fold cross-validation, the data is split into K equal parts; the model is trained on $K-1$ folds and tested on the remaining fold, repeating this process K times so each fold serves as a test set once. The performance is then averaged across all folds. Common choices include $K = 5$ or $K = 10$, with stratified folds that preserve class distributions. This method offers a more reliable and stable evaluation than a single train-test split, helps assess model variance, and makes efficient use of limited datasets [95].

3.4 Related Work and State of the Art

The convergence of ML with optical fiber sensing represents one of the most rapidly evolving frontiers in optical network research, yet polarization-based physical layer security monitoring remains a noticeably underdeveloped domain. While ML techniques have been transforming Quality of Transmission (QoT) estimation and network performance monitoring since the mid-2010s, their application to SOP-based fiber security is emerging only now, creating significant opportunities for novel contributions in adaptive threat detection and real-world deployment validation.

The seminal demonstration in [96] established that SOP monitoring over transoceanic cables can detect seismic activity and ocean swells using standard telecommunication traffic, treating polarization variations in Stokes parameters as a sensing modality without requiring dedicated infrastructure. The work in [62] extended this paradigm to terrestrial networks, demonstrating 85-day continuous SOP and DAS monitoring over 50 km fiber in Southern California, achieving preemptive traffic loss detection with a 160-second decision boundary. However, both foundational works relied on threshold-based statistical analysis of individual Stokes components rather than ML-driven pattern recognition, and neither addressed security applications. The work in [22] explicitly targeted physical layer security using polarization beam splitters with balanced photodetectors and FFT-based signal processing, but employed simple threshold detection, yielding detection without classification and significant susceptibility to environmental noise. This reliance on manual threshold calibration, characteristic of the pre-ML era of polarization monitoring established by foundational tutorials such as [97] on polarization-related impairment monitoring, represents a fundamental limitation that motivates

the development of adaptive, learning-based frameworks.

The work in [98] surveyed ML techniques for optical performance monitoring and modulation format identification in direct-detection and coherent systems, identifying the need for real-time channel state information but focusing exclusively on transmission impairments rather than malicious interference. Papers such as [99] demonstrated Long Short-Term Memory (LSTM)-based OSNR and nonlinear noise estimation, while [100] achieved joint OSNR and modulation format identification using Deep Neural Networks (DNN) —advances that demonstrate the maturity of ML for performance metrics but underscore the absence of equivalent frameworks for polarization-based security. The tutorial in [101] on ML for failure management in optical networks advanced soft failure detection and localization using Bayesian networks and clustering, yet oriented these capabilities toward network reliability rather than threat identification. Only in 2020 the research in [102] bridged the gap between ML and optical network security, proposing frameworks integrating SL, SSL, and USL for threat detection including jamming and polarization scrambling, a pivotal contribution that nonetheless focused on specific service disruption attack methods rather than general physical layer tampering detectable through polarization signatures.

Traditional physical layer security methods face inherent limitations that polarization-based sensing could address. Phase-sensitive OTDR (Φ-OTDR) systems, pioneered in [48] for distributed fiber intrusion detection, achieve 10–20 meter spatial resolution through coherent Rayleigh backscattering but require dark fiber, cannot operate on lit telecommunications links, and suffer from coherent fading and laser frequency drift, causing high false alarm rates. DAS technology, reviewed comprehensively in [103], offers impressive capabilities including 100+ km range and mHz-to-kHz frequency response, but commercial interrogators cost $50,000–$200,000, exhibit environmental sensitivity to wind, rain, and temperature variations, and present a fundamental trade-off between range and resolution. The foundational analysis in [30] identified bent fiber taps as the most easily deployed eavesdropping mechanism, yet demonstrated that OTDR cannot reliably detect subtle taps, while sophisticated evanescent coupling techniques cause minimal signal disturbance altogether. These limitations create a compelling case for complementary SOP-based approaches that can potentially operate on active WDM networks, require no additional hardware beyond the coherent transponders already deployed in

modern optical networks, or at most lower-cost commercial polarimeters or optical analyzers, and exhibit high sensitivity to the mechanical perturbations associated with tampering.

The emerging literature on ML-driven SOP sensing remains sparse but demonstrates proof-of-concept feasibility. The work in [104] applied Vision Transformers to spectrograms derived from SOP measurements, achieving 97% diagnostic accuracy with 6 ms temporal localization on a 2600 km bidirectional link—yet treated overlapping events as single combined signatures rather than developing multi-label classification. Recent work on resilient anomaly detection [105] evaluated KNN, RF, XGBoost, and DT for SOP angular speed analysis to detect malicious vibrations, overlapping disturbances, and fiber tapping, but acknowledged that laboratory datasets do not capture real-world environmental variability.

The intersecting gaps—the absence of ML-driven polarization security frameworks despite mature ML for performance monitoring, the limitations of OTDR/DAS requiring complementary approaches, the lack of comprehensive labeled datasets for polarization signatures, limited real-world validation on operational networks, and the need for adaptive frameworks integrating SL, SSL, and USL for unknown threat detection—define the research opportunity that this thesis addresses. The following chapters develop novel polarization feature extraction and ML classification architectures validated on modulated signals, presenting comprehensive experimental validation on operational network infrastructure with diverse fiber events, and establishing foundations for cross-link generalization and overlapping event detection that current literature has not achieved.

3.5 Chapter Summary

This chapter introduced the ML principles underpinning the detection and classification of SOP-based fiber disturbances. Section 3.1 detailed SL methods, covering classical classifiers, bagging and boosting ensembles, and advanced variants such as RF, ET, GB, XGBoost, and HGB. The SL models are further extended in this section by ANN models, including MLPs, DL architectures, and CNNs, highlighting their ability to learn hierarchical representations of polarization dynamics.

Section 3.2 presented SSL and USL approaches for scenarios with limited

or unlabeled data, focusing on OCSVM for anomaly detection and DBSCAN for clustering-based event discovery. Section 3.3 summarized the performance evaluation tools used in this thesis, including classification metrics, clustering metrics, and cross-validation strategies. Finally, Section 3.4 situated these contributions within the broader research landscape, surveying the state of the art in ML-based optical network security monitoring and identifying the key gaps that this thesis addresses.

Altogether, the methods introduced in this chapter provide the analytical foundation for the monitoring framework developed in the following chapters.

CHAPTER 4

ML-Driven SOP-Based Fiber Sensing: Framework Design and Data Collection

This chapter establishes the methodological and experimental foundations for the research presented in this thesis. It begins by introducing the proposed ML-driven SOP-based fiber sensing framework in Section 4.1 and describing its network telemetry and intelligent analysis components. Section 4.2 provides the specifics of SOP signature acquisition, covering NPSV feature extraction and the applied DSP methods. Section 4.3 then describes the data acquisition campaigns, presenting the extensive datasets collected from controlled laboratory environments and operational field links.

4.1 ML-Driven SOP-based Fiber Sensing Framework

Figure 4.1 illustrates the high-level architecture of our proposed monitoring and learning pipeline for ML-driven SOP-based fiber sensing. The system comprises two coordinated layers: *network telemetry*, detailed in this section, and *ML analysis*, covered in the subsequent chapter.

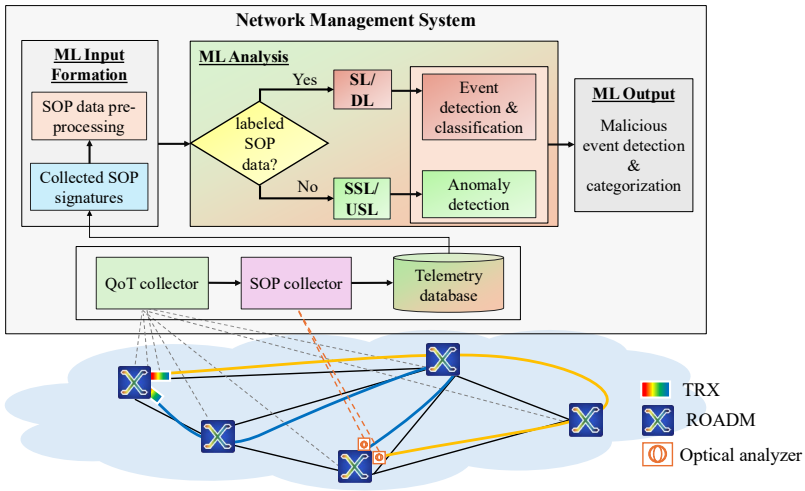


Figure 4.1: The fiber sensing framework for SOP data collection and ML analysis.

The framework is designed to seamlessly integrate with standard optical network architectures, leveraging the optical fiber infrastructure as a sensor. At the physical layer (in the bottom of Figure 4.1), the network consists of a WDM transmission system comprising ROADMs, amplifiers, and Transceivers (TRXs). To enable sensing capabilities, SOP data collection is performed via optical analyzers, either as standalone units placed at critical monitoring nodes or as embedded functionalities within modern coherent transceivers. These devices continuously monitor the Stokes parameters of lightpaths traversing the network, capturing the instantaneous polarization state of optical signals. Simultaneously, a QoT Collector gathers standard optical performance metrics (e.g., optical power, OSNR, Bit Error Rate (BER)) directly from the network elements. The collected raw polarization data is aggregated via an SOP Collector module and, together with the standard performance telemetry data, stored within a centralized Telemetry database within the Network Management System (NMS). Before analysis, the raw SOP data undergoes a pre-processing stage (detailed in Section 4.2) to extract the normalized NPSV and generate spectral SOP signatures, transforming complex raw SOP data into a format suitable for ML algorithmic processing.

The core intelligence of the framework resides in the ML Analysis module

(central part of Figure 4.1), employing a bifurcated logic flow based on data availability. It leverages pre-trained SL and DL models for precise event classification when ground truth labels exist, while defaulting to SSL and USL techniques for anomaly detection when labeled data is scarce or unavailable. The final output of this pipeline is actionable intelligence, categorizing events as harmful or benign, which is fed back to the network management system.

The following sections detail the specific SOP signatures analysis and the data acquisition campaigns used to validate this framework.

4.2 SOP Signature Acquisition

Our polarization analyzer instrument is a commercial polarization sensing module (a "black box") capable of measuring variations in all three Stokes polarization components (S_1, S_2, S_3). This device uses a carefully designed arrangement of passive optical components to project the signal onto different polarization bases. As established in Chapter 3, physical events alter the fiber's birefringence, causing the SOP to traverse specific, rapid trajectories on the Poincaré sphere that correspond to the physical nature of the disturbance. The design of experiments was guided by the principle of capturing a wide spectrum of events that optical fibers might encounter in the field, ranging from benign environmental fluctuations to malicious security threats. For each event, we recorded high-resolution time-series SOP data as it evolved on the Poincaré sphere.

To effectively train ML models, these SOP raw trajectories must be transformed into a discriminative feature set that quantifies the intensity and spectral characteristics of the event. The mathematical formulation and signal processing steps for this transformation are detailed in the following subsections.

Numerical Polarization State Variation (NPSV) Data

A starting point of this thesis is the question of how to analyze the collected SOP data in a way that is suitable for ML processing and preserves the physically meaningful information about fiber disturbances. Instead of collecting three-dimensional SOP trajectories, our optical analyzer directly measured the normalized NPSV metric, which quantifies the angular displacement of

the SOP trajectories on the Poincaré sphere between consecutive time samples. In other words, at each time slot t, the value NPSV_t serves as a scalar indicator of the magnitude of SOP variation between two adjacent sampling points, i.e., during the interval $[t-1, t]$. To quantify this, we define the polarization intensity (the norm of the Stokes vector) at time instance τ as $S_{0,\tau} = \sqrt{S_1^2(\tau) + S_2^2(\tau) + S_3^2(\tau)}$. The normalized polarization magnitude at time t is then given by:

$$A_t = \frac{S_1^2(t) + S_2^2(t) + S_3^2(t)}{S_{0,t}} \tag{4.1}$$

and likewise for the previous sample:

$$A_{t-1} = \frac{S_1^2(t-1) + S_2^2(t-1) + S_3^2(t-1)}{S_{0,t-1}} \tag{4.2}$$

The resulting normalized polarization state variation is computed as:

$$\mathrm{NPSV}_t = A_t - A_{t-1}. \tag{4.3}$$

This transformation converts the complex three-dimensional polarization dynamics into a one-dimensional time series that effectively captures the intensity and temporal characteristics of disturbances.

Digital Signal Processing (DSP) of Raw NPSV Data

To enhance the discriminative potential of the raw NPSV data as the input values for ML models, we applied the following signal processing techniques to the NPSV time series. Starting from the raw normalized NPSV data containing M samples, the data are first divided into time-slot segments (time windows) of length T. Then, each segment undergoes FFT using a Hamming window [106] to obtain its frequency-domain representation with 512 spectral bins, generating time-frequency spectrograms (visualized as "waterfall" plots). The resulting $(\frac{M}{T})\times 512$ dataset referred to as the *SOP signature* for each specific event captures the spectral magnitude of polarization variations over time and constitutes the fundamental input to all subsequent ML-based detection and classification models developed in this thesis. These spectrograms demonstrate that different physical events imprint distinct polarization

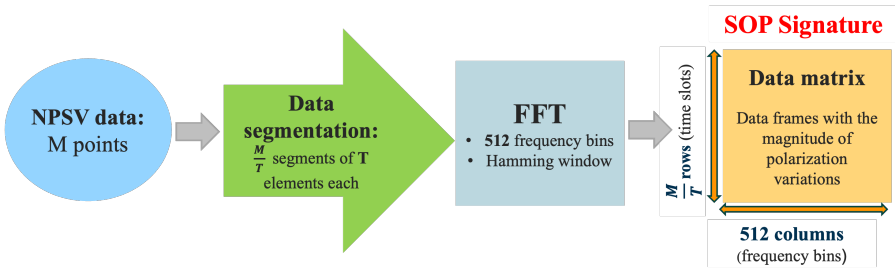
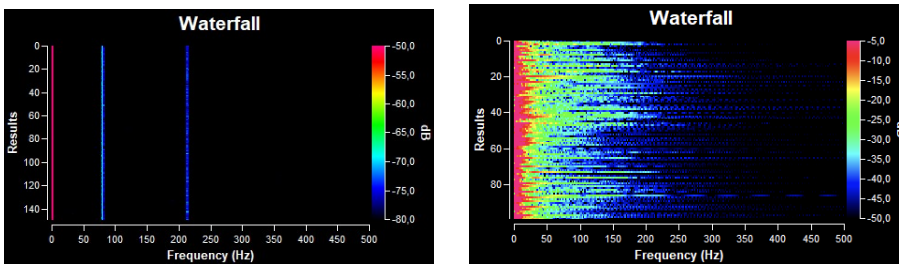


Figure 4.2: DSP pipeline: from raw normalized NPSV data to SOP signatures used for ML analysis.

signatures.



(a) Harmful vibration (80 Hz with 210 Hz overtone)

(b) Eavesdropping event (fiber bending)

Figure 4.3: Representative waterfall spectrograms derived from SOP signatures.

For instance, as depicted in Figure 4.3, a 80 Hz vibration containing an overtone at 210 Hz creates a prominent spectral line at 80 Hz (Figure 4.3(a)), whereas an eavesdropping event results in a broadband spectral distribution induced by fiber bending (Figure 4.3(b)). Such contrasting spectral behaviors form the basis for distinguishing between event types and for training ML models to detect and classify physical-layer anomalies.

4.3 Data Collection

High-quality, representative data form the cornerstone of any ML endeavor. At the beginning of this research, however, a critical gap became evident: the field lacked any comprehensive dataset of polarization signatures capturing a diverse range of fiber disturbance events. While the concept of using SOP variations for sensing had been explored theoretically and demonstrated in limited contexts, there was insufficient empirical data to train and validate ML models over a diverse range of event types and fiber configurations. This created the first major research challenge: *how to generate a diverse, labeled dataset that captures the characteristic polarization responses to different kinds of fiber perturbations under controlled conditions where ground truth is known with certainty?* The absence of such data represented a fundamental impediment to progress in the field. Without sufficient training examples, ML models cannot learn the discriminative patterns that distinguish different event types. Without diversity in the training data, models cannot generalize to new scenarios or fiber types. Without careful experimental control, the ground truth labels may be unreliable, leading to models that learn spurious correlations rather than genuine physical phenomena. Addressing this data gap, therefore, required not just collecting measurements but designing and implementing an experimental methodology aimed at generating representative data suitable for ML analysis of SOP variations.

To address this fundamental need, we designed and implemented a comprehensive experimental testbed in the laboratory.

Controlled Laboratory Dataset (LDS)

We systematically collected three datasets from the controlled laboratory environment. These datasets emulated a wide range of distinct event scenarios across three different fiber cable types: bare fiber, indoor cable, and FOCS. The event categories included normal operation, serving as the baseline, low-amplitude ambient vibrations representing environmental noise, and various harmful events such as intentional fiber bends for simulating eavesdropping taps, and strong mechanical vibrations at different frequencies for emulating construction or excavation activities.

Controlled Laboratory Dataset 1 (LDS1): 13-Class Laboratory Data for Bare, FOCS, and Indoor Cable Types

LDS1 is the first dataset collected in this thesis and serves as its foundational experimental contribution, providing the empirical foundation needed to investigate RQ1 and RQ2 and establish whether SOP signatures can reliably distinguish a broad range of normal, non-harmful, and potentially harmful fiber events for multiple cable types. It also underpins the investigation of RQ5 by providing a rich labeled dataset that can be repurposed for SSL and USL anomaly detection by removing the class labels, thereby simulating a deployment scenario in which the system has no prior knowledge of event types. The SL experiments on LDS1 are reported in **Paper A**, and the SSL/USL experiments are reported in **Paper F**. Full details of the experimental setup and testbed schematic can be found therein.

To form LDS1, we emulated thirteen event scenarios representing normal operation, harmful and non-harmful vibrations, eavesdropping, and overlapping events. We used a controlled testbed with a Continuous Wave Distributed Feedback (CW-DFB) laser emitting at 1310 nm with polarization-maintaining properties, operated at constant power and temperature. The 2 km transmission line comprised a series of FOCS (denoted as *fcs*) cable, indoor cable (*id*), and bare SM G.675 bend-insensitive fiber (*br*), enabling the resulting SOP dynamics to reflect a broad range of practical deployments. Each trace was sampled every millisecond over 20 minutes, producing 1.2 million samples per event. The data were segmented into 1,200 time windows (1,000 samples each), and a 512-bin FFT was applied, yielding 1200×512 arrays SOP signatures for each event. The scenarios included in LDS1 dataset are:

- *Baseline normal* (denoted as *rlx*): relaxed fiber without deliberate mechanical interaction; used to characterize nominal polarization drift and serve as a reference for anomaly detection.

- *Eavesdropping via bending*: fiber bent over 10 mm diameter rod with 4 mm bend radius and 25° bend angle to emulate light extraction via coupler, applied to FOCS (denoted as *b-fc*) and bare fiber (*b-bf*).

- *Non-harmful vibrations*: benign, routine mechanical oscillations occurring at 130 Hz and 155 Hz. These sinusoidal vibrations are generated by a piezoelectric actuator and transmitted through the cable's inner layer

to simulate sources such as fans or nearby street traffic characterized by approximately 7,000 Revolutions Per Minute (RPM) and 9,000 RPM, respectively. Testing was applied to FOCS (denoted as *n-v-fc*), indoor (*n-v-id*), and bare fiber (*n-v-bf*).

- *Harmful vibrations*: sinusoidal acoustic oscillations at 80 Hz, generated by a loudspeaker positioned 5 cm from the bare fiber membrane. This simulates an excavator machine operating at 4,800 RPM (precursor to damage) and was applied to FOCS (denoted as *an-v-fc*), indoor cable (*an-v-id*), and bare fiber (*an-v-bf*).

- *Overlapping events on indoor cable*: combined scenarios testing simultaneous benign and malicious events: (i) eavesdropping with 130 Hz vibration (denoted as *b-n-v-id*), (ii) eavesdropping with 80 Hz vibration (*b-an-v-id*), (iii) eavesdropping with dual 130 Hz and 80 Hz vibrations (*b-dl-v-id*), and (iv) relaxed fiber with dual 130 Hz and 80 Hz vibrations (*wb-dl-v-id*).

Controlled Laboratory Dataset 2 (LDS2): 5-Class Laboratory Data for Indoor Cable

LDS2 was collected to address RQ1 and RQ2 by supporting the development of a focused ML classifier capable of distinguishing harmful from non-harmful perturbations on an indoor fiber link. Unlike LDS1, which spanned three cable types and prioritized breadth of events, LDS2 focuses exclusively on the indoor cable to develop a more specialized model suited for direct deployment on indoor fiber links. It also provides the dataset needed to investigate RQ4 by enabling evaluation of whether the classifier trained under controlled conditions retains its detection accuracy during live monitoring of the indoor fiber link. The event notation in LDS2 follows the naming convention adopted in **Paper B** (*nhrmf*, *hrmf*) and differs from LDS1 (*n-v*, *an-v*) as the two datasets were developed independently; both conventions refer to the same physical distinction between non-harmful and harmful vibration events.

The transmission line consisted of a 1 km coupling fiber connected to an indoor cable segment and followed by a 20 km fiber spool, resulting in a total length of 21 km. Full details of the experimental setup can be found in **Paper B**. The data include five distinct event categories: baseline or normal operation (denoted as *rlx*), non-harmful vibration at 140 Hz (*nhrmf_140Hz_vb*)

generated by a piezoelectric actuator, eavesdropping via bending (*eav*), harmful vibration at 80 Hz (*hrmf_80Hz_vb*) generated by a loudspeaker to simulate excavation machinery, and a combined scenario involving both harmful and non-harmful frequencies vibrations at 80 Hz and 140 Hz (*hrmf_80Hz_140Hz_vb*). Data acquisition involved recording each event for 20 minutes at 1 ms intervals, yielding 1.2 million samples. The final dataset consists of SOP spectral signatures organized into 1200×512 arrays for each event class.

Controlled Laboratory Dataset 3 (LDS3): 14-Class Complex Mechanical Vibration Dataset over Bare Fiber and Patch Cable

LDS3 was collected to provide the empirical basis for addressing RQ1 and RQ2 under more realistic and challenging disturbance conditions, and to enable investigation of RQ8: establishing whether SOP-based ML models can detect and distinguish complex overlapping and mixed-frequency mechanical vibration patterns that occur simultaneously on the fiber. The key motivation for this dataset is that real-world optical fiber installations are routinely exposed to broadband background noise, generated by heavy vehicles, trains, or other persistent machinery passing near the installation, which produces highly variable, overlapping SOP signatures that are difficult to disentangle from deliberate security events. Rather than attempting to capture such naturally occurring events directly, which is operationally impractical, we simulated these conditions using two pseudo-random broadband vibration patterns and combined them systematically with known tampering events. The notation in LDS3 follows the convention established in **Paper H**, where complex vibrations are denoted with A and B, cable types as *br* (bare fiber) and *pc* (patch cable), and tampering events as *eav*, *sb*, and *80vb*.

The experimental testbed utilized a stabilized 1310 nm CW-DFB laser to inject polarized light into a 1 km coupling fiber. This was followed by a sensing region where controlled perturbations were applied to either a bare fiber segment or a patch cable, before propagating through a 20 km fiber spool for a total transmission length of 21 km. To emulate realistic environmental noise, two distinct pseudo-random complex vibration patterns (spanning 0–2000 Hz) were introduced: *Complex A* as a 10-minute sequence and *Complex B* as a 20-minute sequence. These complex background vibrations were combined with specific tampering scenarios, including *eavesdropping* (4 mm radius bend), *soft bending* (2 cm radius), and *harmful vibrations* (80 Hz loudspeaker tone).

Each experiment was sampled at 0.5 ms intervals over 10–20 minute durations, yielding SOP signatures with dimensions of 1200×512 (for Complex A) and 2400×512 (for Complex B). Full details of the experimental setup and classification results can be found in **Paper H**.

The resulting fourteen distinct event classes are summarized in Table 4.1.

Table 4.1: Summary of LDS3 event classes. The dataset captures complex background vibrations (A and B) combined with specific physical perturbations applied to either bare fiber (br) or patch cable (pc).

Event Combination	**Abbreviation**
Complex A/B on bare fiber	A_{br}, B_{br}
Complex A/B + eavesdropping on bare fiber	A_{br+eav}, B_{br+eav}
Complex A/B + 80 Hz vibration on bare fiber	$A_{br+80vb}$, $B_{br+80vb}$
Complex A/B + soft bending on bare fiber	A_{br+sb}, B_{br+sb}
Complex A/B on patch cable	A_{pc}, B_{pc}
Complex A/B + soft bending on patch cable	A_{pc+sb}, B_{pc+sb}
Complex A/B + eavesdropping on patch cable	A_{pc+eav}, B_{pc+eav}

Real-World Dataset (RDS)

While the controlled laboratory experiments provide a fundamental baseline for characterizing isolated event signatures, validation in operational environments is indispensable for developing robust, deployment-ready ML models. Laboratory settings, by design, minimize external interference; however, they cannot fully replicate the complex, stochastic nature of deployed optical networks. Real-world fibers are subject to continuous, unpredictable SOP fluctuations driven by environmental thermal gradients, mechanical vibrations from civil infrastructure (e.g., vehicular traffic, heavy machinery), and acoustic coupling. Furthermore, signal propagation over field-deployed links introduces cumulative physical impairments, including attenuation, CD, and connector-induced noise, which are absent or negligible in short-span laboratory spools. To bridge the gap between theoretical feasibility and operational reality, the following datasets capture these field conditions, ensuring the proposed detection frameworks are capable of generalizing to the noisy dynamic environments of field-deployed networks.

Real-World Dataset 1 (RDS1): 7-Class Field Data over Short and Metro-Scale Links

RDS1 was collected to provide the empirical foundation for investigating RQ1, RQ2, and RQ3 by evaluating whether ML models retain their detection and classification performance when applied to SOP signatures collected from field-deployed fibers carrying live traffic, and by enabling benchmarking of a set of SL classifiers under real-world noise conditions. The collected RDS1 includes seven distinct event scenarios spanning normal operation, benign environmental perturbations, and harmful or malicious activities using the OpenIreland [25] testbed infrastructure at Trinity College Dublin (TCD). A key design decision was to collect data from two links of different lengths simultaneously, enabling direct assessment of how link length and accumulated environmental noise affect classification performance.

The experimental setup utilized a linearly polarized CW-DFB laser emitting at 1310 nm with a spectral width of 0.5 nm to inject optical signals into field-deployed fibers. The transmission lines consisted of two loop-back configurations: a short-link installation (0.15 km single-direction) connecting two TCD locations via an external public road, and a long-link installation (10.5 km single-direction) connecting TCD to Dublin City University (DCU) to emulate a metro-scale scenario. Each experiment lasted 20 minutes with 0.5-ms sampling, resulting in approximately 2.4 million data points, yielding SOP signatures each with $4,800$ samples and 512 frequency bins. Full experimental details are reported in **Paper C**.

The notation in RDS1 uses link length as a suffix (*0.15km*, *10.5km*) to distinguish otherwise identical event types collected on the two links. The complete set of classes is summarized below:

- *Relaxed fiber* (*rlx_0.15km*, *rlx_10.5km*): Baseline scenarios representing normal operation without intentional disturbances.

- *Soft bending* (*sbd_0.15km*, *sbd_10.5km*): Gentle bending of the fiber used to emulate benign handling of the fiber during maintenance or minor environmental stress.

- *Eavesdropping* (*eav_0.15km*): Simulated tapping by applying a controlled bend on the short 0.15 km link to assess detectability of covert signal extraction attempts.

- *Harmful vibrations* (*80vb_0.15km*, *80vb_10.5km*): 80 Hz frequency mechanical vibration generated by a loudspeaker to replicate ground-borne activity such as excavation near buried fibers.

This dataset reflects realistic polarization dynamics where environmental noise and link-specific characteristics influence the measured SOP trajectories.

Real-World Dataset 2 (RDS2): 14-Class Field Data over Short and Metro-Scale Links

RDS2 was collected to provide the empirical foundation for investigating RQ1, RQ2, and RQ3 under more complex and demanding real-world conditions than RDS1. The motivation for expanding beyond RDS1 was that the seven classes in RDS1 did not fully capture the diversity of threats and edge cases encountered in operational networks; in particular, shielded fiber scenarios, bursty versus continuous eavesdropping profiles, and overlapping vibration patterns had not yet been explored in real-world deployments. RDS2 therefore expands the event space, providing a richer and more diverse collection of field-deployed polarization signatures. The notation in RDS2 replaces the distance-based suffixes used in RDS1 (*0.15km*, *10.5km*) with descriptive suffixes (*_short*, *_long*). Additional prefixes *dtc_* and *shld_* are introduced to denote detector-proximate and shielded vibration variants, respectively. Full details of the experimental setup and results are reported in **Paper D**.

The RDS2 dataset preserves the data dimensions of RDS1 (4800×512 SOP signatures). The classes collected for the short (0.15 km) and the long (10.5 km) link are detailed below:

- *Baseline and Maintenance*: Identical to RDS1, this includes relaxed fiber (*rlx_short*, *rlx_long*) and soft bending (*sbd_short*, *sbd_long*) to represent normal operation and routine handling.
- *Harmful vibrations*: The 80 Hz excavator-simulating vibration applied to both links (*80vb_short*, *80vb_long*). For the short link, two new variants were introduced to test detection limits: vibration applied immediately before the detector (*dtc_80vb_short*) and vibration applied to a fiber segment covered by a protective shield (*shld_80vb_short*).
- *Eavesdropping*: Malicious tapping (4 mm radius, 25° angle) captured under two distinct temporal profiles: bursty attacks (10 s active, 20 s

relaxed) for the short link (*eav_short*) and continuous eavesdropping for the long link (*eav_long*).

- *Complex overlapping events*:
  - *Multi-frequency vibrations*: Simultaneous harmful (80 Hz) and benign (155 Hz) vibrations applied to the long link (*80vb_155vb_long*), along with a shielded variant (*shld_80vb_155vb_long*).
  - *Eavesdropping with vibrations*: A high-complexity scenario where continuous eavesdropping occurs simultaneously with both 80 Hz and 155 Hz vibrations, collected for the short (*eav_80vb_155vb_short*) and the long link (*eav_80vb_155vb_long*).

Real-World Dataset 3 (RDS3): Modulated vs. Unmodulated Signatures over Metro Network

RDS3 was collected to provide the empirical foundation for investigating RQ1 and RQ2 under signal modulation conditions, and specifically to enable investigation of RQ6, establishing whether SOP-based ML sensing remains effective when applied to modulated channels carrying live traffic rather than unmodulated signal probes. All prior datasets in this thesis used unmodulated CW signals, which provide clean polarization trajectories but do not reflect the operating conditions of real coherent optical networks. In practice, high-speed modulated signals such as DP-16QAM introduce rapid polarization fluctuations at the symbol rate that are effectively averaged out by the optical analyzer, potentially altering the statistical properties of the measured NPSV and raising the question of whether event-induced SOP signatures remain distinguishable. RDS3 directly addresses this gap by collecting paired modulated and unmodulated signatures under identical physical disturbances on the same fiber link simultaneously, enabling a controlled one-to-one comparison of SOP dynamics for both signal modalities.

The experimental setup utilized a 63.4 km fiber ring with 6 ROADMs within the HEAnet metro network, connected to the OpenIreland testbed. Two distinct optical channels were injected into the link: (i) a 200 Gbps DP-16QAM modulated signal generated by a coherent transceiver, and (ii) an unmodulated External Cavity Laser (ECL) signal. Both signals traversed the 63.4 km field ring followed by a 40 km spool of G.652 fiber in the laboratory, where controlled physical disturbances were applied.

Each event was recorded for 15 minutes with a sampling interval of 0.5 ms, resulting in approximately 1.8 million samples per trace. The NPSV data were processed using an FFT with a Hamming window on 0.5-second segments, yielding SOP spectral signatures with dimensions of 3600×512.

The notation appends *_m* and *_u* suffixes to each event class to distinguish modulated (λ_m) and unmodulated (λ_u) channel measurements of the same physical event. Full details of the experimental setup and classification results are reported in **Paper I**. The dataset comprises eight distinct classes, representing four physical event scenarios applied to the two signal modalities (modulated and unmodulated):

- *Relaxed fiber* (*rlx_m*, *rlx_u*): Baseline measurements capturing the intrinsic polarization drift and environmental noise of the 63.4 km metro link without induced disturbances.
- *Soft bending* (*sbd_m*, *sbd_u*): Simulation of routine maintenance, performed by manually bending the fiber to a 2 cm radius at 10-second intervals to mimic handling by data center technicians.
- *Eavesdropping* (*eav_m*, *eav_u*): A malicious tapping attempt simulated by bending the fiber to a 4 mm radius with a 25° angle.
- *Harmful vibration* (*80vb_m*, *80vb_u*): A mechanical disturbance at 80 Hz generated by a loudspeaker placed 2–4 cm from the fiber.

Real-World Dataset 4 (RDS4): Multi-Band and Multi-Link Generalization Data

RDS4 was collected to provide the empirical foundation for investigating RQ1 and RQ2 across heterogeneous system configurations, and specifically to enable investigation of RQ7, whether SOP-based ML classifiers trained on one spectral band and fiber link can generalize to a different band and link without retraining. This is a critical practical question since polarization dynamics in optical fibers are influenced by wavelength-dependent physical-layer effects such as birefringence and PMD, which can cause the SOP to evolve differently across spectral bands. As a result, features learned in one system may not be directly transferable to another, raising concerns about the applicability of trained models across diverse deployment environments.

RDS4 dataset aggregates SOP signatures from two distinct physical systems, enabling the assessment of whether polarization features learned in one environment can transfer to another without retraining. The notation appends numeric subscripts (*_1* for System 1, *_2* for System 2) to distinguish otherwise identical event types collected on the two systems. Full details of the experimental setup and generalization results are reported in **Paper G**.

The experimental setup utilized the OpenIreland testbed to access two disparate optical links:

- **System 1 (O-band Dark Fiber)**: A 21 km round-trip dark fiber link using standard G.652 SM fiber. The optical source was a CW-DFB laser operating in the O-band.
- **System 2 (C-band Live Metro Network)**: A 63.4 km round-trip path over the live HEAnet metro network as discussed in RDS3 section. The optical source was an ECL operating in the C-band (192.8–193.2 THz).

Data acquisition followed the established protocol of 20-minute recordings with a 0.5 ms sampling interval. The SOP traces were processed into spectral signatures of size 4800×512. The dataset consists of three common event classes collected for both systems:

- *Relaxed fiber* (rlx_1 for O-band, rlx_2 for C-band): Baseline measurements capturing the distinct environmental noise profiles of the dark fiber (System 1) and the active metro network (System 2).
- *Soft bending* (sbd_1, sbd_2): Benign handling events simulated by bending the fiber to a 2 cm radius at 10-second intervals.
- *Eavesdropping* (eav_1, eav_2): Malicious tapping simulated via a clip-on coupler (4 mm radius, 25° angle).

4.4 Chapter Summary

This chapter established the foundational methodology for ML-driven SOP-based fiber sensing. Section 4.1 presented the proposed framework architecture, integrating network telemetry with ML analysis for actionable anomaly

detection in optical networks. Section 4.2 detailed the SOP signature acquisition process, introducing the NPSV description and applying DSP techniques to transform raw polarization data into spectral signatures suitable for ML processing. Section 4.3 described comprehensive data acquisition campaigns spanning three controlled laboratory datasets (LDS1–LDS3) and four real-world datasets (RDS1–RDS4), systematically capturing diverse event scenarios across multiple fiber types, link configurations, and network conditions. These datasets encompass normal operation, benign perturbations, harmful vibrations, eavesdropping attempts, and complex overlapping events, collected from both short-span laboratory testbeds and an operational metro network. Together, these datasets constitute a comprehensive empirical foundation that enables the training and evaluation of ML-based detection and classification models, as presented in the following chapters.

CHAPTER 5

ML Analysis of Polarization Signatures: Performance and Contributions

This chapter presents the main contributions of this thesis, organized according to the research trajectory that progressively addressed the challenges of developing an intelligent, ML-driven optical network monitoring framework based on the analysis of SOP signatures described in Chapter 4. Building on the datasets and feature engineering methodology established there, the research evolved from controlled laboratory experiments to real-world field deployments, systematically exploring SL (including DL), SSL, and USL approaches within the framework depicted in Figure 4.1. The contributions are organized into four sections corresponding to distinct phases of the research and associated with specific publications, collectively addressing the eight RQs defined in Chapter 1.

5.1 SL for Multi-Class Event Detection and Classification

Using the labeled LDS1 and LDS2 datasets described in Chapter 4, this section investigates whether ML algorithms can automatically learn to recognize and classify the different event types from their polarization signatures. This stage addressed **RQ2 and RQ3**: *determining which ML models are most effective for this classification task and evaluating their ability to distinguish between various types of perturbations, including differentiating malicious attacks from accidental disturbances.*

We first approached this problem as a supervised multi-class classification problem, treating each of the 13 event scenarios in LDS1 as a unique class. Data preparation involved randomly dividing each SOP collection into a 70% training subset (840 points) and a 30% testing subset (360 points). By ensuring equal representation of all scenarios, we generated a balanced training dataset of $10,920$ samples and a testing set of $4,680$ samples. To determine the most effective classifier for this specific feature space, we benchmarked a comprehensive suite of SL algorithms from the Scikit-learn library [24]. The evaluated methods represented a wide range of ML paradigms: ensemble-based approaches including RF, ET classifier, XGBoost, and standard GB; kernel-based models such as SVM; linear methods including LR and LDA; distance-based techniques like KNN; and DT classifiers. Our systematic grid search and cross-validation experiments revealed that ensemble methods consistently yielded superior performance.

Figure 5.1 summarizes the performance metrics of the three top-performing classifiers identified during this benchmarking process. Consistent with our findings on ensemble superiority, XGBoost emerged as the most robust model, achieving an overall accuracy of 92.3% and an F1-score of 0.922. RF and GB followed closely, delivering accuracies of 90.9% and 90.4%, respectively. These results underscore the strong discriminative capability of boosting and bagging architectures in capturing the complex, non-linear patterns inherent to the frequency-domain polarization signatures.

To gain a deeper view of classifier performance, Figure 5.2 presents the confusion matrix obtained by the XGBoost model. This matrix details the exact number of correctly classified samples out of the 360 test instances allocated to each class. The matrix illustrates the classifier's ability to separate benign

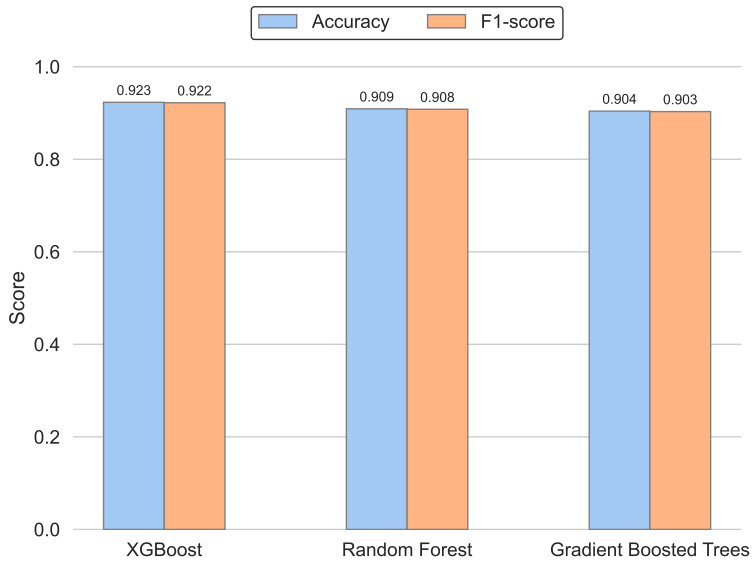


Figure 5.1: Accuracy and F1-score obtained by the three best-performing ML classifiers across the 13 event classes of LDS1.

scenarios (e.g., relaxed fiber, non-harmful vibrations) from covert intrusions (eavesdropping bends) and safety-critical disturbances (harmful vibrations). Most classes are recognized with over 348 out of 360 samples ($> 96\%$ accuracy), underscoring the discriminative strength of polarization-derived features. Some challenges remain in disentangling overlapping indoor scenarios, such as combinations of bending with vibrations, where signature interference leads to modest confusion between neighboring classes.

As the next step, recognizing that, in practical deployments, network operators are primarily concerned with identifying harmful events that require immediate action, we refined our approach to focus on the categorization of event severity in indoor cables, which are widely deployed in practical environments such as data centers, enterprise networks, and access network termination points, where real-time detection of harmful events is operationally critical. While LDS1 was designed for broad multi-class benchmarking over three cable types, LDS2 was purpose-built for this specific task: a focused 5-class dataset collected over a 21 km indoor cable link, designed to train a

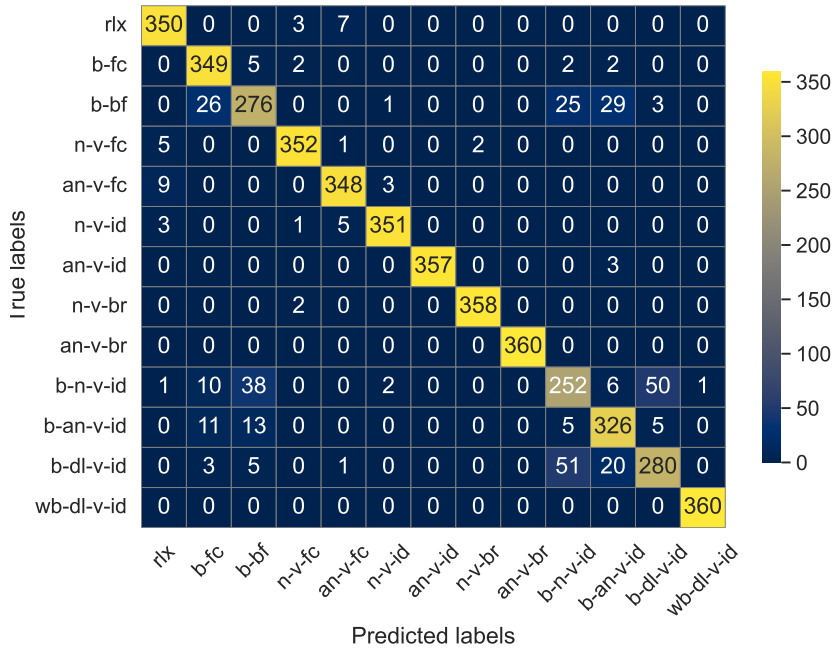


Figure 5.2: Confusion matrix for the XGBoost classifier across the 13 polarization-based event classes of LDS1.

classifier directly suitable for live network deployment. In this context, the HGB classifier proved particularly effective, achieving nearly 98% accuracy in discriminating harmful disturbances from non-harmful events using LDS2.

To validate the operational readiness of these models, we conducted experiments simulating real-time monitoring scenarios. We first trained the HGB model on LDS2 and then tested its ability to detect and classify new, live occurrences of the same events.

Figure 5.3 illustrates a representative example of the live prediction over one-minute monitoring interval. The model accurately recognized an eavesdropping event on the indoor cable, maintaining a detection confidence of 98.2% throughout the test period, while transient background fluctuations were briefly misclassified as relaxed state (1.8%). These results demonstrate the robustness and temporal stability of the proposed ML framework in real-

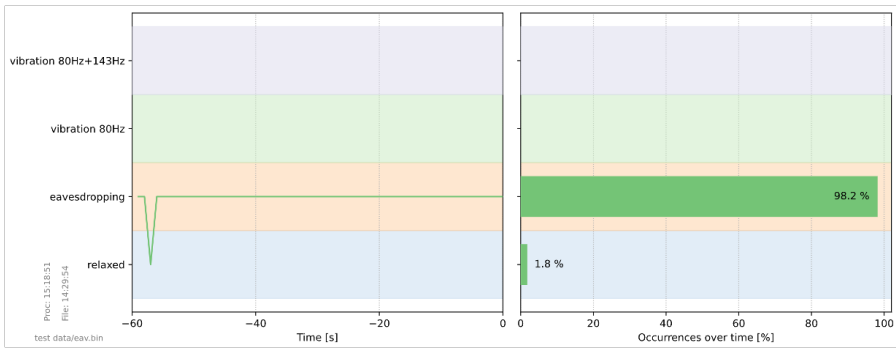


Figure 5.3: Live one-minute prediction demo using the pre-trained HGB model on LDS2 dataset. The model maintained a detection accuracy of 98.2% for an ongoing eavesdropping event, confirming its real-time reliability and robustness.

time operation, validating its potential for continuous fiber security monitoring in practical network environments. The model successfully identified and categorized events with minimal latency, demonstrating its suitability for integration into automated monitoring systems.

This achievement addresses **RQ4**, confirming that *models trained on specific event categories can indeed detect the same categories when they occur in live network conditions.*

Key Publications

The SL methodology and comprehensive algorithm benchmarking were presented in **Paper A**, which demonstrated the 13-class classification results over LDS1. **Paper B** highlighted the trained HGB classifier's superior performance for detecting harmful and non-harmful events in LDS2. The real-time monitoring demonstration described in this section was carried out as part of the CELTIC-NEXT AI-NET-PROTECT project and was presented during the final project review event in Berlin.

In summary, these SL experiments on LDS1 and LDS2 demonstrate that SOP-based monitoring combined with ensemble ML classifiers can reliably detect and categorize a wide range of events, with accuracy levels above 92%. These results provide strong evidence that polarization signatures offer a ro-

bust and scalable foundation for real-time security monitoring in optical networks.

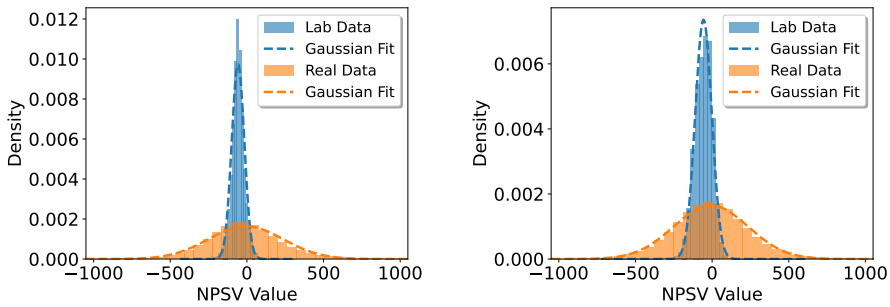
5.2 Extending and Validating the Framework on Real-World Fiber Deployments

While the strong performance of SL classifiers on laboratory data was encouraging, a critical next step was to extend the framework to real fiber networks operating under field conditions and assess *whether the developed models could maintain their effectiveness under the increased complexity and noise of actual deployments.*

Laboratory environments, despite efforts to introduce realistic perturbations, inherently differ from actual deployments. Real fibers experience continuous background polarization drift due to temperature fluctuations, mechanical vibrations from nearby traffic, and connector-induced noise. Furthermore, propagation over long distances introduces attenuation and dispersion effects absent in short laboratory spans. To address this critical validation gap and answer **RQ2 and RQ3** under realistic conditions, we utilized the RDS1 and evaluated the framework's robustness against the stochastic nature of operational environments.

Comparison Between Collected Laboratory and Real-World Data

Before evaluating classification performance, it is essential to quantify the domain shift between controlled and operational environments. We compared the polarization signatures collected from the laboratory setup (LDS1) with those obtained from the field deployment (RDS1). Figure 5.4 illustrates this comparison for two representative scenarios: relaxed state (Figure 5.4(a)) and harmful 80 Hz vibration (Figure 5.4(b)). While both datasets exhibit similar overall polarization variation patterns, confirming the fundamental premise that different physical disturbances produce characteristic SOP signatures that remain recognizable even in field-deployed fibers, clear statistical differences are observable in the distribution of the NPSV values. The laboratory data (LDS1) exhibit narrow NPSV distributions with low variance, indicating stable polarization dynamics. In contrast, the field data (RDS1) display significantly broader distributions with higher variance. This spectral



(a) Relaxed **(b)** 80 Hz vibration

Figure 5.4: Comparison between laboratory and real-world polarization data for (a) relaxed and (b) 80 Hz vibration scenarios, showing increased variance and noise in real-world environments.

broadening is attributed to uncontrolled environmental influences, including traffic-induced micro-vibrations and thermal gradients affecting the buried fiber ducts. These differences highlight the increased complexity and noise present in field-deployed links, emphasizing the need for robust feature extraction to ensure reliable detection performance under realistic operational conditions.

Conventional SL Classification on Field Data

To assess the feasibility of ML-driven SOP-based sensing framework in operational environments, we applied the SL classification methodology to the real-world dataset (RDS1). The dataset was randomly partitioned into a 70% training set (23,520 samples) and a 30% testing set (10,080 samples), ensuring equal representation across the seven event scenarios. We benchmarked several classifiers, including XGBoost, HGB, and SVM, to identify the most robust architecture for handling the increased environmental noise inherent in field deployments. The results confirmed that the proposed framework remains effective outside the laboratory. The HGB classifier emerged as the top-performing model, achieving an overall accuracy of 86.5% and an F1-score of 0.866 across the combined short and long links. XGBoost demonstrated comparable performance, reinforcing the suitability of tree-based ensemble

methods for polarization signature analysis also in real-world environments.

This field validation yielded several important insights. First, it confirmed that the polarization signature approach is fundamentally sound; the core patterns that distinguish different events remain detectable even in the presence of significant background noise. Second, it revealed that, while laboratory training provides a strong foundation, adaptation to specific deployment conditions is essential for maximizing accuracy. Third, it highlighted that the complex, non-stationary nature of real-world environmental noise, arising from a combination of factors such as temperature fluctuations, traffic-induced vibrations, and other uncontrolled field conditions, poses a challenge for traditional feature-based classifiers. Addressing these complexities requires the advanced feature extraction capabilities of DL models, which are the focus of the investigation presented in the next section.

DL Models for Robust Event Detection in Real-World Fiber Links

To overcome the performance limitations of classical SL observed in metro-scale environments deployment, we advanced the framework by developing and evaluating specialized DL architectures. This investigation utilized the expanded RDS2 dataset, which introduces significantly higher complexity through fourteen distinct event classes, including sophisticated overlapping scenarios and varying noise profiles. The objective was to determine whether DL models can learn invariant features well enough to distinguish between the distinctive patterns of SOP signatures and the high-variance background noise of operational networks.

The applied DL architecture comprised 1D CNN layers followed by fully connected dense layers for high-level classification. The CNN architecture was carefully designed to match the characteristics of our data. Using the *Keras Tuner* framework, we conducted an extensive hyper-parameter optimization campaign (100 trials per link type) to tailor the architectures. The results summarized in Table 5.1 demonstrate a substantial performance improvement over the best conventional classifier, HGB, which achieved 86.5% accuracy on the long-link installation using the 7-class RDS1 (**Paper C**). In contrast, the proposed DL model achieves 92.26% on the more challenging 14-class RDS2 long-link dataset and reaches 98.57% on the short-link, matching the ∼98% accuracy previously achieved only under controlled laboratory

Table 5.1: Accuracy of the Best Hyper-Parameter-Tuned 1D CNN Models over the Training, Validation, and Testing Sets.

Installation	Training	Validation	Testing
Short-link	99.81	98.90	98.57
Long-link	95.28	92.77	92.26

conditions (**Paper B**).

These results address **RQ3** by confirming that specialized DL architectures can effectively mitigate environmental noise and generalize to complex, overlapping event scenarios in real-world deployments.

Key Publications and Contributions

The field validation methodology and performance results using classical SL were presented in **Paper C**, while the advanced DL implementation was detailed in **Paper D**. Furthermore, the statistical comparative analysis of the polarization dynamics between LDS1 and RDS1 was comprehensively described in **Paper C** and **Paper E**.

This work is one of the first demonstrations that polarization-based intrusion detection and categorization can function reliably on deployed, in-service fibers. It effectively bridges the gap between controlled laboratory demonstrations and practical network security applications, proving that the proposed ML-driven framework has real-world viability.

5.3 Beyond Supervised Learning: Anomaly Detection for Emerging Threats

The field trials revealed an important limitation of purely SL approaches: they excel at recognizing events similar to those in the training set but may struggle with novel or previously unseen disturbances. In real-world security applications, this is a critical vulnerability. Attackers may develop new tapping techniques, environmental conditions may produce unexpected perturbations, or multiple events may occur simultaneously in ways not represented in the training data. Obtaining labeled examples of every possible future threat scenario is neither practical nor possible.

This realization motivated the exploration of **RQ5**: *how can we detect emerging or previously unseen events when labeled training data is unavailable or scarce?* The answer lies in shifting from classification to anomaly detection, a fundamentally different paradigm where the goal is not to assign a specific label to each event, but rather to identify when something unusual is happening that deviates from the learned profile of normal operation.

We investigated two complementary approaches to anomaly detection on polarization data. The first employed SSL through an OCSVM. This method is trained exclusively on examples of normal fiber operation, learning to characterize the boundary of normal polarization behavior in feature space. Any observation falling outside this learned boundary is flagged as an anomaly. The elegant aspect of this approach is its minimal data requirement; it needs no labeled examples of attacks, only sufficient data to establish what constitutes normal behavior.

The second approach employed USL through DBSCAN. Rather than learning what is normal, DBSCAN directly analyzes the structure of polarization signature data, identifying natural clusters that correspond to different types of events. Points that do not fit well into any cluster are classified as noise or anomalies. Importantly, this method requires no labels whatsoever; it discovers patterns purely from the data distribution.

To evaluate these approaches, we applied them to the comprehensive 13-class LDS1, but treated it as if the event labels were unknown. This simulated a realistic scenario where a monitoring system must detect a wide variety of possible anomalies without prior knowledge of what might occur. The results were highly promising: OCSVM successfully detected the majority of harmful events as anomalies by recognizing they did not conform to the normal operation profile. DBSCAN, remarkably, was able to separate many event types into distinct clusters, effectively rediscovering the event taxonomy purely from the data structure. A particularly exciting finding was that these unsupervised methods proved capable of detecting complex scenarios that were never explicitly included in the training data. For instance, when two disturbances occurred simultaneously on a fiber, both OCSVM and DBSCAN correctly flagged these overlapping events as anomalies or unusual patterns, despite neither model having ever been presented with a labeled example of such combined scenarios: OCSVM was trained exclusively on normal operating conditions, while DBSCAN required no labeled data at all. This demon-

strates a key advantage of anomaly detection approaches: they can recognize that something is wrong even when they have never seen that specific anomaly before.

This work represents the first application of SSL and USL techniques to SOP-based fiber monitoring, filling a significant gap in the literature. It demonstrates that polarization monitoring systems can maintain robust security even in the face of evolving threats and unexpected events. These findings have important implications for real-world deployments, where the monitoring system must be adaptive and not solely reliant on predefined event signatures.

Key Publications

The SSL and USL methodology, including OCSVM and DBSCAN approaches, along with their evaluation on detecting known and emerging threats, was detailed in **Paper F**. This work established the adaptive anomaly detection framework as a core component of the overall monitoring system.

5.4 Evaluating Operational Robustness: Modulation Effects, Model Generalization, and Complex Mechanical Disturbances

As the research progressed from controlled laboratory conditions to real-world deployments and from known events to anomaly detection in Sections 5.1 through 5.3, the foundational validity of the ML-driven SOP sensing framework was asserted. However, these earlier investigations primarily relied on unmodulated optical signals not carrying any data, link-specific training paradigms, and, with the exception of some overlapping vibration scenarios examined in RDS2, predominantly controlled or single-source mechanical perturbations that do not fully reflect the broadband, stochastic nature of real operational environments. To bridge the gap between experimental validation and widespread operational deployment in commercial optical transport networks, it is essential to address the complexities introduced by active data transmission and diverse environmental conditions. Consequently, this section targets three critical dimensions of robust physical-layer monitoring that require specific analysis:

Impact of Signal Modulation on Polarization Sensing

A critical question for the practical deployment of SOP-based monitoring (**RQ6**) concerns *whether the approach works for modulated traffic-carrying signals or only for simple unmodulated probes.* Modern optical networks predominantly use advanced modulation formats such as DP-16QAM for coherent transmission. The modulation process itself encodes digital information by rapidly varying the amplitude and phase of the optical signal, switching between a discrete set of polarization and phase states (symbols) at rates of tens of billions per second in modern coherent systems such as DP-16QAM. Since polarization is one of the dimensions used to carry data in polarization-multiplexed formats, the SOP of a modulated signal no longer remains stable but instead hops rapidly among many states on the Poincaré sphere in a pattern dictated by the transmitted bit sequence. These rapid, symbol-driven polarization fluctuations occur at GHz timescales, which is orders of magnitude faster than the Hz-to-kHz range of SOP changes induced by physical disturbances such as fiber bending or mechanical vibrations. If not properly accounted for, this high-frequency polarization activity could potentially mask or interfere with the detection of the slower, disturbance-induced SOP variations that are the target of the sensing framework.

To address **RQ6**, we performed a comparative analysis of signatures from the RDS3 dataset, which captures identical physical disturbances on a 200 Gbps DP-16QAM traffic channel and a reference CW signal over the 63.4 km HEAnet metro ring. The analysis revealed distinct statistical characteristics between the two signal modalities, as summarized in Table 5.2. The modulated signal consistently exhibited markedly lower standard deviation and near-Gaussian NPSV distributions for all event types, in contrast to the broader, heavy-tailed distributions of the unmodulated signal. For instance, in the relaxed condition, the unmodulated signal shows $\sigma = 18.38$ and kurtosis $\kappa = 276.24$, while the modulated counterpart exhibits $\sigma = 3.60$ and $\kappa = -0.59$. This stabilizing effect arises because the 200 Gbps symbol rate is far beyond the kHz-range sampling bandwidth of the polarization analyzer, causing rapid symbol transitions to be averaged out. As a result, the modulated signal exhibits a smoother SOP trajectory, yet crucially, the SOP signatures of external physical events, such as the 80 Hz mechanical vibration, remain clearly distinguishable for both signal types.

When we applied our ML classification pipeline separately to the modu-

Table 5.2: Statistical properties of NPSV for modulated and unmodulated signals across four event types **Paper_I**.

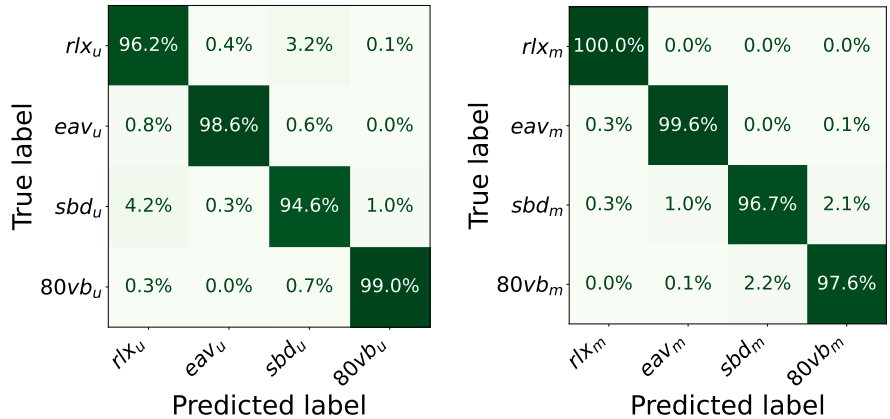
Event	Mean	σ	Skewness	Kurtosis
rlx_u	3.77	18.38	-9.64	276.24
rlx_m	4.01	3.60	-0.29	-0.59
eav_u	3.93	11.00	-4.97	232.43
eav_m	4.03	2.53	-0.12	0.13
sbd_u	-58.12	144.54	-2.52	5.59
sbd_m	4.00	3.01	0.10	-0.14
$80vb_u$	3.90	19.20	-0.08	0.98
$80vb_m$	4.00	3.23	0.01	-0.50

lated and unmodulated datasets, the XGBoost classifier emerged as the best-performing model in both cases. It achieved 97.12% accuracy on the unmodulated signal and 98.47% on the modulated DP-16QAM channel . The corresponding confusion matrices are shown in Figure 5.5, confirming strong class separability in both cases, with minor confusion observed only between the *sbd* and *80vb* events.

This result provides strong empirical evidence that the high-frequency polarization noise inherent to coherent transmission does not obscure the distinct spectral signatures of physical tampering. This work provides the first comprehensive validation that SOP-based security monitoring is applicable to modern coherent optical networks using advanced modulation formats, implying that monitoring systems can operate on channels carrying live traffic without requiring dedicated dark fibers or unmodulated probe signals. The key insight is that physical disturbances and signal modulation operate at fundamentally different timescales and can be effectively separated through appropriate signal processing.

Generalization Across Spectral Bands and Fiber Links

Another relevant practical question (**RQ7**) concerns the generalizability of trained ML models across different deployment scenarios: *can a model trained on data from one spectral band and fiber link accurately classify events occurring on a different band or link?* This question is important because training



(a) Unmodulated signal (97.12% accuracy) **(b)** Modulated DP-16QAM signal (98.47% accuracy)

Figure 5.5: Confusion matrices of the XGBoost classifier for the unmodulated and modulated signal scenarios (RDS3), demonstrating comparable and high classification performance under both signal modalities.

comprehensive models for every possible fiber and wavelength configuration would be prohibitively expensive and time-consuming. Ideally, a monitoring system should exhibit some degree of transfer learning capability.

To investigate RQ7, we utilized the RDS4 dataset, which comprises identical event classes: relaxed, soft bending, and eavesdropping, collected from two distinct physical systems: an O-band signal over a 21 km dark fiber link (referred to as System 1) and a C-band signal in a live metro network link of 63.4 km (System 2). Using the XGBoost classifier, we evaluated performance across three training paradigms:

- **Intra-System:** Training and testing on the same system.
- **Cross-System:** Training on one system and testing on the other.
- **Multi-System:** Training on a combined dataset from both systems and testing on the combined set.

The results, summarized in Table 5.3, reveal a stark contrast between intra-system and cross-system performance. The intra-system scenarios confirmed

Table 5.3: Classification Accuracy for Intra-System, Cross-System, and Multi-System Scenarios using RDS4.

Scenario	Training Domain	Testing Domain	Accuracy
Intra-System	System 1 (O-band)	System 1 (O-band)	88.85%
	System 2 (C-band)	System 2 (C-band)	98.63%
Cross-System	System 1 (O-band)	System 2 (C-band)	8.11%
	System 2 (C-band)	System 1 (O-band)	60.59%
Multi-System	Systems 1 + 2	Systems 1 + 2	**91.11%**

high accuracy (up to 98.6%), validating that the models effectively learn system-specific features. However, the cross-system performance was critically limited. When a model trained on System 1 was applied to System 2, accuracy dropped dramatically to 8.11%, failing to classify any events correctly. This severe degradation indicates that polarization signatures are highly system-dependent; the specific trajectory of SOP variations is strongly affected by the unique birefringence profile, wavelength, and accumulated environmental noise of the specific link. Crucially, however, the Multi-System approach offered a viable solution. By training on a diverse dataset containing samples from both domains, the model achieved a high accuracy of 91.1%. This result demonstrates that, while direct transfer learning is challenging, models trained on sufficiently diverse data can learn more robust, generalizable representations of polarization disturbance patterns.

Detection of Complex and Overlapping Mechanical Vibrations

Another important requirement for operational robustness is the ability to distinguish simultaneous and spectrally complex disturbances (**RQ8**). While earlier validation steps included partially overlapping events, those experiments primarily focused on detecting and categorizing known harmful events under controlled or field-deployed conditions. In contrast, the goal here is to stress-test the framework under significantly more demanding conditions, where pseudo-random broadband vibrations spanning 0–2000 Hz are combined with tampering events, more closely emulating the persistent, spectrally dense background noise generated by heavy traffic, industrial machinery, or environmental noise in real operational environments. Such broadband disturbances

can spectrally overlap with and mask the signatures of malicious interventions like tapping or excavation, posing a greater classification challenge than the targeted single- or dual-frequency scenarios examined previously.

To evaluate the framework's resilience against such stochastic conditions, we utilized the LDS3 dataset, creating a highly challenging landscape of 14 distinct, overlapping event classes. We benchmarked a diverse suite of SL classifiers on this complex dataset. The results confirmed that ensemble-based learning remains highly effective even under these noisy conditions. The HGB classifier emerged as the top-performing model, achieving an overall accuracy of 88.33%. While this is lower than the accuracy achieved on simpler, isolated events, it remains remarkably high given the complexity of the task. More importantly, the model demonstrated the ability to separate overlapping events and distinguish harmful activities from benign disturbances even under these noisy, multi-event conditions. This validates that our ML-based approach is not limited to ideal scenarios but can handle the complexity and ambiguity of real-world fiber environments.

Key Publications

The comprehensive study of modulation effects on polarization sensing was detailed in **Paper I**, providing the first direct comparison of modulated versus unmodulated channels. Cross-band and cross-link generalization was investigated in **Paper G**, revealing the limitations and opportunities for transfer learning. Finally, the detection of complex and overlapping vibrations was presented in **Paper H**, demonstrating robustness under realistic multi-event scenarios.

5.5 Chapter Summary

This chapter presented a comprehensive investigation of ML-driven polarization-based monitoring for optical network monitoring, building on the experimental foundation and datasets established in Chapter 4.

The research progressed through four interconnected phases. Section 5.1 demonstrated that ensemble SL classifiers, particularly XGBoost and HGB, can reliably detect and categorize a wide range of fiber perturbations from their polarization signatures, achieving above 92% accuracy on laboratory

data and validating real-time detection capability on a live indoor fiber link. Section 5.2 extended the framework to field-deployed fibers, confirming its robustness under operational noise conditions and demonstrating that DL architectures further improve classification accuracy on real-world links, reaching 92.26% on a long-link and 98.57% on a short-link. Section 5.3 addressed the challenge of limited labeled data by introducing SSL and USL anomaly detection approaches, establishing the first application of OCSVM and DBSCAN to SOP-based fiber monitoring and demonstrating their ability to detect previously unseen threats. Finally, Section 5.4 evaluated the framework under three additional dimensions of operational realism: the impact of signal modulation, generalization across spectral bands and fiber links, and robustness against complex broadband overlapping disturbances.

Each stage built upon insights from the previous one, advancing the framework from proof-of-concept demonstrations in controlled settings to a validated, adaptive solution capable of addressing the diverse challenges of physical-layer security monitoring in real optical network deployments.

CHAPTER 6

Summary of included papers

This chapter provides a summary of the included papers.

6.1 Paper A

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Marija Furdek
Machine Learning-Based Polarization Signature Analysis for Detection and Categorization of Eavesdropping and Harmful Events
Published in Optical Fiber Communications Conference and Exhibition (OFC), 24-28 March, 2024, pp. 1-3, San Diego, CA, USA.
 ISBN:979-8-3503-7758-3.

This paper presents a ML-based approach to enhance security in optical fiber networks by detecting and categorizing eavesdropping and potentially harmful and non-harmful vibration events over fiber optic installations. It utilizes the SOP variations to identify unique signatures caused by different physical manipulations of the fiber optic transmission line. The methodology includes data collection from 13 experimental scenarios from three different optical cables, including FOCS, indoor cables, and bare SM G.675

bend-insensitive fiber. We conducted experiments over a number of ML algorithms to select the most appropriate classifier for this 13-class classification problem. The XGBoost classifier achieves the best performance with a 92.3% accuracy in distinguishing between normal operations and potentially harmful activities. This approach automates threat detection, providing a scalable and effective solution for securing optical networks.

Leyla Sadighi (LS) led the data analysis by transforming the collected SOP measurements into meaningful representations for ML, conducted the simulations, analyzed the results, and was the primary author of the manuscript. Stefan Karlsson (SK) designed and executed the experimental setup and was responsible for collecting the SOP data. He also collaborated in writing the Introduction and experimental setup sections of the paper. Carlos Natalino (CN) contributed to the evaluation of the ML results and provided substantial input to improve the manuscript. Marija Furdek (MF) initiated the research direction and facilitated the collaboration with FMV, and contributed to the manuscript writing and the analysis of the results.

6.2 Paper B

Leyla Sadighi, Stefan Karlsson, Lena Wosinska, Marija Furdek
Machine Learning Analysis of Polarization Signatures for Distinguishing Harmful from Non-harmful Fiber Events
Published in 24th International Conference on Transparent Optical Networks (ICTON), 14-18 July, 2024, pp. 1-5, Bari, Italy.
 DOI:10.1109/ICTON62926.2024.10648140.

This paper introduced a method for detecting and classifying harmful and non-harmful events in optical fiber networks by leveraging ML to analyze changes in the SOP. We collected SOP signatures by manipulating indoor cables to mimic real-world attack scenarios. Five scenarios were examined, including normal conditions, non-harmful vibrations, eavesdropping attempts, potentially harmful vibrations, and dual-frequency vibrations (both harmful and non-harmful). By generating unique polarization signatures for each event type, we employed various ML classifiers to differentiate these scenarios, with the HGB classifier achieving a high accuracy of 97.94%. This approach significantly improves the identification of physical layer anomalies in optical networks, particularly harmful events such as mechanical vibrations caused

by heavy machinery activities.

LS led the data analysis, conducted the simulations, analyzed the results, and was the primary author of the manuscript. SK initiated the research idea, designed and executed the experimental setup, and was responsible for collecting the SOP data. Lena Wosinska (LW) contributed to the paper writing by providing comments to improve the manuscript. MF contributed to the evaluation of the ML results and contributed to the manuscript preparation.

6.3 Paper C

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek
Detection and Classification of Eavesdropping and Mechanical Vibrations in Fiber Optical Networks by Analyzing Polarization Signatures Over a Noisy Environment
Presented in 50th European Conference on Optical Communication (ECOC), [invited paper]
28–26 September, 2025, pp. 527–530, Frankfurt, Germany.

In this paper, we study how polarization signatures can be recorded and classified, originating from an installed transmission line in a real-life network, OpenIreland, operated by TCD and located under the street in Dublin, Ireland. Real-world data from two separate installations in Dublin with link lengths of 0.15 km and 10.5 km are used to evaluate the method. Our ML analysis uses data from seven real-life network signatures to differentiate between polarization patterns obtained during normal operation and those suggesting malicious vibrations and eavesdropping. We evaluate several ML algorithms to determine a suitable classifier for our seven-class classification problem. The HGB classifier outperforms other models in the real-world dataset, achieving an accuracy of 86.5% and an F1-score of 0.866.

LS formulated the research question, conducted the simulations, analyzed the results, and served as the primary author of the manuscript. SK designed and executed the experimental setup and was responsible for collecting the SOP data. CN contributed to the evaluation and interpretation of the ML results and provided valuable feedback to improve the manuscript. Marco Ruffini (MR) contributed to the data collection process, enabled access to the

real-world OpenIreland infrastructure, and collaborated in writing the experimental setup section. MF initiated the collaboration with MR at TCD, and contributed to the analysis of results and the writing of the manuscript.

6.4 Paper D

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek
Deep Learning for Detection of Harmful Events in Real-World, Noisy Optical Fiber Deployments
Published in IEEE Journal of Lightwave Technology (JLT),
vol. 43, no. 4, pp. 1125–1139, February 2025.
DOI: 10.1109/JLT.2025.3557748,

In this work, we present a DL-based approach for detecting and classifying harmful events in real-world, noisy optical fiber networks by analyzing variations in the SOP of transmitted light. Leveraging a 1D CNN combined with fully connected dense layers, we trained our model on SOP variation data collected from two field-deployed links in Dublin, Ireland, a short 0.15 km single-direction representative of access networks and a 10.5 km single-direction metro-scale, both operating under realistic environmental noise. The experimental scenarios encompassed benign background vibrations, routine cable bending, potentially harmful mechanical vibrations such as those induced by heavy machinery, and deliberate eavesdropping through fiber bending. Following extensive hyperparameter tuning, our models achieved 98.57% classification accuracy on the short link and 92.26% on the long link, demonstrating robustness even under the increased noise of longer spans. Compared to earlier ML methods, which reached around 86.5% accuracy (Paper C), our approach provides a significant performance gain and enables reliable detection of subtle, security-critical events such as fiber tapping. These results highlight the potential of SOP-based monitoring combined with DL as a scalable, non-intrusive, and effective solution for enhancing the security and resilience of deployed optical communication networks.

LS formulated the research question, implemented the DL models, performed hyperparameter tuning, analyzed the results, and prepared the first draft of the paper. SK designed and executed the experimental setup and contributed to writing the experimental setup description in the manuscript.

CN proposed alternative DL models, suggested incorporating the CNN model, and contributed to formulating the research question and writing the paper. LW assisted in writing the manuscript. MR contributed to the data collection process, enabled access to the real-world OpenIreland infrastructure, and collaborated on writing the experimental setup section. MF initiated the collaboration with MR at TCD, contributed to the analysis of the results, provided suggestions to strengthen the findings, and offered valuable feedback to improve the manuscript.

6.5 Paper E

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek
AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges
Published in Optical Fiber Communication Conference (OFC) 2025, Technical Digest Series, [invited paper],
30 March–3 April 2025, San Francisco, CA, USA.
DOI: 10.1364/OFC.2025.M3F.6, © 2025 Optica Publishing Group .

Optical networks are vulnerable to various disturbances that can impact service availability or user privacy. In this paper, we discuss the use of Artificial Intelligence (AI) and ML techniques to analyze SOP changes for cognitive management of such complex disruptions. We explore how SOP, a sensitive indicator of environmental perturbations like physical intrusions or vibrations, can be harnessed as an in-band sensing metric without requiring dedicated hardware. The paper positions AI/ML-enabled SOP monitoring as a promising strategy to supplement traditional sensing methods, offering scalable, non-intrusive, and intelligent detection capabilities. We outline key challenges and open research questions in deploying such approaches, including data variability, event complexity, and model robustness in dynamic network environments.

LS implemented the statistical comparison between the controlled environment and the real-world OpenIreland SOP signatures. SK collected the experimental and real-world data. CN formulated the research question and key challenges and prepared the first draft of the manuscript. LW assisted in writing the manuscript. MR contributed to collecting real-world data. MF

assisted in writing the manuscript.

6.6 Paper F

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Marija Furdek
ML-based State of Polarization Analysis to Detect Emerging Threats to Optical Fiber Security
IEEE Transactions on Network and Service Management (TNSM), TNSM-2024-08489,
DOI: 10.1109/TNSM.2025.3607022 .

In this paper, we present a ML-based approach for monitoring optical fiber networks and detecting emerging threats by analyzing variations in the SOP of signals in an operational link. The work is motivated by the limitations of traditional fiber sensing techniques, which often fail to detect subtle or novel physical-layer anomalies, small bends, low-level vibrations, and similar disturbances, particularly without disrupting normal network traffic. We experimentally collected a comprehensive dataset of polarization signatures from controlled laboratory experiments on three types of fiber: bare fiber, a Fiber Optical tactical Cable System, and an indoor fiber. The dataset covers 13 distinct disturbance scenarios, ranging from normal background vibrations caused by equipment operating at 155 Hz to harmful events such as fiber bending for eavesdropping and 80 Hz vibrations. We also included combined overlapping events to replicate real-world situations in which multiple disturbances occur simultaneously. Two anomaly detection models were applied to these SOP data: a OCSVM used as a SSL detector trained exclusively on normal-operation data, and a DBSCAN algorithm used as an USL detector to automatically identify deviations in SOP patterns indicative of anomalies. To fine-tune the models and ensure consistent performance, we developed tailored evaluation metrics that guided hyperparameter selection and assessed each model's ability to generalize across disturbance types, maintain detection consistency, and minimize false alarms. The results show that OCSVM, after learning a baseline of normal behavior, can detect complex fiber anomalies with high confidence, achieving F1 scores between 0.98 and 0.998 in most scenarios. The DBSCAN method, which requires no labeled training data, also demonstrated strong detection capability with F1 scores up to 0.995, although its performance varied more widely across scenarios, particularly for complex

overlapping events.

LS formulated the research question, implemented the SSL and USL models, performed hyperparameter tuning, analyzed the results, and prepared the first draft of the paper. SK designed and executed the experimental setup. CN and MF proposed additional model options and contributed to refining the research question as well as improving the writing and overall quality of the manuscript.

6.7 Paper G

Leyla Sadighi, Carlos Natalino, Stefan Karlsson, Marco Ruffini, Eoin Kenny, Lena Wosinska, Marija Furdek
Generalizability of ML-Based Classification of State of Polarization Signatures Across Different Bands and Links
Published in 51st European Conference on Optical Communication (ECOC), 28 September to 2 October, 2025, Copenhagen, Denmark.
DOI: 10.1109/ECOC66593.2025.11263096

This paper investigates the generalization capability of ML models for SOP-based event detection across different spectral bands and fiber links. Using experimental data from two real-world systems, a 21 km O-band dark fiber in the OpenIreland testbed and a 77 km C-band link in HEAnet's live metro network, we evaluate the performance of an XGBoost classifier trained on relaxed, eavesdropping, and soft-bending signatures. Intra-system testing yields high accuracy (88.85% in O-band, 98.63% in C-band), while cross-system testing shows significant degradation (8.11% and 60.59%), revealing strong system dependency. Multi-system training improves robustness, achieving 91.11% accuracy, underscoring the importance of system-specific knowledge and the potential of multi-domain approaches for enhancing generalization in SOP-based optical network monitoring.

LS participated in data collection, formulated the research question, implemented the ML models, analyzed the results, and prepared the first draft of the manuscript. SK contributed to the collection of real-world data. MR contributed to the real-world data collection, provided valuable input to refine the research questions, and offered improvements to the writing of the manuscript. Eoin Kenny (EK) assisted in the data collection process by facilitating access

to the HEAnet network via OpenIreland. LW provided comments on the draft of the paper. CN and MF reviewed and discussed the ML model results, suggested improvements for describing the experimental scenarios, and provided valuable feedback that strengthened the writing and overall quality of the manuscript.

6.8 Paper H

Leyla Sadighi, Stefan Karlsson, Marco Ruffini, Marija Furdek
ML-Based Detection and Categorization of Complex Mechanical Vibrations via State of Polarization Analysis in Optical Networks
Published in the 25th International Conference on Transparent Optical Networks (ICTON),
6–10 July, 2025, Barcelona, Spain.
ISBN: 978-1-6654-7164-0,

In this paper, we propose a ML-based framework for detecting and categorizing complex mechanical disturbances in optical fiber networks using SOP monitoring. Addressing the challenge of overlapping and mixed-frequency vibration patterns from benign activities, malicious attacks, or simultaneous events, we collected 14 distinct SOP signatures under controlled laboratory conditions using both bare fiber and patch cable segments. The disturbances include synthetic complex vibration patterns emulating real-world spectral characteristics, combined with targeted events such as soft bending, eavesdropping, and harmful 80 Hz vibrations. SOP variations were captured with an optical analyzer, processed into NPSV features, and transformed into frequency-domain signatures for supervised ML classification. We benchmarked multiple algorithms, with Histogram Gradient Boosting achieving the highest performance at 88.33% accuracy and an F1-score of 0.8828, offering a strong trade-off between predictive accuracy and computational efficiency. The results demonstrate the robustness of SOP-based spectral analysis for distinguishing subtle and overlapping mechanical events, paving the way for real-time monitoring and threat detection in operational optical networks.

LS implemented the ML models and prepared the first draft of the manuscript. SK formulated the research question, designed and executed the experimental setup, and contributed to writing the experimental setup description in the manuscript. MR and MF reviewed and discussed the ML model results

and provided valuable suggestions that improved the overall quality of the manuscript.

6.9 Paper I

Leyla Sadighi, Carlos Natalino, Stefan Karlsson, Lena Wosinska, Eoin Kenny, Venkata Virajit Garbhapu, Marco Ruffini, Marija Furdek
DP-16QAM Modulated vs. Unmodulated Polarization Signatures for Machine Learning-Based Fiber Sensing
Published in IEEE Journal of Lightwave Technology (JLT), February 2026.
DOI: 10.1109/JLT.2026.3660791 .

In this paper, we present the first systematic experimental comparison of SOP signatures obtained from modulated and unmodulated optical signals subjected to identical physical disturbances in a real-world metro network. Using a 77 km C-band link in the HEAnet network, we collected SOP data for eight representative events, including relaxed, soft bending, eavesdropping, and 80 Hz vibration, under both signal modalities. Statistical analysis revealed that modulation suppresses stochastic polarization noise while preserving low-frequency SOP variations critical for event classification. We designed four datasets ranging from isolated to mixed and multi-event/multi-modality scenarios, enabling a thorough assessment of how modulation influences the learnability and separability of SOP features. Benchmarking multiple supervised ML algorithms showed that XGBoost and HGB consistently achieved over 97% accuracy across all scenarios, with no significant degradation due to modulation. The results demonstrate that ML-based SOP analysis remains highly effective in coherent optical networks regardless of signal modality, paving the way for practical, signal-agnostic deployment of polarization-based sensing for real-time detection of benign and malicious fiber events without disrupting live traffic.

LS participated in data collection, formulated the research question, implemented the ML models, analyzed the results, and prepared the first draft of the manuscript. SK contributed to the collection of real-world data. MR contributed to real-world data collection and offered improvements to the manuscript writing. EK assisted in the data collection process and facilitated access to the HEAnet network via OpenIreland. LW supported the literature review for a portion of the introduction section. CN and MF reviewed and dis-

cussed the ML model results, and provided valuable suggestions for improving ML results as well as the overall quality and clarity of the manuscript.

CHAPTER 7

Concluding Remarks and Future Work

7.1 Concluding Remarks

The research presented in this thesis establishes a robust framework for addressing a fundamental concern in modern telecommunications: *ensuring the security and resilience of the physical infrastructure that underpins global connectivity.* As detailed in Chapter 1, the physical layer of these networks is increasingly exposed to a diverse spectrum of threats, spanning from accidental mechanical disturbances to sophisticated eavesdropping attempts. The cumulative impact of this body of work can be synthesized along several key dimensions.

First, from a methodological perspective, this thesis establishes polarization signature analysis as a viable and robust mechanism for physical-layer security. By moving beyond theoretical modelling to rigorous empirical validation across controlled laboratory environments and diverse field settings, we have demonstrated that SOP variations provide a rich, informative feature space for sensing. Second, from an algorithmic perspective, the work proves that a tiered ML strategy—combining SL for precise classification, DL for complex real-world noisy environments, and SSL/USL for detection of

novel anomalies—can handle the full spectrum of monitoring requirements. This comprehensive algorithmic framework addresses the limitations of traditional threshold-based systems, offering the adaptability required to detect both known threats and emerging, unseen anomalies. Ultimately, this thesis demonstrates that intelligent, ML-driven monitoring of optical fiber polarization is not merely a theoretical possibility but a practical approach for real-world deployment.

The proposed framework establishes a critical line of defense against physical-layer vulnerabilities, effectively transforming passive infrastructure into an active, self-monitoring security asset. The research opens new avenues for leveraging AI to safeguard critical network infrastructure, with implications extending beyond fiber-optic security to broader domains of intelligent infrastructure monitoring and protection.

7.2 Future Work

Building on this dissertation's findings, several concrete directions can be pursued to further advance ML-based polarization monitoring for widespread use:

- **Improving Model Generalization:**

  To address cross-link and cross-band limitations, future research can explore transfer learning and domain adaptation techniques. For instance, a base classifier trained on a large amalgamated dataset could be fine-tuned with minimal new data for a specific deployment, reducing the need to retrain from scratch for each fiber route. Investigating the fundamental invariants in polarization patterns across different systems might also lead to identifying features that are less environment-specific. This would move the solution closer to a plug-and-play system for any optical link.

- **Expanded Event Coverage:**

  While our experiments covered common threats (bending taps, vibrations, etc.), an important next step is to broaden the spectrum of detectable events. Future work should incorporate scenarios like fiber pinches, connector tampers, gradual fiber fatigue, and deliberate physical attacks on cable conduits. Expanding the training dataset with such

events (including combinations of multiple simultaneous faults) will enhance the system's preparedness.

- **Precise Event Localization:**

  While this thesis focused on detection and classification, determining the exact geographical location of a disturbance remains a critical challenge for rapid incident response. Future research should investigate methods to estimate the position of an event along the fiber link directly from SOP data. Integrating localization capabilities would transform the framework from a monitoring tool into a complete fault management solution.

- **Real-Time Deployment and Automation:**

  Transitioning from offline analysis to real-time monitoring systems will be a pivotal step. This entails implementing our ML algorithms on streaming data and ensuring they can deliver alerts with low latency. Future studies should develop prototypes that integrate with optical network controllers or Software-Defined Networking (SDN) platforms, allowing automated mitigation actions (like rerouting traffic) when a threat is detected.

Addressing current limitations and pursuing the outlined future directions provides a clear path toward deploying these techniques in production networks. Such deployment would mark an important step toward autonomous optical networks that are not only high-capacity and agile, but also self-protecting against physical-layer disruptions. The continued evolution of this field will ensure that optical networks remain robust in the face of growing threats, safeguarding the critical data highways of the modern world.

References

[1] M. Azadeh, "Fiber optic communications: A review," *Fiber Optics Engineering*, pp. 1–27, 2009.

[2] N. Skorin-Kapov, M. Furdek, S. Zsigmond, and L. Wosinska, "Physical-layer security in evolving optical networks," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 110–117, 2016.

[3] "Internet architecture: Potential risks to fiber infrastructure and impacts on critical services," U.S. Government Accountability Office, Tech. Rep. GAO-22-104560, 2022.

[4] V. Spurny, P. Munster, A. Tomasov, T. Horvath, and E. Skaljo, "Physical layer components security risks in optical fiber infrastructures," *Sensors*, vol. 22, no. 2, p. 588, 2022.

[5] D. Dahan and U. Mahlab, "Security threats and protection procedures for optical networks," *IET Optoelectronics*, vol. 11, no. 5, pp. 186–200, 2017.

[6] O. Nyarko-Boateng, F. E. B. Xedagbui, A. F. Adekoya, and B. A. Weyori, "Fiber optic deployment challenges and their management in a developing country: A tutorial and case study in Ghana," *Engineering Reports*, vol. 2, no. 2, e12121, 2020.

[7] J. Siuzdak, M. Kowalczyk, and M. Marzęcki, "Jamming of optical network operation in physical layer," *International Journal of Electronics and Telecommunication*, vol. 70, no. 4, pp. 969–977, 2024.

[8] Y. Li, Y. Liang, M. Zhang, S. Wei, H. Zhu, Y. Li, Y. Zhao, and J. Zhang, "Fiber eavesdropping detection and location in optical communication system," *Photonics*, vol. 12, no. 5, 2025, ISSN: 2304-6732.

[9] K. T. Kim, H. K. Kim, S. Hwangbo, S. Choi, B. H. Lee, and K. Oh, "Characterization of evanescent wave coupling in side-polished hollow optical fiber and its application as a broadband coupler," *Optics communications*, vol. 245, no. 1-6, pp. 145–151, 2005.

[10] Y. Luo, Q. Wei, Y. Ma, H. Lu, J. Yu, J. Tang, J. Yu, J. Fang, J. Zhang, and Z. Chen, "Side-polished-fiber based optical coupler assisted with a fused nano silica film," *Appl. Opt.*, vol. 54, no. 7, pp. 1598–1605, Mar. 2015.

[11] A. Harris and P. Castle, "Bend loss measurements on high numerical aperture single-mode fibers as a function of wavelength and bend radius," *Journal of Lightwave Technology*, vol. 4, no. 1, pp. 34–40, 1986.

[12] M. Zafar Iqbal, H. Fathallah, and N. Belhadj, "Optical fiber tapping: Methods and precautions," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, 2011, pp. 164–168.

[13] W. Lee, S. I. Myong, J. C. Lee, and S. Lee, "Identification method of non-reflective faults based on index distribution of optical fibers," *Optics express*, vol. 22, no. 1, pp. 325–337, 2014.

[14] K. Abdelli, H. Grießer, C. Tropschug, and S. Pachnicke, "Optical fiber fault detection and localization in a noisy OTDR trace based on denoising convolutional autoencoder and bidirectional long short-term memory," *IEEE Journal of Lightwave Technology*, vol. 40, no. 8, pp. 2254–2264, 2021.

[15] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," *Journal of optical communications and networking*, vol. 14, no. 5, pp. 365–375, 2022.

[16] M. M. Rad, K. Fouli, H. A. Fathallah, L. A. Rusch, and M. Maier, "Passive optical network monitoring: Challenges and requirements," *IEEE Communications Magazine*, vol. 49, no. 2, s45–S52, 2011.

[17] K. V. Stepanov, A. A. Zhirnov, A. O. Chernutsky, K. I. Koshelev, A. B. Pnev, A. I. Lopunov, and O. V. Butov, "The sensitivity improvement characterization of distributed strain sensors due to weak fiber bragg gratings," *Sensors*, vol. 20, no. 22, p. 6431, 2020.

[18] Y. Aono, E. Ip, and P. Ji, "More than communications: Environment monitoring using existing optical fiber network infrastructure," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, W3G.1.

[19] N. Brochier, "Field measurement of polarization fluctuation dynamics and related impact for 40Gbit/s submarine systems," in *Proc. SubOptic 2010 Conference, Yokohama, Japan, May*, 2010.

[20] J. Pesic, E. Le Rouzic, N. Brochier, and L. Dupont, "Proactive restoration of optical links based on the classification of events," in *15th International Conference on Optical Network Design and Modeling-ONDM 2011*, IEEE, 2011, pp. 1–6.

[21] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti, "Detection of abnormal activities on a SM or MM fiber," in *Optical Fiber Communication Conference (OFC)*, 2023, M3Z.6.

[22] A. Tomasov, P. Dejdar, T. Horvath, and P. Munster, "Physical fiber security by the state of polarization change detection," in *Fiber Optic Sensors and Applications XVIII*, SPIE, vol. 12105, 2022, pp. 52–56.

[23] D. Rafique and L. Velasco, "Machine learning for network automation: Overview, architecture, and applications [invited tutorial]," *J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D126–D143, Oct. 2018.

[24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[25] CONNECT Centre for Future Networks and Communications, *Open Ireland Testbed*, Available here.

[26] ASIERA (formerly HEAnet), *Ireland's National Education and Research Network*, Available here.

[27] T. Liu, W. Wang, F. Ouyang, Y. Hao, Y. Li, Y. Zhao, and J. Zhang, "Eavesdropping-aware survivable routing in physical-layer secured optical networks," *J. Opt. Commun. Netw.*, vol. 17, no. 2, pp. 127–138, Feb. 2025.

[28] M. P. Fok, Z. Wang, Y. Deng, and P. R. Prucnal, "Optical layer security in fiber-optic networks," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 725–736, 2011.

[29] F. Futami, K. Tanizawa, and K. Kato, "Y-00 quantum-noise randomized stream cipher using intensity modulation signals for physical layer security of optical communications," *J. Lightwave Technol.*, vol. 38, no. 10, pp. 2774–2781, 2020.

[30] K. Shaneman and S. Gray, "Optical network security: Technical analysis of fiber tapping mechanisms and methods for detection & prevention," in *IEEE MILCOM 2004. Military Communications Conference, 2004.*, vol. 2, 2004, 711–716 Vol. 2.

[31] Y. Li, Y. Liang, M. Zhang, S. Wei, H. Zhu, Y. Li, Y. Zhao, and J. Zhang, "Fiber eavesdropping detection and location in optical communication system," *Photonics*, vol. 12, no. 5, 2025, ISSN: 2304-6732.

[32] V. Spurny, P. Dejdar, A. Tomasov, P. Munster, and T. Horvath, "Eavesdropping vulnerabilities in optical fiber networks: Investigating macrobending-based attacks using clip-on couplers," in *2023 International Workshop on Fiber Optics on Access Networks (FOAN)*, 2023, pp. 47–51.

[33] T. Uematsu, H. Hirota, T. Kawano, T. Kiyokura, and T. Manabe, "Design of a temporary optical coupler using fiber bending for traffic monitoring," *IEEE Photonics J.*, vol. 9, no. 6, pp. 1–13, Dec. 2017, DOI: 10.1109/JPHOT.2017.2762662.

[34] D. Marcuse, "Curvature loss formula for optical fibers," *J. Opt. Soc. Am.*, vol. 66, no. 3, pp. 216–220, Mar. 1976.

[35] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

[36] B. Schneier, *Eavesdropping on a fiber optic cable*, *Schneier on Security Blog*, Accessed: 5 November 2025, 2020.

[37] TorGuard, *It's easier than you think to tap high-speed fiber optics*, *TorGuard Blog*, Accessed: 5 November 2025, 2024.

[38] G. E. UMOH, I. AKPADEN, and A. O. AKPAN, "The vulnerability of fiber-optics communication systems: The role of optical tapping," *J. Inf. Eng. Appl*, vol. 4, pp. 145–153, 2014.

[39] M. Medard, D. Marquis, R. A. Barry, and S. G. Finn, "Security issues in all-optical networks," *IEEE network*, vol. 11, no. 3, pp. 42–48, 2002.

[40] R. Gour, J. Kong, G. Ishigaki, A. Yousefpour, S. Hong, and J. P. Jue, "Finding survivable routes in multi-domain optical networks with geographically correlated failures," *J. Opt. Commun. Netw.*, vol. 10, no. 8, pp. C39–C49, Aug. 2018.

[41] W. D. Grover, *Mesh-based survivable transport networks: options and strategies for optical, MPLS, SONET and ATM networking*. Prentice Hall PTR, 2003.

[42] M. Sena, A. Moawad, R. Emmerich, B. Shariati, M. Geitz, R.-P. Braun, J. Fischer, and R. Freund, "Exploring the potential of longitudinal power monitoring for detecting physical-layer attacks [invited]," *J. Opt. Commun. Netw.*, vol. 17, no. 7, pp. C30–C40, Jul. 2025.

[43] B. Asante and T. Hayford-Acquah, "Causes of fiber cut and the recommendation to solve the problem," *IOSR J. Electron. Commun. Eng.*, vol. 16, pp. 34–51, 2021.

[44] S. Guerrier et al., "Field detection and localization of digging excavator events using MIMO digital fiber sensing over a deployed optical network for proactive fiber break prevention," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, 2024, Tu2J.6.

[45] M. K. Barnoski and S. M. Jensen, "Fiber waveguides: A novel technique for investigating attenuation characteristics," *Appl. Opt.*, vol. 15, no. 9, pp. 2112–2115, Sep. 1976.

[46] A. H. Hartog, *An introduction to distributed optical fibre sensors*. CRC press, 2017.

[47] Z. He and Q. Liu, "Optical fiber distributed acoustic sensors: A review," *Journal of Lightwave Technology*, vol. 39, no. 12, pp. 3671–3686, 2021.

[48] J. C. Juarez, E. W. Maier, K. N. Choi, and H. F. Taylor, "Distributed fiber-optic intrusion sensor system," *Journal of lightwave technology*, vol. 23, no. 6, pp. 2081–2087, 2005.

[49] Y. Muanenda, "Recent advances in distributed acoustic sensing based on phase-sensitive optical time domain reflectometry," *Journal of Sensors*, vol. 2018, no. 1, p. 3 897 873, 2018.

[50] Z. Wang, L. Zhang, S. Wang, N. Xue, F. Peng, M. Fan, W. Sun, X. Qian, J. Rao, and Y. Rao, "Coherent Φ-OTDR based on I/Q demodulation and homodyne detection," *Opt. Express*, vol. 24, no. 2, pp. 853–858, Jan. 2016.

[51] S. Li, Z. Qin, Z. Liu, W. Yang, S. Qu, Z. Wang, and Y. Xu, "Long-distance Φ-OTDR with a flexible frequency response based on time division multiplexing," *Opt. Express*, vol. 29, no. 21, pp. 32 833–32 841, Oct. 2021.

[52] S. Pellegrini, L. Minelli, L. Andrenacci, G. Rizzelli, D. Pilori, G. Bosco, L. D. Chiesa, C. Crognale, S. Piciaccia, and R. Gaudino, "Overview on the state of polarization sensing: Application scenarios and anomaly detection algorithms," *J. Opt. Commun. Netw.*, vol. 17, no. 2, A196–A209, Feb. 2025.

[53] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[54] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[55] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, and A. P. T. Lau, "Optical performance monitoring: A review of current and future technologies," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 525–543, 2015.

[56] R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, W. Freude, and J. Leuthold, "Error vector magnitude as a performance measure for advanced modulation formats," *IEEE Photonics Technology Letters*, vol. 24, no. 1, pp. 61–63, 2012.

[57] M. Furdek, C. Natalino, M. Schiano, and A. D. Giglio, "Experiment-based detection of service disruption attacks in optical networks using data analytics and unsupervised learning," in *Metro and Data Center Optical Networks and Short-Reach Links II*, A. K. Srivastava, M. Glick, and Y. Akasaka, Eds., International Society for Optics and Photonics, vol. 10946, SPIE, 2019, p. 109460D.

[58] C. Natalino, M. Schiano, A. D. Giglio, and M. Furdek, "Root cause analysis for autonomous optical network security management," *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 2702–2713, 2022.

[59] M. Furdek, C. Natalino, F. Lipp, D. Hock, A. D. Giglio, and M. Schiano, "Machine learning for optical network security monitoring: A practical perspective," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2860–2871, 2020.

[60] M. Furdek, C. Natalino, A. Di Giglio, and M. Schiano, "Optical network security management: Requirements, architecture, and efficient machine learning models for detection of evolving threats [invited]," *Journal of Optical Communications and Networking*, vol. 13, no. 2, A144–A155, 2021.

[61] A. D. Kersey, "A review of recent developments in fiber optic sensor technology," *Optical fiber technology*, vol. 2, no. 3, pp. 291–317, 1996.

[62] C. J. Carver and X. Zhou, "Polarization sensing of network health and seismic activity over a live terrestrial fiber-optic cable," *Communications Engineering*, vol. 3, no. 1, p. 91, 2024.

[63] S. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *Journal of Lightwave Technology*, vol. 1, no. 2, pp. 312–331, 2003.

[64] R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Optics Letters*, vol. 5, no. 6, pp. 273–275, 1980.

[65] E. Collett, *Field guide to polarization*. SPIE press Bellingham, 2005, vol. 15.

[66] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," *Transactions of the Cambridge Philosophical Society*, vol. 9, p. 399, 1851.

[67] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th. Cambridge, UK: Cambridge University Press, 1999.

[68] H. Poincaré, *Leçons sur la théorie mathématique de la lumière: professées pendant le premier semestre 1887-1888*. G. Carré, 1889, vol. 1.

[69] Luna Innovations, *POD-001 High Speed In-Line Polarimeter*, `https://lunainc.com/product/pod-001`, Accessed: 2025.

[70] R. M. Azzam, "Stokes-vector and Mueller-matrix polarimetry," *Journal of the Optical Society of America A*, vol. 33, no. 7, pp. 1396–1408, 2016.

[71] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," *Electronics Letters*, vol. 22, no. 19, pp. 1029–1030, 1986.

[72] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of the National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, 2000.

[73] G. Malik, M. U. Masood, M. Cheruvakkadu Mohamed, S. Straullu, S. K. Bhyri, G. Maria Galimberti, J. Pedro, A. Napoli, W. Wakim, and V. Curri, "Machine learning for predictive multi-event detection in fiber optic systems," in *2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, 2025, pp. 1–6.

[74] G. Malik, I. C. Dipto, M. U. Masood, M. C. Mohamed, S. Straullu, S. K. Bhyri, G. M. Galimberti, A. Napoli, J. Pedro, W. Wakim, and V. Curri, "Resilient anomaly detection in fiber-optic networks: A machine learning framework for multi-threat identification using state-of-polarization monitoring," *AI*, vol. 6, no. 7, 2025, ISSN: 2673-2688.

[75] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks/Cole, 1984.

[76] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[77] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[78] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[79] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3–42, 2006.

[80] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[81] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[82] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, 2016, pp. 785–794.

[83] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 3146–3154.

[84] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[85] D. W. H. Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd. Hoboken, NJ: John Wiley & Sons, 2013.

[86] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Institute for Signal and information Processing*, vol. 18, no. 1998, pp. 1–8, 1998.

[87] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[88] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, "Multi-layer perceptrons," in *Computational intelligence: a methodological introduction*, Springer, 2022, pp. 53–124.

[89] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[90] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[91] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[92] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105.

[93] L. M. Manevitz and M. Yousef, "One-class svms for document classification," *Journal of machine Learning research*, vol. 2, no. Dec, pp. 139–154, 2001.

[94] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD)*, Portland, OR, 1996, pp. 226–231.

[95] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 2, Montreal, Canada, 1995, pp. 1137–1143.

[96] Z. Zhan, M. Cantono, V. Kamalov, A. Mecozzi, R. Müller, S. Yin, and J. C. Castellanos, "Optical polarization–based seismic and water wave sensing on transoceanic cables," *Science*, vol. 371, no. 6532, pp. 931–936, 2021.

[97] A. E. Willner, S. Khaleghi, M. R. Chitgarha, and O. F. Yilmaz, "Monitoring and control of polarization-related impairments in optical fiber systems," *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 106–125, 2004.

[98] W. S. Saif, M. A. Esmail, A. M. Ragheb, T. A. Alshawi, and S. A. Alshebeili, "Machine learning techniques for optical performance monitoring and modulation format identification: A survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2839–2882, 2020.

[99] D. Wang, M. Zhang, Z. Li, J. Li, M. Fu, Y. Cui, and X. Chen, "OSNR and nonlinear noise power estimation for optical fiber communication systems using LSTM based deep learning technique," *Optics Express*, vol. 26, no. 16, pp. 21 346–21 357, 2018.

[100] F. N. Khan, C. Lu, and A. P. T. Lau, "Optical performance monitoring in fiber-optic networks enabled by machine learning techniques," *Optics and Fiber Technology*, vol. 37, pp. 1–6, 2017.

[101] F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, "A tutorial on machine learning for failure management in optical networks," *Journal of Lightwave Technology*, vol. 37, no. 16, pp. 4125–4139, 2019.

[102] M. Furdek, C. Natalino, A. Di Giglio, and M. Schiano, "Machine learning for optical network security monitoring: A practical perspective," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2860–2871, 2020.

[103] A. Lv, M. Li, Q. Sun, and J. An, "Phase-sensitive optical time-domain reflectometry with rayleigh enhanced optical fiber," *Photonic Sensors*, vol. 11, no. 1, pp. 79–107, 2021.

[104] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, and P. Layec, "Anomaly detection and localization in optical networks using vision transformer and sop monitoring," in *Optical Fiber Communication Conference (OFC)*, 2024, Tu2J.4.

[105] G. Malik, I. C. Dipto, M. U. Masood, M. C. Mohamed, S. Straullu, S. K. Bhyri, G. M. Galimberti, A. Napoli, J. Pedro, W. Wakim, and V. Curri, "Resilient anomaly detection in fiber-optic networks: A machine learning framework for multi-threat identification using state-of-polarization monitoring," *AI*, vol. 6, no. 7, 2025, ISSN: 2673-2688.

[106] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

Part II

Papers

PAPER A

Machine Learning-Based Polarization Signature Analysis for Detection and Categorization of Eavesdropping and Harmful Events

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Marija Furdek

Published in Optical Fiber Communications Conference and Exhibition (OFC), 24-28 March, 2024, pp. 1-3, San Diego, CA, USA.
 ISBN:979-8-3503-7758-3

The layout has been revised.

Abstract

We propose a methodology that uses State of Polarization (SOP) changes and Machine Learning (ML) to detect and classify eavesdropping, harmful, and non-harmful events in the optical fiber network. Our solution achieves 92.3% accuracy over 13 experimental scenarios.

1 Introduction

Optical fiber infrastructures are critical for handling a broad range of sensitive data, from military intelligence to personal information, across diverse environments such as expansive duct-based installations, submarine routes, and localized indoor networks. Recent years have marked an increase in sabotage attempts on these systems, alongside the ever-present risk of unauthorized data interception, which is exacerbated by advances in quantum computing [1], [2]. Optical fibers are particularly vulnerable to eavesdropping attacks, wherein unauthorized light coupling techniques such as evanescent coupling, V-groove cut, and micro/macro bending [3], [4] can be used to intercept data. While monitoring optical power levels is one way to detect eavesdropping attacks, it may not be applicable against those attacks that cause minimal or undetectable power level drops [5]. A more sophisticated technique than optical power tracking involves monitoring of polarization state changes at the receiver to distinguish normal system variations from eavesdropping attempts. Early work [6] introduced a system using Distributed Fiber Optic Sensing (DFOS) that could detect signatures from touching or manipulating a fence with installed fiber optical cables. However, reliance on Rayleigh and Brillouin backscattering due to fiber impurities made this solution complex. Furthermore, the need for high-speed pulsing lasers to determine the position of a breach based on backscattering pulse delays, coupled with the requirement for diplexers to filter amplified spontaneous noise, contributes to its high costs. The work in [7] investigated polarization signatures of different fiber events as sequences of polarization changes over a specific time and frequency window, derived by processing the polarization state in the Poincaré sphere (refer to Figure. 1.a). The signatures generated from eavesdropping and harmful events are visualized in a unique plot, referred to as a waterfall,

allowing a human security operator to visually distinguish between legitimate and unauthorized activities. This is a simpler and more cost-effective approach to malicious activity detection than the method from [6]. Nevertheless, the visualization-based technique has limited applicability and scalability due to the need of a human specialist analyzing the waterfall plots.

To overcome the scalability and cost limitations of existing human-dependent solutions, we introduce a novel methodology using ML algorithms to analyze polarization signatures. This paper is the first to experimentally collect and analyze a dataset containing eavesdropping attacks and other potentially harmful and non-harmful events for three cable types. Our methodology automates the process of analyzing and categorizing eavesdropping and potentially harmful events from normal operating conditions and non-harmful events, allowing for potential large-scale optical network deployments. The presented methodology successfully segregates signatures with an accuracy of 92.3%.

2 Data Collection and Proposed Methodology

In the act of unauthorized data interception from fiber optic cables, eavesdroppers physically manipulate the cables. These malicious acts generate unique polarization signatures that can be identified using ML techniques. In this study, ML algorithms analyze the signatures derived from Polarization State Variation (PSV)) data generated from experiments mimicking real-world conditions over fiber optic installations, including the risk of cable severance from nearby excavations and eavesdropping by manipulating exposed fibers. The proposed workflow is depicted in Figure.1. For data collection, a Continuous Wave Distributed Feedback (CW-DFB) laser with a polarization-maintaining fiber generates optical power at a specific wavelength. The laser is regulated by a driver that maintains a consistent power level and temperature. Subsequently, the laser emits polarized light into an installed transmission line at a wavelength occupying one channel in the O, E, S, C, or L-band. All other types of optical transmission could occupy the remaining free spectrum. Each external event produces a unique effect on the PSV that can be recorded by the optical analyzer employing the Poincaré sphere analysis technique in the polarization analyzer block (Figure.1.b). The sampling block generates samples of each polarization state on the Poincaré sphere every 1 ms (fulfilling the Nyquist theorem) over a 20-minute recording period, resulting in 1.2

million samples over the entire recording time for each event. The numerical value of the distance between two consecutive polarization states, referred to as Numerical Polarization State Variation (NPSV) (Numerical PSV).

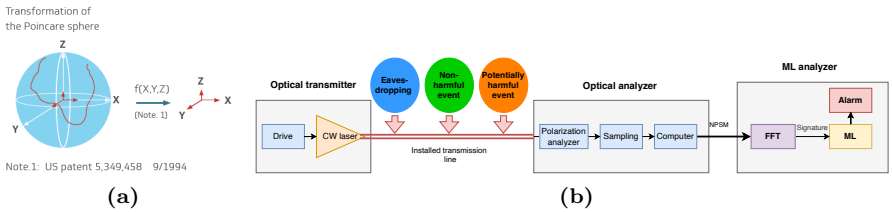


Figure 1: (a) Changes in the Poincaré sphere (b) Proposed methodology for extracting signatures and ML analyzer.

We partition the NPSV data into 1200 time slots of 1000 elements each, and apply a Fast Fourier Transform (FFT) analysis with 512 frequency bins, utilizing a Hamming window [8].The resulting signature for each specific event is power spectrum data of 1200 rows (corresponding to time slots) and 512 columns (corresponding to frequency bins). ML methods then analyze the data to detect the specific signatures and generate an alarm if an eavesdropping attempts or a threat to the installed transmission is identified (Figure.1.b). Our ML analysis uses data from 13 experimental scenarios, summarized in Table 1, aiming to distinguish between normal operational signatures and those from eavesdropping or harmful events. In our test bed, we use 1310 nm signal over a 2 km transmission line that consists of a series of military fiber optical tactical cable systems (Fiber Optical Tactical Cable System (FOCS); *fc*), indoor cables (*id*), and bare single-mode G.675 bend-insensitive fiber (*bf*).

The normal events include the relaxed (*rlx*) fiber without vibrations or eavesdropping, as well as vibrations at 155 Hz and 130 Hz (*n-v*) frequency (the two different values are used for diversity). The considered harmful events include fiber vibrations at 80 Hz (*an-v*), which corresponds to an excavator with an engine running at 4,800 rpm digging close to the cable installation, threatening to cut the cable. We also consider the case of dual vibrations (*dl-v*) at 80 and 130 Hz. The considered eavesdropping attacks are characterized by fiber bending (*b*) over a 10 mm diameter rod. We also consider the case without bending and with dual vibrations (*wb-dl-v*).

Table 1: The considered experimental scenarios

Abbr.	Scenario description	Justification
rlx	Relaxed fiber	Baseline; normal operating conditions
b-fc	FOCS cable bending	Eavesdropping
b-bf	Bare fiber bending	Eavesdropping
n-v-fc	FOCS cable + 155 Hz vibration	Normal operating conditions (non-harmful vibr.)
an-v-fc	FOCS cable + 80 Hz vibration	Harmful; possible cut predecessor
n-v-id	Indoor cable + 155 Hz vibration	Normal operating conditions (non-harmful vibr.)
an-v-id	Indoor cable + 80 Hz vibration	Harmful; possible cut predecessor
n-v-bf	Bare fiber + 155 Hz vibration	Normal operating conditions (non-harmful vibr.)
an-v-bf	Bare fiber + 80 Hz vibration	Harmful; possible cut predecessor
b-n-v-id	Indoor cable bending + 130 Hz vibration	Eavesdropping + non-harmful vibration
b-an-v-id	Indoor cable bending + 80 Hz vibration	Eavesdropping + harmful vibration
b-dl-v-id	Indoor cable bending + 80/130 Hz vib.	Eavesdropping + non-harmful and harmful vibrations
wb-dl-v-id	Indoor cable + 80/130 Hz vibrations	Non-harmful and harmful vibrations

The collected dataset was randomly divided into a 70% training (840 points) and a 30% testing subset (360 points), each with equal representation of the 13 scenarios. This led to a training dataset comprising $10,920$ samples and a testing set of $4,680$ samples. The labeled dataset with 13 distinct classes frames our analysis as a supervised ML problem (classification).

3 Results and Conclusion

We conducted experiments over a number of ML algorithms to select the most appropriate classifier for this 13-class classification problem. Our evaluation included the following classifiers from the Scikit-Learn library: eXtreme Gradient Boosting (XGBoost), Random Forest (RF), Gradient Boosting (GB), Bagging with Decision Trees, Decision Tree (DT), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), K-Nearest Neighbors (KNN), Multi Layer Perceptron (MLP) Neural Network, and Logistic Regression (LR). The classifiers were evaluated based on their accuracy and F1-score over the testing dataset. The final result is summarized in the Figure. 2.a. XGBoost performed the best, achieving an accuracy of 92.3% and an F1-score of 0.92, indicating a balanced performance in terms of false positives and false negatives. Random Forest and Gradient Boosted Trees closely follow the XGBoost performance.

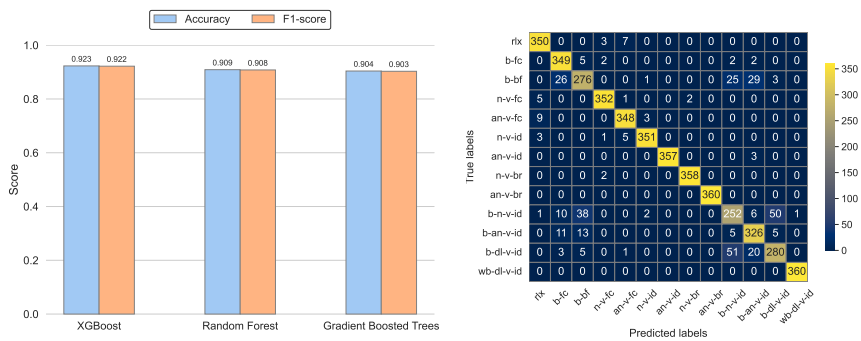


Figure 2: (a) Accuracy and F1 score for the top 3 classifiers (b) Confusion matrix of XGBoost for the test dataset.

As illustrated in Figure. 2.b, the confusion matrix validates the good performance of the XGBoost classifier. The classifier not only demonstrated proficiency in categorizing the relaxed one (*rlx*) and the scenarios without bending combined with dual frequency vibrations for indoor cables (*wb-dl-v-id*), but it also exhibited robust discrimination between harmful vibration events across the three cable types.

This achievement is particularly significant in enhancing optical network se-

curity as the designed classifier effectively distinguishes between typical signal behaviors and those altered due to harmful events and eavesdropping. However, discerning among bending data for bare fiber (*b-bf*), bending and 155 Hz vibration for indoor cable (*b-n-v-id*), and bending and vibrations in two frequencies for indoor cable (*b-dl-v-id*) presented some challenges with evident misclassifications.

In conclusion, this study underscores the critical importance of bolstering security within optical networks, particularly given the escalating vulnerabilities to covert eavesdropping and harmful events. Through an analysis of PSV data signatures from optical devices, we successfully employed ML techniques, specifically the XGBoost classifier, to detect and categorize eavesdropping and harmful events with a high accuracy. To the best of our knowledge, this is the first study that applies ML techniques to detect and categorize harmful and non-harmful events in optical networks with this category of polarization state changes data.

References

[1] Secure the Grid Coalition. "Attacks on fiber networks in california baffle fbi." Accessed: 2025-11-26. [Online]. Available: `https://securethegrid.com/attacks-on-fiber-networks-in-california-baffle-fbi/`.

[2] D. Sabbatini. "Italian navy, telecom provider team up to deter attacks on undersea cables." Accessed: 29-Sept-2023. [Online]. Available: `https://www.defensenews.com/naval/2022/07/14/italian-navy-telecom-provider-team-up-to-deter-attacks-on-undersea-cables/`.

[3] A. Harris and P. Castle, "Bend loss measurements on high numerical aperture single-mode fibers as a function of wavelength and bend radius," *Journal of Lightwave Technology*, vol. 4, no. 1, pp. 34–40, 1986.

[4] M. Zafar Iqbal, H. Fathallah, and N. Belhadj, "Optical fiber tapping: Methods and precautions," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, 2011, pp. 164–168.

[5] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

[6] Y. Aono, E. Ip, and P. Ji, "More than communications: Environment monitoring using existing optical fiber network infrastructure," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, W3G.1.

[7] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti, "Detection of abnormal activities on a SM or MM fiber," in *Optical Fiber Communication Conference (OFC)*, 2023, M3Z.6.

[8] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

PAPER B

Machine Learning Analysis of Polarization Signatures for Distinguishing Harmful from Non-harmful Fiber Events

Leyla Sadighi, Stefan Karlsson, Lena Wosinska, Marija Furdek

Published in 24th International Conference on Transparent Optical Networks (ICTON), 14-18 July, 2024, pp. 1-5, Bari, Italy.
 DOI:10.1109/ICTON62926.2024.10648140

The layout has been revised.

Abstract

Secure and reliable data transmission in optical networks is essential for supporting high-speed internet services. Optical fibers, the enabler of global connectivity for millions of users, are vulnerable to various potentially harmful events, including mechanical failures, like fiber cuts, and malicious physical layer attacks, such as eavesdropping. These incidents can degrade network performance, breach privacy and integrity through unauthorized access to the transmitted data, and cause significant financial and data loss. It is, therefore, crucial to detect and classify the malicious events. Continuous monitoring of polarization state changes, combined with the application of Machine Learning (ML) algorithms, enables the detection of deviations in the polarization patterns caused by the harmful events. In this study, we introduce a method that detects and identifies potentially harmful events in optical networks. By using a Histogram Gradient Boosting classifier within our machine learning framework, we achieve 97.94% detection accuracy of the harmful and non-harmful events.

1 Introduction

The capacity demand in communication networks is growing exponentially, aiming to support the increasing number of users as well as new and emerging online services. To keep up with this development, fiber optic networks offering ultra-high transmission capacity are considered as the future-proof technology. The optical networks carry a lot of data, including sensitive information, ranging from national defense communications to personal private data. Their crucial role in maintaining both short- and long-distance connectivity underscores the importance of safeguarding the integrity of transmitted information. However, the physical nature of these networks makes them vulnerable to sabotage from various sources, which can lead to degraded signal quality, disrupted communication services, and violated data security.

Various external events, regardless of intentions or aims, result with vibrations that impact signals transmitted over fiber optic installations. Examples

include non-harmful events, such as the normal vibrations in buildings, as well as potentially harmful incidents, such as the operation of an excavator near a fiber installation. When an excavator's engine works at 4800 Revolutions Per Minute (RPM), it generates 80 Hz vibrations, which indicates a fiber cut risk. Another severe security threat is eavesdropping and/or information tampering. While encryption algorithms are commonly employed to secure the data, advancements in computing power and the development of quantum computers threaten the efficiency of these security measures. Optical fibers can be eavesdropped relatively easily, and may be difficult to detect by measuring the received optical power [1]. These vulnerabilities underscore the importance of identification and categorization of harmful events, such as eavesdropping and harmful vibrations, to ensure the privacy and integrity as well as reliability of fiber optic communications.

A common way of detecting external fiber events is by monitoring the State of Polarization (SOP), i.e., tracking the polarization state of propagated light. Changes in the polarization state can indicate physical or environmental stress, such as mechanical vibrations and bending [2]. Continuous SOP monitoring is essential in optical networks for early detection of anomalies, enabling proactive measures against fiber damage [3].

A proactive fiber break detection in optical communication systems is proposed in [4]. It employs quaternion time series analysis and ML to classify various mechanical events induced by robotic arm movements, such as bending, shaking, small hits, and up-and-down movements. It utilizes SOP data transformation into quaternion sequences, later re-coded into relational data for event classification using a naive Bayes classifier with over 99% accuracy. The study in [5] introduces a transfer learning approach using a deep convolutional neural network for high-risk event classification given a small amount of SOP data. It demonstrates image-based ML as very efficient for detection and classification of five mechanical events in optical fiber: bending, shaking, small hits, up-and-down movements, and fan ventilation. A novel vision transformer-based method for detecting and localizing mechanical vibrations in optical networks using SOP data is proposed in [6]. In our previous work [7] we used supervised ML to detect and distinguish between 13 different polarization signatures in three cable types.

In this paper, we propose a polarization-based fiber optic sensor, where we apply supervised ML algorithms and the SOP data collected from 5 distinct

events occurring in an indoor optical fiber network. Our sensor can differentiate between potentially harmful and non-harmful events, enhancing the security and reliability of optical network infrastructures.

2 Experimental Setup

We generate polarization signatures by tracking changes in the polarization state on the Poincaré sphere. As depicted in Figure 1 a Continuous Wave Distributed Feedback (CW-DFB) laser serves as the primary light source. This laser inserts light into the optical fiber transmission line, thereby enabling the detailed examination of polarization changes. The optical power from the CW-DFB laser is initially transmitted through a 1 km coupling fiber, which is then connected to an indoor cable and a bare fiber segment used for eavesdropping simulations. Following this, the setup includes a 20 km fiber spool, resulting in a total transmission length of 21 km. The polarization of light transmitted through optical fibers can be changed by mechanical vibrations and eavesdropping attempts.

In our testbed, we collect unique signatures for specific types of manipulations that mimic both harmful and non-harmful events. The received optical signal is analyzed by an optical and an ML analyzer. Polarization signatures are created by deliberate actions applied between the 1 km coupling fiber and the 20 km coupling fiber (see Figure 1). Each type of fiber manipulation makes a distinctive impact on the Polarization State Variation (PSV). Following the method detailed in [7], we obtain a distinct signature for each type of event. The process starts with the sampling block capturing PSM samples from the polarization analyzer on the Poincaré sphere in 1 ms intervals over a 20-minute time period, resulting in 1.2 million samples per event. The system then computes the numerical distances between successive PSV values, thus generating data referred to as Numerical Polarization State Variation (NPSV). These NPSV measurements are organized into sets of 1000, creating distinct time segments. A Fast Fourier Transform (FFT) with frequency size of 512 is subsequently applied to these segments under a Hamming window, yielding a spectral power dataset comprising 1200 rows (each representing a time segment) and 512 columns (each for a frequency bin). With this approach, we generate a unique dataset for each event, creating a distinctive signature for each event type.

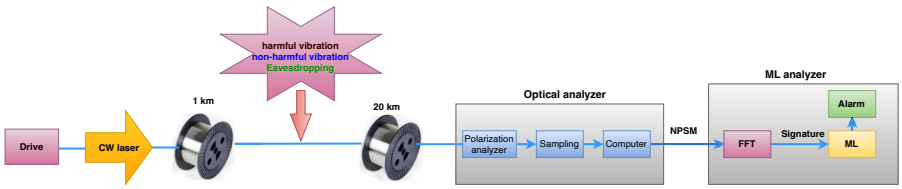


Figure 1: Experimental setup for analysis of polarization signatures. The manipulation is applied between the 1 km and the 20 km fiber spools.

3 Signatures and Data Collection

We consider two types of harmful events in an indoor cable installation: eavesdropping and vibrations at a frequency matching that of an excavator, as well as non-harmful events at a different frequency.

In the eavesdropping (*eav*) scenario, we perform subtle manipulations of the indoor cable, such as subjecting it to a pulling force while it is bent. As referenced in [1], eavesdropping can cause optical power attenuation of less than 0.3 dB, a level typically undetectable by Optical Time Domain Reflectometry (OTDR). The signature for this scenario is shown in Figure 2.b. The second type of considered potentially harmful events involves fiber vibrations at 80 Hz, typically resulting from activities such as an excavator digging near the cable installation and posing a risk of cutting the fiber. The 80 Hz vibration is generated by a loudspeaker, positioned 1 km away from the piezoelectric vibrator. This setup produces vibrations on the indoor cable. We collected signatures of the 80 Hz vibration, denoted as *hrmf_80Hz_vb*, (see signature in Figure 2.d), as well as the combined 80 Hz harmful vibration and 140 Hz non-harmful vibration, denoted as *hrmf_80Hz_140Hz_vb*, (see signature in Figure 2.e) with 210 Hz overtone as potentially harmful events.

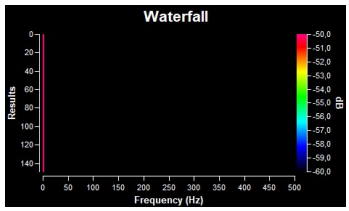
The non-harmful events, denoted as *nhrmf_140Hz_vb*, include any normal vibrations that do not pose a risk to the infrastructure, e.g., vibration caused by a fan in a building. For this study, we collected non-harmful vibration data at a frequency of 140 Hz, generated by a piezoelectric vibrator directly attached to the indoor cable (see signature in Figure 2.c). A baseline scenario, denoted as *rlx*, represents normal operating conditions, characterized by the absence of harmful vibrations, non-harmful vibrations, or eavesdropping activities (see signature in Figure 2.a).

Our ML models use data from the above five use cases to differentiate between normal operation patterns and those suggesting eavesdropping attempts, potential harmful vibrations, and non-harmful vibrations. The dataset is organized into five distinct classes, resulting in a supervised ML classification problem. We randomly divide the dataset into training and testing subsets, with 70% (840 points) and 30% (360 points) respectively, with equal representation across the five scenarios. As a result, the training set comprises 4,200, and the testing set 1,800 samples.

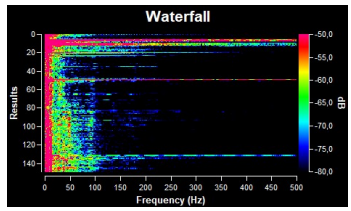
4 Results

We conduct a comprehensive evaluation of multiple supervised ML algorithms to identify the most effective classifier for our five-class classification problem, tailored for our specific dataset. This evaluation utilizes several classifiers from the Scikit-Learn library, selected based on their potential applicability and performance metrics in similar scenarios: eXtreme Gradient Boosting (XGBoost), Random Forest (RF), Bagging with Decision Trees, Decision Tree (DT), Histogram Gradient Boosting (HGB), Gradient Boosting (GB), Support Vector Machine (SVM), Logistic Regression (LR), Extra Trees (ET) Classifier, Bagging Classifier, K-Nearest Neighbors (KNN), Multi Layer Perceptron (MLP) Neural Network, and Linear Discriminant Analysis (LDA). We assess these classifiers based on their accuracy and the F1-score using the testing dataset. The results for the top-performing four classifiers are summarized in Figure 3. the HGB and the GB classifiers outperform the others, achieving an impressive accuracy of 97.94% and an F1-score of 0.9794. However, GB demands ten times longer training time. XGBoost and SVM also deliver robust performance, closely matching that of the HGB.

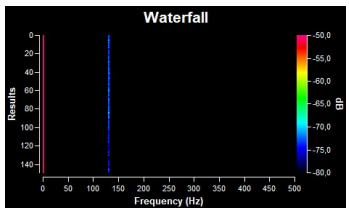
The confusion matrix for the HGB classifier in Figure 4 shows detailed performance across the five classes. The classifier correctly identifies 97.78% of the baseline (rlx) instances, with a minor misclassification rate of 1.39% into 'eavesdropping' (eav) and 0.83% into 'non-harmful vibration' (nhrmf_140Hz_vb) classes. The 'non-harmful vibration' class achieves a high accuracy of 99.44%, with only 0.56% samples misclassified as baseline. The 'eavesdropping' (eav) class was accurately identified in 98.61% of the samples, with a 1.39% misclassification into baseline. For the two potentially harmful vibration events, the classifier correctly identifies 95.0% of the 'hrmf_80Hz_vb' instances, but there



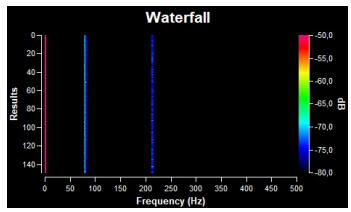
(a) Relaxed fiber



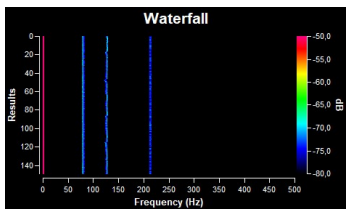
(b) Eavesdropping



(c) Normal vibration 140 Hz



(d) Harmful vibration 80 Hz with 210 Hz overtone



(e) Harmful vibration 80 Hz + normal vibration 140 Hz (with 210 Hz overtone)

Figure 2: Waterfall (visual representation of signatures as defined in [7]) for five considered scenarios.

are some misclassifications: 3.33% of samples is incorrectly identified as baseline, 0.28% as 'nhrmf_140Hz_vb', and 1.39% as 'hrmf_80Hz_140Hz_vb'. The 'hrmf_80Hz_140Hz_vb' class has a high accuracy of 98.89%, with only 1.11% misclassified as 'hrmf_80Hz_vb'. These results highlight the effectiveness of the HGB classifier in distinguishing between different types of events in fiber optic networks. The high accuracy and F1-scores across most classes demonstrate the model's robustness in identifying both harmful and non-harmful events.

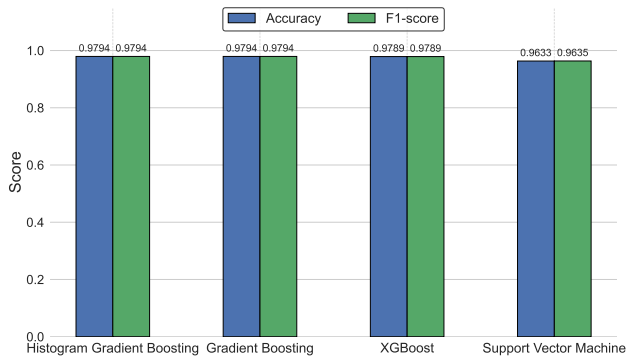


Figure 3: Results of four top-performing ML classifiers

5 Conclusion

In this study, we analyze polarization signatures and propose a method to detect potentially harmful events in optical fiber networks. Among the tested classifiers, the HGB demonstrated superior performance, achieving an accuracy of 97.94% and an F1-score of 0.9794. Detailed analysis using the confusion matrix reveals high precision in distinguishing between the various types of anomalies, particularly in identifying harmful events such as fiber vibrations caused by excavator activities. The results highlight the robustness and reliability of the proposed method in detecting and classifying different types of anomalies in optical networks. Future work will focus on expanding the dataset to include a broader range of scenarios and refining the machine learning models to further improve detection accuracy and computational efficiency.

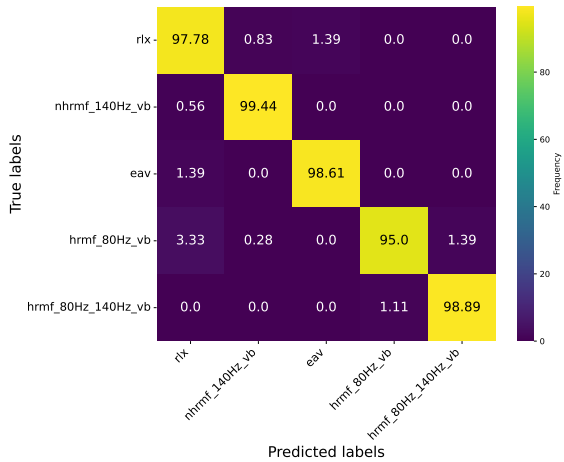


Figure 4: Results of Confusion Matrix for Histogram Gradient Boosting classifier

References

[1] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

[2] J. Pesic, E. Le Rouzic, N. Brochier, and L. Dupont, "Proactive restoration of optical links based on the classification of events," in *15th International Conference on Optical Network Design and Modeling-ONDM 2011*, IEEE, 2011, pp. 1–6.

[3] F. Boitier, V. Lemaire, J. Pesic, L. Chavarría, P. Layec, S. Bigo, and E. Dutisseuil, "Proactive fiber damage detection in real-time coherent receiver," in *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[4] V. Lemaire, F. Boitier, J. Pesic, A. Bondu, S. Ragot, and F. Clérot, "Proactive fiber break detection based on quaternion time series and automatic variable selection from relational data," in *Advanced Analytics and Learning on Temporal Data: 4th ECML PKDD Workshop,*

AALTD 2019, Würzburg, Germany, September 20, 2019, Revised Selected Papers 4, Springer, 2020, pp. 26–42.

[5] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Breaking boundaries: Harnessing unrelated image data for robust risky event classification with scarce state of polarization data," in *European Conference on Optical Communications (ECOC)*, IET, vol. 2023, 2023, pp. 924–927.

[6] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, and P. Layec, "Anomaly detection and localization in optical networks using vision transformer and sop monitoring," in *Optical Fiber Communication Conference (OFC)*, 2024, Tu2J.4.

[7] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

PAPER C

Detection and Classification of Eavesdropping and Mechanical Vibrations in Fiber Optical Networks by Analyzing Polarization Signatures Over a Noisy Environment

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek

Presented in 50th European Conference on Optical Communication (ECOC),
[invited paper]
28–26 September, 2025, pp. 527–530, Frankfurt, Germany.
 ISBN: 978-3-8007-6426-6

The layout has been revised.

Abstract

We propose an Machine Learning (ML)-based method to detect and classify eavesdropping and mechanical vibrations in an optical network based on State of Polarization (SOP) variations. Tests in two real-world installations with links of different lengths demonstrate an accuracy of 86.5% in 7 distinct normal and malicious scenarios.

1 Introduction

Cyber security is widely recognized as a critical concern due to the immense significance of online services and information transmitted over communication networks. In turn, the security of fiber optical network infrastructure as the foundation of global communications is rapidly gaining relevance. The recent years have observed an increasing number of confirmed sabotage attempts on fiber optical installations worldwide[1], [2], [3], which could have a high impact on the global connectivity, economy and defense strategies. The risk of fiber eavesdropping and/or tampering with sensitive information is also becoming severe. An eavesdropper can couple out light from an optical fiber relatively easily. Recent study [4] sheds light on the vulnerability of optical fiber systems to eavesdropping. The information transmitted in the fiber can be detected by an eavesdropper tapping a certain percentage of the light. The success of such eavesdropping attempts strongly depends on the technique employed for tapping the light and the distance of the breach from the transmitter. Gaining access to the optical signal within a certain distance from the transmitter enables an eavesdropper to detect sensitive data. This necessitates the development of eavesdropping detection strategies capable of accurately identifying eavesdropping activities even amidst the prevalent noise in fiber optical networks.

Optical fiber tampering causes changes of the polarization state of the carried light. In general, Polarization State Variation (PSV) data offers crucial insights into the polarization characteristics of light signals in a network. Continuous monitoring of the SOP has been demonstrated as essential for prompt identification of network disruptions, enabling early detection of potential fiber damage[5]. Close examination of SOP changes at the receiver and comparison

with variations associated to normal system behaviour can enable detection of eavesdropping attempts that cannot be discovered by monitoring the received optical power. This technology was demonstrated in [6]. Earlier research looked into naive Bayes classifiers to spot vibrations in optical fibers caused by mechanical stress [5], [7]. Vibrations, created by robotic arm movements, were detected with a coherent receiver. Additionally, [8] suggested a transfer learning method to classify high-risk events using limited SOP data. In [9], we experimentally collected and analyzed 13 distinct polarization signatures using a supervised ML algorithm. The results demonstrated that our model could accurately detect and differentiate between signatures from eavesdropping attacks and other potentially harmful and non-harmful events, achieving an accuracy of 92.3%.

In this paper, we study how polarization signatures can be recorded and classified, originating from an installed transmission line in a real-life network, OpenIreland, operated by Trinity College Dublin (TCD) and located under the street in Dublin, Ireland. Analyzing two separate installations with transmission lengths of 0.15 km and 10.5 km, respectively, we obtained 86.5% accuracy in classifying the signatures using supervised ML.

2 Experimental setup

The experimental setup is illustrated in Fig.1. A Continuous Wave Distributed Feedback (CW-DFB) is used as a transmitter to inject light into the transmission line. The optical power from the CW-DFB laser is first transmitted through a 1 km coupling fiber and then connected to the cable installation. Another 1 km long coupling fiber is added before the receiver. The two fiber cable installations, 0.15 km and 10.5 km long, are used with two fibers connected in a loop, resulting in the total transmission length of either 2.3 km or 23 km. The launch optical power is approximately -10 dBm and the received optical power is in the range of -11 to 21 dBm. The polarization of the transmitted light is affected by the vibrations originating from the street traffic taking place during the experiment. The cable duct consists of a fiber blown into a tube, which partially protects against vibrations but allows direct contact with the tube wall, leading to vibration-induced noise caused by the street traffic.

Accessing the fiber for manipulation requires altering the cable. To reliably

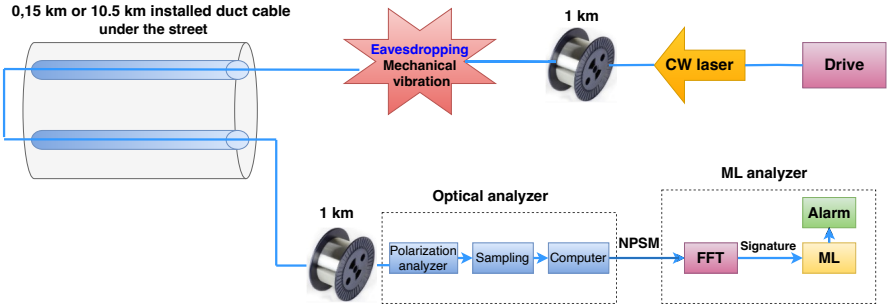


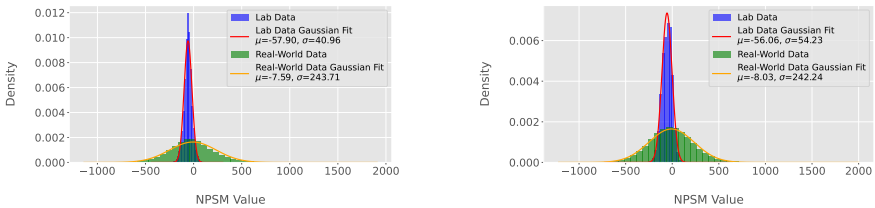
Figure 1: Experimental setup for analyzing polarization signatures in installed cables with 0.15 km and 10.5 km

detect vibration signatures related to this manipulation, they must notably exceed the background noise, which may arise from benign sources like non-harmful vibrations or potentially harmful activities such as eavesdropping and nearby excavation. The polarization signatures are collected by monitoring the PSV on the Poincaré sphere. The received optical signal is analyzed by an optical and an ML analyzer.

We adopt the analytical procedure from [9] to derive a unique signature for each event type. The process begins with the sampling block capturing PSV samples from the polarization analyzer on the Poincaré sphere every 0.5 ms during a 20-minute recording period, resulting in 2.4 million samples for each event. The system then calculates the numerical value of distances between consecutive PSV, generating what is referred to as Numerical Polarization State Variation (NPSV) data. The NPSV values are grouped into batches of 500, forming individual time slots. A Fast Fourier Transform (FFT) analysis is then performed on these segments using a Hamming window [10], producing a power spectrum dataset with 4800 rows (each corresponding to a time slot) and 512 columns (each representing a frequency bin). This dataset forms the unique signature for each specific event. ML techniques are then applied to the data to identify distinct signatures and trigger an alarm if a threat to the transmission line is detected.

3 Definition of signatures and data collection

Our ML analysis uses data from seven real-life network signatures to differentiate between polarization patterns obtained during normal operation and those suggesting malicious vibrations and eavesdropping. We denote normal events as relaxed (*rlx*) and soft bending (*sbd*) fiber scenarios. The *rlx* scenario identifies a baseline scenario without eavesdropping, vibrations, or bending for both the 0.15 km (*rlx-0.15km*) and the 10.5 km (*rlx-10.5km*) fiber installation. *sbd* involves only gentle bending of the cable to assess its resilience to benign environmental stress. Signatures from soft bending events are collected for the 0.15 km (*sbd-0.15km*) and 10.5 km (*sbd-10.5km*) installations. Eavesdropping on the 0.15 km installation (*eav-0.15km*) scenario assesses the ability to detect unauthorized interception attempts by observing subtle manipulations of the cable, such as subjecting it to a pulling force while it is bent. As referenced in[4], eavesdropping can cause optical power attenuation below 0.3 dB, a level typically undetectable by an Optical Time Domain Reflectometry (OTDR). The objective of collecting this signature is to determine if our model can distinguish this eavesdropping activity from soft bending (*sbd-0.15km*). Potentially harmful events considered include fiber vibrations at 80 Hz (*80vb*), typically corresponding to an excavator digging close to the cable installation and threatening to cut the cable, here generated by a loudspeaker. Vibration data at this frequency is gathered for the 0.15 km (*80vb-0.15km*) and the 10.5 km (*80vb-10.5km*) installation. The collected dataset is randomly partitioned into training and testing subsets, containing 70% (3360) and 30% (1440) points, respectively, to ensure equal representation of the seven scenarios. Consequently, the training set consists of 23,520 samples, while the testing set contains 10,080 samples. This dataset, labeled into seven distinct classes, represents a supervised ML classification problem.
Figure 2 shows a comparison of the data collected in the lab [9] and the real-world data for the relaxed (Figure 2.a) and 80 Hz vibration (Figure 2.b) scenarios, revealing distinct behavior. While the lab data exhibits a narrower distribution and lower variability, the real-world data shows a wider spread and increased variability. These differences highlight how environmental conditions affect real-world data by introducing more noise, thereby making the task of detection and classification for ML algorithms more complex.



(a) Relaxed scenario **(b)** 80 Hz vibration

Figure 2: Comparison between the data collected in the lab [9] and the real-world data for the 10.5 km link.

4 Results

We evaluate several ML algorithms to determine a suitable classifier for our seven-class classification problem. The assessment involves the following classifiers from the Scikit-Learn library: eXtreme Gradient Boosting (XGBoost), Histogram Gradient Boosting (HGB), Gradient Boosting (GB), Support Vector Machine (SVM), Logistic Regression (LR), Extra Trees (ET) Classifier, Bagging Classifier, and Linear Discriminant Analysis (LDA). The classifiers are evaluated in terms of their accuracy and F1-score on the testing dataset after training. The final result for the three top-performing classifiers is summarized in Figure 3. HGB outperforms other models in the real-world dataset, achieving an accuracy of 86.5% and an F1-score of 0.866. XGBoost also demonstrates accurate results, closely matching the performance of HGB. The confusion matrix of the HGB classifier in Figure 4 demonstrates good performance, achieving high accuracy across different scenarios. A clear impact of the link length is identifiable from the matrix, where the accuracy is higher for the shorter than for the longer link. For the 0.15 km link, the model achieves 91.04% accuracy for *rlx-0.15km*, 98.47% for *sbd-0.15km*, 88.54% for *eav-0.15km*, and 99.65% for *80vb-0.15km*.Notably, while the 88.54% accuracy for *eav-0.15km* reflects good performance, it is slightly poorer than in the other shorter link scenarios. This suggests that the model is effective at classifying *eav-0.15km*, but a few instances of misclassification, primarily with *rlx-0.15km* (4.44%) and *sbd-0.15km* (3.19%), can occur, possibly due to similarities among these event types at this distance. The accuracy is lower

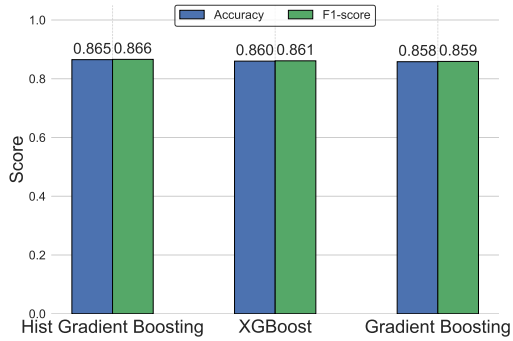


Figure 3: Accuracy and F1-score for the top 3 classifiers

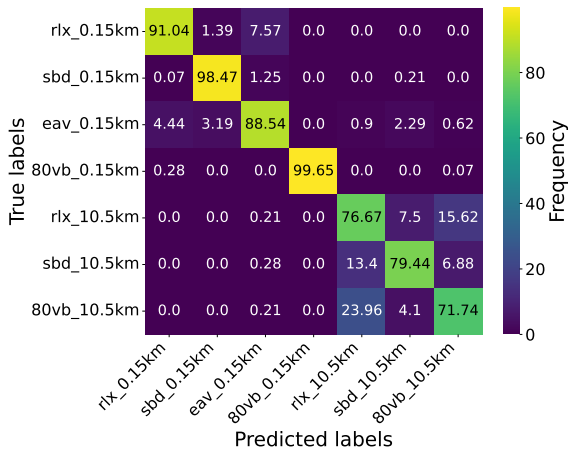


Figure 4: Confusion matrix of Histogram Gradient Boosting (HGB)

for the longer link, with 76.67% for *rlx-10.5km*, 79.44% for *sbd-10.5km*, and 71.74% for *80vb-10.5km*, with notable misclassifications primarily between *rlx-10.5km* and *80vb-10.5km* (15.62% and 23.96%) and between *sbd-10.5km* and *rlx-10.5km* (7.5% and 13.4%). These results highlight the good performance of the model, particularly for short links, while identifying areas for improvement in distinguishing between similar event types over longer distances.

5 Conclusion

This study demonstrated the effectiveness of using PSV data and supervised ML to detect and classify mechanical vibrations and eavesdropping in fiber optic networks. Trained and tested with data collected from a real-world installation in Dublin urban area, our method achieved 86.5% accuracy in event identification, underscoring its practical applicability for enhancing network security. While the classification of a short-distance link showed high accuracy, improvements are needed for longer distances. Therefore, improving the data collection method as well as the applied ML models are considered as the future work.

6 Acknowledgments

Special thanks to Daniel Kilper and Frank Slyne, at Trinity College Dublin for their assistance and for making the testbed available for the experiment. This work was supported by Vetenskapsrådet (2023-05249).

References

[1] Secure the Grid Coalition. "Attacks on fiber networks in california baffle fbi." Accessed: 2025-11-26. [Online]. Available: `https://securethegrid.com/attacks-on-fiber-networks-in-california-baffle-fbi/`.

[2] D. Sabbatini. "Italian navy, telecom provider team up to deter attacks on undersea cables." Accessed: 29-Sept-2023. [Online]. Available: `https://www.defensenews.com/naval/2022/07/14/italian-navy-telecom-provider-team-up-to-deter-attacks-on-undersea-cables/`.

[3] .

[4] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

[5] F. Boitier, V. Lemaire, J. Pesic, L. Chavarría, P. Layec, S. Bigo, and E. Dutisseuil, "Proactive fiber damage detection in real-time coherent receiver," in *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[6] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti, "Detection of abnormal activities on a SM or MM fiber," in *Optical Fiber Communication Conference (OFC)*, 2023, M3Z.6.

[7] V. Lemaire, F. Boitier, J. Pesic, A. Bondu, S. Ragot, and F. Clérot, "Proactive fiber break detection based on quaternion time series and automatic variable selection from relational data," in *Advanced Analytics and Learning on Temporal Data: 4th ECML PKDD Workshop, AALTD 2019, Würzburg, Germany, September 20, 2019, Revised Selected Papers 4*, Springer, 2020, pp. 26–42.

[8] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Breaking boundaries: Harnessing unrelated image data for robust risky event classification with scarce state of polarization data," in *European Conference on Optical Communications (ECOC)*, IET, vol. 2023, 2023, pp. 924–927.

[9] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[10] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

PAPER D

Deep Learning for Detection of Harmful Events in Real-World, Noisy Optical Fiber Deployments

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek

Published in IEEE Journal of Lightwave Technology (JLT),
vol. 43, no. 4, pp. 1125–1139, February 2025.
DOI: 10.1109/JLT.2025.3557748,

The layout has been revised.

Abstract

Optical network infrastructure underpins global communication networks. It is exposed to various physical layer breaches, such as fiber cuts or eavesdropping via fiber bending, that may violate privacy or disrupt services. Analyses of State of Polarization (SOP) variations induced by external events, combined with Machine Learning (ML) techniques, can contribute to early identification of events and categorization of potential threats. However, real-world deployment of automated threat detection and mitigation faces many challenges, including the inconsistencies between controlled laboratory settings, often used for dataset collection for ML training, and real-world, noisy environments. In this paper, we study the detection of external disturbances in real-world fiber installations by analyzing the induced changes in the SOP of optical signals. We develop a suite of Deep Learning (DL) models, including One-Dimension (1D) Convolutional Neural Network (CNN) and fully-connected dense layers, for the detection of harmful events in noisy environments comprising a shorter (300 m) and a longer (21 km) fiber link installation, corresponding to Fiber to the Home (FTTH) and metro-scale optical paths, respectively. The proposed approach employs an optical analyzer to capture SOP changes resulting from mechanical or acoustic vibrations, as well as eavesdropping attempts. Upon careful tuning of the DL models' hyper-parameters, 98.57% accuracy is obtained for the shorter, and 92.26% for the longer link installation.

1 Introduction

Optical fibers are critical to global communication infrastructure. For example, by 2021, over 43% of U.S. and 60% of Canadian households had access to an optical fiber network, reflecting an annual growth rate of 12% [1]. In Sweden, fiber internet penetration has also grown significantly over the past decade, reaching approximately 82.6% of households in 2020, up from 33% in

2010 [2]. This extensive reach underscores the importance of fiber-optic networks, not only in connecting individual households but also in forming the backbone of national and international communication systems. The backbone is the most critical part of the infrastructure, as it interconnects various regions and countries [3], ensuring that data can flow seamlessly across vast distances. Disruptions or breaches in this segment can have widespread effects, impacting numerous users and services simultaneously. Therefore, secure and resilient optical network transmission is of primary concern, where the rippling effects of interruptions or breaches at the physical layer can disrupt a multitude of upper-layer services. Optical fibers are vulnerable to physical tampering, particularly from human activities such as construction, which can compromise network integrity and lead to issues such as fiber cuts. In addition, fiber optic networks are vulnerable to security threats from eavesdropping techniques such as evanescent coupling and fiber bending [4], [5]. These techniques enable attackers to covertly intercept data transmissions without noticeable disruptions to network performance, thereby jeopardizing confidentiality, integrity, and privacy of transmitted information. To address these vulnerabilities, comprehensive monitoring of events that can compromise optical networks—whether caused by human activity or environmental factors—is essential for enhancing network resilience and minimizing the data, financial, or reputational losses caused by disruptions in telecommunication services [6].

To exploit the potential of optical fiber as an environmental sensor and address the diverse monitoring demands of modern optical networks, various fiber-optic sensing methodologies have been developed. While traditional sensing techniques, such as optical interferometry [7], offer high sensitivity with sub-nanometer spatial resolution and precise phase shift detection [8], [9], their reliance on ultra-stable laser sources or specialized hardware limits their practicality in large-scale deployments. Similarly, Optical Time Domain Reflectometry (OTDR) [10], [11] and its coherent counterpart [12] provide robust fault localization but exhibit trade-offs between simplicity, accuracy, and the added complexity and cost of coherent systems. Distributed Acoustic Sensing (DAS) [13], [14] offers higher sensitivity and finer spatial resolution for detecting vibrations and mechanical disturbances compared to traditional fiber sensing techniques like OTDR and interferometric methods [15], [16]. This capability allows DAS to detect subtle environmental changes along the

fiber with high accuracy. However, traditional DAS implementations often depend on expensive, high-power interrogators and were typically deployed on dedicated dark fibers, which precluded simultaneous data transmission and presented significant financial and operational barriers to widespread adoption. Recent advancements, such as L-band DAS systems, have enabled the co-propagation of sensing and data signals over the same fiber without significant interference, thereby enhancing scalability and cost-effectiveness [17], [18]. Nonetheless, the need persists for more scalable and cost-effective fiber-optic sensing solutions that can seamlessly integrate with existing telecommunications infrastructure.

SOP sensing emerges as a transformative approach for addressing the limitations of existing fiber sensing technologies [19]. This approach involves tracking of the Stokes parameters of light, a four-dimensional vector ($\mathbf{S} = [S_0, S_1, S_2, S_3]$) that defines its polarization state over time [20]. Unlike DAS and OTDR-based methods that require costly hardware, dedicated resources, or complex configurations, SOP-based sensing leverages the intrinsic capabilities of carrier-grade coherent optical transceivers, which are designed to meet stringent reliability, performance, and interoperability standards required in commercial telecommunication networks. These transceivers measure SOP as part of their Digital Signal Processing (DSP) for data demodulation, enabling seamless integration of sensing functions into live, traffic-carrying optical fibers [21]. This scalability allows for the accurate monitoring of environmental and mechanical disturbances over extensive distances, making SOP sensing a practical option for addressing the operational and economic constraints of modern optical networks. By utilizing standard network hardware, SOP-based sensing provides a technically viable solution for integrating monitoring capabilities into existing systems without requiring significant additional infrastructure.

To fully leverage SOP sensing in fiber-optic monitoring, advanced data processing is vital to extract useful patterns from complex polarization fluctuations. The evolving threats to transmission systems may appear as events at various frequencies or frequency bands, inducing diverse effects on the SOP variations. Traditional monitoring systems that rely on predefined rules and thresholds often struggle to adapt to these complex and dynamic SOP variations [22]. This is due to the dependence of such systems on manual rule creation and maintenance, which makes them less effective in dynamic environments. To address these challenges, integrating ML techniques into mon-

itoring systems offers a more adaptive and robust solution. ML models can automatically learn to identify intricate patterns in SOP variations from the data, without the need for manual rule definition. This capability enhances the system's ability to detect and classify disturbances, even as threats evolve. ML, and Supervised Learning (SL) in particular, play a critical role by using labeled datasets to train models and achieve accurate classification of disturbances. Traditional SL methods, such as those based on gradient boosting or support vector machine, have been widely employed in the literature to identify and classify different types of tampering and natural disasters over experimental testbeds [23], [24], [25] due to their simplicity and effectiveness in low-dimensional feature spaces.

DL models stand out due to their ability to effectively process complex and high-dimensional datasets, which is a characteristic of SOP data. Several studies have explored the application of DL models, including Deep Neural Networks (DNN) and CNN, for event classification of OTDR traces data [26], [27]. The work in [28] presented a CNN-based approach for detecting and localizing reflective events in the presence of noisy OTDR data. These DL models are implemented over OTDR data, relying on expensive and specialized hardware to measure backscattered light and detect changes in the optical signal. In our study, we address this limitation by utilizing SOP variation data, which monitors temporal variations in the polarization state, providing complementary information on fiber disturbances without relying on backscattered signal analysis or requiring specialized measurement equipment. To effectively address the added complexity from environmental noise, we implement a combination of 1D convolutional and fully-connected dense layers for accurate event classification.

This paper extends our previous work in [29]. Therein, we analyzed SOP variations data for seven different types of events over the same testbed considering two separate installations with transmission lengths of 0.15 km (shorter link) and 10.5 km (longer link) in each direction, respectively. Using an SL technique called Histogram Gradient Boosting (HGB), we achieved an accuracy of 86.5% in event detection and categorization through SOP variation analysis. This result demonstrated that, while the considered ML model achieved reasonable accuracy, its performance was far from that expected for use in production deployments. Building on this foundation, the present study explores the performance of the DL models and focuses on the relationship

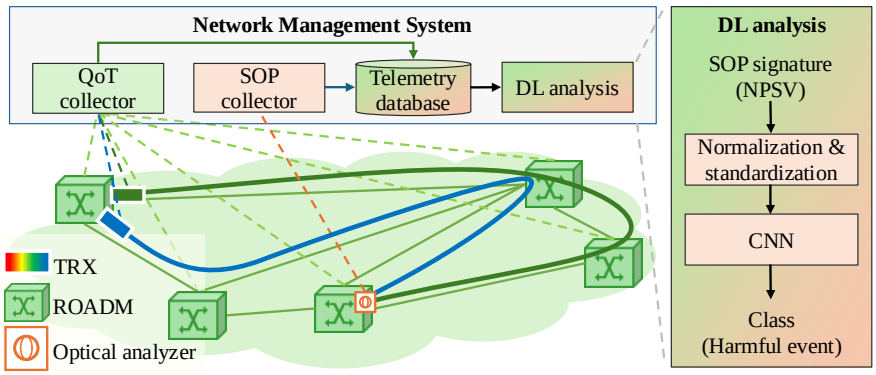


Figure 1: Architecture envisioned for the DL-based analysis of SOP data.

between the complexity of the model, measured by the number of trainable parameters, and accuracy. Additionally, we expand the scope of collected events to include eavesdropping and combinations of eavesdropping with different types of vibrations specifically for the long-link installation, which have not been previously investigated in noisy environments.

The remainder of the paper is structured as follows. Section 2 details the considered DL-based framework used for SOP analysis. Section 3 describes the experimental setup and the data collection process for generating polarization signatures under various conditions. Section 4 details the employed DL models, focusing on the main architecture and hyper-parameter tuning. Section 5 analyzes the experimental results, and Section 6 concludes the paper.

2 Deep Learning-Based Analysis of State of Polarization Data

In this work, we consider a Software-Defined Networking (SDN) scenario, illustrated in Figure 1, where devices form the data plane and are controlled by the control plane. The data plane is composed of traditional infrastructure such as Reconfigurable Optical Add-Drop Multiplexers (ROADMs) and Transceivers (TRXs). Moreover, optical analyzers are strategically deployed at critical network points to directly capture and quantify changes in the SOP of transmitted light. These optical analyzers operate as inline devices in

the transmission path, collecting SOP data and converting it into Numerical Polarization State Variation (NPSV) data that characterizes different events which affect the transmission line. The generated NPSV data is transferred to the control plane in addition to traditional Quality of Transmission (QoT) monitoring data from ROADMs and TRXs, and delivered to a centralized Network Management System (NMS) for detailed analysis. Although the envisioned framework relays additional information from ROADMs and TRXs to the NMS, only the Stokes parameters (obtained from the optical analyzer) are considered in our experimental validation. However, in principle, additional QoT metrics such as measured loss due to bending can also be analyzed. This loss can be evaluated end-to-end or at intermediate ROADMs along a path, and can serve as an indication of the segment where the loss or bending occurred.

At the NMS, an ML-based monitoring system (referred to as *DL analysis* in the figure) processes the collected data to derive actionable insights for network management and security. Pre-processing techniques, such as normalization and standardization, are applied to ensure compatibility of the data with ML algorithms and to enhance the precision of the analysis. While Unsupervised Learning (USL) and Semi-Supervised Learning (SSL) are useful for detecting abnormal patterns and emerging threat types, their lack of precision and limitation in distinguishing exact attack types often limits their applicability. To maintain network reliability and security, it is crucial to differentiate between benign events, such as those caused by routine vibrations, and malicious activities, such as eavesdropping attempts aimed at intercepting data or physical threats like fiber cuts resulting from sabotage or excavation. The current framework emphasizes DL techniques, which excel in detecting and classifying specific event categories, providing actionable insights and enhancing network monitoring efficiency. In [29], we demonstrated robust performance of traditional SL methods in detecting and classifying various events in real-world, noisy environments. Here, we extend our preliminary study by exploring the application of DL models for processing the SOP data collected in optical networks. By leveraging the DL capacity to handle high-dimensional inputs, we aim at improving the accuracy of event classification in real-world, noisy environments.

Ultimately, the anomalies detected by the ML models are conveyed to the NMS, which plays a central role in overseeing network health. The NMS

analyzes the generated alerts, initiates appropriate responses, and ensures that the network remains secure and resilient against potential threats.

3 Experimental Setup

3.1 OpenIreland Testbed

The data used in our experiments is collected from the real-world fiber installation, a part of the OpenIreland testbed infrastructure at Trinity College Dublin (TCD) [30]. The testbed provides the ability to set up a large metro network and links TCD to different field fibers.

In this sensing experiment, we utilized two field-deployed fibers. Figure 2 depicts the approximate routes of these fiber sections. These sections are subject to varying levels of environmental noise, which distinguishes them from a laboratory fiber spool that can be considered nearly noiseless. The fiber links consist of conventional Single Mode (SM) G.652 or G.657 cables. Both fibers are set up in a loop back, so that we can transmit and receive from the same location. The shorter section we considered (referred to as the short-link installation) consists of a 150 m fiber path linking two OpenIreland locations within TCD, but traversing an external public road that is subject to external environmental effects. Since the optical signal is transmitted and received at the same location in the lab, the actual fiber length traversed is 300 m (round-trip).

Similarly, the longer section we considered (referred to as the long-link installation) consists of a fiber connection between TCD and Dublin City University (DCU) on the north side of Dublin. While the length of a one-way link between these sites is 10.5 km, the round-trip distance covered by the optical signal in our experimental setup is 21 km. This longer connection serves as a realistic representation of a metro network environment.

A looped fiber configuration primarily serves to emulate a longer transmission distance, allowing us to analyze the impact of noise and environmental disturbances over an extended fiber length. The key difference between a looped and a one-way transmission lies in the fact that the round-trip setup inherently collects more noise and polarization fluctuations compared to a single-trip scenario, as the signal traverses the fiber distance twice before reaching the receiver. Both installed sections consist of fiber optic cables

housed in a protective liquid-filled tube within ducts. Based on our observations, we believe that the liquid facilitated the transfer of vibrations from external factors to the fibers, compared to empty ducts.

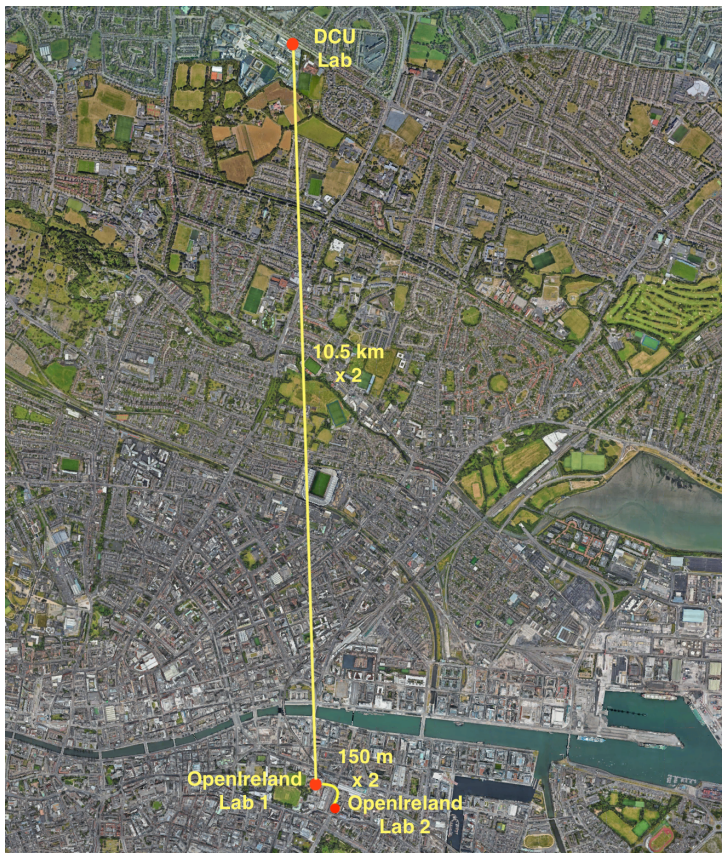


Figure 2: Field fiber routes in the OpenIreland testbed: a 300 m section between two TCD locations and a 21 km section linking TCD to DCU, both in loop back configuration.

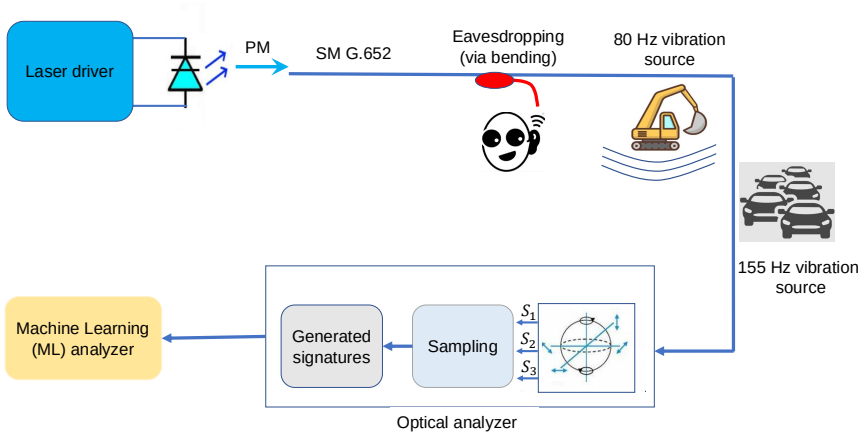


Figure 3: A schematic representation of the system utilized for the collection and analysis of SOP signatures. A linearly polarized laser source is coupled to a Polarization Maintaining (PM) fiber, which is connected to a Single Mode (SM) fiber conforming to the ITU-T G.652 transmission standard.

3.2 Data Collection Process

To collect and analyze signatures generated by various events that affect the short- (300 m) and long-link (21 km) installations, we designed a scheme as illustrated in Figure 3. We use a linearly polarized Continuous Wave Distributed Feedback (CW-DFB) laser capable of transmitting an optical signal that occupies one of the available wavelengths in the O, E, S, C, or L band. This laser source is controlled by a driver, ensuring stable optical power output for transmission. The emitted light is coupled to a Polarization Maintaining (PM) fiber connected to a G.652 SM fiber, where mechanical vibrations induce variations in the SOP. The laser source is set at 1310 nm wavelength and its spectral width was 0.5 nm, measured at a level where the optical power drops to 20 dB below the peak power at the center wavelength. This optical signal is sent through a PM fiber and injected into the installed network. The launch optical power is approximately -16 dBm and adjusted so that the optical analyzer receives power around -22 dBm. The results obtained with this configuration demonstrates a significantly improved Signal-to-Noise Ra-

tio (SNR) compared to the datasets reported in [29]. This improvement is primarily attributed to the reduced output power of the laser, which minimized back-reflected light caused by connector reflections and Rayleigh scattering. Consequently, the laser exhibited enhanced polarization stability.

The transmission line is affected by mechanical vibrations arising from harmless activities such as traffic or operational fans, as well as harmful events, including excavation, sabotage, or eavesdropping attempts. Each action can introduce measurable changes in the SOP of the transmitted light. Intentional disturbances are imposed at the beginning or at the end of the transmission line. The sampling time is set to 0.5 ms, resulting in a Nyquist sampling rate of 1 kHz.

The optical analyzer is responsible for quantifying changes in the SOP and generating characteristic patterns that encapsulate the behavior of each action over the installed network, referred to in this paper as *SOP signatures*. First, for each specific event, the Stokes parameters (S_1, S_2, S_3) are measured and mapped on the Poincaré sphere. S_0 was excluded from our analysis as it does not affect the polarization variation. Then, the sampling process with the sampling time of 0.5 ms spans over a 20-minute recording period for each event, resulting in a total of 2.4 million samples per event. Changes in the SOP are quantified by calculating the numerical distance between the SOP values at two consecutive sampling positions, forming the NPSV as follows:

$$\mathrm{NPSV}_t = \sqrt{(\Delta S_1(t))^2 + (\Delta S_2(t))^2 + (\Delta S_3(t))^2} \tag{D.1}$$

where $\Delta S_i = S_i(t) - S_i(t-1)$. Fast Fourier Transform (FFT) with hamming window [31] and 512 frequency bins is then applied to concatenated 0.5 ms NPSV data segments (time slots) of 500 elements each (i.e., equivalent to 250 ms) to convert the NPSV data from the time domain to the frequency domain. The resulting power spectrum dataset comprises 4,800 time slots (rows) and 512 frequency bins (columns). The collected signatures provide a detailed characterization of the polarization dynamics associated with each event type. The distinct signatures derived from each event type serve as the input for our ML algorithms designed to classify potential anomalies within the network.

3.3 Collected Signatures

This study involved the collection of seven distinct signatures for each of the fiber installations. Tables 1 and 2 summarize the signatures collected for the short- and the long-link installation, respectively.

Two types of normal fiber events are considered: relaxed (*rlx*) and soft bending (*sbd*). The *rlx* event represents the baseline scenario, where only routine background noise is recorded, i.e., no external intentional disturbances like eavesdropping, vibrations, or bending affect either short- (*rlx_short*) or long-link (*rlx_long*) fiber installations. The *sbd* event involves slight bending of the cable to simulate scenarios encountered during routine handling and maintenance of fiber installations, such as within patch panels or during cable routing. To simulate real-life handling of patch panel cables by data center personnel, we define the *sbd* event as handling the cable manually by bending it to a radius of approximately 2 cm. This bending was performed repeatedly with a time period of approximately 10 seconds to reflect realistic conditions. This approach ensures that the experiment accurately captures the effects of minor physical interactions typically encountered during routine maintenance and cable management in fiber installations. Capturing signatures from this type of event for both short- (*sbd_short*) and long-link (*sbd_long*) installations helps to categorize such routine handling apart from malicious fiber bending.

One of the main purposes of this paper is to detect harmful and non-harmful mechanical vibrations acting on the fiber installation. Potentially harmful vibrations, such as those caused by an excavator operating close to an optical fiber, represent a significant threat. Excavators typically produce mechanical vibrations with a frequency spectrum centered around 80 Hz. This frequency is directly related to the rotational speed of the engine, measured in Revolutions Per Minute (RPM). To convert RPM to frequency in hertz (Hz), the value is divided by 60, as there are 60 seconds in a minute. For instance, an excavator operating close to the optical fiber installation poses risk of accidentally cutting the cable and disrupting the traffic. The engine of an excavator normally runsat 4800 RPM, generating an 80 Hz vibration. By detecting this base tone, we can predict that an excavator is operating close to the fiber cable installation, and the network manager can take appropriate action to prevent a cable cut. For this study, 80 Hz vibrations were simulated using a loudspeaker placed 2 to 4 cm from the optical cable at a volume of 60 dBA to model the mechanical wave induced by an excavator. This mag-

Table 1: Collected Signatures for Short-Link Installation

Abbr.	Description	Justification
rlx_short	Relaxed fiber	Normal
sbd_short	Soft bending	Non-harmful (normal)
80vb_short	80 Hz vibration	Harmful; possible cut predecessor
dtc_80vb_short	80 Hz vibration before the detector	Harmful; possible cut predecessor
shld_80vb_short	Shielded fiber + 80 Hz vibration	Harmful; possible cut predecessor
eav_short	Bending	Eavesdropping
eav_80vb_155vb_short	Constant bending + 80 Hz + 155 Hz vibrations	Harmful; constant eavesdropping + non-harmful + harmful vibrations

nitude, equivalent to the volume of normal conversational speech and very likely to be exceeded by an excavator operating near fiber installation, serves as a conservative, worst-case scenario for testing the system's sensitivity. We collected potential harmful vibration signatures for the shorter (*80vb_short*) and the longer link (*80vb_long*).

When generating these two signatures, the distance between the laser and the external vibration sources was between two and three meters, which means that the signal affected by the imposed vibration traversed the full transmission distance and was simultaneously subject to external factors such as vibrations and noise coming from the real-world deployment.

For the short-link installation, the signature *dtc_80vb_short* corresponds to the case where the source of the 80 Hz vibration is placed just before the detector. The *shld_80vb_short* signature represents an 80 Hz vibration event simulated in the presence of a protective shield covering the fiber. This setup allows us to evaluate if the DL models are able to classify events even with the shield's ability to dampen or filter out harmful vibrations. In contrast, non-harmful vibrations often originate from benign sources, such as traffic-induced vibrations transmitted through the ground or other infrastructure.

Table 2: Collected Signatures for Long-Link Installation

Abbr.	Description	Justification
rlx_long	Relaxed fiber	Normal
sbd_long	Soft bending	Non-harmful (normal)
80vb_long	80 Hz vibration	Harmful; possible cut predecessor
80vb_155vb_long	80 Hz + 155 Hz vibrations	Harmful; harmful + non-harmful vibrations
shld_80vb_155vb_long	Shielded fiber + 80 Hz + 155 Hz vibrations	Harmful; non-harmful + harmful vibrations
eav_long	Bending	Eavesdropping
eav_80vb_155vb_long	Bending + 80 Hz + 155 Hz vibrations	Harmful; eavesdropping + non-harmful + harmful vibrations

Recording all potential non-harmful vibration types in real-life scenarios is a complex endeavor. Therefore, this study focuses on a 155 Hz vibration as a representative case for normal, benign activity, such as the vibrations generated by a fan operating at approximately 9,000 RPM. The *80vb_155vb_long* signature represents the combined effect of harmful and non-harmful vibrations on the long-link fiber, enabling us to evaluate the ability of DL models to classify overlapping events. In comparison, the *shld_80vb_155vb_long* signature simulates the impact of shielding on the fiber installation under the same conditions.

To simulate malicious eavesdropping, this study models an eavesdropper breaching the outer layers of the fiber cable and bending the internal fiber at a radius of 4 mm and a 25 degree angle as described in [32]. We consider an eavesdropping device with 3% coupling efficiency and 0.3 dB attenuation, which is difficult to detect using power threshold methods without generating excessive false alarms. Further details on the implications of different bending angles, coupling efficiencies, and attenuation of eavesdropping devices on the overall eavesdropping feasibility can be found in [32]. We consider two types of eavesdropping: continuous and bursty. The *eav_short* signature corresponds to the shorter link scenario where eavesdropping is carried out in short bursts of 10 s duration, with relaxing 20 s periods in between, simulating a

scenario where an attacker intermittently attempts to extract optical signals. The *eav_80vb_155vb_short* signature, on the other hand, refers to a constant eavesdropping event without relaxing periods. This event overlaps with vibrations at 80 Hz and 155 Hz. By introducing periodic interruptions, we assess the ML model's ability to detect and classify both transient and continuous eavesdropping attempts, ensuring its robustness in practical deployment scenarios. Combining the two types of overlapping vibrations simulates a real-world scenario where eavesdropping and mechanical disturbances simultaneously affect the fiber. The *eav_long* signature represents an eavesdropping event occurring at the long-link installation, where the eavesdropping is continuous without short bursts. Analogously, the *eav_80vb_155vb_long* signature models a similar eavesdropping event on the long-link fiber installation, where both 80 Hz and 155 Hz vibrations are present.

3.4 Data Pre-Processing

The pre-processing pipeline for the dataset was designed to ensure efficiency and compatibility with our DL framework. The same data pre-processing approach was applied on both the short- and the long-link dataset. As discussed in Sec. 3.2, each dataset contains 4,800 data points (samples) from each signature, i.e., a total of 33,600 data points per dataset. Each sample represents 0.5 ms of SOP changes characterized by 512 frequency bins (features). The dataset is randomly split into balanced training, validation, and testing sets with an 80-10-10 split ratio. The resulting splits consist of 26,880 samples for training, 3,360 samples for validation, and 3,360 samples for testing. A Z-score normalization technique is applied to scale the data to a uniform range. Additionally, the target labels are one-hot encoded, representing seven distinct classes.

4 Deep Learning Models

4.1 Model Architectures

Our proposed architecture for the classification task in short- and long-link installations integrates convolutional layers followed by fully-connected dense layers for effective sequence classification. The model architecture begins with a 1D convolutional layer, which is designed to learn and extract meaningful

patterns or relationships from the sequential order of the input data. This CNN is followed by a max-pooling layer to reduce spatial dimensionality, accompanied by batch normalization to enhance training stability and dropout layers to prevent overfitting. Subsequently, the output of the CNN is flattened and passed through a series of N fully connected dense layers. These dense layers are followed by batch normalization and dropout layers. Finally, the output of the dense layers is passed through a *softmax* activation function in the output layer, producing a 7-class probability distribution for classification.

4.2 Hyper-Parameter Tuning Algorithm and Setup

To improve the performance of the DL models, a hyper-parameter tuning process was implemented using the random search method of the *Keras Tuner* framework, aimed at exploring the best model architecture. The output of the hyper-parameter tuning process allows us to select a model that balances model complexity and classification accuracy in both short and long-link installations. For the initial convolutional layer, various hyper-parameters were adjusted for enhanced feature extraction. The number of filters in the CNN ranged from 2 to 8, enabling the model to capture diverse patterns at different levels of complexity. For the kernel size, which determines how many consecutive data points the model convolves at once, the considered range was $\{3, 5, 7\}$. Several activation functions, including $\{$*relu*, *swish*, *tanh*, *elu*$\}$ were tested so that the CNN could effectively capture nonlinear patterns in the data. The pooling size was selected from $\{2, 3\}$, and dropout rates were tuned in the range of $[0.1, 0.5]$ with a step size of 0.05.

The flattened output of the last convolutional layer was passed through fully-connected dense layers. We varied the number of these dense layers (N) from 4 to 7 to assess the trade-offs between model complexity and performance. The number of units in these layers was varied between 64 and 2048. Additionally, various activation functions, such as $\{$*relu*, *swish*, *tanh*, *elu*, *selu*, *gelu*$\}$, were examined for the dense layers. To prevent overfitting, dropout rates between these dense layers were tuned in the range $[0, 0.7]$ with a step size of 0.05.

The optimizer plays a critical role in determining how effectively it converges to a local minimum during training. We tested five popular optimizers – *Adam*, *SGD*, *RMSprop*, *Adadelta*, and *Adagrad* – with the learning rate as a key hyper-parameter. The learning rates for each optimizer were sampled from

a log scale ranging from 1×10^{-6} to 1×10^{-2}, to accommodate the dynamic range often necessary for effective training. Since in this work we tackle a multi-class classification problem, we adopt the cross-entropy loss function.

For each dataset comprised of data from the short- or long-link installation, we allow the *Keras Tuner* framework to train and evaluate 100 realizations of the CNN. Each realization was trained for up to 200 epochs, with training guided by an early stopping criterion. Specifically, training stopped if the validation loss did not improve for 30 consecutive epochs, and the best model (comprised of architecture, weights and biases) were stored. We recorded the loss and accuracy for the training and validation dataset across the epochs and selected the model that showed the highest validation accuracy at the end of training.

5 Results

In this section, we present the performance results of the DL models tested over the collected datasets. We first discuss the results of the hyper-parameter tuning step. Then, we analyze the performance of the best models found.

5.1 Tuned Hyper-Parameters

The hyper-parameters of the models were optimized using the *RandomSearch* method of the Keras Tuner framework, targeting validation accuracy as the optimization objective. A total of 100 trials were conducted. The best model configuration was selected based on the highest validation accuracy achieved by the model after training with a particular set of hyper-parameter values. The hyper-parameter values that achieved the best accuracy for the short- and long-link installations are summarized in Table 3.

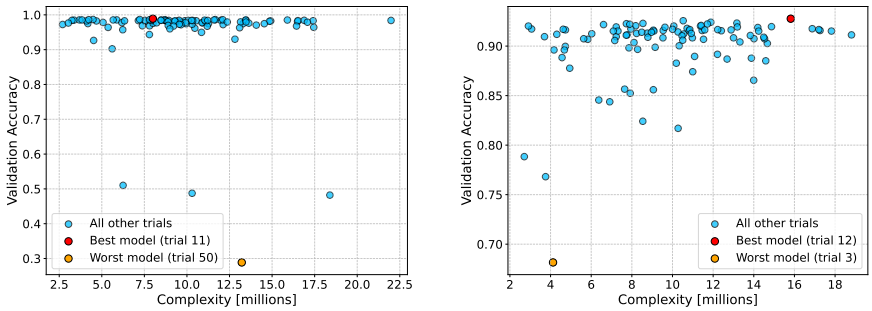
The selected models exhibit unique configurations adapted to the complexity of their respective datasets. Both models consist of a 1D convolutional layer, followed by 7 fully-connected dense layers. For the short-link installation, the model utilizes a slightly higher number of filters (8) in the convolutional layer, while the long-link installation requires a reduced number of filters (6). This difference highlights the need for greater feature extraction capacity in the short-link case, where a lower level of background noise in the signals requires the model to exploit finer details. Additionally, the activa-

Table 3: Hyper-Parameter Values for the Best Models

Hyper-parameter	**Short-link**	**Long-link**
Number of filters in conv. layer	8	6
Kernel size	7	7
Activation function for conv. layer	tanh	swish
Pooling size	2	2
Dropout rate in conv. layer	0.1	0.5
Number of dense layers	7	7
Units in dense layer 0	944	1280
Dropout rate for dense layer 0	0.1	0.35
Units in dense layer 1	240	1600
Dropout rate for dense layer 1	0.0	0.45
Units in dense layer 2	928	1504
Dropout rate for dense layer 2	0.15	0.0
Units in dense layer 3	1024	1296
Dropout rate for dense layer 3	0.5	0.6
Units in dense layer 4	1792	1760
Dropout rate for dense layer 4	0.65	0.0
Units in dense layer 5	400	1376
Dropout rate for dense Layer 5	0.2	0.3
Units in dense layer 6	400	544
Dropout rate for dense layer 6	0.6	0.5
Activation function for dense layers	swish	tanh
Optimizer	adam	adam
Learning rate	0.00478	0.00017
Total trainable parameters	7,986,519	15,818,963

tion functions differ between the two models, with *tanh* proving more effective for the short-link installation and *swish* yielding superior performance for the long-link case. The dropout rates were also tuned independently for each layer, with higher regularization applied to the convolutional layer in the long-link case (0.5). This adjustment suggests a greater need to prevent overfitting due to the increased complexity and noise of the long-link dataset. The disparity in model complexity between the short- and long-link installations is reflected in the total number of trainable parameters. The short-link model contains approximately 7.98 million parameters, while the long-link model requires over 15.8 million. This increase underscores the greater challenge of accurate event

classification for the long-link installation, attributed to higher environmental noise and data variability over the 21 km transmission distance.



(a) Short-link installation **(b)** Long-link installation

Figure 4: (a) Short-link installation (b) Long-link installation: validation accuracy versus model complexity (millions of trainable parameters) across 100 hyper-parameter tuning realizations.

Figure 4 shows the validation accuracy (the percentage of correctly classified samples out of the total number of samples) versus model complexity (measured in terms of the number of trainable parameters) for the 100 realizations of DL models during fine-tuning on the short- and long-link installations. The figure highlights differing requirements for the two considered scenarios. Most of the models tested for the short-link installation shown in Figure 4.a achieve accuracy above 0.9, with many models approaching perfect accuracy. Figure 4.b shows that, for the long-link case, many models fall below 0.9 accuracy. This indicates that the long-link dataset is more challenging. The higher complexity of the longer-link case is corroborated by the fact that the best model for the shorter link is approximately half as complex as the best model for the longer link (around 7.5 M vs. around 16 M, respectively).

The model with the worst performance for the short link in Figure 4.a achieves an accuracy of 0.29, while the worst model for the longer link in Figure 4.b achieves 0.68 accuracy. Interpreting the exact performance values for the worst models is not relevant, since this value is controlled by the random nature of the trainable parameters initialization, and the random hyper-parameter search. On the other hand, it is important to observe that the random search hyper-parameter tuning obtains only a few poorly-performing

Table 4: Accuracy of the Best Models over the Training, Validation, and Testing Sets.

Installation	Training	Validation	Testing
Short-link	99.81	98.90	98.57
Long-link	95.28	92.77	92.26

models, while most of the models perform closely to the best one. This means that the hyper-parameter tuning algorithm is able to quickly find architectures that yield good performance.

Figure 4 also shows that the most complex models are not always able to achieve the highest accuracy. Interestingly, as shown in Figure 4.a, the worst model for the short-link installation has higher complexity than the best model. In summary, the behavior revealed by Figure 4 shows that a thorough hyper-parameter tuning campaign is needed in these scenarios due to the large variations in performance. It also shows that *one-size-fits-all* architectures are not suitable if the goal is to achieve the highest performance possible for individual scenarios such as those considered in this paper.

5.2 Performance Analysis

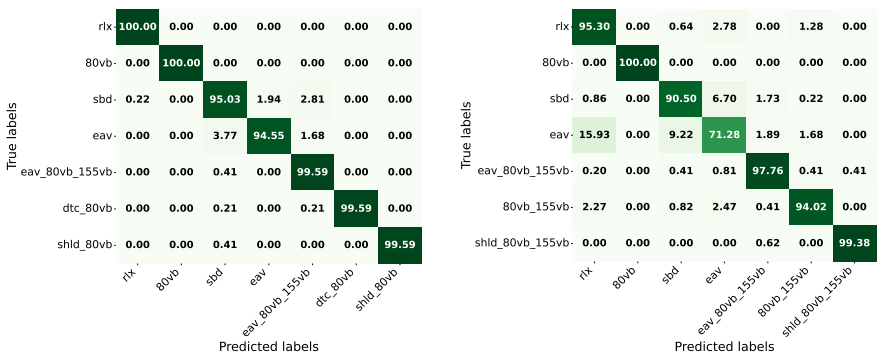
In this section, we analyze the performance of the best DL models for the short- and long-link installations found during the hyper-parameter tuning stage. The evaluation is based on the accuracy metric, followed by an analysis of confusion matrices.

Table 4 summarizes the accuracy results. These results demonstrate the effectiveness of the proposed DL models, achieving high classification accuracy across training, validation, and testing sets for both scenarios. Notably, the model applied to short-link data outperformed the long-link model, achieving an accuracy over the testing set of 98.57% compared to 92.26% for the long-link installation. The results also make it evident that the classification of events over fibers across populated areas becomes increasingly challenging with the distance that they span.

Figure 5.a shows the confusion matrix for the short-link scenario for the testing dataset. It allows us to analyze exactly which classes are affected by inaccuracies. The model has high classification performance in this dataset, with most events identified correctly, and a low misclassification rate. The model

achieved perfect classification for the normal operating condition signature (*rlx_short*). This means that the model presented no false positives (see the first line of the matrix). The confusion matrix also shows that this model presented a small false positive rate of 0.22% (see the first column of the matrix). Similarly, the *80vb_short*, *dtc_80vb_short*, and *shld_80vb_short* events were accurately classified, achieving accuracy of 100%, 99.59%, and 99.59%, respectively. The confusion matrix results for *80vb_short* and *dtc_80vb_short* show that the vibration insertion point does not influence the classification of the 80 Hz vibration signature, i.e., it does not matter significantly whether the vibration source is close to or far from the optical analyzer. For the overlapping event of eavesdropping with dual-frequency vibration denoted as *eav_80vb_155vb_short*, the model accurately detects this malicious event with 0.41% misclassification as a soft bending event. For *sbd_short*, 95.03% of the samples were identified correctly, with minor confusion observed primarily with *eav_short* (1.94%), *eav_80vb_155vb_short* (2.81%), and *rlx_short* (0.22%). Similarly, the *eav_short* signature achieved 94.55% accuracy, with some misclassification into event categories like *sbd_short* (3.77%),

eav_80vb_155vb_short (1.68%). These results highlight that our model performs well detecting constant eavesdropping.



(a) Short-link installation **(b)** Long-link installation

Figure 5: (a) Confusion matrix for the short-link installation (b) Confusion matrix for the long-link installation.

Figure 5.b shows the confusion matrix for the long-link scenario for the

testing dataset. We can observe that the model achieved high classification accuracy for several events, indicating a significant improvement of the results previously reported in **ls_ecoc** handling real-world noisy scenarios. The model classified the normal operating condition (*rlx_long*) with 95.30% accuracy. This means that the model showed a 4.70% false positive rate, with most of the false positives (2.78%) classified as *eav_long* (see the first row of the matrix). The *80vb_long* event was detected with perfect accuracy, confirming the model's ability to reliably identify critical threats. Similarly, the overlapping event identified as *shld_80vb_155vb_long* was classified with 99.38% accuracy, indicating that the model is able to achieve good accuracy even with the presence of a shield which reduces the effects of harmful vibrations on the SOP signature. However, for the soft bending and eavesdropping events (*sbd_long* and *eav_long*), the model faced challenges, with accuracies of 90.50% and 71.28%, respectively. Misclassifications of soft bending (*sbd_long*) occurred primarily as *eav_long* (6.70%), *eav_80vb_155vb_long* (1.73%), *rlx_long* (0.86%), and *80vb_155vb_long* (0.22%). These results suggest that the model has difficulties in distinguishing among these classes, possibly due to subtle overlaps in SOP variations between these events. For *eav_long*, the model showed rather high confusion with *rlx_long* (15.93%), *sbd_long* (9.22%), *eav_80vb_155vb_long* (1.89%), and *80vb_155vb_long* (1.68%), reflecting the complexity of accurately detecting inconstant malicious eavesdropping in noisy environments. The overlapping event *eav_80vb_155vb_long* was detected with 97.76% accuracy, with minimal misclassification into categories such as *shld_80vb_155vb_long* (0.41%), *80vb_155vb_long* (0.41%), *sbd_long* (0.41%), and *eav_long* (0.81%). The *80vb_155vb_long* event itself achieved a lower accuracy of 94.02%, with the majority of misclassifications occurring as *rlx_long* (2.27%) and *eav_long* (2.47%).

6 Conclusions

In this study, we proposed and evaluated DL models for event classification in short- and long-link fiber-optic installations. An extensive hyper-parameter tuning approach using the Keras Tuner framework identified the best model configurations tailored to the unique challenges posed by each installation type. The results demonstrated that our models achieve high classification accuracy, with notable performance differences between the short- and long-

link scenarios due to variations in noise levels and data complexity.

For the short-link installation, the best DL model achieved a testing accuracy of 98.57%, highlighting its ability to effectively classify events in low-noise environments. This performance underscores the suitability of the proposed architecture for scenarios in which the fiber spans shorter distances and experiences less environmental interference. In contrast, the long-link installation presented a more challenging classification problem.The best model for this scenario achieved a testing accuracy of 92.26%, indicating a strong but comparatively poorer performance. These results highlight the need for more robust techniques to address the complexities of long-link installations in noisy environments.

For both installations, the models performed well in detecting critical events such as eavesdropping and vibrations, which are vital for ensuring the security and reliability of optical communication systems. The hyper-parameter tuning revealed that a *one-size-fits-all* approach does not yield the best accuracy results for the use case analyzed in this paper. On the contrary, the best accuracy for each installation is achieved by models with different architectures. The analysis also showed that complexity does not directly translate to accuracy, and the best models were not the ones with the highest complexity. Finally, the performance gap between short-link and long-link installations suggests that future work should explore advanced techniques, such as hybrid architectures, to further enhance classification accuracy, particularly for long-link scenarios.

Acknowledgment

This work was supported by the Swedish Research Council (2023–05249), the European Commission's Digital Europe Programme (101127973) through the 5G-TACTIC project, the European Union's Horizon Europe research and innovation program (10113933) through the ECO-eNET project, and Taighde Éireann – Research Ireland under Grant No. 18/RI/5721: OpenIreland Research Infrastructure.

References

[1] Ashley Schulte, *Fiber Broadband*, https://fiberbroadband.org/2022/01/05/fiber-broadband-enters-largest-investment-cycle-ever/, January 5, 2022.

[2] Bergur, Thormundsson, *Fiber internet penetration rate in households in Sweden 2010-2020*, https://www.statista.com/statistics/598248/fiber-internet-penetration-rate-in-households-in-sweden/, December 19, 2022.

[3] Neos Networks, *What is a Backbone Network?* https://neosnetworks.com/resources/blog/what-is-backbone-network/?utm_source=chatgpt.com, March 24, 2024.

[4] A. Harris and P. Castle, "Bend loss measurements on high numerical aperture single-mode fibers as a function of wavelength and bend radius," *Journal of Lightwave Technology*, vol. 4, no. 1, pp. 34–40, 1986.

[5] M. Zafar Iqbal, H. Fathallah, and N. Belhadj, "Optical fiber tapping: Methods and precautions," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, 2011, pp. 164–168.

[6] C. Pendão and I. Silva, "Optical fiber sensors and sensing networks: Overview of the main principles and applications," *Sensors*, vol. 22, no. 19, p. 7554, 2022.

[7] G. Marra, D. Fairweather, V. Kamalov, P. Gaynor, M. Cantono, S. Mulholland, B. Baptie, J. Castellanos, G. Vagenas, J.-O. Gaudron, et al., "Optical interferometry-based array of seafloor environmental sensors using a transoceanic submarine cable," *Science*, vol. 376, no. 6595, pp. 874–879, 2022.

[8] S. Rerucha, M. Hola, M. Sarbort, J. Oulehla, B. Mikel, J. Lazar, and O. Cip, "Compact interferometric displacement gauge with sub-nanometer resolution and milimeter range," in *2016 IEEE SENSORS*, 2016, pp. 1–3.

[9] D. Grassani, M. Galli, and D. Bajoni, "Active stabilization of a michelson interferometer at an arbitrary phase with subnanometer resolution," *Opt. Lett.*, vol. 39, no. 8, pp. 2530–2533, Apr. 2014.

[10] Q. Bai, Q. Wang, D. Wang, Y. Wang, Y. Gao, H. Zhang, M. Zhang, and B. Jin, "Recent advances in brillouin optical time domain reflectometry," *Sensors*, vol. 19, no. 8, p. 1862, 2019.

[11] G. Wang, Z. Pang, F. Wang, Y. Chen, H. Dai, and B. Wang, "Urban fiber based laser interferometry for traffic monitoring and analysis," *Journal of Lightwave Technology*, vol. 41, no. 1, pp. 347–354, 2022.

[12] L. B. Liokumovich, N. A. Ushakov, O. I. Kotov, M. A. Bisyarin, and A. H. Hartog, "Fundamentals of optical fiber sensing schemes based on coherent optical time domain reflectometry: Signal model under static fiber conditions," *Journal of Lightwave Technology*, vol. 33, no. 17, pp. 3660–3671, 2015.

[13] B. Dong, A. Popescu, V. R. Tribaldos, S. Byna, J. Ajo-Franklin, K. Wu, et al., "Real-time and post-hoc compression for data from distributed acoustic sensing," *Computers & Geosciences*, vol. 166, p. 105 181, 2022.

[14] B. Wang, D. Ba, Q. Chu, L. Qiu, D. Zhou, and Y. Dong, "High-sensitivity distributed dynamic strain sensing by combining rayleigh and brillouin scattering," *Opto-Electronic Advances*, vol. 3, no. 12, pp. 200 013–1, 2020.

[15] Z. Wang, B. Lu, Q. Ye, and H. Cai, "Recent progress in distributed fiber acoustic sensing with Φ-OTDR," *Sensors*, vol. 20, no. 22, p. 6594, 2020.

[16] A. Masoudi, T. Lee, M. Beresna, and G. Brambilla, "10-cm spatial resolution distributed acoustic sensor based on an ultra low-loss enhanced backscattering fiber," *Optics Continuum*, vol. 1, no. 9, pp. 2002–2010, 2022.

[17] L. Shen, W. Yu, R. Xu, Y. Wang, and B. Du, "Co-propagation of distributed acoustic sensing in the L-band and 100-Gb/s WDM coherent communication systems," in *Asia Communications and Photonics Conference (ACP)*, 2022, pp. 629–632.

[18] J. K. Brenne, A. Sladen, P. Pecci, J. P. Morten, J. Pelaez, J. Jacobsen, A. Calsat, P. Plantady, J.-P. Ampuero, D. Rivet, and H. Février, "Non-intrusive das coexisting in telecom networks," in *2024 Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, pp. 1–3.

[19] J. Pesic, E. Le Rouzic, N. Brochier, and L. Dupont, "Proactive restoration of optical links based on the classification of events," in *15th International Conference on Optical Network Design and Modeling-ONDM 2011*, IEEE, 2011, pp. 1–6.

[20] W. H. McMaster, "Polarization and the stokes parameters," *American Journal of Physics*, vol. 22, no. 6, pp. 351–362, 1954.

[21] C. J. Carver and X. Zhou, "Polarization sensing of network health and seismic activity over a live terrestrial fiber-optic cable," *Communications Engineering*, vol. 3, no. 1, p. 91, 2024.

[22] D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J.-P. Elbers, "Cognitive assurance architecture for optical network fault management," *Journal of Lightwave Technology*, vol. 36, no. 7, pp. 1443–1450, 2018.

[23] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Risky event classification leveraging transfer learning for very limited datasets in optical networks," *Journal of Optical Communications and Networking*, vol. 16, no. 7, pp. C51–C68, 2024.

[24] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[25] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, "Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events," in *2024 24th International Conference on Transparent Optical Networks (ICTON)*, 2024, pp. 1–5.

[26] D. Rutigliano, G. Boracchi, P. Invernizzi, E. Sozio, C. Alippi, and S. Binetti, "Event-detection deep neural network for otdr trace analysis," in *International Conference on Engineering Applications of Neural Networks*, Springer, 2021, pp. 190–201.

[27] Z. Ge, H. Wu, C. Zhao, and M. Tang, "High-accuracy event classification of distributed optical fiber vibration sensing based on time–space analysis," *Sensors*, vol. 22, no. 5, p. 2053, 2022.

[28] K. Abdelli, H. Griesser, and S. Pachnicke, "Convolutional neural networks for reflective event detection and characterization in fiber optical links given noisy otdr signals," in *Photonic Networks; 22th ITG Symposium*, VDE, 2021, pp. 1–5.

[29] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[30] CONNECT Centre for Future Networks and Communications, *Open Ireland Testbed*, Available here.

[31] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

[32] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

PAPER E

AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Lena Wosinska, Marco Ruffini, Marija Furdek

Published in Optical Fiber Communication Conference (OFC) 2025, Technical Digest Series, [invited paper],
30 March–3 April 2025, San Francisco, CA, USA.
DOI: 10.1364/OFC.2025.M3F.6,

The layout has been revised.

Abstract

Optical networks are vulnerable to various disturbances that can jeopardize service availability or privacy. We discuss AI/ML-based analysis of the incurred state-of-polarization changes for cognitive management of complex disturbances.

1 Introduction

Optical networks form critical communication infrastructure that supports the majority of digital services in today's society. Their critical role requires very high reliability, security and resilience. These are jeopardized by diverse human activities such as construction work leading to accidental fiber cuts, deliberate sabotage or eavesdropping attempts. Early detection of and reaction to various external disturbances is important to ensuring the high performance and survivability of network services. The recent proliferation of Artificial Intelligence (AI)/Machine Learning (ML) techniques capable of observing intricate changes and trends in optical performance monitoring data has unleashed a tremendous potential for novel diagnostic capabilities. Combined with the advancements in network monitoring techniques, AI/ML drives new use cases related to sensing various environmental changes and disturbances to the optical network.

Traditionally, Optical Time Domain Reflectometry (OTDR) has been used in optical networks as a means to detect and localize anomalies in fiber links. This approach relies on expensive and specialized hardware to detect changes in the optical signal power. Coherent OTDR represents a significant advancement over traditional OTDR by employing coherent detection to capture both the amplitude and phase of backscattered light. A more advanced approach is Distributed Acoustic Sensing (DAS) which leverages Rayleigh backscattering in optical fibers to achieve high spatial resolution for detecting mechanical vibrations along the fiber length. While this technology is highly sensitive and precise, its implementation is costly and presents significant challenges, particularly in integration with legacy optical networks [1]. In recent years, substantial efforts have been made to utilize pre-installed optical fibers for sensing environmental changes [2]. The State of Polarization (SOP) represents the orientation of a light wave's electric field as it propagates through an

optical fiber and comprises values such as Stokes parameters (S_0, S_1, S_2, S_3), which define the light's polarization.

Its high sensitivity to external perturbations, often attributed to stress-induced birefringence [3], makes the SOP a key property for environmental sensing. As a result, it has emerged as a promising technique for monitoring optical fiber networks, garnering significant attention from academia and industry. This solution is particularly interesting due to the possibility of using lightly-modified transceivers to perform the SOP data collection, with promising results reported upon applying AI/ML for the analysis of such data [4], [5]

Despite the high demonstrated potential, widespread adoption of AI/ML-based SOP monitoring in optical networks faces several challenges. This paper overviews the basics of SOP data collection, analysis, and AI/ML processing and discusses several challenges of carrier-grade network-wide SOP monitoring deployment.

2 AI/ML-Based State-of-Polarization Monitoring in Optical Networks

Figure 1 presents a network architecture where traditional monitoring data, e.g., coming from Reconfigurable Optical Add-Drop Multiplexers (ROADMs) and transceivers, is collected in conjunction with SOP data. The data can be collected using standard protocols discussed in the literature and available in commercial products. The figure depicts two channels for which SOP data is collected with the aid of an SOP analyzer co-located with a ROADM at the destination node. The gathered SOP data are sent to the Network Management System (NMS) and stored in a telemetry database, usually implemented as a time-series database to allow queries of network parameter evolution over time. Since the telemetry database stores information about all the SOP channels in the network, it is possible to, in addition to analyzing them individually, correlate the SOP effects collected from multiple channels.

The stored SOP data is analyzed by the AI/ML monitoring module that can, depending on the exact applied technique, perform various functions such as, e.g., detection of anomalies, their classification, localization, or root cause analysis. As detailed in the right-hand side of the figure, pre-processing procedures such as normalization and standardization can be applied to adjust

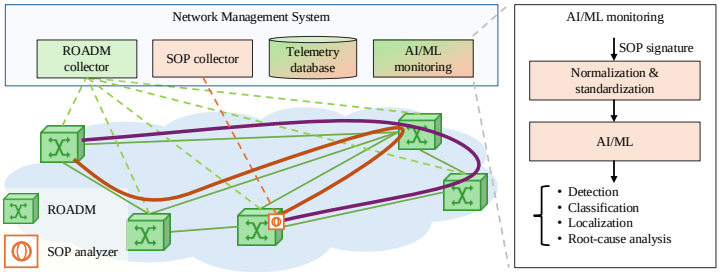


Figure 1: Network architecture with AI/ML monitoring based on traditional Reconfigurable Optical Add-Drop Multiplexer (ROADM) and State of Polarization (SOP) data collection.

the data to the most suitable format for AI/ML processing.

If AI/ML techniques for anomaly detection are applied, the output of the model is a binary variable that indicates whether the provided sample(s) represent an anomaly. This is commonly implemented by using an unsupervised learning algorithm, which does not require prior training or knowledge about the normal or abnormal operating conditions. Deeper insights into the nature of detected anomalies can be obtained by, e.g., root cause analysis [6]. Although the detection of abnormal events already enables several use cases for using SOP data, a more detailed output of the AI/ML model is often desired. This can be achieved by using supervised learning, e.g., a multi-class classifier, which is able to determine the precise category of an anomalous event based on the analyzed SOP data. However, this is only possible if the model is provided with a dataset containing sufficient samples from each category of events that need to be classified, which may be costly or infeasibile.

Finally, another aspect that can be useful is the localization of the event, which can be characterized in several ways. Firstly, *topological* localization refers to the ability of determining which fiber link in the topology is the source of the detected event. Since SOP data is collected only at the receiver, topological localization depends on the physical routes of lightpaths used to collect the SOP data. Secondly, *spatial* localization defines the exact point of an event along the fiber link (e.g., how many kilometers from its origin). SOP lacks spatial resolution in forward propagation alone, as events can occur anywhere along the fiber. However, spatial localization is possible by propagating a reverse signal and correlating the forward and reverse sig-

nals in time, requiring tightly synchronized receivers. The precision of spatial resolution depends on the synchronization accuracy and other factors. The final localization aspect refers to the *temporal* dimension, i.e., determining the start (and possibly the duration) of the detected event, which depends on the monitoring frequency.

3 Open Challenges

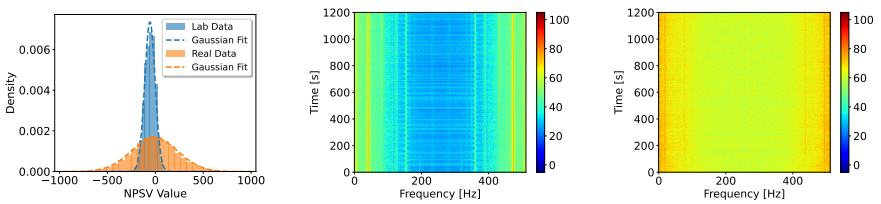
The use of SOP data as an enabler of using the optical network as a sensor for environmental disturbances. However, several remaining challenges should be addressed in order to make the technology suitable for widespread adoption. In the following, we mention a few of the critical challenges.

Data processing: The processing of the SOP data for their use in AI/ML models comprises several steps that involve the detection of analog optical signal, its conversion to the digital domain, the computation of Numerical Polarization State Variation (NPSV), its mapping to the Poincaré sphere, etc. As shown in Figure 1, part of this processing needs to be done at the receiving node, but other parts can be performed at the SOP collector. One of the challenges in this context is to define at which node (local or central) to carry out certain parts of the processing, which can reveal interesting trade-offs in terms of cost, telemetry efficiency, etc. Another challenge to be considered is related to the selection of the data format to be used as an input to the AI/ML model, where different formats may incur a different number of SOP processing steps.

Data collection: Data collection is another challenging aspect whenever the use of AI/ML models is considered, with a strong correlation with the training procedures. There are three aspects related to this challenge. Firstly, it is important to collect data that can be used for the appropriate training procedures. For instance, if multiple events need to be categorized, the collected data must follow the same categorization. Otherwise, if the AI/ML will only perform anomaly detection, a simpler categorization of the collected data can be performed. Secondly, the routing of the channels used for SOP monitoring need to be carefully computed. For instance, in Figure 1, two link-disjoint channels are used for collecting the SOP data, covering six out of the seven fiber links in the topology, and having the SOP analyzer at a single node. For larger topologies, defining paths that cover a representative set of links

at a minimal cost while allowing for complete event disambiguation becomes a challenging problem. Finally, while the collection of traditional monitoring data (e.g., Q-factor, optical signal-to-noise ratio) has been addressed by several protocols in the literature and industry, the protocol and the data format to be adopted for transferring SOP data from the SOP analyzer to the NMS remains an open question that deserves investigation.

AI/ML model training: In optical networks, it is expected that each component will incur a distinct effect on the SOP data. Therefore, it is unlikely that the same AI/ML model can be used for different paths/sections of the network. This means that a specific AI/ML model will be needed to analyse SOP of each channel in the network. An example of this behavior is illustrated in Figure 2, which shows the data related to the same event signature (i.e., 80 Hz vibration) over a lab and a real-world deployment. More details can be found in [7]. Figure 2.a shows the density of data values in both collected datasets, which make it evident that the data collected from the real-world testbed has substantially more noise than the data collected from the lab. Figure 2.b and .c show the waterfall diagrams from the two deployments, respectively, which depict the intensity of the SOP changes across frequency and time. We can observe that the data collected from the lab, Figure 2.b, presents a stronger distinction over frequency than the one from the real-world testbed, Figure 2.c. Moreover, optical fiber infrastructures are always evolving, and the SOP data needs to be continuously analyzed to detect substantial changes such that AI/ML models can be re-trained if necessary.



(a) Data distribution **(b)** Waterfall from the lab **(c)** Waterfall from the real-world testbed

Figure 2: Illustrative response of 80 Hz vibrations collected at a lab and at a real-world deployment.

Performance trade-offs and targets: A critical challenge when designing

a monitoring system is to define how often the monitoring is performed. This aspects defines the overhead of monitoring data transmission in the network, the amount of processing capacity needed to obtain the solution, and the speed of detecting an event. A default assumption that the more frequent monitoring is, the better needs to be revisited when considering AI/ML adoption. AI/ML models can be characterized by metrics such as false positive and false negative rates. These rates are evaluated depending on the number of inferences, i.e., the number of analyzed samples, implying that they scale with the frequency of the data analysis. The more frequently the data is analyzed, the more frequently the system will need to handle incorrect outputs from the AI/ML model. Thus, it is important to consider not only the performance target of AI/ML models, but also the frequency at which these models will be used, in order to clearly evaluate their impact on the overall operational overhead of the network. For instance, a model with a false positive and a false negative rates equal to 0.01 will experience on average two mistakes for every 100 samples evaluated, one false positive and one false negative. If the network is monitored every second, two mistakes will need to be handled every 100 seconds, and this scales linearly with the frequency of monitoring.

4 Conclusion

AI/ML-based SOP monitoring can enable real-time detection of various disturbances affecting optical network infrastructure. Challenges related to the data collection and processing, adaptability to evolving threats, and the computational overhead of real-time AI model deployment need to be overcome through, among other, advanced data gathering, model optimization and integration with NMS to enhance service availability and integrity.

Acknowledgment

This work was supported by the Swedish Research Council (2023–05249), the European Commission's Digital Europe Programme (101127973) through the 5G-TACTIC project, and CONNECT center, OpenIreland, funded by the Science Foundation Ireland grants 13/RC/2077_p2 and 18/RI/5721.

References

[1] Z. He and Q. Liu, "Optical fiber distributed acoustic sensors: A review," *Journal of Lightwave Technology*, vol. 39, no. 12, pp. 3671–3686, 2021.

[2] G. Marra, D. Fairweather, V. Kamalov, P. Gaynor, M. Cantono, S. Mulholland, B. Baptie, J. Castellanos, G. Vagenas, J.-O. Gaudron, et al., "Optical interferometry-based array of seafloor environmental sensors using a transoceanic submarine cable," *Science*, vol. 376, no. 6595, pp. 874–879, 2022.

[3] S. Pellegrini, G. Rizzelli, M. Barla, and R. Gaudino, "Algorithm optimization for rockfalls alarm system based on fiber polarization sensing," *IEEE Photonics Journal*, vol. 15, no. 3, pp. 1–9, 2023.

[4] "Enhancing fiber security using a simple state of polarization analyzer and machine learning," *Optics Laser Technology*, vol. 167, p. 109 668, 2023, ISSN: 0030-3992.

[5] A. Tomasov, P. Dejdar, P. Munster, and T. Horvath, "Utilizing a state of polarization change detector and machine learning for enhanced security in fiber-optic networks," in *CLEO 2024*, Optica Publishing Group, 2024, JTu2A.217.

[6] C. Natalino, M. Schiano, A. D. Giglio, and M. Furdek, "Root cause analysis for autonomous optical network security management," *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 2702–2713, 2022.

[7] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

PAPER F

ML-based State of Polarization Analysis to Detect Emerging Threats to Optical Fiber Security

Leyla Sadighi, Stefan Karlsson, Carlos Natalino, Marija Furdek

IEEE Transactions on Network and Service Management (TNSM),
TNSM-2024-08489,
DOI: 10.1109/TNSM.2025.3607022

The layout has been revised.

Abstract

As the foundation of global communication networks, optical fibers are vulnerable to various disruptive events, including mechanical damage, such as cuts, and malicious physical layer breaches, such as eavesdropping via fiber bending. Traditional monitoring methods often fail to identify subtle or novel anomalies, stimulating the proliferation of Machine Learning (ML) techniques for detection of threats before they cause significant harm. In this paper, we evaluate the performance of Semi-Supervised Learning (SSL) and Unsupervised Learning (USL) approaches for detecting various abnormal events, such as fiber bending and vibrations, by analyzing polarization signatures with minimal reliance on labeled data. We experimentally collect thirteen polarization signatures on three different types of fiber cable and process them using One-Class Support Vector Machine (OCSVM) as an SSL, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) as a USL algorithm for anomaly detection. We introduce tailored evaluation metrics designed to guide hyper-parameter tuning and capture generalization over different anomaly types, detection consistency, and robustness to false positives, enabling practical deployment of OCSVM and DBSCAN in optical fiber security. Our findings demonstrate DBSCAN as a strong contender to detect previously unseen threats in scenarios where labeled data are not available, despite some variability in performance between different scenarios, with F1 score values between 0.615 and 0.995. In contrast, OCSVM, trained on normal operating conditions, maintains high F1 scores of 0.98 to 0.998, demonstrating accurate detection of complex anomalies in optical networks.

1 Introduction

Fiber optic networks form the foundation of modern telecommunications, enabling high-speed, long-distance, and reliable data transmission with minimal

signal loss. In addition to backbone and long-haul networks, fibers are also deployed in the access segment, forming, e.g., Passive Optical Networks (PONs) that deliver broadband connectivity to end users. They connect regions and nations, supporting global connectivity and critical systems such as the Internet, government, financial, and healthcare networks.

Optical networks are vulnerable to physical threats such as fiber cuts, reportedly causing up to 60% of failures [1], [2]. A leading cause of fiber cuts are construction works, where the operation of heavy excavator machinery induces vibrations as a cut predecessor [3]. Fibers are also exposed to covert security risks from evanescent coupling and fiber bending [4], [5], which can compromise data confidentiality via eavesdropping, without affecting signal quality. Another security vulnerability arises from unauthorized signal access through unused or improperly secured ports, such as monitoring outputs or unused branches of optical splitters [6]. These access points, often overlooked during installation or maintenance, can be exploited by malicious actors to tap into the optical signal without introducing noticeable attenuation or disrupting service.

Quick detection and response to a range of anomalies is critical for safeguarding fiber optic networks. Techniques like Optical Time Domain Reflectometry (OTDR), based on Rayleigh backscattering, aid fault detection and localization of large-scale physical faults, such as sharp bends or fiber breaks [7], [8], [9], but face scalability and cost challenges [10], and exhibit limited sensitivity to detect subtle disturbances, such as minor mechanical vibrations and small-radius bends. Alternative methods such as Distributed Fiber Optic Sensing (DFOS) for intrusion detection [11] are effective yet complex and expensive, requiring high-speed lasers, diplexers, and advanced backscattering analysis.

Recent advancements rely on the existing optical fiber infrastructure for environmental sensing [12], successfully detecting natural and human-induced activities [13], [14]. A key enabler for such sensing is the State of Polarization (SOP), which is highly sensitive to mechanical disturbances. SOP is characterized by the Stokes parameters, organized into a four-element Stokes vector, $\mathbf{S} = [S_0, S_1, S_2, S_3]$, and represented on the Poincaré sphere for visual interpretation [2]. Mechanical stress, temperature changes, bending, and vibrations impact fiber birefringence, altering polarized light transmission and modifying the SOP [2]. SOP-based sensing is cost-effective, as mod-

ern Polarization-Multiplexed Quadrature Amplitude Modulation (PM-QAM) coherent receivers inherently track SOP for signal demodulation, eliminating the need for additional hardware [15]. SOP-based analysis leverages the inherent sensitivity of polarization to both benign and malicious disturbances, including covert eavesdropping attempts, and is capable of detecting subtle, short-duration, or overlapping events without relying on backscatter or reflection signatures. Thus, SOP-based monitoring simplifies the detection process, avoiding the complexity and cost of DFOS or OTDR.

Effective monitoring of polarization changes is essential to identify disturbances and maintain network integrity. However, SOP-based monitoring systems that rely on fixed rules and thresholds struggle to handle evolving threats [16], especially those that induce complex changes in the SOP. To overcome these limitations and enable scalable, automated, and adaptive analysis of polarization signatures, recent research has shifted toward data-driven approaches that leverage ML as a critical monitoring technique in SOP-based fiber sensing, enabling models to learn and distinguish complex disturbance patterns from polarization data.

In the context of anomaly detection, polarization signatures can be defined as a sequence of the magnitude of polarization variations in a specific time and frequency, derived after processing the SOP variations data [17]. The derived polarization signatures are plotted in a waterfall diagram. These plots, unique for each event, help differentiate legitimate actions from eavesdropping, offering a cost-effective solution to security challenges. However, the method relies on manual interpretation by technicians, which requires expertise, time, and does not scale. Instead of focusing on image and vision-based analysis of SOP, our prior research [18], [19], [20], [21] employed sampling techniques and Fast Fourier Transform (FFT) processing on SOP and calculated the numerical value of changes in SOP to derive an SOP signature for each event type.

While Supervised Learning (SL) techniques show remarkable potential for the detection and classification of polarization signatures, their reliance on labeled datasets may be limiting for real-world applications. Labeling polarization events, especially in large-scale optical networks, is often labor-intensive, costly, time-consuming, and requires domain expertise, making it impractical for network-wide monitoring and timely anomaly detection. To overcome these challenges, this work investigates the potential of SSL and USL tech-

niques to efficiently analyze and detect optical fiber events with minimal or no reliance on labeled training sets. By leveraging the intrinsic patterns within the data, SSL methods can learn from a small set of normal labeled data, while USL techniques such as clustering can identify patterns and outliers without any prior label knowledge. This shift towards more flexible and scalable ML techniques is essential to improve the robustness and effectiveness of fiber optic network monitoring, particularly in complex and dynamic environments where manual labeling is not feasible, or new threats, previously untrained for, may emerge. In this paper, we consider OCSVM as an SSL technique and DBSCAN as a USL method to tackle the challenge of anomaly detection in fiber optic sensing, where labeled data is scarce or costly to obtain. We apply these methods to the thirteen experimental scenarios described in [18], covering an extensive spectrum of events possibly affecting fiber optic installations across three cable types.

The rest of the paper is structured as follows. Section 2 overviews recent advancements in ML-based anomaly detection for optical networks. Section 3 describes the experimental setup and the data collection process for generating polarization signatures under various conditions. Section 4 details the employed ML methodology, focusing on the data pre-processing and hyperparameter tuning of OCSVM and DBSCAN. Section 5 analyzes experimental results, and Section 6 concludes the paper.

2 Related Work

SL techniques, which rely on labeled data to train models for event detection and classification, have been proliferating in recent studies. In [22], a transfer learning strategy was proposed that leverages unrelated image datasets to improve SOP-based classification performance when labeled data are scarce. A follow-up work introduced a Vision Transformer framework for anomaly detection and localization using SOP time-series inputs, highlighting the effectiveness of deep attention mechanisms in optical monitoring [23]. Most recently, SOP spectrogram representations were combined with Vision Transformers to further enhance anomaly classification and localization accuracy in complex network scenarios [24]. Despite their advantages, SL techniques often struggle with complex scenarios involving overlapping effects, such as operational stress, harmful vibrations, and covert attacks, leading to reduced

accuracy in distinguishing harmful from benign events in real-world deployments.

To address these challenges, our prior work combined robust SOP signature extraction with advanced SL models to improve accuracy in real-world conditions. In [18], polarization signatures from three cable types under thirteen scenarios were analyzed, with eXtreme Gradient Boosting (XGBoost) achieving 92.3% accuracy in classifying eavesdropping and other events. In [19], a polarization-based fiber sensor with supervised ML achieved 97.94% accuracy using Histogram Gradient Boosting (HGB) to distinguish harmful from non-harmful events on an indoor cable. In [20], we analyzed noisy SOP data from OpenIreland's live network [25] in Dublin, achieving 86.5% accuracy using the HGB classifier, demonstrating robustness in real-world conditions. In [21], we applied Deep Learning (DL) models to a broader, noisier dataset, improving accuracy to over 91%. The work in [26] presented a supervised Convolutional Neural Network (CNN)-based scheme for detecting fiber-bending eavesdropping at different bending radii (10.8 mm, 12.1 mm, 15 mm) in coherent optical systems based on the polarization data.

Despite the clear advantages of SSL and USL techniques for optical network monitoring and anomaly detection, to the best of our knowledge, no prior work has explored their application to polarization signature analysis for anomaly detection in optical fibers. The work in [27] proposed a combined USL and SL framework that utilized power spectral density and Signal-to-Noise Ratio (SNR) data to identify anomalies in optical networks. Recent studies, such as [15] and [28], have proposed advanced Digital Signal Processing (DSP)-based techniques for anomaly detection in polarization state, validated on metropolitan fiber links using mechanical shakers to introduce various vibration types. In contrast to these advanced DSP techniques, our study presents a new perspective on anomaly detection in optical fibers. While the work in [15] leveraged an USL approach using bisecting k-means clustering on Optical Performance Monitoring (OPM) data to detect and localize eavesdropping in Wavelength Division Multiplexing (WDM) networks, our work analyzes SOP dynamics to detect a broader range of subtle anomalies, including those not evident through power-level changes. The work in [29] reported on an SOP-based system for detecting physical disturbances in optical fibers using SSL anomaly detection models, including thresholding and autoencoders. In this paper, we employ SSL and USL techniques to detect a

wide range of anomalies in polarization signatures, including eavesdropping, harmful and non-harmful vibrations, and overlapping events directly from raw SOP data, eliminating the need for labeled data and enhancing adaptability to unknown threat types.

3 Testbed and Collected Signatures

The experimental configuration used to generate, record, and analyze polarization signature data is depicted in Figure 1. It consists of a transmitter, transmission line, receiver with a polarization measurement system, and SOP processing units.

The transmitter includes a Continuous Wave Distributed Feedback (CW-DFB) laser emitting light at 1310 nm with polarization-maintaining properties. The laser is controlled by a driver that maintains a constant power level and temperature, ensuring stable operation. The optical signal is transmitted through a transmission line which is a serial structure of three different optical cables: Fiber Optical Tactical Cable System (FOCS) cable, an indoor Single Mode (SM) cable, and a bare G.657.A SM sensitive fiber. The CW-DFB laser is transmitting optical power occupying one of the available wavelengths in the O, E, S, C, or L-band of the network. In order to mimic real-world threat scenarios, we apply targeted manipulations to each cable type. These actions produce polarization signatures that characterize each specific event, crafted to simulate conditions typically linked to eavesdropping attempts, as well as both harmful and benign mechanical vibrations. The receiver obtains the transmitted signal and carries out the SOP analysis using the three steps described as follows.

Step 1: The optical analyzer is used to quantify changes in the SOP. The acquired optical SOP signal is mapped onto the Poincaré sphere to visualize and track polarization variations. To convert the analog signal into a digital format, a sampling mechanism is employed, which captures the SOP on the Poincaré sphere at intervals of 1 ms. This process is carried out over a 20-minute recording period, resulting in a total of 1.2 million samples per event. Subsequently, the numerical vectorial distance between the SOP at two consecutive sampling points is calculated using:

$$\mathrm{NPSV}_t = \sqrt{(\Delta S_1(t))^2 + (\Delta S_2(t))^2 + (\Delta S_3(t))^2} \tag{F.1}$$

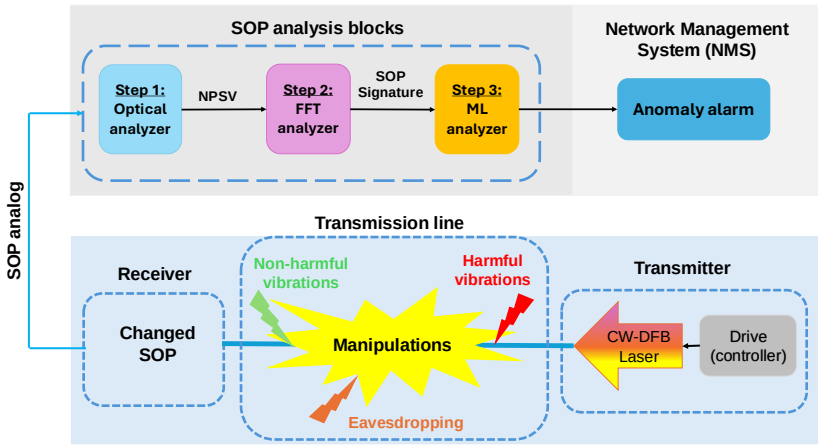


Figure 1: A schematic of the system used to generate and analyze the SOP signatures.

with $\Delta S_i = S_i(t) - S_i(t-1)$.

Step 2: The FFT analyzer further processes the data. FFT with hamming window [30] and 512 frequency bins is applied to concatenated 1-second Numerical Polarization State Variation (NPSV) data segments of 1,000 elements each. The result is a power spectrum data set of 1,200 samples (or rows, corresponding to time slots) and 512 samples (or columns, corresponding to frequency bins). The output of this step characterizes the polarization dynamics for each event type, referred to in this paper as *SOP signatures*.

Step 3: The collected signatures are analyzed by the ML analyzer. It applies techniques for detecting anomalies that may indicate harmful vibrations or eavesdropping attempts. The detected anomalies are communicated to the Network Management System (NMS), which oversees network health, raises alarms, and responds to potential threats. The considered approach is aimed at detecting the presence of anomalies at an affected fiber link, without providing the precise localization along the link.

Using the described experimental setup, we collect data from thirteen experimental scenarios, comprising normal and/or abnormal events. Normal considered events include the relaxed fiber without vibrations or eavesdropping (denoted as *rlx*) as well as non-harmful vibrations which might stem

from equipment like fans, nearby traffic, or benign interactions with the network infrastructure. Since capturing all types of normal vibrations in real-life scenarios is complex, this study simulates normal vibrations using 130 Hz (denoted as *130vb*) and 155 Hz (denoted as *155vb*) frequencies, representing fans with speeds of approximately 7,000 and 9,000 Revolutions Per Minute (RPM). We deliberately selected frequencies close to each other in order to test the sensor system and ML models' ability to distinguish between similar vibration signatures. To generate these vibrations in our testbed, we attached a piezo electrical engine to the optical cable, powered by a sinusoidal 155/130 Hz vibration. This vibration is transferred by the inner layer of the cable and interacts with the optical fiber, inducing changes in the SOP.

The abnormal events encompass potentially harmful vibrations and eavesdropping attempts. The harmful vibrations considered in this study include fiber vibrations at 80 Hz (denoted as *80vb*). An example of such malicious event is an excavator digging close to the installed fiber optical ground cable, which may result in an accidental or deliberate fiber cut, leading to a disruption in network traffic. A common excavator generally runs at 4,800 RPM. Since one minute equals 60 seconds, the fundamental tone of the excavator corresponds to a frequency of 80 Hz (4,800/60 = 80). While the full vibration spectrum includes a range of frequencies around 80 Hz, for detecting the presence of an excavator it is enough to detect the presence of the 80 Hz vibration. In commercial equipment, the entire bandwidth needs to be detected and analyzed in order to detect an excavator running with some other engine at different RPM. For the purpose of this paper, the main issue is to present a fiber optical sensor together with the ML model that can separate examples of malicious events from examples of normal events that can be experienced in a real-life fiber optical network. This vibration was simulated by a loudspeaker producing a sinusoidal tone of 80 Hz at 60 dBA volume (A-weighted sound level). The bare fiber was exposed to the tone at a distance of 5 cm from the membrane of the loudspeaker. The eavesdropping attack signatures (denoted as *eav*) are generated by bending the fiber over a 10 mm diameter rod. The bend radius is chosen to be 4 mm with a bend angle of 25 degrees.

To successfully perform eavesdropping, the eavesdropper must manipulate the fiber optic cable and expose the bare optical fiber that is protected within it. This process generates various signatures that can be detected. In our experiment, we considered three cable configurations: (i) bare (*br*) fiber, (ii) a

Table 1: Collected Signatures for Bare Fiber and FOCS Cable

Abbr.	Scenario	Justification
155vb_br	155 Hz vibration bare	Normal
155vb_fcs	155 Hz vibration FOCS	Normal
80vb_br	80 Hz vibration bare	Harmful; possible cut predecessor
80vb_fcs	80 Hz vibration FOCS	Harmful; possible cut predecessor
eav_br	Bending bare	Eavesdropping
eav_fcs	Bending FOCS	Eavesdropping

standard indoor (*idr*) patch cable, and (iii) a military-grade FOCS (*fcs*) cable. These 3 cable types were deliberately selected to span a range of mechanical isolation levels. The bare optical fiber offers direct mechanical exposure and thus produces strong vibration signatures with high SNRs, serving as a baseline for unshielded conditions. The indoor patch cable, a commonly deployed fiber type, provides minimal isolation, allowing external vibrations to be readily coupled into the fiber and resulting in detectable signatures. In contrast, the FOCS tactical cable is designed for deployment in rugged environments, featuring a robust design that includes multiple protective layers. These layers provide substantial mechanical shielding, enabling the cable to withstand pulling forces up to 2,000 Newtons and resist damage from knotting. This construction significantly attenuates mechanical disturbances before they reach the fiber. As such, we expect the vibration signatures in the FOCS cable to have lower SNR.

For the bare and the FOCS cables, we consider one normal and two abnormal events, summarized in Table 1. The normal event involves non-harmful vibrations at 155 Hz, denoted as *155vb_br* for the bare and *155vb_fcs* for the FOCS cable. The abnormal events include potentially harmful vibrations at 80 Hz, referred to as *80vb_br* and *80vb_fcs* for the two cables, as well as eavesdropping, denoted as *eav_br* and *eav_fcs*.

For the indoor cable, we consider two normal and five abnormal events, outlined in Table 2. The normal events include a relaxed fiber without vibrations or eavesdropping (*rlx_idr*) and with vibrations at 155 Hz (*155vb_idr*). The

Table 2: Collected Signatures for Indoor Cable

Abbr.	Description	Justification
rlx_idr	Relaxed fiber	Normal
155vb_idr	155 Hz vibration	Normal
80vb_idr	80 Hz vibration	Harmful; possible cut predec.
eav_130vb_idr	Bending + 130 Hz vibration	Eavesdropping + non-harmful
eav_80vb_idr	Bending + 80 Hz vibration	Eavesdropping + harmful
eav_80vb_130vb_idr	Bending + 80 Hz + 130 Hz vibrations	Eavesdropping + non-harmful + harmful
rlx_80vb_130vb_idr	Relaxed + 80 Hz + 130 Hz vibrations	Non-harmful + harmful

abnormal events include potentially harmful vibrations at 80 Hz (*80vb_idr*) and a set of overlapping events: a combination of eavesdropping and non-harmful vibration at 130 Hz (*eav_130vb_idr*), a combination of eavesdropping and harmful vibration at 80 Hz (*eav_80vb_idr*), a combination of eavesdropping with dual-frequency vibrations at 80 Hz and 130 Hz (*eav_80vb_130vb_idr*), and a relaxed fiber subjected to both harmful and non-harmful frequency vibrations (*rlx_80vb_130vb_idr*). Overlapping events reflect real-world scenarios with fiber exposed to multiple simultaneous stressors, e.g., routine vibrations generated by fan ventilation or traffic, and intentional malicious activities, such as eavesdropping. They are specifically considered for the indoor cable because this environment is more likely to experience a variety of simultaneous disturbances, making it an ideal candidate to test the robustness of our anomaly detection models.

4 ML-based Anomaly Detection Models

To analyze the obtained SOP data, we use OCSVM as an SSL, and DBSCAN as an USL model. To evaluate the performance of our anomaly detection models, we use standard values derived from the confusion matrix: True Positives (TP), referring to correctly detected anomalies; False Positives (FP), referring to normal instances incorrectly flagged as anomalies; True Negatives (TN), referring to correctly identified normal instances; and False Neg-

atives (FN), referring to anomalies that the model failed to detect. These values enable calculation of key evaluation metrics such as True Positive Rate (TPR) (also known as recall or sensitivity), measuring the proportion of actual anomalies correctly detected as TPR = $\frac{TP}{TP+FN}$; False Positive Rate (FPR) (FPR = $\frac{FP}{FP+TN}$), representing the fraction of normal data mistakenly flagged as anomalous; True Negative Rate (TNR) (also known as specificity), which measures the proportion of normal samples correctly identified (TNR = $\frac{TN}{TN+FP}$); and False Negative Rate (FNR), quantifying the proportion of anomalies that were missed by the model (FNR = $\frac{FN}{TP+FN}$). We also consider the Accuracy (acc), which measures the overall proportion of correctly detected instances (both normal and abnormal), and precision, defined as the ratio of correctly identified anomalies among all detected anomalies. The F1-score, computed as the harmonic mean of precision and recall, offers a balanced perspective on model performance by jointly considering false positives and false negatives. This makes it particularly relevant for anomaly detection tasks, where both Type I (FP) and Type II (FN) errors are critical. In security-sensitive applications, FPR and FNR are especially critical: high FPR can overwhelm operators with false alerts, wasting resources, while high FNR may allow threats to go undetected, potentially leading to breaches or service disruptions. For clustering-based models like DBSCAN, we additionally use the Silhouette Score (SS) to assess the cohesion and separation of detected clusters, and the Adjusted Rand Score (ARS) to measure similarity between clustering results and ground truth labels while accounting for the possibility that some agreement may occur purely by random chance, rather than meaningful structure. Both OCSVM and DBSCAN models require careful data pre-processing of the collected SOP signatures and hyper-parameter tuning, detailed in the following.

4.1 One-Class Support Vector Machine(OCSVM)

OCSVM is a powerful ML algorithm widely used for anomaly detection, particularly in scenarios where the available data (predominantly) represents normal behavior, with the goal of detecting outliers or anomalies that deviate from this norm. OCSVM operates in a semi-supervised manner by learning the boundaries that encapsulates the majority of data points, assuming that they represent normal working conditions. The algorithm maps the input data into a high-dimensional feature space using a kernel function and then

constructs a decision boundary that maximizes the separation between the origin and the data points in this feature space. Data points that fall outside of this boundary are detected as anomalies. The choice of kernel function and hyper-parameters, such as the kernel coefficient gamma (γ) and the regularization parameter (ν), is critical in determining the sensitivity of the model to outliers and its overall performance. During model inference, new data points are detected as normal if they fall within the boundary, and abnormal otherwise.

Data pre-processing for OCSVM

For each of the considered scenarios, the pre-processing phase begins by separating the dataset into subsets corresponding to normal operating conditions (referred to as normal scenario for brevity) and the malicious/harmful events (referred to as abnormal scenario). Table 3 summarizes the separation of training and test data for these scenarios.

Table 3: Summary of the Considered Normal and Abnormal Events for Each Cable Type, and Dataset Separation for OCSVM Model

Cable Type	Normal event	Abnormal events	Training Set	Test Set
Bare	*155vb_br*	*eav_br* *80vb_br* total (2 abnormal events)	900 of *155vb_br*	300 of *155vb_br* + 1,200 of *80vb_br* (1,500 Samples) + 1,200 of *eav_br* (1,500 Samples) + 2,400 of all abnormal events (2,700 Samples)
FOCS	*155vb_fcs*	*eav_fcs* *80vb_fcs* total (2 abnormal events)	900 of *155vb_fcs*	300 of *155vb_fcs* + 1,200 of *80vb_fcs* (1,500 Samples) + 1,200 of *eav_fcs* (1,500 Samples) + 2,400 of all abnormal events (2,700 Samples)
Indoor (Case 1)	*rlx_idr*	*80vb_idr* *eav_130vb_idr* *eav_80vb_idr* *eav_80vb_130vb_idr* *rlx_80vb_130vb_idr* total (5 abnormal events)	900 of *rlx_idr*	300 of *rlx_idr* + 1,200 of *80vb_idr* (1,500 Samples) + 1,200 of *eav_130vb_idr* (1,500 Samples) + 1,200 of *eav_80vb_idr* (1,500 Samples) + 1,200 of *eav_80vb_130vb_idr* (1,500 Samples) + 1,200 of *rlx_80vb_130vb_idr* (1,500 Samples) + 6,000 of all abnormal events (6,300 Samples)
Indoor (Case 2)	*155vb_idr*	*80vb_idr* *eav_130vb_idr* *eav_80vb_idr* *eav_80vb_130vb_idr* *rlx_80vb_130vb_idr* total (5 abnormal events)	900 of *155vb_idr*	300 of *155vb_idr* + 1,200 of *80vb_idr* (1,500 Samples) + 1,200 of *eav_130vb_idr* (1,500 Samples) + 1,200 of *eav_80vb_idr* (1,500 Samples) + 1,200 of *eav_80vb_130vb_idr* (1,500 Samples) + 1,200 of *rlx_80vb_130vb_idr* (1,500 Samples) + 6,000 of all abnormal events (6,300 Samples)

For each cable type, we created a training set of 900 samples from the normal condition and a test set of 1,500 samples, comprising 300 normal samples and 1,200 of the corresponding abnormal scenario samples. In the case of indoor cable, we considered two normal cases: relaxed fiber (case 1) and 155 Hz vibration (case 2). To assess the detector's behavior when exposed to all

abnormal events, we introduced the *total* test sets that contain all respective abnormal events into a single evaluation case. This results in a larger test set comprising 300 normal samples and the cumulative 1,200-sample sets for each abnormal event (e.g., totaling 2,400 samples for bare and FOCS cables, and 6,000 samples for indoor cable cases). After data separation, we evaluated the impact of feature normalization on the performance of our OCSVM model. The results indicated that normalization did not confer any advantages; hence, normalization was not applied in the final analysis.

Tuning of OCSVM hyper-parameters

After data pre-processing, we tuned the main OCSVM hyper-parameters to improve OCSVM anomaly detection performance. This hyper-parameter tuning focused on exploring a grid of three parameters: the kernel type, ν, and γ. The γ parameter controls how much influence each training point has on the model. Smaller γ values produce smoother and more general decision boundaries, while larger values produce boundaries more close to the training samples. In our case, smaller values worked better. We also varied ν, which controls the fraction of data points in the training set allowed to be outliers. Our grid search considered a range of $[0.001, 0.5]$ for ν, $[10^{-7}, 0.5]$ for γ, and {`poly`, `rbf`, `sigmoid`} for the kernel function.

For each hyper-parameter combination, we employed 5-fold cross-validation (CV), assessing the model performance across different training-validation splits. The mean and standard deviation (std) values obtained from the CV folds allows us to evaluate whether a certain hyper-parameter setting results in a model overly tailored to a specific data split. A key element of our model selection process is the use of a novel performance metric that we define in (F.2).

$$\begin{aligned} OCSVM\_perf = avg\Big([TPR - FPR + \mathrm{CV}_{mean} - \mathrm{CV}_{std} \\ - |\mathrm{acc}_{train} - \mathrm{acc}_{test}| + F1]\Big) \end{aligned} \tag{F.2}$$

This metric was designed to balance the trade-offs between different evaluation criteria by averaging the TPR, FPR, CV_{mean}, CV_{std}, F1-score, acc, and the difference in accuracy between the training and the testing sets

Table 4: Tuned Hyper-Parameters of OCSVM for Bare and FOCS Cables

Event type	Cable type	kernel	ν	γ
155vb vs *80vb*	Bare	rbf	0.001	1×10^{-5}
155vb vs *eav*	Bare	rbf	0.02	3×10^{-5}
155vb vs *total*	Bare	rbf	0.035	1×10^{-5}
155vb vs *80vb*	FOCS	rbf	0.2	8×10^{-5}
155vb vs *eav*	FOCS	rbf	0.001	1×10^{-4}
155vb vs *total*	FOCS	rbf	0.05	5×10^{-5}

($|acc_{train} - acc_{test}|$). After calculating the value of OCSVM_$perf$ for each hyper-parameter combination, the models were ranked, and the model with the highest *OCSVM_perf* score was selected.

Results of hyper-parameter tuning for OCSVM

Tables 4 and 5 summarize the selected hyper-parameters for each cable type and event scenario. The best hyper-parameter setting for each normal vs. abnormal scenario, as well as the *total* scenario are shown. The Radial Basis Function (RBF) kernel function consistently achieves the best result across all scenarios. For the bare fiber (Table 4, rows 1–3), the best selected hyper-parameters generally reflect a need for high sensitivity, as indicated by consistently low values of ν and γ, suitable for detecting subtle deviations in polarization caused by events like eavesdropping. In contrast, the FOCS cable (Table 4, rows 4–6), which offers greater insulation from environmental factors, exhibits a wider range of ν values across different scenarios, pointing to a more scenario-dependent sensitivity requirement. Nonetheless, the γ values for both cables remain within a relatively narrow band, reflecting the model's consistent preference for smoother decision boundaries. In *total* scenarios, moderate ν values were selected, indicating a balance between sensitivity and generalization across different types of disturbances.

For the indoor cable (Table 5), where scenarios are more complex and involve overlapping events, ν values generally range from 0.001 to 0.02, with γ values tightly controlled between 1×10^{-5} and 3×10^{-5}. This configuration ensures that the model can effectively distinguish between normal and various abnormal scenarios, including those involving multiple simultaneous vibrations and eavesdropping.

Table 5: Tuned Hyper-Parameters of OCSVM for Indoor Cable

Event type	kernel	ν	γ
155vb vs *80vb*	rbf	0.01	3×10^{-5}
rlx vs *80vb*	rbf	0.02	3×10^{-5}
155vb vs *eav_130vb*	rbf	0.01	1×10^{-5}
rlx vs *eav_130vb*	rbf	0.01	1×10^{-5}
155vb vs *eav_80vb*	rbf	0.02	3×10^{-5}
rlx vs *eav_80vb*	rbf	0.01	1×10^{-5}
155vb vs *eav_130vb_80vb*	rbf	0.01	1×10^{-5}
rlx vs *eav_130vb_80vb*	rbf	0.01	1×10^{-5}
155vb vs *rlx_130vb_80vb*	rbf	0.01	1×10^{-5}
rlx vs *rlx_130vb_80vb*	rbf	0.01	1×10^{-5}
155vb vs *total*	rbf	0.001	1×10^{-5}
rlx vs *total*	rbf	0.009	3×10^{-4}

4.2 Density-Based Spatial Clustering of Applications with Noise(DBSCAN)

The DBSCAN model is an unsupervised ML technique, chosen in this work for its robustness in detecting arbitrarily-shaped clusters and its effectiveness in isolating noise or outliers in unlabeled data. Since DBSCAN does not rely on a training process, it can be directly applied to a sequence of consecutive data points, making it particularly flexible and adaptable to various conditions. The DBSCAN algorithm operates by defining two key parameters: the radius of the neighborhood ϵ and the minimum number of points required to form a dense region $MinPts$. A sample is considered a core point if it has at least $MinPts$ neighbors within the radius ϵ. Clusters are formed by core points that are within ϵ of each other. Points that do not belong to any cluster and have fewer than $MinPts$ neighbors within ϵ are classified as noise or outliers (i.e., anomalies in our case). The choice of ϵ and $MinPts$ values significantly impacts the performance of DBSCAN. A small ϵ may result in classifying many data points as noise, while a large ϵ can lead to the formation of fewer clusters, with potential anomalies remaining undetected. Similarly, a small $MinPts$ value can cause the algorithm to detect too many small clusters, potentially labeling noise points as part of clusters, while a large $MinPts$ value may result in fewer clusters, with some anomalies being overlooked.

Applying DBSCAN to detect anomalies in polarization signatures data, where abnormal patterns manifest as sparse or isolated points in the feature space requires careful tuning of these parameters.

Data pre-processing for DBSCAN

For data pre-processing and hyper-parameter tuning, we consider 1,200 data points for each cable type under normal operating conditions and 1,200 data points corresponding to each abnormal condition. The normal and the abnormal events are defined in the same way as stated in Table 3. The dataset is then normalized through z-score standardization, ensuring that the feature values are consistent and properly scaled for effective application of the DBSCAN algorithm.

Tuning of DBSCAN hyper-parameters

To tune ϵ and $MinPts$, we adopted a controlled sampling strategy to simulate realistic yet varied conditions. We selected fixed-size windows of pre-processed data containing both normal and abnormal samples. The number of normal samples in each window (#N) ranged from 150 to 300, while the number of abnormal samples (#AN) was varied between 10 and 15 to reflect realistic start of threat scenarios with low anomaly prevalence. For each combination of ϵ and $MinPts$, we randomly sampled such windows of consecutive samples and evaluated the detection performance across 20 independent iterations. To assess the effectiveness of a hyper-parameter combination (ϵ, $MinPts$) under varying data conditions, we introduced the metric *DBSCAN_perf*, defined in equation (F.3).

$$DBSCAN\_perf = avg(TPR - FPR + F1 + ARS + SS) \tag{F.3}$$

This metric aggregates five important evaluation criteria into a single scalar value, averaged over the sampling iterations. The rationale behind this formulation is to combine multiple complementary aspects of clustering performance. It combines TPR, F1-score, and FPR to reflect detection performance, while ARS and SS assess the quality of the resulting clusters. Together, *DBSCAN_perf* provides a balanced view of detection accuracy and cluster quality, guiding the selection of hyper-parameters that yield robust clustering performance across multiple randomized trials. After evaluating all

Table 6: Tuned Hyper-Parameters of DBSCAN for Bare and FOCS Cables and the Selected Window Size of Normal (#N) and Abnormal (#AN) Samples.

Event type	Cable type	ϵ	$MinPts$	# N	# AN
155vb vs *80vb*	Bare	19	53	220	11
155vb vs *eav*	Bare	20	78	220	11
155vb vs *total*	Bare	20	78	220	11
155vb vs *80vb*	FOCS	19	11	200	10
155vb vs *eav*	FOCS	15	80	200	10
155vb vs *total*	FOCS	19	11	200	10

possible hyper-parameter combinations, the results were sorted by their *DBSCAN_perf* value, and the hyper-parameter configuration that achieved the highest value was selected. The selected model was then subjected to 50 additional iterations over different sampled windows of normal and abnormal data for further validation, during which the final assessment metrics presented in the next section were calculated.

Results of hyper-parameter tuning for DBSCAN

Tables 6 and 7 summarize the tuned hyper-parameters used for the DBSCAN models across the three cable types for all considered events. For bare fiber (Table 6), due to the unshielded nature of the cable, moderate values of ϵ and $MinPts$ were selected to capture fine-grained distinctions between normal and abnormal behavior. In contrast, the FOCS cable (Table 6) required either more compact or more dispersed clusters depending on the event type, indicating a need to accommodate its more stable but insulated signal profile. The indoor cable (Table 7) presented the most diverse tuning requirements, with a wide spread in both ϵ and $MinPts$ values. This reflects the higher complexity of overlapping event scenarios. Interestingly, the eavesdropping scenario drove the selection of the total hyper-parameters for the bare cable, while 80vb drove the selection for the FOCS cable.

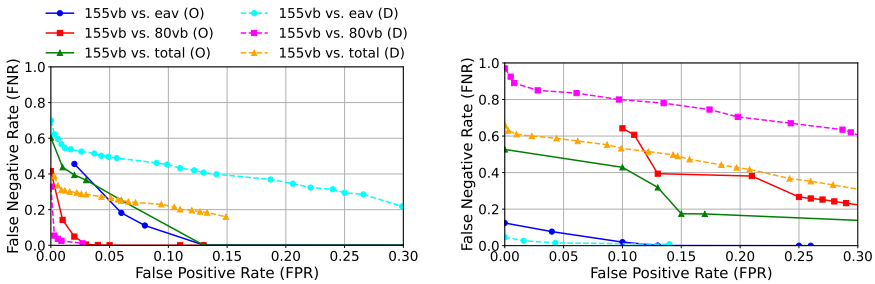
4.3 Hyper-parameter tuning overview

In this section, we summarize the results of the hyper-parameter tuning process by showing the trade-off between FNR and FPR for all hyper-parameter settings tested. The trade-off is illustrated by selecting the hyper-parameter

Table 7: Tuned Hyper-Parameters of DBSCAN for Indoor Cable and the Selected Window Size of Normal (#N) and Abnormal (#AN) Samples.

Event type	ϵ	$MinPts$	# N	# AN
155vb vs *80vb*	24	50	150	10
rlx vs *80vb*	23	60	150	10
155vb vs *eav_130vb*	16	90	150	10
rlx vs *eav_130vb*	15	12	150	10
155vb vs *eav_80vb*	15	85	150	10
rlx vs *eav_80vb*	12	18	150	10
155vb vs *eav_130vb_80vb*	16	80	150	10
rlx vs *eav_130vb_80vb*	16	80	150	10
155vb vs *rlx_130vb_80vb*	21	20	150	10
rlx vs *rlx_130vb_80vb*	24	100	150	10
155vb vs *total*	16	90	150	10
rlx vs *total*	12	18	150	10

settings that show the best (i.e., the lowest) FPR for a given FNR (and vice-versa). Figure 2 shows the Pareto frontier that indicates the trade-off between FPR and FNR for OCSVM and DBSCAN across three evaluation scenarios for bare and FOCS cables. For bare fiber (Figure 13.2(a)), OCSVM consistently demonstrates more favorable detection performance, achieving lower FNR values at comparable or lower FPR across all hyper-parameter settings. OCSVM is particularly effective in the *155vb* vs. *80vb* case, nearly eliminating false negatives while maintaining minimal false positives. In contrast, DBSCAN exhibits higher FNR even at low FPR, especially in the *155vb* vs. *eav* and *155vb* vs. *total* scenarios. A similar trend is observed for the FOCS cable (Figure 13.2(b)). OCSVM continues to outperform DBSCAN across all scenarios, maintaining lower FNR values for the same or lower FPR. In particular, for *155vb* vs. *eav*, it exhibits a near perfect separation with negligible false positives and minimal missed detections. Even in the *155vb* vs. *total* scenario, OCSVM exhibits strong generalization capability. In contrast, DBSCAN exhibits consistently higher FNR values across the same range of FPR levels. Notably, in the *155vb* vs. *80vb* and *155vb* vs. *total* scenarios, DBSCAN fails to reach the same low FNR levels as OCSVM, underscoring its limited adaptability when handling heterogeneous data from multiple abnormal events.



(a) Bare fiber **(b)** FOCS cable

Figure 2: (a) Pareto frontier for bare fiber. (b) Pareto frontier for FOCS cable, showing the trade-off between FNR and FPR during hyper-parameter tuning of OCSVM (O) and DBSCAN (D).

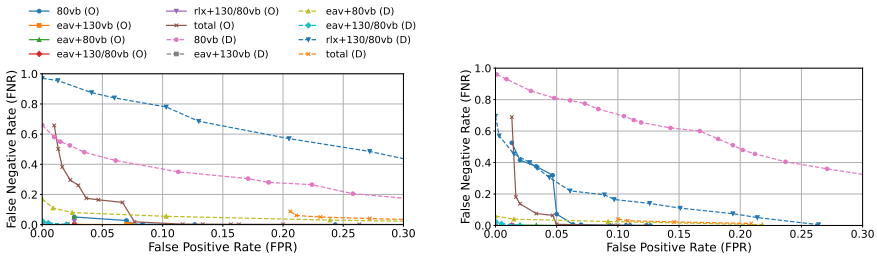
Figure 3 presents the Pareto frontier curves obtained during hyper-parameter tuning of OCSVM and DBSCAN, where each curve corresponds to one evaluated indoor attack scenario. The plots highlight the parameter settings which achieve the most favorable balance between the FPR and FNR. There is a pronounced advantage of OCSVM over DBSCAN, particularly in handling diverse and overlapping abnormal conditions. In both baseline configurations, i.e., *155vb_idr* (Figure 13.3(a)) and *rlx_idr* (Figure 13.3(b)), OCSVM maintains low FNR values across all tested scenarios. The decision boundary learned from normal conditions generalizes well, allowing the model to reject a wide range of abnormal patterns without significantly increasing false alarms. In contrast, DBSCAN exhibits inconsistent behavior in the same scenarios.

5 Results

We first analyze the performance of the two ML models for each fiber type, followed by an assessment of the overall performance across all fiber types. The performance is measured in terms of the TPR, FNR, FPR, and TNR.

5.1 Results for bare fiber

Table 8 depicts the performance of the OCSVM and DBSCAN models in distinguishing between three types of events for the bare fiber model. We



(a) *155vb_idr* **(b)** *rlx_idr*

Figure 3: Pareto frontier curves illustrating the FNR-FPR trade-off during hyperparameter tuning of OCSVM (O) and DBSCAN (D) for the different indoor cable scenarios, considering (a) *155vb_idr* and (b) *rlx_idr* as the normal baseline.

examine the effectiveness in differentiating harmful vibration at 80 Hz (*80vb*), eavesdropping (*eav*), as well as the combined abnormal scenario (*total*) that includes both types of attacks from normal vibration at 155 Hz (*155vb*). The results show that OCSVM achieves below 1% FNR for individual events, but higher FPR of up to 10%. For the combined dataset, OCSVM achieves a high performance, with 97.5% TPR and 95% TNR. The DBSCAN model achieves good results for the *80vb* scenario, but has poor FNR performance for eavesdropping, which also affects the results for the combined dataset. Although DBSCAN achieved an FPR below 1%, demonstrating good ability in detecting the anomalies, its 30.86% FNR means that nearly one-third of actual anomalies were not detected, which reflects a high risk of undetected threats, and indicates the need for further improvements. In practical terms, OCSVM achieves superior performance in detecting harmful vibrations for bare fiber, making it a more reliable choice for environments where the primary concern is preventing physical damage to the fiber.

5.2 Results for FOCS cable

Table 9 presents the performance of the OCSVM and DBSCAN models in distinguishing the three abnormal scenarios, i.e., *80vb*, *eav*, and the combined abnormal scenario (*total*), from the normal scenario (*155vb*) in the FOCS cable.

Based on the multi-layered protection property, we can expect that distin-

Table 8: Results for Bare Fiber Scenarios (Normal vs Abnormal) using the OCSVM (O) and DBSCAN (D) Models (M) in terms of True Positive Rate (TPR), False Negative Rate (FNR), False Positive Rate (FPR), and True Negative Rate (TNR), in %.

Scenario	M	**TPR**	**FNR**	**FPR**	**TNR**
155vb vs *80vb*	O	99.92	0.08	3.33	96.67
155vb vs *80vb*	D	94.55	5.45	0.63	99.37
155vb vs *eav*	O	99.75	0.25	10.67	89.33
155vb vs *eav*	D	46.55	53.45	1.16	98.84
155vb vs *total*	O	97.50	2.50	5.00	95.00
155vb vs *total*	D	69.14	30.86	0.89	99.11

Table 9: Results for FOCS Cable Scenarios (Normal vs Abnormal) using the OCSVM (O) and DBSCAN (D) Models (M) in terms of True Positive Rate (TPR), False Negative Rate (FNR), False Positive Rate (FPR), and True Negative Rate (TNR), in %.

Scenario	M	**TPR**	**FNR**	**FPR**	**TNR**
155vb vs *80vb*	O	79.67	20.33	32.33	67.67
155vb vs *80vb*	D	80.4	19.6	67.36	32.64
155vb vs *eav*	O	99.83	0.17	2	98
155vb vs *eav*	D	92	8	0.47	99.53
155vb vs *total*	O	98.33	1.67	2.00	98.00
textit155vb vs *total*	D	73.00	27.00	33.40	66.60

guishing between the signatures of *80vb_fcs* and *155vb_fcs* will be a more challenging task than in other cable types. This is confirmed in Table 9 where the performance for *155vb* vs. *80vb* drops substantially for both models compared to bare fiber (Table 8). Surprisingly, the eavesdropping detection performance is better than the one observed for bare fiber. This indicates that the protective layers of FOCS cables, while detrimental to the detection of vibrations, may better reveal the effects of the eavesdropping procedures, facilitating its detection by ML models. When considering all anomalous scenarios (*total*), OCSVM shows a strong performance with 98% in both TPR and TNR. However, DBSCAN does not achieve a good performance, with high false negative (27%) and positive (33%) rates.

5.3 Results for indoor cable

The performance of OCSVM and DBSCAN for the indoor cable scenarios is shown in Tables 10 and 11. Different from the two previously considered fiber types where the normal scenario featured 155 Hz vibrations, for the indoor cable we consider two normal scenarios: the 155 Hz vibrations *155vb_idr*, and relaxed fiber *rlx_idr*. In both cases, all other signatures, i.e., *80vb_idr*, *eav_130vb_idr*, *eav_80vb_idr*, *eav_80vb_130vb_idr*, *rlx_80vb_130vb_idr*, and *total* (a comprehensive set of anomalies encompassing all the abnormal scenarios), are considered abnormal. We focus on the ability of the two models to distinguish *(i)* harmful vibrations at 80 Hz, *(ii)* overlapping frequency vibrations at 80 Hz, 130 Hz and *(iii)* eavesdropping combined with single or dual frequency vibrations at 80 Hz and 130 Hz, and *(iv) total* from the two normal cases.

Detection of harmful 80 Hz vibrations

When distinguishing the abnormal 80 Hz vibrations from the normal 155 Hz vibration, as depicted in the first two rows of Table 10, the OCSVM model demonstrates excellent performance, correctly detecting 99.58% of the abnormal instances with a minimal FNR of 0.42%. However, the model obtains an FPR of 7.33%, which suggests that environments with frequent, benign disturbances might generate unnecessary alerts. In comparison, the DBSCAN model presents a high FNR (35.2%) and FPR (9.08%). Similar observations can be made when the relaxed fiber is considered as baseline, as shown in the first two rows of the Table 11. Overall, results show that OCSVM achieves good performance with more than 90% true positive and negative rates in all scenarios. DBSCAN, in contrast, shows a large performance gap. This performance gap in comparison to OCSVM indicates that, while DBSCAN is capable of detecting harmful vibrations, it may not be as effective in distinguishing between closely related vibration patterns in indoor cable.

Detection of eavesdropping overlapping with single and dual-frequency vibrations

The central rows of Tables 10 and 11 (rows 3-8) assess the ability of the models to distinguish scenarios with overlapping events involving eavesdropping and single- or dual-frequency vibrations from two baseline normal con-

Table 10: Results for indoor cable scenarios (Normal vs Abnormal) considering *155vb* as the Normal Class using the OCSVM (O) and DBSCAN (D) Models (M) in terms of True Positive Rate (TPR), False Negative Rate (FNR), False Positive Rate (FPR), and True Negative Rate (TNR), in %.

Scenario	M	**TPR**	**FNR**	**FPR**	**TNR**
155vb / *80vb*	O	99.58	0.42	7.33	92.67
155vb / *80vb*	D	64.80	35.20	9.08	90.92
155vb / *eav_130vb*	O	99.83	0.17	2.67	97.33
155vb / *eav_130vb*	D	97.20	2.80	0.04	99.96
155vb / *eav_80vb*	O	98.17	1.83	7.00	93.00
155vb / *eav_80vb*	D	91.80	8.20	0.49	99.51
155vb / *eav_80vb_130vb*	O	99.92	0.08	2.67	97.33
155vb / *eav_80vb_130vb*	D	96.40	3.60	0.05	99.95
155vb / *rlx_80vb_130vb*	O	98.17	1.83	7.00	93.00
155vb / *rlx_80vb_130vb*	D	91.00	9.00	49.96	50.04
155vb vs *total*	O	97.90	2.10	2.00	98.00
155vb vs *total*	D	6.92	3.08	41.18	58.82

ditions: *155vb_idr* and *rlx_idr*. Across these overlapping-event scenarios, both OCSVM and DBSCAN demonstrate high detection capabilities, with OCSVM generally offering stronger abnormal detection due to its training-based nature. OCSVM's performance remains consistently high regardless of the combination of harmful or non-harmful vibrations, with TPR higher than 97% and TNR higher than 92% in all scenarios. DBSCAN, except for one scenario, shows good performance with false rates below 10%. However, in the *80vb* scenario, its performance is substantially lower than that of OCSVM. In general, these findings are consistent with the previous one, showing that OCSVM has strong capabilities, while DBSCAN may need further enhancements to be suitable for real-world applications.

Table 11: Results for indoor cable scenarios (Normal vs Abnormal) considering *rlx* as the Normal Class using the OCSVM (O) and DBSCAN (D) Models (M) in terms of True Positive Rate (TPR), False Negative Rate (FNR), False Positive Rate (FPR), and True Negative Rate (TNR), in %.

Scenario	M	**TPR**	**FNR**	**FPR**	**TNR**
rlx / 80vb	O	98.83	1.17	5.00	95.00
rlx / 80vb	D	81.60	18.40	40.89	59.11
rlx / eav_130vb	O	100.00	0.00	1.33	98.67
rlx / eav_130vb	D	99.20	0.80	0.00	100.00
rlx / eav_80vb	O	99.58	0.42	1.33	98.67
rlx / eav_80vb	D	97.20	2.80	0.03	99.97
rlx / eav_80vb_130vb	O	100.00	0.00	1.33	98.67
rlx / eav_80vb_130vb	D	98.40	1.60	0.00	100.00
rlx / rlx_80vb_130vb	O	100.00	0.00	1.33	98.67
rlx / rlx_80vb_130vb	D	76.20	23.80	6.91	93.09
rlx vs *total*	O	99.77	0.23	5.00	95.00
rlx vs *total*	D	98.00	2.00	20.02	79.98

Detection of dual-frequency 80 Hz and 130 Hz vibrations

When detecting dual-frequency vibrations in the indoor cable (rows 9-10 of Tables 10 and 11), the OCSVM model consistently shows robust performance. It maintains high sensitivity to anomalies while keeping false positives low across both evaluated baselines. In contrast, the performance of DBSCAN becomes less reliable under the same conditions. Specifically, it exhibits greater difficulty in separating complex overlapping anomalies from normal behavior, especially when the relaxed fiber scenario is used as the baseline. This is evident in the increased misclassification of normal samples and reduced anomaly sensitivity compared to OCSVM. These trends confirm that, while DBSCAN may handle certain overlapping patterns well, OCSVM is more effective in consistently detecting subtle dual-frequency vibration events in more challenging and realistic baseline conditions.

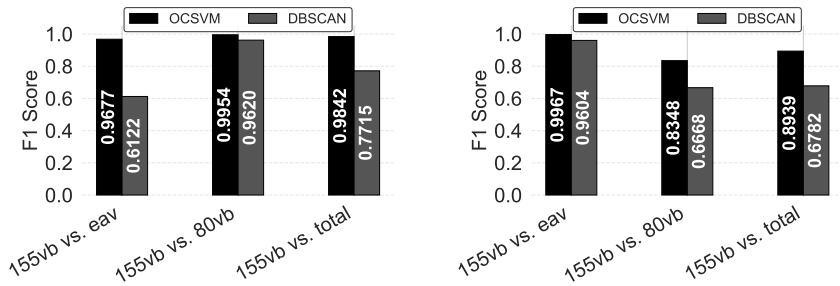
Detection of combined abnormal events

The last two rows in Tables 10 and 11 present the performance of the OCSVM and DBSCAN models in detecting the *total* scenarios. Both models achieve high TPR, but OCSVM consistently offers superior performance, particularly in minimizing false positives and maximizing true negatives. This advantage is more evident when the relaxed fiber scenario is used as the baseline, highlighting the robustness of OCSVM to background variation. In contrast, while DBSCAN maintains strong sensitivity, its performance is limited by higher FPRs. These trends underscore the suitability of OCSVM for comprehensive anomaly detection in indoor fiber environments.

5.4 Overall assessment

We compare the anomaly detection performance of the considered models and scenarios in terms of F1-score values. In practical terms, a high F1-score means that the model not only accurately identifies anomalies (high precision) but also captures the majority of them (high recall).

Fig. 4 shows the F1-scores obtained by OCSVM and DBSCAN when detecting the three abnormal scenarios in bare and FOCS fibers. For bare fiber (Fig. 13.4(a)), OCSVM achieves F1-scores of 0.9677, 0.9954, and 0.9842 for detecting eavesdropping, harmful vibrations, and total anomalies, respectively. DBSCAN is comparatively effective in detecting harmful vibrations with an F1-score of 0.9620, but attains values of only 0.6122 and 0.7715 for eavesdropping and total anomalies, indicating weaker performance in these scenarios. OCSVM demonstrate a strong overall anomaly detection performance for the FOCS cable (Fig. 13.4(b)), achieving an excellent F1-score of 0.9967 for eavesdropping detection. It also performs well in the combined anomaly scenario with an F1-score of 0.8939, and maintains a solid score of 0.8348 when detecting harmful vibrations. DBSCAN, while exhibiting greater variability, achieves an F1-score of 0.9604 for eavesdropping detection, but its performance declines in the *total* and harmful vibration scenarios, with scores of 0.6782 and 0.6668, respectively. These results further highlight the stronger generalization capability of OCSVM. The diminished performance of DBSCAN, particularly for harmful vibrations, is likely influenced by the FOCS cable's multi-layered protective structure, which dampens the physical effects that would otherwise be reflected in polarization changes.



(a) Bare fiber **(b)** FOCS cable

Figure 4: F1-scores for the two abnormal scenarios in bare fiber and FOCS cable.

Results for the more complex indoor cable scenarios with overlapping events are presented in Fig. 5. These results reinforce the consistent high performance of the OCSVM model across all scenarios. Regardless of whether the normal baseline is set to the relaxed fiber (*rlx_idr*) or the 155Hz vibration (*155vb_idr*), OCSVM maintains F1-scores above 0.97 in most scenarios, with peak values reaching 0.9983 in scenarios that involve combined eavesdropping and multi-frequency vibrations. Even in the most challenging *total* anomaly scenario, it achieves F1-scores of 0.9889 (*155vb*) and 0.9976 (*rlx*), demonstrating excellent generalization across diverse threat types. DBSCAN, in contrast, shows greater variability depending on the scenario and baseline condition. It achieves competitive performance in certain scenarios with clear separability, such as eavesdropping combined with dual-frequency vibration, where it attains F1-scores as high as 0.9979. However, its performance degrades for detecting harmful vibrations, where F1-score drops to 0.7112 (*155vb*) and 0.7303 (*rlx*). In the *total* anomaly scenario, DBSCAN achieves moderate to good results, with F1-score of 0.8433 for the *155vb* and 0.9223 for the *rlx* baseline.

6 Conclusion

In this study, we presented a comprehensive analysis of ML-based anomaly detection in optical fiber networks through the collection and processing of SOP data. We evaluated the performance of SSL and USL techniques, specifically OCSVM and DBSCAN. Our goal was to detect various abnormal events,

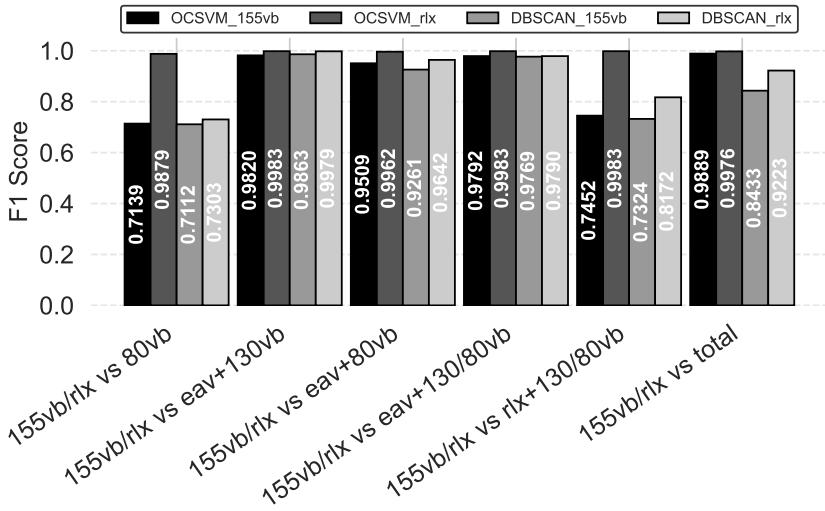


Figure 5: F1-score comparison of OCSVM and DBSCAN performance for the abnormal scenarios in indoor cable considering *155vb_idr/rlx_idr* as normal scenarios.

such as eavesdropping and harmful vibrations, by analyzing polarization signatures of three different cable types: bare fiber, FOCS, and indoor cable. Thirteen polarization signatures were collected under controlled experimental conditions, and careful hyper-parameter tuning was performed. Our findings indicate superior performance of OCSVM in detecting anomalies, with F1-score values exceeding 0.98 in most scenarios. Despite achieving high accuracy in some scenarios, DBSCAN exhibited greater variability in performance and poor performance in complex scenarios with overlapping events. While OCSVM consistently demonstrated high accuracy across most scenarios when trained on well-characterized normal data, our analysis highlights that DBSCAN retains significant value in deployment contexts where such baseline data is scarce or difficult to obtain. In such cases, DBSCAN's unsupervised nature enables it to detect clusters and anomalies without requiring prior labeling, making it a practical alternative when semi-supervised or supervised learning is infeasible. The results suggest a strong potential for the SSL and USL models to aid human security engineers in the process of detecting events

in optical networks, although their performance still shows a non-negligible amount of false positives and false negatives. By using the ML techniques investigated in this paper, human effort can be limited to the analysis of detected anomalies, rather than periodically analyzing the data. This can contribute to more cost-effective security solutions and more sustainable optical network operation. These findings underscore that the choice between SSL and USL techniques should be guided by the availability of training data and the specific operational constraints of the deployment scenario. Moreover, our results suggest that the performance of USL approaches like DBSCAN could be improved through additional post-processing steps, such as window-based temporal analysis of detection outputs. Exploring such enhancements to reduce the performance gap between USL and SSL models represents a promising direction for future research.

References

[1] M. Hoffman, *Cable cuts*, `http://all.net/CID/Attack/papers/CableCuts.html`, Accessed: 7 December 2024.

[2] J. Pesic, E. Le Rouzic, N. Brochier, and L. Dupont, "Proactive restoration of optical links based on the classification of events," in *15th International Conference on Optical Network Design and Modeling-ONDM 2011*, IEEE, 2011, pp. 1–6.

[3] Urbint, *Telecom fiber cuts: Causes, consequences, and prevention*, `https://www.urbint.com/blog/telecom-fiber-cuts-consequences`, Accessed: 23 November 2024.

[4] A. Harris and P. Castle, "Bend loss measurements on high numerical aperture single-mode fibers as a function of wavelength and bend radius," *Journal of Lightwave Technology*, vol. 4, no. 1, pp. 34–40, 1986.

[5] M. Zafar Iqbal, H. Fathallah, and N. Belhadj, "Optical fiber tapping: Methods and precautions," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, 2011, pp. 164–168.

[6] D. Dahan and U. Mahlab, "Security threats and protection procedures for optical networks," *IET Optoelectronics*, vol. 11, no. 5, pp. 186–200, 2017.

[7] W. Lee, S. I. Myong, J. C. Lee, and S. Lee, "Identification method of non-reflective faults based on index distribution of optical fibers," *Optics express*, vol. 22, no. 1, pp. 325–337, 2014.

[8] K. Abdelli, H. Grießer, C. Tropschug, and S. Pachnicke, "Optical fiber fault detection and localization in a noisy OTDR trace based on denoising convolutional autoencoder and bidirectional long short-term memory," *IEEE Journal of Lightwave Technology*, vol. 40, no. 8, pp. 2254–2264, 2021.

[9] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," *Journal of optical communications and networking*, vol. 14, no. 5, pp. 365–375, 2022.

[10] B. Steinar, "Locating disturbances in optical fibres," WO2022185075A1, Sep. 2022.

[11] Y. Aono, E. Ip, and P. Ji, "More than communications: Environment monitoring using existing optical fiber network infrastructure," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, W3G.1.

[12] G. Marra, D. Fairweather, V. Kamalov, P. Gaynor, M. Cantono, S. Mulholland, B. Baptie, J. Castellanos, G. Vagenas, J.-O. Gaudron, et al., "Optical interferometry-based array of seafloor environmental sensors using a transoceanic submarine cable," *Science*, vol. 376, no. 6595, pp. 874–879, 2022.

[13] A. Mecozzi, C. Antonelli, M. Mazur, N. Fontaine, H. Chen, L. Dallachiesa, and R. Ryf, "Use of optical coherent detection for environmental sensing," *Journal of Lightwave Technology*, vol. 41, no. 11, pp. 3350–3357, 2023.

[14] M. Cantono, J. C. Castellanos, V. Kamalov, A. Mecozzi, R. Muller, S. Yin, and Z. Zhan, "Seismic sensing in submarine fiber cables," in *ECOC Conf.*, IEEE, 2021, pp. 1–3.

[15] S. Pellegrini, L. Minelli, L. Andrenacci, D. Pilori, G. Bosco, B. Koch, R. Noé, C. Crognale, S. Piciaccia, and R. Gaudino, "Real-time demonstration of anomalous vibrations detection in a metro-like environment using a SOP-based algorithm," in *OFC Conf.*, 2024, pp. 1–3.

[16] D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J.-P. Elbers, "Cognitive assurance architecture for optical network fault management," *Journal of Lightwave Technology*, vol. 36, no. 7, pp. 1443–1450, 2018.

[17] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti, "Detection of abnormal activities on a SM or MM fiber," in *Optical Fiber Communication Conference (OFC)*, 2023, M3Z.6.

[18] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[19] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, "Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events," in *2024 24th International Conference on Transparent Optical Networks (ICTON)*, 2024, pp. 1–5.

[20] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[21] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Deep learning for detection of harmful events in real-world, noisy optical fiber deployments," *Journal of Lightwave Technology*, vol. 43, no. 13, pp. 6092–6101, 2025.

[22] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Breaking boundaries: Harnessing unrelated image data for robust risky event classification with scarce state of polarization data," in *European Conference on Optical Communications (ECOC)*, IET, vol. 2023, 2023, pp. 924–927.

[23] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, and P. Layec, "Anomaly detection and localization in optical networks using vision transformer and sop monitoring," in *Optical Fiber Communication Conference (OFC)*, 2024, Tu2J.4.

[24] K. Abdelli, M. Lonardi, F. Boitier, D. Correa, J. Gripp, S. Olsson, and P. Layec, "Vision transformers for anomaly classification and localization in optical networks using sop spectrograms," *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1902–1913, 2025.

[25] CONNECT Centre for Future Networks and Communications, *Open Ireland Testbed*, Available here.

[26] W. Qin, Q. Zhang, W. Hou, X. Zhang, and X. Gong, "Convolutional neural networks for fiber-bending eavesdropping attacks detection in coherent optical communication systems," in *2024 International Conference on Ubiquitous Communication (Ucom)*, IEEE, 2024, pp. 342–345.

[27] X. Chen, B. Li, R. Proietti, Z. Zhu, and S. J. B. Yoo, "Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks," *Journal of Lightwave Technology*, vol. 37, no. 7, pp. 1742–1749, 2019.

[28] L. Minelli, S. Pellegrini, L. Andrenacci, D. Pilori, G. Bosco, L. D. Chiesa, A. Tanzi, C. Crognale, and R. Gaudino, "SOP-based DSP blind anomaly detection for sensing on deployed metropolitan fibers," in *ECOC Conference*, vol. 2023, 2023, pp. 519–522.

[29] "Enhancing fiber security using a simple state of polarization analyzer and machine learning," *Optics Laser Technology*, vol. 167, p. 109 668, 2023, ISSN: 0030-3992.

[30] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

PAPER G

Generalizability of ML-Based Classification of State of Polarization Signatures Across Different Bands and Links

Leyla Sadighi, Carlos Natalino, Stefan Karlsson, Marco Ruffini, Eoin Kenny, Lena Wosinska, Marija Furdek

Published in 51st European Conference on Optical Communication (ECOC), 28 September to 2 October, 2025, Copenhagen, Denmark.
DOI: 10.1109/ECOC66593.2025.11263096
 ISBN: 979-8-3315-9531-9

The layout has been revised.

Abstract

We evaluate the Machine Learning (ML) model generalization for State of Polarization (SOP)-based event classification across spectral bands and links. Results show strong intra-system accuracy of up to 98.6% but limited cross-system generalizability, whereas multi-system training improves performance, highlighting the need for specific system-level knowledge.

1 Introduction

Fiber optic networks, essential to modern communications for high-speed, long-distance data transmission, are increasingly used for sensing and security monitoring applications [1]. Optical fiber sensors enable high-precision monitoring by leveraging scattering, interferometric effects, and light propagation changes to detect environmental perturbations [2]. The SOP is particularly sensitive to such perturbations, making it a valuable metric for network monitoring purposes, i.e., detecting external events and anomalies that affect the physical layer in the Open Systems Interconnection (OSI) model, thereby impacting the higher layers involved in data transmission through the fibers [3].

The interest in SOP-based sensing has grown significantly in recent years, largely due to the fact that coherent receivers inherently capture polarization-resolved optical field information. This enables software-based estimation of the SOP, making SOP-based network monitoring a promising and cost-effective approach [4]. Despite this advantage, effectively interpreting SOP signatures remains a challenge. Traditional monitoring systems that rely on rules and thresholds often fail to capture the complexity of real-world fiber events, particularly those that introduce subtle or transient SOP changes [5].

To address these limitations, recent studies explored the potential of ML for detecting SOP variations caused by various disturbances, enabling more robust and intelligent sensing capabilities [6]. For instance, in our prior work, we developed approaches based on Supervised Learning (SL) [7], [8], [9] and Deep Learning (DL) [10] for SOP analysis to detect fiber eavesdropping and mechanical vibrations from the incurred SOP alterations on a channel using

the O-band. Other studies investigated the application of ML to enhance event detection and classification in the C-band using SOP features [6], [11], [12], [13]. In general, existing efforts are confined to SOP data collected from a single spectral band and link. It is however interesting to analyze how an ML model trained on a given link and band behaves when tested on a different setup.

Polarization dynamics in optical fibers is influenced by wavelength-dependent physical-layer effects such as birefringence, Polarization Mode Dispersion (PMD), and varying modal coupling characteristics. These effects can cause the SOP to evolve differently across spectral bands. Hence, knowledge about SOP features extracted in one system may not directly translate to another, raising concerns about the transferability of ML models trained on data in a specific band and link. There is a lack of studies that examine the generalization of ML model performance when they are trained on one band and fiber link and applied to a different link on a different band. This raises a fundamental question: *even if similar physical events produce qualitatively comparable SOP signatures, do ML models for event detection generalize across bands and fiber links without retraining or adaptation?* We address this question by analyzing the generalization capabilities of an ML model, eXtreme Gradient Boosting (XGBoost), in classifying SOP-based signatures across two different fiber links using two spectral bands. We collect three distinct signatures from a real-world, noisy environment of the HEAnet live metro network (which operates in the C-band) [14] and compare the results with those obtained from a field dark fiber in the O-band. XGBoost was selected due to its consistent performance in our prior SOP-based studies [7], [8], [9] and its superior accuracy during preliminary testing. We train one XGBoost model per system and assess their performance when tested either on the same system or on a different one. We compare the performance to a multi-band model trained on an aggregate of both systems. Our study provides the first empirical evidence of band and link dependency of the ML classifier performance, indicating that polarization-based features learned in one system may not be directly transferable to another.

2 Experimental setup

The experimental setup used in this study is illustrated in Figure 1, which depicts two distinct systems designed to investigate the generalization ability of SOP-based event detection. The first system setup (labeled *System1*), operating in the O-band, consists of a 21 km (round-trip) dark fiber link accessed via the OpenIreland testbed. A Continuous Wave Distributed Feedback (CW-DFB) laser is used as the source, and the fiber under test is a standard Single Mode (SM) G.652 fiber. We used a field fiber rather than a lab spool in order to consider a more realistic noisy environment. The second system setup (labeled *System2*), operating in the C-band, is implemented over a live production metro network in the Dublin area, operated by Ireland's National Education and Research Network HEAnet [14]. Access to this network is provided via the OpenIreland laboratory at Trinity College Dublin through a dedicated dark fiber connection. The HEAnet metro ring comprises six Reconfigurable Optical Add-Drop Multiplexer (ROADM) nodes and spans a total fiber length of 77 km, and an External Cavity Laser (ECL) is used as the source. For our experiments, a 400 GHz spectral window ranging from 192.8 to 193.2 THz was allocated as a *Spectrum-as-a-Service* slice. To capture polarization signatures resulting from physical disturbances on the transmission line, we adopt the experimental setup from [10] at a wavelength in the C-band. The transmission line is subjected to various actions that emulate real-world tampering and eavesdropping scenarios capable of inducing measurable changes in the SOP of the transmitted signal. To extract signatures for each specific event, we use the methodology from [10]. For each signature and band, SOP samples are captured every 0.5 ms over a 20-minute interval, yielding 2.4 million samples per event. Numerical Polarization State Variation (NPSV) values are computed as distances between consecutive points on the Poincaré sphere and segmented into 500-sample windows. Each segment undergoes Fast Fourier Transform (FFT) analysis with 512 frequency bins using a Hamming window [15], resulting in a power spectrum dataset with 4,800 time slots (samples) and 512 frequency bins (features). This process generates two datasets, one for each band, which serve as the input for ML-based classification.

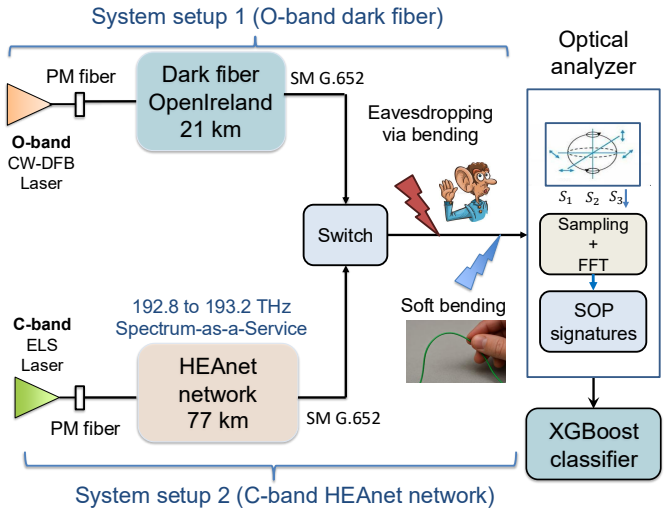


Figure 1: A schematic of the two experimental systems used in the study. System 1 (top) shows the O-band setup and system 2 (bottom) illustrates the C-band deployment.

3 Collected signatures and ML pre-processing

The ML analysis is based on one dataset per band, containing three distinct event signatures: relaxed (*rlx*), eavesdropping (*eav*), and soft bending (*sbd*). *rlx* represents the baseline state of the fiber, with only background environmental noise and no intentional disturbances. *sbd* simulates non-harmful physical handling during routine maintenance where the fiber is gently bent to a radius of approximately 2 cm at 10-second intervals, mimicking typical patch-panel manipulation. *eav* models a malicious eavesdropping attempt, where the fiber is bent using a specialized coupler with a 4 mm radius and a 25-degree angle, resulting in approximately 0.3 dB attenuation and 3% signal coupling, as described in [16].

The C-band dataset from the HEAnet network contains three collected signatures denoted as rlx_2, eav_2, and sbd_2, each initially comprising 4,800 samples, resulting in a total of 14,400 data points. We applied a post-processing filtering step to the eav_2 and sbd_2 classes to remove samples collected between two consecutive events, i.e., samples that represented the fiber in a relaxed

state. As a result, the filtered dataset contains 644 samples for eav_2 and 773 samples for sbd_2. An 80/20 train-test split was applied across all classes, resulting in a total of 4,973 training and 1,244 testing samples.

The O-band dataset contains the same three event signatures, denoted as rlx_1, eav_1, and sbd_1, with 4,800 samples per class, totaling 14,400 data points. This dataset did not require any filtering, as its signatures were collected with clearly separated event intervals, resulting in inherently clean and well-isolated samples. Using the same 80/20 split, this dataset yields 11,520 training and 2,880 testing samples overall.

These datasets are structured for a supervised ML classification task, where each sample is labeled according to its corresponding event class.

4 Results

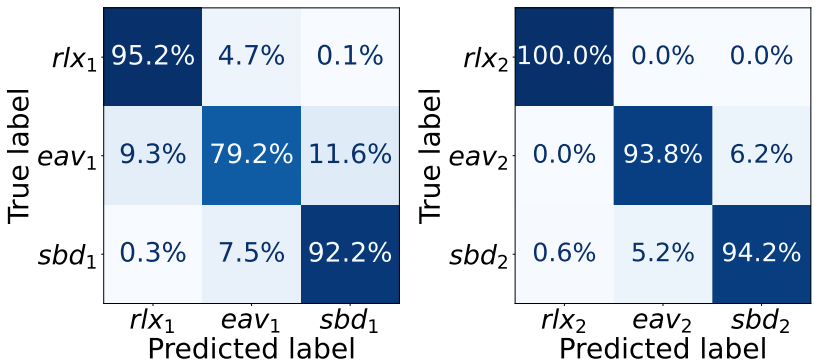
We initially evaluated the performance of ten different ML classifiers from the Scikit-Learn library to identify the most suitable model for SOP-based event classification. Among them, the XGBoost classifier demonstrated superior accuracy across both O-band and C-band datasets. Thus, we limit the scope of this analysis to XGBoost.

We test XGBoost on five scenarios: S_1 (training and testing both in System1), S_2 (training and testing both in System2), S_3 (training on System1 and testing on System2), S_4 (training on System2 and testing on System1), and S_5 (training and testing on a combined dataset from both System1 and System2). Table 1 reports the classification accuracy for each scenario, with confusion matrices shown in Figs. 2 - 4.

Table 1: Classification accuracy (Acc) for the different training/testing scenarios.

Scenario	Training	Testing	Acc.
S_1	System1	System1	88.85%
S_2	System2	System2	98.63%
S_3	System1	System2	8.11%
S_4	System2	System1	60.59%
S_5	System1+2	System1+2	91.11%

1) Intra-system classification (S_1 and S_2): The performance of the model in intra-system scenarios S_1 and S_2 is illustrated in Figure 2. The model achieves high classification accuracy for both systems, i.e., 88.85% in System1 and 98.63% in System2 (see Table 1). In S_1 (System1), the model correctly classifies most samples, although some confusion occurs between the *eav* and the other two classes. Nonetheless, the diagonal dominance of the matrix confirms that the event classes can be distinguished by the model. In S_2 (System2), the model achieves near-perfect classification, with 100% accuracy for *rlx*, i.e., no false positives, and over 93% accuracy for *eav* and *sbd*.



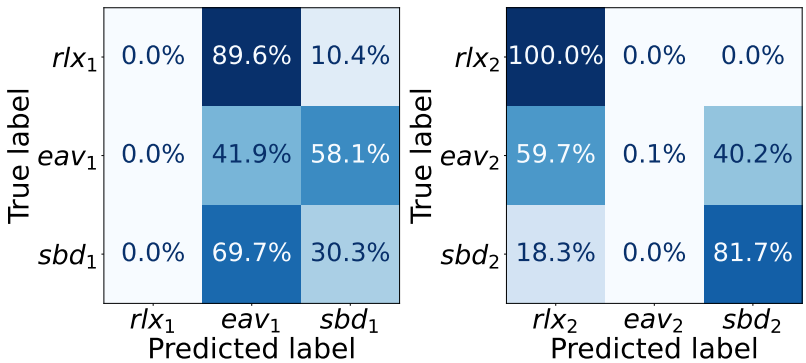
(a) System1 (S_1) **(b)** System2 (S_2)

Figure 2: Confusion matrices for intra-system classification

2) Cross-system generalization (S_3 and S_4):

The performance of the model in cross-system scenarios S_3 and S_4 is illustrated in Figure 3. In S_3 (System1 training $\rightarrow$ System2 testing), the classifier's accuracy drops dramatically to only 8.11% (see Table 1). This performance is worse than a naive classifier, which could theoretically achieve 33.3% accuracy by random guessing. The confusion matrix in Figure 3.a shows that the model fails to correctly classify most of the events, assigning the majority of test samples to the class eav_1. This indicates that the SOP signatures learned by the model from System1 do not apply to System2 data. In contrast, scenario S_4 (System2 training $\rightarrow$ System1 testing) yields an accuracy of 60.59% (see Table 1), suggesting partial generalization, i.e., higher accuracy than a naive classifier. Notably, in both cross-system scenarios, one class is completely mis-

classified: rlx_1 in S_3 and eav_2 in S_4, which indicates a higher susceptibility of certain event signatures to misclassification when transitioning between systems. These results suggest that optical network monitoring based on SOP signatures and ML exhibits a substantial degree of system sensitivity.



(a) System1 → System2 (S_3) **(b)** System2 → System1 (S_4)

Figure 3: Confusion matrices for cross-system generalization

3) Multi-system classification (S_5): The confusion matrix in Figure 4 illustrates strong per-class performance of the model trained and tested on a combined System1 and System2 dataset, achieving an overall accuracy of 91.11%. Despite some misclassifications, particularly between the *eav* and other classes, these results indicate that multi-system training significantly improves robustness and generalization, effectively mitigating the limitations observed in cross-system scenarios.

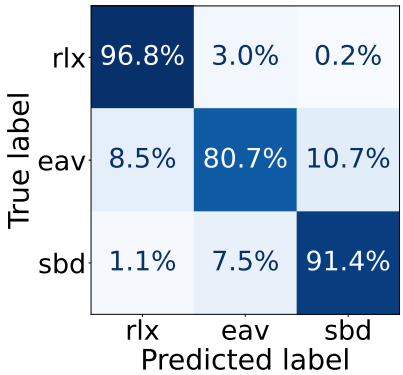


Figure 4: Confusion matrix for multi-system classification

5 Conclusion

Our study on the generalization capability of ML models for SOP-based event detection indicates high accuracy (up to 98.63%) in intra-system classification, a dramatic performance drop and asymmetric generalization in cross-system, and a beneficial trade-off (91.11%) in multi-system training scenarios. To refine generalizability, in our future work we will investigate domain adaptation and transfer learning strategies.

Acknowledgements

This work was supported by the Swedish Research Council (2023–05249), the European Commission's Digital Europe Programme (101127973) through the 5G-TACTIC project, the European Union's Horizon Europe research and innovation program (10113933) through the ECO-eNET project, and by the Research Ireland grant 18/RI/5721.

References

[1] G. Allwood, G. Wild, and S. Hinckley, "Optical fiber sensors in physical intrusion detection systems: A review," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5497–5509, 2016.

[2] P. Lu, N. Lalam, M. Badar, B. Liu, B. T. Chorpening, M. P. Buric, and P. R. Ohodnicki, "Distributed optical fiber sensing: Review and perspective," *Applied physics reviews*, vol. 6, no. 4, 2019.

[3] S. Pellegrini, L. Minelli, L. Andrenacci, G. Rizzelli, D. Pilori, G. Bosco, L. D. Chiesa, C. Crognale, S. Piciaccia, and R. Gaudino, "Overview on the state of polarization sensing: Application scenarios and anomaly detection algorithms," *J. Opt. Commun. Netw.*, vol. 17, no. 2, A196–A209, Feb. 2025.

[4] C. J. Carver and X. Zhou, "Polarization sensing of network health and seismic activity over a live terrestrial fiber-optic cable," *Communications Engineering*, vol. 3, no. 1, p. 91, 2024.

[5] D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J.-P. Elbers, "Cognitive assurance architecture for optical network fault management," *Journal of Lightwave Technology*, vol. 36, no. 7, pp. 1443–1450, 2018.

[6] "Enhancing fiber security using a simple state of polarization analyzer and machine learning," *Optics Laser Technology*, vol. 167, p. 109 668, 2023, ISSN: 0030-3992.

[7] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[8] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[9] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, "Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events," in *2024 24th International Conference on Transparent Optical Networks (ICTON)*, 2024, pp. 1–5.

[10] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Deep learning for detection of harmful events in real-world, noisy optical fiber deployments," *Journal of Lightwave Technology*, vol. 43, no. 13, pp. 6092–6101, 2025.

[11] A. Tomasov, P. Dejdar, P. Munster, and T. Horvath, "Utilizing a state of polarization change detector and machine learning for enhanced security in fiber-optic networks," in *CLEO 2024*, Optica Publishing Group, 2024, JTu2A.217.

[12] K. Abdelli, M. Lonardi, F. Boitier, D. Correa, J. Gripp, S. Olsson, and P. Layec, "Vision transformers for anomaly classification and localization in optical networks using sop spectrograms," *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1902–1913, 2025.

[13] F. Usmani, A. D'Amico, S. Straullu, F. Aquilino, R. Bratovich, E. Virgillito, and V. Curri, "A smart sensing grid for road traffic detection using terrestrial optical networks and attention-enhanced Bi-LSTM," *Journal of Lightwave Technology*, pp. 1–12, 2025.

[14] ASIERA (formerly HEAnet), *Ireland's National Education and Research Network*, Available here.

[15] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

[16] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

PAPER H

ML-Based Detection and Categorization of Complex Mechanical Vibrations via State of Polarization Analysis in Optical Networks

Leyla Sadighi, Stefan Karlsson, Marco Ruffini, Marija Furdek

Published in the 25th International Conference on Transparent Optical Networks (ICTON),
6–10 July, 2025, Barcelona, Spain.
ISBN: 978-1-6654-7164-0,

The layout has been revised.

Abstract

Modern optical networks form the critical backbone of global communications, enabling high-speed data transmission for a wide range of applications. Despite their inherent advantages in bandwidth and scalability, these networks are not immune to physical-layer vulnerabilities. Mechanical disturbances, both accidental and intentional, can compromise service quality or serve as gateways for more severe cyber-physical attacks. Thus, there is a growing need for intelligent, real-time monitoring solutions capable of detecting and interpreting subtle anomalies in optical fiber infrastructures. This paper presents a Machine Learning (ML)-based State of Polarization (SOP) monitoring approach for the identification and classification of complex mechanical vibrations in optical fiber networks. We address the real-world challenge of mixed-frequency and overlapping vibration signatures, arising from benign activities, malicious attacks, or simultaneous events, by collecting 14 distinct polarization signatures under various physical scenarios. A diverse set of supervised ML classifiers is evaluated, with Histogram Gradient Boosting (HGB) achieving the highest performance at 88.33% accuracy.

1 Introduction

Optical networks constitute the core infrastructure for high-capacity and long-distance data transmission, delivering low-loss and high-bandwidth connectivity that is essential to modern communication networks. Their role is critical in a wide range of domains, including telecommunications, healthcare, defense, and data center interconnects. Despite their physical robustness, optical fibers remain susceptible to external mechanical disturbances that can degrade signal quality or compromise security. In particular, construction activities and heavy machinery, such as excavators operating near buried cables, introduce a significant risk of accidental fiber cuts. These machines generate a characteristic frequency spectrum with a dominant low-frequency base tone, which can be detected to enable early warning and preventive action. More alarmingly,

deliberate security breaches such as covert eavesdropping, enabled by deliberately bending the fiber to a specific degree [1] to extract optical signals, raise serious concerns about the confidentiality of transmitted information. Such acts may go undetected without fine-grained monitoring.

In light of these vulnerabilities, recent real-world incidents involving sabotage and tampering of fiber-optic infrastructure have underscored the urgent need for advanced threat detection mechanisms [2]. Parallel to this, emerging research has demonstrated the potential of utilizing existing optical fiber infrastructure for environmental sensing, with proven effectiveness in capturing both natural and human-induced activities [3]. A key enabler of such sensing is the SOP, which is particularly effective in detecting subtle and complex vibration patterns caused by physical tampering. SOP-based sensing offers key advantages over traditional vibration detection methods due to its inherent sensitivity to minute physical perturbations in the optical fiber. Unlike conventional methods, such as Distributed Acoustic Sensing (DAS) or Optical Time Domain Reflectometry (OTDR), which often require specialized infrastructure, backscattering techniques, or high power levels, SOP-based analysis leverages the intrinsic polarization variations of light propagating through standard fiber [4]. Furthermore, when combined with ML techniques, SOP-based sensing enables real-time classification and anomaly detection, offering a scalable and robust solution for protecting critical optical communication infrastructure. Recent studies have extensively investigated ML-based analysis of SOP variations induced by physical disturbances such as eavesdropping attempts and mechanical vibrations [5], [6], [7], [8], [9], [10], [11]. These works demonstrate the effectiveness of SOP for anomaly detection frameworks in reliably identifying and characterizing disruptive events under experimental and real-world deployment scenarios. However, these research works did not consider the challenge of overlapping or mixed-frequency vibration patterns.

In a real-life installation, a normal event such as traffic passing close to the installation causes signatures with a broad-frequency spectrum content. This spectrum could be related to signatures caused by an eavesdropping event, or it can be mixed with other broad-spectrum signatures from normal traffic or from a malicious excavator threatening to cut the fiber optical cable. In order to avoid false alarms, it is therefore essential to classify and separate the frequency spectrum contents from signatures representing potential harmful events. In practical deployments, capturing clean and isolated

signatures of real-world mechanical disturbances, such as those caused by heavy vehicles or trains passing near the installation, is a time-consuming and operationally challenging task. These naturally occurring events often produce highly variable and overlapping vibration patterns that are difficult to reproduce under controlled conditions. To address this, we simulate disturbance environments by generating two complex vibration signatures. These synthetic signatures feature pseudo-random frequency content ranging from 0 to 2000 Hz, effectively emulating the spectral characteristics observed in real-world disturbances. To evaluate the classifier's robustness in differentiating between benign and potentially malicious activities, we further combine these complex vibration signatures with known attack scenarios, including fiber eavesdropping and soft patch cable bending. This paper addresses the gap in the literature by proposing an ML classification technique capable of separating and classifying complex, co-occurring mechanical signatures using SOP-based signatures analysis.

2 Experimental Setup

Our experimental setup, depicted in Figure 1, is designed to generate and analyze polarization signatures resulting from mechanical perturbations applied to an optical fiber transmission line. A stabilized 1310 nm Continuous Wave Distributed Feedback (CW-DFB) laser serves as the optical source, injecting constant polarized light into a 1 km coupling fiber. This coupling fiber connects to the central sensing region, where deliberate physical manipulations are applied to either a bare fiber or a patch cable segment. The perturbed signal then propagates through a 20 km fiber spool, bringing the total optical transmission length to 21 km and emulating a real-world long-distance deployment. Mechanical perturbations acting on the fiber installation induce measurable SOP variations, whose magnitude and frequency content are characteristic of the specific external activity, resulting in distinct polarization signatures. To simulate realistic and challenging monitoring conditions in this research work, we introduce two complex vibration patterns, referred to as complex vibration A (A) and complex vibration B (B), each lasting over 10 and 20 minutes, respectively. These patterns are composed of pseudo-random frequency components spanning the 0–2000 Hz range, mimicking naturally occurring disturbances such as persistent background noise generated by heavy

traffic or trains passing near fiber installations. To simulate overlapping effects, we mix complex vibration A and complex vibration B with known events like eavesdropping, soft bending, and potentially harmful vibrations.

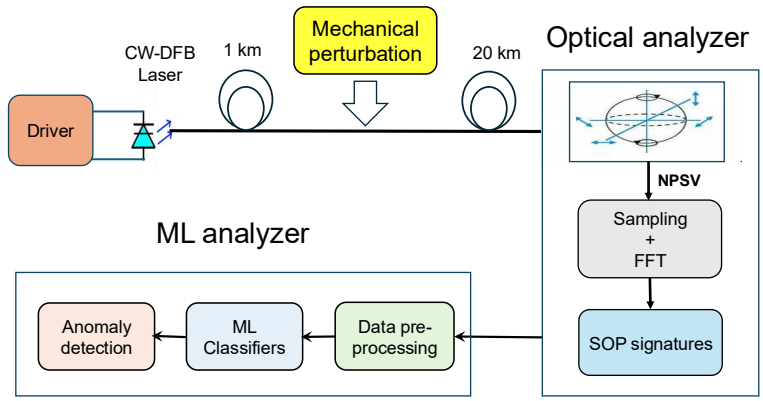


Figure 1: Schematic of the experimental testbed used for polarization signature analysis. Mechanical disturbances are applied between the 1 km coupling fiber and the 20 km fiber spool.

To analyze the impact of mechanical disturbances and generate distinct SOP variation signatures, we follow the approach introduced in [7]. As illustrated in Figure 1, an optical analyzer computes the numerical variation of the SOP, referred to as Numerical Polarization State Variation (NPSV), as time-sequenced samples on the Poincaré sphere at 0.5 ms intervals over 10-minute and 20-minute durations for complex vibrations A and B, respectively, yielding approximately 1.2 million and 2.4 million data points per event. The NPSV is then segmented into windows of 1000 samples and transformed into the frequency domain via an Fast Fourier Transform (FFT) with 512 bins. This results in a time-frequency data matrix (SOP signature) with the shape $[1200, 512]$ for events containing complex vibrations A and $[2400, 512]$ for complex vibrations B, forming the basis of our ML analysis.

The ML analysis is grounded in a comprehensive dataset generated through controlled combinations of complex vibrations and targeted mechanical perturbations applied to both bare fiber and patch cable segments. We consider

three representative tampering scenarios: soft bending (*sb*), eavesdropping (*eav*), and malicious vibrations at 80 Hz (*80vb*). The *sb* condition emulates benign maintenance activity, where fibers are gently bent to a radius of approximately 2 cm at 10-second intervals. The *eav* scenario simulates intentional tapping by introducing a 4 mm-radius, 25-degree bend using a precision coupler, as described in **sk_eav**. The *80vb* scenario replicates ground-borne vibrations from nearby excavation equipment, modeled by applying an 80 Hz, 60 dBA sinusoidal tone from a loudspeaker placed 5 cm from the fiber. These events are overlaid with either complex vibration A or B to emulate challenging real-world environments.

3 Signatures and Data Collection

We collected the following 14 distinct SOP signature involving two types of complex mechanical vibrations, denoted as A and B, applied to both bare fiber and patch cable under various tampering conditions:

- Complex A/B on bare fiber: A_{br}, B_{br}
- Complex A/B on bare fiber + eavesdropping: A_{br+eav}, B_{br+eav}
- Complex A/B on bare fiber + 80 Hz vibration: $A_{br+80vb}$, $B_{br+80vb}$
- Complex A/B on bare fiber + soft bending: A_{br+sb}, B_{br+sb}
- Complex A/B on patch cable: A_{pc}, B_{pc}
- Complex A/B on patch cable + soft bending: A_{pc+sb}, B_{pc+sb}
- Complex A/B on patch cable + eavesdropping: A_{pc+eav}, B_{pc+eav}

The ML analyzer, illustrated in Figure 1, comprises three main stages: data preprocessing, classification, and anomaly detection. In the preprocessing phase, the 14 collected signatures are merged to construct training and testing datasets. Each event involving complex vibration A consists of 1,200 points (8,400 in total across seven scenarios), while those involving vibration B yield 2,400 points (16,800 total). An 80/20 split is applied to each class, resulting in 960 training and 240 testing points per A scenario, and 1,920 training and 480 testing points per B scenario. After aggregation, the final dataset

comprises 20,160 training points and 5,040 testing points. This dataset is then analyzed using supervised ML models to detect anomalies indicative of potentially harmful or malicious events, enabling robust identification of overlapping and complex disturbances in real-world optical fiber installations.

4 Results

We conduct a comprehensive evaluation of multiple supervised ML classifiers to identify the most suitable model for detecting and categorizing mechanical vibration signatures in optical fibers. A diverse range of classifiers from the Scikit-learn library is evaluated, including ensemble-based methods (HGB, eXtreme Gradient Boosting (XGBoost),Gradient Boosting (GB), Random Forest (RF), Extra Trees (ET) Classifier), kernel-based models (Support Vector Machine (SVM)), linear models (Logistic Regression (LR), Linear Discriminant Analysis (LDA)), instance-based learning (K-Nearest Neighbors (KNN)), and Decision Tree (DT). These models are selected for their proven applicability in similar time-frequency classification tasks.

Table 1: Performance comparison of various ML classifiers

Classifier Name	**Accuracy**	**Precision**	**Recall**	**F1-score**	**Training Time (s)**	**Inference Time (s)**
HGB	0.8833	0.8856	0.8833	0.8828	47.1555	0.1349
XGBoost	0.8724	0.8741	0.8724	0.8716	24.9517	0.0368
RF	0.8387	0.8450	0.8387	0.8363	41.6140	0.0859
GB	0.8171	0.8211	0.8171	0.8157	3640.1562	0.1015
SVM	0.8163	0.8289	0.8163	0.8151	36.5793	25.3881
ET	0.7972	0.8076	0.7972	0.7906	7.6353	0.1281
LR	0.7014	0.6992	0.7014	0.6991	53.7773	0.0183
KNN	0.6671	0.7297	0.6671	0.6467	0.0187	0.5907
LDA	0.6609	0.6586	0.6609	0.6517	1.0038	0.0062
DT	0.6583	0.6592	0.6583	0.6585	16.2349	0.0041

Table 1 presents a detailed comparison of classifier performance in terms of accuracy, precision, recall, F1-score, training time, and inference latency. The HGB classifier delivers the highest overall performance, achieving an ac-

curacy of 88.33% and an F1-score of 0.8828. It also offers a favorable trade-off between predictive accuracy and computational efficiency. XGBoost and RF also demonstrate good performance, with accuracies of 87.24% and 83.87%, respectively. Although SVM yields comparable accuracy (81.63%), its inference time of over 25 seconds makes it impractical for real-time applications. Meanwhile, GB, despite being conceptually robust, requires a substantially longer training time (over 3,600 seconds) without a corresponding performance gain. Simpler classifiers such as LR, KNN, LDA, and DT show significantly lower accuracy and are less effective at capturing the complex patterns inherent in SOP-based vibration signatures.

The confusion matrix in Figure 2 further illustrates the effectiveness of the HGB classifier in accurately identifying the 14 mechanical disturbance scenarios. The perfect and near-perfect classification was achieved for $A_{br+80vb}$ and $B_{br+80vb}$ (100.0% and 99.4%), with similarly strong results for baseline scenarios such as A_{br} (95.4%) and B_{br} (99.4%). Some misclassifications occurred in scenarios involving overlapping signatures, particularly those with *sb* or *eav*. For instance, B_{br+eav} showed confusion with B_{pc+eav}, and both A_{pc+sb} and B_{pc+sb} were misclassified as related scenarios such as A_{pc+eav} and B_{br+sb}. These results are not unexpected. Signatures collected from bare fiber (*br*) exhibit a higher signal-to-noise ratio because the fiber is directly exposed to mechanical disturbances, allowing clearer SOP variations. In contrast, the patch cable (*pc*) structure dampens the mechanical vibrations before they reach the fiber core, resulting in weaker SOP signatures. Additionally, *sb* events introduce strong, repetitive SOP fluctuations at consistent time intervals—often stronger than the underlying complex vibration (A or B) present in the signal. This can cause the classifier to prioritize features of the *sb* over the background disturbance, leading to occasional misclassifications between *sb* and related classes. Despite these challenges, the classifier consistently preserved the dominant spectral cues that distinguish each class, confirming its strong effectiveness in SOP-based ML classification for recognizing subtle and overlapping disturbances in optical network environments.

5 Conclusion

This paper proposed an ML-based classification of SOP signatures analysis for identifying and categorizing complex mechanical disturbances in optical fiber

networks. Considering that real-world optical infrastructures are increasingly exposed to both benign and malicious mechanical activities, often occurring simultaneously and with overlapping spectral characteristics, we proposed a robust sensing approach leveraging SOP sensitivity and advanced supervised ML classification. A comprehensive dataset comprising 14 distinct SOP signatures was collected under controlled experimental conditions. Multiple supervised ML classifiers were evaluated for their performance in classifying these events, with HGB emerging as the most accurate and computationally efficient model, achieving an accuracy of 88.33%. Overall, the results validate the effectiveness of SOP-based spectral analysis combined with ML in accurately distinguishing and categorizing co-occurring mechanical disturbances, paving the way for robust monitoring and threat detection in future optical communication networks.

References

[1] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.

[2] BBC News, *France investigates after suspected sabotage hits fibre-optic cables*, `https://www.bbc.com/news/articles/cn52rglxr62o`, 2025.

[3] G. Marra, D. Fairweather, V. Kamalov, P. Gaynor, M. Cantono, S. Mulholland, B. Baptie, J. Castellanos, G. Vagenas, J.-O. Gaudron, et al., "Optical interferometry-based array of seafloor environmental sensors using a transoceanic submarine cable," *Science*, vol. 376, no. 6595, pp. 874–879, 2022.

[4] S. Pellegrini, L. Minelli, L. Andrenacci, G. Rizzelli, D. Pilori, G. Bosco, L. D. Chiesa, C. Crognale, S. Piciaccia, and R. Gaudino, "Overview on the state of polarization sensing: Application scenarios and anomaly detection algorithms," *J. Opt. Commun. Netw.*, vol. 17, no. 2, A196–A209, Feb. 2025.

[5] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures

over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[6] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, "Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events," in *2024 24th International Conference on Transparent Optical Networks (ICTON)*, 2024, pp. 1–5.

[7] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[8] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Deep learning for detection of harmful events in real-world, noisy optical fiber deployments," *Journal of Lightwave Technology*, vol. 43, no. 13, pp. 6092–6101, 2025.

[9] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Breaking boundaries: Harnessing unrelated image data for robust risky event classification with scarce state of polarization data," in *European Conference on Optical Communications (ECOC)*, IET, vol. 2023, 2023, pp. 924–927.

[10] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, and P. Layec, "Anomaly detection and localization in optical networks using vision transformer and sop monitoring," in *Optical Fiber Communication Conference (OFC)*, 2024, Tu2J.4.

[11] K. Abdelli, M. Lonardi, F. Boitier, D. Correa, J. Gripp, S. Olsson, and P. Layec, "Vision transformers for anomaly classification and localization in optical networks using sop spectrograms," *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1902–1913, 2025.

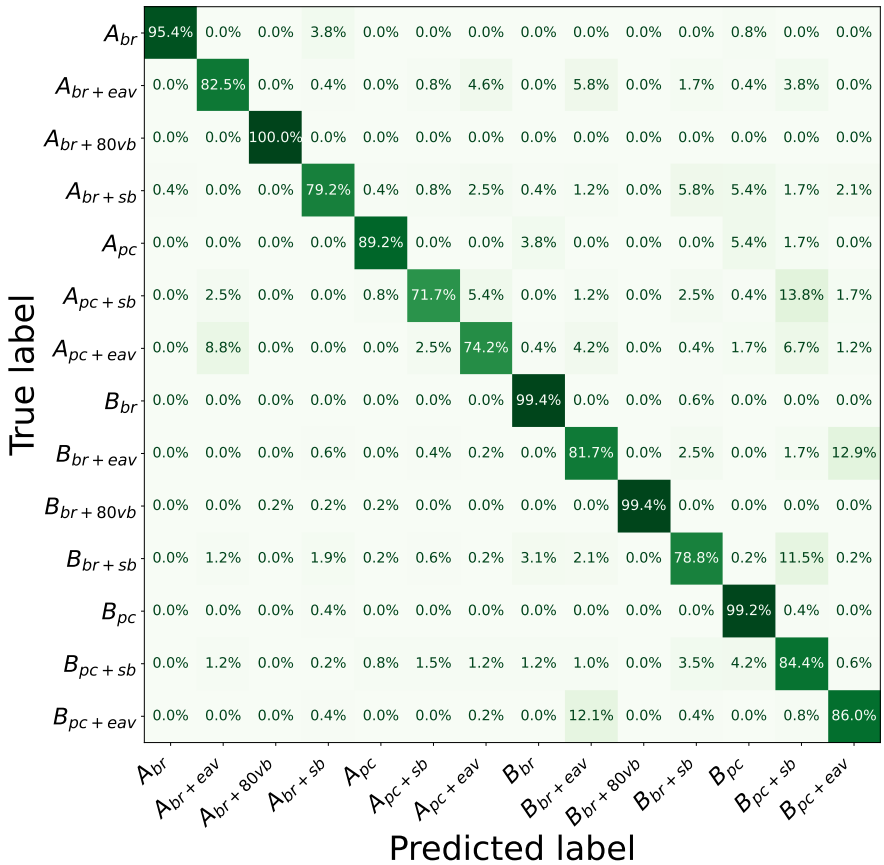


Figure 2: Results of Confusion Matrix for Histogram Gradient Boosting (HGB) classifier

PAPER I

DP-16QAM Modulated vs. Unmodulated Polarization Signatures for Machine Learning-Based Fiber Sensing

Leyla Sadighi, Carlos Natalino, Stefan Karlsson, Lena Wosinska, Eoin Kenny, Venkata Virajit Garbhapu, Marco Ruffini, Marija Furdek

Published in IEEE Journal of Lightwave Technology (JLT), February 2026.
DOI: 10.1109/JLT.2026.3660791

The layout has been revised.

Abstract

Fast and accurate detection of various physical layer threats that target optical networks is key to secure and reliable global communications. Conventional monitoring methods often fail to detect subtle anomalies, which requires advanced sensing techniques. Machine Learning (ML) analysis of the State of Polarization (SOP) of unmodulated signals was shown to successfully detect such disturbances. However, real-world networks typically operate with high-speed modulated signals, which may alter SOP behavior and challenge the applicability of ML techniques developed for unmodulated signals. This paper investigates the implications of signal modulation supporting high-data rates on the interpretability of SOP signatures. We perform the first experimental comparison of anomaly detection approaches based on SOP for Dual-Polarization 16-Quadrature Amplitude Modulation (DP-16QAM)modulated and unmodulated optical signals subjected to identical physical perturbations caused by fiber tapping and vibrations. We analyze four representative events under both signal modalities and assess the impact of modulation on SOP dynamics using a 63.4 km fiber link in a real-world metro network. We design four datasets that isolate, merge, and jointly classify the different signal modalities, and compare the performance of ten best-performing supervised ML techniques in each case. Our findings indicate that modulated signals tend to exhibit smoother SOP trajectories, likely due to the temporal averaging effects introduced by high symbol rates, wherein rapid symbol transitions suppress high-frequency polarization noise. Importantly, this smoothening does not obscure the slower, event-induced polarization drifts observed during physical disturbances, allowing ML models to reliably differentiate between different physical events (e.g., bending, vibrations) and signal modalities (modulated vs. unmodulated), achieving accuracy values between 97.12% and 98.47%.

1 Introduction

Optical fiber networks are the cornerstone of global connectivity, enabling high-capacity and low-latency communication for global internet services, cloud computing, financial systems, and critical infrastructure. As these networks scale in speed and complexity, their role in supporting critical services grows accordingly. Vulnerability of optical fibers to various threats, including mechanical stress, accidental damage, and deliberate intrusions such as fiber tapping [1], poses a risk of security violations and service disruptions. Therefore, ensuring the integrity and security of optical transmission networks is of paramount importance, particularly as cyber-physical security becomes a growing concern in next-generation network deployments.

One of the key components of optical network security management is the timely detection of physical-layer anomalies. Techniques such as Optical Time Domain Reflectometry (OTDR) are commonly used to identify severe faults, like fiber cuts or sharp bends, by analyzing Rayleigh backscattering [2], [3], [4]. They offer relatively precise fault localization and are well established in field deployments. However, their practical use is often constrained by deployment cost and limited scalability [5], and they typically lack the sensitivity needed to detect more subtle disturbances, such as low vibrations or small mechanical deformations. Other approaches, including Distributed Fiber Optic Sensing (DFOS) for intrusion monitoring [6], offer higher sensitivity but require specialized hardware and complex processing, which significantly increases the cost and complicates their adoption. SOP analysis has emerged as a powerful tool for monitoring and securing optical fiber networks. The statistical and dynamic properties of the SOP carry valuable information about environmental disturbances and physical-layer threats such as fiber tapping or mechanical perturbations. Unlike traditional methods that rely on backscatter or reflection signatures, SOP-based techniques exploit the intrinsic polarization sensitivity of optical signals [7]. This enables the detection of subtle anomalies with minimal modifications of the existing infrastructure. As a result, SOP-based monitoring provides a cost-effective and streamlined alternative, avoiding the hardware complexity and deployment overhead associated with DFOS and OTDR, albeit without the ability of event localization along a fiber. SOP-based sensing in coherent transmission systems can be implemented using either an external polarization analyzer, as adopted in our previous work [8], or by leveraging the internal Digital Signal Processing (DSP)

of a coherent receiver. While external analyzers directly measure polarization fluctuations without accessing signal data, coherent receivers inherently estimate and compensate fiber birefringence as part of their DSP pipeline. This internal tracking can be exploited for polarization sensing, as demonstrated in [9], where the authors compare DSP-based phase and polarization sensing in a deployed metro network. Notably, this approach naturally resolves the low-Degree of Polarization (DoP) limitation encountered by standalone analyzers when observing rapidly modulated signals.

Accurate monitoring and detection of polarization variations induced by external events is crucial for identifying fiber disturbances and preserving network reliability. Traditional SOP-based monitoring approaches that depend on static thresholds or heuristic rules, such as the one in [10] often fall short when facing complex or evolving physical-layer threats [11]. To address these limitations, recent research has turned increasingly to data-driven methodologies. By leveraging ML, such approaches enable automated, scalable, and adaptive analysis of polarization behavior, empowering the system to recognize intricate disturbance patterns directly from SOP data.

Despite the proliferation of ML-based techniques for accurate SOP monitoring, most existing studies use simplified experimental setups with limited conditions, i.e., the absence of signal modulation. This exclusive reliance on unmodulated signals raises an important question about the applicability of these methods in practical, real-world systems. Modern optical networks predominantly operate with modulated signals, particularly in coherent systems using modulation formats like Quadrature Phase-Shift Keying (QPSK) or Quadrature Amplitude Modulation (QAM). It should be noted that the modulated signals considered in this paper are coherent polarization-multiplexed 16-ary DP-16QAM channels generated and detected using coherent transceivers. This is fundamentally different from Intensity-Modulated Direct-Detection (IM-DD) systems, which typically employ Non-Return-to-Zero (NRZ) modulation and exhibit nearly fully polarized optical carriers (DoP $\approx$100% at 1 ms integration), resulting in polarization behavior opposite to that of DP-16QAM signals. Furthermore, existing studies lack a direct experimental comparison between SOP signatures of modulated and unmodulated signals under similar physical and environmental conditions. Without such comparison, it remains unclear whether the presence of modulation fundamentally alters the SOP variations in a way that impacts the effectiveness

of monitoring techniques like ML-based SOP fiber sensing.

Unmodulated signals maintain a nearly fixed SOP apart from slow environmental drifts [12], which offer the advantage of relatively clean polarization trajectories, making them particularly amenable to DSP and ML-based classification for anomaly detection. Unlike unmodulated signals, a modulated data signal's SOP can fluctuate on sub-nanosecond timescales due to the rapid symbol transitions at multi-Gbps rates. In effect, SOP is no longer represented by a single, stable point on the Poincaré sphere, but hops among a continuum of states dictated by the bit sequence [13]. The fast polarization fluctuations in modulated signals present new challenges for sensing and monitoring. They can act as a high-frequency noise floor that masks the more gradual SOP rotations caused by physical disturbances. Fiber bends, vibrations, or taps typically induce SOP changes on the Hz-kHz scale, whereas symbol-rate polarization changes occur at the GHz scale. When a modulated channel is observed at a slower timescale than its symbol rate, as is often the case with photodiode-based or low-cost polarization analyzers, the rapidly fluctuating SOP appears scrambled. The analyzer effectively averages over many symbols, and this temporal averaging reduces the measured DoP, resembling the effect of a polarization scrambler [14]. This means that, paradoxically, a high bit rate signal might exhibit a smoother SOP trajectory when viewed in aggregate.

Any anomaly detection scheme must ensure that the residual fast SOP variations in modulated signals do not trigger false alarms or confound true event signatures. To address the gap in the literature, we perform a comprehensive experimental comparison between modulated and unmodulated optical signals under similar physical and environmental conditions. We systematically evaluate the impact of modulation on SOP dynamics and assess its implications for ML-based anomaly detection. The main contributions include detailed statistical analyses comparing the behavior of polarization signatures in modulated and unmodulated signals, offering quantitative evidence of how modulation affects polarization dynamics. We collect a real-world experimental dataset from a 63.4 km fiber link in the HEAnet [15] metro network, capturing eight representative polarization event signatures: relaxed fiber, soft bending, eavesdropping attempt, and 80 Hz vibration, each recorded for modulated and unmodulated signal conditions. These events were selected to represent a diverse mix of normal (e.g., relaxed, soft bending) and potentially harmful or

malicious (e.g., eavesdropping, vibration) fiber conditions commonly encountered in operational environments and security-sensitive scenarios. We then evaluate the performance of a range of supervised ML classifiers for these events and for different signal modalities. We analyze four scenarios, each corresponding to a distinct dataset (i.e., separated, mixed, and joint modality classification), assessing the impact of modulation on the ability of ML models to learn and separate polarization signatures corresponding to modulated and unmodulated signals.

The remainder of the paper is organized as follows. Section 2 reviews prior work on SOP analysis for modulated and unmodulated signals. Section 3 presents the experimental testbed and data collection process. Section 4 details the events used for SOP signature analysis and a statistical comparison of modulated vs. unmodulated signals. Section 5 introduces the dataset configurations and preprocessing. Section 6 reports experimental results, and Section 7 concludes the paper.

2 Related Work

A wide spectrum of ML approaches has been developed to enhance the interpretability and adaptability of SOP-based monitoring. Supervised Learning (SL) techniques are often applied to detect and classify known physical disturbances and malicious activities such as fiber tapping and mechanical intrusions with high accuracy [8], [16], [17], [18], [19], [20]. Unsupervised Learning (USL) methods, such as clustering and outlier detection, allow the identification of previously unseen or emerging anomalies in SOP trajectories without requiring labeled data [21], [22], [23], [24]. Semi-Supervised Learning (SSL) approaches strike a balance by leveraging a small set of labeled normal data along with abundant unlabeled samples, including unknown disturbances, to improve generalization under sparse annotation scenarios [24], [25]. The majority of ML-driven SOP analyses rely on Continuous Wave (CW) light sources, particularly unmodulated Distributed Feedback (DFB) lasers to simplify the polarization behavior and avoid the complexities introduced by high-speed modulation [7], [8], [18], [19], [26], [27]. The review of polarization-based fiber sensing methods and anomaly detection algorithms [7] demonstrated that polarization measurements from a CW source, obtained via a polarimeter, reflect only the disturbance-induced SOP changes, whereas modulated signals

introduce additional SOP estimation noise, confirming that unmodulated light yields clean polarization signatures that can be readily interpreted for anomaly detection. Our prior experimental studies [8], [19], [28] demonstrated successful use of supervised ML techniques to detect and classify disturbances like harmful vibrations and eavesdropping in controlled setups with unmodulated signals, with the impact of noise in real-world environments considered in [18], [26].

Modulation-induced polarization effects have received comparatively little attention, primarily approached from the communications perspective. Recent work by Karlsson *et al.* demonstrated SOP-based event detection using coherent transceivers in live networks. A real-time monitoring system detected polarization precursors before a cable break [29]. Follow-up analysis showed that mechanical disturbances produce distinct SOP patterns separable from noise [30]. In coherent communications, polarization tracking algorithms follow the time-varying SOP of a high-speed modulated signal so that the receiver can correctly demultiplex the polarization-multiplexed channels. For instance, a stochastic polarization drift model that treats the SOP evolution as a random walk on the Poincaré sphere can be found in [31]. Empirical studies have shown that environmental factors can indeed induce rapid SOP changes in deployed systems. SOP fluctuations on ms time scale in aerial fibers under varying climatic conditions were reported in [32], while cyclic SOP oscillations in overhead fiber cables caused by wind gusts and power-line electromagnetic interference were observed in [33]. Lightning-induced abrupt, fast polarization rotations, were shown to necessitate sub-ms reaction times in coherent receivers [34]. Real-time polarization tracking in modern coherent transceivers' DSP, e.g., the adaptive filters that continuously correct for polarization rotation and mode dispersion, allowing the receiver to lock on the signal's SOP [35], [36], treat polarization fluctuations purely as impairments to be corrected, rather than indicators of physical network disturbances. The effect of modulation on polarization sensing has not been explicitly addressed, and existing research lacks a systematic investigation of how modulated signals compare to unmodulated lights in terms of polarization disturbance detectability. In this work, our objective is to identify statistically meaningful patterns in SOP dynamics that correlate with external disturbances, which is key to understanding the role of ML-based polarization monitoring as a sensing tool.

3 Experimental Setup

3.1 Network topology and setup

To evaluate the effects of optical signal modulation on the behavior of polarization signatures under a realistic noisy environment, we use the experimental setup in Figure 1, illustrating the 63.4 km test channel route of the HEAnet Dublin metro ring, connected to the OpenIreland [37] testbed infrastructure. The route traverses six Reconfigurable Optical Add-Drop Multiplexers (ROADMs) in sequence: TCD-1 (Lloyd Institute, Trinity College Dublin (TCD)), UCD-2 and UCD-1 (two nodes located at University College Dublin (UCD), Belfield Campus), CWT-2 (Citywest), PW-2 (Park West), and TCD-2 (Pearse Street, TCD), with the corresponding fiber segment lengths indicated in Fig. 1. The optical signal is added at TCD-1 and dropped at TCD-2 before returning to the OpenIreland Lab via a cross-campus fiber patch, where controlled disturbances are applied. The HEAnet ring is lightly loaded, and the experiment runs over a 400 GHz Optical Spectrum as a Service (OSaaS) window in the C-band (192.8–193.2 THz). The SOP sensing is carried out over two wavelength channels injected from a Lumentum ROADM in OpenIreland into the HEAnet entry ROADM. The first channel (λ_m, centered at 193 THz) is a coherent DP-16QAM 200 Gbps channel generated by an Adtran Teraflex transceiver, while the other (λ_u, centered at 193.1 THz) is an unmodulated signal generated by an External Cavity Laser (ECL) source. The two channels are within the same 400 GHz OSaaS window, allowing for a fair comparison focusing on the impact of modulation. Only one of the two signals is active at a time.

The signals traverse the same path, which ensures that both unmodulated and modulated channels experience similar propagation conditions. Here, the signals are exposed to controlled physical disturbances while traveling on a fiber patch that is part of the equipment shown in Figure 1, including soft bending, eavesdropping by bending the fiber, and 80 Hz vibrations, which are introduced to evaluate the impact on polarization signatures. The signals are then demultiplexed by a second Lumentum ROADM (DEMUX) and sent to the polarization analyzer. Our polarization analyzer instrument is a commercial polarization sensing module [8] (a "black box") capable of measuring all three Stokes polarization components (S_1, S_2, S_3) variations. This device uses a carefully designed arrangement of passive optical components

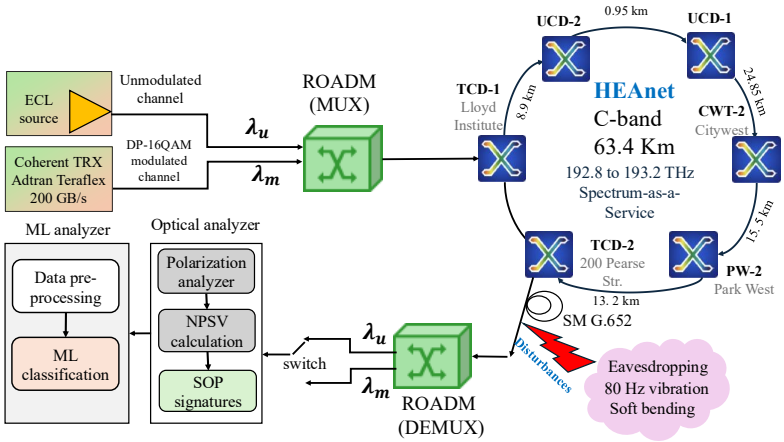


Figure 1: Schematic of the experimental and analytical setup used to investigate the impact of signal modalities on SOP dynamics under controlled perturbations.

to project the signal onto different polarization bases. The complete processing pipeline, including additional optical analyzer components and analysis stages, is described in the following subsection.

3.2 Data Collection Process

Upon transmission over the same path and exposure to similar physical disturbances, each optical signal is directed to the optical analyzer. This analyzer implements a structured processing pipeline for extracting and analyzing the Numerical Polarization State Variation (NPSV) data and generating SOP signatures for each disturbance event and signal modality. In the first stage of processing, the optical polarization analyzer captures the temporal evolution of the SOP variations by projecting the received signal onto the Poincaré sphere. For each experimental run, the signal is continuously sampled for 15 minutes at 0.5 ms intervals, resulting in approximately 1.8 million NPSV samples per experiment. At each time slot t, the value NPSV_t serves as a scalar indicator of the magnitude of SOP variation between two adjacent sampling points, i.e., during the interval $[t-1, t]$. To quantify this, we define the polarization intensity (the norm of the Stokes vector) at time in-

stance τ i.e $S_{0,\tau}$ as $S_{0,\tau} = \sqrt{S_1^2(\tau) + S_2^2(\tau) + S_3^2(\tau)}$, where S_1, S_2, and S_3 are the normalized Stokes parameters corresponding to the horizontal/vertical, diagonal/anti-diagonal, and right-/left-handed circular polarization components, respectively. The normalized polarization magnitude at time t is then given by

$$A_t = \frac{S_1^2(t) + S_2^2(t) + S_3^2(t)}{S_{0,t}} \tag{I.1}$$

and likewise for the previous sample:

$$A_{t-1} = \frac{S_1^2(t-1) + S_2^2(t-1) + S_3^2(t-1)}{S_{0,t-1}} \tag{I.2}$$

The resulting normalized polarization state variation is computed as

$$\text{NPSV}t = A_t - A_{t-1}. \tag{I.3}$$

Although the 200 Gbps Dual-Polarization (DP)-16-Quadrature Amplitude Modulation (16QAM) modulated signal exhibits rapid polarization fluctuations at the symbol rate (tens of GHz), these are effectively averaged out by our optical analyzer, which operates at a sampling rate of 2 ksamples/s and bandwidth of 1–2 kHz. As a result, our measurements isolate only the slow SOP drifts caused by environmental or mechanical disturbances below 5 kHz. These low-frequency effects are preserved and comparable in both modulated and unmodulated channels. In both cases, these vibrations cause a rotation of the polarization state along the fiber, causing again a movement of the NPSV SOP state, which we can detect using the variation of S1, S2 and S3 parameters, which are measured by our device. To analyze the spectral features of the SOP variations and transfer the signal from the time to the frequency domain, each 0.5-second segment (comprising 500 NPSV values) undergoes Fast Fourier Transform (FFT) processing with a Hamming window [38]. Each FFT generates 512 frequency components. This results in a time-frequency representation of size $3{,}600 \times 512$, where each row corresponds to a 0.5-second time slot and each column to a specific frequency bin. These spectral profiles form what we refer to as *SOP signatures*. Figure 2 summarizes this data-processing pipeline, illustrating the transformation from raw NPSV measurements (1.8 million samples) into 3,600 time segments, followed by FFT-based spectral decomposition with 512 frequency bins. The resulting

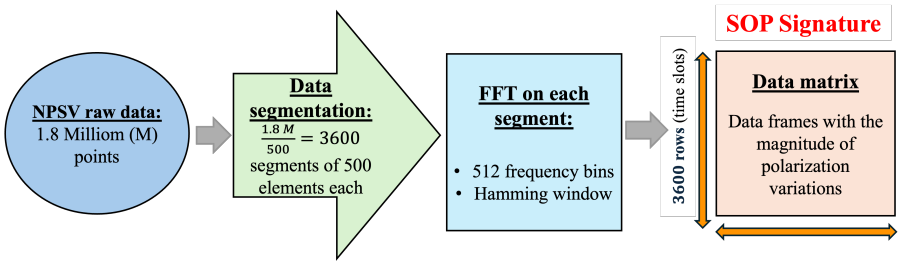


Figure 2: Data collection and DSP processing pipeline, from raw NPSV data to the SOP signature used for ML classification.

3,600×512 dataset captures the spectral magnitude of polarization variations over time and constitutes the final SOP signature used in the analysis.

The generated SOP signatures are forwarded to the ML analyzer for event and signal modality classification. The ML analysis is conducted in two stages. First, a data pre-processing pipeline aggregates and prepares the input SOP signatures for ML investigation, including partitioning of the dataset into training and testing subsets. In the second stage, a suite of supervised ML classification algorithms, ranging from ensemble methods to kernel-based and linear models, is evaluated to identify the most suitable classifier for each dataset. The selection is based on the highest classification accuracy and overall performance across precision, recall, and F1-score, as detailed in Section 6. The top-performing model is then used to infer the class of unseen samples and compute the corresponding performance metrics.

4 Experimental Scenarios

4.1 Types of Disturbances

Using the described experimental setup, we collect SOP signatures for eight types of events encompassing normal operating conditions and abnormal, harmful events. We consider two classes of normal fiber activity: relaxed fiber (*rlx*) and soft bending (*sbd*). The relaxed condition serves as a baseline reference, capturing only routine background noise in the absence of any deliberate physical disturbances. In contrast, soft bending reflects benign mechanical interactions that are typically encountered during routine handling

and maintenance of fiber installations. To simulate such conditions typically found in patch panels, a fiber segment was gently bent by hand to a curvature radius of approximately 2 cm. This action was repeated at 10-second intervals, mimicking handling behavior commonly exhibited by data center technicians.

The abnormal events we consider encompass eavesdropping attempts and potentially harmful vibrations. To simulate an eavesdropping attempt (*eav*), we adopt the mechanical interaction described in [39], where an attacker breaches the outer layer of a standard G.652 fiber and introduces a controlled bend in the internal core with a curvature radius of 4 mm and a bending angle of 25 degree. This controlled deformation enables signal leakage suitable for covert interception. This configuration allows for successful signal tapping while remaining nearly undetectable under standard power monitoring.

To simulate harmful vibrations, we introduce a mechanical disturbance at 80 Hz (*80vb*), which is equivalent to a frequency commonly associated with heavy machinery such as excavators. These machines present a tangible threat to optical fiber infrastructure, as their activity may unintentionally damage or sever fiber cables. The dominant vibration frequency stems from the rotational speed of the engine, which typically operates at around 4,800 Revolutions Per Minute (RPM), equivalent to 80 Hz. To replicate this real-world scenario in a controlled environment, we positioned a loudspeaker 2 to 4 cm away from the fiber under test and generated an 80 Hz acoustic signal. The signal's intensity corresponds to normal human conversation and is a conservative approximation, considering that actual excavators generate significantly higher vibration amplitudes. The goal of this setup is to evaluate the system's ability to detect early-stage mechanical interference before it escalates into service disruption.

4.2 Statistical Analysis of the SOP Signatures of Unmodulated vs. Modulated Signals

To evaluate the effect of signal modulation on the statistical behavior of polarization signatures, we perform a comparative analysis of the four collected signatures for unmodulated (denoted by rlx_u, eav_u, sbd_u, $80vb_u$) and modulated (denoted by rlx_m, eav_m, sbd_m, $80vb_m$) signal configurations. We analyze the NPSV distribution for each event and extract four key statistical parameters from these distributions: the mean, standard deviation, skewness, and kurtosis, which offer a comprehensive description of the shape and dynamics of the distribution. The mean μ quantifies the average magnitude of variation

Table 1: Statistical properties of NPSV obtained for modulated and unmodulated signals in different scenarios.

Event	**Mean** (μ)	**Standard Deviation** (σ)	**Skewness** (γ)	**Kurtosis** (κ)
rlx_u	3.77	18.38	-9.64	276.24
rlx_m	4.01	3.60	-0.29	-0.59
eav_u	3.93	11.00	-4.97	232.43
eav_m	4.03	2.53	-0.12	0.13
sbd_u	-58.12	144.54	-2.52	5.59
sbd_m	4.00	3.01	0.10	-0.14
$80vb_u$	3.90	19.20	-0.08	0.98
$80vb_m$	4.00	3.23	0.01	-0.50

and serves as a measure of the central tendency of the distribution. The standard deviation σ captures the extent of dispersion around the mean, reflecting the variability in polarization dynamics. Skewness γ describes the asymmetry of the distribution; negative values indicate a longer tail on the low NPSV values (i.e., more frequent low-magnitude deviations), while positive values suggest a longer tail on high NPSV values. Finally, kurtosis κ describes how sharply peaked a distribution is and how often extreme values occur. It measures the tendency of a distribution to produce outliers by describing the heaviness of its tails relative to a normal distribution ($\kappa = 0$). High kurtosis value means that the distribution has a sharp peak and more values far from the average (i.e., more outliers). Low or negative kurtosis value means that the distribution is flatter, with fewer extreme values.

Table 1 summarizes the statistical characteristics of the NPSV distributions for the modulated and unmodulated signal modalities under the four collected event signatures. In all conditions, the modulated signals consistently exhibit markedly lower standard deviations, reduced kurtosis, and skewness values closer to zero than those of the unmodulated counterpart. These patterns suggest that modulation plays a significant role in suppressing stochastic polarization fluctuations, yielding more symmetric and approximately Gaussian-like NPSV distributions. This suppression effect is not a consequence of intrinsic polarization behavior, but rather stems from limitations in the resolution of the polarization analyzer. The mechanical disturbances investigated in this

paper, such as fiber bending and vibrations, induce SOP variations at much lower frequencies at rates in the kHz, which is a very slow variation compared to the 200 Gbps high-speed modulation scheme of DP-16QAM. As a result, modulation indirectly contributes to the reduction of apparent stochasticity by decorrelating slowly varying polarization states and leading to smoother NPSV signatures in this case.

In the relaxed condition, the unmodulated signal displays considerable variability ($\sigma = 18.38$) and an extremely peaked and heavy-tailed ($\kappa = 276.24$), indicating heavy tails and the presence of rare, large deviations from the mean. In contrast, the modulated counterpart is significantly more stable ($\sigma = 3.60$) and near-Gaussian ($\kappa = \breve{}0.59$), with a slightly elevated mean ($\mu = 4.01$ vs. 3.77). Similar stabilizing effects of modulation are observed in the eavesdropping and 80 Hz vibration scenarios, where modulation reduces standard deviation by approximately 77% and 83%, respectively.

The soft bending event reveals the most striking contrast. The unmodulated signal exhibits a dramatically negative mean ($\mu = -58.12$) and very high dispersion ($\sigma = 144.54$), reflecting intense and irregular polarization disturbances. Conversely, the modulated case maintains a stable distribution with a baseline mean ($\mu = 4.00$) and low variance ($\sigma = 3.01$), closely resembling the relaxed condition. However, this superficial similarity in statistics can be misleading: while the modulated signal appears stable in terms of mean and variance, the full polarization trajectory may still contain subtle but relevant temporal patterns.

Figure 3 illustrates the statistical trends discussed in Table 1 by showing the NPSV histograms and their Gaussian fits for modulated and unmodulated signals under the four events. The plot has been scaled uniformly in the two axis to allow for a comparative analysis. In each subplot, the unmodulated distributions exhibit broader, often asymmetric profiles with heavier tails. By contrast, the modulated signal distributions are consistently narrower, more symmetric, and closely aligned with their Gaussian approximations, indicating reduced variability and improved stability in capturing polarization behavior. These visual observations support the numerical findings in Table 1 and highlight the benefits of modulation in achieving consistent polarization signature profiles, indicating that it might be more suitable for anomaly detection.

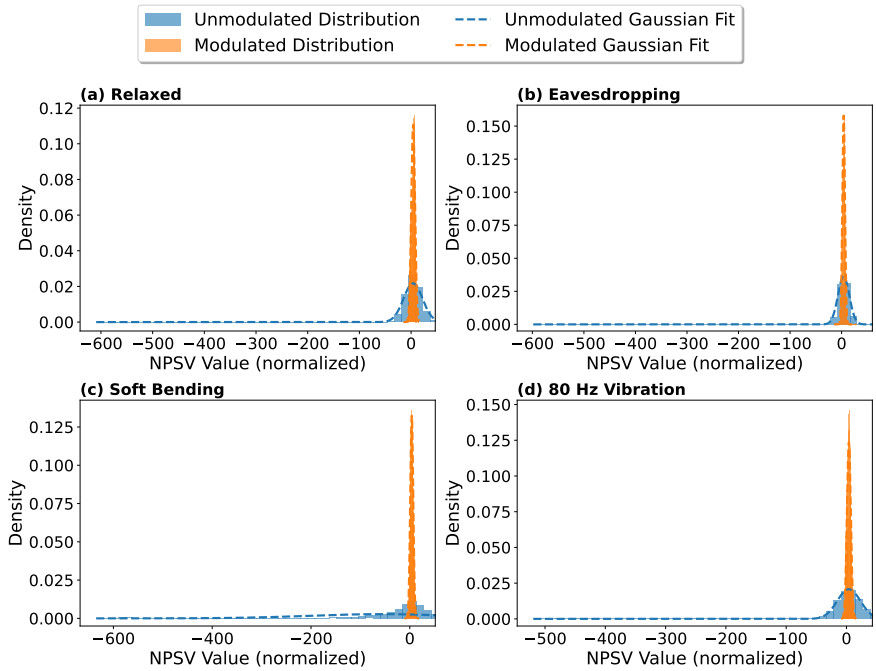


Figure 3: NPSV histograms and Gaussian fits for unmodulated and modulated conditions in four event types: (a) relaxed, (b) eavesdropping, (c) soft bending, and (d) 80 Hz vibration.

5 Dataset Definition and Pre-Processing

This section introduces four systematically designed datasets, each developed to investigate a distinct aspect of ML-driven classification for SOP-based optical fiber monitoring. The goal is to assess how signal modulation influences the effectiveness of ML techniques in distinguishing physical events, learning polarization-based features, and generalizing for different signal modalities. The datasets are constructed to evaluate the performance of the ML classification task: from separated analysis of signal types, to mixed-modality configuration, and finally to a fully discriminative and signal-aware framework. For each dataset, a corresponding data pre-processing pipeline is implemented to support its specific structure and objective. The four dataset configurations

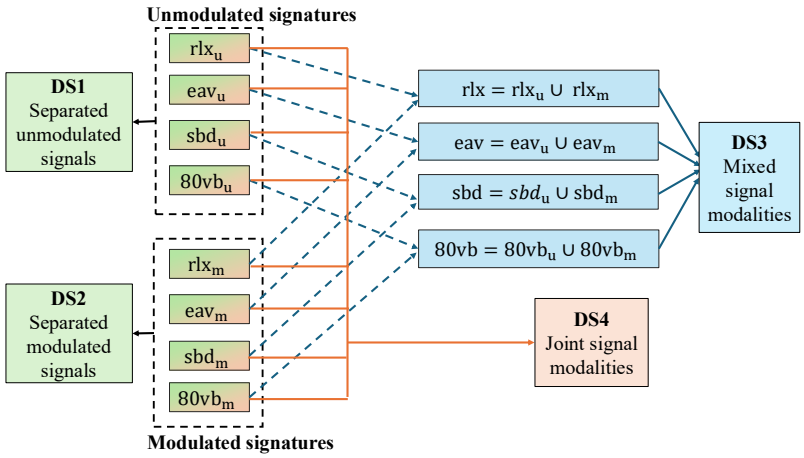


Figure 4: The four datasets for evaluating ML-based classification of SOP signatures in modulated and unmodulated signals.

are illustrated in Figure 4 and are described as follows.

5.1 Datasets 1 and 2: Separate Signal Modalities

Datasets 1 (DS1) and 2 (DS2) represent a scenario where the modulated and unmodulated optical signals are analyzed separately. Each dataset contains four SOP event signatures. This separation enables a direct comparison of how signal modulation influences the discriminability of physical disturbances in the SOP signatures and the learnability of their spectral patterns by ML models. The research question we address with these datasets is: *Do the inherent differences in polarization dynamics between modulated and unmodulated signals lead to measurable variations in the performance of ML-based classification of SOP signatures?*

An identical pre-processing pipeline was applied separately to the modulated and unmodulated signal sets. As detailed in Section 3.2, each signature comprises 3,600 samples per event, with each sample corresponding to 0.5 ms of SOP variations and represented by 512 frequency-domain features derived from spectral analysis. This results in 14,400 samples for each signal type. The two datasets were independently shuffled and partitioned using an 80/20 split, yielding 11,520 training samples and 2,880 testing samples per set.

5.2 Dataset 3: Mixed Signal Modalities

Dataset 3 (DS3) extends the scope of the classification pipeline by combining modulated and unmodulated polarization signatures into a unified dataset. In this setup, the model is trained and assessed on a combined scenario comprising all eight collected SOP signatures, grouped into four distinct event classes, each defined by the union of its modulated and unmodulated instances.

Unlike DS1 and DS2, DS3 introduces mixed signal modalities during training and inference, i.e., equivalent classes from the modulated and unmodulated scenarios are merged into a single class. The goal is to evaluate whether a unified classifier can effectively learn class boundaries from both signal types, regardless of their underlying characteristics. This mixed modality setup reflects practical deployment scenarios where the signal format may vary or be unknown. The research question we aim to address by models trained with DS3 is: *Can a single ML model accurately classify SOP signatures when modulated and unmodulated signals are present during training and inference?* The results of a model trained with this dataset will provide insights into the performance of ML-driven SOP-based fiber sensing in heterogeneous signal environments.

To prepare DS3, each of the eight collected SOP signatures was first independently partitioned into training and testing subsets using an 80/20 split. This yielded 2,880 training samples and 720 testing samples per signature. After splitting, samples from modulated and unmodulated signals were merged within each event type to form four final event classes, each comprising 5,760 training samples and 1,440 testing samples. This results in a unified dataset containing 23,040 training samples and 5,760 testing samples.

5.3 Dataset 4: Joint Signal Modalities

Dataset 4 (DS4) contains eight classes, where each collected SOP signature is treated as a distinct class, explicitly distinguishing both the event type and its associated signal modality in the class. Unlike the DS3, which merges modulated and unmodulated variants of the same event into a single class, DS4 should capture fine-grained distinctions between signatures influenced by signal modality. This structure reflects a deployment scenario where both the nature of the event and the signal type are relevant and unknown thereby requiring classification. The central research question to be answered by train-

ing a model with DS4 is: *Can an ML model simultaneously distinguish both event type and signal modality when polarization signatures from both domains are presented as separate classes?* This approach will allow for a detailed investigation into the model's ability to distinguish event variations introduced by modulation effects.

To prepare DS4, the eight collected SOP signatures were partitioned individually using the same pre-processing pipeline of DS1 and DS2where each signature was partitioned into training and testing subsets using an 80/20 split, yielding 2,880 training samples and 720 testing samples per class. In total, the dataset contains 23,040 training and 5,760 testing samples.

6 Results

This section presents the evaluation results for four scenarios trained with the datasets introduced in Section 5. To identify the most suitable classification algorithm for each scenario, we conduct a comprehensive benchmarking of ten supervised ML classifiers available in the Scikit-learn library. The evaluated methods span a diverse range of learning paradigms, including ensemble learners: Random Forest (RF), Extra Trees (ET) Classifier, Histogram Gradient Boosting (HGB), eXtreme Gradient Boosting (XGBoost), Gradient Boosting (GB); kernel-based models: Support Vector Machine (SVM); linear classifiers: Logistic Regression (LR), Linear Discriminant Analysis (LDA); distance-based techniques: K-Nearest Neighbors (KNN); and tree-based learners: Decision Tree (DT). These classifiers are selected based on their previously documented performance in time-frequency domain tasks and their compatibility with the spectral representations extracted from the SOP data. For each dataset, the classifiers were trained and tested independently using 5-fold cross-validation to ensure robust performance assessment. The ML classifier yielding the highest test accuracy was selected as the representative classifier for that dataset, and its confusion matrix is reported in the corresponding subsection.

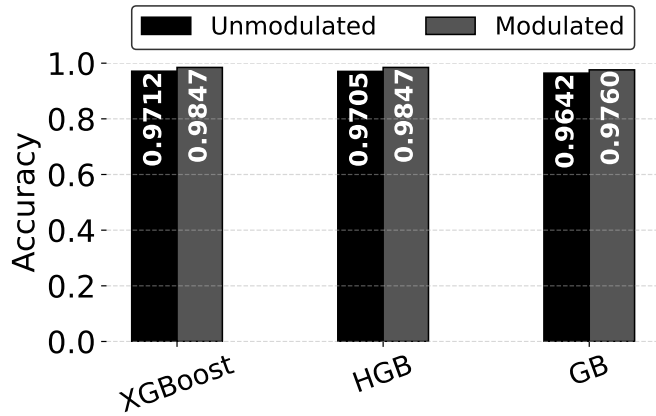
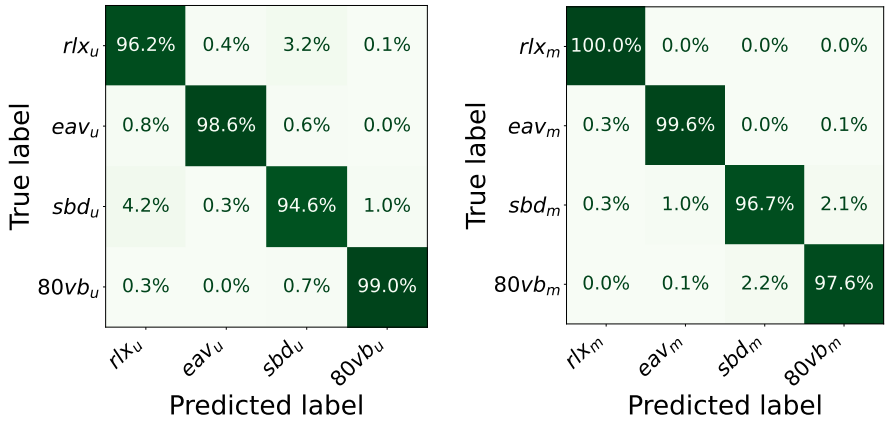


Figure 5: Classification accuracy on separated signal modalities scenario for the top three performing models: eXtreme Gradient Boosting (XGBoost), Histogram Gradient Boosting (HGB), and Gradient Boosting (GB).

6.1 Scenarios 1 and 2: Classification Performance for Separated Signal Modalities (DS1 and DS2)

Scenarios 1 and 2 evaluate the performance of ML-based SOP classification when polarization signatures are analyzed separately for modulated and unmodulated signals. Independent classifiers were trained using their default hyperparameters for DS1 and DS2. Figure 5 shows the classification accuracy in the test set for the three ML classifiers that achieve the highest accuracy across both DS1 and DS2: XGBoost, HGB, and GB. Among them, XGBoost demonstrates superior performance, attaining an accuracy of 97.12% for unmodulated and 98.47% for modulated signals. A deeper insight into the performance of XGBoost is offered by the confusion matrices in Figure 6. In both cases, the classifiers exhibit strong class separability with minimal confusion. Few misclassifications were observed primarily between the *sbd* and *80vb* events. Nonetheless, the impact of these errors on overall detection reliability is negligible.

These results indicate that signal modulation, such as the DP-16QAM 200 Gbps signal used in the experiment, does not hinder the ability of ML classifiers to distinguish between SOP signatures associated with different physical



(a) Unmodulated **(b)** Modulated

Figure 6: Confusion matrices of the eXtreme Gradient Boosting (XGBoost) classifier for (a) unmodulated and (b) modulated signals.

events. In fact, modulated signals in this study exhibited more stable SOP behavior than their unmodulated counterparts, contributing to consistently high classification performance. While modulated signals are often associated with higher structural complexity at the symbol level, their temporal averaging effect appears to suppress high-frequency polarization noise, leading to smoother SOP trajectories. Consequently, low-frequency SOP variations caused by external disturbances remain clearly distinguishable. These findings reinforce the feasibility of ML-based SOP analysis for anomaly detection in coherent optical networks, where modulation is an inherent aspect of system operation.

6.2 Scenario 3: Classification Performance for Event Classification in Mixed Signal Modalities (DS3)

Scenario 3 investigates the performance of ML classifiers in a unified classification task where both modulated and unmodulated signals are included for each physical event class. Table 2 summarizes the performance of the tested classifiers in terms of accuracy, precision, recall, F1-score, and computational

Table 2: Performance benchmarking of supervised ML classifiers for scenario 3 (mixed signal modalities classification).

Classifier	Accuracy	Precision	Recall	F1-score	Training Time (s)	Inference Time (s)
HGB	0.9809	0.9809	0.9809	0.9809	18.38	0.057
XGBoost	0.9793	0.9794	0.9793	0.9793	7.87	0.024
GB	0.9641	0.9640	0.9641	0.9640	1555.49	0.042
RF	0.9616	0.9617	0.9616	0.9615	40.46	0.066
ET	0.9498	0.9509	0.9498	0.9490	7.29	0.109
SVM	0.9479	0.9477	0.9479	0.9476	55.99	17.99
DT	0.9028	0.9030	0.9028	0.9028	18.89	0.005
KNN	0.8365	0.8417	0.8365	0.8298	0.023	0.655
LR	0.8193	0.8194	0.8193	0.8193	40.51	0.021
LDA	0.7814	0.7813	0.7814	0.7810	1.70	0.012

cost (training and inference times). The classifiers are sorted in descending order of accuracy. Compared to the other models, HGB and XGBoost achieved the two highest accuracy values of 98.09% and 97.93%, respectively. The confusion matrix of the best-performing HGB is shown in Figure 7. The model achieves excellent class separability for all four event types. The relaxed (*rlx*) and eavesdropping (*eav*) classes are recognized with near-perfect accuracy (98.8% and 99.0%, respectively). Minor confusion is observed between soft bending (*sbd*) and 80 Hz vibration (*80vb*), with at most 2.2% of misclassified samples.

These results demonstrate that ML models are able to learn discriminative features of both unmodulated and modulated signals. This supports the feasibility of deploying signal-agnostic SOP-based ML monitoring systems in network environments with unmodulated and modulated signals.

6.3 Scenario 4: Classification Performance for Joint Signal Modalities (DS4)

Scenario 4 evaluates the classification performance in a fine-grained, eight-class scenario, where each unique event-modality combination (e.g., rlx_u, rlx_m) is treated as a distinct class. Table 3 presents the performance of all evaluated classifiers for this setup. Among the tested models, XGBoost achieved

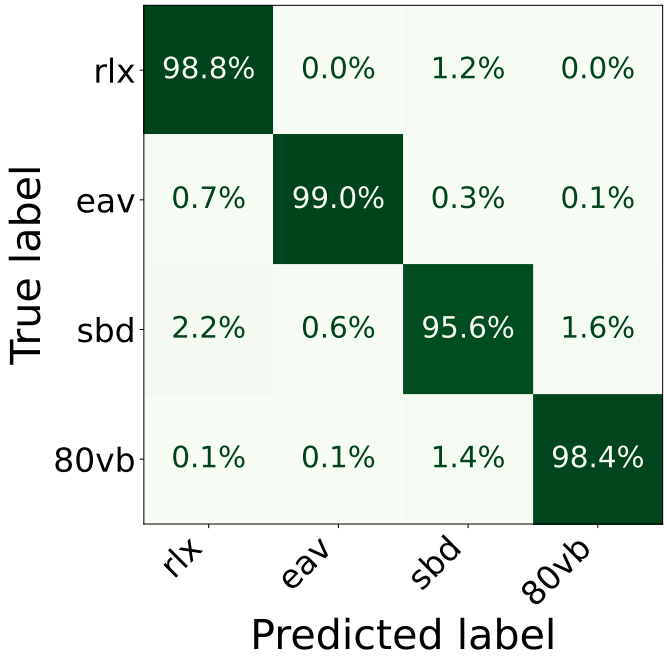


Figure 7: Confusion matrix of the Histogram Gradient Boosting (HGB) classifier for mixed signal modalities classification scenario.

the highest overall accuracy, followed closely by HGB, and GB. In addition to delivering the best classification performance, XGBoost maintained a relatively modest computational footprint, requiring 13.07 s for training and only 0.04 s for inference. The confusion matrix of XGBoost is shown in Figure 8.

XGBoost achieves high discrimination for nearly all classes. The modulated signature classes have near-perfect accuracy above 98.9%, with 100% accuracy for the rlx_m class. The worst-case misclassification in the unmodulated signals is observed for the sbd_u class, where 3.6% of samples are incorrectly labeled as rlx_u. In the modulated signal, the highest confusion occurs between sbd_m and $80vb_m$, with a misclassification rate of 2.4%. These errors indicate strong class separability even when both signal modality and event type are jointly classified. Interestingly, there is no confusion between unmodulated and modulated signals (100% accuracy), i.e., misclassification is only observed within

Table 3: Performance benchmarking of supervised ML classifiers for Scenario 4 (joint signal modalities classification).

Classifier	Accuracy	Precision	Recall	F1-score	Training Time (s)	Inference Time (s)
XGBoost	0.9804	0.9804	0.9804	0.9804	13.0668	0.0401
HGB	0.9793	0.9794	0.9793	0.9793	21.8151	0.0631
GB	0.9688	0.9687	0.9688	0.9687	2394.0043	0.0598
RF	0.9658	0.9661	0.9658	0.9658	39.8580	0.0678
SVM	0.9483	0.9483	0.9483	0.9481	24.5029	17.2565
ET	0.9467	0.9488	0.9467	0.9462	7.4505	0.1216
DTD	0.9057	0.9057	0.9057	0.9056	17.8165	0.0053
LDA	0.8807	0.8817	0.8807	0.8794	1.0988	0.0115
LR	0.8806	0.8807	0.8806	0.8805	56.5408	0.0043
KNN	0.8365	0.8458	0.8365	0.8300	0.0272	0.8149

the same modality. These results confirm that it is possible for a classifier to accurately classify both modality and physical events simultaneously. The findings reinforce the conclusion that modulation does not significantly affect the performance of ML-based classifiers.

7 Conclusions

This paper investigates the impact of optical signal modulation on ML classification of harmful events by using SOP signatures. While prior studies have largely focused on unmodulated light sources, our work addresses the critical gap of evaluating SOP-based monitoring under practical, high-speed modulated conditions. Using a real-world experimental setup on a metro network, we collected polarization signatures from modulated and unmodulated signals subjected to identical physical disturbances. Statistical analysis revealed that modulation stabilizes the SOP distribution by suppressing fluctuations. A suite of classifiers was assessed on four datasets with separated, mixed, and joint modalities. The assessed ML algorithms consistently achieved high classification accuracy in benign and malicious events, confirming that the spectral features extracted from SOP dynamics remain distinguishable by ML algorithms regardless of modulation effects.

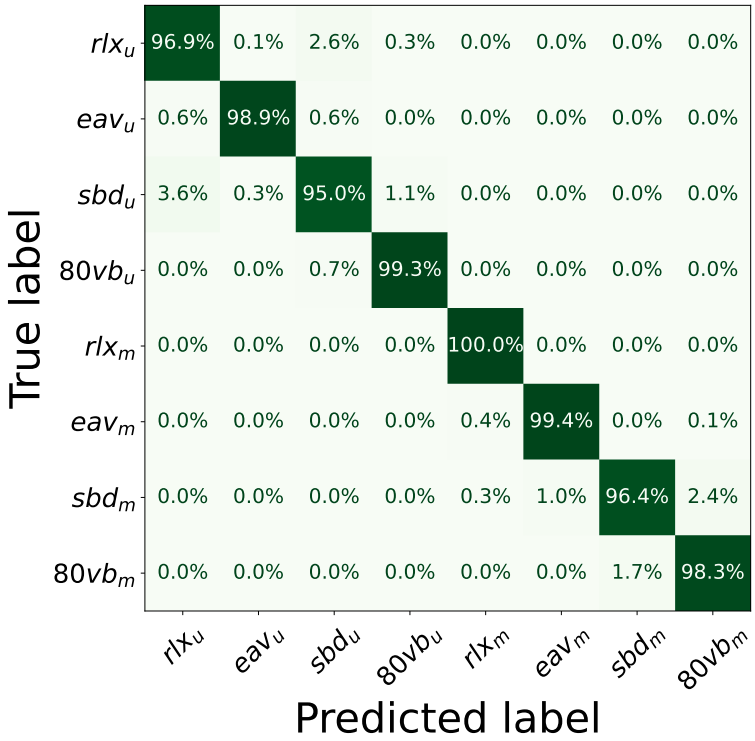


Figure 8: Confusion matrix for scenario 4 showing classification performance using the eXtreme Gradient Boosting (XGBoost) classifier.

References

[1] D. Dahan and U. Mahlab, "Security threats and protection procedures for optical networks," *IET Optoelectronics*, vol. 11, no. 5, pp. 186–200, 2017.

[2] W. Lee, S. I. Myong, J. C. Lee, and S. Lee, "Identification method of non-reflective faults based on index distribution of optical fibers," *Optics express*, vol. 22, no. 1, pp. 325–337, 2014.

[3] K. Abdelli, H. Grießer, C. Tropschug, and S. Pachnicke, "Optical fiber fault detection and localization in a noisy OTDR trace based on denois-

ing convolutional autoencoder and bidirectional long short-term memory," *IEEE Journal of Lightwave Technology*, vol. 40, no. 8, pp. 2254–2264, 2021.

[4] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," *Journal of optical communications and networking*, vol. 14, no. 5, pp. 365–375, 2022.

[5] B. Steinar, "Locating disturbances in optical fibres," WO2022185075A1, Sep. 2022.

[6] Y. Aono, E. Ip, and P. Ji, "More than communications: Environment monitoring using existing optical fiber network infrastructure," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, W3G.1.

[7] S. Pellegrini, L. Minelli, L. Andrenacci, G. Rizzelli, D. Pilori, G. Bosco, L. D. Chiesa, C. Crognale, S. Piciaccia, and R. Gaudino, "Overview on the state of polarization sensing: Application scenarios and anomaly detection algorithms," *J. Opt. Commun. Netw.*, vol. 17, no. 2, A196–A209, Feb. 2025.

[8] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, M1H.1.

[9] L. Andrenacci, D. Pilori, S. Pellegrini, L. Minelli, G. Bosco, C. Crognale, S. Piciaccia, and R. Gaudino, "Comparison between phase and polarization sensing using coherent transceivers over deployed metro fibers," in *Optical Fiber Communication Conference*, Optica Publishing Group, 2024, M2K–2.

[10] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti, "Detection of abnormal activities on a SM or MM fiber," in *Optical Fiber Communication Conference (OFC)*, 2023, M3Z.6.

[11] D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J.-P. Elbers, "Cognitive assurance architecture for optical network fault management," *Journal of Lightwave Technology*, vol. 36, no. 7, pp. 1443–1450, 2018.

[12] EXFO Inc, *Tunable DFB Laser Sources – Application Note 012*, Available here.

[13] B. Szafraniec, T. S. Marshall, and B. Nebendahl, "Performance monitoring and measurement techniques for coherent optical systems," *Journal of Lightwave Technology*, vol. 31, no. 4, pp. 648–663, 2013.

[14] L. Moeller and B. Bakhshi, *Polarization modulation of supervisory signals for reducing interference with data signals*, US Patent 10,230,472, Mar. 2019.

[15] ASIERA (formerly HEAnet), *Ireland's National Education and Research Network*, Available here.

[16] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Breaking boundaries: Harnessing unrelated image data for robust risky event classification with scarce state of polarization data," in *European Conference on Optical Communications (ECOC)*, IET, vol. 2023, 2023, pp. 924–927.

[17] K. Abdelli, M. Lonardi, F. Boitier, D. Correa, J. Gripp, S. Olsson, and P. Layec, "Vision transformers for anomaly classification and localization in optical networks using sop spectrograms," *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1902–1913, 2025.

[18] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment," in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[19] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, "Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events," in *2024 24th International Conference on Transparent Optical Networks (ICTON)*, 2024, pp. 1–5.

[20] W. Qin, Q. Zhang, W. Hou, X. Zhang, and X. Gong, "Convolutional neural networks for fiber-bending eavesdropping attacks detection in coherent optical communication systems," in *2024 International Conference on Ubiquitous Communication (Ucom)*, IEEE, 2024, pp. 342–345.

[21] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, "Unsupervised anomaly detection and localization with generative adversarial networks," *arXiv preprint arXiv:2409.03657*, 2024.

[22] X. Chen, B. Li, R. Proietti, Z. Zhu, and S. J. B. Yoo, "Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks," *Journal of Lightwave Technology*, vol. 37, no. 7, pp. 1742–1749, 2019.

[23] H. Song, R. Lin, L. Wosinska, P. Monti, M. Zhang, Y. Liang, Y. Li, and J. Zhang, "Cluster-based unsupervised method for eavesdropping detection and localization in WDM systems," *Journal of Optical Communications and Networking*, vol. 16, no. 10, F52–F61, 2024.

[24] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, "ML-based state of polarization analysis to detect emerging threats to optical fiber security," *TechRxiv*, 2025, DOI: 10.36227/techrxiv.175099843.35967234.v1.

[25] "Enhancing fiber security using a simple state of polarization analyzer and machine learning," *Optics Laser Technology*, vol. 167, p. 109 668, 2023, ISSN: 0030-3992.

[26] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Deep learning for detection of harmful events in real-world, noisy optical fiber deployments," *Journal of Lightwave Technology*, vol. 43, no. 13, pp. 6092–6101, 2025.

[27] C. J. Carver and X. Zhou, "Polarization sensing of network health and seismic activity over a live terrestrial fiber-optic cable," *Communications Engineering*, vol. 3, no. 1, p. 91, 2024.

[28] L. Sadighi, C. Natalino, S. Karlsson, L. Wosinska, M. Ruffini, and M. Furdek, "AI/ML-based state-of-polarization monitoring in optical networks: Concepts and challenges," in *Optical Fiber Communication Conference (OFC) 2025*, 2025, M3F.6.

[29] M. Mazur, D. Wallberg, L. Dallachiesa, E. Börjeson, R. Ryf, M. Bergroth, B. Josefsson, N. K. Fontaine, H. Chen, D. T. Neilson, J. Schröder, P. Larsson-Edefors, and M. Karlsson, "Real-time monitoring of cable break in a live network using a coherent transceiver prototype," in *2024 Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, pp. 1–3.

[30] A. Rode, M. Farsi, V. Lauinger, M. Karlsson, E. Agrell, L. Schmalen, and C. Häger, "Machine learning opportunities for integrated polarization sensing and communication in optical fibers," *Optical Fiber Technology*, vol. 90, p. 104 047, 2025, ISSN: 1068-5200.

[31] C. B. Czegledi, M. Karlsson, E. Agrell, and P. Johannisson, "Polarization drift channel model for coherent fibre-optic systems," *Scientific Reports*, vol. 6, no. 1, p. 21 217, 2016, DOI: 10.1038/srep21217.

[32] D. Waddy, P. Lu, L. Chen, and X. Bao, "Fast state of polarization changes in aerial fiber under different climatic conditions," *IEEE Photonics Technology Letters*, vol. 13, no. 9, pp. 1035–1037, 2001.

[33] J. Wuttke, P. Krummrich, and J. Rosch, "Polarization oscillations in aerial fiber caused by wind and power-line current," *IEEE Photonics Technology Letters*, vol. 15, no. 6, pp. 882–884, 2003.

[34] P. M. Krummrich, D. Ronnenberg, W. Schairer, D. Wienold, F. Jenau, and M. Herrmann, "Demanding response time requirements on coherent receivers due to fast polarization rotations caused by lightning events," *Opt. Express*, vol. 24, no. 11, pp. 12 442–12 457, May 2016.

[35] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, no. 2, pp. 804–817, Jan. 2008.

[36] K. Kikuchi, "Performance analyses of polarization demultiplexing based on constant-modulus algorithm in digital coherent optical receivers," *Opt. Express*, vol. 19, no. 10, pp. 9868–9880, May 2011.

[37] CONNECT Centre for Future Networks and Communications, *Open Ireland Testbed*, Available here.

[38] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, pp. 1–7, 2014.

[39] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, "Eavesdropping G. 652 vs. G. 657 fibres: A performance comparison," in *2022 International Conference on Optical Network Design and Modeling (ONDM)*, IEEE, 2022, pp. 1–3.