CHALMERS
UNIVERSITY OF TECHNOLOGY

Automated, Interpretable and Efficient ML Models for Real-World Lightpaths' Quality of Transmission Estimation

Citation for the original published paper (version of record):
Aladin, S., Wosinska, L., Tremblay, C. (2025). Automated, Interpretable and Efficient ML Models for Real-World Lightpaths' Quality of
Transmission Estimation. IEEE Open Journal of the Communications Society, 6: 9785-9801.
http://dx.doi.org/10.1109/OJCOMS.2025.3635533

Received 25 October 2025; accepted 16 November 2025. Date of publication 19 November 2025; date of current version 1 December 2025.

Digital Object Identifier 10.1109/OJCOMS.2025.3635533

Automated, Interpretable and Efficient ML Models for Real-World Lightpaths' Quality of Transmission Estimation

SANDRA ALADIN [1] (Member, IEEE), LENA WOSINSKA [2] (Senior Member, IEEE), AND CHRISTINE TREMBLAY [1] (Senior Member, IEEE)

[1]Network Technology Lab, Department of Electrical Engineering, École de technologie supérieure, Montréal H3C 1K3, QC, Canada
[2]Electrical Engineering Department, Chalmers University of Technology, 412 96 Gothenburg, Sweden

CORRESPONDING AUTHOR: C. TREMBLAY (e-mail: christine.tremblay@etsmtl.ca)

This work was supported in part by the National Sciences and Engineering Research Council of Canada under Grant RGPIN 2019-03972; in part by the ÉTS Program for the Development of International Research Collaborations; in part by European Commission's Digital Europe Programme (DIGITAL) Research and Innovation Program under Grant 101127973; and in part by the 5G-TACTIC 5G Trusted and Secure Network Services Project.

ABSTRACT Fast and accurate estimation of lightpaths' quality of transmission (QoT) is crucial for ensuring quality of service (QoS) and seamless operation in real-world optical networks. Machine learning (ML) algorithms are promising tools for QoT estimation of lightpaths before their establishment. In multi-domain optical networks, where learned QoT estimation models must be transferred between heterogeneous environments with limited target data, deep neural networks (DNNs) have shown promising results. However, DNN-based transfer learning (TL) approaches using fine-tuned artificial neural networks (ANNs) and convolutional neural networks (CNNs), offer limited interpretability. Consequently, little insight into the decision-making process is provided, and large labeled datasets as well as high computational resources are required, limiting their suitability for real-time, large-scale deployment in production networks. To address these challenges, we propose a novel lightweight and interpretable TL framework that integrates the Boruta-SHAP algorithm for automated feature selection (FS) together with two domain adaptation (DA) techniques: Feature Augmentation and Correlation Alignment. In contrast to the existing approaches based on DNN, our strategy leverages interpretable and efficient ML models to enhance the adaptability across diverse datasets in real-world network environments. We show that our random forest (RF)-based models achieve better performance than the ANN-based models, without sacrificing the classification accuracy. The FS via Boruta-SHAP allows for reducing dimensionality as well as training and inference times up to 70.68%, and 41.64%, respectively. Our proposed framework outperforms DA baseline models achieving 99.35% accuracy improvement in domain shift. Moreover, it offers 86% accuracy with a 99.83% reduction in the size of the target domain.

INDEX TERMS Artificial neural networks, deep learning, domain adaptation, explainable artificial intelligence, machine learning, optical fiber communication, performance metrics, quality of transmission, random forests, support vector machines, transfer learning, WDM networks.

I. INTRODUCTION

RELIABLE optical networks are crucial for the rapid development of emerging online services such as the Internet of Things, 5G/6G applications, and edge computing. These technologies require high bandwidth and low latency connectivity while demanding excellent transmission performance. Coherent optical networks with flexible modulation and tunable transceivers enable reliable high-speed and long-distance data transmission. However, these technologies entail complex network architectures and dynamic changes of parameters. Quality of transmission (QoT) estimation of lightpaths before they are established is an important task to guarantee the required connection reliability performance. Analytical solutions, based on

mathematical models, have been developed to estimate linear and nonlinear impairments (NLI) of lightpaths before their establishment. Two commonly used methods are the split step Fourier method (SSFM) and the Gaussian Noise (GN) model. The SSFM has been proven to accurately estimate the QoT but it suffers from the long computation time and high complexity. Authors in [1] developed an optical fiber simulator based on SSFM. The proposed software was tested using graphical processing units (GPU) computation on a standard desktop computer and a more robust server. Their simulations showed good accuracy and speed improvement in estimating the signal-to-NLI ratio on the server compared to a baseline result obtained by central processing unit (CPU) computation. Moreover, authors in [2] used the same SSFM to generate data set through simulation to test a predictive models proposed in their study. On the other hand, the GN model, although less accurate, offers faster computation than the SSMF method. In [3], the authors presented the validation of an open-source network simulation application that is called Gaussian noise simulation in Python (GNPy). The tool is based on the GN model and is used to generate QoT data in many research studies [4], [5], [6], [7], [8], [9]. Due to its favorable trade-off between computational efficiency and accuracy, the GN model is more commonly employed than the SSFM for analytical QoT estimation of unestablished lightpaths. However, the accuracy of analytical QoT estimation methods is fundamentally constrained by the quality of input parameters–such as span loss, dispersion, and amplifier noise figures–whose precision and availability are often limited in real-world network environments.

Recently, QoT estimation based on machine learning (ML) gained high attention. In ML-based approaches, algorithms learn the complex mapping between input features – such as candidate lightpath characteristics and network configuration parameters - and the corresponding QoT outcomes. Thus, estimating the QoT of unestablished lightpaths involves leveraging these features to predict the feasibility and expected performance of the connection. By comparing the predicted QoT against the receiver sensitivity threshold, the viability of a potential lightpath can be assessed, which is an essential step for efficient service provisioning. An underestimated QoT may result in underutilized resources, while overestimating it can lead to connection failures and violations of service level agreements (SLAs). ML classifiers based on K-nearest neighbors (KNN), random forest (RF), and support vector machine (SVM) were proposed in [10], [11], [12]. Additionally, a QoT estimation model using artificial neural network (ANN) was introduced in [13]. All these models were trained with synthetic Bit-Error-Rate (BER) data generated by using the additive Gaussian white noise (AGWN) model. While the results demonstrated the potential of ML techniques for efficient and automated lightpath provisioning, these studies relied exclusively on proprietary QoT datasets generated by individual research groups. Consequently, the challenges related to reproducibility and independent validation of results persist.

In 2022, the Fraunhofer Heinrich Hertz Institute (HHI) has made available a collection of four QoT datasets based on the Core Optical Network of the United States (CONUS) and Telefónica Spain National Network (TSNN) topologies, intended for evaluating QoT estimation models [14]. The use of the publicly accessible QoT datasets provides significant benefits, enabling researchers to validate their findings, benchmark models against existing methods, and compare performance effectively. So far different approaches using open QoT datasets to estimate the QoT for unestablished lightpaths have been explored. However, most proposed solutions rely on complex ML techniques that primarily focus on improving predictive performance, often overlooking critical aspects such as computation time (during both training and inference) and model interpretability. While interpretability is essential for building trust in model's decisions, inference time is directly linked to decision-making speed and can significantly impact lightpath provisioning in real-world applications. For network operators, short training times are crucial for adapting quickly to changing network conditions in dynamic, high-capacity optical networks. In addition, during a real-time provisioning, low-latency inference is essential to ensure SLA compliance and optimal resource utilization. Therefore, minimizing ML model training and inference times not only enhances operational efficiency but also supports the deployment of cognitive management systems in real-time optical network environments. Additional strategies such as transfer learning (TL), active learning, and meta-learning have been proposed to enable practical deployment of ML-based QoT estimation models in real-world networks [15]. These approaches address the scarcity of field QoT datasets, as effective ML-based QoT estimation models require both high-quality and large volume of data. Active learning optimizes the selection of informative samples, meta-learning extracts transferable knowledge across tasks, and TL leverages prior knowledge from related domains to improve learning efficiency and generalization in new scenarios. Many TL solutions are based on neural networks pre-trained on source domain data and fine-tuned on target domain data [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26]. While these models can achieve high accuracy, their deployment in production networks is hindered by computational complexity, large labeled data requirements [20], and limited generalizability due to heterogeneous optical system configurations, as highlighted in [7]. As a result, models often require full retraining for each new system, leading to significant time and resource overhead.

In this work, we propose a novel TL framework that includes a feature selection stage before applying DA techniques with lightweight classifiers, improving model efficiency and generalizability across different optical network scenarios. We adopt an incremental approach. In the first step, we select the best-performing low-complexity classifier through a comparative performance analysis of three ML techniques – KNN, SVM, and RF - with a more advanced

deep learning ANN-based approach. In the second step, we propose applying a feature selection strategy using the recursive feature elimination with cross-validation (RFECV) and Boruta-SHAP where SHAP (SHapley. Additive exPlanations) is a game-theoretic approach used for feature importance. A comparative evaluation of model accuracy and computation times across the four datasets enables finding the most effective feature selection method. The feature selection methods automatically identify the most relevant features for the RF-based solution for each dataset. In the last step, leveraging the RF classifiers, we implement a transfer learning strategy that combines domain adaptation (DA) techniques with feature space derived from the automated feature selection process in the preceding phase.

This paper is organized as follows: Section II summarizes the related works on lightpath QoT estimation. Section III describes the methodology, as well as the dataset and proposed framework for automated feature selection and domain adaptation guided by reduced features. Section IV details the results with a comparative analysis of the different models as well as the performance achieved by our proposed domain adaptation guided by feature selection framework in an inter-network scenario. Finally, Section V offers some conclusive remarks as well as future research directions.

II. RELATED WORKS

In this section, we review prior research on ML-based lightpath QoT estimation, starting with general regression and classification approaches, followed by studies using the Fraunhofer QoT datasets. We then examine efforts to enhance model trustworthiness through explainable artificial intelligence (XAI) and to improve adaptability for addressing data scarcity challenges in real-world deployment scenarios.

ML applications in management of optical networks have gained interest in recent years. Leveraging recent advancements in technologies and data accessibility, ML uses algorithms to enable cognition in optical networks. Four main categories of ML techniques are used in the field, i.e., supervised learning, semi-supervised learning, reinforcement learning and unsupervised learning. In supervised learning models are trained using labeled data to map input parameters to correct outputs. It includes classification and regression tasks. Semi-supervised learning trains models using a combination of labeled and unlabeled data. Reinforcement learning requires an agent to learn through trial-and-error interactions within an environment and make decisions, which can lead to rewards or penalties. Finally, unsupervised learning allows to find hidden patterns by using clustering and dimensionality reduction.

ML-based lightpath QoT estimation can be achieved through a regression or a classification strategy. Supervised learning is the most used approach to predict categorical or continuous QoT indicators. Reference [27] presents an overview of studies on the topic categorized into three main groups: traditional regression methods, classification approaches and deep learning-based regression architectures such as recurrent neural networks (RNN)s and ANNs. Numerous studies have proposed both regression [18], [28], [29], [30], [31] and classification approaches [32], [33], [34], [35], [36] for estimating the QoT of unestablished lightpaths. However, as previously noted, these solutions were developed using proprietary datasets, limiting the reproducibility and independent validation of their results.

Most research on ML-based QoT estimation has focused primarily on the predictive performance of the models. In [37], the authors employed a regression approach to predict the signal-to-noise ratio (SNR) of synthesized lightpaths, using data generated via the enhanced Gaussian Nonlinear (EGN) model. This method is similar to the one proposed in this paper. However, their solution suffers from the comparability issue discussed in [14], while it is addressed in the solution presented in this paper.

Moreover, researchers have often adopted complex ML techniques such as deep learning, which offers improved performance but come at the cost of interpretability and transparency in the model's decision-making process prior to deployment. In [38], the authors addressed interpretability and transparency challenges by proposing an Extreme Gradient Boosting (XGBoost) model for lightpath QoT estimation which incorporates XAI techniques. They used SHAP as a framework for feature selection to identify a subset of features sufficient for accurate QoT estimation.

Recognizing the benefits of reproducibility, validation, and fair comparison, many researchers have leveraged the publicly available Fraunhofer QoT datasets to propose a variety of solutions for lightpath QoT estimation. Studies such as [14], [39], [40], [41] introduced complex models based on ANN, deep convolution neural networks (DCNN), vertical federated learning (VFL) and secure multiparty computation (SMPC) techniques. In contrast, works like [43], [44] proposed an AI/ML-as-a-service framework for optical network automation, utilizing the Automated ML (AutoML) libraries that favor simpler models such as RF. However, RF was not selected as part of the final ensemble applied in their framework. The evaluation of the framework's inference performance revealed persistent challenges, particularly regarding inference speed, explainability and model trustworthiness.

Although the authors in [38] also utilized the Fraunhofer datasets to propose an XGBoost-based solution for QoT estimation – incorporating explainable XAI techniques – they employed a heuristic-based feature selection approach and did not address computational efficiency in their proposed solution. Furthermore, XGBoost is generally considered more complex than RF, due to its greater number of tuning parameters, more complex architecture, higher computational cost, and lower interpretability. While RF trains decision trees in parallel and aggregates predictions through averaging, XGBoost builds tree training sequentially using gradient boosting and regularization step, which - although potentially more accurate - can hinder optimization and interpretability.

In production networks, adaptability - alongside interpretability and computational efficiency - is crucial for overcoming the scarcity of labeled QoT data. Many deep learning-based TL solutions have been proposed to adapt models from source domains with abundant labeled data to target networks with limited datasets. Although these approaches can achieve high prediction accuracy, their substantial computational complexity often limits their practical deployment in real-time or resource-constrained environments. Reference [20] proposed a hybrid framework combining knowledge distillation (KD) with TL to transfer knowledge from a large, complex model to a smaller more efficient one, thereby reducing computational complexity, storage overhead, and dependence on large-scale training datasets. In [45], a TL approach leveraging three standard ML algorithms, i.e., SVM, RF, and logistic regression (LR), was evaluated using two DA methods for transferring knowledge between two network domains, with results showing that DA outperformed simple mixing of source and target data. In [24], a comparative study of DA and active learning with Gaussian Processes (GPs) showed that active learning achieved similar QoT estimation performance improvements while requiring fewer labeled samples than the domain adaptation, highlighting its effectiveness in data-scarce scenarios.

In TL, DA facilitates knowledge transfer across network domains, enabling models trained on large source datasets to perform effectively in target domains with limited data. This capability is vital for the practical deployment of QoT estimation solutions in real-world optical networks where labeled data are scarce. Our approach combines automated feature selection with DA to enhance the generalization of lightweight, interpretable models.

In this work, we address three key aspects of ML-based lightpath QoT estimation with the goal of facilitating real-world deployment by focusing on model complexity, performance, benchmarking, interpretability as well as adaptability. First, we present RF as a meaningful, interpretable, and computationally efficient alternative for lightpath QoT estimation. Rather than introducing yet another baseline, our focus is on simpler ML models that have been relatively overlooked in the prior works, especially in comparison with more complex methods such as ANNs and XGBoost [14], [38]. Although RF is a well-established ML technique, our contribution lies in demonstrating its effectiveness and practicality in the QoT estimation context, highlighting how its architectural simplicity translates into both operational efficiency and performance advantages. This contribution is significant as it provides practical insights into the operational feasibility and model selection criteria for QoT estimation tasks in real-world optical networks. Second, in contrast to the prior studies, we employ a fully automated and interpretable feature selection methodology, which enhances model transparency, reduces manual intervention in the modeling process, and supports scalable deployment across diverse network environments. Our implementation of RFECV and Boruta-SHAP represents a novel advancement over the heuristic-based method used in [38], providing a replicable framework for other researchers working with publicly available datasets. Third, we introduce a feature-based DA approach that enhances alignment between source and target domains by leveraging automated feature selection to identify domain-invariant and relevant features prior to applying TL. This contrasts with the source domain baseline (SDB) approach described in [24], [45], [46] and proposed in [14], to address the limited availability of datasets related to topology, transceivers, and network status. Additionally, our proposed solution enables more stable transfer between domains by only sharing relevant features while applying the DA techniques.

To support multi-domain infrastructure deployments, network operators need fast, scalable and accurate decision-making for real-time lightpath provisioning, as well as model interpretability for easier troubleshooting and refinement. Keeping in mind the above mentioned preferences, in this paper we demonstrate the advantages of simpler models in terms of performance, interpretability, and adaptability across varied optical network architectures. Hence, the main motivation and impact of our contribution is the relevance for network operators, which has not yet been sufficiently addressed by other existing works.

III. METHODOLOGY

This section outlines the methodology, beginning with a description of the dataset, model selection, and comparative assessment strategy. We then detail the automated feature selection methods and DA techniques applied in this paper. Finally, we present the procedures for data preprocessing, feature selection, and feature-guided DA.

Fig. 1 shows the overall process of our proposed QoT estimation approach with subprocesses such as data preprocessing, model selection, automated feature selection including the visualization of how the important features impact the classification of lightpath instances. The figure also illustrates the implementation of DA guided by the Boruta-SHAP-based feature selection, and the performance analysis of our proposed solution in inter-networks context.

A. DATASET DESCRIPTION

This study is conducted using the QoT datasets that are made available to the research community in [14]. The synthetic datasets are generated using the Fraunhofer Institute's planning tool for optical networks (PLATON). More information about the configuration and procedures to generate the data are available in [14]. A QoT estimation (QOTE) module uses the Gaussian-Noise-based nonlinear channel model proposed by [47] to calculate the Optical Signal-to-Noise (OSNR) on each link which helps estimating the QoT of lightpaths. The QOTE module estimates the QoT of new lightpaths through a provisioning procedure triggered by a traffic request. Based on the estimated QoT values, the new lightpath can be either established or blocked. QoT metrics of established and

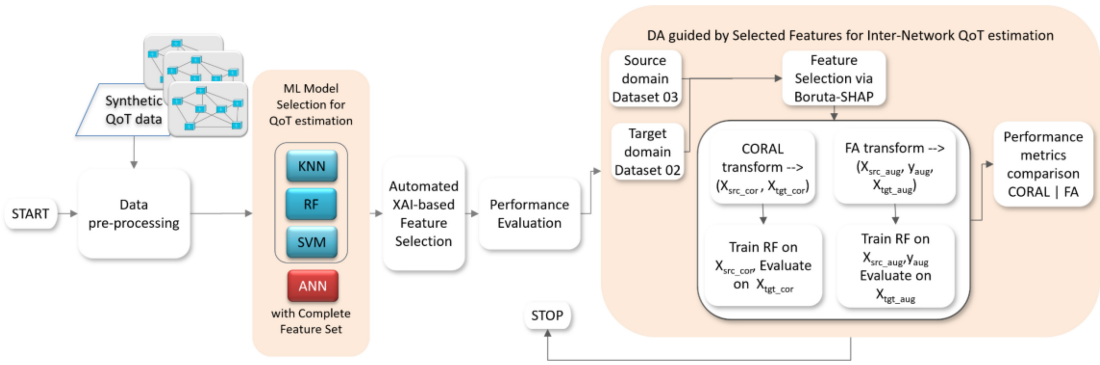

FIGURE 1. Overall process to implement automated, interpretable and adaptable lightpath QoT estimation models. X_{src_cor} and X_{tgt_cor} are source and target subsamples transformed with CORAL; X_{src_aug}, y_{aug}, and X_{tgt_aug} are source subsamples with labels and target subsamples transformed by feature augmentation.

rejected connection requests are documented in PLATON along with numerous statistics in its Traffic Engineering Database (TED).

The simulation scenarios behind the gathered QoT data are based on elastic optical networks using the C-band. With a fixed-grid flex-rate network scheme and a fixed channel spacing of 37.5 GHz, the configuration allows for 96 channels carrying various data rates. Nominal central frequencies range from 192.2 to 195.7625 THz with the frequency grid anchored to 194 THz [14]. Possible modulation formats include polarization-multiplexed binary phase-shift keying (PM-BPSK), polarization-multiplexed quadrature phase-shift keying (PM-QPSK), and higher-order quadrature amplitude modulation (QAM) schemes such as PM-8QAM, PM-16QAM, PM-32QAM, and PM-64QAM. Physical link parameters are set as follows: the EDFA noise figure to 5 dB, the amplifier gain to 16 dB, the fiber attenuation coefficient to 0.2 dB/km, the fiber dispersion to 17 ps/nm/km, the fiber nonlinear coefficient to 1.3 1/W/km, the noise bandwidth to 12.5 GHz, and the reference wavelength to 1550 nm. The symbol rate is set to 28 GBd and the channel launch power to –3 dBm. All links consist of one or more 80-km spans of ITU-T G.652 SMF. Erbium-Doped Fiber Amplifiers (EDFAs) are used to compensate for span loss. A Hard Decision Forward Error Correction (HD-FEC) is implemented for error correction with a required minimum pre-FEC BER of 3.8×10^{-3}. More details about the simulations to build the four datasets can be found in [14], included in the documentation provided with each dataset. With almost 1.2 million samples each, these datasets provide a wide range of simulated networking scenarios. Each dataset contains QoT metrics such as the OSNR, SNR, and BER as well as a binary class label indicating QoT compliance. Although the GN model used in PLATON includes neither inter-channel stimulated Raman scattering (ISRS) nor modulation format correction, these datasets include consistent structure across multiple simulated network scenarios and offer rich feature representation with topology-related, network-related and transceiver-related characteristics to support both network-wide and lightpath-based QoT estimation tasks, enabling diverse ML applications.

The lightpath dataset comprises 35 scalar features with different numbers of instances. Each subset includes 31 lightpaths' characteristics with 4 simulation-related metadata attributes. Four target variables are associated to each instance: a binary class label based on a predefined BER threshold ($BER_{th} = 0.0038$), the BER, the OSNR and the SNR of the lightpath. Binary class labels can be positive ($y = 1$) or negative ($y = 0$). More information about the features and labels can be found in [14].

For each dataset, we adopt the same strategy as in [14] to have class-balanced subsets of 100,000 instances. We keep the same names for Datasets 01, 02, and 04 which derive from the CONUS optical network as well as Dataset 03 which is based on the TSNN. Table 1 shows a complete list of the features used as the input to our models.

B. MODEL SELECTION AND COMPARATIVE ASSESSMENT STRATEGY

As mentioned in the Introduction, we first implement twelve models leveraging the four subsets of dataset 01, 02, 03, and 04 using the three classical ML techniques KNN, SVM and RF. After comparing their classification accuracies, the RF models outperform the two other models for dataset 01, 02, and 04, while the SVM model performed better with the dataset 03.

Table 2 shows a summary of the results obtained for each scenario.

To ensure a fair comparative evaluation with the first formulation of Occam's Razor principle as stated in [48], we compare models based on both classification performance and computational efficiency within a consistent execution environment. As in [14], we implement ANN models for comparison with the RF classifiers. We use the *SciKeras* library version 0.13.0 with the module *scikeras.wrappers* to develop the classifiers. In the hyperparameter optimization step, we consider 100 epochs with 128/256 hidden units, the learning rate in {0.001, 0.01}, 0.1/0.2 for the model dropout rate and the batch size in {50, 100}. The rectified linear unit (ReLU) and the hyperbolic tangent (tanh) are the two activation functions tested for the hidden layer. For the RF models, we use the *scikit-learn* library version

TABLE 1. Description of features representing lightpath instances in the datasets with their measurement units.

Feature	Description
Path Len (km)	Path length
Avg Link Len (km)	Average link length
Min Link Len (km)	Minimum link length
Max Link Len (km)	Maximum link length
Num Links	Number of links
Num Spans	Number of spans
Freq (THz)	Central frequency
Mod Order	Cardinality of the modulation format
LP Linerate (Gb/s)	Lightpath line rate
Source Node Degree	Source node degree
Destination Node Degree	Destination node degree
Min Link Occ	Minimum link occupation
Max Link Occ	Maximum link occupation
Avg Link Occ	Average link occupation
Std Link Occ	Standard deviation of the link occupation
Sum Link Occ	Sum of the spectral occupation of the links the lightpath traverses
Max BER	Maximum BER of interfering lightpaths
Min BER	Minimum BER of interfering lightpaths
Avg BER	Average BER of interfering lightpaths
Min Mod Order (L)	Minimum cardinality of the modulation format (left)
Max Mod Order (L)	Maximum cardinality of the modulation format (left)
Min Mod Order (R)	Minimum cardinality of the modulation format (right)
Max Mod Order (R)	Maximum cardinality of the modulation format (right)
Min LP Linerate (L) (Gb/s)	Minimum lightpath line rate (left)
Max LP Linerate (L) (Gb/s)	Maximum lightpath line rate (left)
Min LP Linerate (R) (Gb/s)	Minimum lightpath line rate (right)
Max LP Linerate (R) (Gb/s)	Maximum lightpath line rate (right)
Min BER (L)	Minimum BER (left)
Max BER (L)	Maximum BER (left)
Min BER (R)	Minimum BER (right)
Max BER (R)	Maximum BER (right)

TABLE 2. Classification accuracy (%) .

Dataset	*KNN Train/Test*	*RF Train/Test*	*SVM Train/Test*
01	88.10/84.60	99.90/98.50	98.20/97.80
02	97.90/97.20	99.90/99.60	99.60/99.50
03	97.30/96.70	99.90/99.20	99.30/99.30
04	95.70/94.80	99.80/99.50	99.50/99.30

1.7.1. In the hyperparameter tuning for the RF classifiers we take two values for number of estimators {50, 75}, as well as for the maximum depth {27, 30} and three values for the maximum leaf nodes {66, 72, 80}. We keep the default values for the other parameters. We consider a 5-fold cross-validation using *GridSearchCV* from the *scikit-learn* library for both techniques. The performance results in terms of classification accuracy and computation time for

TABLE 3. Performance comparison of RF and ANN models.

Dataset	*ANN*		*RF*	
	Time (s) Train/Test	*Accuracy (%) Train/Test*	*Time (s) Train/Test*	*Accuracy (%) Train/Test*
01	884.76/1.45	98.45/98.10	85.86/0.21	99.12/98.59
02	881.27/1.43	99.57/99.60	26.65/0.12	99.79/99.64
03	887.61/1.27	99.18/99.34	63.30/0.20	99.41/99.24
04	872.62/1.34	99.45/99.45	35.40/0.14	99.75/99.50

the new ANN and RF models are shown in Table 3. Only the classification performances are available for the models presented in [14]. Our proposed RF models outperform both their and our ANN models for dataset 01, 02 and 04. However, our ANN model for dataset 03 performs better than both the ANN proposed in [14] and our RF models (with classification accuracy of 99.34%, 99.03% and 99.24% respectively).

Moreover, as expected, the RF models consistently have faster computation times than the ANN models in all scenarios. The computation time presented in Table 3 for each scenario is achieved on a workstation using an Intel Core i5-8250U 1.6 GHz CPU and 16 GB RAM. These findings reveal that RF-based models significantly reduce training time (by 10 to 33 times) and inference time (by 6 to 11 times) while maintaining classification accuracy.

We also perform a comparison of our RF-based models with the XGBoost models. To validate the replication of the solution proposed in [38], we use a similar approach with the whole set of features for Datasets 03 and 04, including the simulation-related meta-data attributes, and without any pre-processing of the class-imbalanced instances in the two datasets. The replicated XGBoost models perform similarly to [38] with accuracies of 99.60% and 99.84%, compared to 99.70% and 99.80% for Datasets 03 and 04, respectively. Moreover, to allow a fair comparison between our simpler RF-based solution and the proposed XGBoost-based models, we use our class-balanced sub-samples with the selected features from [38]. The RF-based models achieve almost similar accuracies with differences of 0.01% and 0.02% for Datasets 03 and 04, respectively, making it a formidable choice to assess how a simple approach compares to a complex one in a context of lightpath's QoT estimation, leveraging the available QoT datasets proposed in [14].

C. AUTOMATED FEATURE SELECTION METHODS AND DOMAIN ADAPTATION TECHNIQUES

In [38], the authors applied the recursive feature elimination (RFE) method to select important features from the datasets used to build their classifiers. But this method is only recommended when prior knowledge of the optimal number of features is available. We implement an automated approach using the following two methods.

RFECV is a feature selection method that automatically determines the most relevant features for a specific model.

An RFE selector is used on different cross-validation splits. The number of selected features is set to the number of features that maximize the cross-validation score, obtained by evaluating the performance for different numbers of selected features and averaged across folds. This algorithm uses a supervised learning estimator that returns information about feature importance.

Boruta-SHAP is proposed here in order to leverage the robustness of the Boruta method and the interpretability of SHAP values. Boruta is a tree-based method working for tree models and other classification models that output feature importance measure. The algorithm shuffles original features to create shadow attributes at each iteration. The algorithm iteratively compares the importance of the features with the importance of the shadow attributes. Features are consecutively dropped or admitted when their importance compared to the shadow ones is worse or better, respectively. The algorithm stops either when all features are covered, or it reaches a specified limit. SHAP values help to determine feature importance based on principles of cooperative game theory. They allow a better understanding of the model's decisions by measuring the contribution of each feature to the classifiers' performance.

In [37], the authors used another automated version of the techniques, namely BoostRFE, and BoostBoruta, applying SHAP values, but the solutions were designed for gradient boosting models such as light gradient boosting (LGBM) and XGBoost. With our RF-based models, we explore the two automated methods for the feature selection, i.e., RFECV and Boruta-SHAP, which both are compatible with RF.

The fully automated feature selection process retains only statistically significant features, eliminating the need to manually specify their number. This brings a notable improvement over the manual approaches used in the previous studies, e.g., in [38].

We evaluate two DA methods as in [24], [45], [46].

Feature Augmentation (FA) encodes the domain of a sample by augmenting its feature vector. In particular, the length of the original feature vector $\mathbf{x}$ is tripled with a rule that depends on the domain. This way, both commonalities between the two domains and the unique characteristics of each domain are captured. While for sample coming from the source domain S, the triple feature vector is defined as $x' = \langle x,\ x,\ 0\rangle$, for sample coming from the target domain T, the triple feature vector is defined as $x' = \langle x,\ 0,\ x\rangle$. It is a supervised DA technique.

Correlation Alignment (CORAL) transforms the features in the source domain to match the second-order statistics of the features in the target domain. This allows minimizing the domain shift. A transformation is applied that re-colors the whitened features of the samples in S with the covariance computed from the feature distribution of the samples in T. The transformed data is used to train the model. This method does not require any information about the labels of samples in T; thus, it is considered an unsupervised DA technique.

In this study, we implement the source domain baseline (SDB) where the model is trained only with source domain data, for comparison with the two DA techniques. The other baseline scenarios such as target domain baseline (TDB) or dataset mixing baseline (DMB) are not considered in our assessment.

To improve the model's ability to find patterns, FA adds domain-specific information, such as statistical summaries or domain markers, to the original feature collection. CORAL applies a linear adjustment to align the second-order statistics (covariance) of the source and target feature distributions, minimizing domain shift [45].

D. DATA PREPROCESSING, AUTOMATED FEATURE SELECTION AND THE FEATURE-SELECTION-BASED DOMAIN ADAPTATION

The four gathered class-balanced subsets of data used in this work include 50,000 instances of each class, forming four sub-samples of 100,000 instances as in [14]. The *StandardScaler* class from the *sklearn* package in Python is used to remove the mean and scale the data to unit variance as a way of standardizing the features. It is worth noting that while data scaling benefits techniques such as KNN, SVM, and ANN, it is intentionally not applied to the RF technique in order to preserve the interpretability of feature importance. Data scaling can improve gradient convergence in ANN, facilitate the construction of decision boundaries in SVM, and enhance distance calculations in KNN for more accurate neighbor selection. In contrast, RF does not rely on distance-based or gradient-based computations and therefore does not benefit from feature scaling. Furthermore, scaling the data can obscure the real-world meaning of feature values, making SHAP-based explanations less intuitive.

1) Automated Feature Selection: Reducing the number of features when building ML models simplifies the resulting models. This can lead to faster training and prediction time as well as easier interpretation and reduced risk of overfitting. Two methods, compatible with RF, are explored for an automated feature selection to have a better insight to how the parameters affect the lightpaths' QoT. We applied the RFECV and Boruta-SHAP methods to the four datasets and built RF-based classifiers using the reduced data sets obtained with the two methods.

After the model selection process, we perform the automated feature selection process. We start with the RFECV method, where we apply the scikit-learn ML library in Python, using the *RFECV* class. We pass our previously created RF-based estimator into the RFE with cross-validation instance to find the optimal number of features for each of the four class-balanced sub-samples. We keep the number of folds (cv parameter) to 5, its default value. The most relevant features are obtained by fitting the RFECV model. We then transform the training and test data using the RFECV model with the '*transform()*' method. Next, we train the final model using the transformed dataset with selected features to perform final classification of the

TABLE 4. Performance comparison of RFECV and Boruta-SHAP.

Dataset	*Boruta-SHAP*		*RFECV*	
	Time (s) Training	*Accuracy (%)*	*Time (s) Training*	*Accuracy (%)*
01	17.98	98.55	7.14	98.59
02	8.89	99.68	10.3	99.67
03	14.91	99.25	7.82	99.29
04	10.01	99.56	15.62	99.55

instances in the feature-reduced test dataset. For the Boruta-SHAP implementation, we install the package and import the method. Similarly to the RFECV implementation, we pass our RF-based estimator into the BorutaSHAP instance. We set the importance measure parameter to '*shap*', and classification to '*True*'. We transform the training and test datasets in this case by dropping the less relevant features: we use the attribute 'features_to_remove' from the *BorutaSHAP* class which tracks the rejected features during the selection process. And for the final steps, as with the RFECV method, we train the model with the selected features then we predict the class for the transformed test dataset. These final steps allow us to compare the two strategies and establish our choice. While RFECV is inherently more efficient when used with large datasets, our findings show that simpler models built using class-balanced sub-samples can achieve better overall classification accuracy and reduced computation time. This makes them a more practical option for real-world applications where both efficiency and performance are critical.

To interpret the ML models built with reduced feature sets, we install the SHAP framework and use its explainer interface. The interface supports both local and global explanations: local explanations allow to visualize and interpret the model's decision for a specific instance, while global explanations reveal the contribution of each feature to the models' predictions. In our study, we use 'beeswarm' for global explanations and 'forceplot' plots for local explanations.

2) Selected-Feature-Based DA: We propose a DA framework based on the Boruta-SHAP feature selection performed in previous step. We integrate the Boruta-SHAP-based feature selection with the DA transformation into an efficient, explainable, and adaptable RF-based solution, addressing computational complexity and opacity in DL-based models while lowering the requirement for large amounts of training data in the development of supervised learning models. The pseudo-code of our approach is shown in Algorithm 1. We determine the source and target domains based on the worst-case scenario for domain adaptation inter-networks, namely with Dataset 03 from TSNN as the source domain and the Dataset 02 from CONUS as the target domain. Then we assess the two methods FA and CORAL, commonly used for transfer learning in QoT estimation to test our selected-feature-based models' generalization between different networks. In this case, we consider the scenario where the source domain is gathered from the small network TSNN, and the target domain from the more comprehensive network CONUS. We pick three different target sample sizes of 50, 100, and 1000 instances. After the first transformation performed using the Boruta-SHAP-based feature selection, we perform the second transformation using the CORAL method, then we fit the RF-based model using the transformed data. We predict the class for new data instances from the target domain. For the FA method, the outputs from the first transformation are used along with the labels to build the augmented feature vectors. The RF-based model is then fit using the new augmented data, and the resulting model is used to predict the class for unseen instances from the target dataset. Moreover, we implement the SDB scenario, where only source domain data is used to build the model, with the reduced feature set. While the feature sets were arbitrarily chosen to build the models in [45], our solution leverages a more comprehensive feature set, where the automated feature selection Boruta-SHAP allows transferring only relevant attributes.

We test our models using different sizes of target domain datasets and we present different performance parameters, along with the classification accuracy, to have better insight on the models' decision.

IV. NUMERICAL RESULTS AND COMPARATIVE ANALYSIS

This section highlights the performance of automated feature selection and evaluates the effectiveness of the proposed feature-guided DA framework in an inter-network scenario.

A. BORUTA-SHAP-BASED FEATURE SELECTION

The two aforementioned feature selection methods, i.e., Boruta-SHAP and RFECV, are used to select relevant features from the four datasets. In the context of lightpath QoT estimation, both the classification accuracy and the training time are good parameters to evaluate models' performance.

As shown in Table 4 the Boruta-SHAP method achieves better performance than the RFECV method with slightly higher classification accuracy and faster training time for Datasets 02 (99.68% and 8.89 seconds against 99.67% and 10.3 seconds) and 04 (99.56% and 10.01 seconds against 99.55% and 15.62 seconds), while the RFECV method performs better for Datasets 01 (98.59% and 7.14 seconds against 98.55% and 17.98 seconds) and 03 (99.29% and 7.82 seconds against 99.25% and 14.91 seconds). We pick the Boruta-SHAP method for the rest of this study because it provides superior global and local interpretability compared to the alternative RFECV method used in this context of explainable models for lightpath's QoT estimation using the Fraunhofer QoT datasets. Indeed, while RFECV focuses on eliminating features based on the model's internal feature importance, Boruta-SHAP uses the SHAP values for consistent feature attributions and interpretability.

TABLE 5. Selected features from the most to the least significant in model's decision for each dataset.

RANK	1	2	3	4	5	6	7	8	9
D01	Mod Order	Num Spans	Path Len	Avg Link Occ	Freq	Min Link Occ	Sum Link Occ	Max Link Occ	Max BER (R)
D02	Mod Order	Num Spans	Path Len	Avg Link Occ	Freq	Min Link Occ	Sum Link Occ	Min Mod Order (L)	Max BER (R)
D03	Mod Order	Num Spans	Sum Link Occ	Min Link Occ	Freq	Min Mod Order (R)	Path Len	Max BER (R)	Avg Link Occ
D04	Mod Order	Num Spans	Path Len	Avg Link Occ	Min Link Occ	Freq	Sum Link Occ	Min BER (R)	Max BER (R)
	10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	
D01	Min BER (R)	Avg Link Len	Min Mod Order (R)	Min Link Len	Min Mod Order (L)				
D02	Min BER (L)								
D03	Min BER (R)	Max Link Len	Min BER	Std Link Occ	Min Mod Order (L)	Max BER (L)	Min BER (L)	Max LP Linerate (L)	
D04	Source Node Degree								

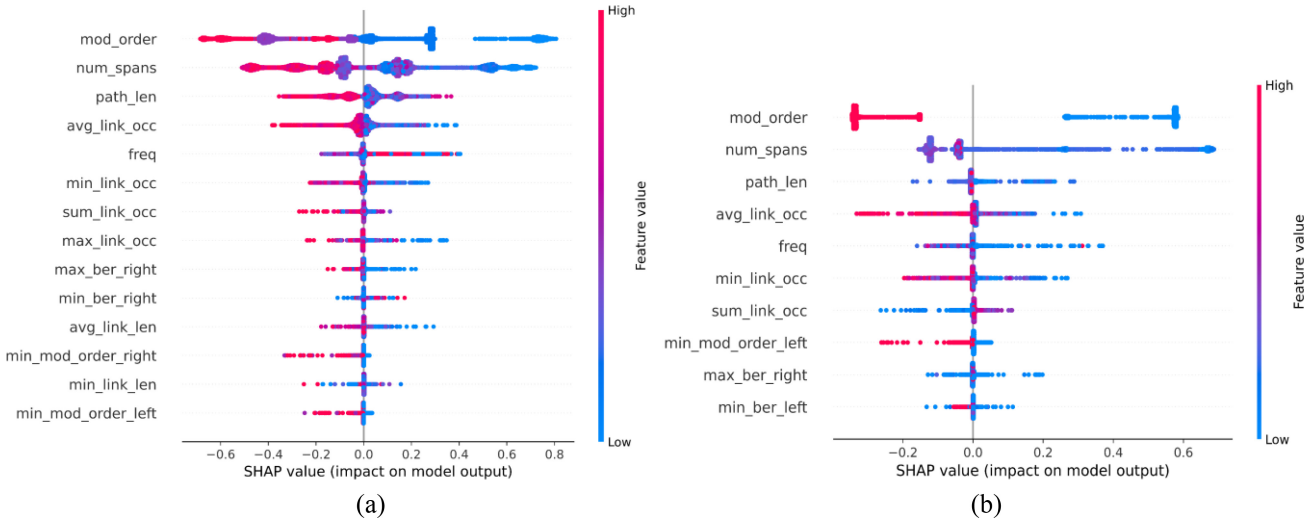

FIGURE 2. SHAP summary plots for the RF-based models using: (a) Dataset 01 and (b) Dataset 02.

The resulting sets of features obtained with Boruta-SHAP are of 14 features, 10 features, 17 features, and 10 features for Dataset 01, 02, 03, and 04 respectively. The rankings of selected features for the four datasets are displayed in Table 5.

With the class-balanced sub-samples, we can see that the same more important attributes appear for the four datasets with similar ranking for the first four features, with the exception for Dataset 03 where the sum of the spectral occupation of the links traversed by the lightpath (*Sum Link Occ*) is ranked 3rd while it is ranked 7th for the other 3 datasets. And the path length (*Path Len*) is ranked 7th for Dataset 03 while it is third for all other datasets. This can be supported by the path length distributions per modulation format and data rate displayed in [14] for the two network topologies CONUS and TSNN. The first four important features for the sub-samples gathered from the CONUS topology are the cardinality of the modulation format (*Mod Order*), the number of spans *(Num Spans*), the path length *(Path Length*), and the average link occupation (*Avg Link Occ*). And for Dataset 03 from TSNN topology, the four most important features are *Mod order*, *Num Spans*, *Sum Link Occ*, and minimum link occupation (*Min Link Occ*). This confirms that datasets from CONUS exhibit greater variation in *Path Len* than the one from TSNN, and the distributions of *Sum Link Occ* and *Min Link Occ* are more heterogeneous in Dataset 03. Despite using class-balanced subsamples of the datasets and a fully automated feature selection process, we obtained similar results for the top three most important features across Dataset 03 (*Mod Order*, *Num Spans*, *Sum Link Occ)* and Dataset 04 (*Mod Order*, *Num Spans*, and *Path Len)*, aligning closely with the outcomes reported by the solution proposed in [38]. This highlights one of the key advantages of using publicly available datasets: it increases our confidence in the outcomes of the proposed solution. More importantly, the process enables us to focus on the most informative features, reducing feature-space noise and improving computational efficiency. Additionally, leveraging only relevant features enhances model interpretability.

Fig. 2 and Fig. 3 show the SHAP summary plots for the models built with the reduced feature sets for global explanation purposes. The figures show only the most

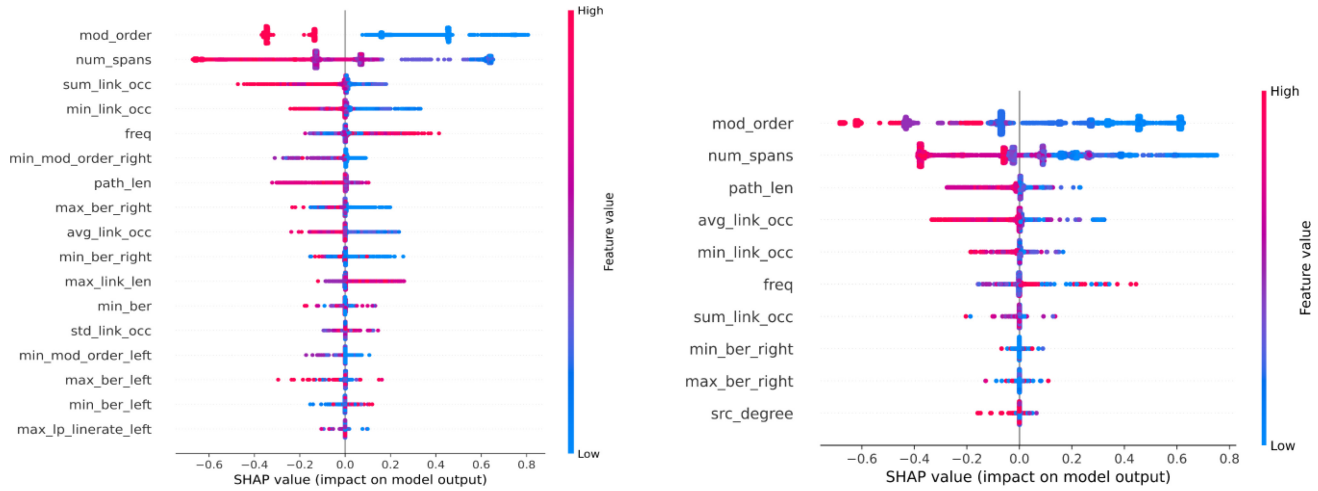

FIGURE 3. SHAP summary plots for the RF-based models using: (a) Dataset 03 and (b) Dataset 04.

relevant features, amongst the 31 attributes, for lightpath QoT estimation using the four datasets. These two figures display bee swarm plots, with features ordered by their impact on the classifier's decisions as measured by SHAP values. Each dot represents an individual observation, with color indicating whether the corresponding feature is high or low relative to other instances in the dataset. Negative SHAP values indicate a negative impact on the model's decision, while positive values correspond to a positive impact. For all four datasets, high values of *Mod Order* contribute negatively to the classification. In other words, a higher modulation format leads the model to assign a negative class label (class 0) to the corresponding instance. Similarly, high values of *Num Spans* exhibit a similar pattern across all datasets, with the exception of Dataset 02 shown in Fig. 2 (b). This aligns with the observation that, in this specific subsample, all instances with a class 0 label have low *Num Spans* values ranging from 23 to 39, while the instances with a positive class label (class 1) have *Num Spans* values between 1 to 102. Another interesting observation in Figures 2 (a), 2 (b), and 3 (b), concerns the feature *Path Len*, which ranks as the third most important feature across all datasets, with the exception of Dataset 03 in Fig. 3 (a). The visualization proposed in [14] shows *Path Len* values ranging from 84 to 1382 km for Dataset 03 while the feature values range from 24 to more than 7000 km for the three other datasets. Although this feature is ranked seventh in Dataset 03, its contribution to the model's decisions remains consistent across Dataset 01 and Dataset 04. High values of *Path Len* are associated to class 0 while low values tend to lead the model toward class 1, as seen in Fig. 2 (a), Fig. 3(a) and (b). For the model built with Dataset 02, we can see in Fig. 2 (b) that high values of *Path Len* have no significant impact on the model's decision. This is consistent with the observation that most of the instances in the subsample with high values of *Path Len* (values greater than 2640 km) are labeled as class 1, as clearly shown in Fig. 4 published in [14].

In real-world scenarios, this Boruta-SHAP-based feature selection integration allows not only to identify the relevant features for QoT estimation of unestablished lightpaths while preserving the classification performance, but with the knowledge provided on the impact of important features on each model, network operators can develop domain-specific strategies to address different SHAP patterns for different network topologies. Also, Figures 2 and 3 can help in prioritizing features with high impact on the models in provisioning decisions.

The *shap* package in Python allows displaying local explanation about the classifier's decision built with selected features using the Boruta-SHAP method. Fig. 4 shows examples of misclassified instances for Dataset 01 for a false positive and a false negative case respectively, namely local explanation examples. Fig. 4 (a) illustrates a false positive case in which the modulation order, the path length and the number of spans increase the model's confidence in predicting class 1. For this particular instance, only the *Avg Link Occ* contributes negatively to the model's prediction score. A lightpath instance, characterized by a *Path Len* of 3462 km, *Num Spans* of 48 and *Mod Order* of 8, supports a prediction of "class 1". However, an *Avg Link Occ* of 66.22 reduces the model's confidence in this prediction. Notably, the ground truth label for this instance is class 0. Fig. 4 (b) shows a false negative case in which *Mod Order* of 8 increases the model's confidence in predicting class 1. But the values for *Num Spans* of 51, *Avg Link Occ* of 53.22, *Path Len* of 3686.93 km, and *Freq* of 193.88 THz all contribute negatively to the model's prediction score, leading to a "class 0" label. The ground truth label in this case is class 1. The length of the bar, representing a feature in the figure, indicates the magnitude of its impact on the classifier's decision. The *Mod Order* is the most influential feature contributing to the false positive shown in Fig. 4(a), while the *Num Spans*, *Avg Link Occ*, and *Path Len* have a greater impact on the false negative shown in Fig. 4 (b).

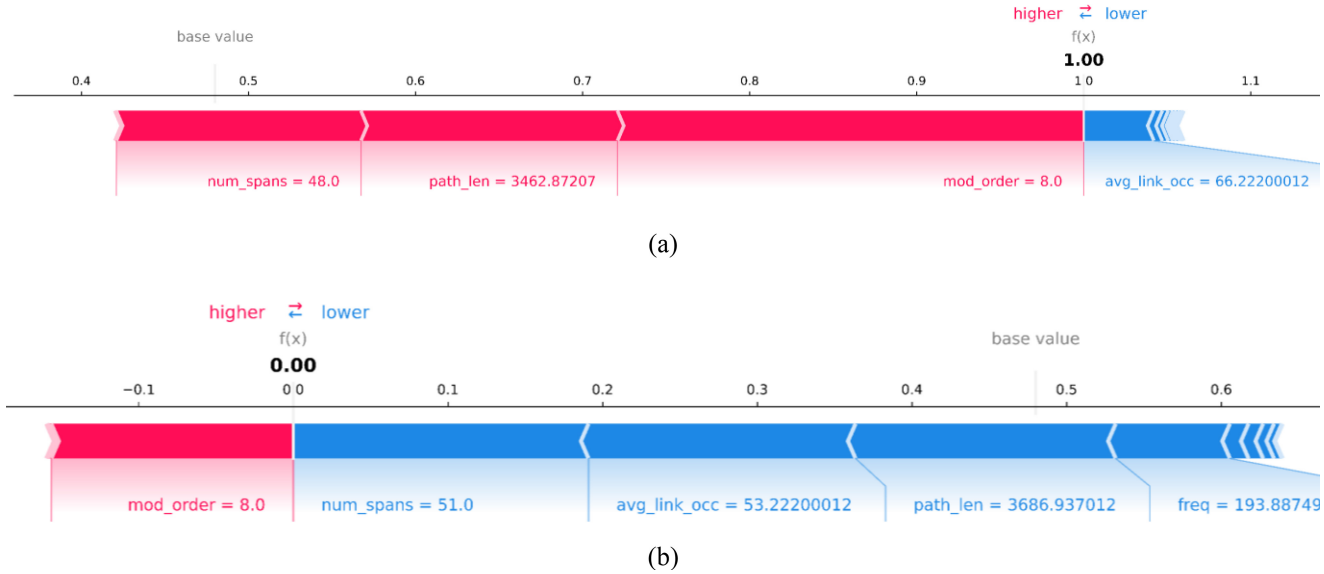

FIGURE 4. Examples of misclassification from Dataset 01: (a) False positive and (b) False negative.

This transparency in the decision-making process enables a deeper understanding of model misclassification and supports further analysis and strategies to complement the automation provided by ML. A step added to the lightpaths' provisioning process where the knowledge provided by the Boruta-SHAP feature selection can be used to review lightpaths' requests presenting specific feature values as the ones shown in Figures 4 (a) and (b). Reviewed requests can be added back to the training set to improve the ML models.

The RF-based model constructed with the 14 selected features outperforms the model built with all 31 features for Dataset 01, with accuracies of 98.64% and 98.6% respectively. And the classifier built with 10 features from Dataset 02 performs slightly better than the 31-feature classifier, with 99.65% and 99.64% respectively. The model built with 17 features from Dataset 03 achieves a classification accuracy of 99.27%, which is better than the performance of the model with all features, which is 99.24%. Lastly, we observe a difference of 0.02% between the model built with 10 features from Dataset 04 and the one built with 31 features (99.52% and 99.5%, respectively).

Although there is only a slight improvement in the classification accuracy for each model, while running on a workstation using an Intel Core i7-9700T 2.00GHz CPU and 16 GB RAM, they show clear reductions in the training times: from 21.98 to 10.38 seconds for Dataset 01, from 11.44 to 4.43 seconds for Dataset 02, from 16.24 to 11.14 seconds for Dataset 03, and from 21.18 to 6.21 seconds for Dataset 04. This represents reductions of 52.78%, 61.28%, 31.40%, and 70.68%, respectively. Moreover, inference times decrease from 73.8 to 71.8 milliseconds for Dataset 01, from 47.8 to 33 milliseconds for Dataset 02, from 60.8 to 41.9 milliseconds for Dataset 03, and from 71.8 to 41.9 milliseconds for Dataset 04. This represents reductions of 2.71%, 30.96%, 31.09%, and 41.64% for reduced feature sets of Dataset 01, 02, 03, and 04 respectively.

TABLE 6. Performance of the RF-based models using complete vs. reduced features sets.

Dataset	*Training Time (s)*		*Inference Time (ms)*	
	Complete	*Reduced*	*Complete*	*Reduced*
01	21.98	**10.38**	73.80	**71.80**
02	11.44	**4.43**	47.80	**33**
03	16.24	**11.14**	60.80	**41.90**
04	21.18	**6.21**	71.80	**41.19**

Table 6 shows a summary of the training time and inference speed of the classical RF-based classifiers built with the complete sets of data and with the reduced feature sets for comparison. These findings show that, with automated feature selection, simpler RF-based models can achieve accuracy comparable to full-feature models while significantly reducing training and inference times. This highlights the practical advantages of using lightweight, interpretable models in time-sensitive applications such as real-time QoT estimation without sacrificing classification performance.

The inference times in Table 6 are achieved with 30,000 test samples. Consequently, the times from 33 to 72 ms presented are the total inference times for the whole test datasets. Thus, on a workstation using an Intel Core i7-9700T 2.00GHz CPU and 16 GB RAM, the average inference times per lightpath vary from 1.1 μs to 2.4 μs. These values are acceptable for practical real-world scenarios if we consider the target of well below one second per lightpath set for the computational time by the Open Optical & Packet Transport–Physical Simulation Environment (OOPT–PSE) group within the Telecom Infra Project (TIP) [3].

Algorithm 1: Feature-Selection-Based Domain Adaptation

Input:
  Source dataset: X_src, y_src
  Target dataset: X_tgt, y_tgt
  Target sample sizes: TS = {50, 100, 1000}
Output:
  {accuracy, precision, recall, f1_score, auc, computation time}
Feature selection phase
Initialize Boruta-SHAP selector
Selected_features ← BorutaSHAP.fit_transform(X_src, y_src)
$X\_src\_sel \leftarrow X\_src[:, Selected\_features]$
$X\_tgt\_sel \leftarrow X\_tgt[:, Selected\_features]$
Domain adaptation phase
for each *num* in TS **do**
  Generate target samples based on the size X_tgt_num, y_tgt_num
  Method 1: CORAL Domain Adaptation
  Initialize CORAL transformer COR
  Fit COR model using (X_src_sel, X_tgt_num)
  Transform the two data samples
  Train the RF-based model using transformed source samples and labels
  Predict classes for transformed target samples
  Method 2: FA Domain Adaptation
  Generate domain indicators for source and target domains
  Concatenate the source samples with the source domain indicators and the target samples with the target domain indicators
  Combine the obtained augmented data samples
  Combine the source and target domain labels
  Train the RF models using the augmented data and combined labels
  Predict classes for transformed target samples
  Jump to performance evaluation
end for
Performance evaluation
for each method $m \in$ {CORAL, FA} **do**
  Compute performance_metrics_m
  Store metrics in results[num][m]
end for

It is worth noting that the order of magnitude of the average inference times per lightpath achieved by our RF-based classifiers is comparable to those reported in [37], where the models were evaluated using a more powerful environment (T4 GPU as the primary hardware accelerator, 13 GB of RAM, and approximately 80 GB of disk space) and a smaller, in-house dataset of 10,000 instances. The CatBoost model with a FS using the model's built-in feature importance ranking in [37], achieves the lowest inference time of 0.49 ms for a 2000-instance test dataset (average 0.24 μs per lightpath), while the lowest inference time achieved by our RF-based model is 33 ms for a 30,000-instance test dataset (in average 1.1 μs per lightpath).

TABLE 7. Classification accuracies of the RF-based models trained on source datasets on target datasets.

Source \ *Target*	*Dataset01*	*Dataset02*	*Dataset03*	*Dataset04*
Dataset01	-	**99.33**	98.62	99.17
Dataset02	**70.21**	-	49.93	65.39
Dataset03	**60.35**	49.93	-	57.98
Dataset04	98.19	**99.50**	98.30	-

B. RF-BASED TRANSFER LEARNING GUIDED BY SELECTED FEATURES

In this section, we describe the case where we apply a TL approach by utilizing the selected features to construct four source models, each corresponding to one of the datasets. As mentioned in Section III-B, we apply a feature-representation transfer strategy, where the features selected from each source domain with the pre-trained RF models are used to transform the target domain. We perform a transformation of the target domain to find shared representation in the data, as explained in the section about the DA in shallow or classical learning presented in [42]. This feature transformation method corresponds to the SDB presented in [45]. We then evaluate each pre-trained RF model against the three other datasets with reduced feature sets. The outcome of this study is shown in Table 8.

Our results confirm that Dataset 01 is the most comprehensive of the four subsamples, with classification accuracies above 99% on Dataset 02 and Dataset 04

In addition, with the selected features the proposed RF-based models achieve 98.62% on Dataset 03, which is better than the performance obtained in [14]. Moreover, models built with Dataset 02 and Dataset 03 achieve classification accuracies below 71% on each of the other datasets. It is worth noting that our RF-based model trained with Dataset 04 achieves an accuracy of 98.19% on Dataset 01, while the model trained with Dataset 01 achieves 99.17% accuracy on Dataset 04. Similarly to the results in [14], the model trained with Dataset 01 achieves more than 99% accuracy in classifying test instances from Dataset 04, even though the BPSK modulation format does not appear in Dataset 01. However, we could see an increase in classification accuracy (from 73.3% to 98.19%) for the transfer learning on the target domain Dataset 01 using the model trained with Dataset 04, as compared to the performance obtained in [14]. This may be because of the applied feature selection process which, with the class-balanced subsamples, allows to find new feature combinations to gain better insight on how to apply feature-based transfer learning as well as which attributes are important for QoT estimation in optical networks.

TABLE 8. Performance metrics of the DA models with Dataset 03 as source domain and Dataset 02 as target domain.

	Accuracy (%)	*Precision (%)*	*Recall (%)*	*F1-Score (%)*	*AUC (%)*	*Train/Test Time (s)/(ms)*
			Target domain size = 50			
CORAL	*68.00*	*73.33*	*73.33*	*73.33*	*69.33*	*46.19/10.2*
FA	*82.00*	*95.65*	*73.33*	*83.01*	*94.75*	*12.482/10.1*
FS-CORAL	*82.00*	*83.87*	*86.66*	*85.24*	*89.50*	*29.04/10.1*
FS-FA	*86.00*	*96.00*	*80.00*	*87.27*	*94.16*	*10.42/10.1*
			Target domain size = 100			
CORAL	*80.00*	*82.00*	*78.84*	*80.39*	*88.10*	*44.41/11.9*
FA	*97.00*	*98.03*	*96.15*	*97.08*	*99.41*	*12.94/18.3*
FS-CORAL	*55.00*	*53.93*	*92.30*	*68.08*	*78.14*	*28.02/11.9*
FS-FA	*98.00*	*100.00*	*96.15*	*98.03*	*99.75*	*10.68/10.9*
			Target domain size = 1000			
CORAL	*84.30*	*84.70*	*81.27*	*82.95*	*91.42*	*44.60/18.1*
FA	*99.00*	*99.78*	*98.08*	*98.92*	*99.89*	*13.02/10.1*
FS-CORAL	*64.70*	*57.72*	*92.97*	*71.23*	*78.47*	*28.34/14.9*
FS-FA	*99.10*	*100.00*	*99.03*	*99.03*	*99.84*	*10.91/15.9*

These results show that RF can face generalization challenges when DA is performed across heterogeneous domains using the SDB approach, in contrast to the conclusions reported in [45]. Nonetheless, the results obtained using the FA and CORAL feature-selection-based methods exhibit a significant improvement in classification accuracy for one of the worst-case scenarios presented in Table 6. This worst-case scenario corresponds to TL from source Dataset 03, representing the smaller TSNN network, to target Dataset 02, associated with the larger CONUS network.

Our novel TL framework integrates the feature selection and domain adaptation into a unified algorithm. We use the Boruta-SHAP technique to build a feature importance mask that captures the most relevant features from the source domain while providing insights on classification decision made by the model. By using the mask to transform both the source and target domain data, we ensure that adaptation is performed only over the most meaningful and statistically stable features. Following the first transformation using the Boruta-SHAP to select only relevant features, the solution further adjusts the joint source and target data to minimize domain shift, allowing the model to learn transferable relationships. Rather than applying the two components independently, our approach combines them into an integrated transformation that aligns domains, improves generalization on small target samples, and provides interpretability of classification outcomes.

Overall better accuracies of 86%, 98%, and 99.10% are achieved with FS-FA as compared to 82%, 97%, and 99% obtained with the standard feature augmentation technique for target domains of 50, 100, and 1000 instances respectively as shown in Table 8.

We test our proposed feature-selection-based feature augmentation (FS-FA) DA framework with the same source and target domain sizes used with the SDB approach. The proposed solution achieved a classification accuracy of 99.54% with Dataset 03 as the source domain and Dataset 02 as the target domain. That is an increase of 99.35% in classification accuracy when we compare it to the 49.93% achieved with the source domain baseline approach.

The results show that increasing in the number of samples in the target domain allows the combined (FS-FA) framework to achieve better performance, with a slight increase in training time ranging from 10.42 to 10.68 and 10.91 seconds for 50, 100, and 1000 instances, respectively. The same trend is observed for the inference time which varies from 10.1 to 10.9, and 15.9 milliseconds for target domain sizes of 50, 100, and 1000 instances, respectively. For a better view of trade-offs between sensitivity and specificity, we calculate the area under the curve (AUC) for each scenario. A similar trend is observed for target domain sizes of 50, 100, and 1000 instances with AUC values of 94.16%, 99.75%, and 99.84% respectively. The proposed framework achieves faster training times in the three scenarios and faster inference times in average, along with the combined feature selection with CORAL with 12.3 seconds as compared to 12.8 and 13.4 milliseconds for standard FA and CORAL respectively.

Moreover, while the model built with the CORAL method exhibits reduced training times after combining it with the feature selection (FS-CORAL) compared to its counterpart without feature selection, the results indicate that its performance degrades as the size and the complexity of the associated feature-space of the target domain increases. The

classification performance for the CORAL method decreases from 80% to 55% and from 84.3% to 64.7% for the model without feature selection and the feature-selected-based model with target domain sizes of 100 and 1000 respectively.

Ultimately, we obtain overall better results while applying the proposed (FS-FA) DA framework using RF-based model than when applying the standard DA methods (CORAL or FA) without prior feature selection. Thus, we conclude that the proposed FS-FA framework significantly outperforms standard DA methods, which use FA and CORAL without applying the FS method, in classification performance as well as computational efficiency for QoT estimation of lightpaths before their establishment. We perform additional tests using Dataset 01 as the source domain and subsamples from the three other datasets (Dataset 02, 03, and 04) as target domains. An ablation style comparison of the standard DA techniques (standard FA) with our proposed framework (FS-FA) shows consistent improvements in the computation times with the small subsamples of 50 instances. The training time decreases from 17.41 to 10.20 seconds and the inference time from 8.97 to 8.49 ms for DA between Dataset 01 and Dataset 02. The same trend is seen for DA between Dataset 01 to Dataset 03 and Dataset 04 with the training time decreasing from 17.48 to 10.16 seconds and 17.0 to 10.39 seconds, and the inference time reductions from 9.97 to 8.97 ms and 9.47 to 8.99 ms, respectively, without sacrificing the classification accuracy. Our Boruta-SHAP-based method yields shorter computation times while maintaining the same classification accuracy of 98% for both DA approaches, when using the standard feature augmentation and the Boruta-SHAP-based feature augmentation schemes, with Dataset 01 as the source domain and Datasets 02, 03, and 04 as target domains.

In contrast to the hybrid solution integrating KD with TL proposed in [20] where knowledge is transferred from a bigger "teacher" ANN model to a smaller "student" ANN model, our proposed approach leverages RF-based classifiers with Boruta-SHAP-based feature selection providing model transparency with lower complexity. Our TL framework automatically identifies important features while in previous works the authors manually build feature vector with characteristics of the considered lightpath as well as the ones of the neighboring channels [45], and other scholars filter samples by manually selecting length and number of adjacent channels features [21].

Our framework achieves reduction in both training and inference times while preserving the classification accuracy with 99.83% reduction in the size of the target domain. The Boruta-SHAP-based feature augmentation (FS-FA) achieves reduction of the training times by 41.41%, 41.89%, and 38.88%, as well as reduction of the inference times by 5.35%, 10.03%, and 5.07% for Dataset 01 to Dataset 02, 03, and 04, respectively. Although the authors of [20] used a workstation with a 12th Generation Intel Core i7 12700F 2.10 GHz CPU and 32GB RAM, which is superior to our workstation with an Intel Core i7-9700T 2.00GHz CPU and 16 GB RAM, their proposed KD+TL model, with their 6,000-instance target data samples, achieves training time of 6.7 seconds and prediction time of 0.09 seconds while our framework achieves training and inference times of 2.51 and 0.04 seconds when we use a 12,000-instance subsample from Dataset 01 as source domain and a 6,000-instance subsample from Dataset 03 as target domain. Moreover, while a direct comparison with [44] is not entirely feasible due to differences in the experimental environments, datasets, and ML methods, our lightweight RF-based model nonetheless achieves a better inference time - by an order of magnitude - than the AI/ML-as-a-service framework reported therein. In [44], the average inference time achieved by the ANN model, tested with a 10,000-instance subsample gathered from Dataset 01 on an Ubuntu 22.04 AMD Ryzen Threadripper 3960X 24-Core Processor is 40 ms while their proposed AI/ML-as-a-service framework achieves 553 ms per inference, due in part, to the ensemble of models implemented in the framework. We demonstrate that our framework achieves similar order of magnitude for both the training and the inference times when compared to [20] for domain adaptation. When compared to [44], our lightweight RF-based model achieves better inference times than the proposed AI/ML-as-a-service framework. This shows the benefits of lightweight RF-based models compared to ANN-based models for QoT estimation of unestablished lightpaths.

V. CONCLUSION AND FUTURE RESEARCH DIRECTION

We develop an interpretable and efficient QoT estimation framework tailored in particular for the real-world use cases, where the training datasets are scarce and the training time is a critical factor. A comparative analysis using the four publicly available datasets from the Fraunhofer Institute with different representative techniques clearly demonstrates the advantages of our approach. In particular, we show effectiveness of classical RF-based models in estimating the QoT of unestablished lightpaths, offering enhanced efficiency and interpretability while preserving classification accuracy.

We propose an efficient automated approach to select the most important features, guiding the models in classifying test instances across each class-balanced sub-dataset. Our approach reduces the dimensionality of the training data (from 31 features to 14, 10, 17, and 10 for Datasets 01, 02, 03, and 04, respectively), thereby significantly lowering processing time while preserving, or even slightly improving, classification accuracy. Moreover, the combination of SHAP values with the Boruta technique, referred to as the *Boruta-SHAP method,* provides deeper insights into how important features influence the model's classification decisions for test instances. Basic visual error analysis helps uncover potential explanations for the two misclassification cases presented in Fig. 4 (a) and (b). The RF-based models, built with reduced feature sets, show clear reductions in the training times of 52.78%, 61.28%, 31.40%, and 70.68% for the four sub-samples obtained from Datasets 01, 02, 03,

and 04, respectively. Additionally, the inference times also show reductions of 2.71%, 30.96%, 31.09%, and 41.64% for the four sub-samples, respectively. This demonstrates the practical applicability of RF-based models in real-world scenarios.

Furthermore, we implement a feature-selection-driven feature augmentation (FS-FA) domain adaptation framework to evaluate the cross-domain performance of models trained on one dataset and tested on others. In this approach, Boruta-SHAP-based feature selection is applied prior to the feature transformation step for transfer learning. Depending on the subsample used as the source domain, a pre-trained models achieve accuracies ranging from 49.93% to 99.50% with an SDB approach. Our framework reaches a classification accuracy of 99.54% with the complete 30,000 samples target domain used with the SDB approach and of 86% with a small target domain of 50 samples. Moreover, our framework shows a reduction in training time of 77.44%, 16.52%, and 64.11% for the small target dataset, when compared to the standard CORAL, the standard FA, and the FS-CORAL methods, respectively. The framework also exhibits fast inference time of 10.10 milliseconds, which makes it an efficient solution for deployment in real-world optical networks.

As a future research direction, we plan to explore more advanced DA strategies to improve the generalizability of QoT estimation models across diverse network scenarios. And we will perform simulations with other target domain sizes to determine appropriate sizes for real-world implementations.

REFERENCES

[1] D. Pilori, M. Cantono, A. Carena, and V. Curri, "FFSS: The fast fiber simulator software," in *Proc. 19th Int. Conf. Transparent Optical Netw. (ICTON)*, Girona, Spain, 2017, pp. 1–4, doi: 10.1109/ICTON.2017.8025002.

[2] J. Wass, J. Thrane, M. Piels, R. Jones, and D. Zibar, "Gaussian process regression for WDM system performance prediction," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, Los Angeles, CA, USA, 2017, pp. 1–3.

[3] A. Ferrari et al., "GNPy: An open source application for physical layer aware open optical networks," *J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. C31–C40, Jun. 2020, doi: 10.1364/JOCN.382906.

[4] I. Khan et al., "QoT estimation for light-path provisioning in unseen optical networks using machine learning," in *Proc. 22nd Int. Conf. Transparent Opt. Netw. (ICTON)*, Bari, Italy, 2020, pp. 1–4, doi: 10.1109/ICTON51198.2020.9203364.

[5] C. Manso, R. Vilalta, R. Muñoz, R. Casellas, and R. Martínez, "Scalable for cloud-native transport SDN controller using GNPy and machine learning techniques for QoT estimation," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, San Francisco, CA, USA, 2021, pp. 1–3.

[6] F. Usmani et al., "Evaluating cross- feature trained machine learning models for estimating QoT of unestablished lightpaths," in *Proc. Int. Conf. Elect., Commun., Comput. Eng. (ICECCE)*, Kuala Lumpur, Malaysia, 2021, pp. 1–6, doi: 10.1109/ICECCE52056.2021.9514154.

[7] F. Usmani, I. Khan, M. U. Masood, A. Ahmad, M. Shahzad, and V. Curri, "Transfer learning aided QoT computation in network operating with the 400ZR standard," in *Proc. Int. Conf. Opt. Netw. Design Model. (ONDM)*, Warsaw, Poland, 2022, pp. 1–6, doi: 10.23919/ONDM54585.2022.9782856.

[8] F. Usmani et al., "Performance analysis of transfer-learning approaches for QoT estimation of network operating with 400ZR," in *Proc. Asia Commun. Photonics Conf. (ACP)*, Shenzhen, China, 2022, pp. 1238–1242, doi: 10.1109/ACP55869.2022.10088462.

[9] M. K. Naik, R. K. Jeyachitra, and P. Doss M, "Machine learning based multi-parameter light path quality of transmission estimation and prediction in optical networks," in *Proc. 5th Int. Conf. Elect., Comput. Commun. Technol. (ICECCT)*, Erode, India, 2023, pp. 1–6, doi: 10.1109/ICECCT56650.2023.10179761.

[10] S. Aladin and C. Tremblay, "Cognitive tool for estimating the QoT of new lightpaths," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, San Diego, CA, USA, 2018, pp. 1–3.

[11] C. Tremblay and S. Aladin, "Machine learning techniques for estimating the quality of transmission of lightpaths," in *Proc. IEEE Photon. Soc. Summer Topical Meeting Series (SUM)*, Waikoloa, HI, USA, 2018, pp. 237–238, doi: 10.1109/PHOSST.2018.8456791.

[12] S. Allogba, S. Aladin, and C. Tremblay, "Machine-learning-based lightpath QoT estimation and forecasting," *J. Lightw. Technol.*, vol. 40, no. 10, pp. 3115–3127, May 15, 2022, doi: 10.1109/JLT.2022.3160379.

[13] S. Aladin, A. V. S. Tran, S. Allogba, and C. Tremblay, "Quality of transmission estimation and short-term performance forecast of lightpaths," *J. Lightw. Technol.*, vol. 38, no. 10, pp. 2807–2814, May 15, 2020, doi: 10.1109/JLT.2020.2975179.

[14] G. Bergk, B. Shariati, P. Safari, and J. K. Fischer, "ML-assisted QoT estimation: a dataset collection and data visualization for dataset quality evaluation," *J. Opt. Commun. Netw.*, vol. 14, no. 3, pp. 43–55, Mar. 2022, doi: 10.1364/JOCN.442733.

[15] Z. Ouyang, X. Chen, Z. Liu, X. Chen, and Z. Zhu, "Overview of ML-aided QoT estimation in optical networks: A perspective of model generalization," in *Proc. IEEE 24th Int. Conf. Commun. Technol. (ICCT)*, Chengdu, China, 2024, pp. 130–136, doi: 10.1109/ICCT62411.2024.1094652.

[16] Z. Cai et al., "Domain adversarial adaptation framework for few-shot QoT estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 16, no. 11, pp. 1133–1144, Nov. 2024, doi: 10.1364/JOCN.530915.

[17] Y. Zhou, Z. Gu, J. Zhang, and Y. Ji, "Evolutionary neuron-level transfer learning for QoT estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 16, no. 4, pp. 432–448, Apr. 2024, doi: 10.1364/JOCN.514618.

[18] P. Paudyal, S. Shen, S. Yan, and D. Simeonidou, "Toward deployments of ML applications in optical networks," *IEEE Photon. Technol. Lett.*, vol. 33, no. 11, pp. 537–540, Jun. 1, 2021, doi: 10.1109/LPT.2021.3074586.

[19] C.-Y. Liu, X. Chen, R. Proietti, and S. J. B. Yoo, "Performance studies of evolutionary transfer learning for end-to-end QoT estimation in multi-domain optical networks [Invited]," *J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B1–B11, Apr. 2021, doi: 10.1364/JOCN.409817.

[20] F. Usmani, I. Khan, Mehran, A. Ahmad, and V. Curri, "Integrating knowledge distillation and transfer learning for enhanced QoT-estimation in optical networks," *IEEE Access*, vol. 12, pp. 156785–156802, 2024, doi: 10.1109/ACCESS.2024.348599.

[21] Z. Gu, T. Qin, Y. Zhou, J. Zhang, and Y. Ji, "Sample-distribution-matching-based transfer learning for QoT estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 15, no. 9, pp. 649–663, Sep. 2023, doi: 10.1364/JOCN.493053.

[22] J. Pesic, M. Lonardi, E. Seve, N. Rossi, and T. Zami, "Transfer learning from unbiased training data sets for QoT estimation in WDM networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Brussels, Belgium, 2020, pp. 1–4, doi: 10.1109/ECOC48923.2020.9333399.

[23] Y. Zhou, Z. Gu, J. Zhang, and Y. Ji, "Neuron-level transfer learning for ANN-based QoT estimation in optical networks," in *Proc. Asia Commun. Photon. Conf. Int. Photon. Optoelectron. Meetings (ACP/POEM)*, Wuhan, China, 2023, pp. 1–4, doi: 10.1109/ACP/POEM59049.2023.10368618.

[24] D. Azzimonti, C. Rottondi, A. Giusti, M. Tornatore, and A. Bianco, "Active vs transfer learning approaches for qot estimation with small training datasets," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, San Diego, CA, USA, 2020, pp. 1–3.

[25] H. Tariq, F. Usmani, I. Khan, M. U. Masood, A. Ahmad, and V. Curri, "Iterative transfer learning approach for QoT prediction of lightpath in optical networks," in *Proc. 23rd Int. Conf. Transparent Opt. Netw. (ICTON)*, Bucharest, Romania, 2023, pp. 1–4, doi: 10.1109/ICTON59386.2023.10207173.

[26] F. Usmani, I. Khan, M. U. Masood, A. Ahmad, M. Shahzad, and V. Curri, "QoT-estimation assisted by transfer learning in extended C-band network operating on 400ZR," in *Proc. IEEE Photon. Soc. Summer Topicals Meeting Series (SUM)*, Cabo San Lucas, Mexico, 2022, pp. 1–2, doi: 10.1109/SUM53465.2022.9858280.

[27] S. Cruzes, "Overview of quality of transmission estimation in optical networks," TechRxiv, Aug. 6, 2024. doi: 10.36227/techrxiv.171340665.50082312/v2.

[28] M. Ibrahimi et al., "Machine learning regression for QoT estimation of unestablished lightpaths," *J. Opt. Commun. Netw.*, vol. 13, no. 4, pp. B92–B101, Apr. 2021, doi: 10.1364/JOCN.410694.

[29] I. Khan, M. Bilal, and V. Curri, "Advanced formulation of QoT-estimation for un-established lightpaths using cross-train machine learning methods," in *Proc. 22nd Int. Conf. Transparent Opt. Netw. (ICTON)*, Bari, Italy, 2020, pp. 1–4, doi: 10.1109/ICTON51198.2020.9203334.

[30] M. A. Amirabadi, M. H. Kahaei, S. A. Nezamalhosseini, F. Arpanaei, and A. Carena, "Deep neural network-based QoT estimation for SMF and FMF links," *J. Lightw. Technol.*, vol. 41, no. 6, pp. 1684–1695, Mar. 15, 2023, doi: 10.1109/JLT.2022.3225827.

[31] Q. Wang, Z. Cai, A. P. T. Lau, Y. Li, and F. Nadeem Khan, "Invariant convolutional neural network for robust and generalizable QoT estimation in fiber-optic networks," *J. Opt. Commun. Netw.*, vol. 15, no. 7, pp. 431–441, Jul. 2023, doi: 10.1364/JOCN.488689.

[32] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," *J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A286–A297, Feb. 2018, doi: 10.1364/JOCN.10.00A286.

[33] A. Caballero et al., "Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems," in *Proc. 38th Eur. Conf. Exhibition Opt. Commun.*, Amsterdam, The Netherlands, 2012, pp. 1–3.

[34] R. M. Morais and J. Pedro, "Machine learning models for estimating quality of transmission in DWDM networks," *J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D84–D99, Oct. 2018, doi: 10.1364/JOCN.10.000D84.

[35] T. Panayiotou, G. Savva, B. Shariati, I. Tomkos, and G. Ellinas, "Machine learning for QoT estimation of unseen optical network states," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, San Diego, CA, USA, 2019, pp. 1–3.

[36] A. A. Diaz-Montiel, S. Aladin, C. Tremblay, and M. Ruffini, "Active wavelength load as a feature for QoT estimation based on support vector machine," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, 2019, pp. 1–6, doi: 10.1109/ICC.2019.8761369.

[37] H. Fawaz et al., "Reducing complexity and enhancing predictive power of ML-based lightpath QoT estimation via SHAP-assisted feature selection," in *Proc. Int. Conf. Opt. Netw. Design Model. (ONDM)*, Madrid, Spain, 2024, pp. 1–6, doi: 10.23919/ONDM61578.2024.10582777.

[38] O. Ayoub et al., "Towards explainable artificial intelligence in optical networks: The use case of lightpath QoT estimation," *J. Opt. Commun. Netw.*, vol. 15, no. 1, pp. A26–A38, Jan. 2023, doi: 10.1364/JOCN.470812.

[39] P. Safari, B. Shariati, G. Bergk, and J. K. Fischer, "Deep convolutional neural network for network-wide QoT estimation," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, San Francisco, CA, USA, 2021, pp. 1–3.

[40] N. Hashemi, P. Safari, B. Shariati, and J. K. Fischer, "Vertical federated learning for privacy-preserving ML model development in partially disaggregated networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Bordeaux, France, 2021, pp. 1–4, doi: 10.1109/ECOC52684.2021.9605846.

[41] P. Safari, B. Shariati, G. Bergk, and J. K. Fischer, "Secure multi-party computation and statistics sharing for ML model training in multi-domain multi-vendor networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Bordeaux, France, 2021, pp. 1–4, doi: 10.1109/ECOC52684.2021.9606082.

[42] C. Natalino, N. Mohammadiha, and A. Panahi, "Machine-learning-as-a-service for optical network automation," in *Proc. Opt. Fiber Commun. Conf. Exhibition (OFC)*, San Diego, CA, USA, 2023, pp. 1–3, doi: 10.1364/OFC.2023.W4G.

[43] C. Natalino, A. Panahi, N. Mohammadiha, and P. Monti, "AI/ML-as-a-service for optical network automation: Use cases and challenges [invited]," *J. Opt. Commun. Netw.*, vol. 16, no. 2, pp. A169–A179, Feb.2024, doi: 10.1364/JOCN.50070.

[44] C. Rottondi, R. di Marino, M. Nava, A. Giusti, and A. Bianco, "On the benefits of domain adaptation techniques for quality of transmission estimation in optical networks," *J. Opt. Commun. Netw.*, vol. 13, no. 1, pp. A34–A43, Jan. 2021, doi: 10.1364/JOCN.40191.

[45] R. Di Marino, C. Rottondi, A. Giusti, and A. Bianco, "Assessment of domain adaptation approaches for QoT estimation in optical networks," in *Proc. Opt.Fiber Commun. Conf. Exhibition (OFC)*, San Diego, CA, USA, 2020, pp. 1–3.

[46] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012, doi: 10.1109/JLT.2012.2217729.

[47] P. Domingos, "The Role of Occam's Razor in knowledge discovery," *Data Min. Knowl. Disc.*, vol. 3, pp. 409–425, Dec. 1999. [Online]. Available: https://doi.org/10.1023/A:1009868929893

[48] P. Singhal, R. Walambe, S. Ramanna, and K. Kotecha, "Domain adaptation: Challenges, methods, datasets, and applications," *IEEE Access*, vol. 11, pp. 6973–7020, 2023, doi: 10.1109/ACCESS.2023.3237025.

SANDRA ALADIN (Member, IEEE) received the B.Eng. degree in electronics from the Université d'Etat d'Haiti, Port au Prince, Haiti, in 2009, and the M.A.Sc. degree in telecommunications networks engineering from the École de technologie supérieure, Montréal, QC, Canada, in 2018, where she is currently pursuing the Ph.D. degree. Her Master's thesis focused on machine-learning-based quality of transmission estimation in optical networks. From 2018 to 2021, she was a Research Assistant of Machine Learning for Optical Networking with the Network Technology Laboratory, Department of Electrical Engineering, École de technologie supérieure, where she is working on smart optical networks enabled by performance monitoring and machine learning.

LENA WOSINSKA (Senior Member, IEEE) received the Ph.D. degree in photonics and the Docent degree in optical networks from the KTH Royal Institute of Technology, Sweden, where she became a Full Professor of Telecommunication. At KTH, she established and led a highly recognized research group working on optical networks. She is currently a Research Professor with the Chalmers University of Technology, Sweden, where she moved together with her team in the beginning of 2019. Her research broadly concerns optical communications and networks, with a focus on fiber access and 5G/6G transport networks, energy and cost efficiency, optical data center networks, photonics in switching, network control, reliability, security, and survivability, including application of ML/AI. She has been involved in many expert assignments, including serving in the panels evaluating research project proposals for many funding agencies, editorship of leading IEEE, OSA, Elsevier, and Springer journals, serving as the General Chair and the Co-Chair of several IEEE, OSA, and SPIE conferences and workshops. She was also involved and had a leading role in several EU projects, as well as coordinating national and international research projects. She received the IEEE ComSoc ONTC Outstanding Technical Achievement Award for Pioneering Research in Optical Networks and the OPTICA Jane M. Simons Memorial Speakership Award for the numerous contributions to optical networking research and for mentoring and inspiring a generation of leaders in optical networking.

CHRISTINE TREMBLAY (Senior Member, IEEE) received the B.Sc. degree in engineering physics from Université Laval, Quebec City, Canada, in 1984, the M.Sc. degree in energy from INRS-Énergie, Varennes, Canada, in 1985, and the Ph.D. degree in optoelectronics from the École Polytechnique de Montréal, Canada, in 1992, where she is a Full Professor with the Department of Electrical Engineering and Director for the Master's Program in Electrical Engineering and an Associate Director for the Ph.D. Program. She is the Founding Researcher and a Head of the Network Technology Lab. Before joining ÉTS, she was a Research Scientist with the National Optics Institute, where she conducted research on integrated optical devices for communication and sensing applications. She held Senior Research and Development and Technology Management positions for several organizations. As an Engineering Manager with EXFO and a Director of Engineering with Roctest, she was responsible for the development of fiber-optic test equipment. She also served as the Product Manager at Nortel for DWDM systems. Her team pioneered the research on filterless optical networking. Her current research interests include machine learning for optical communications, optical performance monitoring, fiber sensing, and quantum optical networking. She has been a Co-Instructor for the OFCnet SC528 hands-on short course since 2024, following her earlier role as Co-Instructor for OFC hands-on short courses SC314 (2009–2015) and SC210 (2011–2015). She served as the Program Chair for the Photonics Networks and Devices (NETWORKS) Topical Meeting in 2020 and 2021, as well as a Program Committee Member for the OFC 2020–2023 Subcommittee N3: Architecture and Software-Defined Control for metro and core networks. She is a member of the Optical Society of America and STARaCom and COPL Strategic Clusters of FRQNT. She has been a member of OFCnet Committee since 2024.