Thesis for The Degree of Licentiate of Engineering

Toward Enabling Robotic Visual Perception for Assembly Tasks

An Application in Wire Harness Assembly onto Electric Vehicles

Hao Wang

Department of Industrial and Materials Science
Chalmers University of Technology
Gothenburg, Sweden, 2024

Toward Enabling Robotic Visual Perception for Assembly Tasks
An Application in Wire Harness Assembly onto Electric Vehicles

Hao Wang

This thesis has been prepared using LaTeX.

Licentiate Thesis at Chalmers University of Technology
Technical Report No IMS-2024-5

Department of Industrial and Materials Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31 772 1000

Printed by Chalmers Digital Printing,
Gothenburg, Sweden 2024

To Mom and Dad.

Contents

Toward Enabling Robotic Visual Perception for Assembly Tasks
An Application in Wire Harness Assembly onto Electric Vehicles

Hao Wang

Department of Industrial and Materials Science
Chalmers University of Technology

Abstract

Industry faces an urgent need for prospective solutions to scale up assembly automation, a challenge that requires immediate attention. In contemporary manufacturing, industrial robots need more intelligence to qualify for increasingly demanding flexible automation tasks. Research in artificial intelligence, computer vision, and robotics paints a promising picture of the future, where intelligent robots play a significant role in fostering sustainable and resilient manufacturing. However, academia and industry have yet to realize the potential of intelligent robots in production fully.

This thesis plays a pivotal role in advancing the development of intelligent robots for flexible automation tasks, a crucial area of research in automation and robotics. Toward this goal, this thesis investigates perception, a prerequisite of intelligence, and mainly focuses on visual perception, a critical contactless perception approach. A multi-method research approach, comprising a qualitative literature study and a quantitative experimental study, was adopted to explore the challenges and prospective technical solutions to enabling robotic visual perception for assembly tasks.

The research has identified four key challenges in enabling robotic visual perception for assembly tasks, particularly in developing and integrating vision systems in practical production. Additionally, the research has proposed six prospective directions for developing technical solutions, focusing on computer vision algorithms, dataset and benchmark, practical evaluation, human-robot collaboration, and product design.

Keywords

Robotic visual perception, Computer vision, Artificial intelligence, AI, Human-robot collaboration, HRC, Flexible automation, Assembly, Automotive industry

Acknowledgment

This work would only have been realized with much help and support from different people and organizations. I want to acknowledge those who have contributed to bringing this work to fruition.

Firstly, I am deeply grateful to my examiner, Johan Stahre, my main supervisor, Björn Johansson, and my co-supervisor, Anders Skoogh. Their guidance and support have been instrumental in instilling in me a passion for my study and research and in helping me realize this work.

I also appreciate the rest of the Production Systems division for the many beneficial discussions and friendly chats. A relaxed and friendly working environment is genuinely an incubator of inspiration and collaboration toward success. Likewise, I thank all other people in academia or industry with whom we have talked, collaborated, and exchanged ideas, especially those who co-authored the papers appended to this thesis.

Besides, I greatly appreciate the significant impact of all the partners we have collaborated with on research projects. The invaluable support from the funding organizations, Vinnova, Produktion2030, the European Commission, and Chalmers Production Area of Advance, has also been crucial and greatly acknowledged.

Lastly, my eternally special gratitude goes to my most loved ones. To my mom and dad, for bringing me the courage to explore the unknowns. To my fiancée, Shan, for casting the brightest light on my way forward.

Acronyms and Abbreviations

2D	Two-Dimensional
2DOD	2D Object Detection
3D	Three-Dimensional
3DOD	3D Object Detection
6D	Six-Degree-of-Freedom
6DOPE	6D Object Pose Estimation
AI	Artificial Intelligence
BDLO	Branched Deformable Linear Object
CAD	Computer-Aided Design
CIRP	The International Institution of Production Engineering Research
CNN	Convolutional Neural Network
DARE	The Database of Abstracts of Reviews of Effects
DETR	Detection Transformer
DINO	DETR with Improved Denoising Anchor Boxes
DLO	Deformable Linear Object
DLON	Deformable Linear Object Network
DOO	Deformable One-dimensional Object
DPM	Deformable Part Model
DSR	Design Science Research
DSRM	Design Science Research Methodology
EV	Electric Vehicle
FPN	Feature Pyramid Network
HOG	Histogram of Oriented Gradients
HRC	Human-Robot Collaboration
LiDAR	Light Detection And Ranging
mAP	mean Average Precision
NHS	The National Health Service
PASCAL	Pattern Analysis, Statistical Modelling and Computational Learning
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
R-CNN	Region-based CNN
RGB	Red-Green-Blue
RGB-D	Red-Green-Blue-Depth
ROI	Region of Interest
RQ	Research Question
SDLO	Semi-Deformable Linear Object
SPPNet	Spatial Pyramid Pooling Network
SSD	Single-Shot Detector
TRL	Technology Readiness Level
VOC	Visual Object Classes
YOLO	You Only Look Once

List of Publications

Appended Publications

This thesis is based on the following publications:

[**Paper 1**] **Overview of Computer Vision Techniques in Robotized Wire Harness Assembly: Current State and Future Opportunities**

Hao Wang, Omkar Salunkhe, Walter Quadrini, Dan Lämkull, Fredrik Ore, Björn Johansson, Johan Stahre

Presented at the 56th CIRP Conference on Manufacturing Systems (CIRP CMS 2023), Cape Town, South Africa, 24-26 October 2023.

Published in: *Procedia CIRP*, 2023.

`https://doi.org/10.1016/j.procir.2023.09.127`

Hao Wang initiated and wrote the paper with other authors' contributions and reviews. The research design as well as data collection and analysis were conducted together with Omkar Salunkhe and Walter Quadrini.

[**Paper 2**] **A Systematic Literature Review of Computer Vision Applications in Robotized Wire Harness Assembly**

Hao Wang, Omkar Salunkhe, Walter Quadrini, Dan Lämkull, Fredrik Ore, Mélanie Despeisse, Luca Fumagalli, Johan Stahre, Björn Johansson

Under the second revision by *Advanced Engineering Informatics*, April 2024.

Hao Wang initiated and wrote the paper with other authors' contributions and reviews. The research design as well as data collection and analysis were conducted together with Omkar Salunkhe and Walter Quadrini.

[**Paper 3**] **Deep Learning-Based Connector Detection for Robotized Assembly of Automotive Wire Harnesses**

Hao Wang, Björn Johansson

Presented at IEEE 19th International Conference on Automation Science and Engineering (CASE 2023), Auckland, New Zealand, 26-30 August 2023.

Published in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, 2023.

`https://doi.org/10.1109/CASE56687.2023.10260619`

Hao Wang initiated and wrote the paper, collected and analyzed the dataset, and designed and conducted the experiments, with Björn Johansson's contribution and review.

Other Publications

The following publications were published during my PhD studies, or are currently in submission/under revision. However, they are not appended to this thesis due to contents not sufficiently related to the thesis.

[**Paper a**] **Manufacturing Challenges and Opportunities for Sustainable Battery Life Cycles**

Björn Johansson, Mélanie Despeisse, Jon Bokrantz, Greta Braun, Huizhong Cao, Arpita Chari, Qi Fang, Clarissa A. González Chávez, Anders Skoogh, Henrik Söderlund, **Hao Wang**, Kristina Wärmefjord, Lars Nyborg, Jinhua Sun, Roland Örtengren, Kelsea Schumacher, Laura Espinal, Katherine Morris, Jason Nunley, Yusuke Kishita, Yasushi Umeda, Federica Acerbi, Marta Pinzone, Hanna Persson, Sophie Charpentier, Kristina Edström, Daniel Brandell, Maheshwaran Gopalakrishnan, Hossein Rahnama, Lena Abrahamsson, Anna Öhrwall Rönnbäck, Johan Stahre

Under review by a journal, March 2024.

Hao Wang contributed to the content related to computer vision techniques for automation and human-robot collaboration in battery manufacturing with other authors' contributions, content, and reviews.

[**Paper b**] **Review of Current Status and Future Directions for Collaborative and Semi-Automated Automotive Wire Harnesses Assembly**

Omkar Salunkhe, Walter Quadrini, **Hao Wang**, Johan Stahre, David Romero, Luca Fumagalli, Dan Lämkull

Presented at the 56th CIRP Conference on Manufacturing Systems (CIRP CMS 2023), Cape Town, South Africa, 24-26 October 2023.

Published in *Procedia CIRP*, 2023.

`https://doi.org/10.1016/j.procir.2023.09.061`

Hao Wang contributed to the content related to computer vision applications in automated automotive wire harness assembly and conducted data collection and analysis.

[**Paper c**] **Battery Production Systems: State of the Art and Future Developments**

Mélanie Despeisse, Björn Johansson, Jon Bokrantz, Greta Braun, Arpita Chari, Xiaoxia Chen, Qi Fang, Clarissa A. González Chávez, Anders Skoogh, Johan Stahre, Ninan Theradapuzha Mathew, Ebru Turanoglu Bekar, **Hao Wang**, Roland Örtengren

Presented at IFIP International Conference on Advances in Production Management Systems (APMS 2023), Trondheim, Norway, 17-21 September 2023.

Published in *Advances in Production Management Systems. Production Management Systems for Responsible Manufacturing, Service, and Logistics Futures*, 2023.

`https://doi.org/10.1007/978-3-031-43688-8_36`

Hao Wang contributed to the content related to computer vision techniques for automation and human-robot collaboration in battery manufacturing with other authors' contributions, content, and reviews.

Part I

Summary

Chapter 1

Introduction

This chapter begins with the research background and the core problem investigated. Then, the vision and aim of the research are elaborated, followed by the formulation of the research questions. Further, this chapter sets the scope and delimitation of the research. Lastly, the thesis outline is delineated.

1.1 Background

The advent of automation and robotics revolutionized the overall industry. In the third industrial revolution, the development in electronics and information technology promoted the massive adaptation of automation in different sectors of the modern industry significantly [1]. Implementing automation can better quality [2], promote productivity [3], optimize resource allocation [2], and improve working conditions [4]. The development in electronics and information technology also laid the foundation of robotics and has propelled robots to become a critical role in human life [5]. Specifically, industry has witnessed a remarkable expansion of robotic automation applications over the years [6]. Since its birth in the 1950s, industrial robots have become a significant enabler of automation and have changed the manufacturing industry radically [7].

Despite the extensive usage of industrial robots, there are still places in the manufacturing industry where a higher degree of robotic automation is desired but still needs to be achieved. As a representative area, the final assembly has long been anticipated to reach the automation rate between 35% to 75% of overall final assembly operations [8]. However, a high degree of manual operations remains common in contemporary assembly [9], albeit the degree of automation is rising [4]. The high complexity of assembly processes constrains the automation of the final assembly [10]. Human-based teamwork organizations have demonstrated superiority over machines in flexibility [11] and efficiency [12] for assembly tasks. Nonetheless, the global demographic change urges industry to pursue solutions to compensate for the potential shortage of workforce [12]. On the other hand, this high degree of manual operations causes production problems concerning the business aspect (e.g., quality and productivity [12]) and the human aspect (e.g., safety and ergonomics [13]). Hence, more automation is desired, especially considering the consistent and increasing pursuit of improvement in quality, efficiency, and sustainability in society [14], [15].

Currently, industry is exploring more automation solutions to address these remaining issues in production [16]. More industrial robots are expected to be deployed to enable and facilitate automation [6]. Meanwhile, industry needs industrial robots with more autonomy to handle assembly tasks that are even more challenging [17]. Particularly in the final assembly, the ongoing industrial paradigm shift from mass production to mass customization and personalization leads to small batch size and considerable variations in the assembly line [18]. This shift stimulates the transformation toward more flexible automation to fulfill increasingly complex manufacturing tasks and the enlarging market of customized products [19]. Numerous tasks require robots to generalize their skills to adapt to specific task scenarios and handle various activities [7]. However, it is infeasible for conventional industrial robots in current production to fulfill this requirement on generalization due to their lack of autonomy [10]. Recent robotics research has investigated human-robot collaboration (HRC) and its applications in industrial tasks to exploit the combination of robots' advantages on repeatability, accuracy, and physical strength and humans' superiority in cognitive abilities and flexibility [20]. Nevertheless, intelligent robots are needed to understand the surroundings in unstructured environments [16], especially for human-centered robot applications [21]. Therefore, industrial robots need advanced cognitive abilities to be competent at flexible automation tasks in the new era [1].

Perception is a prerequisite to enabling intelligent robots with cognitive abilities [22]. It consists of obtaining sensory input and interpreting it meaningfully [23]. Robotics has long been considered a discipline that studies "the intelligent connection of perception to action" [24]. A robot with advanced cognitive abilities can be regarded as a specific intelligent robotic agent that can perceive its environment through sensors and react to that environment through actuators [25]. As one of the fundamental components in robotic manipulation, knowing the position and orientation of an object is the premise of accomplishing the following manipulation operations upon the object. The autonomous recognition of positions and orientations of objects of interest is especially critical when product variants are involved in assembly tasks and/or the positions and orientations of objects of interest are difficult or impossible to define by humans in advance [26]. Therefore, enabling industrial robots to identify the parts to be assembled autonomously is significant to facilitating the robotization of complex assembly tasks [27].

Among other sensory inputs, vision is instrumental for recognizing objects and obtaining their positions and orientations [5]. Visual machine perception is one of the significant tasks studied in artificial intelligence (AI) and robotics, where many computer vision techniques have been adapted [22]. Previous research in computer vision, AI, and robotics has discussed various solutions for enabling robotic visual perception [28] and indicated the potential to promote intelligent robotic automation toward enabling flexible automation [1] and intelligent manufacturing [29]. Nevertheless, research on robotic visual perception for assembly tasks remains primary in laboratory scenarios [30]. Further research is required to reveal the challenges of enabling robotic visual perception for assembly tasks and prospective technical solutions to bring intelligent robotic assembly to fruition.

1.2 Problem Description

This thesis investigates the problem of enabling robotic visual perception to automate wire harness assembly in automobiles' final assembly. Specifically, this thesis explores the challenges and potential technical solutions to enabling robotic visual perception. Enabling visual perception capabilities will contribute to increasing industrial robots' autonomy and making them competent at wire harness assembly. With this, this thesis can provide insights into problems that should be addressed in academia. This thesis may also help industry decision-makers understand the potential challenges and opportunities of promoting robotization in the final assembly. In the long term, theoretical research can be realized satisfactorily in production.

Robotizing all or part of the assembly operations of wire harnesses in the final assembly is desired to address problems in production due to quality, efficiency, safety, ergonomics, and demographic change. It is essential to guarantee a high-quality installation of wire harnesses onto vehicles because: 1) they are fundamental elements within an automotive electronic system; 2) they are widely distributed in automobiles; and 3) they are responsible for quality-essential and safety-critical functions of automobiles [31]. The manual assembly in the current production of automobiles makes it challenging to guarantee consistent assembly quality [32]. It is also fundamental to assure the efficiency of wire harness assembly. On the one hand, the automotive industry persists in a continuous demand for productivity to promote competitiveness and acquire market share. On the other hand, the usage of wire harnesses in modern vehicles has been enlarging remarkably over the years, and industry expects the continued growth of wire harnesses installed in future automobiles [33]. The current manual assembly has been identified as one of the significant bottlenecks of automobile production promotion [32], [34], [35]. Moreover, it is crucial to ensure safety and improve ergonomics for human operators when assembling wire harnesses. Some manual assembly procedures could be more ergonomic for human operators, such as heavy lifting, high-pressure pressing, far-reaching operation, and repetitive movements [36]. These operations can cause severe musculoskeletal disorders and occupational safety and health issues in the workforce [37]. There are also high-voltage wire harnesses installed in automobiles, especially in electric vehicles (EVs), which demands more careful object handling regarding safety, assembly quality, and reliability [38], [39]. In addition, previous research has indicated a potential shortage of either skilled or unskilled workforce willing to work in automotive factories [12]. Therefore, assuring assembly quality and safety and promoting productivity while improving ergonomics and optimizing resource utilization is desired [36]. Implementing robotic assembly automation is one of the prominent approaches [7].

Automotive wire harness assembly has remained manual over the years and is challenging to automate mainly due to the high complexity of assembly processes [12]. This high complexity stems from various reasons, e.g., the considerable product variants due to the shift toward mass customization and personalization [14], the mix of rigid and non-rigid wire harness components and the deformation of wire harnesses [40], the limited process time in practical production [41], and safety concerns on robot deployment [42]–[44]. Among them, the considerable product variants and the deformation of wire harnesses make it unwieldy to program industrial robots in advance by hand. Industrial robots need to be able to perceive the whole or part of the assembly task and figure out their movements autonomously to be competent at either fully or semi-automated wire harness assembly [36], [40]. Visual perception is a fundamental contactless approach for robots to extract information from the surrounding environment [5]. However, in industrial applications, vision-based robotic assembly of wire harnesses has yet to succeed [40]. Enabling robotic visual perception is, thus, an important aspect to investigate when automating wire harness assembly.

1.3 Vision and Aim

This thesis envisions a sustainable manufacturing industry where robots are intelligent and cognizant of their tasks, the surrounding environment, and the humans nearby. With such intelligence, robots can handle all tasks that are either non-value-adding or not ergonomic to human operators. Robots can also adapt and react flexibly to tasks and situations with minimum human intervention requirements. Realizing and applying such intelligent robots will contribute to the symbiosis of humans and robots toward highly efficient production without problems regarding quality, safety, and ergonomics.

Toward such a vision, this thesis aims to facilitate enabling robotic visual perception to promote the degree of autonomy of industrial robots. With visual machine perception enabled, industrial robots can be improved to achieve higher levels of autonomy to handle more robotic manipulation required in flexible automation applications. With a better robotic perception, a robot can adapt and react to non-predefined situations and accomplish new tasks under unstructured physical configurations in the final assembly.

1.4 Research Questions

Although the significance of visual machine perception for increasing robotic autonomy has been recognized in previous research and by industry, the extensive application of vision-based robotic assembly has yet to succeed in practice [45]. To promote industrial robots' autonomy and competence in robotic assembly, this thesis investigates the aspect of visual machine perception and explores potential solutions to enabling robotic visual perception for assembly tasks.

The ultimate goal of this thesis is to suggest technical solutions to enabling robotic visual perception for assembly tasks. Nevertheless, it is necessary to understand the challenges of enabling robotic visual perception before solutions can be suggested and tested. Hence, the first research question (RQ1) is formulated to discover the challenges and indicate prospective directions for the following studies on technical solutions:

RQ1: What are the challenges of enabling robotic visual perception for assembly tasks?

This thesis intends to answer this research question by providing an overview of the challenges of adapting computer vision techniques for the robotic assembly of deformable objects in the final assembly.

With the challenges understood, the next step is to explore the opportunities for potential research and identify the prospective solutions to enabling robotic visual perception toward more autonomy on industrial robots. Research is needed to reveal how the identified challenges can be addressed and to study prospective technical solutions for enabling robotic visual perception, which would increase industrial robots' autonomy and competence in robotic assembly tasks. This leads to the second research question (RQ2):

RQ2: How can robotic visual perception be enabled for assembly tasks?

This research question is formulated to identify potential research opportunities and explore prospective vision-based approaches for promoting industrial robots' autonomy and competence in handling deformable objects in assembly tasks.

1.5 Scope and Delimitation

This thesis's topic lies at the intersection of automation, robotics, computer vision, and AI, particularly in the context of the final assembly. The scope of this thesis is meticulously defined, focusing on understanding the challenges and exploring potential opportunities and technical solutions for enabling robotic visual perception. This research aims to enhance the autonomy of industrial robots in robotic assembly. Though targeting to facilitate the overall manufacturing industry, the application area that this thesis researched is the final assembly in the automotive industry. This thesis explores robotic visual perception. It delves into potential technical solutions based on existing theoretical research in computer vision, AI, and robotics. In addition, this thesis only considers the technical aspect on the experimental level under simplified laboratory configurations. Thus, the delimitation of this thesis is defined as follows.

This thesis does not evaluate the extent of the improvement in robotic autonomy nor discusses the metrics for such an evaluation. This thesis does not analyze the practicality of potential technical solutions for enabling robotic visual perception for applications in actual production. This thesis does not investigate human-robot collaboration. Though critical for practical application, the human factors that may affect the application of technologies, such as the organizational culture and employees' attitudes, are also not studied in this thesis.

1.6 Thesis Outline

This thesis is organized in the following structure.

Chapter 1 Introduction delineates the background of the research in this thesis and the core problem focused in this research. The vision and aim are described, followed by the research questions of this research. Lastly, the scope and delimitation of this research and the thesis outline are explained.

Chapter 2 Theoretical Framework describes the theoretical framework of this thesis, including the background knowledge on robotic assembly and robotic perception and the state of the art of robotized wire harness assembly.

Chapter 3 Research Approach elaborates on the design of the research approach of this research, including the adopted philosophical worldview, research design, and research methods, followed by the methods adopted for guaranteeing the research quality.

Chapter 4 Summary of Appended Papers briefly summarizes each of the three appended papers, including each paper's core problem, methodology, and contribution. This chapter concludes with a summary of each appended paper's contribution to each research question.

Chapter 5 Discussion discusses the main findings in each study of this research and the answers to the research questions formulated in Section 1.4 in Chapter 1. Then, this chapter analyzes the contribution of this thesis to academia and industry. Reflections upon the research follow regarding the limitations of this thesis, the research methodology, ethics, and sustainability. Prospective research for the future is envisaged at the end of this chapter.

Chapter 6 Conclusion recapitulates the findings and contributions of this thesis.

Chapter 2

Theoretical Framework

This chapter describes the theoretical framework of this thesis. First, this chapter discusses robotic assembly, the background information and key concepts relevant to the context of this research, concerning assembly automation and robotic automation. Then, this chapter introduces robotic visual perception by elaborating the background information and related research on robotic perception, computer vision, and artificial intelligence. Lastly, this chapter summarizes the state-of-the-art research on robotized wire harness assembly.

2.1 Robotic Assembly

Robotic assembly is an essential application of robotic automation in assembly toward assembly automation. It is also the context of the research in this thesis. Hence, this section introduces the background knowledge and specifies terms relevant to robotic assembly.

2.1.1 Assembly Automation

Automation

Automation, a concept that has undergone diverse definitions in various contexts, holds immense significance in the field of production engineering. Linguistically, *automation* is defined as "the action or process of introducing automatic equipment or devices into a manufacturing or other process or facility; (also) the fact of making something (as a system, device, etc.) automatic" in Oxford English Dictionary [46]. Practically for research and industrial applications, *automation* is defined as "the conversion of a procedure, a process, or equipment to an automatic operation without intervention by a human operator" by the International Institution for Production Engineering Research (CIRP) [47] and as "the creation and application of technology to monitor and control the production and delivery of products and services" by International Society of Automation[1]. These definitions of automation underscore its potential to liberate the workforce from specific production scenarios by significantly reducing the need for human control or intervention.

Automation research in production has been conducted regarding physical and cognitive automation [48]. While physical automation aims more to automate physical operations [49], [50], cognitive automation targets more to automate the cognitive processes that are currently performed by human operators [51]. This thesis is dedicated to exploring technical solutions that can revolutionize manual assembly operations in current production, thereby emphasizing the practical application of physical automation.

[1]`https://www.isa.org/about-isa/what-is-automation`

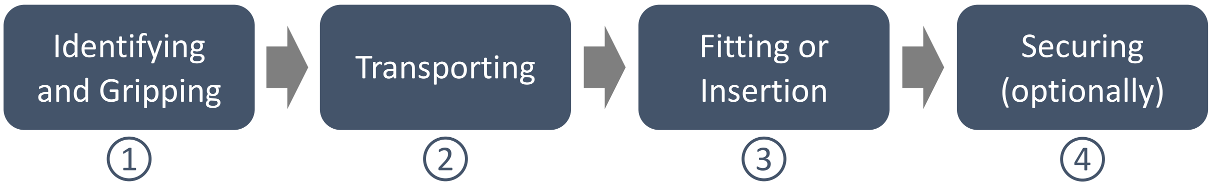


Figure 2.1: Sub-procedures consisted in assembly tasks, adapted from Lien [26].

Assembly

Assembly is a critical constituent part of production and has been researched from two major perspectives: 1) assembly as process and 2) assembly as product [52]. This thesis investigated problems in assembly regarding the former point of view, and particularly the final assembly in the automotive industry, where all sub-parts or units (e.g., engines, body frames, wire harnesses, glasses, and wheels) are fitted together to form the final product (automobiles) of production (the automotive production). In this scenario, as shown in Figure 2.1, the workflow of assembly tasks can be divided into several sub-procedures: 1) part identification and gripping; 2) part transportation to target assembly position; 3) part fitting or insertion; and 4) part securing (optionally) [26].

As the process where manufactured parts are fitted together into a complete product of any kind, the assembly can be performed manually [26], semi-automated [53], or fully automated [54]. The choice of assembly method is affected by multiple factors, such as the product design, the required production rate, the availability of labor, and the market life of the product [55].

Automation in Assembly

In assembly research, assembly automation describes introducing automatic machines to convert manual assembly operations into ones without needing human controls [54]. Assembly automation is desired and of benefit to both business (e.g., higher quality, more optimized resource allocation, and higher efficiency [2], [3]) and human operators (e.g., safer working space and better ergonomics [4], [56]). However, the scale of automation in assembly remains limited, especially in the final assembly [9]. From the perspective of assembly tasks, the complexity of assembly operations inhibits the automation of assembly [57]. From the perspective of technical solutions, assembly automation is constrained due to the inability of automatic systems to handle complex assembly tasks and the insufficient flexibility to handle product variants [12].

2.1.2 Robotic Automation

Robotics

Robotics is one of the subjects boosted by the research and application of automation [58]. Through research over decades, robotics has become an interdisciplinary subject studying the science and technology related to robots and similar automatic devices, including the design, construction, operation, and usage [59]. Already in the early stage of robotics research, Brady [24] had indicated the significance of concerning robotics as “the intelligent connection of perception to action”, which persists in the core of contemporary robotics research [60].

Industrial Robots

Industrial robots have been one of the core interests of robotics since the beginning of robotics research [61]. A robot, in general, can be seen as a physical agent that manipulates the physical world with equipped effectors based on the environmental information perceived through equipped sensors [62]. Notably, an *industrial robot* can be defined as "an automatically controlled, reprogrammable multipurpose manipulator programmable in three or more axes" [63]. In the context of production, "industrial robot" and "robot" are used interchangeably by convention [63], which is also adopted in this thesis hereinafter except where otherwise stated.

The first industrial robot is widely acknowledged to be the Unimate from Unimation, founded by George Devol and Joseph Engelberger in the late 1950s [64]. In 1961, the Unimate was first deployed industrially for unloading the finished castings in a General Motors plant in Trenton, New Jersey, the United States of America [64]. Since then, industrial robots have become significant in factory automation and radically changed the manufacturing industry [60]. Among other automation technologies, robots are widely adopted in the modern manufacturing industry due to their superiority in conducting repetitive and unergonomic tasks fast and precisely [7], [20]. Commonplace industrial application scenarios of industrial robots include spot welding, spray painting, part handling, packaging, and palletizing [63]. Industrial robots are also applied to automate assembly tasks, which, though highly favored [17], remains a small portion of the robotic automation application [10].

Parallel to the upsurge of industrial robot applications promoting industry automation, academia is deepening research to continuously improve robots' capabilities. While previous research on industrial robots focused more on the aspect of robotic action and mainly investigated kinematic calibration, motion planning, and control laws, the primary research interest has veered toward the aspect of intelligence to improve the flexibility and enhance the autonomy of industrial robots [61]. Siciliano and Khatib [60] projects that "the new generation of robots is expected to safely and dependably co-habitat with humans in homes, workplaces, and communities, providing support in services, entertainment, education, healthcare, manufacturing, and assistance". Such new-generation robots need to not only succeed in their actions but also be capable of perceiving the environment, learning, and reasoning their choices of actions [5].

Strengths and Weaknesses of Conventional Industrial Robots for Assembly

Compared to entirely manual operations, robotic assembly's better precision, repeatability, transparency, and comprehensibility can enable more rigorous, safer, and more ergonomic-friendly manufacturing with better quality and higher productivity [20]. The increase in their degrees of freedom and payload promotes industrial robots to take over tasks harmful to humans from human operators, e.g., operations in dirty and dangerous work environment [65], repetitive operations [5], demanding and tedious operations [66], and unergonomic operations [20].

However, conventional industrial robots are only superior in repetitive and familiar industrial configurations but brittle when the assembly process involves increasing product variants and/or requires more flexibility in unstructured environments [5]. There are also limitations to expanding the usage of conventional industrial robots in assembly, such as the complexity and flexibility of assembly [10], [67] and the safety of humans [68]. Typically based on scripted trajectory planning, conventional industrial robots can only accomplish simple tasks with monotonous operations in structured scenarios, where robot programmers must specify the positions and orientations of objects in advance [69]. However, with the increasing complexity of assembly tasks and the industrial paradigm shift toward mass customization, contemporary assembly systems handle increasingly more product variants [18]. Particular assembly tasks also involve manipulating non-rigid objects, whose positions and orientations are impossible to pre-define due to their deformation with almost infinite degrees of freedom [70]. The increasingly demanding production requirements demand industrial robots with higher flexibility potentials and a higher degree of autonomy [14]. Thus, this situation has steered the robotics research toward developing adaptive and intelligent systems [61].

2.2 Robotic Visual Perception

The manufacturing industry has witnessed the significant promotion of the application of automation and the adoption of industrial robots since the third industrial revolution [1]. Nevertheless, the contemporary manufacturing industry calls for intelligent industrial robots to fulfill emerging production requirements of flexible robotic automation [61]. A critical task is to improve the capability of robots to reach a higher degree of autonomy so that robots can handle unstructured tasks in an unstructured environment [60].

Russell and Norvig [25] defines an agent as "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators". An industrial robot able to adapt to variations can be considered an intelligent agent. It perceives the environment through equipped sensors, interprets the perceived information with intelligent systems, and then interacts with the surrounding environment via end effectors. Generally, an intelligent industrial robot can only select its actions at any time based on its embedded knowledge and the information perceived to date [25]. Robots can acquire knowledge through either hard coding by humans or learning based on self-perceived information [71].

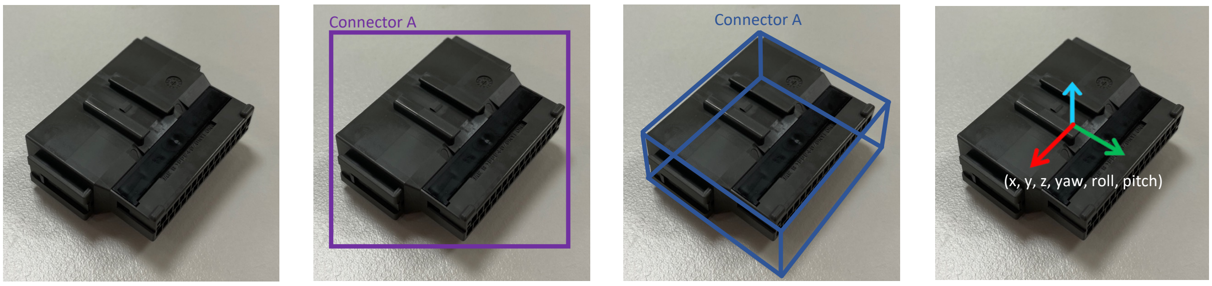
2.2.1 Robotic Perception

Perception is a premise of intelligent agents of any kind [25]. Russell and Norvig [62] defines *robotic perception* as "the process by which robots map sensor measurements into internal representations of the environment". Regarding different types of sensory data input, robot systems can perceive the outer environment typically through visual, range, force/torque, and tactile perception [59]. Among approaches based on sensory data input, force/torque-based approaches are critical for robotic control at a low level, while recognition, measurement, and learning of robots at a higher level heavily rely on visual and range data-based perception [72].

Visual perception indicates the organization, identification, and interpretation of information from visual inputs, which has been extensively attractive in the academic field and industrial applications [73]. Specifically for robotic assembly, a robot needs to recognize *what* the object to be manipulated is and localize *where* the object is, among acquiring other physical properties of the object, so that the robot can reach, grasp, and manipulate the object to accomplish the assembly task. Vision is a fundamental source of information for object recognition and localization [5]. The vision system is one of the most common sensory systems integrated into robots to enable automatic operations [63].

2.2.2 Computer Vision

Visual perception is one of the significant topics studied within computer vision research [74]. Computer vision is also extensively related to automation, robotics, and manufacturing [45].



(a) Raw input (b) 2DOD (c) 3DOD (d) 6DOPE

Figure 2.2: Expected results of 2D object detection (2DOD), 3D object detection (3DOD), and 6D object pose estimation (6DOPE) regarding an RGB image of a connector.

Computer vision has a dual goal: 1) from the perspective of biological science, developing computational models of the human visual system; 2) from the perspective of engineering, developing autonomous systems capable of tasks performed by human visual system or even superior [75].

In alignment with the engineering perspective, this thesis investigates the challenges and solutions to developing robotic visual perception systems to extract useful information from visual inputs [76]. The robot system with a vision system will be able to convert the perceived visual information of a scene into a symbolic description [77]. With this description, the robot can understand the scene and decide the next operations [45]. As a thriving field, computer vision research encompasses diverse topics, where two-dimensional (2D) object detection, three-dimensional (3D) object detection, and six-degree-of-freedom (6D) object pose estimation are closely related to robotics and robotic assembly [72]. Figure 2.2 illustrates the expected results of these three tasks regarding a visual input.

2D Object Detection

In 2D space, object recognition implies image classification, object localization, and object detection. Image classification deals with classifying the principal object in an image, involving assigning a class label to an image [78], [79]. Object localization deals with locating object instance(s) of a given category in an image, involving drawing a bounding box around one or more objects in an image [80]. Object detection is a process of image classification and object localization [81]. As shown in Figure 2.2(b), object detection involves drawing a bounding box around each object instance of interest in the image (localization) and assigning each localized object instance a class label (classification) [82], i.e., recognizing what objects are where [83]. The year of 2014 marked a milestone in the research of 2D object detection with the advent of R-CNN [84]. It symbolized the division of research and breakthroughs into two historical periods: the traditional object detection period and the deep learning-based object detection period [85].

Traditional methods were designed based on handcrafted features and various feature descriptors [83]. The pipeline of traditional object detection methods generally comprises three steps: 1) informative region proposal, 2) feature extraction, and 3) classification and bounding box regression [82]. Traditional methods required the design of sophisticated feature representations and diverse speedup skills due to the lack of effective image representation at that time [85]. Classic traditional detectors include Viola–Jones detector [86], [87], Histogram of Oriented Gradients (HOG) detector [88], and deformable part model (DPM) [89]. However, traditional detection methods have significant flaws, e.g., slow speed, low accuracy, arduous manual feature engineering, and poor generalizability, which have gradually been replaced by deep learning-based methods [81].

Deep learning-based methods with the structure of deep convolutional neural networks (CNNs) dominate the latest research on 2D object detection [81]. Deep learning-based detectors can be classified into two major groups: two-stage detectors and one-stage detectors [85]. Motivated by the attentional mechanism of the human brain, two-stage detectors first scan the whole scenario coarsely and then focus on regions of interest (ROIs) to distinguish the object [83]. Quintessential two-stage detectors include the R-CNN family (R-CNN [84], [90], Fast R-CNN [91], and Faster R-CNN [92]), SPPNet [93], and FPN [94]. One-stage detectors were initiated to address the constraint of two-stage detectors on speed and computation [85]. Localization and classification are accomplished all at once through the backbone network. Typical one-stage detectors include YOLO [95] and its successors [96]–[99], SSD [100], RetinaNet [101], CornerNet [102], and CenterNet [103]. Recent research on Transformer models [104] also initiated the design of new detectors [105]. Typical Transformer-based detectors include DETR [106], Deformable DETR [107], DINO [108], and Mask DINO [109].

3D Object Detection

3D object detection is more challenging than 2D object detection. Different from 2D object detection, 3D object detection emphasizes the importance of recovering the amodal bounding box of the exact object instance, i.e., the minimum 3D bounding box enclosing the object of interest [110], as shown in Figure 2.2(c). Therefore, 3D object detection needs to obtain the size and direction of the object of interest in 3D space in addition to its position [111].

Previous research on 3D object detection explored diverse solutions regarding various types of input data, including RGB images [81], point clouds [112], RGB-D data [110], and multi-modal data [113]. Approaches based on RGB images can be divided into four categories: monocular-based, stereo-based, pseudo-LiDAR-based, and multi-view-based [81]. Approaches based on point clouds can be divided into two categories: region proposal-based and single-stage [112]. Approaches based on RGB-D data can be divided into two categories: 2D object detection-driven and data fusion-driven [110]. Approaches based on multi-modal data can be divided regarding the adopted data fusion method, e.g., early fusion, late fusion, and deep fusion [113].

6D Object Pose Estimation

Six-degree-of-freedom (6D) object pose estimation refers to determining the 6D pose of an object in 3D space. The 6D pose of an object is the combination of position, (x, y, z), and orientation, $(yaw, roll, pitch)$, of an object in 3D space [114]. Figure 2.2(d) illustrates an expected result of 6D pose estimation on a connector. Previous studies on 6D object pose estimation can be clustered into two primary settings: **instance-level 6D pose estimation** and **category-level 6D pose estimation**, regarding whether the computer-aided design (CAD) model of each object instance is a prerequisite [115].

Instance-level 6D pose estimation mainly addresses the pose estimation of objects whose 3D models are available [110]. Similar to research in object detection, research in instance-level 6D pose estimation can be divided into traditional methods and deep learning-based methods. Traditional methods address the problem based on CAD models [116] or 2D images synthesized from CAD models [117]. Regarding the modality of the input data, previous deep learning-based methods can be divided into three sub-groups: RGB-based, point cloud/depth-based, and RGB-D-based [115]. Research on RGB-based methods is extensive, thanks to the widespread use and affordability of RGB cameras. However, these RGB-based methods suffer from occlusion, illumination variations, objects without distinctive visual features, real-time performance, and generalizability [115]. Point cloud/depth-based methods process inputs in the format of point clouds or depth data acquired by 3D scanners or depth cameras. Point clouds or depth data include spatial information compared to RGB images, which facilitates the recovery of 3D information, especially for objects without distinctive texture features. However, this group of methods suffers from the expensive manual labeling of training data and heavy computing consumption, especially for point cloud-based methods. RGB-D-based methods consider RGB inputs and depth data jointly, which can promote the pose estimation performance in complex configurations with features extracted from both modalities of data [115]. However, it requires further investigation on efficient approaches for fusing features extracted from RGB data and depth data [115].

Category-level 6D pose estimation aims to achieve generalization to unseen instances when recovering the poses of objects [110]. Previous research efforts on category-level 6D pose estimation can also be divided into traditional methods and deep learning-based methods, in alignment with instance-level 6D pose estimation. Traditional methods address the problem based on images of natural objects but hardly remain in recent research efforts due to the time-consuming and challenging data collection [117]. Previous deep learning-based methods for category-level 6D pose estimation can be summarized into two major sub-groups: regression-based and prior-based [115]. Category-level pose estimation methods do not require accurate CAD models but are challenging due to the unavailability of ground truth data. Category-level pose estimation methods are also susceptible to the prominent appearance and/or shape variance across instances.

Computer Vision in Manufacturing

Computer vision techniques were already applied to industrial applications in the early 1970s [118] but remained scarcely commercialized in practical manufacturing until the 1990s due to the limitation of computing resource [119]. The research on industrial applications of computer vision techniques is also categorized as research in the field of **machine vision** [120] or industrial vision[2] [121]. While computer vision research is more methodology-oriented, machine vision research is more application-oriented and represents the particular implementation of computer vision for industrial purposes [120]. There are two major approaches to addressing industrial vision problems [122]–[124]. One is to address all problems with a general-purpose system, and the other is to design an ad-hoc system for each application scenario [121].

Throughout decades of development, computer vision techniques have become a vital booster of industrial manufacturing systems toward a higher level of informatization, digitalization, and intelligence [45]. In the manufacturing industry, the application of computer vision techniques can be classified regarding different criteria, e.g., application tasks [125] or stages of the product life cycle in the entire manufacturing process, including product design, modeling and simulation, planning and scheduling, production process, inspection and quality control, assembly, transportation, and disassembly [45].

As a critical stage in manufacturing, the assembly has drawn long-lasting attention to applying computer vision techniques [118]. Previous research mainly investigated applying computer vision techniques for automatic assembly, assembly quality control, and other assembly applications [45]. Researchers explored improving the performance of industrial robots in assembly tasks in unstructured environments using visual perception and learning techniques [27]. There are also research efforts on making industrial robots more adaptive to unknown scenarios using machine vision [126]. Nevertheless, machine

[2] Diverse industrial vision applications can be found on `https://www.cs.ubc.ca/~lowe/vision.html`

vision is mainly employed for quality-related purposes but is advocated to be expanded into assembly scenarios [127]. Promoting the adaptation of computer vision techniques in industrial applications remains challenging regarding computer vision algorithms, data, and benchmarks [45].

Particularly for assembly tasks, the accuracy of identifying the position and orientation of objects to be assembled is crucial to robotic grasping and the following manipulation operations. However, it is difficult to fulfill the demanding requirement with traditional computer vision techniques currently used in actual manufacturing systems [45]. On the other hand, it also remains challenging to identify objects' accurate position and orientation with existing deep learning-based methods [72]. The performance of visual recognition can also suffer from diverse problems in actual production environments, e.g., occlusions [128], illumination conditions [129], and the camera movement [130].

The dataset is essential for learning-based computer vision techniques [131] and scalable learning-based computer vision applications in manufacturing [132]. However, it is challenging to collect high-quality data in practical manufacturing scenarios and arduous to effectively preprocess and efficiently label the collected data [45].

Using benchmarks is critical to evaluating performance across different computer vision and robotic systems [133]. However, specific manufacturing cases require particular benchmarks for evaluating computer vision techniques in different scenarios [45].

There are also other concerns for the implementation of AI-driven computer vision systems in industry, e.g., the cost of implementing new systems in existing systems [45], industry's lack of trust in AI systems due to the lack of interpretability and explainability of the decision-making of AI [70] and the safety of human-robot collaboration [45].

2.2.3 Artificial Intelligence

The term *artificial intelligence* (AI) was first introduced in 1955 for proposing a workshop at Dartmouth College[3] to "proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" [134]. Throughout decades of research and development, AI has been evolving into a flourishing field encompassing diverse, active research topics and practical applications, ranging from general topics, e.g., learning and perception, to specific application scenarios, e.g., object detection, machine translation, weather forecasting, and robotics [22].

Machine Learning

One significant task in enabling machine intelligence is to equip machines with knowledge. Generally, machines acquire knowledge via either knowledge hard-coding by humans or self-learning [71]. Approaches based on hard-coded knowledge previously gained limited breakthroughs due to the inability to enumerate all scenarios or describe specific scenarios in formal languages [135]. This limitation suggested the significance of the machine's capability of self-learning. This capability is known as machine learning, with which machines observe data and extract patterns from the observed data [135].

Deep Learning

Deep learning is a broad family of techniques for machine learning and has significantly boosted the latest surge of public interest in AI [136]. The advent of deep learning radically

[3]`https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth`

reshaped the research and development in AI-relevant fields, e.g., computer vision, natural language processing, and robotics [137]. Deep learning enables computers to learn data representations with a hierarchy of abstraction via computational models consisting of multiple processing layers [71]. Adapting deep learning techniques can relieve the humans' burden of formally specifying all the knowledge required by computers in advance [71]. The genuine reasons for the success of deep learning remained obscured [137]. Nonetheless, deep learning-based approaches demonstrate their superiority over traditional approaches based on manually designed features, especially for tasks with high-dimensional input data, e.g., images, videos, and speech signals [136].

Convolutional Neural Network

Deep learning stemmed from early trials on mathematically modeling networks of neurons in the brain by McCulloch and Pitts [138], thus naming the networks trained by deep learning approaches neural networks [137]. Convolutional neural network (CNN) is "a specialized kind of neural network for processing data that has a known grid-like topology" [71]. CNN is characterized by using convolution operations instead of general matrix multiplication in at least one of their layers [139]. A typical CNN architecture usually consists of layers alternating between convolutional and pooling layers [140]. Since the early 1990s, CNN has achieved numerous practical successes and has been widely adopted in computer vision research, e.g., object detection, time series prediction, and human action recognition [141].

Artificial Intelligence in Manufacturing

Academia and industry have promisingly anticipated that the achievement in AI research will remarkably improve manufacturing systems in terms of productivity, quality, and profitability [142]. However, regardless of recent rapid advances in AI, previous research uncovered that "the application of AI technology in industry lags far behind the development of the AI technology" [143]. This gap promotes the research on AI applications, more specifically in industrial scenarios, also known as Industrial AI [143], aiming to address the problems and needs coming from industry. AI is applied or envisaged to be applied in diverse industrial application areas, e.g., supply chain management and production planning on the systems level, safe human-robot collaboration and robotic motion planning on the workstation level, and quality monitoring, tool wear prediction on the manufacturing process level [144]. For robotic assembly, AI has the potential to support different tasks, e.g., robotic perception, robotic motion, manipulation planning, and safety assurance [145].

2.3 Robotized Wire Harness Assembly

Wire harnesses can be theoretically generalized as deformable linear objects (DLOs) [70], [146], [147]. Wire harness assembly, thus, can be regarded as a specific task of DLO manipulation [148]. Challenges for automating wire harness assembly arise from product and production aspects. The development of robotic assembly of wire harnesses has attracted research over the years but has yet to succeed [36], [40]. This section elaborates on the research on robotic manipulation of DLO in general, the challenges for automating wire harness assembly, and the state-of-the-art research on robotic perception in the robotic assembly of wire harnesses.

2.3.1 Robotic Manipulation of Deformable Linear Objects

Objects manipulated by robots can be classified as rigid or non-rigid objects depending on shape-changing caused by an external physical force applied to objects [149]. Previous research also called non-rigid objects [70], [150] as deformable objects [149], [151], [152] or flexible materials [148]. Thus, in alignment with the existing research, this thesis uses these three terms interchangeably hereinafter except where otherwise stated.

Enabling robots to manipulate deformable objects can greatly benefit various application scenarios, e.g., flexible printed circuit board handling in manufacturing industry [153], tomato grasping in food industry [154], wound suturing in medical surgery [155], and garment manipulation in daily activity [156]. However, throughout decades of efforts in the robotics community, the study of non-rigid objects has yet to reach the compatible level of maturity of rigid object-centered research [149]. Notably, robotic manipulation of rigid objects primarily concerns the change of objects' pose (position and orientation) and the avoidance of collisions [157]. Differently, manipulation of non-rigid objects must further consider the shape-changing, which usually causes the changes in the objects' geometry and/or topology and lead to potential robotic failures [158]. Therefore, strategies designed for manipulating regular rigid objects cannot be adapted for deformable object manipulation directly [149].

Deformable objects can be more specifically classified into three groups: one-dimensional (or linear), two-dimensional, and three-dimensional (or volumetric) [148]. Further considering the physical properties, deformable two-dimensional objects can be separated into planar objects and cloth-like objects [149]. In line with the existing literature, wire harnesses can be categorized as deformable linear objects (DLOs). DLOs are also called deformable one-dimensional objects (DOOs) [159], [160]. Hence, wire harness assembly is a specific industrial application of DLO manipulation [148].

The DLO manipulation has also been a significant concern in industry over the years [70], [148], [149], e.g., the wire insertion in the electrical industry [161], [162] and the assembly of cables in the automotive industry [163], [164]. Various robotic tasks are involved in the robotic manipulation of DLOs, including modeling, perception, and manipulation [149]–[151]. Research has been conducted in investigating diverse deformable object models and integrating different sensors and AI into robots to endow robots with fast, accurate, and multi-modal perception capabilities [165] and adaptive modeling and control capabilities [151]. Regardless of the remarkable advancement of robotics, robotic manipulation of DLOs remains challenging in robotic flexible automation [11], [166], [167]. Specific challenges of accomplishing robotic DLO manipulation exist in object detection, deformation state estimation, object modeling, robotic motion planning, and robotic manipulation [149], [151], [152], [157], [167], [168].

Robotic perception is a prerequisite for accomplishing complex robotic manipulation tasks [165]. Particularly for robotic manipulation of DLOs, perceiving DLOs' physical properties, e.g., geometry, topology, deformation, and strain, before and during the robotic manipulation is required for modeling, motion planning, and manipulation planning [149], [152], [157]. Robotic perception for robotic manipulation of DLOs essentially involves perception based on independent or multi-modal sensing data, including visual, sound, force, tactile, and range data [70], [150]–[152]. Tactile perception is often adopted to obtain shape and contact information on the local level [152], while the global information of DLOs on a large scale, e.g., geometry, topology, and deformation, is often obtained via visual perception [169], [170].

2.3.2 Challenges for Automating Wire Harness Assembly

The automotive industry has implemented automation in assembly over the years to fulfill the increasingly demanding production requirements [14], [171]. While the body shop and the final assembly gather the majority of assembly operations, a remarkably higher level of automation has been achieved in the assembly of the body in white in the body shop than other assembly operations in the final assembly line [14]. Specifically, the installation of wire harnesses in the final assembly stage remains manual extensively and laborious to automate due to obstacles regarding the product and the production [36].

The customization and deformation of wire harnesses demand intelligent robots to handle flexible and agile automation tasks in the robotized assembly of wire harnesses. The automotive industry is consistently seeking solutions to robotize the overall or part of the assembly operations of wire harnesses in final assembly [36], [148]. However, no practical solution has been witnessed in actual production yet as equipment and technologies in the current production of automobiles are inadequate to accomplish the demanded flexible and agile automation tasks [7], [14], [54], [172]. Industrial robots deployed in current production are good at specialized tasks but weak at handling variants flexibly due to the lack of cognitive abilities [7]. Besides reducing the assembly complexity by simplifying the harness architecture [173], the robotization of wire harness assembly can be achieved by improving robots, i.e., making industrial robots more intelligent with more autonomy.

Industrial robots cannot handle variations due to limited perception and cognitive capabilities. For wire harness assembly, the deformability of wire harnesses further challenges the robotic perception [32], [34], [174]. This indicates that industrial robots need to have more autonomy and be more intelligent.

Challenges Regarding the Product

As a specific industrial application of DLO manipulation, the robotic assembly of wire harnesses inherits the challenges of robotic perception, modeling, and control in DLO manipulation. Robotizing wire harness assembly requires robotic systems to recognize the geometry and topology of wire harnesses, estimate the state of manipulation, and track the deformation so that they can model the wire harnesses and adapt their control strategies to handle the flexibility [174], [175]. Even if the robot achieves successful perception and modeling, the deformability of wire harnesses makes it complex to plan the robotic motion [175] and manipulation [157].

On the other hand, the robotic assembly of wire harnesses is more challenging than the generic robotic manipulation of DLOs. A bunch of wire harnesses is more complex than generic DLOs, considering its tree-like structure [176]. With multiple DLOs bound into bundles, it is necessary to address the interaction and constraints among different branches of wire harnesses while manipulating them. Besides the deformable cables, wire harnesses consist of rigid objects, e.g., connectors and clamps. This difference has also been identified in previous research, where wire harnesses were categorized more specifically as semi-deformable linear objects (SDLOs) [166], branched deformable linear objects (BDLOs) [32], [34], [177] or DLO networks (DLONs) [178]. Even with the mature research in robotic manipulation of rigid objects, the physical properties of some rigid wire harness components, e.g., the small sizes and the complex structures, exacerbate the arduous robotization of wire harness assembly [179].

Additionally, in mass customization, multiple variants of products are commonly produced on the same production line. This situation causes the wire harnesses installed onto each vehicle to be different. This increases the complexity of automation system design and challenges its adaptiveness and agility regarding different product variants.

Challenges Regarding the Production

To be deployed in actual production, technologies need to address challenges stemming from the actual production environments.

First, the proposed automation solution needs to be effective. However, Jiang, Nagaoka, Ishii *et al.* [180] indicated the challenge of the effectiveness of many proposals in actual production due to the required extremely tight position accuracy in assembly operations and the lack of precise contactless measurement to the state of the target wire in real-time. The actual production environment also challenges the effectiveness of proposals. The proposed system also needs to be reliable and robust in industrial environments. Extremely tight position accuracy in some assembly operations requires precise robotic perception of the object to be manipulated in real-time, which challenges the reliability of robotic perception capability in actual production [180]. The diverse and dynamic physical environments in actual production further challenge the robustness of the automation system. The automation systems also need to function reliably and efficiently in actual production to fulfill the demanded product rate and maintain the manufacturer's competitiveness.

Introducing new robotic systems brings challenges to safety and risk management. Physical equipment, such as steel fences and laser curtains, is typically required in industrial robotic applications to safeguard human operators [42], [43]. The growing applications of human-robot collaboration also require careful consideration of the safety aspect [44]. With new robotic systems introduced in actual production, a systematic re-design of the workspace and the human-robot interaction may be necessary, which poses challenges to safety and risk management within the existing system.

Robotic systems deployed in the final assembly stage need to deal with moving assembly lines. The final assembly lines in the automotive industry are typically non-stop, which requires robots to move in synchronization with the moving assembly line while executing assembly operations [10]. The mobility of robots and the synchronization between robots and assembly lines pose a challenge to the development of robotic assembly solutions.

Additionally, developing a universal solution is challenging due to the diverse production requirements across different productions and sectors, even within the same industry. In the automotive industry, for example, the production requirements of passenger, heavy, and special vehicles are different in terms of the physical production environments, required production quality, and production rate. These different production requirements set different criteria for developing automation solutions in terms of effectiveness, efficiency, reliability, and robustness. They also demand heterogeneous solutions, which increases the workload of automation solution development.

2.3.3 Robotic Perception in Robotic Assembly of Wire Harnesses

Robotic perception is pivotal to the robotization of wire harness assembly. Preliminary to accomplishing robotized wire harness assembly, robots demand advanced perception capability to recognize different wire harnesses and obtain their position and movement. Previous research efforts studied different types of perception based on diverse sensing data to perceive the physical properties of wire harnesses [36], [40], [148]. Visual perception is a critical contactless measurement to perceive the global information of wire harnesses on a large scale [169], [170].

Vision is instrumental for object localization, classification, and tracking [5]. The vast amount of information embedded in visual input [181], [182], the significance of computer vision on the perception for robotic manipulation [183], and numerous applications in different manufacturing scenarios [45] demonstrate the potential of applying computer vision techniques to facilitate robotized wire harness assembly. Previous research suggested the promising performance of vision-based approaches in robotic manipulation of DLOs [169], [170], [184] However, the application of vision-based robotized assembly of wire harnesses in actual production has not succeeded yet [45], [185], [186]. The vision-based robotization of wire harness assembly has drawn enduring attention and research effort [163], [178], but the task remains challenging to accomplish in actual production [34].

Besides, previous research has explored obtaining the physical properties and contact information of wire harnesses on the local level based on tactile [180], [187] and sound data [188]. Though also critical, this part of the research is out of the scope of this thesis.

Chapter 3

Research Approach

This chapter elaborates on the design of the research approach of this thesis. The philosophical worldview of the author of this thesis is analyzed first, followed by the research design and methods adapted in this thesis. The method to guarantee the research quality is elaborated in the end.

3.1 Philosophical Worldview

Though concealed mainly in research, philosophical worldviews influence the research practice and need to be identified [189], [190]. Representing researchers' fundamental beliefs about the nature of knowledge, reality, and human behavior, philosophical worldviews guide researchers' actions in determining research approaches and conducting studies [190]–[192]. Revealing the espoused philosophical worldviews can not only facilitate researchers elaborating the reason behind choices on specific research approaches but also help readers better interpret the research with a clearer mind on the biases and the researcher's particular stance [190]. Multiple factors contribute to developing individuals' worldviews, e.g., externally, discipline orientations, research communities, advisors, mentors, and internally, personal education background and experiences in culture and research [193].

The education and research experiences in electrical engineering, computer science, artificial intelligence, and computer vision formulated the author of this thesis's gravitation on the empirical postpositivist worldview. Researchers with positivist worldviews deem the existence of the absolute truth of knowledge [194]. Positivists aim to determine the connection between cause and effect [195]. They coincide with empiricists when believing sense experience is indispensable with the inquiry of knowledge [196]. The postpositivist worldview complements the positivist worldview by stressing the reflection upon the potential personal bias of a researcher's claim of knowledge [197]. With the empirical postpositivist worldview, the author of this thesis focused more on studying problems related to the identification and evaluation of causes of particular effects through empirical observation [198]. Specifically, the author initiated the research in this thesis from the hypothesis that applying computer vision techniques will enable robotic visual perception on industrial robots for assembly tasks. The author needed to identify and assess, based on empirical evidence, the causes that may influence the effectiveness of applying computer vision techniques.

Nonetheless, researchers are seldom limited to one worldview and may involve different worldviews in parts of the research regarding the discipline orientations and research communities [190]. Before beginning the research, the author of this thesis realized the necessity of identifying remaining research problems and specifying research objectives based on an understanding of the state of the art of research. Constructivists aim to forge a theory inductively based on individual participants' view of a specific scenario being studied [199]. Thus, a constructivist worldview was also possessed to gain such knowledge from scientific literature.

3.2 Research Design

Within research approaches, research designs indicate specific directions for conducting procedures of inquiries [190]. The specific research design in a study is also a guideline for executing specific research methods to collect, analyze, and interpret data [200]. As elaborated in Section 3.1, the research in this thesis intended to identify problems, define objectives, and investigate technical solutions in sequence. Therefore, the author of this thesis recognized design science research methodology [201], [202] as the guideline for the high-level design of this research. This research also adopted a multiple-method design [203] encompassing two studies. The first study was a qualitative literature study with a qualitative descriptive design, mainly aiming to identify problems and define objectives. The second study was a quantitative experimental study with a quantitative experimental design, mainly aiming to investigate technical solutions.

Design Science Research Methodology

As defined in Hevner and Chatterjee [202], design science research (DSR) is a research paradigm where researchers, as designers, contribute scientific knowledge by creating innovative artifacts that are both useful and fundamental to understanding and addressing human problems. An artifact can be anything that can contribute to the transformation from the current state to the desired one, e.g., constructs, models, methods, and instantiations [204], [205]. Peffers, Tuunanen, Rothenberger *et al.* [201] proposed the design science research methodology (DSRM) comprising the following six steps:

1. Problem identification and motivation (Defining a problem and justifying the need)
2. Objective definition (Defining objectives for a solution)
3. Design and development (Designing and developing a solution)
4. Demonstration (Demonstrating the effectiveness of the solution)
5. Evaluation (Evaluating the solution)
6. Communication (Disseminating the findings to relevant audiences)

The framework of design science research methodology proposed by Peffers, Tuunanen, Rothenberger *et al.* [201] inspired the research design in this thesis, as delineated in Figure 3.1. This thesis initiated the research with a study (study A) focusing on understanding the current state of research and the challenges of enabling robotic visual perception on industrial robots for assembly tasks. This study was closely related to problem identification and motivation, as well as defining the objectives of a solution. Then, the research continued with a study (study B) exploring potential technical solutions for enabling robotic visual perception on industrial robots for assembly tasks. This study was to design and develop computer vision-based solutions to address specific challenges and objectives identified in the first study. Quantitative experiments were needed to demonstrate the developed solutions' effectiveness and evaluate the developed solutions' performance. Lastly, the research findings of each study were disseminated through publications (Paper 1, Paper 2, and Paper 3 appended to this thesis).

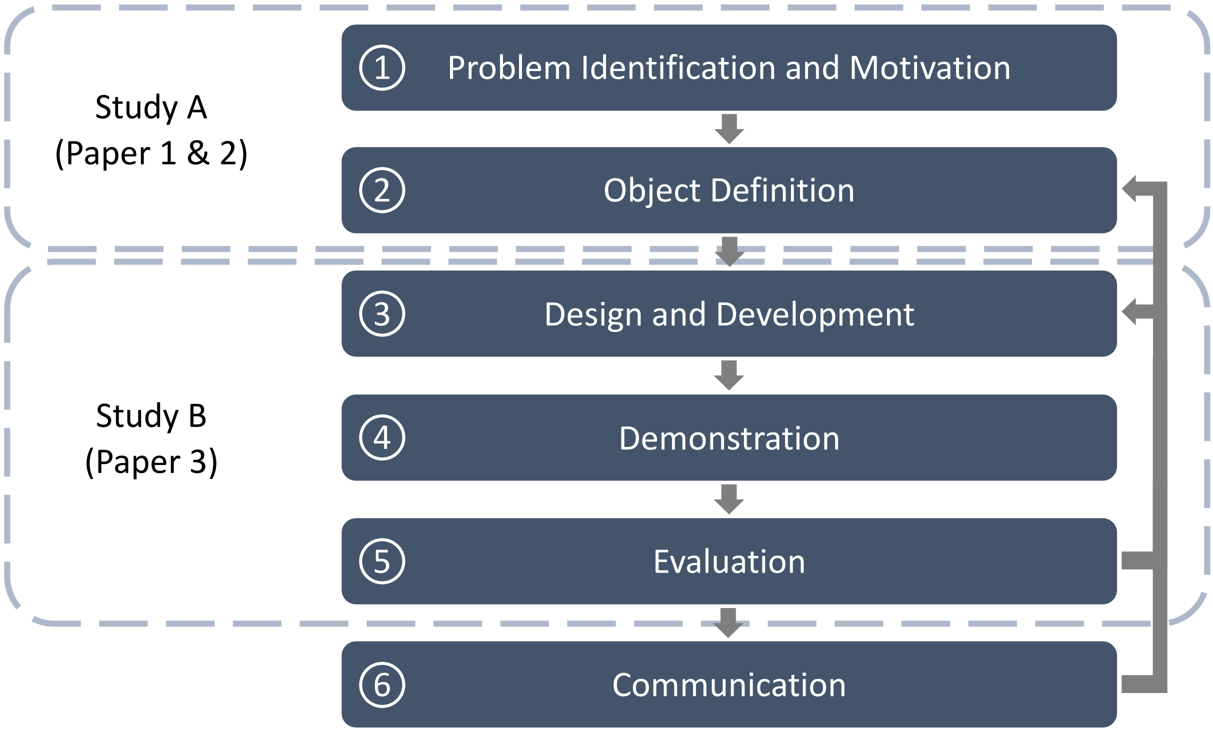


Figure 3.1: The framework of DSRM adapted from Peffers, Tuunanen, Rothenberger *et al.* [201]. The research in this thesis comprised two studies. Study A was designed to identify and motivate the problem of robotic visual perception in the robotic assembly of wire harnesses and define the objective for a vision-based solution. Study A led to the appended Paper 1 and Paper 2. Study B was designed to design and develop a technical solution, demonstrate its effectiveness, and evaluate its performance. Study B led to the appended Paper 3.

Multiple-Method Design

As interpreted in Morse [203], research with a multiple-method design consists of multiple self-contained and complete studies using different research methods to "address the same research question or different parts of the same research question or programmatic goal". This research adopted a multiple-method design, including a literature study (study A) following a qualitative descriptive design and an experimental study (study B) following a quantitative experimental design.

First, a qualitative descriptive design was adopted for the literature study to analyze the current state of research and clarify the prospective research directions. This literature study aimed to address the RQ1 (on challenges) by recognizing the challenges for developing technical solutions based on the analysis of the current state of research and to contribute to answering the RQ2 (on technical solutions) by discussing opportunities for further research and defining objectives for potential technical solutions. The research findings of this literature study were included in Paper 1 and Paper 2.

Then, a quantitative experimental study was conducted to assess the performance of designed technical solutions in addressing the challenges and objectives recognized in the previous literature study. Thus, this experimental study contributed to addressing the RQ2 (on technical solutions) by providing insights on potential solutions supported by quantitative experimental results. The potential positive results would demonstrate the effectiveness of the designed solutions. The potential negative results would, on the other hand, reveal constraints on the performance of the designed solutions, which would also contribute to addressing the RQ1 (on challenges). Considering the derivation of

Figure 3.2: A framework of positivist research design, adapted from Williamson, Burstein and McKemmish [195].

postpositivism from positivism, a framework of positivist research design [195], as shown in Figure 3.2, was referred to specify the stages of this experimental study. The research findings of this literature study were included in Paper 3.

3.3 Research Methods

Various research methods were implemented in each studies leading to the appended papers, as outlined in Table 3.1. The summary of the appended papers stemmed from the studies, as well as their contributions, will be elaborated in Chapter 4. Nonetheless, this section briefly summarizes the research methods implemented in each study to provide readers a quick reference.

The literature study was designed first to acknowledge the state-of-the-art research on enabling robotic visual perception of industrial robots for assembly tasks, particularly for wire harness assembly in the final assembly of automobiles. Based on the knowledge of the current state of research, this study intended to identify and motivate the problem of interest, i.e., to identify and motivate existing challenges demanding further inquiries. With challenges identified, this study would explore potential opportunities to enable robotic visual perception on industrial robots for the robotic assembly of wire harnesses in the final assembly of automobiles. With a constructivist worldview, a literature study was selected in this study through a qualitative approach with a descriptive design. The literature study was conducted following a systematic literature review methodology, considering systematic literature review as an instrumental methodology for comprehensively understanding the state of the art of a subject and identifying the gaps requiring future research [211]–[213]. A review protocol, including literature search and selection strategies, was determined first by a group of researchers. Then, literature data was collected by searching literature databases using a pre-defined search string. The collected literature was further selected with the consent of multiple researchers regarding pre-defined criteria. Additionally, the strategy of “snowballing” [209], including reference tracking and citation tracking, was conducted on the selected articles to identify potentially missing studies in the literature searching process. With the identified literature, this literature study adopted text coding based on a pre-defined scheme, followed by theme and pattern interpretation. More detailed research methodology of this study can be found in Section 4.1 and Section 4.2, as well as Paper 1 and Paper 2.

Based on the knowledge acquired from the literature study, an experimental study was designed to examine vision-based object detection solutions that may contribute to enabling robotic perception on industrial robots for assembly tasks. A quantitative

Table 3.1: The research design, research methods, and measures to ensure research quality of each appended paper.

Paper	Research design	Research method	Research quality measures
1	Qualitative approach Descriptive design Literature study	Systematic literature review - Database searching using a pre-defined string - Literature selection based on pre-defined criteria - Text coding using a pre-defined scheme - Inductive reasoning	Adopting consistent methods DARE criteria [206] Investigator triangulation [207] Peer debriefing [208]
2	Qualitative approach Descriptive design Literature study	Systematic literature review - Database searching using a pre-defined string - Literature selection based on pre-defined criteria - "Snowballing" [209] on selected literature - Text coding using a pre-defined scheme - Inductive reasoning	Adopting consistent methods DARE criteria [206] Investigator triangulation [207] Peer debriefing [208] Expert review [210]
3	Quantitative approach Experimental design Experimental study	Empirical measurement Statistical analysis Deductive reasoning	Adopting consistent methods Stratified sampling Statistical metrics Peer debriefing [208]

approach with an empirical postpositivist worldview was selected for this experimental study. Though positive that computer vision techniques are adequate, the performance of technical solutions needs to be examined and analyzed empirically [121]. Besides verifying the effectiveness of technical solutions based on quantitative evaluation, this study intended to discuss why the solutions failed on specific samples. A quantitative experimental study was selected to accomplish this objective. In this quantitative experimental study, numeric data was collected and analyzed based on statistical metrics, such as the rate of precision of detection. More detailed research methodology of this study can be found in Section 4.3 as well as Paper 3.

3.4 Research Quality

Validity and reliability are two significant criteria for research quality evaluation [193], [214]–[217]. Validity, including internal validity and external validity, suggests to what extent the study findings represent the truth among similar population outside the study [193], [211], [217]. Internal validity is the cornerstone of external validity and is defined as the extent to which systematic errors can be prevented in the design and conduct of the research, while external validity reflects the generalizability and applicability of the research outcomes outside the study [211], [217]. Reliability indicates the extent of consistency of the research approach in a study with the one in other studies by other researchers [218]. As shown in Table 3.1, each study adopted different methods to guarantee the quality of the research.

Qualitative research should concern both validity and reliability [193]. Researchers should adopt the research approach consistent across different researchers and projects to guarantee the reliability of the qualitative study and examine the accuracy of the research findings through specific procedures to ensure the validity [218]. A methodology for planning and conducting systematic literature reviews was suggested in Kitchenham [211], which has been adopted in various systematic literature reviews in computer science and engineering [219]–[227]. Thus, the systematic literature review in the qualitative literature study of this thesis was conducted following the methodology suggested in Kitchenham [211] to strengthen the reliability of the qualitative literature study. Besides, various methods were considered to reinforce the validity of the literature study. Funded by the Department of Health and the National Institute for Health Research of the United Kingdom, the Database of Abstracts of Reviews of Effects (DARE) and the NHS Economic Evaluation Database provide access to over 35000 systematic reviews in the field of health and social care interventions, whose quality was assessed based on publicly available criteria (shortened as "DARE criteria" hereinafter in this thesis) [206]. DARE criteria [206] qualify a systematic review considering[1]: 1) whether inclusion/exclusion criteria were reported; 2) whether the search was adequate; 3) whether the included studies were synthesized; 4) whether the quality of the included studies was assessed; and 5) whether sufficient details about the individual included studies were reported. According to DARE criteria [206], a systematic review is qualified only if it fulfills the first three and at least one of the fourth and fifth criteria. Following Kitchenham, Pearl Brereton, Budgen *et al.* [228] and Saleem, Khan, Zafar *et al.* [229], this thesis adopts DARE criteria [206] to evaluate the quality of the systematic literature review in the qualitative study. Besides, to reduce the subjective bias on data collection, analysis, and interpretation in the qualitative literature study, investigator triangulation [207] was adopted to strengthen the impartiality and mitigate the personal bias on the design of the review protocol and the judgment on literature selection and interpretation. Peer debriefing is also a helpful technique to strengthen the research quality by improving the study with feedback from colleagues and other researchers who are familiar with the research topic [208]. Multiple researchers were involved in conducting this qualitative literature study and reviewing different parts of the study. Additionally, expert review [210] were also adopted in the systematic literature review. Specifically, experts in relevant subjects from academia and industry (some of them as co-authors) were involved in calibrating the research methods and cross-validate the findings and interpretation.

For quantitative research, there are potential threats to the research validity embedded in, for example, for internal validity, experimental procedures, treatments, or participants' experiences, and for external validity, the generalization of research findings to the outside of the study [193], [230]. The quantitative experimental study in this thesis adopted the assessment suggested by Hammersley [230] to evaluate the reliability and validity of the quantitative experimental study concerning three aspects of the process of research: 1) whether measurement procedures were reliable and valid; 2) whether the findings can be generalized to a larger populations; and 3) whether variables were controlled effectively and sufficiently. In the quantitative experimental study in this research, the computer was the measure to collect data, which, with an experiment plan designed referring to relevant studies, strengthened the reliability and validity of the measurement. Then, stratified sampling was planned to separate the dataset regarding the distribution of data and the ratio among different data samples to ensure the generalizability of the treatment. Further, to guarantee the quality of the control of variables, the research intended to allocate sampled data randomly to treatment and control groups. In addition, statistical analysis was implemented to mitigate potential researchers' subjective bias in evaluating the treatment's performance. This study also adopted peer debriefing [208] to enhance the quality of this study.

[1] https://www.crd.york.ac.uk/CRDWeb/

Chapter 4

Summary of Appended Papers

This chapter briefly summarizes the three appended papers, including the core problem, the research methodology, and the contribution of each paper. This chapter also summarizes each appended paper's contribution to the research questions inquired in this thesis.

4.1 Paper 1

Title: Overview of Computer Vision Techniques in Robotized Wire Harness Assembly: Current State and Future Opportunities

Problem

Improving ergonomics and optimizing resource utilization while improving the assembly quality and assuring safety is desired for the installation of wire harnesses in the final assembly of automobiles [33], [176], [231]. Robotic assembly is a significant enabler and facilitator for achieving this goal, considering its superiority in replicability, transparency, and explainability over manual operations [7]. However, robotizing the assembly of wire harnesses remains laborious in actual production [129], [179]. The high degree of customization on wire harnesses and their deformation exacerbate the complexity of the assembly task, requiring robots to perceive and react to the surrounding environment and manipulate the object adaptively.

Vision is a fundamental source of information for object localization and recognition [5]. Previous research explored vision-based robotized assembly in different sectors by enabling the adaptive robotic visual perception [45], [185], [232]–[234]. However, the automotive industry has yet to identify any practical solution to robotize the assembly of wire harnesses in actual production. Therefore, further research is needed to understand the challenges of enabling robotic visual perception for robotizing wire harness assembly. Moreover, future research opportunities should also be identified to promote technical solutions for vision-based robotic assembly of wire harnesses.

Methodology

This paper's primary objective was to provide a comprehensive overview of the existing research on vision-based robotized wire harness assembly. Additionally, it aimed to identify crucial opportunities for future research in enabling visual perception of industrial robots to assemble wire harnesses. A qualitative literature study with a constructivist worldview was conducted to achieve these objectives.

The study initiated an inquiry on the Scopus database. The search string was TITLE-ABS-KEY((wir* OR cabl*) AND (harness* OR bundl*) AND assembl*). Then, three researchers examined the search results thoroughly and selected articles focused on vision-based robotized wire harness assembly in final assembly. The study only included articles in English for the analysis. Besides, the study excluded secondary studies, i.e., review articles and conference reviews. The selected articles were then grouped, considering multiple attributes of each study, such as the task of the operation, the object of interest, the type and location of vision systems, and the number of cameras. Lastly, the grouped articles were analyzed and interpreted through inductive reasoning to understand the current state of research and identify future research directions.

Contribution

This paper provides an overview of the existing research on vision-based robotized wire harness assembly. Table 1 and Table 2 in the appended Paper 1 summarized existing studies on vision-based robotic manipulation of wire harness components and visual machine perception of wire harness structures, respectively. This paper also discussed future research opportunities toward a more practical vision-based robotized wire harness assembly, including:

- Investigating the use of learning-based computer vision algorithms
- Evaluating proposed vision systems in actual production scenarios regarding the practicality and reliability
- Exploring new product designs of wire harnesses to facilitate robotic visual perception

4.2 Paper 2

Title: A Systematic Literature Review of Computer Vision Applications in Robotized Wire Harness Assembly

Problem

This article extended the study in the appended Paper 1 to summarize the state-of-the-art research and systematically discuss the challenges and future research directions.

Methodology

A systematic literature review was conducted in this study to answer the following three research questions:

1. What computer vision-based solutions have been proposed for robotized wire harness assembly?
2. What are the challenges for computer vision applications in robotized wire harness assembly?
3. What are the required future research activities and fields for developing more efficient and practical computer vision-based robotized wire harness assembly?

This study followed the methodology for planning and conducting a systematic literature review suggested by Kitchenham [211]. Following the methodology of investigator triangulation [207], three co-authors of this article collaborated continuously through all aspects of this study to ensure the quality of different stages of this literature study.

A review protocol was developed first to ensure a systematic and reproducible review method, as shown in Table 1 in the appended Paper 2. The following string was defined for the literature search within the field of *Article title*, *Abstract*, *Keywords* on Scopus on September 6, 2023: (wir* OR cabl*) AND (harness* OR bundl*) AND assembl*. The subject area was limited to *Engineering*, *Computer Science*, *Decision Sciences*, *Multidisciplinary*, and *Business, Management and Accounting*. The language of the article was limited to English. Finally, the literature search returned 662 articles for literature selection.

Next, three researchers conducted a two-step screening jointly regarding the inclusion and exclusion criteria shown in Table 1 in the appended Paper 2 to select the qualified literature for data synthesis and analysis. The article selection process was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [235], as shown in Figure 4 in the appended Paper 2. The first screening was based on the title and abstract of each article, which selected 22 articles for the second screening. The second screening was based on the full text of each article, which further sifted out 13 articles. Moreover, "snowballing" [209], including *reference tracking* and *citation tracking*, was implemented on the selected 13 articles for analysis to identify other relevant articles missed in the original search, which returned 2 more articles. Therefore, 15 peer-reviewed scientific articles were included for further data analysis and interpretation.

The selected articles were then grouped, considering multiple attributes of each study, such as the task of the operation, the object of interest, the type and location of vision systems, and the number of cameras. Lastly, the grouped articles were analyzed and interpreted through inductive reasoning to understand the current state of research and identify future research directions.

In addition, this systematic literature review adopted DARE criteria [206] to evaluate the research quality.

Contribution

This study reviewed the state-of-the-art research on vision-based robotized wire harness assembly. This systematic literature review identified 15 relevant studies that discussed vision-based robotized wire harness assembly. Table 2 in the appended Paper 2 summarizes the contribution of each relevant study from the perspective of computer vision applications.

The identified 15 studies indicated the existing research on enabling robotic visual perception on different levels of the constituent structure of wire harnesses to facilitate the robotization of wire harness assembly [35], [129], [166], [174], [175], [179], [180], [186], [236]–[242].

Regarding the components of wire harnesses the study focused on, the identified 15 studies were categorized into four groups (section 4.1 in Paper 2):

- Four studies on clamps [129], [174], [180], [237]
- Seven studies on connectors [166], [179], [236], [238]–[241]
- Three studies on cables [175], [186], [242]
- One study on wire harness bags [35]

Regarding the sub-tasks of the assembly, the identified 15 studies were categorized into three groups:

- Robotic manipulation of wire harness components [35], [129], [166], [174], [179], [180], [236]–[241]

- Monitoring sub-processes of the assembly [175], [186], [236], [238], [240], [242]
- Fault detection during the assembly [236]

The proposed vision systems in the identified 15 studies also contributed to different operations in the current installation of wire harnesses in the final assembly of electric passenger vehicles (section 4.2 in Paper 2), including preparation [35], untanglement [175], routing [32], [34], [129], [174], [186], [237], [242], and assembly [129], [166], [174], [179], [180], [236]–[241]. Nevertheless, the automotive industry had yet to identify adequate vision-based solutions to robotize the assembly of wire harnesses in the final assembly of automobiles.

The study in Paper 2 identified challenges for computer vision applications in robotized wire harness assembly, including:

- Achieving compatible robustness compared with the human vision system, especially considering the demanding production rate and intricate production environments
- Accomplishing visual recognition based on intrinsic physical properties of different components of wire harnesses

The study in Paper 2 identified future research opportunities for introducing computer vision applications in robotized wire harness assembly more efficiently and effectively, including:

- Adapting learning-based visual recognition algorithms to exploit intrinsic features and multi-modality data of wire harnesses
- Investigating the adaptation of vision-based solutions proposed for robotized manufacture of wire harnesses
- Assessing vision-based solutions under practical production conditions in terms of practicality, robustness, reliability, and sustainability
- Considering vision-based HRC and exploring solutions to address different assembly operations
- Exploring new product designs to facilitate visual recognition

4.3 Paper 3

Title: Deep Learning-Based Connector Detection for Robotized Assembly of Automotive Wire Harnesses

Problem

It is challenging to design and implement rule-based connector detection as the manual feature engineering can be unwieldy due to the significantly various and complex structures of wire harness connectors. Extensive experimental results have demonstrated the superiority of learning-based object detection over the rule-based methods, especially in this task scenario [85], [136]. However, as identified in Paper 1 and Paper 2, there is a lack of research on deep learning-based solutions for recognizing wire harness components, e.g., connectors, in previous studies.

Paper 3 aimed to verify the effectiveness of deep learning-based object detection on wire harness components. Wire harness connectors were focused in particular, considering connectors as essential components for connecting wire harnesses and transmitting signals and power. Datasets are essential to training and evaluating learning-based object detection [243]–[246]. However, no publicly available dataset fulfilled the need of the study. Thus, creating a dataset of connectors was fundamental before investigating learning-based connector detection for robotized wire harness assembly in this study.

Methodology

The objectives of this study were to verify the effectiveness of deep learning-based object detection on wire harness connectors and identify constraints on its effectiveness. With a postpositivist worldview, a quantitative experimental study was conducted to achieve these objectives.

Given the absence of a publicly available connector dataset, this study first created a dataset comprising 360 images of single and multiple connectors. This was achieved by utilizing 20 different automotive wire harness connectors, a process detailed in Fig. 2 in the appended Paper 3. This study captured 60 images of mixed connectors and 300 images of each connector. In particular, this study captured 6 profile images for each connector, covering the front, back, left, right, top, and bottom views. Additionally, 9 images were captured from other random views of each connector, ensuring a comprehensive dataset. Fig. 3, Fig. 4, and Fig. 5 in the appended Paper 3 present example images in the collected dataset. The image annotation procedure of the collected connector dataset followed the methodology implemented in the PASCAL visual object classes (VOC) challenge 2007 [243]. The number of annotated object instances in the collected connector dataset is presented in TABLE I and Fig. 6 in the appended Paper 3.

Then, this study trained and evaluated a two-stage object detector, Faster R-CNN [92], [247], and a one-stage object detector, YOLOv5 [248], to investigate the effectiveness of deep learning-based object detection on connector detection using the collected dataset. Data augmentation methods were implemented to inflate the created dataset with artificially generated images. The experimental results were analyzed and interpreted based on statistical metrics, including the rate of precision and mean Average Precision (mAP), as reported in TABLE II and TABLE III in the appended Paper 3.

Contribution

This study investigated deep learning-based object detection on automotive wire harness connectors to facilitate the robotized wire harness assembly. A dataset of automotive wire harness connectors was collected for training and evaluation of a two-stage object detector, Faster R-CNN [92], [247], and a one-stage object detector, YOLOv5 [248], respectively. The experiment results verified the effectiveness of deep learning-based connector detection for automotive wire harness assembly.

Besides, this study indicated two potential hindrances to learning-based visual recognition:

- Insufficient amount of training data
- Confusing features due to product designs of connectors

The experiment results encouraged future research on collecting a benchmark dataset for better training and more consistent and rigorous evaluation across various object detectors. The results also encouraged three approaches to promote the detection performance:

- Designing new object detectors to extract nuance features on objects
- Conducting multi-view or video-based detection to extract more distinguishable features
- Optimizing the product design, especially the appearance, to facilitate the vision-based object detection

4.4 Contributions of Appended Papers

This section amalgamates each appended paper's contributions to this thesis's research questions formulated in Chapter 1. As shown in Table 4.1, all three appended papers contribute to both research questions of this thesis to different extents.

Table 4.1: Summary of each appended paper's contribution to each research question.

Paper	Contribution to RQ1 (Challenge)	Contribution to RQ2 (Solution)
1	**Minor contribution** 1) Visually recognizing and tracking objects of interest without using additional artificial fiducial markers 2) Obtaining spatial information in 3D space visually 3) Ensuring the practicality and reliability of vision systems in actual production	**Major contribution** 1) Adopting learning-based approaches 2) Evaluating vision systems under practical production conditions 3) Exploring new product designs that can facilitate visual recognition
2	**Major contribution** 1) Visual recognition exploiting intrinsic features of objects of interest instead of additional artificial fiducial markers 2) Recognizing and tracking the structure and topology of deformable linear objects 3) Obtaining spatial information in 3D space for robotic operations 4) Guaranteeing the practicality, reliability, robustness, and sustainability in actual production	**Major contribution** 1) Adopting learning-based approaches 2) Learning from intrinsic features and multi-modality data 3) Collecting benchmark dataset for training and evaluating learning-based solutions 4) Evaluating vision systems under practical production conditions 5) Developing vision-based HRC 6) Investigating new product designs to facilitate visual recognition
3	**Minor contribution** 1) High-precision position and orientation acquisition 2) Visual recognition addressing occluded features and high-similarity features	**Major contribution** 1) Adopting learning-based approaches 2) Collecting benchmark dataset for training and evaluating vision-based solutions 3) Investigating multi-view visual recognition 4) Exploring new product designs to facilitate visual recognition

Chapter 5

Discussion

This chapter discusses the main research findings toward answering the research questions of this thesis and analyzes this thesis's contribution to academia and industry. This chapter also reflects on the limitations, the research quality, and the aspects of ethics and sustainability of the research. Lastly, this chapter envisions future research.

5.1 Answers to Research Questions

5.1.1 The Answer to RQ1

RQ1: What are the challenges of enabling robotic visual perception for assembly tasks?

The Answer in a Nutshell

On the level of objects, robotic visual perception needs to address the challenge of visually recognizing objects of interest by exploiting their intrinsic features instead of being assisted by affixed artificial fiducial markers. It is also challenging to visually recognize and track non-rigid objects' structure and topology during robotic manipulation. Moreover, there are challenges for robotic visual perception caused by product designs, such as small sizes, complex structures, and highly similar appearances of objects of interest. The need for precise positions and orientations of objects of interest in 3D space is not just a requirement but a constant necessity for the success of robotic operations. However, this task remains a challenge for up-to-date computer vision techniques. Translating theoretical results into practical industrial applications is a task that demands meticulous consideration of physical production environments and practical business requirements. This consideration is crucial to ensure technology's efficient, effective, and safe application. This requirement poses challenges in enabling robotic visual perception, which would enhance industrial robots' autonomy and make them adequate for demanding flexible automation tasks.

The Implication of the Answer

In essence, enabling the visual perception capability of industrial robots for robotizing assembly operations is to make industrial robots capable of recognizing objects of interest visually. Given the objects recognized, industrial robots can recognize the task and adapt their actions to accomplish required assembly tasks. Hence, recognizing objects of interest first is of utmost importance. Beyond recognizing objects, industrial robots demand precise positions and orientations of objects before accomplishing required sub-tasks, such as reaching, grasping, and manipulating. Additionally, as a practical application in final assembly, physical production environments and practical business issues should be considered to guarantee an efficient, effective, and safe implementation of technologies in actual production.

On the object level, industrial robots handle both rigid and non-rigid objects in final assembly. Specifically, as elaborated in section 2.3, non-rigid objects involved in the robotized wire harness assembly are branched deformable linear objects, i.e., the part of wire harness cables. Clamps for fixing wire harnesses and connectors for connecting wires are examples of rigid objects involved in the robotized assembly of wire harnesses. This thesis identify that the intrinsic physical properties of these two types of objects lead to different hindrances to industrial robots' robotic visual perception. Non-rigid objects in final assembly create the challenge of the visual recognition and tracking of the structure and the topology of non-rigid objects due to their deformation during robotic manipulation, as recognized in Paper 1 and Paper 2. Paper 1 and Paper 2 also unearthed that rigid objects create the challenge for industrial robots to visually recognize objects of interest by exploiting intrinsic features of the objects instead of being assisted by affixed artificial fiducial markers. Moreover, rigid objects with small sizes, complex structures, confusing colors, and family designs exacerbate these challenges. The experimental results in Paper 3 indicated the challenge on differentiating products with family designs from particular views due to their highly similar appearances or occluded distinguishable features.

For robotic operations, obtaining precise positions and orientations in 3D space is critical for designing the robotic control strategy but unwieldy to achieve, as elaborated in section 2.2. Robots require more than just recognizing the position and orientation of an object to accomplish their operations. The recognition must be accurate within tolerances. Otherwise, problems can arise with robotic grasping, leading to failed manipulation and assembly. Paper 1 and Paper 2 revealed that obtaining positions and orientations of objects with enough precision is arduous, especially for objects with small sizes and complex structures. Previous research discussed visual recognition of objects of interest using traditional rule-based computer vision techniques. However, objects with complex structures may cause extreme difficulty in feature engineering when manually designing rule-based vision systems. Previous research exploited artificial fiducial markers with specific patterns to facilitate feature engineering. However, the effectiveness of this kind of rule-based visual recognition solution can be impaired, and the quality of visual recognition will be affected if these artificial fiducial markers are occluded or have low visibility in actual production. The limited space for assembly onto the final products and the large number of objects demanding affixing artificial fiducial markers make this approach more troublesome to implement in industrial applications. Thus, accomplishing visual recognition by exploiting the intrinsic features of objects of interest without attaching artificial fiducial markers is desired.

Further considering industrial applications in practice elaborated in section 2.3, Paper 1 and Paper 2 identified the challenge of guaranteeing a successful integration of visual machine perception into industrial robots for practical applications, regarding the practicality, reliability, robustness, and sustainability of vision systems. Plenty of studies suggested the potential of computer vision techniques for enabling robotic visual perception and the prospective success of vision systems in practical industrial applications [45]. Nevertheless, the practicality and reliability of vision systems in practical applications still need to be discovered. The robustness of vision systems in actual production has yet to be compatible with humans, especially considering the demanding production rate and intricate production environments, such as the background and illumination conditions of visual inputs and the moving production line. In addition, more consideration should be given to the sustainability aspect of vision systems, which is increasingly critical in the current industry and society.

5.1.2 The Answer to RQ2

RQ2: How can robotic visual perception be enabled for assembly tasks?

The Answer in a Nutshell

Learning-based computer vision techniques should be considered and adopted to enable robotic visual perception, increase industrial robots' autonomy, and make them adequate for demanding flexible automation tasks. Beyond 2D RGB images, multi-modality data, such as depth images and point clouds, can be integrated into learning and acquiring the necessary spatial information for robotic operations through multimodal learning. Multi-view visual inputs can improve visual recognition performance when distinguishable features are invisible from particular views. Meanwhile, the benchmark dataset is fundamental for training and evaluating different learning-based solutions, not only for comparing different algorithms' performance but also for evaluating vision systems under practical production configurations. Moreover, evaluating the developed vision systems in practical production configuration is necessary to examine the practicality, reliability, robustness, and sustainability of the developed vision-based solutions. In addition to developing vision systems, new product designs are worth investing in to facilitate robotic visual perception.

The Implication of the Answer

Recent research in computer vision and deep learning has remarkably promoted the performance of visual recognition with learning-based techniques [136]. This indicates the potential of learning-based computer vision techniques to exploiting intrinsic features of objects of interest. Section 2.2 elaborated the superiority of learning-based techniques over rule-based techniques in handling vision tasks around objects with complex features. Paper 1 and Paper 2 advocated to develop and implement learning-based visual recognition algorithms to enable robotic visual perception to utilize intrinsic features of objects of interest. The outcome of Paper 3 further verified the effectiveness of deep learning-based computer vision techniques on rigid object detection in wire harness assembly.

Paper 3 also discussed the potential invisibility of distinguishable features on objects of interest from particular views. Multi-view or video-based visual recognition can address this challenge. Notably, multi-view visual inputs can capture multiple views of an object besides the specific view where distinguishable features are invisible. The same idea is shared with video-based approaches as multi-view information will be captured in videos shooting around objects. The visual recognition algorithm can better recognize the object with more features observed. In addition to processing 2D RGB visual inputs, research in multimodal learning indicates the prospective solution for strengthening visual recognition performance by exploiting various modalities of data [113]. Advancing sensor technology has made it much more convenient to obtain data other than 2D RGB images, such as depth images and point clouds. As elaborated in Paper 1 and Paper 2, these modalities of data have spatial information embedded compared to 2D RGB images, which can facilitate the robotic visual perception of objects of interest in 3D space.

In parallel to the research on learning-based vision systems, research on benchmark datasets is also instrumental, considering the significance of datasets for learning-based computer vision techniques [243]–[246] and the application of computer vision techniques in industry [132]. However, as indicated in Paper 3, there is a lack of datasets for specific scenarios in enabling visual perception of industrial robots for assembly tasks. Therefore, in the experimental study in Paper 3, a dataset of connectors was created initially by taking pictures of connectors and annotating the positions of connectors manually. More studies investigating the aspect of the dataset are desired to improve the visual recognition performance further and lay the foundation for consistent evaluations of various technical solutions for enabling robotic visual perception for assembly tasks.

Moreover, as suggested in Paper 1 and Paper 2, evaluation under practical production configurations is desired and necessary to validate the practicality, reliability, robustness, and sustainability. This evaluation is a prerequisite for guaranteeing that the developed vision systems fulfill the practical requirements of actual production. Additionally, Paper 1, Paper 2, and Paper 3 identified various challenges due to product designs. Inspired by the philosophy of "Design for X" [249], novel designs on objects of interest are desired to facilitate visual recognition.

5.2 Contributions of This Thesis

As elaborated in Chapter 1, more research is required to understand why visual perception capabilities have yet to be enabled on industrial robots for assembly tasks in production and explore prospective technical solutions to achieve it. This thesis ventures into wire harness assembly operations for electric vehicles and the untapped potential of enabling visual perception of industrial robots for such assembly tasks. Through the designed studies, this thesis identified challenges for enabling robotic visual perception for assembly tasks and initiated exploring potential vision-based approaches to address the identified challenges. This thesis provides a practical lens for industry decision-makers, illuminating the potential challenges and opportunities of promoting the vision-based robotic assembly, thereby offering tangible insights for real-world applications. This thesis also provides academia with a road map to address the problems, thereby facilitating the translation of research on robotic visual perception into practical industrial applications.

To Academia

This thesis contributes to the applied research in computer vision, artificial intelligence, robotics, and automation. This thesis's outcome also contributes to research on the robotic assembly of wire harnesses and the robotic manipulation of deformable linear objects in production elaborated in section 2.3. The research conducted in Paper 1 and Paper 2 revealed the increasing research effort across countries. It also underscored the crucial role of vision systems in facilitating the automation of manual wire harness assembly in production. Despite research over a decade, industry has yet to implement practical solutions to address the problem [36]. Therefore, the outcome of Paper 1 and Paper 2 deepened the understanding of the existing research and challenges of enabling robotic visual perception for assembly tasks through a systematic literature review. These insights helped lay the foundation for designing and developing more efficient and effective solutions to enable robotic visual perception for assembly tasks. Afterward, the research conducted in Paper 3 dug deeper into deep learning-based computer vision techniques to verify the effectiveness of deep learning-based object detection for vision-based robotic assembly of wire harnesses through a quantitative experimental study. The computer vision, artificial intelligence, and robotics community discussed the potential applicability of deep learning-based vision systems in industrial applications [22]. The outcome of Paper 3 provided quantitative and qualitative verification of the potential applicability of deep learning-based vision systems in industrial applications. It also hinted at the transformative impact they could have on assembly tasks. On the other hand, the experimental results unearthed more challenges demanding further inquiries into enabling industrial robots' visual perception capabilities for assembly tasks.

To Industry

The technology readiness level (TRL) scale is widely adopted in complex system development to assess the maturity of technologies [250]. The research outcome of this thesis promotes the development of robot vision systems for complex assembly tasks toward higher TRLs. By the beginning of this thesis's research, industry had yet to implement automation solutions for complex assembly tasks that demand flexible automation. The outcome of Paper 1 and Paper 2 indicated that research in relevant fields had been conducted but led to a minor impact on industry. Therefore, the TRL was estimated to be 2 to 3. Afterward, the outcome of Paper 3 verified the effectiveness of deep learning-based techniques to sub-tasks of assembly. Hence, it is estimated that these results contribute to promoting the TRL to 3 to 4.

Specifically, the outcome of Paper 1, Paper 2, and Paper 3 revealed that, beyond research on computer vision algorithms, enabling robotic visual perception for assembly tasks can also be constrained by product designs, practical assembly environment, and production requirements. These insights have indicated the negative effect of several assembly environment conditions on robotic visual perception, which contributes to the assembly sector in the manufacturing industry regarding, for example, designing assembly lines that facilitate robotic visual perception. The research also identified the necessity of optimizing product designs to facilitate visual machine perception, which encourages product designers and manufacturers to optimize designs to facilitate robotic visual perception. Additionally, the experimental study in Paper 3 indicates the effectiveness of deep learning-based computer vision techniques. It suggests a general workflow for deploying a deep learning-based solution to enable robotic visual perception in practice, including dataset collection, data processing, deep learning model training, and evaluation.

5.3 Limitations of This Thesis

Though targeting to make the research findings as generalizable as possible, the author of this thesis foresees potential limitations to the outcome of this thesis, particularly regarding the object and application context.

The realm of robotic perception of the external environment is complex, encompassing not only objects to be manipulated but also objects in the surrounding environment and the presence of humans nearby. However, this thesis generalized all the things a robot needs to perceive as objects of interest and simplified the scope to the objects to be manipulated, i.e., wire harnesses. Before beginning the research, it was premised that wire harnesses could be seen as representative objects in assembly tasks because they consist of rigid and non-rigid objects. However, as discussed in section 2.3, research can generalize wire harnesses as specific DLOs, which is a sub-group of non-rigid object manipulation [70]. Additionally, wire harnesses typically used in industry are made of opaque materials. Given the complexity of the research area, further studies are crucial to examine the generalizability of this thesis's conclusions to other types of objects involved in practical assembly tasks.

Context-wise, the research in this thesis heavily relies on the Swedish automotive industry. Though anticipating the potential of generalizing the research findings to other industrial contexts, the outcome of this thesis related to practical production requirements, e.g., productivity, quality, ergonomics, and safety, may be interpreted differently in other manufacturing industries. Therefore, the research findings may need to be calibrated considering specific application scenarios.

5.4 Reflections on Research Quality

Validity and reliability are two significant criteria for research quality evaluation [193], [214]–[217]. This research consists of a qualitative literature study and a quantitative experimental study. As explained in section 3.4, various methods were adopted in each study of this thesis intended to guarantee the quality of the research, whereas there is a necessity to reflect on the assessment of the research quality of each study.

Following [228], the qualitative literature study in this thesis adopted DARE criteria [206] to evaluate the quality of the systematic literature review, as explained in section 3.4. Correspondingly, the review protocol and the literature searching and selection processes were documented in the appended Paper 2. Specifically, Paper 2 reported the inclusion/exclusion criteria, justified the adequacy of the search, and synthesized the included studies. The quality of each included study was assessed during the literature selection. Details about the included studies were also presented by summarizing each study's purpose, method, and result. The literature study also adopted investigator triangulation [207] to reduce the negative impact of personal bias on data selection, analysis, and interpretation. Specifically, three researchers collaborated through research planning, literature searching, and literature selection. Other researchers were also involved in this research to provide comments and reviews to improve the research quality through peer debriefing [208] and expert review [210].

Following [230], the reliability and validity of the quantitative experimental study in this thesis can be assessed concerning three aspects of the research process: measurement, generalization, and the control of variables. Regarding measurement, the research design and experiment settings were documented in Paper 3. Besides, the dataset was created following the methodology in peer-reviewed research to assure the reproducibility of the study on different occasions by different researchers. Adopting these methods contributed to the reliability of the measurement technique and strategy employed and the validity of the findings of this measurement process. Regarding generalization, the quantitative experimental study in this thesis did not involve sampling from a larger population nor demand inquiry on the generalization of the deep learning-based computer vision techniques. Such theoretical analysis on the generalizability of deep learning techniques can be found in other research [251]–[253]. In addition, the experimental study randomly allocated the dataset for training and testing following stratified sampling and conducted statistical analysis on the experimental results. Adopting both approaches reduced the potential impact of personal bias and enhanced the variable control's reliability and validity. Nevertheless, the author of this thesis foresaw a potential threat to the validity of this quantitative experimental study in the dataset collected for experiments. Remarkably, the research quality can be affected by the quality of images and the annotation quality. The camera and the image-capturing strategy can affect different aspects of the image quality, e.g., the image resolution, the noise, the illumination condition, and the background. This problem may be relieved by using a more sophisticated camera to shoot images using a well-designed strategy. The annotation quality may be affected by the strategy and accuracy of drawing the bounding boxes around objects, especially when the bounding box does not fit the object perfectly and includes parts of other objects. This problem may be relieved by increasing the number of images in the dataset to provide more data covering more scenarios or choosing different annotation strategies in future research. Hence, the author advocated a follow-up study on the benchmark dataset. The quantitative experimental study also adopted peer debriefing [208] to further enhance the quality of this study with reviews and comments from other researchers.

5.5 The Aspect of Ethics

Researchers should anticipate the potential ethical issues and endeavor to address them during their studies [254]–[257]. This section reflects on this thesis's ethics aspect with the intention of introspection.

Prior to beginning the studies, authorship for publication was first negotiated. The risk of breaching privacy is low for data collection, analysis, and interpretation, as this research only collected publicly available literature and images of products with the providers' consent, and this thesis did not collect data from or about people. In the literature study, papers were selected based on their quality and content to avoid potential bias against any author or organization of any author. In the experimental study, both good and bad results were analyzed and presented to avoid confirmation bias [258] or disclosing only positive results. Besides, the details of the research with the study design were released in each appended paper to make it possible for readers to assess the credibility of each study by themselves, following suggestions from [259]. Each study's raw data and other materials were also archived for retracing.

5.6 The Aspect of Sustainability

Attention to sustainability has been expanding immensely globally. Reflecting on this aspect is necessary to strengthen the inclusiveness of research. This section analyzes the impact of the output of this research regarding the aspect of sustainability, considering the triple bottom line of sustainability: economic, social, and environmental [260].

Economic Sustainability

Currently, some assembly tasks are desired to be automated with industrial robots but remain accomplished manually due to the high complexity of tasks. These manual operations need to be improved in terms of quality and productivity. By harnessing the power of robotic visual perception, industrial robots can achieve higher levels of autonomy and become more capable of handling assembly tasks that are currently accomplished manually due to their high complexity. This holds the promise of significant improvements in assembly quality and productivity. In this manner, production can achieve better assembly quality and higher productivity, strengthening economic sustainability.

Social Sustainability

The outcome of this research also has an impact on social sustainability. As explained in Chapter 1, some manual operations in assembly tasks in current production cause ergonomic problems and safety concerns to human operators. With robotic autonomy improved by enabling robotic visual perception, industrial robots will become more intelligent and capable of handling complex assembly tasks, especially those involving ergonomic problems or dangerous operations. In this manner, realizing this research will contribute to the transformation toward safer and more ergonomic-friendly workspace. Nevertheless, it is noteworthy that industrial robots will become more intelligent, with better perception capabilities, and increasingly capable of performing more tasks. Intelligent industrial robots may take over some current manual tasks from human operators. On the other hand, they will also create new challenges and responsibilities for human operators to address. Hence, workforce development through upskilling and reskilling, though out of the scope of this research, should be considered seriously, involving researchers and practitioners from different fields to prepare the organization and business for the forthcoming new industry.

Environmental Sustainability

Enabling the visual perception of industrial robots can indirectly strengthen environmental sustainability. For example, robotic automation in the final assembly may improve assembly quality. Better assembly quality can reduce the waste from the assembly process regarding material and rework. The reduced waste will result in better resource efficiency and less negative environmental impact.

5.7 Future Research

Collecting Better Benchmark Datasets

Although a connector dataset was collected to train and evaluate learning-based object detectors in Paper 3, the visual recognition performance remains unsatisfactory. One cause identified was the insufficient amount of data in the dataset. Therefore, the author of this thesis intends to collect better benchmark datasets for better training and evaluation. Besides 2D RGB images collected in the connector dataset in Paper 3, other modalities of data can also be considered, such as the depth images, to enable industrial robots' capabilities to perceive and infer 3D spatial information.

Investigating Other Vision Tasks

The literature study identifies the need to acquire positions and orientations of objects of interest to enable visual perception in industrial robots for assembly tasks. Moreover, the experimental study verified the effectiveness of learning-based object detectors in acquiring the positions of objects. Future research will include more studies on acquiring orientations, such as 3D object detection and 6D object pose estimation. Additionally, semi-automation based on human-robot collaboration is a direction worth further research to promote robotic automation for assembly tasks. Human-robot collaboration also involves diverse vision tasks for developing efficient, effective, and safe human-robot interaction.

Conducting Evaluation Under Industrial Configurations

The practical aspect of vision-based solutions for enabling robotic visual perception is also instrumental, considering the ultimate goal of integrating the developed vision systems into practical applications in production. Assessing vision-based solutions under practical manufacturing scenarios is critical to validating the proposed vision systems' practicality, reliability, robustness, and sustainability. In the future, the author of this thesis intends to explore how such evaluations can be accomplished.

Chapter 6

Conclusion

Inspired by design science research methodology, this thesis delved into the challenges of enabling robotic visual perception for assembly tasks while exploring potential opportunities and technical solutions through a multiple-method research approach.

This thesis identified four challenges hampering the enabling of robotic visual perception for assembly tasks:

- Visual recognition based on intrinsic features of objects of interest
- Recognizing and tracking the structure and topology of non-rigid objects
- Acquiring high-precision positions and orientations of objects of interest
- Ensuring effective, efficient, and safe applications in practical production

This thesis also identified six prospective directions for developing technical solutions to enable robotic visual perception for assembly tasks:

- Adopting learning-based computer vision techniques
- Developing vision systems based on multi-view and/or multi-modality visual inputs
- Collecting benchmark datasets to facilitate algorithm development and assessment
- Evaluating proposed vision systems under practical manufacturing scenarios
- Developing vision-based human-robot collaboration
- Exploring new product designs to facilitate visual recognition

This thesis has theoretical and practical implications. Theoretically, the results of this thesis are not just an endpoint but a stepping stone for further research. Notably, this thesis provides empirical evidence of challenges and opportunities for enabling visual perception of industrial robots for assembly tasks. These outcomes lay the foundation for developing intelligent robots for industrial applications, opening up new avenues for exploration and innovation in the field. Practically, practitioners can directly adapt the results of this thesis to analyze the specific challenges and opportunities of enabling more intelligent robotic assembly or robotic automation in their task scenarios. Notably, the findings of this thesis can serve as a solid foundation for developing specific machine vision-based robot systems in industry, thereby enhancing the intelligence and adaptability of industrial robots and contributing to more intelligent and sustainable production.

Bibliography

[1] X. Xu, Y. Lu, B. Vogel-Heuser and L. Wang, 'Industry 4.0 and industry 5.0-inception, conception and perception,' *Journal of Manufacturing Systems*, vol. 61, pp. 530–535, 2021. DOI: 10.1016/j.jmsy.2021.10.006 (cit. on pp. 3, 4, 12).

[2] F. J. Riley, *Assembly automation: a management handbook*, 2nd ed. New York, NY, USA: Industrial Press Inc., 1996 (cit. on pp. 3, 10).

[3] J. Krüger, T. Lien and A. Verl, 'Cooperation of human and machines in assembly lines,' *CIRP Annals*, vol. 58, no. 2, pp. 628–646, 2009. DOI: 10.1016/j.cirp.2009.09.009 (cit. on pp. 3, 10).

[4] S. J. Hu, J. Ko, L. Weyand *et al.*, 'Assembly system design and operations for product variety,' *CIRP Annals*, vol. 60, no. 2, pp. 715–733, 2011. DOI: 10.1016/j.cirp.2011.05.004 (cit. on pp. 3, 10).

[5] A. Billard and D. Kragic, 'Trends and challenges in robot manipulation,' *Science*, vol. 364, no. 6446, eaat8414, 2019. DOI: 10.1126/science.aat8414 (cit. on pp. 3–5, 11, 12, 20, 29).

[6] C. Müller, 'World robotics 2023 - industrial robots,' IFR Statistical Department, Tech. Rep., 2023 (cit. on p. 3).

[7] M. T. Mason, 'Toward robotic manipulation,' *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 1–28, 2018. DOI: 10.1146/annurev-control-060117-104848 (cit. on pp. 3–5, 11, 19, 29).

[8] T. Arai, 'Forecast of assembly automation in the automobile industry: Technological progress in robotics,' *Technological Forecasting and Social Change*, vol. 35, no. 2, pp. 133–148, 1989. DOI: 10.1016/0040-1625(89)90051-6 (cit. on p. 3).

[9] Å. Fast-Berglund, F. Palmkvist, P. Nyqvist, S. Ekered and M. Åkerman, 'Evaluating cobots for final assembly,' *Procedia CIRP*, vol. 44, pp. 175–180, 2016. DOI: 10.1016/j.procir.2016.02.114 (cit. on pp. 3, 10).

[10] H. Chen, B. Zhang and G. Zhang, 'Robotic assembly,' in *Handbook of Manufacturing Engineering and Technology*, A. Y. C. Nee, Ed., London, UK: Springer, 2015, pp. 2347–2401. DOI: 10.1007/978-1-4471-4670-4_105 (cit. on pp. 3, 4, 11, 20).

[11] K. E. Stecke and R. P. Parker, 'Flexible automation,' in *Encyclopedia of Production and Manufacturing Management*, P. M. Swamidass, Ed., Boston, MA, USA: Springer, 2000, pp. 213–217. DOI: 10.1007/1-4020-0612-8_343 (cit. on pp. 3, 18).

[12] T. Pardi, 'Fourth industrial revolution concepts in the automotive sector: Performativity, work and employment,' *Journal of Industrial and Business Economics*, vol. 46, no. 3, pp. 379–389, 2019. DOI: 10.1007/s40812-019-00119-9 (cit. on pp. 3, 5, 10).

[13] V. D. Pasquale, S. Miranda, W. P. Neumann and A. Setayesh, 'Human reliability in manual assembly systems: A systematic literature review,' *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 675–680, 2018. DOI: 10.1016/j.ifacol.2018.08.396 (cit. on p. 3).

[14] G. Michalos, S. Makris, N. Papakostas, D. Mourtzis and G. Chryssolouris, 'Automotive assembly technologies review: Challenges and outlook for a flexible and adaptive approach,' *CIRP Journal of Manufacturing Science and Technology*, vol. 2, no. 2, pp. 81–91, 2010. DOI: 10.1016/j.cirpj.2009.12.001 (cit. on pp. 3, 5, 12, 19).

[15] European Commission, Directorate-General for Research and Innovation and J Müller, *Enabling Technologies for Industry 5.0: results of a workshop with Europe's technology leaders*. Luxembourg: Publications Office of the European Union, 2020. DOI: 10.2777/082634 (cit. on p. 3).

[16] J. Leng, W. Sha, B. Wang *et al.*, 'Industry 5.0: Prospect and retrospect,' *Journal of Manufacturing Systems*, vol. 65, pp. 279–295, 2022. DOI: 10.1016/j.jmsy.2022.09.017 (cit. on pp. 3, 4).

[17] L. Rozo, A. G. Kupcsik, P. Schillinger *et al.*, 'The e-bike motor assembly: Towards advanced robotic manipulation for flexible manufacturing,' *Robotics and Computer-Integrated Manufacturing*, vol. 85, p. 102 637, 2024. DOI: 10.1016/j.rcim.2023.102637 (cit. on pp. 3, 11).

[18] H. ElMaraghy, L. Monostori, G. Schuh and W. ElMaraghy, 'Evolution and future of manufacturing systems,' *CIRP Annals*, vol. 70, no. 2, pp. 635–658, 2021. DOI: 10.1016/j.cirp.2021.05.008 (cit. on pp. 4, 12).

[19] F. Jovane, Y. Koren and C. Boër, 'Present and future of flexible automation: Towards new paradigms,' *CIRP Annals*, vol. 52, no. 2, pp. 543–560, 2003. DOI: 10.1016/S0007-8506(07)60203-0 (cit. on p. 4).

[20] L. Wang, S. Liu, H. Liu and X. V. Wang, 'Overview of human-robot collaboration in manufacturing,' in *Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing*, L. Wang, V. D. Majstorovic, D. Mourtzis, E. Carpanzano, G. Moroni and L. M. Galantucci, Eds., Cham, Switzerland: Springer, 2020, pp. 15–58. DOI: 10.1007/978-3-030-46212-3_2 (cit. on pp. 4, 11).

[21] W. He, Z. Li and C. L. P. Chen, 'A survey of human-centered intelligent robots: Issues and challenges,' *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 602–609, 2017. DOI: 10.1109/JAS.2017.7510604 (cit. on p. 4).

[22] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach* (Pearson series in artificial intelligence), 4th ed. Hoboken, NJ, USA: Pearson, 2021 (cit. on pp. 4, 16, 38).

[23] K. M. Galotti, 'Perception: Recognizing patterns and objects,' in *Cognitive Psychology In and Out of the Laboratory*, 5th ed., SAGE Publications, Inc., 2014, ch. 3, pp. 38–64 (cit. on p. 4).

[24] M. Brady, 'Artificial intelligence and robotics,' *Artificial Intelligence*, vol. 26, no. 1, pp. 79–121, 1985. DOI: 10.1016/0004-3702(85)90013-X (cit. on pp. 4, 10).

[25] S. J. Russell and P. Norvig, 'Intelligent agents,' in *Artificial intelligence: a modern approach*, ser. Pearson series in artificial intelligence, 4th ed., Hoboken, NJ, USA: Pearson, 2021, ch. 2, pp. 36–62 (cit. on pp. 4, 12).

[26] T. K. Lien, 'Manual assembly,' in *CIRP Encyclopedia of Production Engineering*, L. Laperrière and G. Reinhart, Eds., Berlin, Heidelberg, Germany: Springer, 2014, pp. 825–828. DOI: 10.1007/978-3-642-20617-7_6624 (cit. on pp. 4, 10).

[27] M. Peña-Cabrera, I. Lopez-Juarez, R. Rios-Cabrera and J. Corona-Castuera, 'Machine vision approach for robotic assembly,' *Assembly Automation*, vol. 25, no. 3, pp. 204–216, 2005. DOI: 10.1108/01445150510610926 (cit. on pp. 4, 15).

[28] H. A. Pierson and M. S. Gashler, 'Deep learning in robotics: A review of recent research,' *Advanced Robotics*, vol. 31, no. 16, pp. 821–835, 2017. DOI: 10.1080/01691864.2017.1365009 (cit. on p. 4).

[29] S. W. Kim, J. H. Kong, S. W. Lee and S. Lee, 'Recent advances of artificial intelligence in manufacturing industrial sectors: A review,' *International Journal of Precision Engineering and Manufacturing*, pp. 1–19, 2022. DOI: 10.1007/s12541-021-00600-3 (cit. on p. 4).

[30] C.-Y. Lee, H. Lee, I. Hwang and B.-T. Zhang, 'Visual perception framework for an intelligent mobile robot,' in *2020 17th International Conference on Ubiquitous Robots (UR)*, 2020, pp. 612–616. DOI: 10.1109/UR49135.2020.9144932 (cit. on p. 4).

[31] H. G. Nguyen, M. Kuhn and J. Franke, 'Manufacturing automation for automotive wiring harnesses,' *Procedia CIRP*, vol. 97, pp. 379–384, 2021. DOI: 10.1016/j.procir.2020.05.254 (cit. on p. 5).

[32] M. Wnuk, C. Hinze, M. Zürn, Q. Pan, A. Lechler and A. Verl, 'Tracking branched deformable linear objects with structure preserved registration by branch-wise probability modification,' in *2021 27th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, 2021, pp. 101–108. DOI: 10.1109/M2VIP49856.2021.9665147 (cit. on pp. 5, 19, 32).

[33] S. M. Research, 'Automotive wiring harness market: By material (copper, aluminum, others), by vehicle type (passenger cars, two-wheelers, commercial vehicles), by application (chassis harness, engine harness, body & lighting harness, dashboard/cabin harness, sunroof harness, battery harness, hvac harness, seat harness, door harness), by geography, size, global share, forecast, 2021-2030,' Strategic Market Research, Tech. Rep., 2022 (cit. on pp. 5, 29).

[34] M. Zürn, M. Wnuk, A. Lechler and A. Verl, 'Topology matching of branched deformable linear objects,' in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 7097–7103. DOI: 10.1109/ICRA48891.2023.10161483 (cit. on pp. 5, 19, 21, 32).

[35] B. L. Žagar, A. Caporali, A. Szymko *et al.*, 'Copy and paste augmentation for deformable wiring harness bags segmentation,' in *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2023, pp. 721–726. DOI: 10.1109/AIM46323.2023.10196168 (cit. on pp. 5, 31, 32).

[36] O. Salunkhe, W. Quadrini, H. Wang *et al.*, 'Review of current status and future directions for collaborative and semi-automated automotive wire harnesses assembly,' *Procedia CIRP*, vol. 120, pp. 696–701, 2023. DOI: 10.1016/j.procir.2023.09.061 (cit. on pp. 5, 17, 19, 20, 38).

[37] S. A. Pascual and S. Naqvi, 'An investigation of ergonomics analysis tools used in industry in the identification of work-related musculoskeletal disorders,' *International Journal of Occupational Safety and Ergonomics*, vol. 14, no. 2, pp. 237–245, 2008. DOI: 10.1080/10803548.2008.11076755 (cit. on p. 5).

[38] C. Fischer, J. Bönig, J. Franke, M. Lušić and R. Hornfeck, 'Worker information system to support during complex and exhausting assembly of high-voltage harness,' in *2015 5th International Electric Drives Production Conference (EDPC)*, 2015, pp. 1–7. DOI: 10.1109/EDPC.2015.7323211 (cit. on p. 5).

[39] S. Olbrich and J. Lackinger, 'Manufacturing processes of automotive high-voltage wire harnesses: State of the art, current challenges and fields of action to reach a higher level of automation,' *Procedia CIRP*, vol. 107, pp. 653–660, 2022. DOI: 10.1016/j.procir.2022.05.041 (cit. on p. 5).

[40] H. Wang, O. Salunkhe, W. Quadrini *et al.*, 'Overview of computer vision techniques in robotized wire harness assembly: Current state and future opportunities,' *Procedia CIRP*, vol. 120, pp. 1071–1076, 2023. DOI: 10.1016/j.procir.2023.09.127 (cit. on pp. 5, 17, 20).

[41] J. Shi, B. Hamner, R. Simmons and S. Singh, 'Mobile robotic assembly on a moving vehicle,' in *International Symposium on Flexible Automation*, vol. 45110, 2012, pp. 117–123. DOI: 10.1115/ISFA2012-7193 (cit. on p. 5).

[42] International Organization for Standardization, *Iso 10218-1:2011 robots and robotic devices — safety requirements for industrial robots — part 1: Robots*, 2011 (cit. on pp. 5, 20).

[43] International Organization for Standardization, *Iso 10218-2:2011 robots and robotic devices — safety requirements for industrial robots — part 2: Robot systems and integration*, 2011 (cit. on pp. 5, 20).

[44] International Organization for Standardization, *Iso/ts 15066:2016 robots and robotic devices — collaborative robots*, 2016 (cit. on pp. 5, 20).

[45] L. Zhou, L. Zhang and N. Konz, 'Computer vision techniques in manufacturing,' *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 1, pp. 105–117, 2023. DOI: 10.1109/TSMC.2022.3166397 (cit. on pp. 6, 12, 13, 15, 16, 20, 21, 29, 36).

[46] O. E. Dictionary, *Automation (n.), sense 2*, 2023. DOI: 10.1093/OED/5111105470 (cit. on p. 9).

[47] CIRP, 'Fundamental terms of manufacturing/grundlegende begriffe der produktion/termini fondamentali della produzione,' in *Dictionary of Production Engineering III – Manufacturing Systems Wörterbuch der Fertigungstechnik III – Produktionssysteme Dizionario di Ingegneria della Produzione III – Sistemi di produzione: Trilingual Edition Dreisprachige Ausgabe Edizione completa trilingue*, Berlin, Heidelberg, Germany: Springer, 2020, ch. 1, pp. 1–59. DOI: 10.1007/978-3-662-53334-5_1 (cit. on p. 9).

[48] Å. Fasth-Berglund and J. Stahre, 'Cognitive automation strategy for reconfigurable and sustainable assembly systems,' *Assembly automation*, vol. 33, no. 3, pp. 294–303, 2013. DOI: 10.1108/AA-12-2013-036 (cit. on p. 9).

[49] Å. Fasth, 'Quantifying levels of automation - to enable competitive assembly systems,' Ph.D. dissertation, 2012 (cit. on p. 9).

[50] D. Romero, J. Stahre and M. Taisch, 'The operator 4.0: Towards socially sustainable factories of the future,' *Computers & Industrial Engineering*, vol. 139, p. 106 128, 2020. DOI: 10.1016/j.cie.2019.106128 (cit. on p. 9).

[51] D. Thurman, D. Brann and C. Mitchell, 'An architecture to support incremental automation of complex systems,' in *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 2, 1997, pp. 1174–1179. DOI: 10.1109/ICSMC.1997.638109 (cit. on p. 9).

[52] S. J. Hu, 'Assembly,' in *CIRP Encyclopedia of Production Engineering*, L. Laperrière and G. Reinhart, Eds., Berlin, Heidelberg, Germany: Springer, 2014, pp. 50–52. DOI: 10.1007/978-3-642-20617-7_6616 (cit. on p. 10).

[53] T. Lien and F. Rasch, 'Hybrid automatic-manual assembly systems,' *CIRP Annals*, vol. 50, no. 1, pp. 21–24, 2001. DOI: 10.1016/S0007-8506(07)62062-9 (cit. on p. 10).

[54] G. Reinhart, 'Assembly automation,' in *CIRP Encyclopedia of Production Engineering*, L. Laperrière and G. Reinhart, Eds., Berlin, Heidelberg, Germany: Springer, 2014, pp. 52–54. DOI: 10.1007/978-3-642-20617-7_6617 (cit. on pp. 10, 19).

[55] G. Reinhart, 'Assembly line,' in *CIRP Encyclopedia of Production Engineering*, L. Laperrière and G. Reinhart, Eds., Berlin, Heidelberg, Germany: Springer, 2014, pp. 55–60. DOI: 10.1007/978-3-642-20617-7_8 (cit. on p. 10).

[56] G. Boothroyd, *Assembly automation and product design*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2005. DOI: 10.1201/9781420027358 (cit. on p. 10).

[57] H. Bley, G. Reinhart, G. Seliger, M. Bernardi and T. Korne, 'Appropriate human involvement in assembly and disassembly,' *CIRP Annals*, vol. 53, no. 2, pp. 487–509, 2004. DOI: 10.1016/S0007-8506(07)60026-2 (cit. on p. 10).

[58] M. T. Mason, 'Creation myths: The beginnings of robotics research,' *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 72–77, 2012. DOI: 10.1109/MRA.2012.2191437 (cit. on p. 10).

[59] B. Siciliano and O. Khatib, Eds., *Springer Handbook of Robotics*, 2nd ed. Cham, Switzerland: Springer, 2016. DOI: 10.1007/978-3-319-32552-1 (cit. on pp. 10, 12).

[60] B. Siciliano and O. Khatib, 'Robotics and the handbook,' in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., 2nd ed., Cham, Switzerland: Springer, 2016, ch. 1, pp. 1–6. DOI: 10.1007/978-3-319-32552-1_1 (cit. on pp. 10–12).

[61] E. Garcia, M. A. Jimenez, P. G. De Santos and M. Armada, 'The evolution of robotics research,' *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 90–103, 2007. DOI: 10.1109/MRA.2007.339608 (cit. on pp. 11, 12).

[62] S. J. Russell and P. Norvig, 'Robotics,' in *Artificial intelligence: a modern approach*, ser. Pearson series in artificial intelligence, 4th ed., Hoboken, NJ, USA: Pearson, 2021, ch. 26, pp. 925–980 (cit. on pp. 11, 12).

[63] T. K. Lien, 'Robot,' in *CIRP Encyclopedia of Production Engineering*, L. Laperrière and G. Reinhart, Eds., Berlin, Heidelberg, Germany: Springer, 2014, pp. 1068–1076. DOI: 10.1007/978-3-642-20617-7_6628 (cit. on pp. 11, 12).

[64] L. A. Ballard, S. Sabanovic, J. Kaur and S. Milojevic, 'George charles devol, jr. [history],' *IEEE Robotics & Automation Magazine*, vol. 19, no. 3, pp. 114–119, 2012. DOI: 10.1109/MRA.2012.2206672 (cit. on p. 11).

[65] S. Yun, M.-S. Choi, M.-Y. Cho *et al.*, 'Next-generation furniture assembly by ai and robots: Team sk2y: A winner of the furniture assembly competition at ai-robot challenge 2021,' *IEEE Robotics & Automation Magazine*, vol. 30, no. 2, pp. 96–108, 2023. DOI: 10.1109/MRA.2022.3188214 (cit. on p. 11).

[66] D. Jiang, H. Wang and Y. Lu, 'Mastering the complex assembly task with a dual-arm robot: A novel reinforcement learning method,' *IEEE Robotics & Automation Magazine*, vol. 30, no. 2, pp. 57–66, 2023. DOI: 10.1109/MRA.2023.3262461 (cit. on p. 11).

[67] C. Heyer, 'Human-robot interaction and future industrial robotics applications,' in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 4749–4754. DOI: 10.1109/IROS.2010.5651294 (cit. on p. 11).

[68] R. R. Galin and R. V. Meshcheryakov, 'Human-robot interaction efficiency and human-robot collaboration,' in *Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms*, ser. Studies in Systems, Decision and Control, A. G. Kravets, Ed., vol. 272, Cham, Switzerland: Springer, 2020, pp. 55–63. DOI: 10.1007/978-3-030-37841-7_5 (cit. on p. 11).

[69] A. Grau, M. Indri, L. L. Bello and T. Sauter, 'Industrial robotics in factory automation: From the early stage to the internet of things,' in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, 2017, pp. 6159–6164. DOI: 10.1109/IECON.2017.8217070 (cit. on p. 11).

[70] S. Makris, F. Dietrich, K. Kellens and S. J. Hu, 'Automated assembly of non-rigid objects,' *CIRP Annals*, vol. 72, no. 2, pp. 513–539, 2023. DOI: 10.1016/j.cirp.2023.05.003 (cit. on pp. 12, 16–18, 39).

[71] I. Goodfellow, Y. Bengio and A. Courville, *Deep learning*. Cambridge, MA, USA: MIT press, 2016 (cit. on pp. 12, 16, 17).

[72] R. Li and H. Qiao, 'A survey of methods and strategies for high-precision robotic grasping and assembly tasks—some new trends,' *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 6, pp. 2718–2732, 2019. DOI: 10.1109/TMECH.2019.2945135 (cit. on pp. 12, 13, 16).

[73] J. Yang, C. Wang, B. Jiang, H. Song and Q. Meng, 'Visual perception enabled industry intelligence: State of the art, challenges and prospects,' *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204–2219, 2021. DOI: 10.1109/TII.2020.2998818 (cit. on p. 12).

[74] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Cham, Switzerland: Springer, 2022. DOI: 10.1007/978-3-030-34372-9 (cit. on p. 12).

[75] T. S. Huang, 'Computer vision: Evolution and promise,' in *1996 CERN School of Computing*, C. E. Vandoni, Ed., 1996, pp. 21–25. DOI: 10.5170/CERN-1996-008.21 (cit. on p. 13).

[76] S. J. D. Prince, *Computer Vision: Models, Learning, and Inference*. Cambridge, UK: Cambridge University Press, 2012. DOI: 10.1017/CBO9780511996504 (cit. on p. 13).

[77] B. K. P. Horn, *Robot vision*. Cambridge, MA, USA: MIT press, 1986 (cit. on p. 13).

[78] A. Krizhevsky, I. Sutskever and G. E. Hinton, 'Imagenet classification with deep convolutional neural networks,' in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou and K. Weinberger, Eds., vol. 25, Curran Associates, Inc., 2012 (cit. on p. 13).

[79] A. Krizhevsky, I. Sutskever and G. E. Hinton, 'Imagenet classification with deep convolutional neural networks,' *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017. DOI: 10.1145/3065386 (cit. on p. 13).

[80] H. Harzallah, F. Jurie and C. Schmid, 'Combining efficient object localization and image classification,' in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 237–244. DOI: 10.1109/ICCV.2009.5459257 (cit. on p. 13).

[81] W. Chen, Y. Li, Z. Tian and F. Zhang, '2d and 3d object detection algorithms from images: A survey,' *Array*, vol. 19, p. 100 305, 2023. DOI: 10.1016/j.array.2023.100305 (cit. on pp. 13, 14).

[82] A. B. Amjoud and M. Amrouch, 'Object detection using deep learning, cnns and vision transformers: A review,' *IEEE Access*, vol. 11, pp. 35 479–35 516, 2023. DOI: 10.1109/ACCESS.2023.3266093 (cit. on p. 13).

[83] Z.-Q. Zhao, P. Zheng, S.-T. Xu and X. Wu, 'Object detection with deep learning: A review,' *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212–3232, 2019. DOI: 10.1109/TNNLS.2018.2876865 (cit. on pp. 13, 14).

[84] R. Girshick, J. Donahue, T. Darrell and J. Malik, 'Rich feature hierarchies for accurate object detection and semantic segmentation,' in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587. DOI: 10.1109/CVPR.2014.81 (cit. on pp. 13, 14).

[85] Z. Zou, K. Chen, Z. Shi, Y. Guo and J. Ye, 'Object detection in 20 years: A survey,' *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, 2023. DOI: 10.1109/JPROC.2023.3238524 (cit. on pp. 13, 14, 32).

[86] P. Viola and M. Jones, 'Rapid object detection using a boosted cascade of simple features,' in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. 511–518. DOI: 10.1109/CVPR.2001.990517 (cit. on p. 13).

[87] P. Viola and M. J. Jones, 'Robust real-time face detection,' *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004. DOI: 10.1023/B:VISI.0000013087.49260.fb (cit. on p. 13).

[88] N. Dalal and B. Triggs, 'Histograms of oriented gradients for human detection,' in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 2005, pp. 886–893. DOI: 10.1109/CVPR.2005.177 (cit. on p. 13).

[89] P. Felzenszwalb, D. McAllester and D. Ramanan, 'A discriminatively trained, multiscale, deformable part model,' in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8. DOI: 10.1109/CVPR.2008.4587597 (cit. on p. 13).

[90] R. Girshick, J. Donahue, T. Darrell and J. Malik, 'Region-based convolutional networks for accurate object detection and segmentation,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, 2016. DOI: 10.1109/TPAMI.2015.2437384 (cit. on p. 14).

[91] R. Girshick, 'Fast r-cnn,' in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448. DOI: 10.1109/ICCV.2015.169 (cit. on p. 14).

[92] S. Ren, K. He, R. Girshick and J. Sun, 'Faster r-cnn: Towards real-time object detection with region proposal networks,' in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama and R. Garnett, Eds., vol. 28, Curran Associates, Inc., 2015, pp. 91–99 (cit. on pp. 14, 33).

[93] K. He, X. Zhang, S. Ren and J. Sun, 'Spatial pyramid pooling in deep convolutional networks for visual recognition,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015. DOI: 10.1109/TPAMI.2015.2389824 (cit. on p. 14).

[94] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, 'Feature pyramid networks for object detection,' in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 936–944. DOI: 10.1109/CVPR.2017.106 (cit. on p. 14).

[95] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, 'You only look once: Unified, real-time object detection,' in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788. DOI: 10.1109/CVPR.2016.91 (cit. on p. 14).

[96] J. Redmon and A. Farhadi, *Yolov3: An incremental improvement*, 2018. DOI: 10.48550/arXiv.1804.02767 (cit. on p. 14).

[97] A. Bochkovskiy, C.-Y. Wang and H.-Y. M. Liao, *Yolov4: Optimal speed and accuracy of object detection*, 2020. DOI: 10.48550/arXiv.2004.10934 (cit. on p. 14).

[98] J. Redmon and A. Farhadi, 'Yolo9000: Better, faster, stronger,' in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6517–6525. DOI: 10.1109/CVPR.2017.690 (cit. on p. 14).

[99] C.-Y. Wang, A. Bochkovskiy and H.-Y. M. Liao, 'Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,' in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 7464–7475. DOI: 10.1109/CVPR52729.2023.00721 (cit. on p. 14).

[100] W. Liu, D. Anguelov, D. Erhan *et al.*, 'Ssd: Single shot multibox detector,' in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe and M. Welling, Eds., Cham, Switzerland: Springer, 2016, pp. 21–37. DOI: 10.1007/978-3-319-46448-0_2 (cit. on p. 14).

[101] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, 'Focal loss for dense object detection,' in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2999–3007. DOI: 10.1109/ICCV.2017.324 (cit. on p. 14).

[102] H. Law and J. Deng, 'Cornernet: Detecting objects as paired keypoints,' in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu and Y. Weiss, Eds., Cham, Switzerland: Springer, 2018, pp. 765–781. DOI: 10.1007/978-3-030-01264-9_45 (cit. on p. 14).

[103] X. Zhou, D. Wang and P. Krähenbühl, *Objects as points*, 2019. DOI: 10.48550/arXiv.1904.07850 (cit. on p. 14).

[104] A. Vaswani, N. Shazeer, N. Parmar *et al.*, 'Attention is all you need,' in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio *et al.*, Eds., vol. 30, Curran Associates, Inc., 2017 (cit. on p. 14).

[105] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan and M. Shah, 'Transformers in vision: A survey,' *ACM Computing Surveys*, vol. 54, no. 10s, 2022. DOI: 10.1145/3505244 (cit. on p. 14).

[106] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov and S. Zagoruyko, 'End-to-end object detection with transformers,' in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox and J.-M. Frahm, Eds., Cham, Switzerland: Springer, 2020, pp. 213–229. DOI: 10.1007/978-3-030-58452-8_13 (cit. on p. 14).

[107] X. Zhu, W. Su, L. Lu, B. Li, X. Wang and J. Dai, 'Deformable detr: Deformable transformers for end-to-end object detection,' in *ICLR2021 - 9th International Conference on Learning Representations*, 2021 (cit. on p. 14).

[108] H. Zhang, F. Li, S. Liu *et al.*, 'Dino: Detr with improved denoising anchor boxes for end-to-end object detection,' in *ICLR2023 - 11th International Conference on Learning Representations*, 2023 (cit. on p. 14).

[109] F. Li, H. Zhang, H. Xu *et al.*, 'Mask dino: Towards a unified transformer-based framework for object detection and segmentation,' in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 3041–3050. DOI: 10.1109/CVPR52729.2023.00297 (cit. on p. 14).

[110] C. Sahin, G. Garcia-Hernando, J. Sock and T.-K. Kim, 'A review on object pose recovery: From 3d bounding box detectors to full 6d pose estimators,' *Image and Vision Computing*, vol. 96, p. 103 898, 2020. DOI: 10.1016/j.imavis.2020.103898 (cit. on pp. 14, 15).

[111] S. Hoque, M. Y. Arafat, S. Xu, A. Maiti and Y. Wei, 'A comprehensive review on 3d object detection and 6d pose estimation with deep learning,' *IEEE Access*, vol. 9, pp. 143 746–143 770, 2021. DOI: 10.1109/ACCESS.2021.3114399 (cit. on p. 14).

[112] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu and M. Bennamoun, 'Deep learning for 3d point clouds: A survey,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 12, pp. 4338–4364, 2021. DOI: 10.1109/TPAMI.2020.3005434 (cit. on p. 14).

[113] W. Liang, P. Xu, L. Guo, H. Bai, Y. Zhou and F. Chen, 'A survey of 3d object detection,' *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29 617–29 641, 2021. DOI: 10.1007/s11042-021-11137-y (cit. on pp. 14, 37).

[114] Y. Zhu, M. Li, W. Yao and C. Chen, 'A review of 6d object pose estimation,' in *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, vol. 10, 2022, pp. 1647–1655. DOI: 10.1109/ITAIC54216.2022.9836663 (cit. on p. 14).

[115] J. Guan, Y. Hao, Q. Wu, S. Li and Y. Fang, 'A survey of 6dof object pose estimation methods for different application scenarios,' *Sensors*, vol. 24, no. 4, p. 1076, 2024. DOI: 10.3390/s24041076 (cit. on pp. 14, 15).

[116] F. Gorschlüter, P. Rojtberg and T. Pöllabauer, 'A survey of 6d object detection based on 3d models for industrial applications,' *Journal of Imaging*, vol. 8, no. 3, p. 53, 2022. DOI: 10.3390/jimaging8030053 (cit. on p. 14).

[117] Z. He, W. Feng, X. Zhao and Y. Lv, '6d pose estimation of objects: Recent technologies and challenges,' *Applied Sciences*, vol. 11, no. 1, p. 228, 2021. DOI: 10.3390/app11010228 (cit. on pp. 14, 15).

[118] G. J. Agin, 'Computer vision systems for industrial inspection and assembly,' *Computer*, vol. 13, no. 5, pp. 11–20, 1980. DOI: 10.1109/MC.1980.1653613 (cit. on p. 15).

[119] A. Soini, 'Machine vision technology take-up in industrial applications,' in *ISPA 2001. Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis. In conjunction with 23rd International Conference on Information Technology Interfaces (IEEE Cat.*, 2001, pp. 332–338. DOI: 10.1109/ISPA.2001.938651 (cit. on p. 15).

[120] M. Ejiri, 'Machine vision technology: Past, present and future,' in *EEE International Workshop on Intelligent Robots and Systems, Towards a New Frontier of Applications*, vol. 1, 1990, pp. XXIX–XXXX. DOI: 10.1109/IROS.1990.262354 (cit. on p. 15).

[121] J. Jolion, 'Computer vision methodologies,' *CVGIP: Image Understanding*, vol. 59, no. 1, pp. 53–71, 1994. DOI: 10.1006/ciun.1994.1004 (cit. on pp. 15, 27).

[122] J. M. Jolion, 'Methodology for design of image analysis systems: An application to electron microscopy,' Ph.D. dissertation, Lyon, France, 1987 (cit. on p. 15).

[123] H. Freeman, Ed., *Machine vision: algorithms, architectures, and systems* (Perspectives in Computing). San Diego, CA, USA: Academic Press, Inc., 1988, vol. 22. DOI: 10.1016/B978-0-12-266720-6.X5001-3 (cit. on p. 15).

[124] H. Freeman, Ed., *Machine vision for inspection and measurement* (Perspectives in Computing). San Diego, CA, USA: Academic Press, Inc., 1989, vol. 24. DOI: 10.1016/B978-0-122-66719-0.X5001-5 (cit. on p. 15).

[125] H. Golnabi and A. Asadpour, 'Design and application of industrial machine vision systems,' *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 6, pp. 630–637, 2007. DOI: 10.1016/j.rcim.2007.02.005 (cit. on p. 15).

[126] V. Kuts, T. Otto, T. Tähemaa, K. Bukhari and T. Pataraia, 'Adaptive industrial robots using machine vision,' in *Proceedings of the ASME 2018 International Mechanical Engineering Congress and Exposition*, ser. ASME International Mechanical Engineering Congress and Exposition, vol. 2, Pittsburgh, PA, USA: ASME, 2018, V002T02A093. DOI: 10.1115/IMECE2018-86720 (cit. on p. 15).

[127] F. K. Konstantinidis, S. G. Mouroutsos and A. Gasteratos, 'The role of machine vision in industry 4.0: An automotive manufacturing perspective,' in *2021 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2021, pp. 1–6. DOI: 10.1109/IST50367.2021.9651453 (cit. on p. 16).

[128] Y. Guo, M. Bennamoun, F. Sohel, M. Lu and J. Wan, '3d object recognition in cluttered scenes with local surface features: A survey,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2270–2287, 2014. DOI: 10.1109/TPAMI.2014.2316828 (cit. on p. 16).

[129] X. Jiang, K. mo Koo, K. Kikuchi, A. Konno and M. Uchiyama, 'Robotized assembly of a wire harness in a car production line,' *Advanced Robotics*, vol. 25, no. 3-4, pp. 473–489, 2011. DOI: 10.1163/016918610X551782 (cit. on pp. 16, 29, 31, 32).

[130] S. Ojha and S. Sakhare, 'Image processing techniques for object tracking in video surveillance- a survey,' in *2015 International Conference on Pervasive Computing (ICPC)*, 2015, pp. 1–6. DOI: 10.1109/PERVASIVE.2015.7087180 (cit. on p. 16).

[131] L. Alzubaidi, J. Zhang, A. J. Humaidi *et al.*, 'Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions,' *Journal of Big Data*, vol. 8, no. 1, p. 53, 2021. DOI: 10.1186/s40537-021-00444-8 (cit. on p. 16).

[132] H. G. Nguyen, R. Habiboglu and J. Franke, 'Enabling deep learning using synthetic data: A case study for the automotive wiring harness manufacturing,' *Procedia CIRP*, vol. 107, pp. 1263–1268, 2022. DOI: 10.1016/j.procir.2022.05.142 (cit. on pp. 16, 37).

[133] K. Kimble, J. Albrecht, M. Zimmerman and J. Falco, 'Performance measures to benchmark the grasping, manipulation, and assembly of deformable objects typical to manufacturing applications,' *Frontiers in Robotics and AI*, vol. 9, 2022. DOI: 10.3389/frobt.2022.999348 (cit. on p. 16).

[134] J. McCarthy, M. L. Minsky, N. Rochester and C. E. Shannon, 'A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955,' *AI Magazine*, vol. 27, no. 4, p. 12, 2006. DOI: 10.1609/aimag.v27i4.1904 (cit. on p. 16).

[135] S. J. Russell and P. Norvig, 'Learning from examples,' in *Artificial intelligence: a modern approach*, ser. Pearson series in artificial intelligence, 4th ed., Hoboken, NJ, USA: Pearson, 2021, ch. 19, pp. 651–720 (cit. on p. 16).

[136] Y. LeCun, Y. Bengio and G. Hinton, 'Deep learning,' *nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539 (cit. on pp. 16, 17, 32, 37).

[137] S. J. Russell and P. Norvig, 'Deep learning,' in *Artificial intelligence: a modern approach*, ser. Pearson series in artificial intelligence, 4th ed., Hoboken, NJ, USA: Pearson, 2021, ch. 21, pp. 750–788 (cit. on p. 17).

[138] W. S. McCulloch and W. Pitts, 'A logical calculus of the ideas immanent in nervous activity,' *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943. DOI: 10.1007/BF02478259 (cit. on p. 17).

[139] X. Zhang, X. Zhang and W. Wang, 'Convolutional neural network,' in *Intelligent Information Processing with Matlab*. Singapore: Springer, 2023, pp. 39–71. DOI: 10.1007/978-981-99-6449-9_2 (cit. on p. 17).

[140] S. Cong and Y. Zhou, 'A review of convolutional neural network architectures and their optimizations,' *Artificial Intelligence Review*, vol. 56, no. 3, pp. 1905–1969, 2023. DOI: 10.1007/s10462-022-10213-5 (cit. on p. 17).

[141] Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, 'A survey of convolutional neural networks: Analysis, applications, and prospects,' *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, 2022. DOI: 10.1109/TNNLS.2021.3084827 (cit. on p. 17).

[142] I. Kovalenko, K. Barton, J. Moyne and D. M. Tilbury, 'Opportunities and challenges to integrate artificial intelligence into manufacturing systems: Thoughts from a panel discussion [opinion],' *IEEE Robotics & Automation Magazine*, vol. 30, no. 2, pp. 109–112, 2023. DOI: 10.1109/MRA.2023.3262464 (cit. on p. 17).

[143] J. Lee, 'Definition and meaning of industrial ai,' in *Industrial AI: Applications with Sustainable Performance*. Singapore: Springer, 2020, pp. 33–61. DOI: 10.1007/978-981-15-2144-7_3 (cit. on p. 17).

[144] G. Chryssolouris, K. Alexopoulos and Z. Arkouli, *A Perspective on Artificial Intelligence in Manufacturing*, 1st ed. Cham, Switzerland: Springer, 2023. DOI: 10.1007/978-3-031-21828-6 (cit. on p. 17).

[145] A. Simeth and P. Plapper, 'Artificial intelligence based robotic automation of manual assembly tasks for intelligent manufacturing,' in *Smart, Sustainable Manufacturing in an Ever-Changing World*, K. von Leipzig, N. Sacks and M. Mc Clelland, Eds., Cham, Switzerland: Springer, 2023, pp. 137–148. DOI: 10.1007/978-3-031-15602-1_11 (cit. on p. 17).

[146] K. Galassi and G. Palli, 'Robotic wires manipulation for switchgear cabling and wiring harness manufacturing,' in *2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*, 2021, pp. 531–536. DOI: 10.1109/ICPS49255.2021.9468128 (cit. on p. 17).

[147] N. Lv, J. Liu and Y. Jia, 'Dynamic modeling and control of deformable linear objects for single-arm and dual-arm robot manipulations,' *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2341–2353, 2022. DOI: 10.1109/TRO.2021.3139838 (cit. on p. 17).

[148] M. Saadat and P. Nan, 'Industrial applications of automatic manipulation of flexible materials,' *Industrial Robot: An International Journal*, vol. 29, no. 5, pp. 434–442, 2002. DOI: 10.1108/01439910210440255 (cit. on pp. 17–20).

[149] J. Sanchez, J.-A. Corrales, B.-C. Bouzgarrou and Y. Mezouar, 'Robotic manipulation and sensing of deformable objects in domestic and industrial applications: A survey,' *The International Journal of Robotics Research*, vol. 37, no. 7, pp. 688–716, 2018. DOI: 10.1177/0278364918779698 (cit. on p. 18).

[150] F. Nadon, A. J. Valencia and P. Payeur, 'Multi-modal sensing and robotic manipulation of non-rigid objects: A survey,' *Robotics*, vol. 7, no. 4, p. 74, 2018. DOI: 10.3390/robotics7040074 (cit. on p. 18).

[151] H. Yin, A. Varava and D. Kragic, 'Modeling, learning, perception, and control methods for deformable object manipulation,' *Science Robotics*, vol. 6, no. 54, eabd8803, 2021. DOI: 10.1126/scirobotics.abd8803 (cit. on p. 18).

[152] J. Zhu, A. Cherubini, C. Dune *et al.*, 'Challenges and outlook in robotic manipulation of deformable objects,' *IEEE Robotics & Automation Magazine*, vol. 29, no. 3, pp. 67–77, 2022. DOI: 10.1109/MRA.2022.3147415 (cit. on p. 18).

[153] X. Li, X. Su and Y.-H. Liu, 'Vision-based robotic manipulation of flexible pcbs,' *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 6, pp. 2739–2749, 2018. DOI: 10.1109/TMECH.2018.2869147 (cit. on p. 18).

[154] M. C. Gemici and A. Saxena, 'Learning haptic representation for manipulating deformable food objects,' in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 638–645. DOI: 10.1109/IROS.2014.6942626 (cit. on p. 18).

[155] L. Cao, X. Li, P. T. Phan *et al.*, 'Sewing up the wounds: A robotic suturing system for flexible endoscopy,' *IEEE Robotics & Automation Magazine*, vol. 27, no. 3, pp. 45–54, 2020. DOI: 10.1109/MRA.2019.2963161 (cit. on p. 18).

[156] F. Zhang and Y. Demiris, 'Learning garment manipulation policies toward robot-assisted dressing,' *Science Robotics*, vol. 7, no. 65, eabm6010, 2022. DOI: 10.1126/scirobotics.abm6010 (cit. on p. 18).

[157] P. Jiménez, 'Survey on model-based manipulation planning of deformable objects,' *Robotics and Computer-Integrated Manufacturing*, vol. 28, no. 2, pp. 154–163, 2012. DOI: 10.1016/j.rcim.2011.08.002 (cit. on pp. 18, 19).

[158] F. Guo, H. Lin and Y.-B. Jia, 'Squeeze grasping of deformable planar objects with segment contacts and stick/slip transitions,' in *2013 IEEE International Conference on Robotics and Automation*, 2013, pp. 3736–3741. DOI: 10.1109/ICRA.2013.6631102 (cit. on p. 18).

[159] S. Javdani, S. Tandon, J. Tang, J. F. O'Brien and P. Abbeel, 'Modeling and perception of deformable one-dimensional objects,' in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 1607–1614. DOI: 10.1109/ICRA.2011.5980431 (cit. on p. 18).

[160] A. Keipour, M. Bandari and S. Schaal, 'Deformable one-dimensional object detection for routing and manipulation,' *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4329–4336, 2022. DOI: 10.1109/LRA.2022.3146920 (cit. on p. 18).

[161] C. Chen and Y. Zheng, 'Deformation identification and estimation of one-dimensional objects by using vision sensors,' in *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, vol. 3, 1991, pp. 2306–2311. DOI: 10.1109/ROBOT.1991.131538 (cit. on p. 18).

[162] W. Wang, D. Berenson and D. Balkcom, 'An online method for tight-tolerance insertion tasks for string and rope,' in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 2488–2495. DOI: 10.1109/ICRA.2015.7139532 (cit. on p. 18).

[163] J.-E. Byun and T. i Nagata, 'Determining the 3-d pose of a flexible object by stereo matching of curvature representations,' *Pattern Recognition*, vol. 29, no. 8, pp. 1297–1307, 1996. DOI: 10.1016/0031-3203(95)00165-4 (cit. on pp. 18, 21).

[164] G. E. Navas-Reascos, D. Romero, J. Stahre and A. Caballero-Ruiz, 'Wire harness assembly process supported by collaborative robots: Literature review and call for r&d,' *Robotics*, vol. 11, no. 3, p. 65, 2022. DOI: 10.3390/robotics11030065 (cit. on p. 18).

[165] M. A. Lee, Y. Zhu, P. Zachares *et al.*, 'Making sense of vision and touch: Learning multimodal representations for contact-rich tasks,' *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 582–596, 2020. DOI: 10.1109/TRO.2019.2959445 (cit. on p. 18).

[166] H. Zhou, S. Li, Q. Lu and J. Qian, 'A practical solution to deformable linear object manipulation: A case study on cable harness connection,' in *2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM)*, 2020, pp. 329–333. DOI: 10.1109/ICARM49381.2020.9195380 (cit. on pp. 18, 19, 31, 32).

[167] D. Andronas, Z. Arkouli, N. Zacharaki *et al.*, 'On the perception and handling of deformable objects - a robotic cell for white goods industry,' *Robotics and Computer-Integrated Manufacturing*, vol. 77, p. 102 358, 2022. DOI: 10.1016/j.rcim.2022.102358 (cit. on p. 18).

[168] T. Hermansson, R. Bohlin, J. S. Carlson and R. Söderberg, 'Automatic assembly path planning for wiring harness installations,' *Journal of Manufacturing Systems*, vol. 32, no. 3, pp. 417–422, 2013. DOI: 10.1016/j.jmsy.2013.04.006 (cit. on p. 18).

[169] A. Nair, D. Chen, P. Agrawal *et al.*, 'Combining self-supervised learning and imitation for vision-based rope manipulation,' in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 2146–2153. DOI: 10.1109/ICRA.2017.7989247 (cit. on pp. 18, 20, 21).

[170] M. Yan, Y. Zhu, N. Jin and J. Bohg, 'Self-supervised learning of state estimation for manipulating deformable linear objects,' *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2372–2379, 2020. DOI: 10.1109/LRA.2020.2969931 (cit. on pp. 18, 20, 21).

[171] S. Heydaryan, J. Suaza Bedolla and G. Belingardi, 'Safety design and development of a human-robot collaboration assembly process in the automotive industry,' *Applied Sciences*, vol. 8, no. 3, p. 344, 2018. DOI: 10.3390/app8030344 (cit. on p. 19).

[172] M. Despeisse, B. Johansson, J. Bokrantz *et al.*, 'Battery production systems: State of the art and future developments,' in *Advances in Production Management Systems. Production Management Systems for Responsible Manufacturing, Service, and Logistics Futures*, E. Alfnes, A. Romsdal, J. O. Strandhagen, G. von Cieminski and D. Romero, Eds., Cham, Switzerland: Springer, 2023, pp. 521–535. DOI: 10.1007/978-3-031-43688-8_36 (cit. on p. 19).

[173] S. Chandra, I. J. Chung, A. Esmail, M. Blum and R. Bhandari, *Wiring system architecture*, 2022 (cit. on p. 19).

[174] K. mo Koo, X. Jiang, K. Kikuchi, A. Konno and M. Uchiyama, 'Development of a robot car wiring system,' in *2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2008, pp. 862–867. DOI: 10.1109/AIM.2008.4601774 (cit. on pp. 19, 31, 32).

[175] X. Zhang, Y. Domae, W. Wan and K. Harada, 'Learning efficient policies for picking entangled wire harnesses: An approach to industrial bin picking,' *IEEE Robotics and Automation Letters*, vol. 8, no. 1, pp. 73–80, 2023. DOI: 10.1109/LRA.2022.3222995 (cit. on pp. 19, 31, 32).

[176] E. Aguirre and B. Raucent, 'Economic comparison of wire harness assembly systems,' *Journal of Manufacturing Systems*, vol. 13, no. 4, pp. 276–288, 1994. DOI: 10.1016/0278-6125(94)90035-3 (cit. on pp. 19, 29).

[177] M. Zürn, A. Kienzlen, L. Klingel *et al.*, 'Deep learning-based instance segmentation for feature extraction of branched deformable linear objects for robotic manipulation,' in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, 2023, pp. 1–6. DOI: 10.1109/CASE56687.2023.10260646 (cit. on p. 19).

[178] T. Toner, V. Molazadeh, M. Saez, D. M. Tilbury and K. Barton, *Sequential manipulation of deformable linear object networks with endpoint pose measurements using adaptive model predictive control*, 2024. DOI: 10.48550/arXiv.2402.10372 (cit. on pp. 19, 21).

[179] B. Sun, F. Chen, H. Sasaki and T. Fukuda, 'Robotic wiring harness assembly system for fault-tolerant electric connectors mating,' in *2010 International Symposium on Micro-NanoMechatronics and Human Science*, 2010, pp. 202–205. DOI: 10.1109/MHS.2010.5669533 (cit. on pp. 19, 29, 31, 32).

[180] X. Jiang, Y. Nagaoka, K. Ishii, S. Abiko, T. Tsujita and M. Uchiyama, 'Robotized recognition of a wire harness utilizing tracing operation,' *Robotics and Computer-Integrated Manufacturing*, vol. 34, pp. 52–61, 2015. DOI: 10.1016/j.rcim.2014.12.002 (cit. on pp. 20, 21, 31, 32).

[181] X. Wang, X. L. Wang and D. M. Wilkes, 'An automated vision based on-line novel percept detection method for a mobile robot,' *Robotics and Autonomous Systems*, vol. 60, no. 10, pp. 1279–1294, 2012. DOI: 10.1016/j.robot.2012.06.005 (cit. on p. 20).

[182] E. Martinez-Martin and A. P. del Pobil, 'Vision for robust robot manipulation,' *Sensors*, vol. 19, no. 7, p. 1648, 2019. DOI: 10.3390/s19071648 (cit. on p. 20).

[183] M. T. Shahria, M. S. H. Sunny, M. I. I. Zarif, J. Ghommam, S. I. Ahamed and M. H. Rahman, 'A comprehensive review of vision-based robotic applications: Current state, components, approaches, barriers, and potential solutions,' *Robotics*, vol. 11, no. 6, p. 139, 2022. DOI: 10.3390/robotics11060139 (cit. on p. 20).

[184] S. Zhaole, H. Zhou, L. Nanbo, L. Chen, J. Zhu and R. B. Fisher, 'A robust deformable linear object perception pipeline in 3d: From segmentation to reconstruction,' *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 843–850, 2024. DOI: 10.1109/LRA.2023.3337695 (cit. on p. 21).

[185] K. Nottensteiner, A. Sachtler and A. Albu-Schäffer, 'Towards autonomous robotic assembly: Using combined visual and tactile sensing for adaptive task execution,' *Journal of Intelligent & Robotic Systems*, vol. 101, no. 3, p. 49, 2021. DOI: 10.1007/s10846-020-01303-z (cit. on pp. 21, 29).

[186] J. Guo, J. Zhang, Y. Gai, D. Wu and K. Chen, 'Visual recognition method for deformable wires in aircrafts assembly based on sequential segmentation and probabilisitic estimation,' in *2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)*, vol. 6, 2022, pp. 598–603. DOI: 10.1109/ITOEC53115.2022.9734432 (cit. on pp. 21, 31, 32).

[187] F. Duan, J. T. C. Tan and T. Arai, 'A new human-robot collaboration assembly system for cellular manufacturing,' in *Proceedings of the 30th Chinese Control Conference*, 2011, pp. 5468–5473 (cit. on p. 21).

[188] N. S. Joshi, S. Singh, M. Krugh and L. Mears, 'Background noise mitigation of dual microphone system for defect detection in electrical cable connection,' *Procedia Manufacturing*, vol. 26, pp. 1287–1295, 2018. DOI: 10.1016/j.promfg.2018.07.139 (cit. on p. 21).

[189] B. D. Slife and R. N. Williams, *What's Behind the Research? Discovering Hidden Assumptions in the Behavioral Sciences*. Thousand Oaks, CA, USA: SAGE Publications, Inc., 1995. DOI: 10.4135/9781483327372 (cit. on p. 23).

[190] J. W. Creswell and J. D. Creswell, *Research design: qualitative, quantitative, and mixed methods approaches*, 6th ed. Thousand Oaks, CA, USA: SAGE Publications, Inc., 2022 (cit. on pp. 23, 24).

[191] R. Hirschheim, 'Information systems epistemology: An historical perspective,' *Research methods in information systems*, pp. 9–33, 1985 (cit. on p. 23).

[192] E. G. Guba, 'The alternative paradigm dialog,' in *The Paradigm Dialog*, SAGE Publications, Inc, 1990, pp. 17–30 (cit. on p. 23).

[193] J. W. Creswell, *Research design: qualitative, quantitative, and mixed methods approaches*, 4th ed. Thousand Oaks, CA, USA: SAGE Publications, Inc., 2014 (cit. on pp. 23, 27, 28, 40).

[194] J. Paley, 'Positivism,' in *The SAGE Encyclopedia of Qualitative Research Methods*, L. M. Given, Ed., vol. 2, Thousand Oaks, CA, USA: SAGE Publications, Inc., 2008, pp. 647–650. DOI: 10.4135/9781412963909 (cit. on p. 23).

[195] K. Williamson, F. Burstein and S. McKemmish, 'The two major traditions of research,' in *Research Methods for Students, Academics and Professionals*, ser. Topics in Australasian Library and Information Studies, K. Williamson, A. Bow, F. Burstein *et al.*, Eds., 2nd ed., Chandos Publishing, 2002, ch. 2, pp. 25–47. DOI: 10.1016/B978-1-876938-42-0.50009-5 (cit. on pp. 23, 26).

[196] J. Laird, 'Xi.-positivism, empiricism, and metaphysics,' *Proceedings of the Aristotelian Society*, vol. 39, no. 1, pp. 207–224, 1939. DOI: 10.1093/aristotelian/39.1.207 (cit. on p. 23).

[197] D. C. Phillips and N. C. Burbules, *Postpositivism and educational research*. Lanham, MD, USA: Rowman & Littlefield Publishers, Inc., 2000 (cit. on p. 23).

[198] N. J. Fox, 'Postpositivism,' in *The SAGE Encyclopedia of Qualitative Research Methods*, L. M. Given, Ed., vol. 2, Thousand Oaks, CA, USA: SAGE Publications, Inc., 2008, pp. 660–664. DOI: 10.4135/9781412963909 (cit. on p. 23).

[199] T. E. Costantino, 'Constructivism,' in *The SAGE Encyclopedia of Qualitative Research Methods*, L. M. Given, Ed., vol. 2, Thousand Oaks, CA, USA: SAGE Publications, Inc., 2008, pp. 116–120. DOI: 10.4135/9781412963909 (cit. on p. 24).

[200] E. Bell, A. Bryman and B. Harley, *Business research methods*, 6th ed. Oxford, UK: Oxford University Press, 2022 (cit. on p. 24).

[201] K. Peffers, T. Tuunanen, M. A. Rothenberger and S. Chatterjee, 'A design science research methodology for information systems research,' *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007. DOI: 10.2753/MIS0742-1222240302 (cit. on pp. 24, 25).

[202] A. Hevner and S. Chatterjee, 'Design science research in information systems,' in *Design Research in Information Systems: Theory and Practice*, ser. Integrated Series in Information Systems, vol. 22, Boston, MA, USA: Springer, 2010, ch. 2, pp. 9–22. DOI: 10.1007/978-1-4419-5653-8_2 (cit. on p. 24).

[203] J. Morse, 'Procedures and practice of mixed method design: Maintaining control, rigor, and complexity,' in *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, A. Tashakkori and C. Teddlie, Eds., 2nd ed., SAGE Publications, Inc., 2010, pp. 339–352. DOI: 10.4135/9781506335193 (cit. on pp. 24, 25).

[204] S. Gregor and A. R. Hevner, 'Positioning and presenting design science research for maximum impact,' *MIS Quarterly*, vol. 37, no. 2, pp. 337–355, 2013. DOI: 10.25300/MISQ/2013/37.2.01 (cit. on p. 24).

[205] A. R. Hevner, S. T. March, J. Park and S. Ram, 'Design science in information systems research,' *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004. DOI: 10.2307/25148625 (cit. on p. 24).

[206] M. Petticrew, F. Song, P. Wilson and K. Wright, 'Quality-assessed reviews of health care interventions and the database of abstracts of reviews of effectiveness (dare),' *International Journal of Technology Assessment in Health Care*, vol. 15, no. 4, pp. 671–678, 1999. DOI: 10.1017/S0266462399015469 (cit. on pp. 27, 28, 31, 40).

[207] N. K. Denzin, *The research act: A theoretical introduction to sociological methods*, 1st ed. New York, NY, USA: Routledge, 2009. DOI: 10.4324/9781315134543 (cit. on pp. 27, 28, 30, 40).

[208] V. J. Janesick, 'Peer debriefing,' in *The Blackwell Encyclopedia of Sociology*, G. Ritzer, Ed., John Wiley & Sons, Ltd, 2015. DOI: 10.1002/9781405165518.wbeosp014.pub2 (cit. on pp. 27, 28, 40).

[209] T. Greenhalgh and R. Peacock, 'Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources,' *BMJ*, vol. 331, no. 7524, pp. 1064–1065, 2005. DOI: 10.1136/bmj.38636.593461.68 (cit. on pp. 26, 27, 31).

[210] M. W. Tracey, 'Design and development research: A model validation case,' *Educational Technology Research and Development*, vol. 57, no. 4, pp. 553–571, 2009. DOI: 10.1007/s11423-007-9075-0 (cit. on pp. 27, 28, 40).

[211] B. Kitchenham, 'Procedures for performing systematic reviews,' Keele University, Keele, Staffs, UK, Tech. Rep. TR/SE-0401, 2004 (cit. on pp. 26–28, 30).

[212] J. Rowley and F. Slack, 'Conducting a literature review,' *Management research news*, vol. 27, no. 6, pp. 31–39, 2004. DOI: 10.1108/01409170410784185 (cit. on p. 26).

[213] D. Denyer and D. Tranfield, 'Producing a systematic review,' in *The SAGE Handbook of Organizational Research Methods*, D. A. Buchanan and A. Bryman, Eds., Sage Publications Ltd., 2009, pp. 671–689 (cit. on p. 26).

[214] H. K. Mohajan, 'Two criteria for good measurements in research: Validity and reliability,' *Annals of Spiru Haret University*, vol. 17, no. 3, pp. 58–82, 2017. DOI: 10.26458/1746 (cit. on pp. 27, 40).

[215] L. Sürücü and A. Maslakci, 'Validity and reliability in quantitative research,' *Business & Management Studies: An International Journal*, vol. 8, no. 3, pp. 2694–2726, 2020. DOI: 10.15295/bmij.v8i3.1540 (cit. on pp. 27, 40).

[216] R. Adcock and D. Collier, 'Measurement validity: A shared standard for qualitative and quantitative research,' *American Political Science Review*, vol. 95, no. 3, pp. 529–546, 2001. DOI: 10.1017/S0003055401003100 (cit. on pp. 27, 40).

[217] C. M. Patino and J. C. Ferreira, 'Internal and external validity: Can you apply research study results to your patients?' *Jornal brasileiro de pneumologia*, vol. 44, no. 3, p. 183, 2018. DOI: 10.1590/S1806-37562018000000164 (cit. on pp. 27, 40).

[218] G. R. Gibbs, *Analyzing Qualitative Data*. London, UK: SAGE Publications, Ltd, 2007. DOI: 10.4135/9781849208574 (cit. on pp. 27, 28).

[219] K. Petersen, S. Vakkalanka and L. Kuzniarz, 'Guidelines for conducting systematic mapping studies in software engineering: An update,' *Information and Software Technology*, vol. 64, pp. 1–18, 2015. DOI: 10.1016/j.infsof.2015.03.007 (cit. on p. 28).

[220] A. Botta, W. de Donato, V. Persico and A. Pescapé, 'Integration of cloud computing and internet of things: A survey,' *Future Generation Computer Systems*, vol. 56, pp. 684–700, 2016. DOI: 10.1016/j.future.2015.09.021 (cit. on p. 28).

[221] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto and S. G. S. Alcalá, 'A systematic literature review of machine learning methods applied to predictive maintenance,' *Computers & Industrial Engineering*, vol. 137, p. 106 024, 2019. DOI: 10.1016/j.cie.2019.106024 (cit. on p. 28).

[222] J. Dalzochio, R. Kunst, E. Pignaton *et al.*, 'Machine learning and reasoning for predictive maintenance in industry 4.0: Current status and challenges,' *Computers in Industry*, vol. 123, p. 103 298, 2020. DOI: 10.1016/j.compind.2020.103298 (cit. on p. 28).

[223] O. Ali, M. Ally, Clutterbuck and Y. Dwivedi, 'The state of play of blockchain technology in the financial services sector: A systematic literature review,' *International Journal of Information Management*, vol. 54, p. 102 199, 2020. DOI: 10.1016/j.ijinfomgt.2020.102199 (cit. on p. 28).

[224] A. Zuiderwijk, Y.-C. Chen and F. Salem, 'Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda,' *Government Information Quarterly*, vol. 38, no. 3, p. 101 577, 2021. DOI: 10.1016/j.giq.2021.101577 (cit. on p. 28).

[225] A. F. Borges, F. J. Laurindo, M. M. Spínola, R. F. Gonçalves and C. A. Mattos, 'The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions,' *International Journal of Information Management*, vol. 57, p. 102 225, 2021. DOI: 10.1016/j.ijinfomgt.2020.102225 (cit. on p. 28).

[226] C. Collins, D. Dennehy, K. Conboy and P. Mikalef, 'Artificial intelligence in information systems research: A systematic literature review and research agenda,' *International Journal of Information Management*, vol. 60, p. 102 383, 2021. DOI: 10.1016/j.ijinfomgt.2021.102383 (cit. on p. 28).

[227] M. Haghi Kashani, M. Madanipour, M. Nikravan, P. Asghari and E. Mahdipour, 'A systematic review of iot in healthcare: Applications, techniques, and trends,' *Journal of Network and Computer Applications*, vol. 192, p. 103 164, 2021. DOI: 10.1016/j.jnca.2021.103164 (cit. on p. 28).

[228] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman, 'Systematic literature reviews in software engineering - a systematic literature review,' *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, 2009. DOI: 10.1016/j.infsof.2008.09.009 (cit. on pp. 28, 40).

[229] S. Saleem, N. F. Khan, S. Zafar and N. Raza, 'Systematic literature reviews in cyberbullying/cyber harassment: A tertiary study,' *Technology in Society*, vol. 70, p. 102 055, 2022. DOI: 10.1016/j.techsoc.2022.102055 (cit. on p. 28).

[230] M. Hammersley, 'Assessing validity in social research,' in *The SAGE Handbook of Social Research Methods*, P. Alasuutari, L. Bickman and J. Brannen, Eds., London, England: SAGE Publications Ltd, 2008, ch. 4, pp. 42–53. DOI: 10.4135/9781446212165 (cit. on pp. 28, 40).

[231] J. Tilindis and V. Kleiza, 'The effect of learning factors due to low volume order fluctuations in the automotive wiring harness production,' *Procedia CIRP*, vol. 19, pp. 129–134, 2014. DOI: 10.1016/j.procir.2014.05.019 (cit. on p. 29).

[232] A. Hormann and U. Rembold, 'Development of an advanced robot for autonomous assembly,' in *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, vol. 3, 1991, pp. 2452–2457. DOI: 10.1109/ROBOT.1991.131992 (cit. on p. 29).

[233] C. Feng, Y. Xiao, A. Willette, W. McGee and V. R. Kamat, 'Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites,' *Automation in Construction*, vol. 59, pp. 128–138, 2015. DOI: 10.1016/j.autcon.2015.06.002 (cit. on p. 29).

[234] R. Ahmad and P. Plapper, 'Safe and automated assembly process using vision assisted robot manipulator,' *Procedia CIRP*, vol. 41, pp. 771–776, 2016. DOI: 10.1016/j.procir.2015.12.129 (cit. on p. 29).

[235] M. J. Page, J. E. McKenzie, P. M. Bossuyt *et al.*, 'The prisma 2020 statement: An updated guideline for reporting systematic reviews,' *BMJ*, vol. 372, no. n71, 2021. DOI: 10.1136/bmj.n71 (cit. on p. 31).

[236] P. Di, J. Huang, F. Chen, H. Sasaki and T. Fukuda, 'Hybrid vision-force guided fault tolerant robotic assembly for electric connectors,' in *2009 International Symposium on Micro-NanoMechatronics and Human Science*, 2009, pp. 86–91. DOI: 10.1109/MHS.2009.5352078 (cit. on pp. 31, 32).

[237] X. Jiang, K.-m. Koo, K. Kikuchi, A. Konno and M. Uchiyama, 'Robotized assembly of a wire harness in car production line,' in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 490–495. DOI: 10.1109/IROS.2010.5653133 (cit. on pp. 31, 32).

[238] P. Di, F. Chen, H. Sasaki, J. Huang, T. Fukuda and T. Matsuno, 'Vision-force guided monitoring for mating connectors in wiring harness assembly systems,' *Journal of Robotics and Mechatronics*, vol. 24, no. 4, pp. 666–676, 2012. DOI: 10.20965/jrm.2012.p0666 (cit. on pp. 31, 32).

[239] T. Tamada, Y. Yamakawa, T. Senoo and M. Ishikawa, 'High-speed manipulation of cable connector using a high-speed robot hand,' in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2013, pp. 1598–1604. DOI: 10.1109/ROBIO.2013.6739695 (cit. on pp. 31, 32).

[240] H.-C. Song, Y.-L. Kim, D.-H. Lee and J.-B. Song, 'Electric connector assembly based on vision and impedance control using cable connector-feeding system,' *Journal of Mechanical Science and Technology*, vol. 31, no. 12, pp. 5997–6003, 2017. DOI: 10.1007/s12206-017-1144-7 (cit. on pp. 31, 32).

[241] F. Yumbla, M. Abeyabas, T. Luong, J.-S. Yi and H. Moon, 'Preliminary connector recognition system based on image processing for wire harness assembly tasks,' in *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, 2020, pp. 1146–1150. DOI: 10.23919/ICCAS50221.2020.9268291 (cit. on pp. 31, 32).

[242] P. Kicki, M. Bednarek, P. Lembicz *et al.*, 'Tell me, what do you see?-interpretable classification of wiring harness branches with deep neural networks,' *Sensors*, vol. 21, no. 13, p. 4327, 2021. DOI: 10.3390/s21134327 (cit. on pp. 31, 32).

[243] M. Everingham, L. Van Gool, C. K. Williams, J. Winn and A. Zisserman, 'The pascal visual object classes (voc) challenge,' *International Journal of Computer Vision*, vol. 88, pp. 303–308, 2009. DOI: 10.1007/s11263-009-0275-4 (cit. on pp. 33, 37).

[244] T.-Y. Lin, M. Maire, S. Belongie *et al.*, 'Microsoft coco: Common objects in context,' in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele and T. Tuytelaars, Eds., Cham, Switzerland: Springer, 2014, pp. 740–755. DOI: 10.1007/978-3-319-10602-1_48 (cit. on pp. 33, 37).

[245] M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn and A. Zisserman, 'The pascal visual object classes challenge: A retrospective,' *International Journal of Computer Vision*, vol. 111, pp. 98–136, 2015. DOI: 10.1007/s11263-014-0733-5 (cit. on pp. 33, 37).

[246] O. Russakovsky, J. Deng, H. Su *et al.*, 'Imagenet large scale visual recognition challenge,' *International Journal of Computer Vision*, vol. 115, pp. 211–252, 2015. DOI: 10.1007/s11263-015-0816-y (cit. on pp. 33, 37).

[247] S. Ren, K. He, R. Girshick and J. Sun, 'Faster r-cnn: Towards real-time object detection with region proposal networks,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017. DOI: 10.1109/TPAMI.2016.2577031 (cit. on p. 33).

[248] G. Jocher, *Yolov5 by ultralytics*, version 7.0, 2020. DOI: 10.5281/zenodo.3908559 (cit. on p. 33).

[249] M.-C. Chiu and G. l. E. Okudan, 'Evolution of design for x tools applicable to design stages: A literature review,' in *Proceedings of the ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 6, 2010, pp. 171–182. DOI: 10.1115/DETC2010-29091 (cit. on p. 38).

[250] J. C. Mankins, 'Technology readiness assessments: A retrospective,' *Acta Astronautica*, vol. 65, no. 9, pp. 1216–1223, 2009. DOI: 10.1016/j.actaastro.2009.03.058 (cit. on p. 39).

[251] B. Neyshabur, S. Bhojanapalli, D. Mcallester and N. Srebro, 'Exploring generalization in deep learning,' in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio *et al.*, Eds., vol. 30, Curran Associates, Inc., 2017 (cit. on p. 40).

[252] C. Zhang, S. Bengio, M. Hardt, B. Recht and O. Vinyals, 'Understanding deep learning (still) requires rethinking generalization,' *Communications of the ACM*, vol. 64, no. 3, pp. 107–115, 2021. DOI: 10.1145/3446776 (cit. on p. 40).

[253] K. Kawaguchi, Y. Bengio and L. Kaelbling, 'Generalization in deep learning,' in *Mathematical Aspects of Deep Learning*, P. Grohs and G. Kutyniok, Eds., Cambridge, UK: Cambridge University Press, 2022, ch. 2, pp. 112–148. DOI: 10.1017/9781009025096.003 (cit. on p. 40).

[254] B. L. Berg, *Qualitative research methods for the social sciences*, 7th ed. Boston, MA, USA: Allyn & Bacon, 2009 (cit. on p. 41).

[255] S. N. Hesse-Biber, *The Practice of Qualitative Research*, 3rd ed. Thousand Oaks, CA, USA: SAGE Publications, Inc, 2016 (cit. on p. 41).

[256] K. F. Punch, *Introduction to social research: quantitative and qualitative approaches*, 2nd ed. Thousand Oaks, CA, USA: SAGE Publications, Inc, 2005 (cit. on p. 41).

[257] J. E. Sieber, 'Planning ethically responsible research,' in *The SAGE Handbook of Applied Social Research Methods*, L. Bickman and D. J. Rog, Eds., 2nd ed., Thousand Oaks, CA, USA: SAGE Publications, Inc., 2009, ch. 4, pp. 106–142. DOI: 10.4135/9781483348858 (cit. on p. 41).

[258] R. S. Nickerson, 'Confirmation bias: A ubiquitous phenomenon in many guises,' *Review of General Psychology*, vol. 2, no. 2, pp. 175–220, 1998. DOI: 10.1037/1089-2680.2.2.175 (cit. on p. 41).

[259] W. L. Neuman, *Social research methods: qualitative and quantitative approaches*, 7th ed. Boston, MA, USA: Allyn & Bacon, 2011 (cit. on p. 41).

[260] J. Elkington, 'The triple bottom line,' in *Environmental management: Readings and cases*, M. V. Russo, Ed., 2nd ed., SAGE Publications, Inc., 1997, pp. 49–66 (cit. on p. 41).