2023-09-21

Vision-Based Automated Hole Assembly System with Quality Inspection

Kim, Doowon

UNIVERSITY OF CALGARY

Vision-Based Automated Hole Assembly System with Quality Inspection

by

Doowon Kim

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MECHANICAL ENGINEERING

CALGARY, ALBERTA

SEPTEMBER, 2023

Abstract

Automated manufacturing, driven by rising demands for mass-produced products, calls for efficient systems such as the peg-in-hole assembly. Traditional industrial robots perform these tasks but often fall short in speed during pick-and-place processes. This study presents an innovative mechatronic system for peg-in-hole assembly, integrating a novel peg insertion tool, assembly mechanism and control algorithm. This combination achieves peg insertion with a 200 μm tolerance without the need for pick-and-place, meeting the requirements for high precision and rapidity in modern manufacturing. Dual cameras and computer vision techniques, both traditional and machine learning (ML)-based, are employed to detect workpiece features essential for assembly. Traditional methods focus on image enhancement, edge detection and circular feature recognition, whereas ML verifies workpiece positions. This research also introduces a novel statistical quality inspection, offering an alternative to standard ML inspections. Through rigorous testing on varied workpiece surfaces, the robustness of the methods is affirmed. The assembly system demonstrates a 99.00% success rate, while the quality inspection method attains a 97.02% accuracy across diverse conditions, underscoring the potential of these techniques in automated assembly, defect detection and product quality assurance.

Acknowledgements

I wish to sincerely express my gratitude to those instrumental in bringing this thesis to fruition. First and foremost, my profound thanks go to my supervisor, Dr. Jihyun Lee, who not only granted me this invaluable opportunity but also furnished me with unwavering guidance and support throughout this endeavour. Equally, I am indebted to my co-supervisor, Dr. Simon Park, for his academic counsel and mentorship. Their combined wisdom and support have been the cornerstones of my journey.

I extend my gratitude to the dedicated members of the Intelligent Automation Research Laboratory (fondly known as iAR Lab). I would especially like to highlight the contributions of Ebrahim Mahairi, Zhanhao Wang, Mitchell Weber, Jeff Shin, Ali Khishtan, Omer Sajjad, Dr. Erfan Shojaei Barjuei and Navid Moghtaderi. My collaboration with the team at the Multi-functional Engineering Dynamics Automaton Lab (MEDAL) has been both insightful and rewarding. Special thanks to Dr. Majid TabkhPaz, Mr. Sherif Hassan and GN Corporations Inc. Their invaluable research opportunities and technical expertise have profoundly influenced my engineering journey.

On a personal note, the unwavering support from my parents, brother and close friends has been a foundation throughout this research journey. Sincerely, thank you all.

Finally, a significant portion of this research was made feasible by the support from NSERC. Their backing has been pivotal to the advancement of this project.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Challenge and motivation

Hole assemblies, such as peg-in-hole, are manufacturing processes used in the automotive, oil and gas, aerospace, and construction fields. One prime example is the frac plug used in the oil and gas sector. These plugs, essential for hydraulic fracturing—a key operation for extracting oil from unconventional shale oilfields—feature a cylindrical body riddled with multiple holes intended for peg assembly. With shale oilfields becoming dominant contributors to the global oil supply in the last two decades [1], the demand for frac plugs has surged. Despite this, many factories still rely on manual labor for hole assembly, struggling to keep pace with increasing product demands. This labor-intensive approach, combined with the diversity in workpiece dimensions and hole configurations, has hindered the seamless integration of comprehensive automated assembly systems. However, the pressing need to improve production efficiency and reduce costs has spurred the demand for automated solutions in the manufacturing sector [2, 3]. As a result, the development of novel mechatronics systems that combine advanced automation, computer vision (CV) techniques and machine learning (ML) algorithms has emerged as a solution to address the challenges posed by peg-in-hole assemblies, allowing for efficient and reliable automated assembly processes across diverse industries.

Flexible automation in manufacturing and assembly provides an adequate solution for many manufacturers in meeting their high demand while dealing with different product variations and a changing environment. Sensors play a vital role in flexible automation because they enable

automated systems to detect and read the changes and behave accordingly. Hence, sensor-based technologies such as radiography [4], laser scanners [5] and vision [6] have been used for automated systems. Although radiography and laser scanners provide measurements with high precision and accuracy, their integration with automated systems is challenging due to additional requirements. In Canada, for example, Health Canada has established Safety Code 34, which outlines the radiation protection and safety requirements for industrial X-ray equipment [7]. This code mandates the creation of a controlled environment for radiographic operations, a provision that can sometimes be complex to implement. Therefore, alternative sensors are being explored to ensure manufacturing without requiring modifications to existing manufacturing processes. One common approach is to use CV-based automation, which offers several advantages, such as cost-effective devices and various techniques for different applications. These benefits have accelerated the use of CV in automated systems.

Customized mechatronics systems for similar assembly processes have been proposed in previous studies [8, 9, 10]. Most of these systems employ robotic systems with a gripper to perform the assembly. The detection system is mainly composed of an image acquisition device, connected to an electronic device for image processing and robot control. However, these robotic systems are often time-consuming, especially during the pick-and-place process. For instance, a robot arm must first move to the peg, grasp it, then navigate back to the designated hole for assembly. When faced with multiple holes, this sequence must be executed repeatedly, underscoring the need to enhance the efficiency of this mechanism. Additionally, expenses can escalate due to complex joint control involving multiple sensors to ensure accurate assembly. Therefore, an advanced

mechatronics system with a simple control, low system costs, and high assembly efficiency is necessary to automate the peg-in-hole assembly with high efficiency and repeatability.

In addition to manufacturing and assembly, CV techniques for quality inspection (QI) have been developed and discussed in numerous studies for product inspection [11]. They provide low-cost and high-accuracy solutions compared to human labour. The CV techniques typically involve steps of image data acquisition, image preprocessing, image segmentation and region of interest extraction, and image processing. Lately, the integration of ML methods within CV systems has gained traction due to their broad application potential [12, 13, 14]. However, they necessitate highly trained experts to generate the appropriate model [15]. Dataset preparation for training often requires intensive manual work, such as labelling or modifying each image [16]. The model must be retrained if the system requires additional classifications, with use of additional images that have been prepared with manual work. As such, ML methods might not be universally optimal, especially if iterative upgrades are expected. Systems based on traditional CV techniques can parallel the functionalities of ML methods. However, they often require extensive customization for each application, which can introduce uncertainty during method development. In contrast, traditional CV techniques are computationally efficient, and their results are more traceable for subsequent analysis. Thus, developing adaptable methods with traditional CV techniques can provide an alternative solution for such systems.

1.2. Objectives

The primary goal of this research is to develop a vision-assisted automated peg-in-hole assembly system, complemented by a QI algorithm. This goal is achieved through the following specific objectives:

Objective 1: Design of a novel peg-insertion machine for improved production efficiency

This objective introduces a peg-insertion tool that automates the hole assembly process. Unlike existing models that employ robotic grippers, this tool aims to insert pegs into specified holes with precision. A thorough exploration of the peg-insertion process highlights challenges and prerequisites. Factors such as peg dimensions, hole specifics, insertion pressures and alignment parameters guide the machine's development.

Key components, such as the three-axis machine, rotary roller and syringe pumps, are integrated into the system. The design balances durability and compactness, making it fit for industrial contexts. An advanced automation control system integrates into the design. By merging microcontrollers with visual feedback mechanisms, the machine offers precise movement during operations. Following its development, the machine undergoes tests to assess its functionality, reliability and precision in peg-insertion processes.

Objective 2: Incorporating both traditional and ML approaches in a hole assembly vision algorithm

To further enhance the peg-insertion machine's functionality, CV techniques integrate into the system. Each stage, from image acquisition to tool path generation, is carefully managed. A suitable camera system captures images of assembly components. Once obtained, preprocessing methods improve image quality and support the subsequent feature extraction process. Utilizing the You Only Look Once (YOLOv5) architecture, an object recognition system identifies assembly workpieces. Feature extraction then employs established techniques such as edge detection and the Hough transform. The algorithm's performance undergoes evaluation in various assembly scenarios, using metrics that underscore its reliability and efficiency.

Objective 3: Design of a CV algorithm for QI in hole assemblies

The aim is to develop a CV algorithm that differentiates between positive and negative assembly defects. Using the established image capture system, images are processed to ensure quality for defect detection and classification. Three prediction models are in development for QI. The statistical model incorporates traditional CV techniques and statistical analysis. The CNN model uses convolutional neural networks, and the residual network (ResNet50) model applies deep learning for defect detection. These models are in training and validation phases, with a focus on enhancing their defect detection and classification capabilities.

1.3. Organization of thesis

This thesis is structured into several chapters to provide a comprehensive understanding of the research conducted. 0 offers a thorough literature review encompassing various topics that are relevant to this thesis. The aim is to familiarize the reader with the terminology and concepts associated with each topic, as well as present an overview of the current progress in these areas and their relevance to the research at hand.

In Chapter 3, the setup of the assembly system is described, along with the development of the CV algorithm and the control algorithm specifically designed for this system. The focus is on providing detailed insights into the technical aspects of the system's construction and operation.

Moving on to Chapter 4, the vision-based QI system developed for this project is explained. The chapter explores the specifics of the algorithm employed and outlines the development and implementation of the QI system.

Chapter 5 presents the results of tests conducted with the prototype system and engages in a discussion of these findings.

Chapter 6 serves as a conclusion to the entire project, providing a synthesis of the research outcomes and highlighting key insights. Additionally, it offers recommendations for future work, outlining potential avenues for further exploration and improvement.

Chapter 2. Literature Survey

This chapter provides a comprehensive literature survey of two distinct research areas related to this study. The primary objective is to establish a foundational understanding of key concepts and terminology within each field, while also emphasizing their current advancements. The first section focuses on the domain of assembly automation, specifically centred around peg-in-hole assemblies. It outlines the fundamental components typically found in peg-in-hole assembly systems and discusses devices employed in such systems. In the second section, an extensive background to CV techniques is presented. This includes a classification of CV algorithms and a thorough exploration of the disparities between traditional CV techniques and ML-based approaches. Furthermore, the chapter summarizes the various methods utilized for feature extraction and quality control in CV. Finally, a comprehensive summary is provided, highlighting the pertinence of these topics to the ongoing research.

2.1. Peg-in-hole assembly automation

2.1.1. Robotic arm-based assembly

The development of automated peg-in-hole assembly processes has embraced the utilization of robotic arms due to their high flexibility. Consequently, significant research has been conducted in the field of robotic control to facilitate the peg-in-hole assembly. Numerous techniques, including feedback [9] and compliant control [3], elastic displacement device [9], and impedance control [17], have been extensively studied to adopt robotic systems for the assembly. Notably, Song et al. [18] presented a peg-in-hole assembly system using a robot arm (manipulator) equipped

with a force sensor, as shown in Fig. 2.1. The new assembly strategy was proposed that does not need exact relative pose data between the assembled parts. This system did not consist of sensors like camera that can read the surroundings. Instead, the dragging teaching mode with manual training process of the robot arm is used to drag the peg to the vicinity of the hole. Instead, a "dragging teaching mode" was employed, wherein a manual training process guides the robot arm to bring the peg close to the hole. However, this assembly technique begins with the presumption that assembly proceeds in a singular direction, indicating potential areas for enhancement.

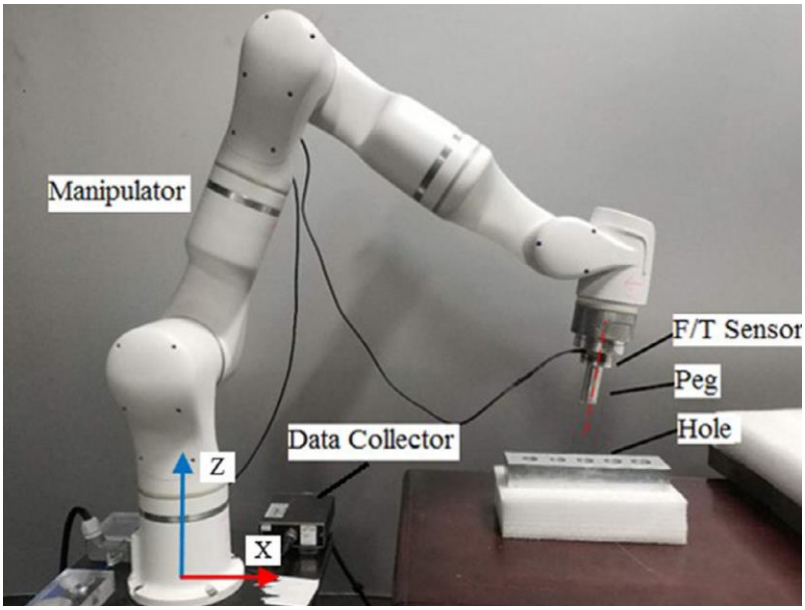

Figure 2.1. Automated peg-in-hole assembly system with compliance control system [18].

The additional sensors like cameras for the peg-in-hole assemblies are adopted to easily approximate the position and orientation of the hole. Utilizing visions systems that uses cameras combined with image processing algorithms offers an efficient method to capture detailed visual information from the assembly area. This visual feedback, when integrated with the control

mechanisms of the robotic arm or manipulator, facilitates enhanced precision and adaptability, especially in environments where the hole's position may be subjected to minute variations or disturbances. Nigro et al. [8] presented an innovative assembly system in which a camera was affixed to the gripper of a robotic manipulator, as illustrated in Fig. 2.2. Their approach utilized the You-Only-Look-Once version 3 (YOLOv3) object detector for hole detection, complemented by a 3D surface reconstruction method. The data acquired from this detection phase steered the robot's approach towards the hole. Nonetheless, certain positional inaccuracies stemming from factors such as reprojection, reconstruction, and merging were identified. Following the initial positioning, the assembly's peg insertion phase was executed, adopting an admittance control that conferred the required compliance to the peg. Yet, the process still manifested an augmented error rate during insertion, largely attributed to the aforementioned positional discrepancies and the absence of any subsequent hole detection after the initial approach. An effective solution to counteract these issues could lie in the conceptualization of a feedback mechanism, continually drawing from camera observations to rectify any positional errors.

Figure 2.2. Automated peg-in-hole assembly system with a robot manipulator and a camera [8].

Other control techniques for the peg-in-hole assembly are developed to compensate for the positional errors. Combination of force control with the vision system is one of the most widely adopted control techniques because it provides a direct measure of the contact situation between the peg and the hole. Deng et al. [19] used force control in addition to the vision control to solve the position error in its assembly. However, attempts to simplify the process by reducing the number of sensors is continually being made by researchers. Consequently, researchers have investigated the use of motion planning algorithms and trajectory optimization techniques to achieve precise and reliable assembly. As a part of this, Park et al. [3] proposed an assembly strategy (Fig. 2.3) that consists of an analysis of the state of contact between the peg and the hole and unit of motions to resolve different states and replace expensive devices such as sensors or remote compliance mechanisms.

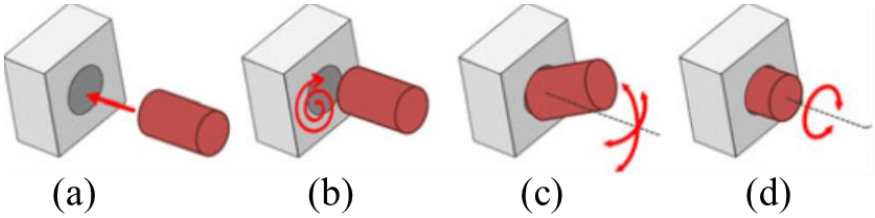

Figure 2.3. Unit motions for peg-in-hole assembly. (a) Pushing. (b) Rubbing. (c) Wiggling. (d) Screwing. [3]

The techniques reviewed in this section revolve around the utilization of robotic arms and grippers to perform the peg-in-hole assembly process. However, this design approach imposes a constraint on the number of pegs that can be assembled simultaneously, limiting it to one or fewer than five pegs. After inserting a peg, the gripper needs to retrieve second peg from storage and return to the hole, which introduces the runtime for the pick-and-place mechanism. It is desirable to minimize this runtime, and therefore, additional research is being conducted to explore optimized systems for efficient peg-in-hole assembly.

2.2. Computer vision

CV has emerged as a prominent field of research, driven by the extensive exploration of vision sensors. Vision sensors provide unparalleled richness of information compared to other sensory parameters. This inherent advantage has propelled the evolution of CV and machine vision technology, finding widespread applications in various automation industries. Furthermore, the recent surge in ML has significantly broadened the capabilities of CV, revolutionizing the field by enabling advanced and accurate analysis of visual data.

2.2.1. *Image acquisition*

The implementation and design of CV technology starts with the selection of an image acquisition system. Charge-coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) are the two most commonly used image sensors in cameras. CCD offers several benefits, including a broad dynamic range, exceptional sensitivity and resolution, minimal distortion, and compact size. Meanwhile, CMOS is known for its cost-effectiveness, low power consumption and high level of integration [20]. In addition, the selection of camera type plays a vital role in the performance of CV. The camera type is decided by the arrangement of the photosensitive units. Line scan cameras consist of photosensitive units distributed linearly in a line, whereas area scan cameras consist of units distributed widely in two dimensions [21].

Two separate image acquisition units computationally merged can create 3D cloud images using stereo vision techniques. This enables users to capture the depth and positional information of the scene. Similarly, light detection and ranging and time-of-flight sensors acquire images that provide depth information about the scene. They measure distance by emitting an artificial light signal such as laser or LED and measuring the time for the reflected light to return to the receiver. By measuring the distance, they can offer a wide array of applications such as shape analysis [22], 3D reconstruction [23] and navigation of outdoor robots [24].

2.2.2. *Image processing*

2.2.2.1. *Traditional computer vision techniques*

The images acquired by the sensors undergo a series of image processing steps to extract the necessary features and information. The initial step in image processing is often image denoising,

which aims to reduce the noise present in raw images. By applying denoising techniques, such as averaging, Gaussian and Kalman filters, the noise can be effectively suppressed, enhancing the quality of the image and facilitating subsequent analysis [25, 26, 27, 28]. Once denoising is performed, region of interest (ROI) identification is commonly carried out to focus on the relevant information while eliminating redundant background details, thereby enhancing system accuracy and computational efficiency. This technique, although simple in its calculation, has proven to be highly effective. Zhang et al. [29] utilized ROI identification to develop an off-axis vision monitoring method using a high-speed camera, demonstrating its applicability in practical scenarios. However, it is important to note that ROI identification relies on user-defined constraints, which may limit its adoption in real-world scenarios.

Feature extraction plays a crucial role in enabling the system to identify distinctive traits in the image for subsequent analysis. Various methods are available for feature extraction, including thresholding, edge detection and segmentation techniques. Thresholding divides the pixels of an image into different categories by setting a threshold value. Over the years, numerous thresholding methods have been proposed, particularly for images with varying grayscale values between the target and background [30]. These methods encompass fixed thresholding, histogram-based techniques and adaptive thresholding [31, 32, 33]. The fixed threshold method is characterized by its speed and suitability for cases where there is a significant difference between the image background and the target. It involves a simple comparison between the threshold value and the pixel values to determine whether to retain or discard them based on the threshold value [31]. Histogram-based methods, such as histogram equalization, adjust the contrast of the image using its histogram, making them among the most widely employed image enhancement techniques [32].

Adaptive thresholding is similar to the fixed thresholding method but dynamically adjusts the threshold value across the image, resulting in a more localized and adaptive approach [33]. The edge detection methods are another commonly employed technique for the feature extraction. The edge detection methods use the discontinuity at the edge pixels in their grayscale values to identify edges through derivation. The first derivative edge detection operators mainly include Sobel and Prewitt, and the second derivative operators mainly include Canny, Laplacian, Laplacian of Gaussian and Difference of Gaussian [34].

The enhanced and edge-identified images are further processed to extract various features, including texture, transform domain and shape. Texture features capture the visual patterns and variations in the image, allowing for texture-based analysis and classification. Common texture feature extraction methods include local binary patterns (LBP) [35], gray-level co-occurrence matrices (GLCM) [36] and Gabor filters [37]. Transform domain features involve transforming the image into a different domain, such as the frequency domain using Fourier transform or wavelet transform. These transformations reveal the distribution of frequencies or scale-related information in the image, enabling the extraction of frequency-based features or multi-resolution analysis [38]. Shape features describe the geometric characteristics and contours of objects within the image. These features can be extracted using techniques such as contour tracing [39], Hough transform [40] or mathematical shape descriptors such as Hu moments or Zernike moments [41, 42].

Although traditional CV techniques have proven to be reliable and widely adopted in various applications, they do have drawbacks when compared to ML-based CV techniques. One of the main limitations of traditional methods is their reliance on handcrafted features and heuristics, which can be labour-intensive and require domain-specific knowledge for fine-tuning [43]. This

manual feature engineering process may not always capture all possible intricate patterns and complexities present in the image, potentially leading to reduced performance and accuracy in challenging scenarios.

Furthermore, traditional CV techniques may not perform as designed when confronted with variations or when adapting to new and unseen data. These techniques often lack the flexibility to generalize well to diverse and dynamic environments because they heavily depend on predefined rules and assumptions [43]. Interpreting and explaining the decisions made by traditional CV algorithms can also be challenging. As the number of traditional CV techniques employed within an algorithm increases, the reasoning behind the results may become less transparent, posing difficulties in understanding the underlying processes and potentially limiting their applications in safety-critical or highly regulated domains [43].

However, it is important to recognize the continued significance of traditional CV techniques alongside the emergence of ML-based algorithms. Traditional techniques serve as the foundation for various algorithms and possess distinct advantages, including lower computational requirements and easily interpretable behaviours. These advantages make traditional techniques a subject of ongoing research and practical implementation [43]. Despite the advancements in ML, traditional CV methods continue to contribute to the field, providing optimized solutions and insights that complement the capabilities of ML algorithms.

2.2.2.2. Machine learning-based computer vision techniques

Over the past decade, ML-based algorithms have gained significant popularity and proven their effectiveness in various CV tasks, including region detection, feature extraction and image classification. These algorithms have demonstrated exceptional performance in identifying and

categorizing specific objects or entities within images, enabling applications such as person detection and vehicle recognition. By leveraging the power of ML, these algorithms can learn complex patterns and features from large datasets, leading to improved accuracy and performance [43].

Fig. 2.4 illustrates the evolutionary journey of ML-based CV techniques, particularly focusing on object detectors. Over time, these detectors have undergone significant advancements and improvements, driven by research and technological developments in the field of ML. The early stages of ML object detectors are characterized by simpler architectures and limited capabilities. However, with the introduction of machine learning and CNNs, the performance of object detectors has significantly improved, enabling them to achieve higher accuracy and handle more complex tasks.

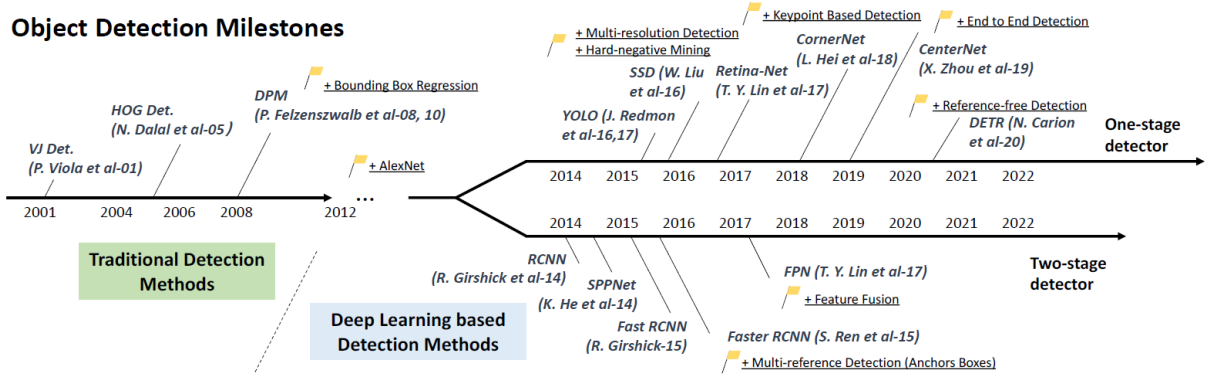

Figure 2.4. A road map of object detection [44].

As ML algorithms continued to evolve, various architectures and techniques emerged, such as Viola Jones (VJ) detectors, faster region-based convolutional neural network (RCNN), YOLO and single shot multibox detector. These architectures are designed to address specific challenges, such as faster inference speed, better real-time performance and improved accuracy. For earlier

detectors, such as the VJ Detector, the algorithms are built based on the handcrafted features to compensate for the lack of effective image representations. VJ Detector is the first real-time human face detector without any constraints. It used the sliding window technique to go through all possible locations and scales in an image to localize the human face. It used three important techniques, namely integral image, feature selection and detection cascades, to outperform other detectors at the time [45].

Following the earlier stages, a significant advancement in ML object detection occurred with the groundbreaking work of Krizhevsky et al. [46], who presented AlexNet, a deep learning model based on the CNN architecture. Prior to AlexNet, while CNN architectures had existed, none had demonstrated the same level of efficacy on large-scale datasets. AlexNet's design was unique, consisting of multiple convolutional layers, followed by max-pooling, fully connected layers, and a final softmax layer, making it substantially deeper than previous models. It also introduced the efficient usage of the Rectified Linear Unit (ReLU) as its activation function, which accelerated training times by mitigating the vanishing gradient problem. A significant aspect of AlexNet's success was its performance in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), where it dramatically reduced the error rates compared to prior methods. This wasn't just a victory, but a demonstrative showcase of deep learning's capability, particularly with CNN architectures, in handling image classification at scale. This remarkable success of AlexNet in image classification tasks inspired researchers to explore the potential of CNNs in object detection.

The success of AlexNet led to the development of CNN-based two-stage detectors that use a multi-stage approach to object detection. They involve a separate region proposal step to identify potential object regions in the image, followed by a classification and regression step to refine the

proposed regions and classify the objects [47]. This approach allowed for more accurate and robust object detection, especially in complex scenarios with multiple objects and overlapping instances.

RCNN, developed by Girshick et al. [48] is one of the first two-stage detectors to leverage CNNs for object detection. They proposed an innovative approach that involved generating region proposals using selective search and then performing CNN-based classification and bounding box regression on these regions. Despite its significant progress in accuracy, RCNN had some limitations, particularly in terms of speed and computational complexity, due to the need to process a large number of region proposals.

In response to these challenges, Ren et al. [49] introduced Faster RCNN as an improvement over RCNN. It incorporated a Region proposal network that learned to generate region proposals directly from the convolutional feature maps, eliminating the need for external region proposal methods. This innovation led to faster and more efficient object detection, with improved accuracy and reduced computational burden.

Another key advancement is the introduction of Feature pyramid networks (FPN), presented by Lin et al [50]. FPN aimed to address the issue of scale variation in object detection. FPN proposed a top-down architecture that allowed feature maps of different scales to be merged and used for object detection. By leveraging multi-scale features, FPN significantly enhanced the ability of detectors to detect objects of varying sizes, leading to more robust and accurate performance.

One-stage detectors, commonly referred to as single-shot detectors, perform object detection in a single forward pass through the neural network. Unlike their two-stage counterparts, which adopt a multi-stage approach, one-stage detectors directly predict location and size (bounding

boxes), and class scores for potential objects in the input image without requiring a separate region proposal step [47].

The efficiency and simplicity of one-stage detectors stem from their direct prediction process, wherein they simultaneously determine object locations and associated class labels. This characteristic enables them to achieve real-time performance, making them ideal for applications where speed is critical. Additionally, one-stage detectors are well-suited for scenarios involving numerous objects in cluttered environments because they can efficiently detect multiple instances in a single pass.

YOLO is one of the major one-stage detectors presented by Redmon et al [51]. The YOLO architecture is designed to achieve real-time performance, making it highly suitable for time-critical applications, such as video surveillance and autonomous vehicles. By leveraging a unified framework, YOLO can handle multiple objects simultaneously, effectively detecting and classifying objects even in cluttered environments. Its single-shot nature allows YOLO to strike a balance between speed and accuracy, making it an attractive choice for various CV tasks including manufacturing.

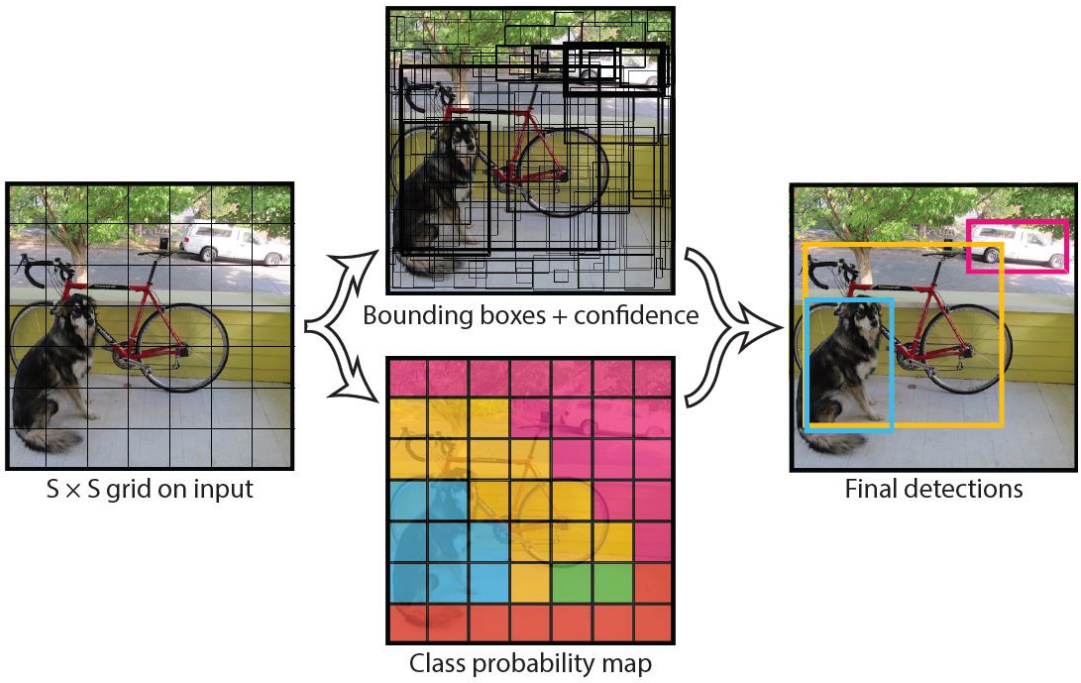

Figure 2.5. YOLO prediction model [51].

The continual advancement of YOLO has led to the publication of multiple versions, including v2, v3, v4 and v5. The successive iterations of YOLO have addressed various challenges and further optimized its detection accuracy and processing speed. As a result, YOLO has garnered widespread attention and adoption in the CV community, driving innovations in real-time object recognition and advancing the capabilities of CV systems in diverse practical scenarios.

As the technology continues to advance, it is expected that ML-based CV techniques will play an increasingly critical role in ensuring consistent and high-quality manufacturing outputs across various industries. Additionally, the integration of transfer learning and fine-tuning methodologies allowed ML CV techniques to be more adaptable and efficient because they could leverage pretrained models to handle new and diverse datasets with reduced training time. Furthermore, the continuous expansion of annotated datasets and the availability of powerful hardware accelerated

the progress of ML object detectors, enabling them to excel in numerous applications. With the advent of edge computing and embedded systems, ML-based CV techniques have become increasingly deployable on resource-constrained devices, making them suitable for on-site applications. As research and development in the field of ML continue, it is anticipated that ML object detectors will undergo further advancements, unlocking new possibilities and applications in manufacturing and beyond.

Although ML-based CV techniques have shown remarkable progress and achieved state-of-the-art performance in various applications, they are not without their downsides compared to traditional CV techniques. One of the main disadvantages of ML-based approaches is their high computational requirements. Training and deploying complex ML models demands substantial computational resources, making them less suitable for resource-constrained environments [43]. In contrast, traditional CV techniques often rely on handcrafted algorithms and heuristics, which generally have lower computational overheads, enabling faster and more efficient processing [43].

Another drawback of ML-based techniques is the need for extensive labeled datasets for training. Building large and diverse datasets can be time-consuming and labour-intensive, hindering the adoption of ML models in domains with limited data availability [43]. Traditional CV techniques, conversely, can be more easily adapted and customized for specific tasks without the need for extensive data-driven training. Additionally, the interpretability of ML models can be a concern, especially in safety-critical applications. Though traditional CV techniques often produce interpretable results, ML models, especially deep neural networks, can be regarded as "black boxes," making it challenging to understand the reasoning behind their decisions. This lack

of transparency can raise questions about the reliability and trustworthiness of ML-based systems [43].

Despite these challenges, both ML and traditional CV techniques have their unique strengths, and the choice between them depends on the specific requirements and constraints of the application. Striking a balance between the advantages and limitations of each approach is crucial in developing effective and practical CV solutions.

In practice, there is often a synergy between traditional CV techniques and ML algorithms. By combining these approaches, researchers and practitioners can leverage the strengths of each method. For instance, a hybrid approach may involve using traditional techniques for initial feature extraction and preprocessing, followed by ML algorithms for further analysis and classification. This hybridization can lead to more comprehensive and accurate analysis of visual data while providing an optimized computational efficiency [43]. For instance, image preprocessing steps with traditional CV algorithms are often performed on an input image for a object detector that performs classification based on an ML algorithm to achieve an upgraded detection performance while reducing the computational time [52].

2.2.3. Computer vision in manufacturing and assembly

CV technology has made significant strides in the field of manufacturing and assembly, playing a pivotal role in component and feature detection, and aiding in the planning of tool paths. In manufacturing processes, accurate detection of components is crucial for ensuring proper assembly and maintaining production efficiency. Traditional approaches rely on manual identification or physical templates, which can be time-consuming and prone to errors. CV systems offer a more efficient and reliable solution by automatically detecting and recognizing components

within an image or video stream. By leveraging advanced algorithms, these systems analyze visual data to identify specific components based on their shape, colour, texture or other distinguishing features. This automated component detection not only saves time but also improves accuracy, reducing the risk of assembly errors and enhancing overall productivity. Once components are detected, planning tool paths around them can be used for optimizing the assembly process. CV can assist in this aspect by providing information about the positions and orientations of components. By utilizing cameras and sensors, the system can accurately determine the location of each component or important feature of a component within the manufacturing environment. This information can then be used to plan efficient tool paths that minimize the distance travelled by robotic arms or other automated systems, reducing assembly time and improving production throughput. Additionally, CV can help identify obstacles or potential collisions, enabling the system to plan safe and collision-free tool paths. Tsai et al. [53] demonstrated vision technology that recognizes the weld seam on gold-club heads and automatically generates a welding path for the robot (Fig. 2.6). For some of the early works, Solvang et al. [54] presented a robot path generation strategy that uses the vision system to track a line that has been drawn by an operator with a marker pen.

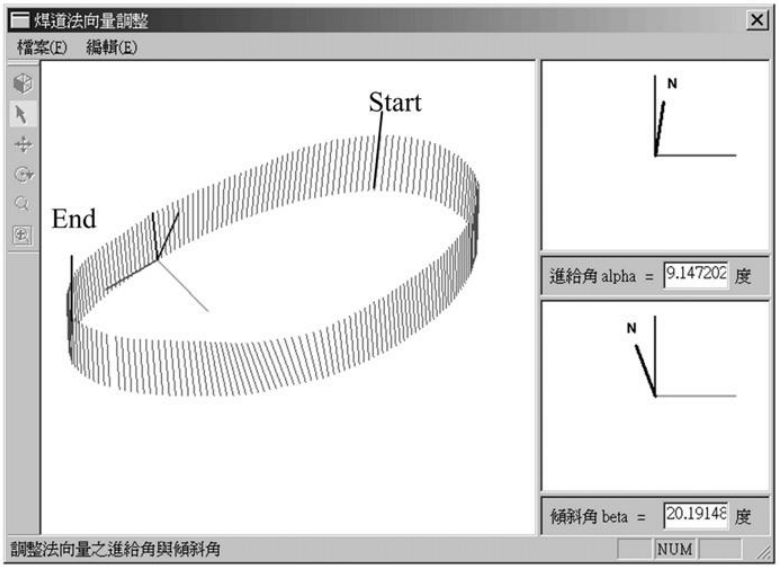

Figure 2.6. Generated welding path of a golf club from the system using computer vision [53].

Various traditional CV techniques are employed in manufacturing and assembly applications. These include image segmentation algorithms for isolating components from the background, feature extraction algorithms for capturing relevant geometric or textural information and pattern recognition algorithms for component identification. However, limits of the traditional CV techniques are clear as mentioned earlier in section 2.2.2, and they generally need human intervention to make the final classification. Hence, they are mostly used as a supplementary tool that helps operators working at the site. In addition, some techniques require references to specific features or tools for calibration or boundary definition, adding complexity in adopting the techniques. De Araujo et al. [55] presented an image processing algorithm based on only traditional CV techniques to find the workpiece position on three-axis machine to aid machining workers in their operations. A physical reference point called part zero is used to zero the object's position at the beginning of every machining operation.

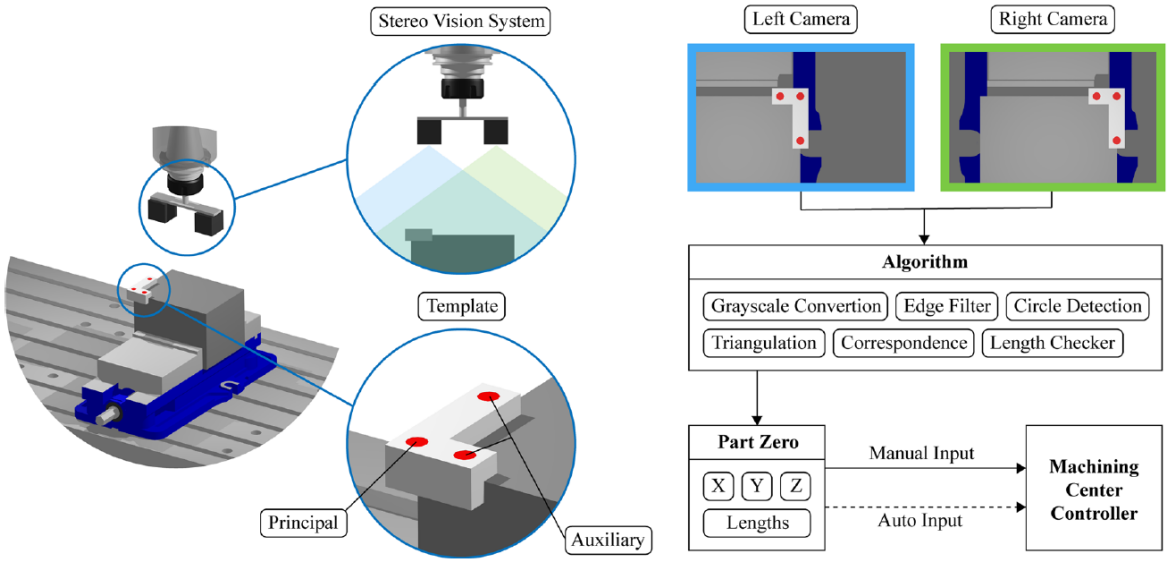

Figure 2.7. Workpiece referencing computer vision system [55].

The utilization of ML object detectors in CV systems enables them to recognize specific components, eliminating the need for calibration or explicit boundary definition, which are major constraints of traditional CV techniques. ML object detectors also possess the capability to adapt to variations in lighting conditions or component appearance. By training these algorithms on diverse datasets that encompass a wide range of lighting scenarios, camera perspectives and component variations, CV systems can acquire the ability to generalize and make accurate predictions in real-world manufacturing environments. The adaptability of ML algorithms is a significant advantage because their performance can continuously improve through continuous learning and exposure to new data. Consequently, ML object detectors provide the necessary flexibility and adaptability to effectively address the challenges posed by dynamic environments and varying conditions in manufacturing fields.

Derived from these benefits, ML has proven to be highly successful in various process optimization, monitoring and control applications [56]. For instance, Jin et al. [57] developed a real-time monitoring and autonomous correction system using an ML model to modify 3D-printing parameters iteratively and adaptively in additive manufacturing processes. This system can detect in-plane printing conditions and in-situ correct defects faster than a human can.

The integration of CV into manufacturing and assembly processes brings several benefits. It enhances productivity by automating component detection and reducing manual intervention. By providing accurate component information, CV facilitates precise tool path planning, minimizing assembly time and maximizing efficiency. Moreover, the ability of CV systems to adapt to variations in component appearance or environmental conditions ensures robust performance and flexibility in dynamic manufacturing environments.

2.2.4. Vision-based quality inspection

The techniques utilized in vision-based QI in manufacturing are similar to those employed in CV for manufacturing and assembly. However, the primary distinction lies in the final output with QI applications providing classification predictions based on input images, whereas manufacturing and assembly applications focus on feature extraction for localization and positioning purposes.

QI systems that exclusively rely on traditional CV techniques have demonstrated their effectiveness and practicality in various manufacturing applications. These methods offer several advantages, including lower computational requirements. With manual feature engineering, these systems can be tailored to specific manufacturing processes and product characteristics, allowing for precise defect detection and classification. Additionally, traditional CV techniques have a proven track record of reliability and accuracy in detecting common defects and flaws in

manufacturing processes. By leveraging these well-established methods, manufacturers can benefit from cost-effective and efficient QI solutions that have stood the test of time.

Furthermore, traditional CV techniques are often more interpretable, providing valuable insights into the detected defects and contributing to a better understanding of the production process. As a result, these systems continue to play a crucial role in ensuring product quality, enhancing manufacturing efficiency and maintaining customer satisfaction. For instance, Hocenski et al. [58] developed an automated QI system for ceramic tiles using only traditional vision techniques, including Canny edge detection and histogram subtraction method. Their research demonstrated the successful application of traditional CV techniques in detecting defects and ensuring the quality of ceramic tiles (Fig. 2.8), showcasing the practicality and reliability of these methods in real-world manufacturing scenarios.

Figure 2.8. Automated QI for ceramic tiles using only traditional vision techniques for detecting cracks [58].

ML-based CV techniques have gradually started being employed for QI purposes in manufacturing. These detectors leverage the power of ML algorithms, such as CNNs, to identify defects, anomalies and deviations accurately and efficiently in manufactured products. By training

on large datasets of both defect-free and defective products, these detectors can learn the distinguishing features and patterns associated with different types of defects. As a result, they can quickly and reliably identify and classify defects, ensuring that only products of the highest quality reach the market. Moreover, ML-based object detectors offer the advantage of adaptability, allowing them to handle variations in product appearance and environmental conditions, making them suitable for dynamic manufacturing environments.

For instance, Ren et al. [59] applied a CNN-based algorithm to perform defect classification for surface inspection tasks, encompassing diverse surfaces such as Northeastern University surface defect database, weld, wood and micro-structure surfaces. This algorithm effectively built a classifier on image patches by utilizing features transferred from a pretrained network. Subsequently, pixel-wise prediction is conducted by convolving the trained classifier over the input image, enabling image classification and defect segmentation. The algorithm's performance surpassed its benchmarks, exhibiting an accuracy improvement ranging from 0.66–25.50%, depending on the specific task and defect type.

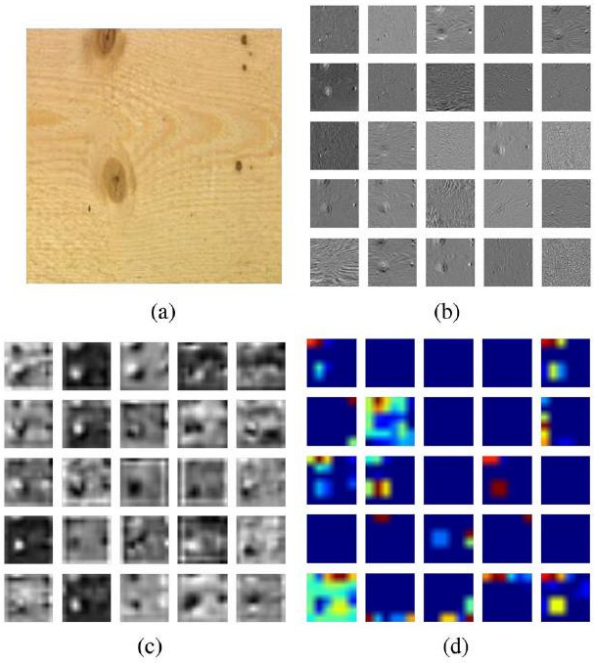

Figure 2.9. Visualization of features in different layers of the CNN algorithm presented by Ren et al. (A) Original image. (B) Feature map from first convolutional layer. (C) Feature map from third convolutional layer. (D) Feature map from fifth convolutional layer [59].

Similarly, in the domain of hot-rolled strip steel inspection, Feng et al. [60] developed an ML-based surface defect classifier employing ResNet50 architecture augmented with FcaNet and Convolutional block attention module. This well-designed solution achieved an impressive classification accuracy of 94.11%. By integrating advanced network architectures and attention mechanisms, the model demonstrated its ability to effectively capture intricate details and discriminative features essential for accurate defect classification.

These research endeavours exemplify the capabilities of ML-based CV techniques in addressing complex QI tasks. The integration of ML-based CV techniques into manufacturing processes not only improves product quality but also streamlines inspection procedures, reduces manual labour and enhances overall production efficiency.

2.3. Summary

The literature survey in this chapter provides a comprehensive review of two key research areas related to this study: assembly automation and CV in manufacturing and assembly. In the domain of assembly automation, the focus is on peg-in-hole assemblies, a critical process in various industries such as robotics, electronics and automotive. The chapter concerns the significance of fast and precise assembly and explores various control techniques. It also addresses the constraint of the current design approach, which limits the speed of the assembly, and highlights the ongoing research to optimize systems for efficient peg-in-hole assembly.

The second part of the chapter explores CV techniques, categorizing them into ML-based and traditional approaches. The development of object detectors, exemplified by versions of YOLO, underlines the potential and growth of ML-based CV techniques. This suggests a possible increasing reliance on such methods for consistent manufacturing outcomes in the future.

Traditional CV techniques remain valuable due to their low computational requirements and the interpretability of their results. These methods have consistently supported various manufacturing tasks such as component detection and tool path planning. Their straightforward nature and ability to be adjusted for specific manufacturing scenarios make them an integral part of the industry’s tool kit.

QI methods, though rooted in traditional CV techniques, have started to explore ML-based approaches. Although the traditional methods continue to offer reliable solutions, ML techniques introduce a different dimension to defect detection and classification, particularly in terms of adaptability and handling varied conditions.

In conclusion, the literature presents a scenario where traditional CV and modern ML techniques coexist. This combination is likely to provide the manufacturing sector with diverse tools to address varying needs. As research progresses, further insights and improvements in both areas can be expected.

Chapter 3. Peg-in-Hole Assembly System

This chapter focuses on the design and implementation of the prototype for the proposed assembly system, encompassing both hardware and software components. The main objective of the system is to assemble the workpiece with the pegs, as illustrated in Fig. 3.1. The workpiece itself possesses a cylindrical geometry with a diameter of 93.0 mm. It incorporates 20 surface holes, each with a diameter of 9.7 mm and a depth of 7.1 mm. The assembly process involves the utilization of pegs, available in black and white colours (represented as yellow in Fig. 3.1 simply to improve the visibility). These pegs are standardized, measuring 9.5 mm in diameter and 6.9 mm in height. Both the workpiece and peg dimensions come with a tolerance of ±0.1 mm.

A key aspect of this assembly process is ensuring a 200 µm clearance between the pegs and holes. The system is therefore tailored to uphold this clearance. An additional step involves applying glue to firmly anchor the pegs. Lastly, the assembly process does not require any specific patterns or sequences for the utilization of different peg types.

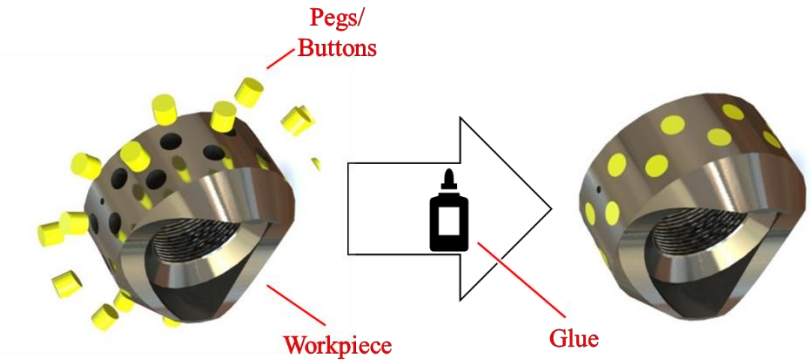

Figure 3.1. Peg-in-hole assembly task.

This assembly (Fig. 3.1) is the frac plug, a crucial component in the oil and gas industry. Surprisingly, there have been no automated systems introduced for its assembly thus far. This has led to the reliance on manual labor for what is, essentially, a basic and repetitive task. Such a reliance can be considered a suboptimal use of skilled labor, especially when these professionals could be more effectively utilized in performing complex and intricate operations. Recognizing this gap, the thesis aims to develop a system to automate the assembly process, significantly reducing the need for human intervention.

The hardware design integrates the base machine with the peg insertion tool, and the software design emphasizes image processing techniques for automating the assembly process. As depicted in Fig. 3.2, the assembly system prototype's primary components are the central computer, microcontroller, two cameras, and the assembly machine. Throughout the chapter, the design process will be explored and discussed. These iterations showcase the evolution of the system design, highlighting the improvements and enhancements made to enhance its functionality and performance.

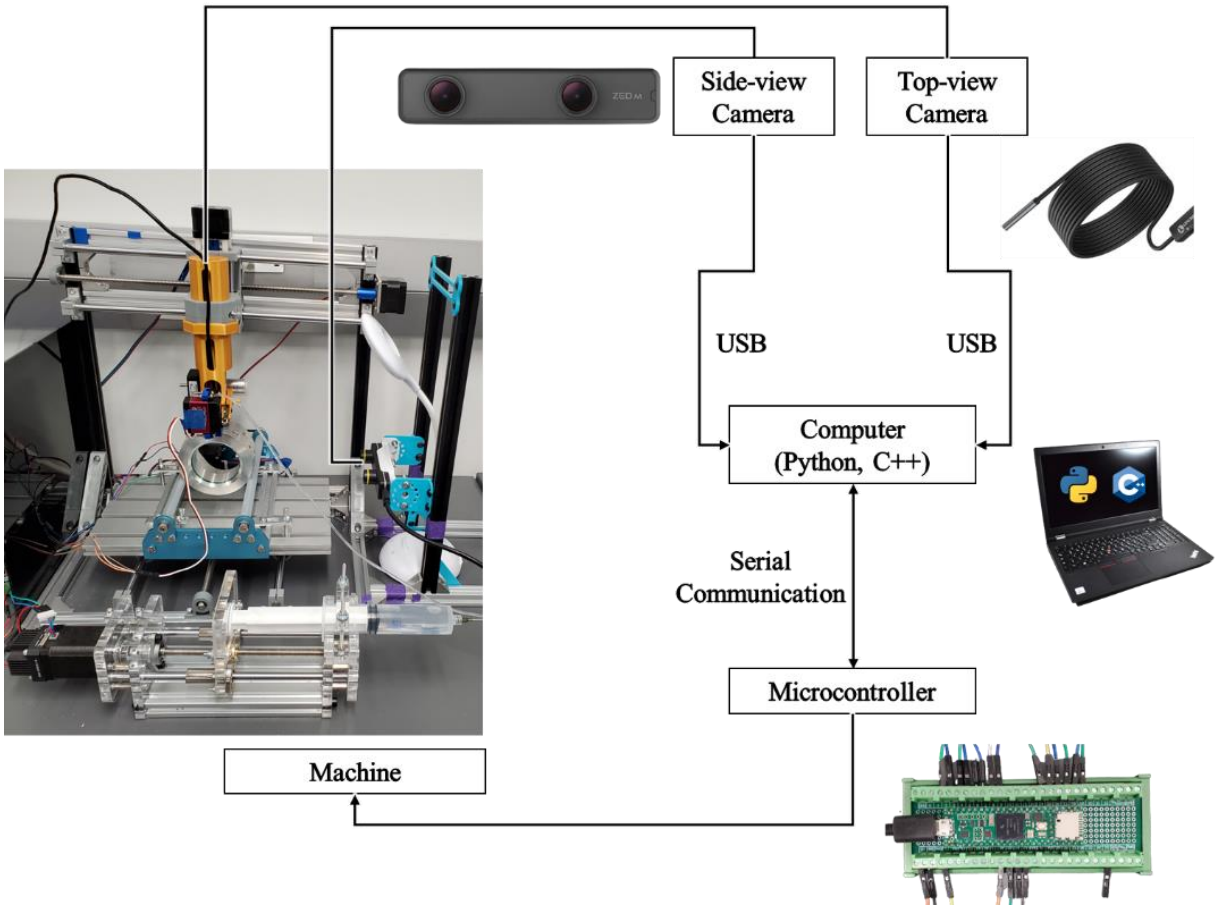

Figure 3.2. Configuration of the assembly system.

3.1. Hardware Design

This section presents the detailed hardware design of the proposed automated peg-in-hole assembly system prototype. The hardware components are carefully selected and integrated to ensure smooth and accurate operation of the system. The key elements of the hardware design include the base machine and the peg insertion tool.

The key objective of the design of the hardware is to achieve full control over the workpiece. For this, the three-axis machine (CNCTOPBAOS 3018 Pro CNC Router Kit) is selected as the

base machine, and the rotary roller (Genmitsu Laser Rotary Roller Engraving Module) is installed at its worktable to enable four-axis control as demonstrated in Fig. 3.3.

The hardware selection process for the system design prioritized the use of commercially available components to ensure the simplicity and accessibility of the overall solution. The system incorporates four NEMA 17 stepper motors (1.3A), with each motor dedicated to controlling a specific axis of motion.

To drive these stepper motors, the system utilizes TB6600 stepper motor drivers. The TB6600 drivers are H-bridge bipolar constant phase flow drivers that provide micro-stepping capability, allowing for smoother and more precise motor control. In this system, the motors are operated at a current of 1.0A. The motors accountable for y-axis and r-axis movements (Fig. 3.3), which are crucial for workpiece alignment, are set to run on a 32 micro-step setting. The remaining motors are configured to operate on the full step setting. By carefully selecting and configuring the components, the system can achieve the desired level of control and accuracy in its motion operations.

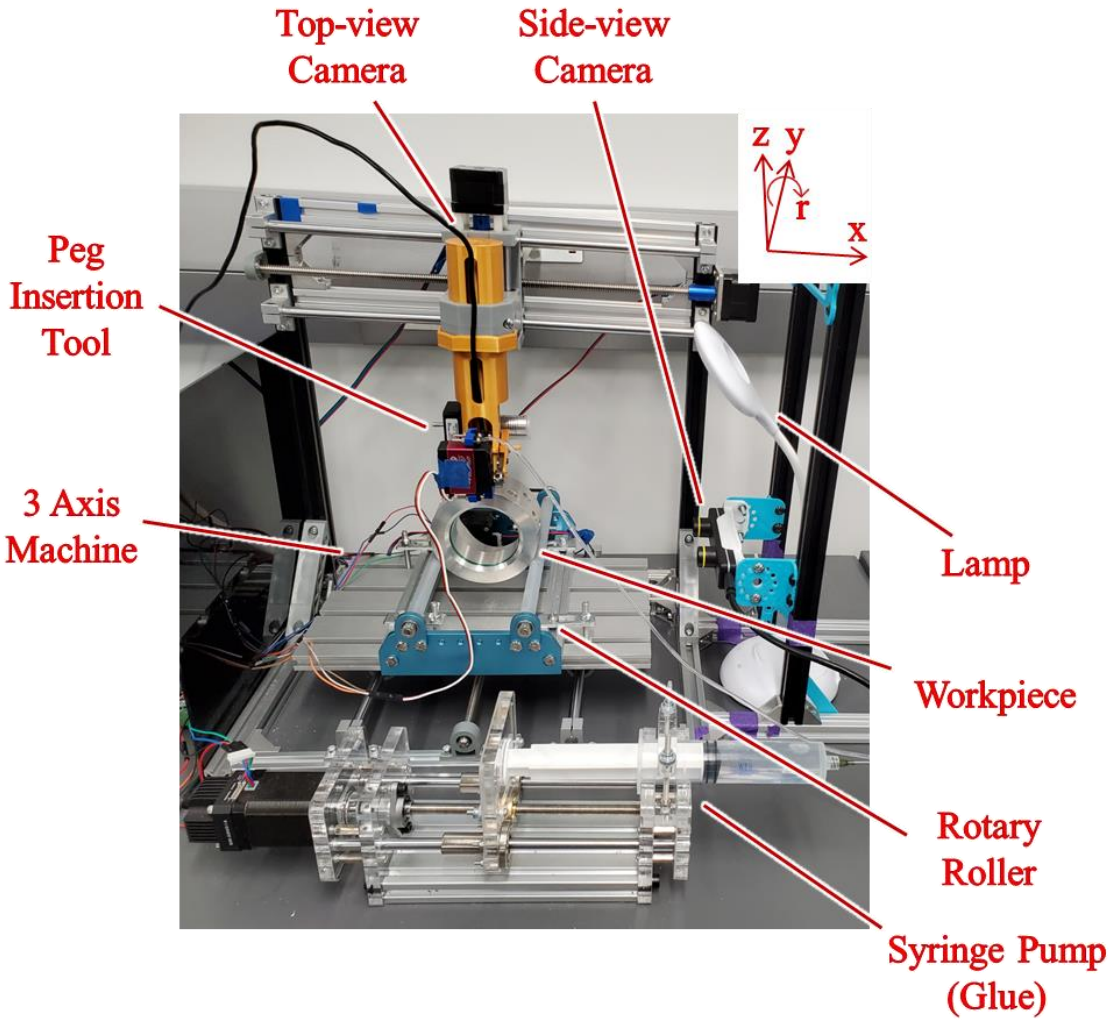

Figure 3.3. Hardware configuration of the assembly machine prototype.

To ensure consistency and repeatability during the assembly process, the peg insertion tool is securely attached to the tool mount, while the camera mount remains fixed, maintaining a constant relative position between the camera and the tool. Detailed information regarding the peg insertion tool can be found in Section 3.1.1.

To achieve glue metering, the system incorporates a syringe pump (DIY Syringe Pump, RobotDigg) that is controlled by a stepper motor. This syringe pump is driven by the TB6600

driver, utilizing a 32 micro-step setting to ensure controlled metering of the glue application. The syringe is mounted onto the syringe pump, and a needle adapter is connected to PVC tubing. The other end of the tubing is attached to a blunt tip needle. Connections between the tubing is established using hose barbs and Luer-lock fitting adapters. The selection of the needle size (25Ga), tube material (PVC transparent hose vinyl tubing) and magnitude of the syringe pump actuation is determined through iterative testing to ensure controllability and sustainability. The needle tip, responsible for dispensing the glue, is positioned within the peg insertion tool where the pegs are launched.

3.1.1. Peg insertion tool

The peg insertion tool, shown in Fig. 3.4, performs two main roles: storing and launching pegs. It features a long loading path designed to hold multiple pegs in a sequential stack. This path intersects the pusher path near its base, where the launching mechanism operates. Powered by a stepper motor (NEMA 14, 0.4A), the pusher employs a linear back-and-forth movement. When actuated in one direction, it releases one peg from the base of the stack. In the reverse motion, it prepares the next peg for dispatch. This pusher and its corresponding path have undergone iterative design and tests, ensuring its mechanical integrity and consistent performance. Once the peg reaches the end of the pusher path, termed the peg launch zone, it descends into its designated hole, completing the peg launch. Consequently, the need for a pick-and-place mechanism is eliminated, thereby enhancing efficiency.

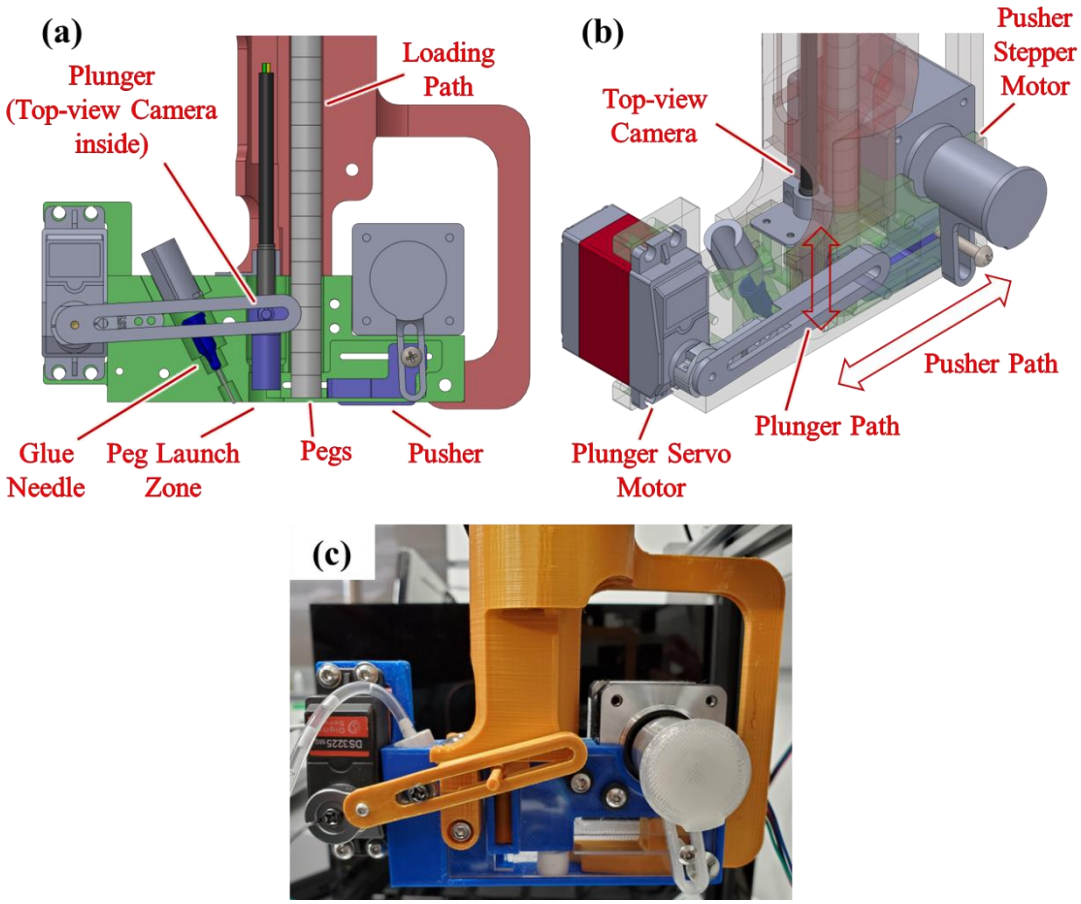

Figure 3.4. Peg insertion tool. (a) Cross-sectional view. (b) Isometric view. (c) 3D printed prototype.

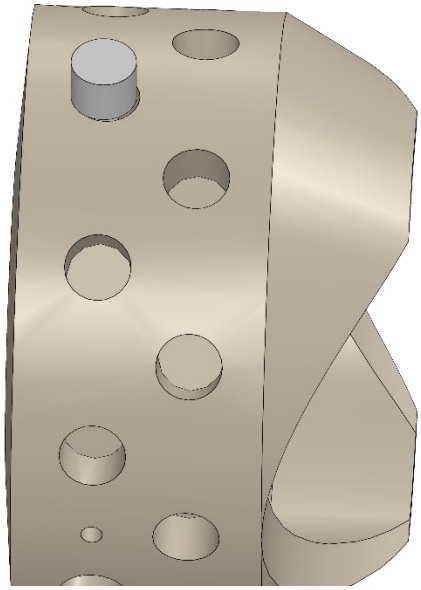

Figure 3.5. Assembly state after peg launch.

Fig. 3.6a illustrates the strategic design and positioning of the glue needle and top-view camera within the peg insertion tool. The glue needle, linked to a syringe pump, is situated at a distance of 7.41 mm from the peg launch zone. This configuration ensures that the machine can alternate between applying glue and executing peg assembly within the glue's curing time of 10 seconds. Furthermore, the top-view camera is housed within the plunger, enabling direct visualization of the target hole during the assembly process. This placement allows the plunger to exert downward pressure on the launched peg without necessitating any machine repositioning. For this purpose, the plunger has been designed to be hollow, as depicted in Fig. 3.6b.

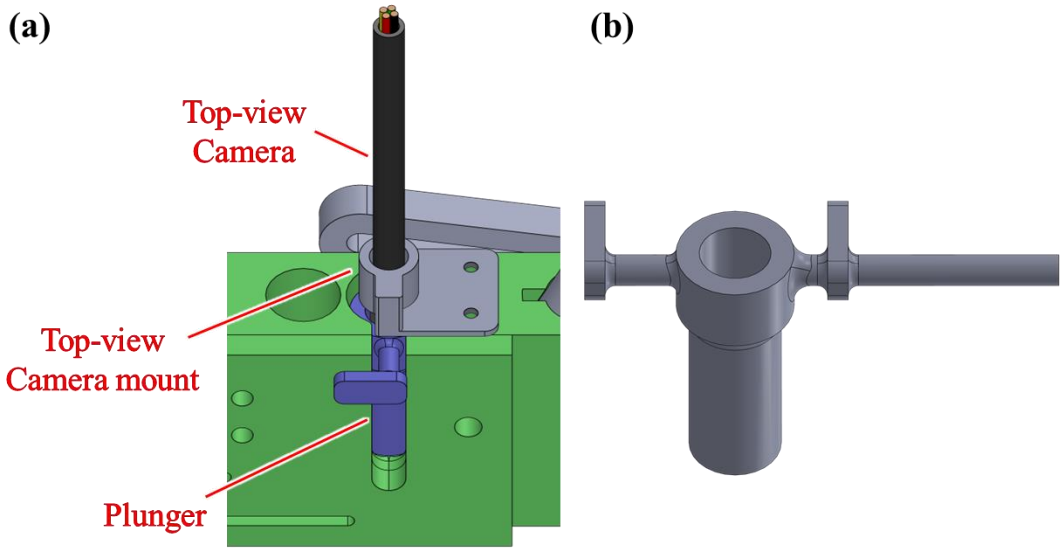

Figure 3.6. Components of the peg insertion tool. (a) Combined assembly of the plunger with the top-view camera, highlighting their positioning on the peg insertion tool. (b) Detailed view of the plunger.

The illustrations in Fig. 3.7a and b showcase the pusher's rear and isometric views, respectively, highlighting the presence of the pusher guide. Designed to fit within the pusher guide path, which is essentially a groove depicted in Fig. 3.7c, this guide ensures the pusher maintains linear oscillation without veering off its dedicated course. Furthermore, the peg contact area on the pusher features a contoured design, facilitating an enhanced grip when interfacing with the peg. This design maximizes the contact surface area therefore prevents the peg from rebounding upon initial contact. Consequently, the peg remains in consistent touch with the pusher throughout its launch trajectory. Such a design prohibits any unwanted rotational movement of the peg within the tool and ensures finer control of the peg drop position from the peg launch zone.

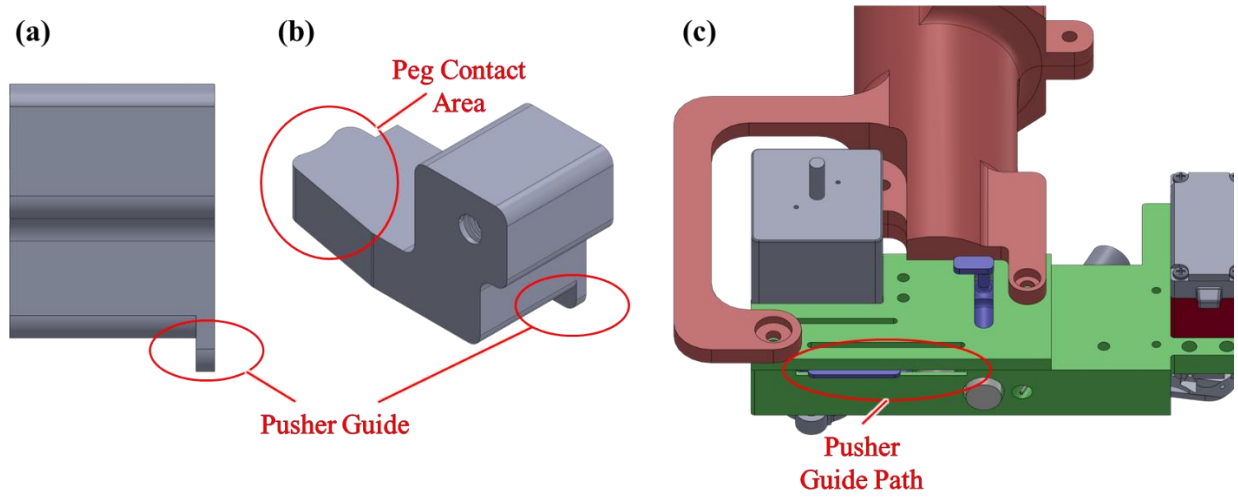

Figure 3.7. Components of the peg insertion tool. (a) Rear view of the pusher. (b) Isometric view of the pusher. (c) Combined assembly of the pusher, highlighting its positioning on the peg insertion tool.

3.1.2. Reloader

The reloader, an additional tool designed to be used with the peg insertion tool, serves the purpose of reloading pegs into the peg insertion mechanism. This tool comprises two stepper motors (NEMA 17, 1.5A), each responsible for actuating a sorter that rotates to align the orientation of the pegs correctly. Additionally, the reloader features a vibration motor that facilitates the movement of pegs down the funnel, ensuring a smooth and continuous feeding process into the peg insertion tool. To minimize any undesirable effects of vibration on the peg-in-hole assembly, a spring mount is incorporated between the reloader mount and the funnel, isolating the vibration and maintaining the stability of the overall system. The reloader enhances the efficiency of the peg reloading process, streamlining the assembly workflow and minimizing interruptions.

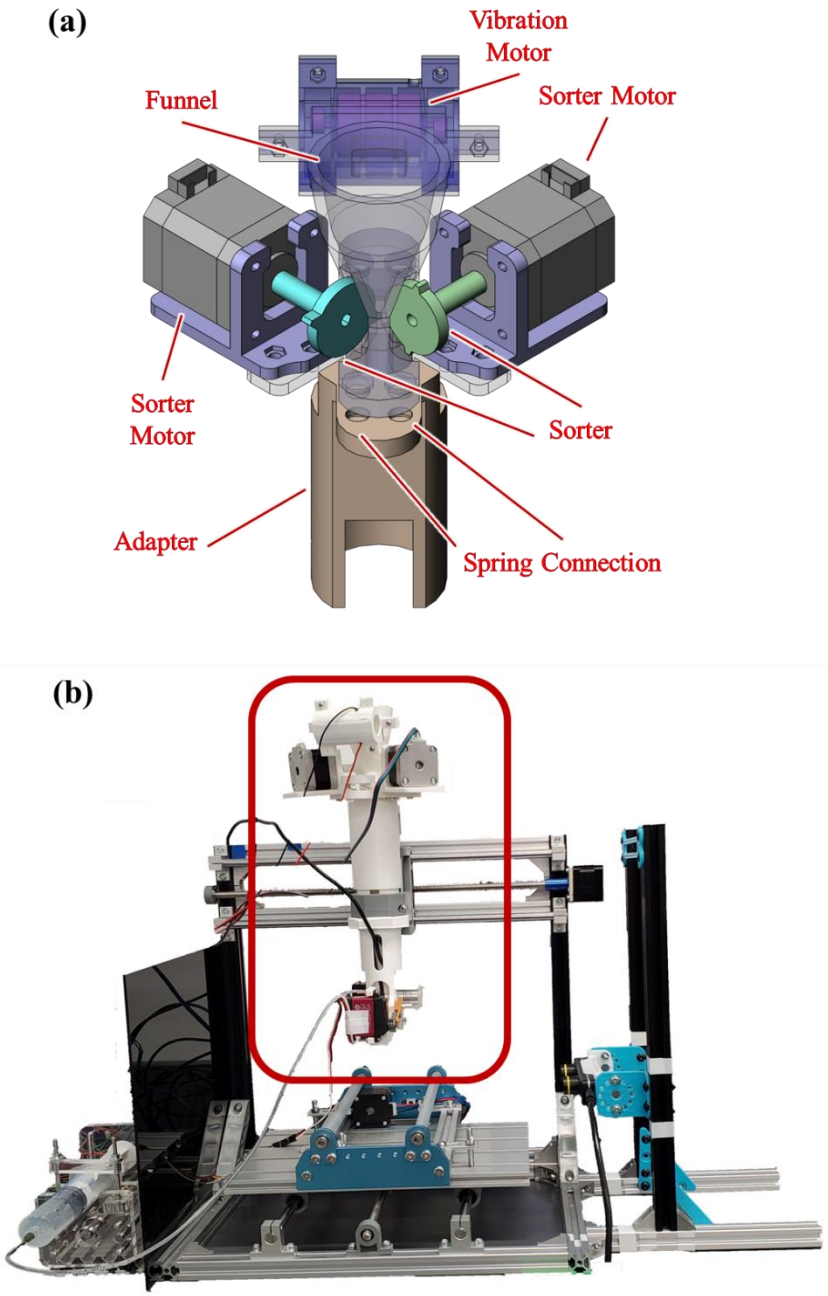

Figure 3.8. Reloader. (a) Auxiliary view. (b) Photo of the integrated machine with the reloader.

3.1.3. Summary of the tools

The integration of the peg insertion tool and the reloader, with their specialized storage and launching features, streamlines the automated peg-in-hole assembly process. By employing the wiggling and pushing mechanisms, the system adeptly addresses the challenges of insertion, guaranteeing precise and stable assembly. Using a 3D printer (Mega X, Anycubic) and PLA filament (1.75mm), prototypes of both the peg insertion tool and the reloader are produced.

3.2. Software Design

This section provides an in-depth exploration of the software design of the proposed automated peg-in-hole assembly system prototype. The software design encompasses critical aspects such as the communication protocol, assembly process, camera selection, and image processing techniques.

3.2.1. Computer specifications and serial communication

The software operates on a central computer, powered by an 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz CPU and an NVIDIA RTX A2000 Laptop GPU with 12 GB RAM. Python (Version 3.9.10) is the chosen language, focusing on image processing and system control algorithms.

The Python program interacts with the Teensy 4.1 microcontroller through serial communication, directing the machine's movements. This microcontroller, programmed in C++, produces control signals based on the instructions from the Python program. Notably, the microcontroller employs digital signals, encompassing digital high, low, and PWM signals, for

motor control. A core aspect of the design is the communication protocol between the microcontroller and the central computer.

The communication between the central computer and the microcontroller is facilitated through Universal Asynchronous Receiver-Transmitter (UART) methodology. In this method, data is transmitted serially, byte by byte. The chosen baud rate for this system is 9600 bit/sec.

The communication protocol is designed to streamline the interaction between the Python program (central computer) and the microcontroller. Commands sent from the Python program to microcontroller primarily include parameters including: the motor ID (indicating axis of movement), direction, and magnitude of the desired movement.

Upon receipt of a command, the microcontroller acknowledges by sending back a Boolean value: "False" or "True". The transmission of a "False" response signifies that the microcontroller has received the command and is initiating the requisite motor control sequence based on the provided instructions. Once the motor operation concludes, the microcontroller returns a "True" signal, denoting that it is ready to accept and execute the subsequent command.

This feedback mechanism is pivotal as it ensures command serialization, eliminating the risk of command overlaps. Such serialization is crucial as it prevents potential errors that might arise from concurrent computations during machine operation.

3.2.2. Assembly process

This section presents the implementation and programming of the assembly process presented in Fig. 3.9. The algorithm begins by waiting for the placement of the workpiece on the base machine’s worktable, which is detected by the side-view camera. Upon confirmation of the workpiece’s presence, the algorithm proceeds to position the system based on the workpiece’s

centre. Subsequently, a series of image processing techniques is applied to locate the holes. By comparing the displacement of all circles with the predefined target location, the algorithm calculates the tool path based on the nearest hole. The target location is determined considering the peg launch position and the camera's location. The algorithm aligns the hole using the computed tool path and applies glue. The final step involves launching and inserting the peg into the hole, and the process repeats until all holes are filled. Throughout the assembly, the algorithm ensures peg alignment and insertion by coordinating the synchronized motion of the base machine, peg insertion tool and other system components. Feedback from vision sensors is integrated into the algorithm to validate peg insertion and trigger subsequent steps. The subsequent sections will explore the selection of cameras and the application of specific techniques to facilitate the assembly process, providing a comprehensive understanding of the software design aspects of the automated assembly system.

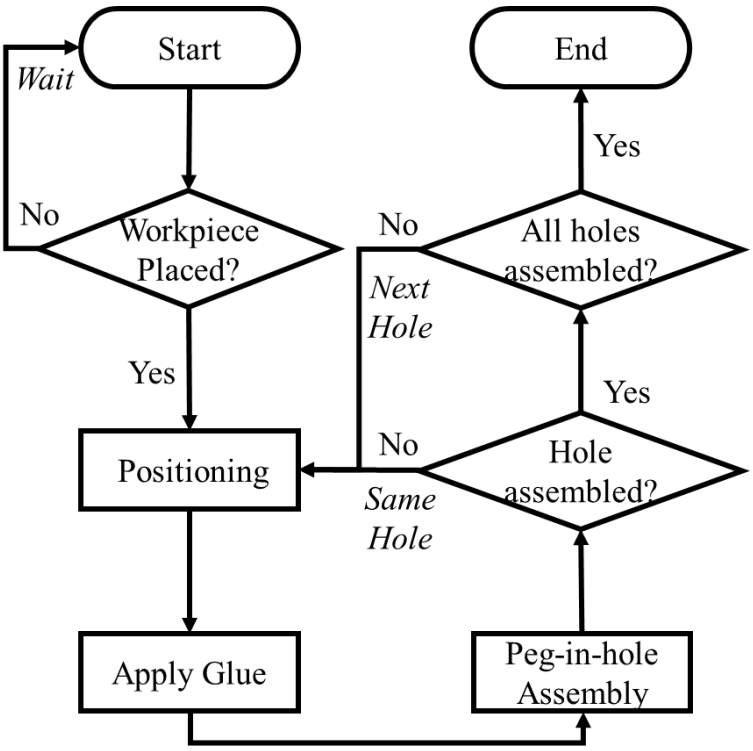

Figure 3.9. Assembly process flow diagram.

3.2.3. Camera selection and integration

Consideration is given to the selection and integration of cameras into the software design. The chosen cameras are positioned to capture different perspectives of the workpiece and the peg insertion tool as shown in Fig. 3.2. The Zed mini camera, chosen as the side-view camera, offers depth sensing capabilities and comes with a software development kit that facilitated the custom software development for this project. Positioned at a distance of 125–150 mm from the centre of the worktable and the rotary roller, the side-view camera provides a comprehensive view of the entire worktable. To minimize external interference from the background, an acrylic board is placed at the opposite end of the machine. However, due to the resolution limitations and the

substantial distance between the side-view camera and the workpiece, the accuracy is insufficient to compensate for the 200 µm clearance between the peg and the hole. To address this challenge, a top-view camera is incorporated into the system, offering a direct and consistent view of the workpiece and the hole within a range of 15 mm. It also serves as a verifier, confirming successful peg insertion before the algorithm proceeds to the next hole. The top-view camera selected is an endoscope with a 5.5 mm outer diameter, installed inside the peg insertion tool using a custom clamp mount, as illustrated in Fig. 3.4a. The technical specifications of the cameras used are provided in Table 3.1, offering a comprehensive overview of their capabilities and functionalities within the automated assembly system.

Table 3.1. Camera technical specifications.

Attribute	**Side-view camera**	**Top-view camera**
Camera	Zed mini camera	T TAKMLY endoscope 5.5-HD
Type	Stereo camera	Endoscope
Output resolution	1920x1080 @30fps	1280x720
RGB sensor type	1/3” 4MP CMOS	5MP CMOS
Active array size	2688x1520 pixels (4MP)	2592x1944 pixels (5MP)
Focal length	2.8*mm* (0.11”) - f/2.0	N/A
Shutter	Electronic synchronized rolling shutter	N/A
V FOV	70°	66°
H FOV	100°	66°

3.2.4. Image processing

The software design incorporates a four-step image processing algorithm, as shown in Fig. 3.10. This algorithm is incorporating multiple techniques to detect the workpiece, define the ROI, extract the hole locations and accurately position the machine based on the calculated distance between the peg launch zone and the extracted holes. To accomplish these tasks, the algorithm extensively utilizes the OpenCV library [61], a popular CV library known for its comprehensive image processing capabilities. The algorithm follows the following sequential steps.

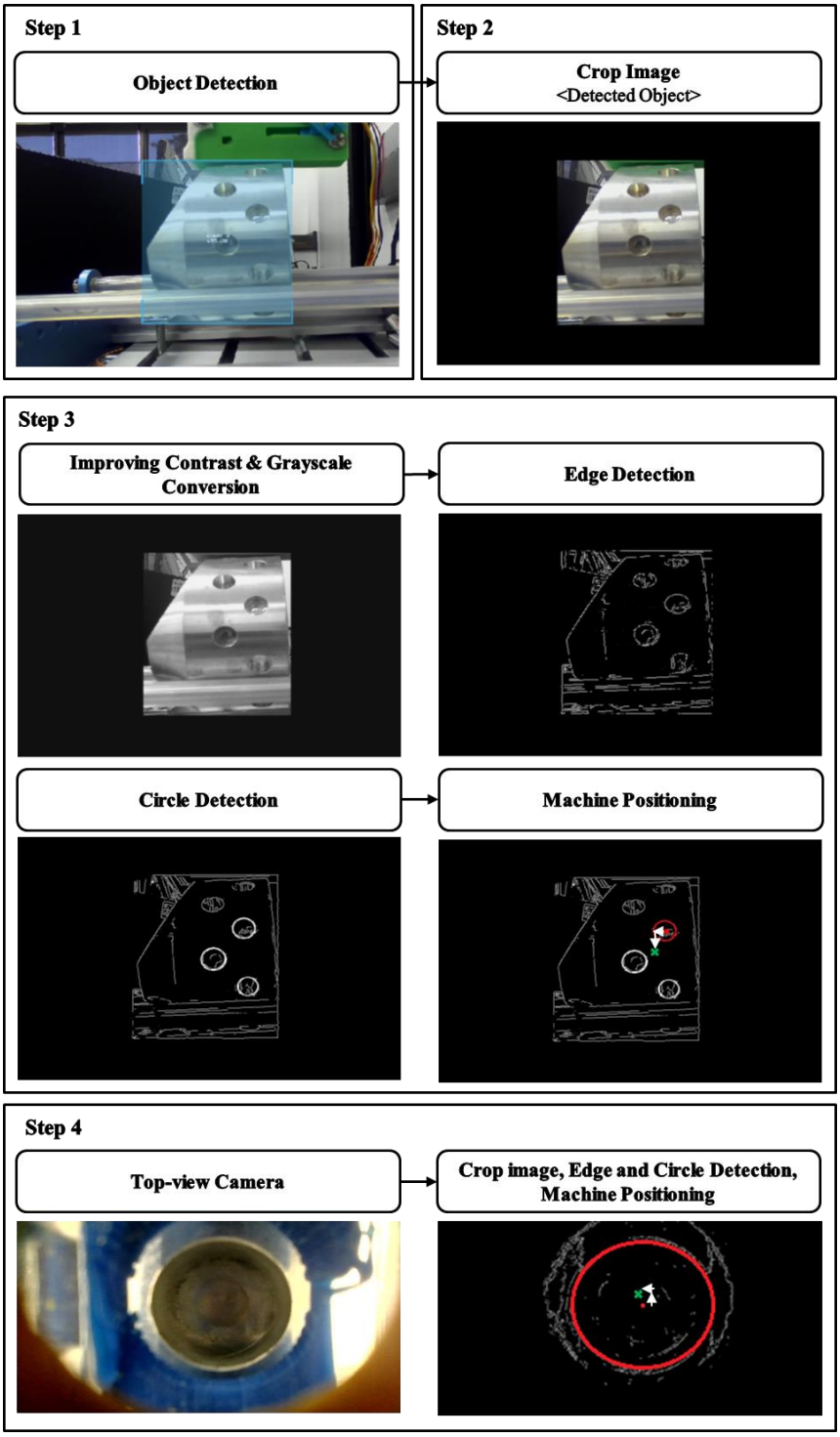

Figure 3.10. Overview of the image processing and machine positioning algorithm. Red circle: Detected closest circle; Red dot: Centre of the detected closest circle; Green "×" mark: Target point.

In the first step of the image processing algorithm, the system remains in a standby state until the side-view camera, which provides a comprehensive view of the workspace, detects the presence of the workpiece. To achieve this, the YOLOv5 object detector [62] is employed, and a customized model is trained using PyTorch (Version 1.10.1 + CU113). This model is specifically trained to accurately detect the workpiece within the captured images.

The dataset used for model training comprises 697 raw images of the workpiece. To improve the training process and enhance the model's ability to generalize, data augmentation techniques are applied. These techniques introduce variations and diversity into the dataset, resulting in a total of 7520 augmented images. The augmentation process includes transformations such as cut out, sharpen, flip and darken, as shown in Fig. 3.11. This process creates a more comprehensive dataset that covers different lighting conditions and perspectives.

To evaluate and validate the trained model, the augmented dataset is divided into three subsets: train, validation and test datasets. The train subset is the largest, containing 6893 images, and is used as the primary dataset for training the model. The validation subset consists of 313 images, which are utilized during the training process to assess the model's performance, fine-tune its parameters and prevent overfitting. Finally, the test subset comprises 314 images, serving as an independent dataset for evaluating the final performance and generalization ability of the trained model.

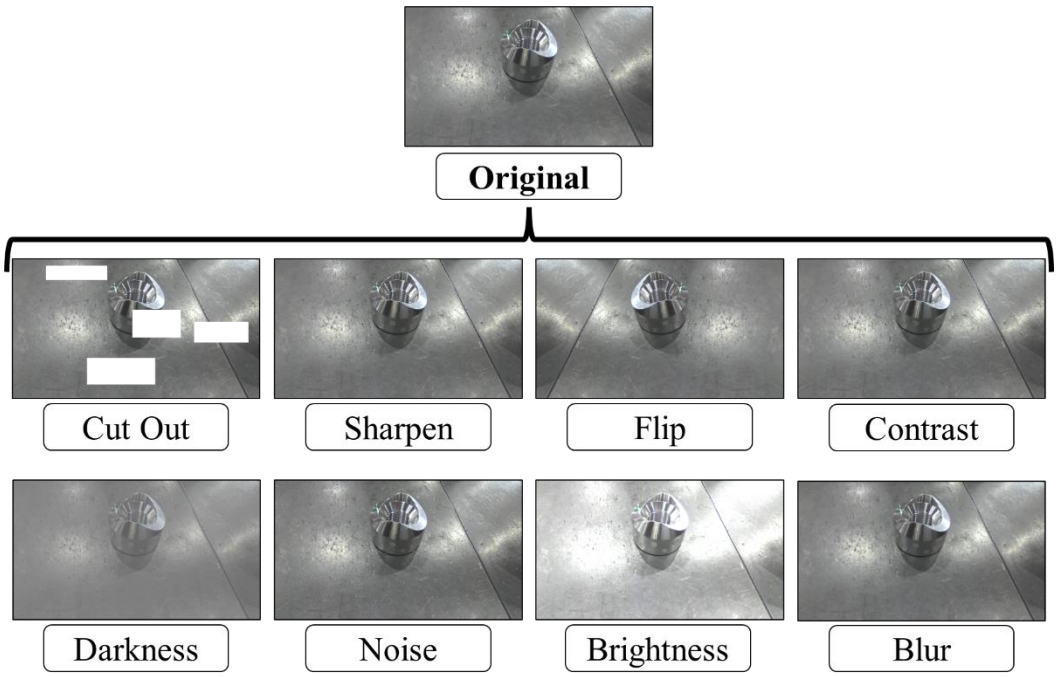

Figure 3.11. Data augmentation for YOLOv5 model training.

To facilitate the training process, the YOLOv5l model is selected as a pretrained model, providing a strong foundation for further training. Several parameters are optimized to ensure efficient training within the available GPU memory. The training is conducted over 100 epochs, with a batch size of 13 and an image size of 416. The box loss specifically gauges the algorithm's ability to accurately locate the centre of an object and ensure that the predicted bounding box adequately encompasses the object. The objectness loss measures the probability of an object's existence within a proposed ROI, providing valuable insights into the likelihood of an image window containing an object. The classification loss evaluates the algorithm's capability to predict the correct class of a given object. During the training process (Fig. 3.12), convergence is observed in all the loss metrics, including objectness, box and classification losses. These losses gradually

approach a value close to 0, indicating that the algorithm effectively locates the centre of objects, accurately predicts bounding boxes and correctly classifies objects within the proposed regions of interest. Notably, there is a rapid decline in these losses until approximately epoch 25, showcasing substantial improvement in the training and validation datasets. These metrics serve as essential indicators of the model's accuracy and ability to correctly identify and classify objects. As a result, the model trained in this stage is considered apt for incorporation into the assembly system and is subsequently integrated.

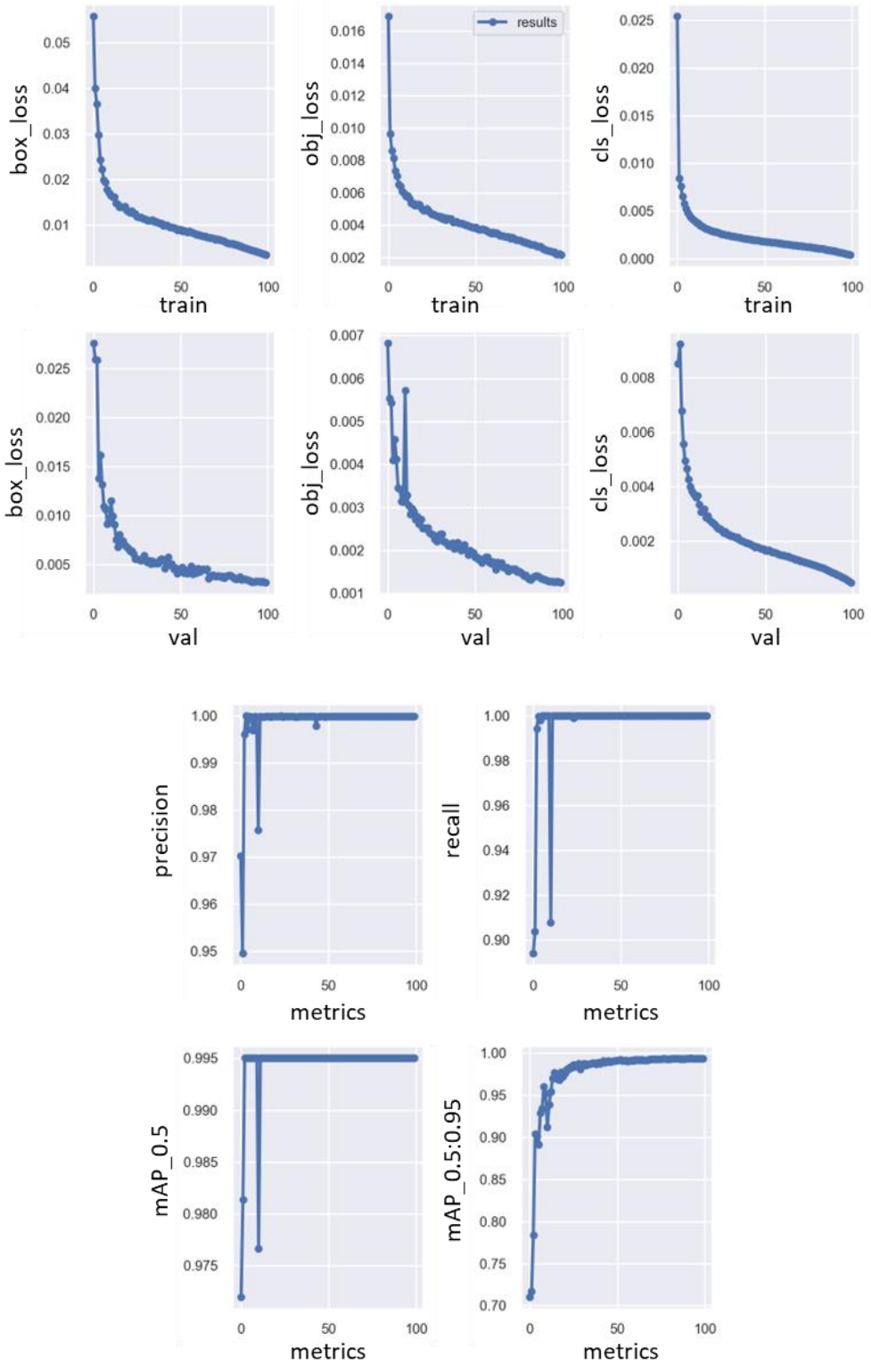

Figure 3.12. Custom YOLOv5 model training results.

The integration of the object detection model into the assembly system allows for the detection of the workpiece as soon as it enters the field of view of the side-view camera. Upon detection, the system proceeds to draw a boundary box around the identified object using the custom trained YOLOv5 model. However, the system does not immediately progress to the next stage of the assembly process. It instead invokes a placement verification step to confirm the workpiece's positioning within the designated working zone. During this verification, the system observes the workpiece for a continuous span of four seconds. This duration affords the operator sufficient time to position the workpiece on the workspace subsequent to its initial detection by the system.

In the second step of the image processing algorithm, the machine performs necessary movements to bring the workpiece to the location of the peg insertion tool. This step also involves an image cropping process to prepare future steps. Prior to image cropping, the following algorithm is implemented to centre the detected workpiece with respect to the peg insertion tool location:

i. The algorithm establishes the relationship between the image coordinate system and the machine coordinate system to facilitate accurate positioning.
ii. Based on the peg insertion tool location, a target point is defined, and the centre point of the detected workpiece is computed.
iii. The target and centre points are then converted from the image coordinate system to the machine coordinate system using the previously identified relationship.
iv. The relative position between the two points is calculated.
v. The machine is moved until the target and centre points overlap, ensuring the workpiece is centred correctly.

From the perspective of the side-view camera (Fig. 3.13), the vertical position of the detected workpiece is assumed to be fixed. Only the horizontal position within the image coordinate system is taken into account for the centring process, based on the workpiece location. The object detector returns the location of the four corners of the boundary box. The workpiece centre is computed by summing the coordinate value of all four corners and dividing it by four. Then, the algorithm calculates the displacement in pixels by subtracting the target location from the centre location of the workpiece. For precision, measurements from five frames are captured. The centre location closest to the target position is chosen for path creation. Prior to initiating machine movements, the pixel displacement is converted to displacement in millimetres using a conversion relation established in Fig. 3.14 and Equation (3-1).

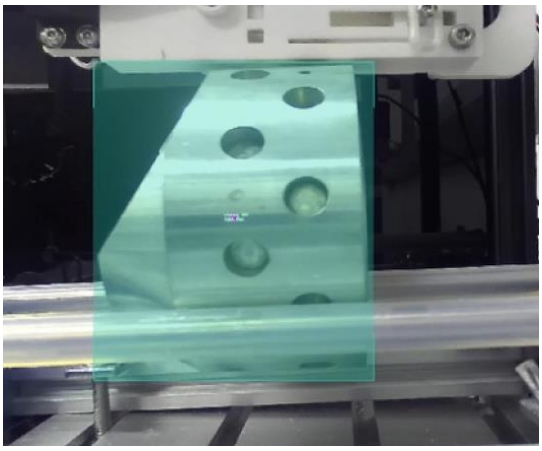

Figure 3.13. Visual representation of the assembly workspace and the workpiece as captured by the side-view camera.

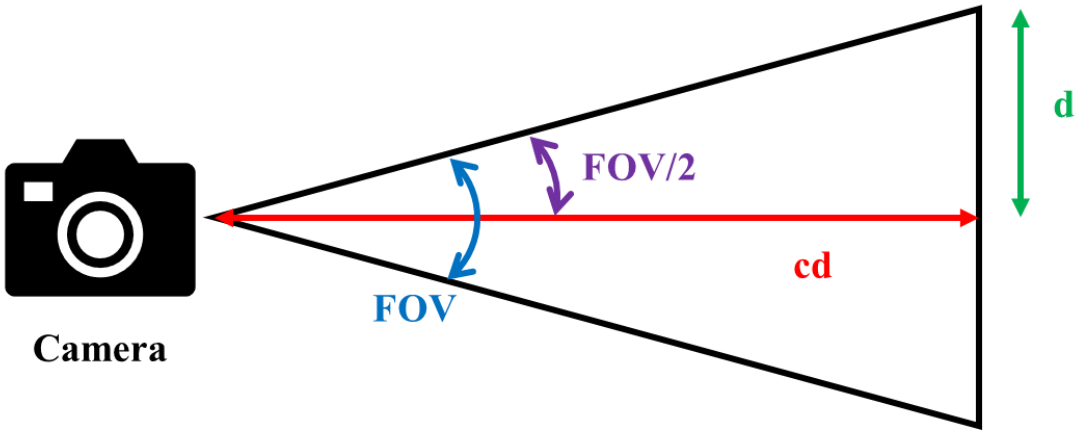

Figure 3.14. Pixel to mm conversion.

$$d\ (mm) = cd * \tan\left(\frac{FOV}{2}\right)$$

$$\frac{mm}{pixel} = \frac{2d}{Total\ \#\ of\ pixels\ in\ the\ axis} \tag{3-1}$$

where d is the distance from the center of the captured image to its edge, measured in millimeters. cd is camera distance, denoting the gap between the workpiece and the camera. FOV is field of view, which is the visible scene captured by the camera lens. The cd is determined by the mount system demonstrated in Fig. 3.2. For the side-view camera, the cd is continuously verified using the depth measuring feature based on stereography. The FOV is provided in the camera datasheets as listed in Table 3.1. Additionally, the camera datasheets specify the total number of pixels in the vertical and horizontal axes. For instance, in the case of 1080p resolution, the vertical and horizontal axes utilize 1080×1920 pixels, respectively. By utilizing these parameters, including cd, FOV and pixel dimensions from the camera datasheets, the algorithm accurately converts the displacement in pixels to displacement in millimetres. It is important to note that separate conversion magnitudes are utilized for the side-view and top-view cameras,

tailored to their specific parameters. This consideration accounts for the differing characteristics of each camera and ensures precise and reliable measurements in the respective camera views.

Upon obtaining the precise measurements using the aforementioned technique, the assembly system proceeds to send commands to the microcontroller and the machine for executing the required movement. Once the movement is completed, the vision system enters a verification phase to assess the alignment accuracy. This verification is achieved by comparing the position of the object centre with the target location. To establish a feedback loop and ensure precise alignment, the system iterates through this process continuously until the object centre is within a predefined threshold distance of the target location. In this particular case, the threshold value is set at 0.8 mm, indicating that the system aims to achieve a high level of accuracy, with the object centre being within 1.2 mm of the desired target location.

Once the positioning is completed, the system crops the image by setting the boundary box as the ROI. The background is replaced with a black image of the same size. This step prevents the background information from being processed in subsequent image processing steps while maintaining the correspondence between the machine and image coordinate systems.

In the third step of the image processing algorithm, the holes are extracted, and the machine is positioned and prepared for peg insertion. The image processing techniques are applied sequentially in the following order: Contrast Limited Adaptive Histogram Equalization (CLAHE), Grayscale Conversion, Canny Edge Detector and Hough Circle Transform [63, 64, 65].

The workpiece used in the assembly process is made of a reflective metallic material, which often results in the appearance of light marks on its surface. These light marks can be similar in colour to the white pegs, occasionally obstructing the algorithm’s ability to detect the holes when

they are assembled with white pegs. To address this issue, CLAHE is employed. CLAHE improves the contrast of the image based on the local grayscale distribution, making the edge boundaries of the holes more distinctive while minimizing the impact on other regions of the image [63].

Next, the image is converted to grayscale as a prerequisite step for edge detection. Grayscale conversion simplifies the image representation by eliminating colour information while retaining the necessary intensity values. Then, the Gaussian blurring is applied to the input grayscale image, which aims to smooth the image [66]. The Canny Edge Detector is then applied to the grayscale image. This edge detection technique uses the calculus of variations to convert the image into a binary image, highlighting the edges only [64, 67]. By emphasizing the edges, the algorithm can focus on the relevant features for hole detection.

The image, post-processing via the Canny Edge Detector, undergoes evaluation using the Hough Circle Transform to identify circles. The Hough Circle Transform is a specialized algorithm devised for circle detection in images. For every edge pixel from the input image, probable circle centers are deduced based on a designated range of radii in conjunction with the pixel's gradient direction. Votes are accumulated in a 3D matrix for these potential centers and their respective radii. Predominant vote accumulations in this matrix denote probable circle centers with their associated radii. By specifying a radius range, the algorithm narrows its focus to circles within that size range, bolstering both computational efficiency and precision of detection [65, 68]. The side-view and top-view cameras may require different parameter values to achieve the better accuracy in circle detection. Therefore, individualized parameters are designated for each camera. By implementing the Hough Circle Transform using optimal parameters, the algorithm isolates circles from binary images, providing circle data for subsequent stages in the assembly process.

Recent studies have explored the use of ML-based techniques for hole detection, which offer high accuracy with reduced false positives [8]. However, these techniques tend to be more complex and computationally intensive compared to the image processing algorithm proposed in this study. The algorithm presented in this study, which includes image cropping, CLAHE and Hough Circle Transform, achieves a high level of accuracy while imposing a lighter computational load.

After finding circles with the Hough Circle Transform, the algorithm calculates the distance between each circle and the target point, following a similar methodology as described in step 2. To enhance precision, data from two consecutive frames are factored into this computation. Ultimately, the machine adjusts its position until the nearest identified hole aligns with the target point. However, it is crucial to acknowledge the presence of potential uncertainties in the final positioning due to the inherent vibrations that may cause shifts in the relationship between the image and machine coordinate systems, due to the linkage limitation of the side-view camera as discussed in Section 3.2.3. These vibrations can introduce variations and compromise the accuracy of the positioning, thus introducing a level of uncertainty in the assembly process. In response to this challenge, step 4 is introduced wherein the algorithm is applied to the top-view camera, which offers improved accuracy and provides reliable data for the assembly process.

In step 4, the algorithm seamlessly transitions to the top-view camera, which is positioned inside the peg insertion tool to provide a direct and unobstructed view of the assembled hole. This configuration ensures a stationary relative position between the top-view camera and the peg launch zone, reducing the impact of uncertainties compared to the side-view camera, which involves multiple links in between. Similar to the side-view camera, the top-view camera undergoes grayscale conversion, Canny edge detection and Hough Circle Transform techniques

for hole extraction. These image processing techniques enable the identification and extraction of the holes with high accuracy. The relative position between the closest hole and the newly defined target point for the top-view camera is calculated. To facilitate this calculation, a new relationship between the top-view camera's image coordinate system and the machine coordinate system is established as mentioned in step 2, enabling the conversion of the relative position to the machine coordinate system.

Once the relative position is determined, the machine adjusts its position until the hole aligns with the target point. For the top-view camera, the data from three frames are used to find the closest circle to the target location. For this step, the alignment threshold value of 0.3 mm is adopted. After the alignment, the glue application, peg launch and the two assembly mechanisms described in Section 3.1.1 are activated to complete the hole assembly process.

To address potential failures in the peg assembly process, an additional verification step is implemented using the top-view camera. This verification step becomes crucial, especially considering the tight tolerance requirement of 200 μm. When an insertion failure occurs, the peg remains in the peg insertion zone directly beneath the top-view camera, obstructing its view, as illustrated in Fig. 3.15. To overcome this challenge, the verification process focuses on detecting circles after the assembly mechanisms have been executed. If the top-view camera fails to detect any circles, it indicates that the peg has not been successfully inserted, despite the alignment being correct. Leveraging this characteristic, the system repeats the assembly mechanisms until circles are detected by the top-view camera once again. However, if the elapsed time during this process exceeds eight seconds, indicating a prolonged unsuccessful attempt, the system takes an alternative approach. It clears the peg from the peg launch zone and proceeds to repeat the entire sequence,

starting from circle alignment and going to glue application, peg launch and two assembly mechanisms. This allows sufficient time for the glue to cure without the peg being present. By incorporating this additional verification step and implementing the necessary adjustments in case of prolonged failures, the system enhances the reliability and accuracy of the peg assembly process, ensuring the successful and secured assembly of the peg within the specified tolerance.

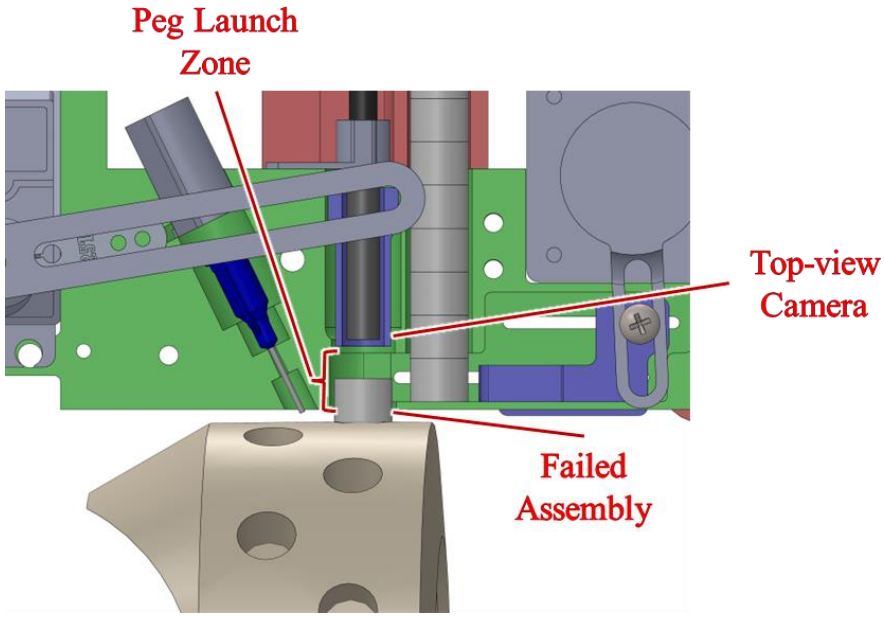

Figure 3.15. Potential assembly failure.

After the successful completion of the verification step, the workpiece undergoes rotation to orient the tool toward the next hole, followed by repetition of steps 3 and 4. This iterative process continues until all the holes in the workpiece have been effectively assembled.

The integration of the top-view camera plays a crucial role in this assembly procedure. By capturing the necessary visual information and executing the sequential steps outlined earlier, the algorithm ensures the accurate detection of hole locations and precise alignment. This integration

of the top-view camera, combined with the systematic execution of the assembly steps, results in a reliable and efficient process for the assembly of the workpiece.

Overall, this approach enables the automated assembly system to consistently and accurately assemble the workpiece, providing a high level of reliability and efficiency in the assembly process.

3.3. Summary

In this chapter, the image processing techniques employed in the automated assembly system for precise hole detection and alignment are discussed. The system utilizes a combination of CV algorithms and machine control to achieve accurate and reliable assembly of workpieces.

The chapter began by introducing the image processing steps involved in the detection and extraction of holes. The process starts with capturing images from a side-view camera and a top-view camera. The captured images are then processed using various techniques to identify the holes accurately.

The first step in the image processing algorithm involves obtaining the gradient magnitude and direction using the Canny edge detection algorithm. The gradient magnitude represents the rate of change of pixel values in the image and helps identify areas with significant intensity variations, corresponding to edges. The gradient direction is quantized into four angles, simplifying the representation of edge directions for further processing steps.

After obtaining the gradient magnitude and direction, non-maximum suppression is applied to eliminate redundant pixels that do not represent edges. This step ensures that only the pixels corresponding to the sharpest change in intensity, the true edges, are preserved. Hysteresis

thresholding is then used to classify pixels as sure edges or non-edges based on their intensity gradients, further refining the edge representation.

By employing these image processing techniques, a binary edge image is generated, where white pixels represent detected edges, and black pixels represent non-edges. This binary edge image provides a precise representation of the edges in the original image, enabling subsequent steps in the assembly process.

The next step in the algorithm involves applying the Hough Circle Transform technique to detect circles in the image. This technique utilizes specific parameters such as the minimum distance between circle centres, edge intensity threshold and circle radius range to identify and extract circles accurately. Separate lines of code are used for the side-view and top-view cameras, considering their distinct optimal parameter values.

The integration of the top-view camera is particularly crucial in the assembly procedure because it offers an unobstructed view and reduces uncertainties associated with vibrations. Similar to the side-view camera, the top-view camera undergoes grayscale conversion, Canny edge detection and Hough Circle Transform techniques for hole extraction. The relative position between the closest hole and the target point is calculated, facilitating precise alignment.

To address potential failures in the peg assembly process, an additional verification step is implemented using the top-view camera. This step focuses on detecting circles after the assembly mechanisms have been executed. If no circles are detected, indicating a failed insertion, the system repeats the assembly mechanisms until circles are detected again. If a prolonged unsuccessful attempt occurs, the system clears the peg and restarts the entire sequence, ensuring reliable assembly.

The chapter concludes by emphasizing the reliability and efficiency achieved through the integration of the top-view camera and the systematic execution of assembly steps. The combination of CV algorithms and machine control enables consistent and accurate assembly of workpieces, ensuring successful peg insertion within specified tolerances. Overall, the image processing techniques discussed in this chapter provide a robust foundation for automated assembly systems, facilitating precise hole detection, alignment and reliable assembly processes.

Chapter 4. Vision-Based Quality Inspection

The goal of the proposed QI system is to detect positive defects, specifically unfilled holes, during the assembly process. To achieve this, the QI system utilizes the hardware system, various image processing techniques and the side-view camera integrated into the assembly system, as described in Chapter 3. In this chapter, we evaluate two QI methods: a statistical method and ML-based methods.

Both approaches directly require the cropped image of the hole as their input. Thus, this chapter includes a prerequisite step, the hole-cropping algorithm, which identifies and extracts the hole images for further analysis. Statistical method is a novel approach developed specifically for this hole assembly process, utilizing traditional CV techniques. We subsequently compare this method's performance and feasibility with ML-based methods, including traditional CNN and ResNet architectures. The comparative analysis aims to determine the most effective and suitable method for defect identification in our assembly process.

4.1. Hole-cropping algorithm

The hole-cropping algorithm is explained in Fig. 4.1. The objective of this algorithm is to crop out the hole images from the assembly image while it is being assembled. Traditional image processing techniques including grayscale conversion, Canny Edge Detector and Hough Circle Transformation are first used in the process to detect the circles similar to the process elaborated in Section 3.2.4. However, these techniques often detect false positives, which will always produce the wrong results once input into the QI methods. Hence, they should always be filtered before the

QI predictions. The false positives cause more detrimental effects in QI compared to the assembly process. The assembly process receives continual update from the real-time view of the workspace, hence it can self-correct and mitigate the effect of the false positives from updating frames. However, the QI system takes in a single image as an input to make the predictions. Hence, it does not possess the ability to self-correct. To resolve this problem, the hole-cropping algorithm is developed.

Fig. 4.1 presents the hole-cropping algorithm, designed to extract hole images from the assembly image during the assembly process. The algorithm utilizes traditional image processing techniques, such as grayscale conversion, Canny edge detection and Hough circle transformation, similar to the assembly system image processing process detailed in Section 3.2.3. However, these techniques often generate false positives, which can lead to inaccurate results when fed into the QI methods. As the assembly process continuously updates from views of the workspace, it can self-correct and mitigate the impact of false positives in updating frames. Conversely, the QI system takes a single image as input for predictions and lacks the ability to self-correct. To address this limitation, the hole-cropping algorithm is developed.

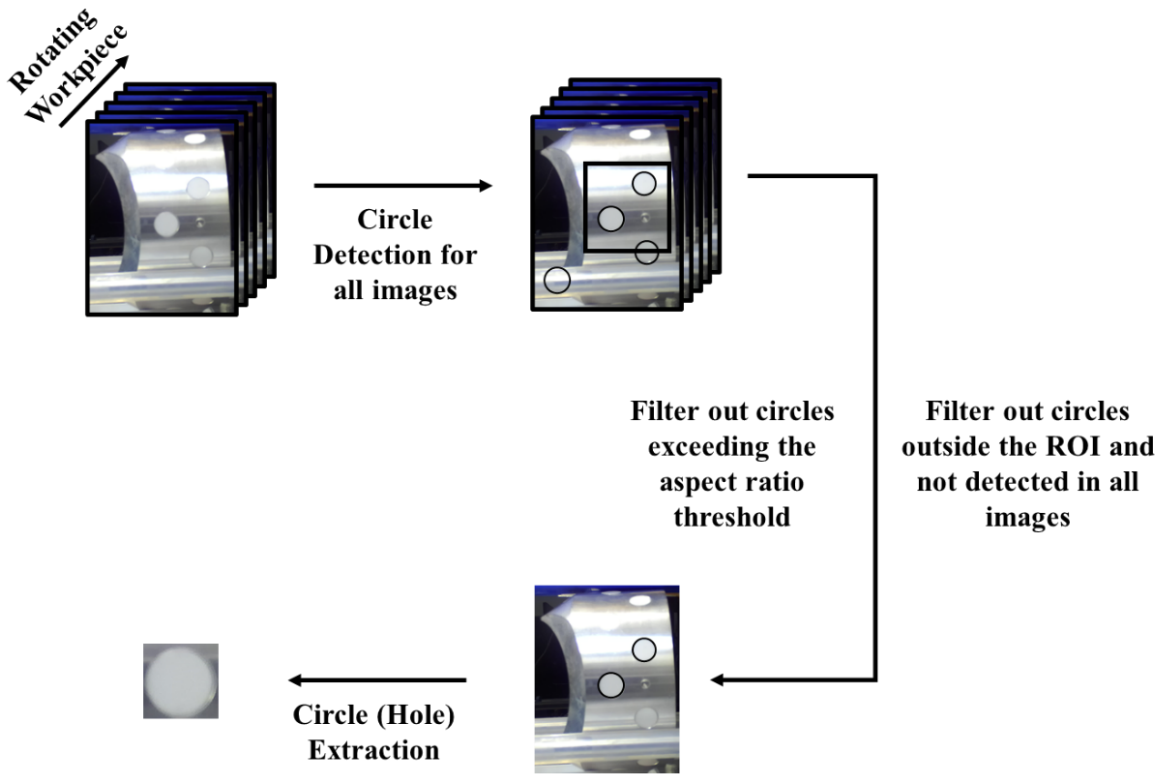

Figure 4.1. Hole-cropping algorithm.

The hole-cropping algorithm functions in the following sequence:

i. Five images of the rotating workpiece are acquired, each with an approximate 1° rotation interval between them.

ii. Holes are detected in all images using image processing techniques.

iii. Holes located outside the new ROI and those exceeding the threshold aspect ratios (defined as 0.75:1 and 1.25:1) are filtered out. The new ROI encompasses regions that provide undisturbed and undistorted hole images, determined based on their relative position to the centre of the boundary box. Holes with distorted aspect ratios, falling outside this range, cannot be accurately predicted using QI methods during further analysis.

iv. The closest detected hole from the preceding image is identified, starting from the second image.

v. Holes with a distance exceeding the threshold distance (equivalent to 1° rotation) to the closest hole in the preceding image are eliminated. For instance, if a hole is detected in the second image, it is linked to the closest circle in the first image. If the relative distance between two holes exceeds the 1° equivalent distance, both circles are filtered out. However, if the relative distance is less than or equal to 1°, the circle in the second image is retained for the next step.

vi. Steps iii–v are repeated until all acquired images are processed.

vii. The remaining holes from the processed images are cropped and subjected to the QI methods.

Through this process, the holes are effectively detected and cropped during the assembly process using traditional image processing techniques. By applying rigorous filtering, it eliminates false positives and ensures accurate hole identification. The algorithm's iterative approach and new ROI creation and aspect ratio thresholding result in reliable data for QI predictions. This enhances the QI system's performance, making it a valuable tool for quality control in assembly and related applications.

4.2. Statistical method

The statistical method capitalizes on variations in light reflection intensity levels observed at the holes during assembly (negative defects) or when they are not assembled (positive defects). The conventional CV reflection model is described by the following formula [69].

$$I = E\rho h(\theta_i, \phi_i, \theta_r, \phi_r)\cos\theta_i \quad (4\text{-}1)$$

where I represents the light reflection intensity level, E denotes the light source intensity, (θ_i,ϕ_i) and (θ_r,ϕ_r) represent the incident angle and the viewing angle relative to the surface normal, ρ is the surface reflectance and $h(\theta_i,\phi_i;\theta_r,\phi_r)$ is the reflectivity function.

In the context of our assembly process, the workpiece is constructed from magnesium, which exhibits a high value for surface reflectance (ρ). Additionally, the unassembled holes in the workpiece create three additional surfaces with different surface normals. Consequently, the light is reflected in multiple directions when the holes are unassembled, leading to rapid fluctuations in the intensity level within the hole region. Conversely, in the case of the assembled workpiece, the holes are filled with ceramic pegs that possess significantly lower ρ values. As a result, light is primarily reflected in the direction normal to the workpiece surface, leading to a more consistent light intensity level with minimal fluctuations on the filled holes.

The high ρ value of magnesium and the presence of multiple surfaces contribute to the distinct variation in light reflection intensity observed in the unassembled holes, making them identifiable as positive defects. Conversely, the reduced ρ value and uniform reflection in the direction normal to the workpiece surface make the assembled holes distinguishable, indicating negative defects in the assembly process. By exploiting these variations in light reflection intensity, the statistical method offers a viable approach for accurate defect detection and QI during the assembly process.

The statistical method encompasses various image processing steps to capture and mathematically express specific behaviours, as depicted in Fig. 4.2. Initially, the Sobel filter is employed to smoothen the hole image and eliminate noise. Unlike its typical use for edge

detection, we adapt the Sobel filter with a box blur kernel to prioritize noise removal over edge sharpening. This adjustment is significant, as modifying filter component values might influence subsequent analyses, specifically changing the attributes at the boundary between areas of varying intensity levels within the hole image. The resulting gradient array computes the gray value fluctuations numerically, with edges of the hole exhibiting higher magnitudes due to intensity discontinuities [70]. Because our method focuses on measuring intensity fluctuations in the hole images, these edge gradient values carry significant information and cannot be disregarded. However, to emphasize inner region information and eliminate edge effects, we remove the first and last 20% of pixel values from the gradient array.

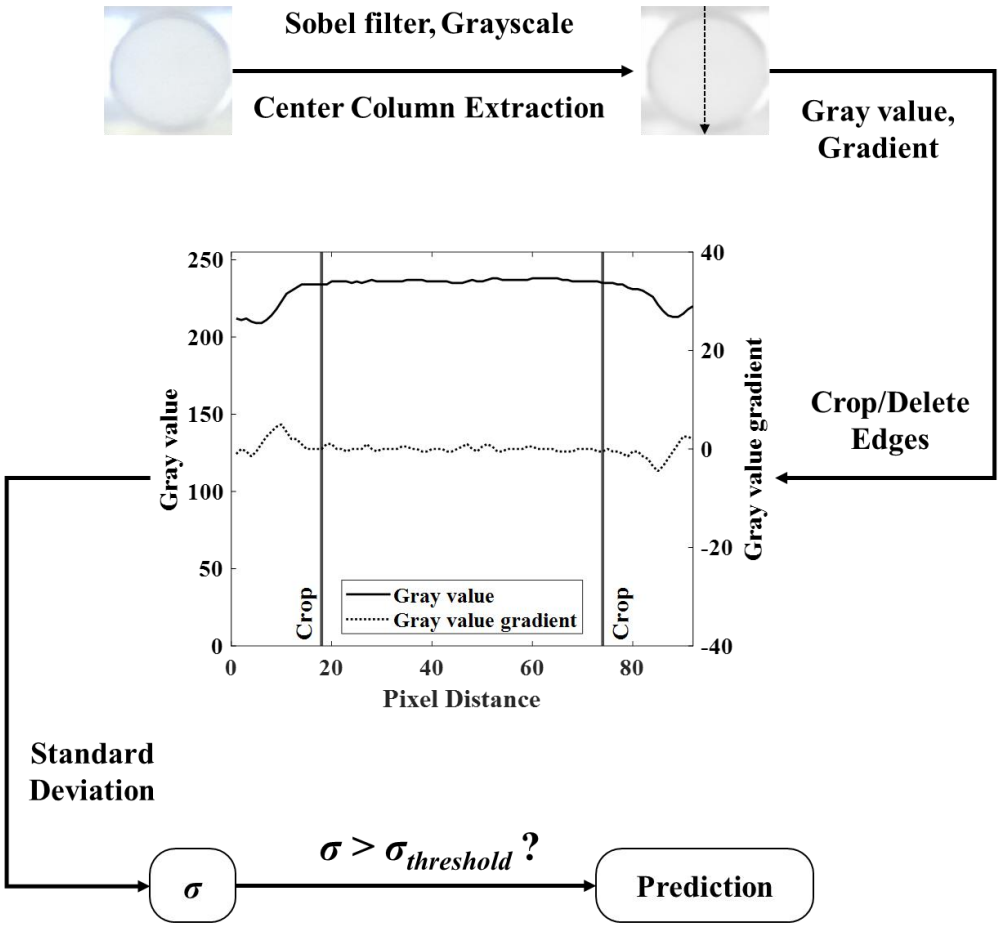

Figure 4.2. Statistical method.

To characterize the intensity fluctuation in the hole images quantitatively, we calculate the standard deviation (σ) of the resulting gradient values. Fig. 4.3 illustrates exemplary results obtained using the statistical method based on the hole status. When white or black pegs are inserted, the gray value graph of negative defect holes remains generally consistent within the cropped region, with changes limited to the ±10 range. Consequently, their gradient values remain close to 0. Conversely, positive defect holes exhibit greater fluctuations in the gray value graph, with a significant leap corresponding to the boundary between the shaded and unshaded area

within the hole. This results in an unstable behaviour throughout the gradient graph, with a sharp spike evident at the shaded area. In summary, negative defect holes yield σ values close to 0, whereas positive defect holes produce σ values greater than 1. Hence, the statistical method effectively classifies positive and negative defects based on the σ values, enabling accurate defect identification in the assembly process.

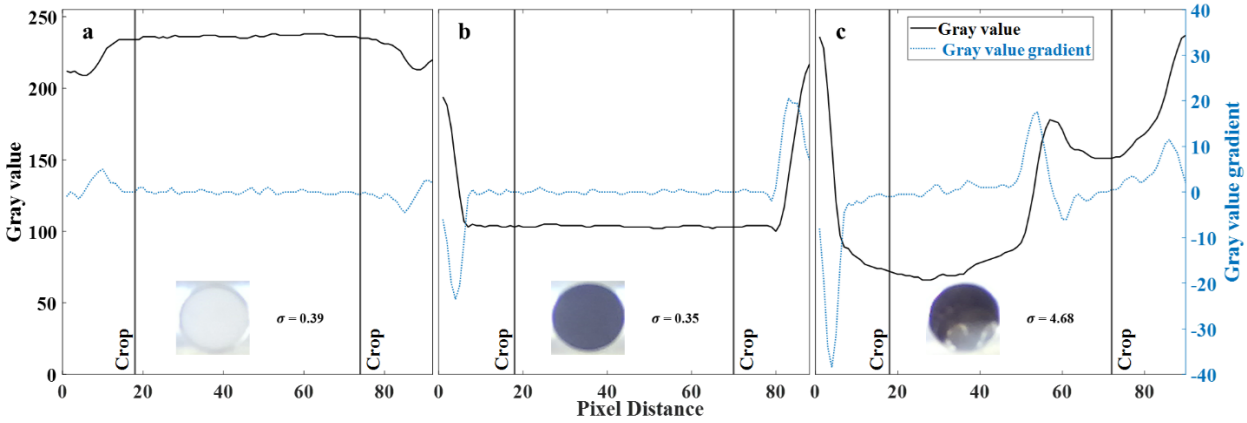

Figure 4.3. Statistical method results based on hole status. (a) Negative defect - white peg inserted. (b) Negative defect - black peg inserted. (c) Positive defect.

The threshold σ value ($\sigma_{threshold}$) denotes the specific numerical value utilized to differentiate positive and negative defects. To determine $\sigma_{threshold}$, the σ values of 100 hole images for each hole status are computed, and their corresponding maximum and minimum values are identified, as presented in Table 4.1. A negative defect hole assembly is defined as one that does not impede the workpiece's rotation on the system's rotary roller. An obstruction is any halt or bump caused by a peg that is only partially inserted, which may protrude and interfere with the rotary roller. Thus, only the assembled workpieces that rotate without such obstructions are used to source negative

defect hole images. Contrastingly, a positive defect hole assembly is defined by its vacancy, which permits the insertion of an additional peg to conform to the negative defect criteria previously described. The minimum σ value for positive defect holes (1.05) exceeds the maximum σ value for negative defect holes (0.72). To establish a balanced threshold, the mean of the minimum and maximum σ values is adopted as $\sigma_{threshold}$ ($\sigma_{threshold}$ = 0.89).

Table 4.1. Minimum and maximum of hundred σ values for each hole status.

Hole status	**Minimum σ**	**Maximum σ**
Negative defect - White peg inserted	0.13	0.72
Negative defect - Black peg inserted	0.20	0.51
Positive defect	1.05	7.32

Consequently, during hole analysis, a hole is categorized as a positive defect if its σ value surpasses $\sigma_{threshold}$, and as a negative defect if its σ value is equal to or less than $\sigma_{threshold}$:

- $\sigma > \sigma_{threshold} \rightarrow$ Positive defect
- $\sigma \leq \sigma_{threshold} \rightarrow$ Negative defect

It is important to emphasize that $\sigma_{threshold}$ is determined based on the maximum value that accounts for both white and black peg-inserted negative defects, solely focusing on the classification of positive and negative defects.

4.3. Machine learning-based method

4.3.1. Convolutional neural network

CNN has gained widespread adoption as a powerful algorithm for QI [71, 72, 73]. Its success in various CV tasks can be attributed to its ability to effectively extract essential features from input images. A typical CNN consists of multiple convolutional layers, followed by a neural network component. The convolutional layers play a critical role in feature extraction by applying filters to the input images, enabling them to capture intricate patterns and distinctive characteristics relevant to the inspection process. Furthermore, the inclusion of pooling layers between convolutional layers is instrumental in reducing the activation map size of the output data. Pooling helps in down-sampling the feature maps, facilitating more efficient processing and contributing to the network's ability to handle larger and more complex datasets. During the feature extraction process, the convolutional layers transform the extracted values into nonlinear values through activation functions. This nonlinearity introduces essential degrees of freedom to the model, enhancing its capacity to learn and distinguish between intricate patterns in the inspection data [46].

The use of CNN for QI is justified by its capability to automatically learn and adapt to the specific features and patterns related to the inspected items. This makes CNN a versatile and robust algorithm for tasks such as defect detection, classification and other quality assessment applications. Additionally, its ability to handle large amounts of data and its capacity for deep learning make CNN a suitable choice for addressing complex inspection challenges and achieving high accuracy in quality control processes. As a result, CNN has become a popular choice for researchers and practitioners in various industries seeking reliable and efficient QI solutions [44].

For this thesis, the CNN architecture shown in Fig. 4.4 is utilized. A hole image, as shown in Fig. 4.5, is extracted from the hole-cropping algorithm and resized to 300×300 for training. The CNN output is of Boolean data type, where “true” denotes positive defect and “false” denotes negative defect holes. Leveraging the capabilities of CNN in learning intricate patterns and distinguishing between defects, this architecture is well-suited for addressing the specific QI requirements of the assembly process and its associated hole defects.

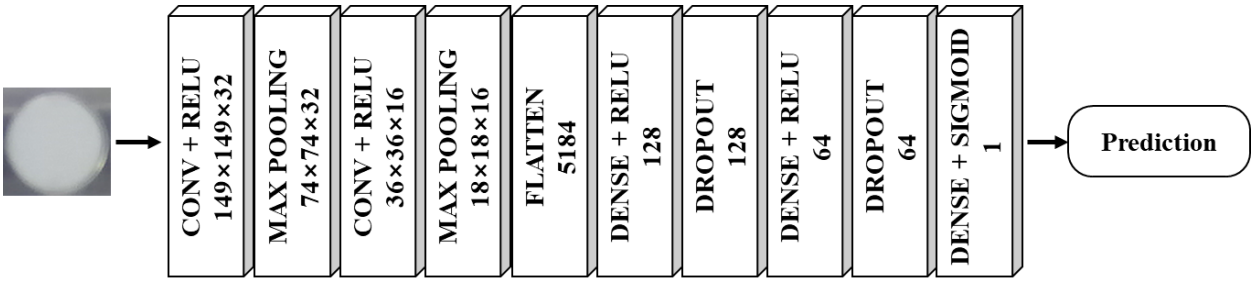

Figure 4.4. CNN architecture used for QI.

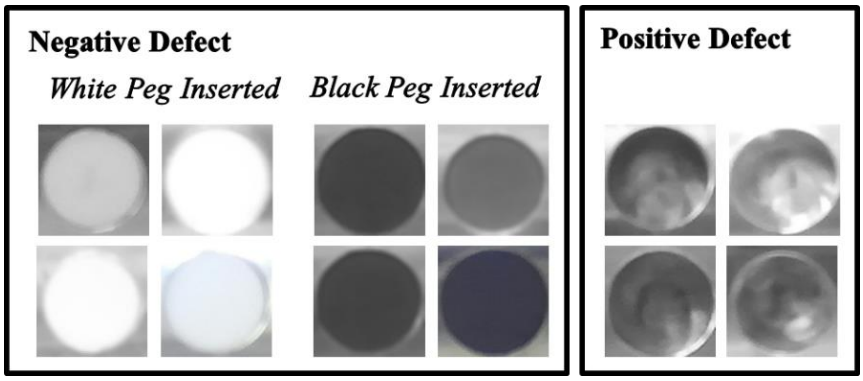

Figure 4.5. Sample hole images used for ML-based QI training.

4.3.2. Residual network

ResNet, another widely recognized ML algorithm, has been extensively studied and proven effective in various research works [15, 60, 74]. ResNet is built upon the CNN structure but introduces skip connections that directly connect inputs to the back layers. This unique architecture allows the back layers to directly learn the residuals, resulting in more efficient training and mitigating the issues associated with vanishing gradients [72].

For this thesis, the ResNet50 model, as illustrated in Fig. 4.6, is utilized to extract the features from hole images. The model is trained using hole images of size 300x300, similar to the sample images displayed in Fig. 4.5. Upon processing a hole image, the ResNet50 model produces a 1x2 size array as its output. Each element in this array represents a class in the model, specifically denoting positive and negative defect holes.

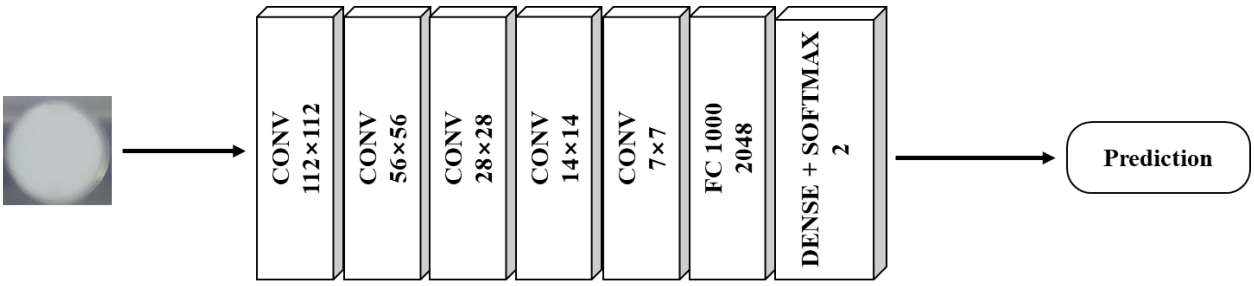

Figure 4.6. ResNet architecture used for QI.

4.4. Summary

In this chapter, the vision-based QI system is presented, with the primary objective of detecting positive defects, specifically unfilled holes, during the assembly process. The QI system utilizes the hardware system, various image processing techniques and the side-view camera

integrated into the assembly system for defect identification. Two distinct QI methods are evaluated: the statistical method and ML-based methods, including CNN and ResNet.

The hole-cropping algorithm serves as a prerequisite step to extract and identify hole images from the assembly images. Traditional image processing techniques, such as grayscale conversion, Canny Edge Detection and Hough Circle Transformation, are employed in this algorithm. By effectively filtering false positives and generating reliable data, the algorithm contributes to enhanced hole identification accuracy during the assembly process.

The statistical method capitalizes on variations in light reflection intensity levels observed at the holes during assembly and when not assembled. By quantifying the intensity fluctuations using the standard deviation (σ) of gradient values, positive and negative defect holes are effectively classified. The threshold σ value ($\sigma_{threshold}$) determines to differentiate between positive and negative defects, further enhancing the accuracy of the statistical method.

In contrast, the ML-based methods, CNN and ResNet, demonstrate their efficacy in feature extraction and defect classification. CNN's ability to learn intricate patterns and handle complex datasets, along with ResNet's unique skip connections and efficient training, make them suitable choices for addressing QI challenges in the assembly process.

By combining the strengths of both the statistical method and ML-based methods, the vision-based QI system provides a comprehensive approach to defect detection in the assembly process. The application of these methods facilitates accurate identification and classification of defects, enhancing the quality control process and ensuring the production of high-quality assembled products.

Chapter 5. Results and Discussions

This chapter presents the comprehensive results obtained from the implementation and evaluation of the proposed automated peg-in-hole assembly system prototype and the QI system. The previous chapters have discussed the design considerations, software development and image processing techniques employed to achieve accurate and reliable hole detection and alignment. In addition to the assembly process, this chapter explores the outcomes of the QI system, which plays a critical role in assessing the quality and accuracy of the assembled workpieces. The QI system encompasses vision sensor (camera) in combination with traditional CV techniques or ML algorithms to verify the alignment and securing of each peg within the hole. The presented results not only focus on the successful execution of the assembly process but also highlight the effectiveness of the QI system in detecting any potential defects or misalignments, contributing to a more efficient and precise manufacturing workflow. Throughout this chapter, an analysis and discussion of the obtained results describe the system's performance, addressing any challenges encountered during the testing phase and proposing potential refinements to optimize both the assembly and QI processes. By critically examining the system's strengths and limitations, this chapter aims to provide valuable insights toward the advancement of automated assembly systems, fostering enhanced accuracy, reliability and quality control in manufacturing processes.

5.1. Assembly system results

The proposed assembly system is evaluated with a focus on assessing its robustness and performance under various light settings. A total of 100 assembly trials are conducted, with each light setting tested 20 times, resulting in a diverse range of lighting conditions achieved by manipulating the lamp's brightness level and controlling the room light environment, as illustrated in Fig. 5.1. Throughout the experiments, a metallic surface workpiece (Fig. 8a) and white pegs are consistently employed to ensure uniformity and consistency in testing conditions. Each assembly trial consisted of one workpiece assembly. Therefore, it was recorded as a successful assembly when all 20 holes of the workpiece was assembled. Hence, a total of 2000 hole assemblies is attempted during the 100 assembly trails. Thus, over the course of the 100 trials, a total of 2000 hole assemblies are attempted. Following each assembly, the workpiece underwent a complete rotation on the machine. A trial's success was contingent upon this rotation proceeding without hindrance from any protruding pegs, in adherence to the definitions of negative and positive defects delineated in Section 4.2. This evaluation aimed to examine the system's ability to operate effectively under real-world scenarios and its performance across varying lighting conditions.

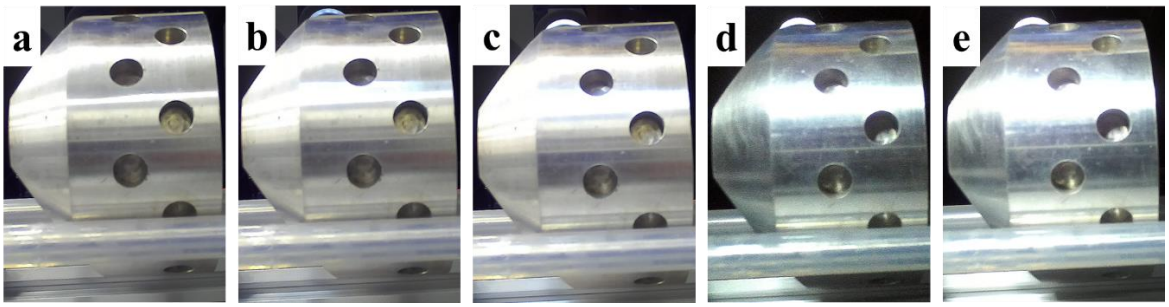

Figure 5.1. Assembly experimental light settings (a) Room. (b) Room + lamp brightness Level 1. (c) Room + lamp brightness Level 2. (d) Lamp brightness Level 1. (e) Lamp brightness Level 2.

The key metrics analyzed include the success rate, representing the proportion of successful assembly trials, and the assembly rate, measuring the average time required for a single workpiece assembly. The results, presented in Table 5.1, demonstrate a high level of robustness and reliability in the system's operation, with consistent success rates and assembly times across all light settings. The overall success rate achieved 99%, signifying the system's capability to consistently achieve accurate and reliable assembly under different lighting conditions. Additionally, the system exhibited an average assembly rate of 386.00 seconds per workpiece, highlighting its efficiency in completing the assembly process. These findings underscore the system's potential for industrial applications, offering a reliable and efficient solution for automated assembly processes with good success rates and assembly rate. In conclusion, the experimental evaluation presents crucial insights into the system's performance, validating its efficacy and suitability for practical implementation in diverse manufacturing settings.

Table 5.1. Assembly system results on the light settings (Fig. 5.1) with 20 trials per setting. The successful completion of an assembly trial is determined when all 20 holes on the workpiece are filled with the pegs.

Light setting	**Success rate (%)**	**Assembly rate (sec/workpiece)**
Room	100.00	384.00
Room + lamp brightness Level 1	100.00	387.00
Room + lamp brightness Level 2	95.00	386.00
Lamp brightness Level 1	100.00	385.00
Lamp brightness Level 2	100.00	386.00
Average	99.00	386.00

Additional experiments are conducted to assess the adaptability of the assembly system using different types of workpieces, as shown in Fig. 5.2. The first variation involves a metallic workpiece wrapped in black vinyl film (Fig. 5.2b). To ensure the tolerance of the holes and pegs, the black vinyl film is removed around the holes. The second variation comprises a workpiece 3D printed using blue PLA filament (Fig. 5.2c). Each type undergoes 20 experiments under identical light settings of room + lamp brightness level 2, as shown in Fig. 5.1c. Similar to previous experiments on varied lighting conditions, a trial's success is contingent upon the complete insertion of pegs into all 20 holes of the workpiece without any interference during a full machine rotation. The results, presented in Table 5.2, indicate a success rate of 90% and above for all workpiece types. However, there are notable differences in the assembly rate among the variations,

with the metallic surface exhibiting the fastest rate of 386.00 seconds/workpiece, followed by the surface wrapped in black vinyl film at 420.00 seconds/workpiece, and the 3D printed workpiece at 493.00 seconds/workpiece.

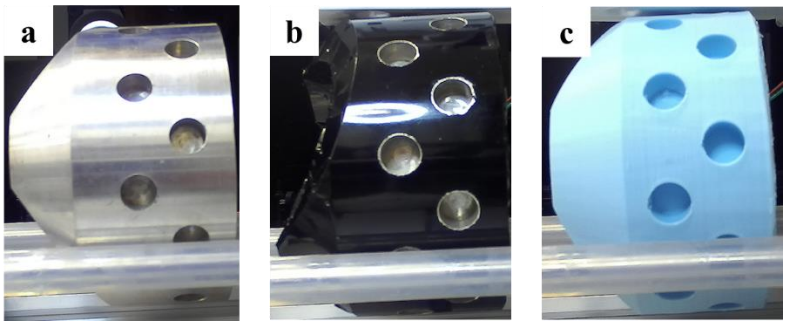

Figure 5.2. Experimented workpiece types (a) Metallic surface. (b) Metal surface wrapped in black vinyl film. (c) 3D printed with blue PLA plastic filament.

Table 5.2. Assembly system results on the workpiece types (Fig. 5.2) with 20 trials per type. The successful completion of an assembly trial is determined when all 20 holes on the workpiece are filled with the pegs.

Workpiece type	**Success rate (%)**	**Assembly rate (sec/workpiece)**
Metallic surface	100.00	386.00
Metal wrapped in black vinyl film	95.00	420.00
3D printed with blue PLA plastic	90.00	493.00

Throughout all experiments, no programming errors are observed, demonstrating the system's stability and reliability. The system proved capable of continuous operation for more than two

consecutive days without requiring any system restarts, further affirming its robustness and efficiency in an industrial setting.

5.2. Assembly system discussions

The proposed assembly system is tailored to efficiently assemble metallic surface workpieces with fixed characteristics, such as cylindrical geometry and specific hole and peg dimensions and colours. However, the system's adaptability to different workpiece properties is evident from the results obtained. Although optimized for metallic surfaces, the system can effectively detect workpieces with diverse characteristics and accurately extract hole locations using image processing techniques. The system has demonstrated high success rates when assembling variation types, including surfaces wrapped in black vinyl film and 3D printed workpieces with blue PLA plastic filament. However, the optimization impact is observable in the assembly rate, with the metallic surface achieving the fastest rate, followed by the surface wrapped in black vinyl film and the 3D printed workpiece. The variation in assembly rates primarily results from the additional time taken in the hole detection process for different workpiece types.

Adapting the assembly system to new workpieces with varying diameters or a different number of holes is straightforward. This process involves adjusting specific parameters, including the number of holes, workpiece diameter, side-view camera target point and tool height. Modifications to the peg insertion tool, including adjustments to its dimensions and circle detection parameters, are required if the dimensions of the hole and peg change. Additionally, changes in workpiece material properties necessitate adjustments to image processing algorithms. When dealing with different surface geometries, the design of a new fixture and re-design of the machine

positioning algorithm are necessary. The presented prototype is designed for components with dimensions not exceeding 300×180×120 mm. For larger components, the system can be adapted by utilizing a machine with a larger work area. The versatility of the assembly system allows it to accommodate varying workpiece characteristics and dimensions, making it an adaptable and efficient solution for a wide range of manufacturing needs.

The proposed assembly system is specifically designed to efficiently assemble metallic surface workpieces with fixed characteristics, including cylindrical geometry and specific hole and peg dimensions and colours. However, the system's adaptability to different workpiece properties is evident from the obtained results. Despite being optimized for metallic surfaces, the system effectively detects workpieces with diverse characteristics and accurately extracts hole locations using image processing techniques. The system demonstrates high success rates in assembling variation types, encompassing surfaces wrapped in black vinyl film and 3D printed workpieces with blue PLA plastic filament. However, the optimization impact is observable in the assembly rate, with the metallic surface achieving the fastest rate, followed by the surface wrapped in black vinyl film and the 3D printed workpiece. The variation in assembly rates primarily results from the additional time taken in the hole detection process for different workpiece types.

In the event of introducing new workpieces with different diameters or varying numbers of holes, the presented assembly system can be easily adapted for automated assembly. These modifications involve adjusting specific parameters, including the number of holes, workpiece diameter, side-view camera target point and tool height. Additionally, alterations to the peg insertion tool, such as adjusting its dimensions and circle detection parameters, are necessary when the dimensions of the hole and peg change. Furthermore, changes in workpiece material properties

require adjustments to the image processing algorithms. In cases of different surface geometries, the design of a new fixture and a re-design of the machine positioning algorithm become necessary. It is worth noting that the presented prototype is specifically designed for components with dimensions not exceeding 300×180×120 mm. For larger components, the system's adaptability can be extended by utilizing a machine with a larger work area. The assembly system's versatility enables it to accommodate a wide range of workpiece characteristics and dimensions, establishing it as an adaptable and efficient solution for diverse manufacturing needs.

5.3. QI results

The ML models employed in this study are trained using a dataset comprising 2503 hole images exclusively obtained from the metallic surface similar to images shown in Fig. 5.3a. Out of these 2503 images, 500 are reserved as a validation dataset, of which 285 belong to the negative defect class and 215 to the positive defect class. The training dataset comprises the remaining 2003 hole images, with 1141 from the negative defect class and 862 from the positive defect class. The negative defect class is split evenly between images of black peg insertions and white peg insertions. The criteria for selecting negative defect hole images in the dataset aligns with the approach discussed in Section 4.2, ensuring that they do not impede the workpiece's rotation within the system. The training process is conducted three times for each architecture, with distinct test datasets utilized for evaluation. The test datasets are divided into three categories, corresponding to different workpiece types: metallic surface, surface wrapped in black vinyl film and blue PLA plastic surface, as illustrated in Fig. 5.3. Each test dataset encompassed 615 hole images, prepared using the hole-cropping algorithm as elaborated in Section 4.1. Consistent light settings,

characterized by room + lamp brightness level 2 in Fig. 5.1, including the direction of the lamp light, are maintained throughout the testing phase.

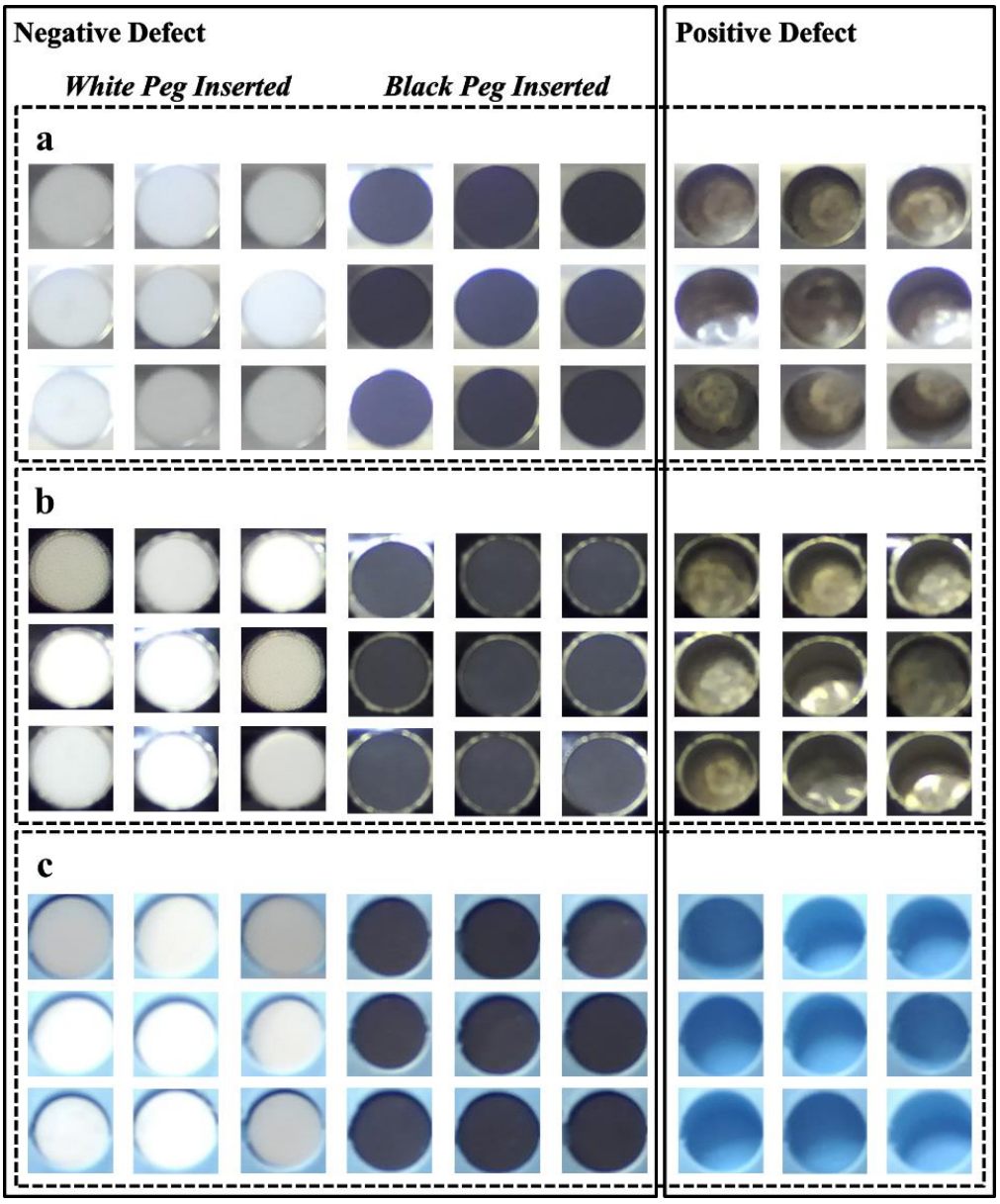

Figure 5.3. Hole image sample for QI testing. Samples are extracted from the following workpiece types. (a) Metallic surface. (b) Metal surface wrapped in black vinyl film. (c) 3D printed with blue PLA plastic filament.

It is important to note that no identical hole images are employed in both the training or test datasets, nor in the images used to calculate the $\sigma_{threshold}$ in Section 4.2. For the training of the ResNet50 model, a batch size of 64 and a learning rate of 1.0×10^{-3} are adopted for 128 epochs. The traditional CNN utilized the same batch size and number of epochs, but with a decaying learning rate with a factor of 0.2, patience set at 3 and a minimum learning rate of 1.0×10^{-4}. The training histories of the models are depicted in Fig. 5.4 for the metallic surface, Fig. 5.5 for the metal surface wrapped in black vinyl film, and Fig. 5.6 for the 3D printed with blue PLA plastic filament workpiece types, respectively. The final test accuracies recorded upon completion of model training during cross-validation are used to represent the accuracy of each ML model. To derive statistical test accuracies, each test dataset is inputted into the program separately, which systematically analyzed each hole image and generated the prediction results for further evaluation.

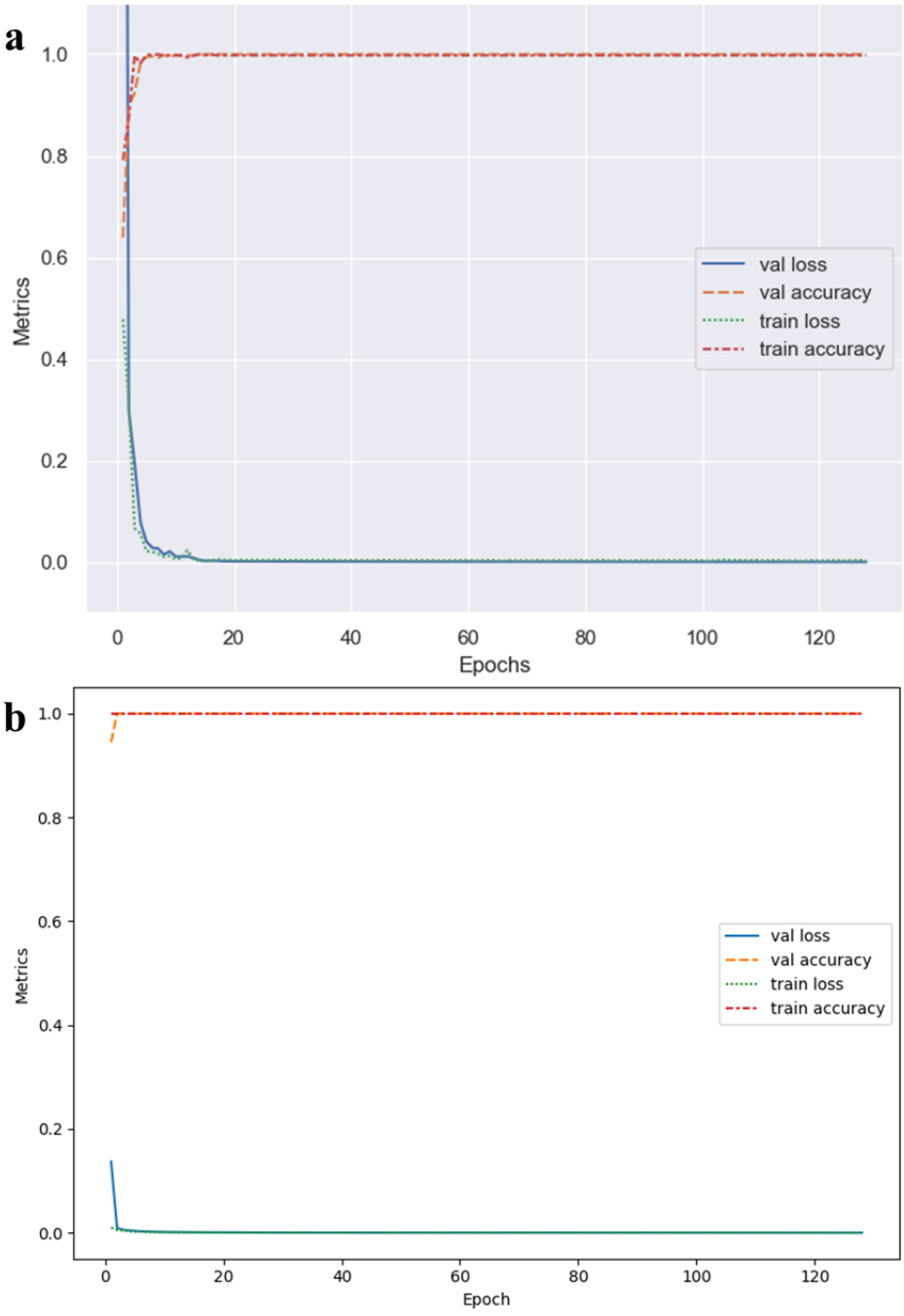

Figure 5.4. QI testing results on hole images from the metallic surface workpiece (Fig. 5.3a) using different ML models. (a) Traditional CNN. (b) ResNet.

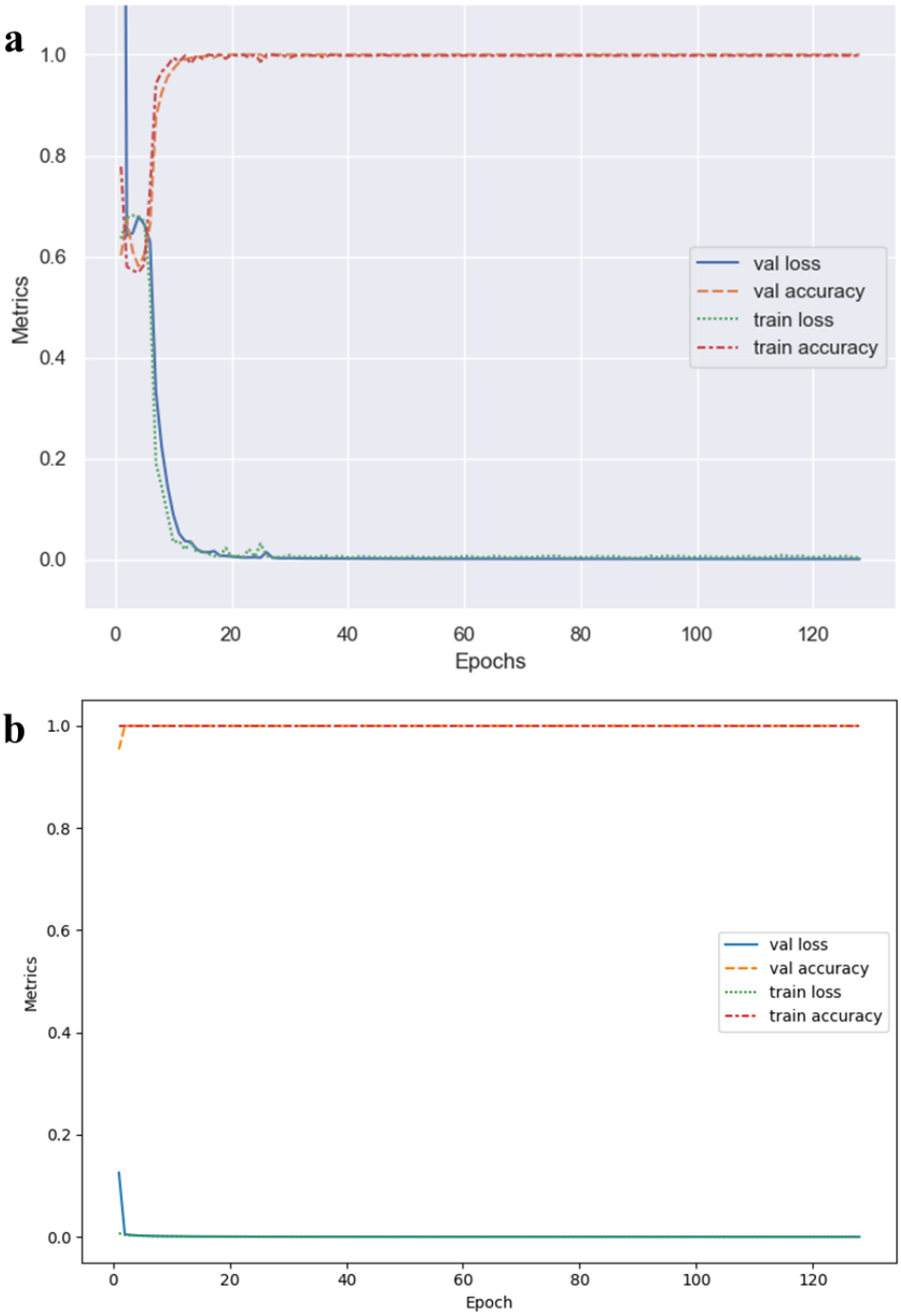

Figure 5.5. QI testing results on hole images from the metal surface wrapped in black vinyl film workpiece (Fig. 5.3b) using different ML models. (a) Traditional CNN. (b) ResNet.

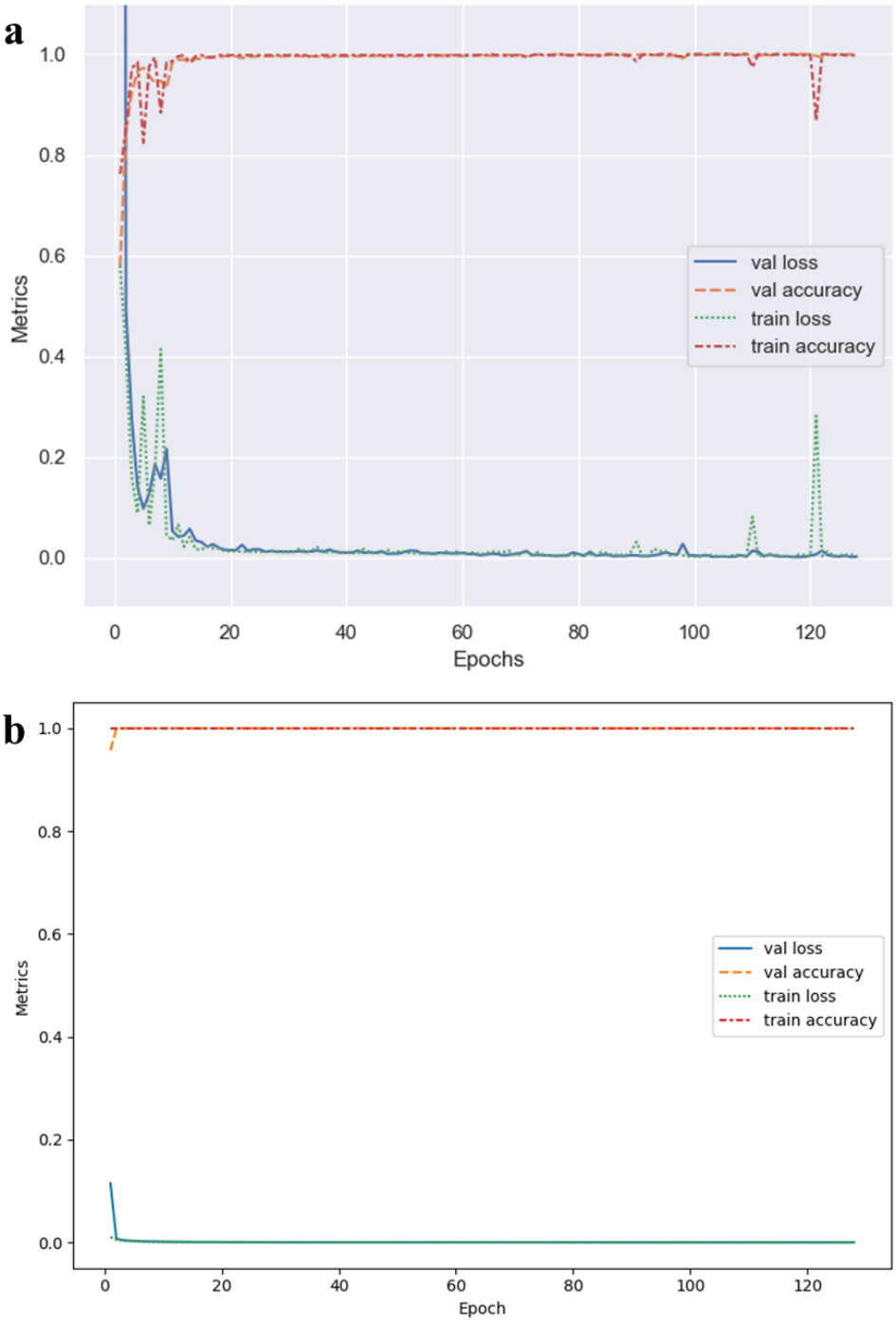

Figure 5.6. QI testing results on hole images from the 3D printed with blue PLA plastic filament workpiece (Fig. 5.3c) using different ML models. (a) Traditional CNN. (b) ResNet.

Table 5.3 presents the experimental accuracy of the three QI methods discussed in this study. Among them, the statistical method achieves the highest average experimental accuracy at 97.02%. This outperforms the ResNet50 model, which achieves 93.98%, and the traditional CNN model at 89.65%. All methods demonstrate a perfect accuracy of 100% when tested on the metallic surface. However, when tested on the black vinyl-wrapped surface workpiece, there is a decline in accuracy for each method, though all still exceed 85%. The most significant drop in accuracy is observed when testing on a blue PLA plastic surface. Here, the ResNet50 and traditional CNN models yield test accuracies of 82.28% and 79.19% respectively, marking a decrease of 17.72% and 20.81% from their performance on metallic surfaces. In contrast, the statistical method experiences a more modest drop in accuracy, with a decline of only 8.13% when compared to its performance on the metallic surface.

Table 5.3. Performance of three QI methods for each workpiece type (Fig. 5.3) with 615 hole images per each type test. Identical hole images are used for each QI method.

QI method	**Test accuracy (%)**
Metallic surface	
Statistical	100.00
ResNet50	100.00
Traditional CNN	100.00
Black vinyl wrapped surface	
Statistical	99.19
ResNet50	99.67
Traditional CNN	89.76
Blue PLA plastic surface	
Statistical	91.87
ResNet50	82.28
Traditional CNN	79.19
Average	
Statistical	97.02
ResNet50	93.98
Traditional CNN	89.65

To further evaluate the statistical method, sensitivity analyses are conducted. For these tests, the same dataset of 615 hole images from the metallic surface workpiece (Figure 5.3a) is used, as in the previous QI assessments. The $\sigma_{threshold}$ incrementally adjusts from 0 ($\sigma_{threshold}$ = 0.89) to ± 50% ($\sigma_{threshold}$ = 0.45, $\sigma_{threshold}$ = 1.33), with accuracies logged based on this dataset of 615 hole images (refer to Table 5.4). In all instances, test accuracy surpasses 90%.

Table 5.4. $\sigma_{threshold}$ adjustment sensitivity analysis test results on the statistical method with 615 hole images per each adjustment.

$\sigma_{threshold}$ [%]	**Test accuracy (%)**	$\sigma_{threshold}$ (%)	**Test accuracy (%)**
0.93 [+5%]	100.00	**0.84 [-5%]**	100.00
0.97 [+10%]	100.00	**0.79 [-10%]**	100.00
1.02 [+15%]	99.67	**0.75 [-15%]**	100.00
1.06 [+20%]	99.67	**0.71 [-20%]**	100.00
1.10 [+25%]	99.02	**0.66 [-25%]**	100.00
1.33 [+50%]	97.56	**0.44 [-50%]**	91.38

Moreover, testing occurs under various lighting conditions illustrated in Fig. 5.1, excluding the room setting (Fig. 5.1a) due to its lack of the angled light source, the lamp. A fresh dataset, also containing 615 hole images but representing each lighting condition, is used for this purpose. Results are documented in Table 5.5, and consistently, the statistical method yields an accuracy rate exceeding 90% across all settings.

Table 5.5. Different light setting (Fig. 5.1) sensitivity analysis test results on the statistical method with 615 hole images per each setting.

Light setting	**Test accuracy (%)**
Room + lamp brightness Level 1	100.00
Room + lamp brightness Level 2	100.00
Lamp brightness Level 1	93.50
Lamp brightness Level 2	92.68

5.4. QI discussions

The study's findings indicate that the statistical method consistently performs well across different surface conditions, whereas the traditional CNN and ResNet50 architectures used exhibit reduced adaptability. The QI methods evaluated in this study are optimized mainly for inspecting metallic workpieces and black and white pegs, taking into account consistent characteristics in terms of diameter, height, surface and material properties. Given this focus, hole images exclusively from metallic surfaces are utilized both for calculating the $\sigma_{threshold}$ of the statistical method and for training the ML models. As a result, all methods demonstrate high test accuracies on both metallic and black vinyl-wrapped surfaces due to the similarities in hole characteristics. Nonetheless, there is a discernible decline in accuracy for the black vinyl-wrapped surfaces, which can be attributed to their differences compared to the purely metallic surfaces used for model training and $\sigma_{threshold}$ calculation.

When it came to assessing the blue PLA plastic surface, which has unique colour, material and reflective properties, all methods exhibited a marked reduction in accuracy. Yet, the statistical

methods maintained relatively high-performance levels. In contrast, the traditional CNN and ResNet50 models saw a more pronounced decrease in accuracy. This reduction might stem from overfitting, given the vast number of training parameters and the specificity of the dataset, which primarily consists of hole images. Notably, the statistical method, being devoid of ML training intricacies, successfully avoids such overfitting issues. Furthermore, the statistical method consistently showcases test accuracies above 90% in all sensitivity tests, underscoring its robustness.

In conclusion, our experimental findings underscore that the statistical method outperforms both the ResNet50 and traditional CNN models in terms of average test accuracy. This finding suggests that, within this specific context, the statistical method offers a more efficient and effective solution. Unlike its counterparts, it eliminates the need for model training, sidestepping the inherent complexities that often accompany ML-based methods. Moreover, the statistical method overcomes challenges related to training, such as overfitting, hyperparameter optimization and data overload [75, 76], making it a well-balanced approach in terms of accuracy and adaptability.

However, the statistical method does have some limitations. First, obtaining several sample data is necessary to calculate the $\sigma_{threshold}$, and the inspection accuracy heavily depends on it. Additionally, finding the $\sigma_{threshold}$ can be challenging if the pegs and workpiece have similar reflective properties. Second, a consistent directional light source, as shown by the angled lamp in Fig. 3.3, must be provided to create shade gradients inside the holes, which is crucial for defect classification.

5.5. Summary

In this chapter, a detailed exploration of the assembly system's capabilities and adaptability is presented. Initially tailored for metallic workpieces, the system demonstrates commendable versatility by efficiently handling different workpiece variations, such as those wrapped in black vinyl film and those 3D printed with blue PLA filament. Notably, the rate of assembly shows variability based on workpiece type. An essential feature of the system's robustness is its uninterrupted operation capacity, running for over two consecutive days without the necessity of restarts and maintaining an error-free programmatic operation.

For the QI methods, a deep dive into a statistical method that uses a series of image processing techniques, and ML models, specifically the traditional CNN and ResNet50, is undertaken. These methods are rigorously trained with a comprehensive set of hole images from metallic surfaces. Among the methods evaluated, the statistical QI approach stands out, consistently outperforming the ML models across different workpiece surfaces. This suggests the superior adaptability of the statistical method, especially when the training dataset is specialized, as is the case in this study. However, it is also worth noting that the statistical method comes with its inherent challenges, particularly its dependency on substantial sample sizes for $\sigma_{threshold}$ calculation and its reliance on consistent lighting.

In conclusion, this chapter highlights the system's impressive adaptability across varied workpiece types. When compared to its ML peers, the statistical QI method emerges as a more balanced solution, offering superior accuracy and adaptability, despite certain limitations.

Chapter 6. Conclusions

This thesis aimed to develop an automated assembly and vision-based QI system for industrial peg-in-hole applications. The research focused on designing and implementing a lab-scale prototype, as well as exploring the feasibility and effectiveness of various assembly techniques and image processing. Throughout the study, several key objectives were achieved, and important findings obtained.

The development of an efficient image processing algorithm lies at the core of the automated peg-in-hole assembly system. In this regard, the algorithm is designed to capture images from a side-view camera and a top-view camera, serving as valuable inputs for the assembly process. Utilizing a custom trained YOLOv5 object detection model, the images obtained from the side-view camera underwent thorough processing to detect the workpiece and accurately determine its position. To optimize computational efficiency, the side-view camera image was cropped, eliminating redundant background information and enabling focused analysis on the workpiece region. Additionally, the Canny edge detection algorithm was employed to calculate the gradient magnitude and direction of the captured images from both cameras. This facilitated the identification of critical features such as edges and circles that corresponded to the holes in the workpiece.

By successfully implementing the developed image processing algorithm, the feasibility of the automated peg-in-hole assembly system was demonstrated through the construction of a prototype. The integration of key components, including the peg insertion tool, three-axis machine and rotary roller, played a crucial role in achieving precise positioning and seamless assembly of

the workpiece. The system exhibited exceptional capabilities in hole detection, peg alignment, glue application and peg launching, effectively completing the assembly process with the average success rate of 99 and 95% at different light settings and different workpiece types, respectively. These results reinforce the potential and viability of the automated peg-in-hole assembly system in industrial applications, providing a strong foundation for further research and development in the field.

The vision-based image processing methods discussed in this thesis aim to address the classification of positive and negative defects in assemblies produced by the automated assembly system introduced earlier. Three distinct methods are presented, all of which rely on the proposed hole image extraction process to obtain the necessary input hole images. The hole image extraction process utilizes a combination of conventional image processing techniques, such as the Canny edge detector and Hough circle transform, to detect and extract circular features. To mitigate the occurrence of false positives, which are commonly encountered in conventional image processing techniques, additional conditions such as boundary box and aspect ratio are introduced. These conditions successfully generated cropped hole images containing only true positive defects, which were subsequently fed into the QI model for classification.

Among the three methods proposed in this thesis, the statistical model represents a novel approach that builds upon conventional image processing techniques. Initially, the input image undergoes a series of preprocessing steps, including noise removal using the Sobel filter, grayscale conversion and extraction of pixel values along the centre column. Leveraging the observation that positive defects exhibit greater variations in pixel values due to light reflection and shading, the statistical method calculates a numerical representation of such behaviour by computing the

gradient of the grayscale values along the centre column and subsequently calculating its standard deviation. Through an iterative process involving 300 trials, a threshold value is determined, which is then used to compare against the standard deviation value for prediction. The statistical method achieved an average test accuracy of 97.02%, showcasing its effectiveness in defect classification.

Meanwhile, the traditional CNN and ResNet methods represent ML-based approaches for QI. These architectures were selected due to their wide usage and proven effectiveness in image classification tasks. Both models were trained on a dataset comprising 2503 hole images, encompassing both positive and negative defects. The training process involved optimizing the model parameters to minimize the classification error and maximize the accuracy of defect identification. The traditional CNN and ResNet models achieved average test accuracies of 95.74% and 92.85%, respectively.

By incorporating these vision-based image processing methods into the automated assembly system, a comprehensive framework for assembly and defect detection can be established. The combination of statistical analysis and traditional CNN and ResNet models provides a diverse range of approaches to cater to various scenarios and requirements. The achieved results highlight the potential of these methods in accurately identifying and classifying defects, thereby enhancing the overall quality assurance of the automated assembly process.

6.1. Contributions

This research work has made several significant contributions to the field of automated peg-in-hole assembly and QI. The key contributions of this thesis are summarized as follows:

i. *Development of a novel automated peg-in-hole assembly system*

The system integrated several key components, including a side-view camera, a top-view camera, a custom trained YOLOv5 object detection model, and a feature detection algorithm. Through the integration of these components, the system achieved the automation of the assembly process, enabling precise positioning and assembly of the workpiece. The system effectively detected the holes, aligned the pegs and applied glue before launching the pegs into the corresponding holes, resulting in successful assembly. Notably, the presented peg insertion tool introduced a novel concept by incorporating peg storage and the sequential launch of one peg at a time. This approach eliminated the need for conventional pick-and-place mechanisms, enhancing the efficiency and reliability of the assembly process.

ii. Hybrid image processing techniques for effectively executing the hole assemblies

This thesis makes significant contributions to the advancement of image processing techniques for automated assembly systems. The developed hole image extraction process effectively utilizes conventional techniques, such as the Canny edge detector and Hough circle transform, to detect and extract hole features. To enhance the accuracy of defect classification, additional conditions are introduced to filter out false positives, resulting in the generation of cropped hole images that accurately represent true positives. These enhancements improve the overall effectiveness and reliability of the defect classification system, further contributing to the advancement of automated assembly processes.

iii. Proposed vision-based QI methods

This thesis presents three distinct methods for classifying positive and negative defects in assembly products. The statistical model introduces a novel approach to defect classification by

leveraging conventional image processing techniques. The traditional CNN and ResNet models provide ML-based approaches for QI. Through rigorous training and testing, these methods demonstrate promising results, highlighting their potential for accurate defect identification. The statistical model utilizes gradient and standard deviation calculations to capture pixel value variations, achieving an average test accuracy of 97.02%. The CNN and ResNet models showcase their effectiveness in defect classification, contributing to the advancement of vision-based defect detection in automated assembly systems.

6.2. Limitations and assumptions

Though this research work has made significant progress in the development of the automated peg-in-hole assembly system and the vision-based defect classification methods, it is important to acknowledge certain limitations and assumptions inherent in this study.

The workpieces and pegs are assumed to be defect-free components, adhering to specified dimensions within a tolerance of ±0.1 mm. The surface finish across all workpieces and pegs is also presumed to be consistent. Environmental conditions such as temperature and humidity during the assembly operations are held constant and fall within typical industrial parameters, ensuring no external influence on the material properties or the assembly process itself. Additionally, potential external disturbances like vibrations that could affect the peg-in-hole assembly process are not accounted for in this study.

The assembly system is designed to accommodate various workpiece variations. For optimal operation, specific parameters, such as the number of holes per workpiece, the number of rows of holes, the displacement between rows, and any patterns associated with hole positions, must be adjusted in the program to reflect these variations. In its present configuration, the system can

manage workpieces with a maximum diameter of 125 mm and a length of up to 150 mm. To facilitate larger workpieces, there would be a need to expand the machine's dimensions. Additionally, if pegs of different dimensions are to be used, the peg insertion tool must undergo modifications. As it stands, the tool is specially designed to cater to the peg of a specific size. Any deviation from this size would necessitate design adjustments to ensure seamless operation and accurate assembly.

The performance of the image processing algorithms and defect classification methods discussed in this research is dependent on the availability of a minimum light source. Adequate illumination is essential to ensure clear and high-quality image capture, which is crucial for accurate detection of workpieces, holes and defects. Insufficient lighting conditions can lead to reduced image clarity, increased noise and inaccurate identification of assembly components. Therefore, it is recommended to maintain a minimum light source level that meets the system’s requirements to ensure optimal performance and reliable results. Future researchers can explore techniques to enhance the system’s adaptability to varying lighting conditions and develop algorithms that can compensate for suboptimal illumination scenarios.

The developed automated assembly system and the proposed defect classification methods are implemented and evaluated using a prototype. The performance and robustness of the system may vary when scaled up to an industrial-level machine. Further research and development are required to address the challenges associated with scaling up the system, including the integration of industrial motors, enhancing the reloader’s capacity and organizing the electrical wiring for improved robustness, reliability and safety.

The defect classification models assume that all defects can be accurately represented and identified through the proposed image processing and ML techniques. However, certain complex defects or variations in the assembly process may not be adequately captured or classified by the current models. For instance, no data on partially inserted pegs are provided, hence the models are classifying between fully inserted pegs and completely empty holes. The performance of the defect classification methods may be affected by the complexity and variability of defects encountered in real-world assembly scenarios.

Despite the efforts made to improve computational efficiency by implementing techniques such as cropping the side-view camera image and employing optimized algorithms, there are potential areas for further improvement. This study did not include specific metrics to quantitatively measure the extent of computational efficiency enhancements achieved, hindering a comprehensive evaluation of the system's performance. Therefore, it is essential to explore and develop numerically comparable metrics to accurately assess the improvement in computational efficiency. Additionally, it is crucial to address the processing speed and computational requirements of the image processing and defect classification methods to ensure performance in high-speed assembly environments. Furthermore, considerations should be given to prevent computer overheating because prolonged operation under demanding conditions may pose risks. Future researchers should focus on optimizing these aspects, taking into account both efficiency and system reliability, to enhance the overall effectiveness and practicality of the automated peg-in-hole assembly system in industrial settings.

By recognizing these limitations and assumptions, future researchers can build upon this work and address these challenges to advance the field of automated assembly and QI. These

considerations provide valuable insights into the practical implementation and further development of the proposed system and methodologies.

6.3. Future work

The focus of this thesis has been the development of a lab-scale prototype for automated peg-in-hole assembly, aimed at industrial applications. However, to achieve the ultimate goal of an industrial-scale machine capable of robust, complete and reliable long hours of operation to meet high demands, several tasks need to be accomplished.

i. QI system continuous development

As discussed in 0, CV techniques, particularly ML-based methods, remain at the forefront of extensive research across various institutions. This dynamic and rapidly evolving field continually introduces improved models each year, driven by advancements in deep learning, neural networks and data augmentation methodologies. It is worth noting that the ML architectures utilized for QI in this study represent fundamental models, and it is evident that numerous advanced models have already surpassed their performance. Consequently, a significant aspect of our future work will focus on exploring and implementing these cutting-edge techniques to enhance the QI system's capabilities.

ii. Integrated assembly and QI system

Although both the assembly and QI systems have been developed independently as presented, the integration of these two tasks into a single system remains pending. Additional algorithms are required to enable simultaneous assembly and QI, including a mechanism for counting the number of assembled holes to trigger the appropriate sequence for QI.

iii. *Production of an industrial-scale system*

To fulfill the project's scope of an industrial-scale machine, suitable for installation and implementation at a factory level, the lab-scale prototype must be upgraded.

Key subtasks involve:

a. Production of a metallic peg insertion tool for improved robustness and reliability.
b. Integration with a robotic arm and gripper to enable automatic placement of the workpiece on the assembly machine, facilitating a fully automated assembly line.
c. Replacement of motors with industrial-grade versions to enhance robustness and reliability.
d. Increasing the funnel capacity of the reloader to extend continuous operational time.
e. Addition of a gauge sensor to the reloader for automatic pausing and resuming of operations based on button fill status.
f. Organization of electrical wires into an electrical panel, enhancing robustness, reliability and safety.

iv. *Validation and optimization of the industrial-scale system*

Upon completion of the upgrades and production of the industrial-scale machine, validation and optimization steps must be undertaken. This includes rigorous testing of the system's performance, assessing its robustness and reliability under various operating conditions, and fine-tuning the algorithms and parameters to maximize efficiency and accuracy.

By addressing these future work areas, the development of an industrial-scale automated peg-in-hole assembly system will be advanced, providing a highly robust, reliable and efficient solution for meeting industrial production demands.

References

[1] L. Zhao, T. R. Dunne, J. Ren and P. Cheng, "Dissolvable Magnesium Alloys in Oil and Gas Industry," in *Magnesium Alloys - Processing, Potential and Applications*, Rijeka, IntechOpen, 2023.

[2] J. Su, R. Li, H. Qiao, J. Xu, Q. Ai and J. Zhu, "Study on dual peg-in-hole insertion using of constraints formed in the environment," *Industrial Robot: An International Journal*, vol. 44, no. 6, pp. 730-740, 2017.

[3] H. Park, J. Park, D.-H. Lee, J.-H. Park, M.-H. Baeg and J.-H. Bae, "Compliance-Based Robotic Peg-in-Hole Assembly Strategy Without Force Feedback," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 8, pp. 6299-6309, 2017.

[4] K. Tan, N. Watanabe and Y. Iwahori, "X-ray radiography and micro-computed tomography examination of damage characteristics in stitched composites subjected to impact loading," *Composites Part B: Engineering*, vol. 42, no. 4, 2011.

[5] S. Son, H. Park and K. H. Lee, "Automated laser scanning system for reverse engineering and inspection," *International Journal of Machine Tools and Manufacture, Volume 42, Issue 8*, pp. 889-897, 2002.

[6] J. Beyerer, F. P. León and C. Frese, Machine Vision: Automated Visual Inspection: Theory, Practice and Applications, Springer Berlin Heidelberg, 2015.

[7] Health Canada, "Radiation Protection and Safety for Industrial X-Ray Equipment," authority of the Minister of Health, Ottawa, 2003.

[8] M. Nigro, M. Sileo, F. Pierri, K. Genovese, D. D. Bloisi and F. Caccavale, "Peg-in-Hole Using 3D Workpiece Reconstruction and CNN-based Hole Detection," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, 2020.

[9] S. Wang, G. Chen, H. Xu and Z. Wang, "A Robotic Peg-in-Hole Assembly Strategy Based on Variable Compliance Center," *IEEE Access,* vol. 7, no. 7, pp. 167534-167546, 2019.

[10] R. K. Jain, S. Saha and S. Majumder, "Development of piezoelectric actuator based compliant micro gripper for robotic peg-in-hole assembly," in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, 2013.

[11] S. K. Sinha and P. W. Fieguth, "Automated detection of cracks in buried concrete pipe images," *Automation in Construction,* vol. 15, no. 1, pp. 58-72, 2006.

[12] X. Zheng, J. Chen, H. Wang, S. Zheng and Y. Kong, "A deep learning-based approach for the automated surface inspection of copper clad laminate images," *Applied Intelligence,* vol. 51, no. 3, p. 1262–1279, 2021.

[13] S. A. Singh and K. A. Desai, "Automated surface defect detection framework using machine vision and convolutional neural networks," *Journal of Intelligent Manufacturing,* vol. 34, no. 4, p. 1995–2011, 2023.

[14] F. Wei, G. Yao, Y. Yang and Y. Sun, "Instance-level recognition and quantification for concrete surface bughole based on deep learning," *Automation in Construction*, vol. 107, 2019.

[15] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[16] D. Wang and Y. Shang, "A new active labeling method for deep learning," in *2014 International Joint Conference on Neural Networks (IJCNN)*, 2014.

[17] N. Hogan, "Impedance Control: An Approach to Manipulation: Part II—Implementation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 8-16, 1985.

[18] J. Song, Q. Chen and Z. Li, "A peg-in-hole robot assembly system based on Gauss mixture model," *Robotics and Computer-Integrated Manufacturing*, vol. 67, 2021.

[19] W. Deng, C. Zhang, Z. Zou, M. Gao, X. Wang and W. Yu, "Peg-in-hole assembly of industrial robots based on Object detection and Admittance force control," in *2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2021.

[20] N. Waltham, M. C. E. Huber, A. Pauluhn, J. L. Culhane, J. G. Timothy, K. Wilhelm and A. Zehnder, "CCD and CMOS sensors," in *Observing Photons in Space: A Guide to Experimental Space Astronomy*, New York, NY, Springer, New York, NY, 2013, pp. 423-442.

[21] K. C. P. Wang and X. Li, "Use of Digital Cameras for Pavement Surface Distress Survey," *Transportation Research Record,* vol. 1675, no. 1, pp. 91-97, 1999.

[22] N. M. Tuan, Y. Kim, J.-Y. Lee and S. Chin, "Automatic Stereo Vision-Based Inspection System for Particle Shape Analysis of Coarse Aggregates," *Journal of Computing in Civil Engineering,* vol. 36, no. 2, 2022.

[23] A. Belhaoua, S. Kohler and E. Hirsch, "Estimation of 3D reconstruction errors in a stereo-vision system," in *Proc. SPIE 7390, Modeling Aspects in Optical Metrology II, 73900X*, 2009.

[24] K. Schmid, T. Tomic, F. Ruess, H. Hirschmüller and M. Suppa, "Stereo vision based indoor/outdoor navigation for flying robots," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[25] J. P. Kowalski, J. Peksinski and G. Mikolajczak, "Detection of Noise in Digital Images by Using the Averaging Filter Name COV," in *Intelligent Information and Database Systems*, Berlin, Heidelberg, 2013.

[26] A. Jain and R. Gupta, "Gaussian filter threshold modulation for filtering flat and texture area of an image," in *2015 International Conference on Advances in Computer Engineering and Applications*, 2015.

[27] J. Pan, X. Yang, H. Cai and B. Mu, "Image noise smoothing using a modified Kalman filter," *Neurocomputing,* vol. 173, pp. 1625-1629, 2016.

[28] M. Piovoso and P. A. Laplante, "Kalman filter recipes for real-time image processing," *Real-Time Imaging,* vol. 9, no. 6, pp. 433-439, 2003.

[29] Y. Zhang, J. Y. Fuh, D. Ye and G. S. Hong, "In-situ monitoring of laser-based PBF via off-axis vision and image processing approaches," *Additive Manufacturing,* vol. 25, pp. 263-274, 2019.

[30] P. Sahoo, S. Soltani and A. Wong, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing,* vol. 41, no. 2, pp. 233-260, 1988.

[31] V. Baligar, L. Patnaik and G. Nagabhushana, "Low complexity, and high fidelity image compression using fixed threshold method," *Information Sciences,* vol. 176, no. 6, pp. 664-675, 2006.

[32] R. P. Singh and M. Dixit, "Histogram Equalization: A Strong Technique for Image Enhancement," *International Journal of Signal Processing, Image Processing and Pattern Recognition,* vol. 8, no. 8, pp. 345-352, 2015.

[33] P. Roy, S. Dutta, N. Dey, G. Dey, S. Chakraborty and R. Ray, "Adaptive thresholding: A comparative study," in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2014.

[34] D. Ziou and S. Tabbone, "Edge detection techniques-an overview," *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, vol. 8, pp. 537-559, 1998.

[35] Z. Guo, L. Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification," *IEEE Transactions on Image Processing,* vol. 19, no. 6, pp. 1657-1663, 2010.

[36] F. Roberti de Siqueira, W. Robson Schwartz and H. Pedrini, "Multi-scale gray level co-occurrence matrices for texture description," *Neurocomputing,* vol. 120, pp. 336-345, 2013.

[37] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological Cybernetics,* vol. 61, no. 2, pp. 103-113, 1989.

[38] S. Hemalath, U. D. Acharya, A. Renuka and P. R. Kamath, "A Secure Color Image Steganography In Transform Domain," *CoRR,* 2013.

[39] J. Seo, S. Chae, J. Shim, D. Kim, C. Cheong and T.-D. Han, "Fast Contour-Tracing Algorithm Based on a Pixel-Following Method for Image Sensors," *Sensors,* vol. 16, p. 353, 2016.

[40] J. Illingworth and J. Kittler, "A survey of the hough transform," *Computer Vision, Graphics, and Image Processing,* vol. 44, no. 1, pp. 87-116, 1988.

[41] R. K. Sabhara, C.-P. Lee and K.-M. Lim, "Comparative study of hu moments and zernike moments in object recognition," *SmartCR,* vol. 3, no. 3, pp. 166-173, 2013.

[42] Y. Mingqiang, K. Kidiyo and R. Joseph, "A survey of shape feature extraction techniques," *Pattern recognition,* vol. 15, no. 7, pp. 43-90, 2008.

[43] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan and J. Walsh, "Deep Learning vs. Traditional Computer Vision," in *Advances in Computer Vision*, 2020.

[44] 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.

[45] 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*, 2001.

[46] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, 2012.

[47] A. Lohia, K. D. Kadam, R. R. Joshi and A. M. Bongale, "Bibliometric Analysis of One-stage and Two-stage Object Detection," *Library Philosophy and Practice*, pp. 1-32, 2021.

[48] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[49] S. Ren, K. He, R. B. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[50] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2117-2125, 2017.

[51] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[52] S. Lu, B. Wang, H. Wang, L. Chen, M. Linjian and X. Zhang, "A real-time object detection algorithm for video," *Computers & Electrical Engineering*, vol. 77, pp. 398-408, 2019.

[53] M. J. Tsai, H.-W. Lee and N.-J. Ann, "Machine vision based path planning for a robotic golf club head welding system," *Robotics and Computer-Integrated Manufacturing*, vol. 27, no. 4, pp. 843-849, 2011.

[54] B. Solvang, G. Sziebig and P. Korondi, "Vision Based Robot Programming," in *2008 IEEE International Conference on Networking, Sensing and Control*, 2008.

[55] P. R. M. de Araujo and R. G. Lins, "Computer vision system for workpiece referencing in three-axis machining centers," *The International Journal of Advanced Manufacturing Technology*, vol. 106, no. 5, pp. 2007-2020, 2020.

[56] T. Wuest, D. Weimer, C. Irgens and K.-D. Thoben, "Machine learning in manufacturing: advantages, challenges, and applications," *Production & Manufacturing Research*, vol. 4, no. 1, pp. 23-45, 2016.

[57] Z. Jin, Z. Zhang and G. X. Gu, "Autonomous in-situ correction of fused deposition modeling printers using computer vision and deep learning," *Manufacturing Letters*, vol. 22, pp. 11-15, 2019.

[58] Z. Hocenski, S. Vasilic and V. Hocenski, "Improved Canny Edge Detector in Ceramic Tiles Defect Detection," in *IECON 2006 - 32nd Annual Conference on IEEE Industrial Electronics*, 2006.

[59] R. Ren, T. Hung and K. C. Tan, "A Generic Deep-Learning-Based Approach for Automated Surface Inspection," *IEEE Transactions on Cybernetics*, vol. 48, no. 3, pp. 929-940, 2018.

[60] X. Feng, X. Gao and L. Luo, "A ResNet50-Based Method for Classifying Surface Defects in Hot-Rolled Strip Steel," vol. 9, no. 19, 2021.

[61] OpenCV, "Introduction," OpenCV, [Online]. Available: https://docs.opencv.org/4.x/d1/dfb/intro.html. [Accessed 27 June 2023].

[62] Ultralytics, "YOLOv5," Github, [Online]. Available: https://github.com/ultralytics/yolov5. [Accessed 14 September 2023].

[63] A. M. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 38, no. 1, pp. 35-44, 2004.

[64] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vols. PAMI-8, no. 6, pp. 679-698, 1986.

[65] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, p. 11–15, 1972.

[66] OpenCV, "Smoothing Images," OpenCV, [Online]. Available: https://docs.opencv.org/4.x/d4/d13/tutorial_py_filtering.html. [Accessed 27 June 2023].

[67] OpenCV, "Canny Edge Detection," OpenCV, [Online]. Available: https://docs.opencv.org/4.x/da/d22/tutorial_py_canny.html. [Accessed 27 June 2023].

[68] J. Illingworth and J. Kittler, "The Adaptive Hough Transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vols. PAMI-9, no. 5, pp. 690-698, 1987.

[69] H.-C. Lee, E. J. Breneman and C. P. Schulte, "Modeling light reflection for computer color vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 12, no. 12, pp. 402-409, 1990.

[70] O. R. Vincent and O. Folorunso, "A Descriptive Algorithm for Sobel Image Edge Detection," in *Proceedings of Informing Science & IT Education Conference (InSITE) 2009,* 2009.

[71] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010.

[72] W. Lihao and D. Yanni, "A Fault Diagnosis Method of Tread Production Line Based on Convolutional Neural Network," in *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS),* 2018.

[73] P. Wei, C. Liu, M. Liu, Y. Gao and H. Liu, "CNN-based reference comparison method for classifying bare PCB defects," *The Journal of Engineering,* vol. 2018, no. 16, pp. 1528-1533, 2018.

[74] S. Abdallah, W. M. Elmessery, M. Shams and N. Al-Sattary, "Deep learning model based on ResNet-50 for beef quality classification," *Information Sciences Letters,* vol. 12, no. 1, 2023.

[75] C. Gambella, B. Ghaddar and J. Naoum-Sawaya, "Optimization problems for machine learning: A survey," *European Journal of Operational Research*, vol. 290, no. 3, 2021.

[76] A. L'Heureux, K. Grolinger, H. F. Elyamany and M. A. M. Capretz, "Machine Learning With Big Data: Challenges and Approaches," *IEEE Access*, vol. 5, pp. 7776-7797, 2017.

[77] OpenCV, "cv::CLAHE Class Reference," OpenCV, [Online]. Available: https://docs.opencv.org/4.x/d6/db6/classcv_1_1CLAHE.html. [Accessed 27 June 2023].