applied sciences

MDPI

Article

Validating the Use of Smart Glasses in Industrial Quality Control: A Case Study

José Silva [1,*], Pedro Coelho [2], Luzia Saraiva [2], Paulo Vaz [1], Pedro Martins [1] and Alfonso López-Rivero [3]

[1] Research Centre in Digital Services (CISeD), Instituto Politécnico de Viseu, 3504-510 Viseu, Portugal; paulovaz@estgv.ipv.pt (P.V.); pedromom@estgv.ipv.pt (P.M.)
[2] Department of Informatics Engineering, Instituto Politécnico de Viseu, 3504-510 Viseu, Portugal; estgv16312@alunos.estgv.ipv.pt (P.C.); estgv17334@alunos.estgv.ipv.pt (L.S.)
[3] Computer Science Faculty, Pontifical University of Salamanca, 37002 Salamanca, Spain; ajlopezri@upsa.es
* Correspondence: jsilva@estgv.ipv.pt

Abstract: Effective quality control is crucial in industrial manufacturing for influencing efficiency, product dependability, and customer contentment. In the constantly changing landscape of industrial production, conventional inspection methods may fall short, prompting the need for inventive approaches to enhance precision and productivity. In this study, we investigate the application of smart glasses for real-time quality inspection during assembly processes. Our key innovation involves combining smart glasses' video feed with a server-based image recognition system, utilizing the advanced YOLOv8 model for accurate object detection. This integration seamlessly merges mixed reality (MR) with cutting-edge computer vision algorithms, offering immediate visual feedback and significantly enhancing defect detection in terms of both speed and accuracy. Carried out in a controlled environment, our research provides a thorough evaluation of the system's functionality and identifies potential improvements. The findings highlight that MR significantly elevates the efficiency and reliability of traditional inspection methods. The synergy of MR and computer vision opens doors for future advancements in industrial quality control, paving the way for more streamlined and dependable manufacturing ecosystems.

Keywords: quality inspection; smart manufacturing; object detection; computer vision

Citation: Silva, J.; Coelho, P.; Saraiva, L.; Vaz, P.; Martins, P.; López-Rivero, A. Validating the Use of Smart Glasses in Industrial Quality Control: A Case Study. *Appl. Sci.* **2024**, *14*, 1850. https://doi.org/10.3390/app14051850

Academic Editor: Nikolaos Papakostas

Received: 8 January 2024
Revised: 19 February 2024
Accepted: 21 February 2024
Published: 23 February 2024

1. Introduction

Quality control (QC) holds paramount importance in industrial manufacturing, serving as a linchpin for operational efficiency, product reliability, and overall customer satisfaction [1–3]. The intricate nature of manufacturing processes involving diverse components and the critical safety implications of the final product underscores the indispensable role of QC within the industry. Swift and precise quality inspections not only have a profound impact on an organization's financial well-being but also have a significant influence on the broader trajectory of the industry [4].

The belated discovery of defects, especially after the manufacturing process or product release, exacts a heavy toll, encompassing substantial costs such as reworking, recalls, and a detrimental impact on the manufacturer's reputation. Thorough inspections meticulously integrated across all production stages play a pivotal role in mitigating these costs, optimizing both productivity and profitability. Moreover, a robust quality control system functions as a bulwark, instilling confidence in the brand and thereby enhancing sales and fortifying competitiveness in the market [5,6].

Quality control acts as a safeguard against the late-stage identification of defects, safeguarding against financial losses and reputational damage. Its multifaceted benefits not only ensure the reliability and safety of the end product but also contribute significantly to the organizational bottom line, shaping a positive trajectory for the entire manufacturing industry.

As manufacturing environments evolve, traditional inspection methods face challenges in keeping pace with the increasing intricacies of modern products and efficiency demands. Innovative approaches, such as integrating smart glasses and advanced image recognition, become crucial for enhancing precision, speed, and the overall effectiveness of quality control processes. This study delves into such advancements, exploring the potential of mixed reality and cutting-edge computer vision algorithms to propel quality control into a new era of efficiency and reliability.

Traditionally, manual inspection processes relying on human experience and judgment have dominated industry practices [7,8]. However, the intricacy of modern products and efficiency demands have stretched these traditional methods. Manual inspections are time-consuming and error-prone and may struggle to detect subtle or complex defects, leading to inconsistencies among inspectors [9–11].

In the pursuit of enhanced efficiency and precision, industries globally are turning to technology. Mixed reality (MR), characterized by the augmentation of the real world with digital information, emerges as a promising solution to QC challenges. MR facilitates real-time interactive overlays of data in the user's field of view, potentially revolutionizing inspection processes [12].

The integration of MR into quality control processes offers several advantages. Firstly, it boosts efficiency [13–16], allowing inspectors to swiftly identify and assess parts without time-consuming manual checks against plans or specifications. Moreover, MR improves accuracy [16] by leveraging sophisticated imaging and recognition technology to detect defects missed by the human eye.

MR technology introduces the real-time tracking of inspection cycles and tasks, providing insights into manufacturing process efficiency [17,18]. This aids in identifying bottlenecks and improving workflow designs, which are crucial in industries like automotive manufacturing for understanding operational efficiency and productivity [16,19].

Despite these potential benefits, the integration of MR into industrial quality control processes is in its early stages, raising questions about practical feasibility and effectiveness. Motivated by the need for more efficient and reliable quality control methods and inspired by MR's potential to transform industrial inspections, this study explores how MR can be effectively employed in the QC industry.

In the context of this project, assembly quality is the verification that components are in their correct position. This context underlines the critical importance of assembly quality to ensure seamless integration and functionality of a smart glasses-based quality control system. The verification process involves meticulous checks carried out by the system to confirm that each component is placed in the specified positions and meets the standards defined in the project.

The primary objective is to develop an MR-based quality control system using smart glasses integrated with a server-based image recognition system and assess its effectiveness in a controlled environment. This study also aims to evaluate the system's performance and viability under real-world conditions, identifying problems and possibilities for improvement and verifying the feasibility of implementing this MR system on a larger operational scale in the manufacturing industry.

This exploration aims to provide a practical roadmap for adopting MR in industrial quality control and contribute to the understanding and advancement of this technology. The subsequent sections offer an overview of object detection, detail the proposed methodology, present the pilot test results in a controlled environment, and conclude by summarizing essential findings, and emphasizing the significance of the proposed methodology for improving the quality control systems.

The remainder of this paper is organized as follows. The next section provides an overview of object detection. Section 3 details the proposed methodology. Section 4 presents the pilot test carried out in a controlled environment where the intention is to evaluate the performance of the system, identify possible problems and evaluate its scalability.

Section 5 concludes the paper by summarizing the essential findings and highlighting the significance of the proposed methodology in improving quality control system.

2. Object Detection

Object detection plays a central and crucial role in the context of quality control processes enhanced by mixed reality. It serves as the foundation for seamlessly integrating real-world and digital environments, enabling meticulous quality assessment and assurance.

In the realm of mixed reality, which combines elements from both physical reality and virtual environments, there is a need for a robust mechanism to identify, localize, and analyze objects within this merged context. Object detection technologies, often powered by advanced computer vision techniques, play a key role in precisely recognizing and categorizing relevant items of interest.

The effectiveness of quality control processes relies heavily on the accurate and timely identification of specific items, defects, or anomalies. Object detection, positioned at the core of the framework, empowers the system to carry out these functions with precision. Through the real-time detection of objects and the overlay of pertinent digital information, such as annotations or diagnostic insights, the quality control process achieves a heightened level of thoroughness and accuracy.

The evolution of object detection algorithms has progressed significantly over the years, spurred by advancements in deep learning and the demand for more accurate and efficient solutions. A pivotal distinction in this evolution is the categorization of algorithms into two-stage and one-stage approaches [20,21].

2.1. Two-Stage Object Detection Algorithms

The advent of two-stage object detection algorithms marked a significant shift in the field, bringing about transformative changes. This approach introduced a distinct separation between region proposal and object classification stages, addressing computational inefficiencies prevalent in earlier methods [22]. The family of R-CNN (Region-Based Convolutional Neural Network) algorithms, including R-CNN, Fast R-CNN, and Faster R-CNN, played a pivotal role in establishing and refining this innovative approach.

R-CNN, presented by Girshick et al. in 2014 [23], introduced a novel strategy by proposing regions of interest (ROIs) using a selective search and then classifying objects within these regions through a convolutional neural network (CNN). While promising in terms of accuracy, this method required faster processing speeds due to the sequential nature of region proposal and classification. The subsequent refinement, Fast R-CNN, addressed this challenge by merging the region proposal and feature extraction stages, sharing computation across different regions, and enhancing efficiency [24].

The culmination of this two-stage approach occurred with the introduction of Faster R-CNN by Ren et al. in 2015 [25]. This approach seamlessly integrated the region proposal process into the network architecture using a Region Proposal Network (RPN), enabling end-to-end training. Faster R-CNN not only achieved remarkable accuracy but also significantly reduced the processing time, establishing itself as a cornerstone in object detection [26,27].

2.2. One-Stage Object Detection Algorithms

While two-stage algorithms achieved impressive accuracy, the computational overhead of region proposals limited their real-time practicality. In response to this challenge, one-stage object detection algorithms emerged, aiming to perform object classification and localization in a single pass.

This one-stage approach is exemplified by the YOLO (You Only Look Once) family of algorithms and the SSD (Single Shot MultiBox Detector) algorithm. YOLO, introduced by Redmon et al. in 2016 [28], divided the input image into a grid and directly predicted bounding boxes and class probabilities from grid cells. This design significantly reduced the computation time, enabling real-time object detection [29,30]. Although YOLO demonstrated competitive accuracy, its weakness lay in accurately detecting small objects due to the inherent nature of its grid-based architecture.

By contrast, SSD, introduced by Liu et al. in 2016 [31], employed a different strategy. It utilized a set of predefined default bounding boxes with varying aspect ratios and sizes, predicting object classes and offsets for these boxes. This approach allowed SSD to effectively address the challenge of detecting objects at different scales, enhancing its accuracy in small object detection. However, SSD still faces the task of matching the accuracy of two-stage approaches for larger and more complex scenes.

The YOLO series, culminating in YOLOv8 in 2023, has shown significant improvements in accuracy and speed, making it an increasingly competitive solution compared to two-stage methods [32–34].

Figure 1 presents a concise timeline, highlighting the historical dates corresponding to various versions of the YOLO algorithm from its inception to the latest version, YOLOv8.

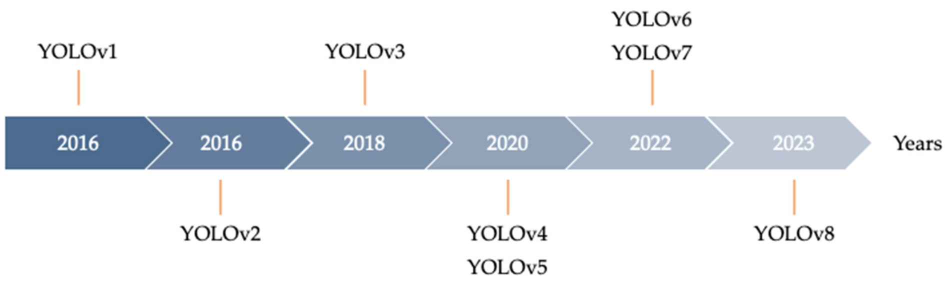


Figure 1. YOLO timeline.

Throughout the historical evolution of the YOLO algorithm, specific versions have stood out, playing pivotal roles in shaping the landscape of object detection. YOLOv1 introduced the concept of one-stage object detection, revolutionizing the field by directly predicting bounding boxes and class probabilities in a single pass. Despite its relative simplicity, YOLOv1 showcased the potential of real-time object detection, setting the stage for further advancements.

YOLOv3 marked a substantial leap in accuracy and versatility by introducing a feature pyramid network and multiple detection scales. This version improved detection performance across various object sizes, solidifying YOLO's reputation as a robust object detection algorithm [35].

YOLOv4 represented a significant milestone by substantially improving both accuracy and speed through optimized architecture and advanced techniques. Achieving state-of-the-art performance, YOLOv4's holistic approach to accuracy and speed propelled it to the forefront of object detection algorithms [36,37].

While YOLOv8 may not be as widely recognized as its predecessors, it has gained prominence due to reported advancements in its accuracy. Positioned as the latest effort to refine the algorithm, YOLOv8 demonstrates a continued commitment to enhancing object detection performance, particularly in scenarios involving small objects and complex scenes [36].

The strategic selection of YOLOv8 for this study is supported by its robust performance and specific advantages that align with the requirements of the proposed quality control application, namely [38–42].

– Advancements in Accuracy:

YOLOv8 has demonstrated notable advancements in accuracy compared to its predecessors. The algorithm has been refined to achieve higher precision in object detection. This increased accuracy is particularly crucial in quality control applications, where the identification of components and defects demands a high level of reliability.

– Suitability for Near-Real-Time Processing:

YOLOv8 strikes a balance between accuracy and speed, making it suitable for near-real-time processing. In industrial quality control scenarios, immediate feedback is essential for timely corrective actions. YOLOv8's efficiency ensures that the system can process video streams swiftly, contributing to the overall speed and responsiveness of the quality control process.

– Advanced Feature Extraction Capabilities:

YOLOv8 boasts advanced feature extraction capabilities, allowing it to discern intricate details within images. This is particularly advantageous in quality control, where the identification of subtle defects or components with nuanced characteristics is paramount. The algorithm's ability to extract relevant features contributes to the system's overall precision.

– Holistic View of Images:

YOLO's architecture processes images as a whole, avoiding the need for a two-step approach involving region proposals and subsequent classification. This holistic view allows YOLOv8 to contextualize information about classes and their appearance, enhancing its capability to precisely detect objects within the image. This is advantageous in scenarios involving complex assembly processes.

– Generalization and Adaptability:

YOLOv8's architecture and training processes enable effective generalization, meaning it can perform well in various scenarios and environments. In the context of industrial quality control, where the system may encounter diverse manufacturing conditions, YOLOv8's adaptability enhances its overall performance.

– Proven Performance in Real-world Applications:

The choice of YOLOv8 is substantiated by its proven performance in real-world applications. The algorithm has been successfully employed in diverse domains, including industrial settings, showcasing its reliability and effectiveness. This track record supports its suitability for the proposed quality control system.

The strategic selection of YOLOv8 for this study is grounded in its advancements in accuracy, its suitability for near-real-time processing, advanced feature extraction capabilities, holistic view of images, generalization, and proven performance in real-world applications. These specific advantages make YOLOv8 well-suited for the intricacies of industrial quality control, providing a robust foundation for a proposed mixed reality (MR)-based quality control system.

2.3. Evolutionary Trends

The evolution of object detection algorithms has been a dynamic interplay between the two-stage and one-stage approaches, each catering to specific requirements and challenges. While two-stage algorithms excel in accuracy but often sacrifice speed, one-stage algorithms offer real-time capabilities at the expense of accuracy, especially for small objects and complex scenes [43,44].

Object detection, a fundamental task in computer vision, plays a crucial role in various applications, including autonomous driving, video surveillance, and medical imaging. Traditional object detectors, such as R-CNN, Faster R-CNN, and YOLO, have revolutionized image understanding by enabling the accurate detection of objects with upright orientations. However, these methods encounter significant challenges in handling objects with arbitrary rotations, which are prevalent in real-world scenarios. To address this challenge, researchers have delved into the domain of rotated object detection, exploring novel techniques to effectively identify and localize rotated objects.

A significant step forward is the introduction of adaptive rotated convolution (ARC) [45]. ARC addresses the limitations of standard convolutions, which treat all input pixels equally, regardless of their orientation. ARC kernels dynamically rotate to match the orientations of objects, enabling the network to extract more informative features. This adaptive rotation mechanism significantly enhances the accuracy of rotated object detection, surpassing the performance of traditional methods.

To handle the increased complexity of rotated objects efficiently, Yang et al. [46] proposed an adaptive object detection system based on early-exit neural networks. This system employs a multi-scale feature extraction backbone and early-exit branches to progressively refine object detection results. Early-exit branches discard redundant features, reducing computational costs while maintaining detection accuracy. This adaptive architecture demonstrates a remarkable performance when detecting rotated objects compared to traditional single-scale detectors.

Zhuang et al. introduced Rank-DETR [47], a method that synergistically combines object detection and instance segmentation techniques to achieve high-quality object detection results. Rank-DETR utilizes a dual-stage architecture as follows: first, it identifies object proposals and then refines their locations and identifies their boundaries. This iterative refinement process ensures precise object localization and segmentation.

Additionally, Rank-DETR employs a novel ranking loss function that prioritizes accurate object detection, further boosting performance.

These studies highlight the transformative impact of rotated object detection research. ARC, the adaptive convolution operation, addresses the challenge of arbitrary object orientations by dynamically rotating convolution kernels, leading to substantial improvements in rotated object detection accuracy. The adaptive object detection system based on early-exit neural networks demonstrates the efficient handling of rotated objects by refining detections at multiple scales and reducing computational resources without compromising accuracy. Rank-DETR, with its dual-stage architecture and ranking loss function, achieves high-quality object detection by prioritizing precise localization and segmentation. These advancements in rotated object detection open new avenues for the development of more robust and versatile object detection systems that are capable of effectively handling real-world scenarios with a wide range of object orientations.

Additional references [48] propose a novel feature extraction module incorporating rotation information, improving feature representation, and enhancing rotated object detection performance.

Reference [49] introduces a rotation-aware feature extraction mechanism that utilizes attention modules to focus on salient features, leading to more accurate rotated object detection. In [50], a multi-stage approach that fuses features extracted from different orientations and refines detections iteratively is presented, achieving improved performance for rotated object detection.

Reference [51] explores the intricacies of visual inspection within assembly processes, leveraging state-of-the-art techniques rooted in contrastive and transfer learning. The authors address the challenges associated with inspecting small components, a critical aspect of assembly quality control, and propose innovative approaches to enhance the precision and efficiency of the inspection process. With a focus on leveraging advanced visual methodologies, this paper is poised to enrich our understanding of contemporary strategies in small component inspections.

Zhao et al. [52] introduced an innovative approach to online assembly inspection, merging a lightweight hybrid neural network with precise positioning box matching techniques. Navigating the complexities of real-time assembly assessments, the authors delivered a comprehensive and integrated methodology promising to enhance the accuracy and efficiency of product assembly inspections significantly.

These additional studies underscore the continuous progress in rotated object detection research, with researchers continuously exploring innovative techniques to address the challenges of detecting objects with arbitrary orientations in real-world environments. The development of robust and versatile rotated object detection systems has the potential to revolutionize various applications, including autonomous navigation, video surveillance, and medical imaging analysis.

3. Methodology

The proposed architecture for the server-side image recognition system operates in the following two main stages: training the model and real-time processing. These stages are crucial for the effective implementation of an MR-based quality control system.

3.1. Model Training

In the initial stage, as illustrated in Figure 2, we outline the proposed process for model training.

Figure 2. Proposed framework for model training.

The process for training a custom image dataset comprises five main steps. Firstly, data collection and preprocessing are essential. In the context of quality control in industrial manufacturing, especially in the automotive industry, a custom image dataset is curated directly from real-world field environments. This industry's intricate tasks necessitate a dataset reflecting the complexity of actual operations. Utilizing images from natural field settings ensures that the dataset becomes a potent tool for training models to accurately detect defects, anomalies, and quality variations. This approach aligns with the unique challenges of the automotive quality control sector, resulting in robust and industry-relevant solutions.

Images should ideally encompass both faulty and non-faulty examples for each product and component, collected under various lighting conditions and angles and featuring different parts and defect types. Preprocessing involves resizing, normalizing, and augmenting the data to enhance the model's ability to generalize and perform well on real-world data. Augmentation simulates different scenarios by applying transformations like rotation, scaling, and cropping.

Next is the crucial step of image annotation, where each image is meticulously reviewed, and relevant objects are marked or labeled with details about the part and nature of the defect. While manual and time-consuming, this annotation process is paramount for creating a reliable training dataset, providing the model with ground truth data points to learn.

With the dataset ready, the appropriate models are selected. A convolutional neural network (CNN) is suitable for part identification due to its ability to learn the spatial hierarchies of image features. In the case of object detection, the framework leverages YOLOv8, a state-of-the-art object detection model known for its accuracy and speed in detecting and localizing multiple objects in a single pass, as indicated in the previous section.

The subsequent step involves model training as an iterative and multi-step process. The dataset is split into training and validation sets, hyperparameters are fine-tuned, and various optimization algorithms are tested. Post-training, models are evaluated on a separate test set, refining them based on the results. Once trained, considered, and refined, the models are deployed to the server, ready for the second stage, which is real-time processing.

3.2. Real-Time Processing

The real-time processing stage of the MR-based quality control system involves creating an architectural design for analyzing and processing the video stream captured by smart glasses. The architecture can be conceptually divided into two main segments: the client-side MR system and the server-side image recognition system.

The client-side MR system utilizes a smart glasses device, which the quality inspector interacts with. This MR headset captures the video stream of the real-world environment, primarily components inspected for quality control, and streams it to the server. Equipped with depth sensors and multiple cameras, the device captures a rich, three-dimensional view crucial for accurately overlaying augmented reality elements on the real-world scene.

On the server side, the system analyzes the incoming video stream, identifying objects and detecting defects. The server runs a machine learning model and utilizes a framework for real-time video processing and computer vision tasks. Combining an MR headset with a powerful server-based image recognition system, this architecture enables a robust and efficient MR-based quality control system that surpasses traditional manual inspection in speed and accuracy. Figure 3 illustrates the proposed framework for real-time processing and visualization within a smart glasses environment for quality inspection.

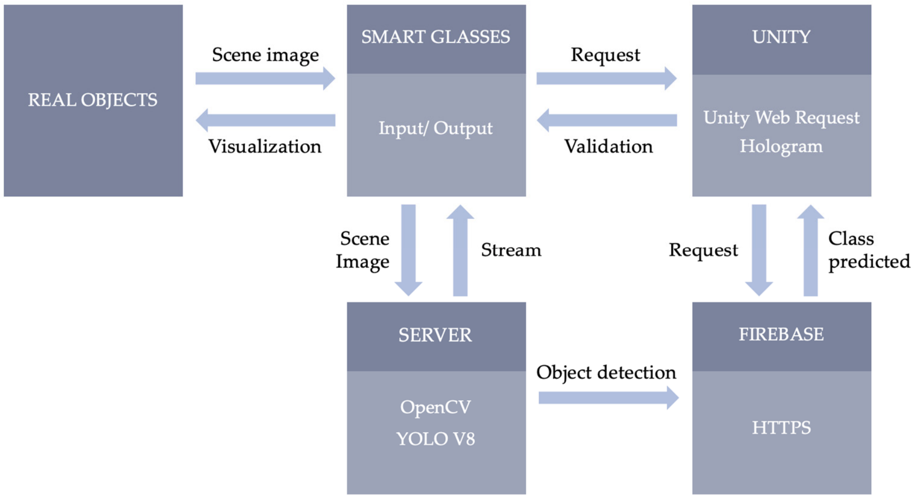


Figure 3. Proposed framework for real-time processing.

The process begins with natural objects on the assembly line, serving as subjects for quality inspection. The smart glasses act as intermediaries, capturing images from the scene and live videos of the assembly process. They not only capture the feed but also serve as the display for visual output, allowing the user to see results superimposed on the real-world view.

In this methodology, OpenCV, an open-source computer vision library, assumes a critical role as the initial gateway for processing the incoming video stream. Its multifaceted capabilities are harnessed to undertake a range of image-processing tasks vital for the preliminary assessment of the assembly process. As the continuous video stream is received, OpenCV takes charge, parsing it into discrete frames and thereby laying the foundation for meticulous frame-by-frame analysis.

Once the frames are extracted, OpenCV seamlessly integrates with the pre-trained part identification model to conduct a granular examination of the individual components and parts within each frame. The versatility of OpenCV shines in its ability to handle diverse image processing operations, ensuring an accurate identification of components even in complex visual scenarios. This preliminary analysis sets the stage for the subsequent phases of the methodology.

Following OpenCV's comprehensive pre-processing, the torch is passed to the YOLOv8 object detection system, which comes into play as the primary engine for identifying and discerning between different components. YOLOv8, having been pre-trained on an extensive dataset specific to the assembly process under investigation, employs its neural network architecture to execute rapid and precise object detection. Its proficiency lies in distinguishing correct assemblies from incorrect ones with exceptional accuracy and speed, which is a characteristic vital for real-time analysis.

OpenCV is chosen for its versatility, high-performance features, and robustness, making it an ideal tool for analyzing the assembly process in this methodology. Its comprehensive suite of image-processing functions facilitates the parsing of continuous video streams into individual frames, allowing for diverse operations such as filtering, transformation, and enhancement. The capability for frame-by-frame analysis is particularly valuable in scenarios requiring a detailed inspection, as commonly encountered in assembly processes [53–55].

The interaction between OpenCV and YOLOv8 is synergistic. OpenCV's role in framing, preprocessing, and initial part identification lays the groundwork for YOLOv8 to execute its specialized task of high-precision object detection. OpenCV not only facilitates the extraction of frames but also enhances the quality of input data for YOLOv8, optimizing the latter's performance. Together, they form a robust pipeline wherein OpenCV's capabilities seamlessly complement the strengths of YOLOv8, collectively contributing to the methodology's efficacy in identifying parts, detecting defects, and providing real-time feedback to inspectors.

In parallel, a secure HTTPS protocol establishes a connection to Firebase, which is the backbone for real-time database management. It logs server actions, including video processing results and relevant metadata, ensuring a persistent state of operation and historical data tracking.

The Unity engine plays an essential role in the visualization of results. Unity sends requests to Firebase to fetch the predictive classification results and generates a holographic display based on this information. An 'OK' hologram is displayed to the operator if the assembly is correct. Should there be any discrepancies, a 'NOT OK' hologram appears, alerting the operator to the specific issue.

The selection of Unity [56–59] for visualization in the MR-based quality control system is underpinned by its versatile capabilities, seamless integration, and ability to elevate the overall user experience. As a robust cross-platform game engine, Unity accommodates the dynamic integration of predictive classification results, enabling the generation of holographic displays that signify the correctness or issues with assemblies. This adaptability ensures that the system's visualization component can effectively communicate real-time quality control results to operators, contributing to a comprehensive and responsive user interface.

Unity's ease of integration with Firebase is a pivotal factor, streamlining the process of fetching predictive classification results. This efficient data flow enhances the system's ability to display timely and accurate holographic representations. This seamless integration facilitates a smooth communication flow between the analysis components of the system and the visualization module, ensuring that operators receive instant feedback on the quality of assemblies.

This entire cycle, from the capture of real objects to the visualization of the hologram, is designed to operate in real-time, minimizing any delay that could impact the efficiency of the assembly process. The integration of real-time streaming, advanced image processing, secure data management, and an interactive holographic display exemplifies the convergence of cutting-edge technologies to enhance the precision and reliability of manufacturing quality control.

4. Results

This case study delves into the practical application of smart glasses for quality control in pneumatic cylinder assembly. To assess its effectiveness in providing real-time feedback on assembly quality, a prototype was developed in a controlled laboratory setting, and the setup is illustrated in Figure 4.

The prototype setup focused on assembling a pneumatic cylinder within a laboratory environment, employing the Microsoft HoloLens 2 smart glasses as the primary tool for the mixed reality-based quality control system [60]. The selection of these glasses was based on their advanced capabilities in capturing and overlaying mixed reality elements in a real-world setting.

The pneumatic cylinder assembly process encompassed various components, such as the cylinder barrel, end caps, piston, tie rods, and nuts. Participants were guided to assemble the cylinder under different conditions, incorporating both correct and incorrect assembly scenarios.

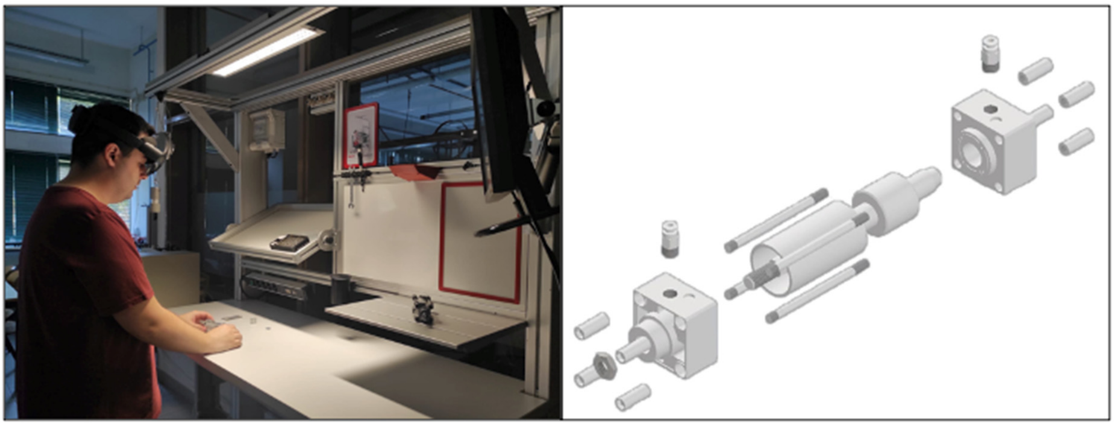


Figure 4. Prototype setup.

The primary objective was to evaluate the prototype system's effectiveness in accurately assessing the quality of the assembly process and providing real-time feedback to the assembler. Through analysis, the smart glasses displayed appropriate holographic overlays, indicating whether the assembly was correct or not. The results of this case study contribute valuable insights into the practical implementation of smart glasses for quality control in pneumatic cylinder assembly.

4.1. Dataset

The initial phase of model training for the quality control system relies on gathering images from the pneumatic cylinder prototype. The quality and diversity of this image dataset form the bedrock of the system's accuracy and robustness. The assembly and disassembly of the pneumatic cylinder are captured under various scenarios, including correct and incorrect assemblies, missing parts, and misaligned components. These images showcase different assembly stages, individual components, partial assemblies, and fully assembled cylinders. Figure 5 displays captured images showcasing the fully assembled pneumatic cylinder under various lighting conditions and from different viewing angles.

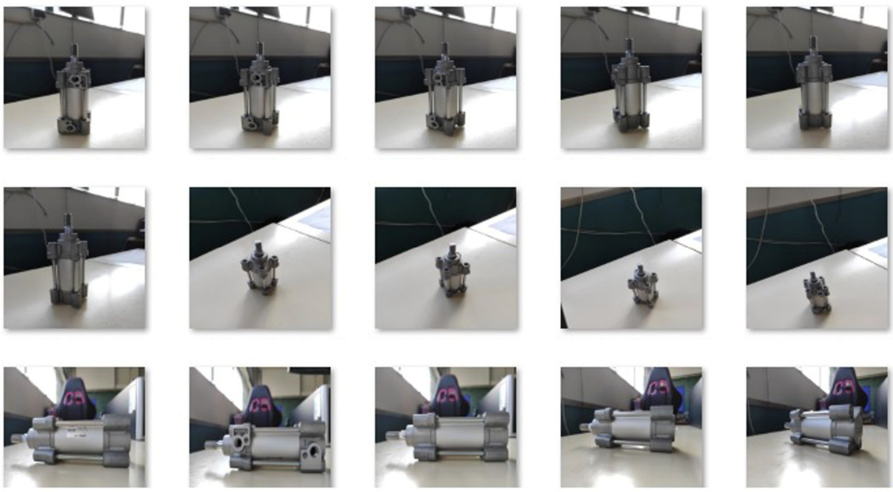


Figure 5. Captured images of the fully assembled pneumatic cylinder (complete).

This comprehensive image collection ensures that the dataset encompasses all potential variations that might occur during pneumatic cylinder assembly. Once collected, the images are categorized based on assembly stage, component type, and assembly correctness. Proper categorization simplifies the annotation process as a crucial step in training the machine learning model. The web-based platform Roboflow [61] was employed for dataset management and image annotation due to its streamlined interface, comprehensive toolset, and efficient handling of large datasets.

The dataset comprises over 2290 annotated images, categorized into the following six distinct classes: cylinder, end caps, piston, tie rods, nuts, and complete assemblies. The pneumatic cylinder prototype dataset was carefully selected, covering various instances essential for training a robust quality control system. It is structured to facilitate not only the classification of various components but also the precise detection and location of these elements in assembly scenarios.

Notably, these instances stand out for their balanced distribution across the defined classes, ensuring that the training process is not biased towards any specific category. Furthermore, their normalized sizes contribute to consistency and computational efficiency during the learning phase.

Figure 6 illustrates the proof of correlation between different instances or features within the dataset through the correlogram.

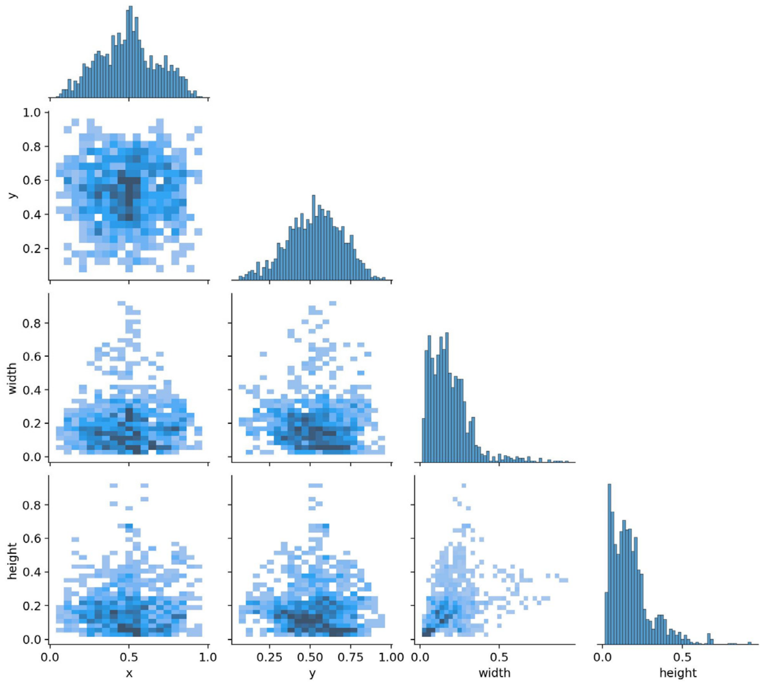


Figure 6. Correlogram of the different instances in the dataset.

A correlation analysis, represented in Figure 6, visually demonstrates the relationships between different instances or features within the dataset. This correlogram serves as a valuable tool for understanding possible redundancies or distinguishing features, guiding subsequent stages of model development and refinement.

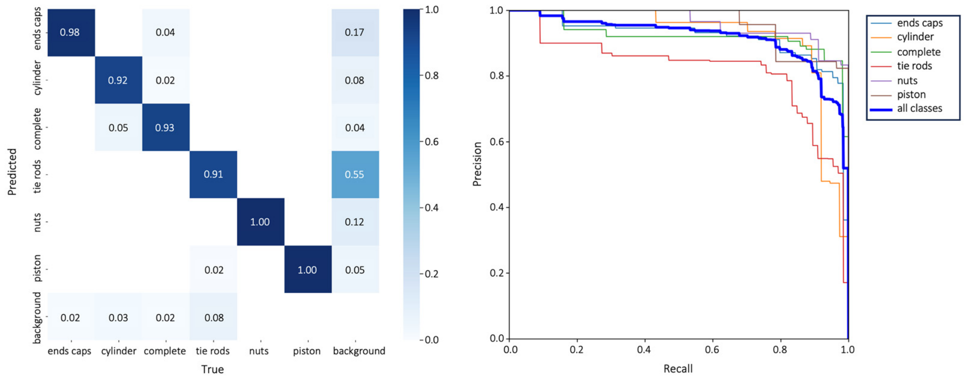
It is crucial to note that in the YOLO algorithm, the presence of correlated features among the chosen elements, known as anchor boxes, can be advantageous. Anchor boxes in YOLO play an essential role in detecting objects across different scales and aspect ratios. These predefined bounding box shapes aid the algorithm during the object detection process, contributing to the algorithm's accuracy and effectiveness.

The dataset employed in this study is characterized by its extensive size, diverse representation of assembly conditions, the inclusion of realistic environmental settings, consideration of lighting conditions and viewing angles, and meticulous normalization. These characteristics collectively contribute to a dataset that not only reflects the complexities of pneumatic cylinder assembly but also provides a solid and versatile foundation for training a machine-learning model tailored to the challenges of industrial quality control.

4.2. Experimentation and Results

The model training for the pneumatic cylinder prototype utilized the YOLOv8 object detection architecture over 50 epochs, with a dataset partitioned into 70% for training, 20% for testing, and 10% for validation. Consistency was maintained by resizing all input images to a resolution of 640 × 640 pixels, providing a standardized input format that enhances training efficiency.

The metrics showcased in Figure 7 show the commendable performance of the trained model when applied to the dataset.



(**a**) (**b**)

Figure 7. (**a**) Confusion matrix and (**b**) precision/recall curve.

These metrics, serving as quantifiable evidence of the model's capabilities, lend credibility to the approach. These satisfactory outcomes validate the efforts in the design and training processes and emphasize the model's potential.

The performance of the proposed detector is quantified using the metrics recall, mAP50, and mAP50-95. Specifically, mAP50-95 represents the mean average precision (mAP) computed over an Intersection over Union (IoU) range from 0.5 to 0.95, offering a holistic assessment of the detector's accuracy across varied thresholds. The training results for all classes based on the proposed method are shown in Table 1.

Table 1. Training results for all classes.

Class	R	mAP50	mAP50-95
All	0.917	0.916	0.777
End caps	0.953	0.919	0.82
Cylinder	0.865	0.921	0.82
Complete	0.875	0.921	0.854
Tie rods	0.879	0.821	0.701
Nuts	0.94	0.959	0.619
Piston	0.991	0.957	0.847

These results provide valuable insights into the performance of the object detector in distinct classes and overall. For all combined classes, we observed that a recall of 0.917 indicates proficiency in identifying objects; since the value is greater than 90%, we considered it a very good value to stop the training process. An mAP50 score of 0.916 further attests to the model's precision and recall balance at an IoU threshold of 0.5. While lower, the mAP50-95 score of 0.777 still signifies strong performance across varied IoU thresholds, underscoring the model's adaptability and effectiveness in different scenarios.

These findings collectively suggest that the proposed object detector exhibits substantial efficacy in identifying and classifying objects of distinct classes with commendable accuracy and consistency. The specific strengths and areas for improvement highlighted by these metrics can inform focused refinements to enhance the model's performance further.

To address false positives, an analysis of the confusion matrix and precision/recall curves in Figure 7a,b is crucial. These values suggest that while the model demonstrates good performance overall, there is room for improvement.

In real-time processing, Microsoft HoloLens 2 smart glasses capture the video stream, which is transmitted to the server for analysis by the trained YOLO model. Combined with the YOLO model, OpenCV ensures that the system can swiftly and accurately identify components and assess the assembly quality.

The analysis results are then sent back to the smart glasses, where the custom Unity application processes the information and generates an MR overlay on the inspector's view.

Figure 8 shows the information presented by the Hololens in the validation process of the pneumatic cylinder's assembly quality control.

(**a**) (**b**)

Figure 8. Validation process of the assembly quality control: (**a**) NOT OK; (**b**) OK.

If the assembly is correct, a holographic indication of "OK" appears on the glasses. If the assembly is found to be incorrect or incomplete, a holographic representation of "NOT OK" appears. This real-time feedback mechanism allows the inspector to instantly identify and correct any errors, enhancing the efficiency and accuracy of the assembly process.

The validation process focuses on the external result of the components' assembly. The system compares the observed assembly configuration with the expected configuration. The assembly is considered successful if the observed configuration matches the expected one. On the other hand, if there is any deviation between the observed and expected configurations, the assembly is classified as incorrect. The validation of operational functionality and component assembly sequence will be analyzed in future work.

Several tests were carried out with similar results. The consistency and accuracy of the server-based image recognition system were evident, contributing to high levels of quality control.

The prototype confirmed the feasibility and potential benefits of the proposed quality control methodology, highlighting the advantages of incorporating mixed reality and computer vision technologies in the inspection process.

Three people assembled the pneumatic cylinder 4 h a day for 5 days to evaluate the system's performance, totaling 818 assemblies. The system detected 11 incorrect assemblies, constituting 1.34% of the total pneumatic cylinder assemblies during the test period.

This result highlights the system's effectiveness in identifying and rectifying errors in the assembly process. The system's ability to quickly identify errors substantially reduces defective products, promoting excellent product reliability and increasing customer satisfaction. The real-time functionality of the system, seamlessly integrated into the assembly process, enhances operational efficiency by promptly detecting errors and enabling immediate corrective measures. This proactive approach curtails the propagation of faults throughout the manufacturing line, establishing the system as a linchpin in pursuing high-level quality control standards.

The integration process was not without its challenges. Several notable obstacles were encountered and addressed during the development and testing phases.

– Hardware Limitations:

The HoloLens 2, while a state-of-the-art device, poses challenges in terms of computational power and memory constraints. Optimizing the object detection model and implementing efficient data processing is crucial to ensure real-time performance on the device.

– Network Latency:

Ensuring low-latency communication between the HoloLens and the server is critical for real-time feedback. It was necessary to implement robust protocols to handle potential delays.

– User Interface Design:

Designing an intuitive and user-friendly interface for HoloLens requires careful consideration. The placement and visualization of holographic overlays needed to be precise and unobtrusive, enhancing the user experience without hindering the inspector's workflow.

To overcome these challenges, a collaborative approach involving multidisciplinary expertise was adopted. Hardware optimizations, network protocol refinements, and iterative user interface testing were conducted. Continuous user feedback during simulated scenarios allowed for rapid adjustments, ensuring that the final integrated system met user expectations and performance standards.

Our study has made notable progress in demonstrating the feasibility and advantages of integrating mixed reality (MR) technologies into industrial quality control. However, it is imperative to recognize that specific challenges linked with introducing MR technologies into industrial settings still need to be addressed. These challenges include potential user distractions, security risks, eyestrain, and cybersickness. Previous authors [62,63] have noted that using smart glasses may induce visual fatigue and cybersickness, manifesting as symptoms like disorientation and dizziness.

In our specific study, visual fatigue and cybersickness were not thoroughly explored. This limitation arises due to the short duration of each cylinder assembly, lasting less than 5 min, and the reduced continuous assembly time in the laboratory environment. Nevertheless, recognizing the significance of this aspect, we acknowledge that the potential impact of smart glasses on visual comfort and overall well-being is an important consideration. Consequently, investigating these potential challenges, including visual fatigue and cybersickness, is an area to be analyzed in future work.

5. Conclusions

The primary goal of this study was to create a quality control system by integrating smart glasses with a server-based image recognition system and assess its effectiveness in a controlled environment. Additionally, this study aimed to evaluate the system's performance and feasibility for larger-scale implementation in the manufacturing industry, providing insights to facilitate the adoption of smart glasses in industrial quality control.

In achieving these objectives, this study demonstrated notable success. The system effectively improved the quality control process in controlled conditions, showcasing the benefits of real-time feedback, reduced human error, and enhanced observation capabilities facilitated by smart glasses.

The case study underscored the efficacy of combining YOLO and OpenCV for constructing robust real-time object detection systems. Such integrated systems have the potential to automate tasks, enhance accuracy, and streamline operations, leading to significant efficiency gains.

Through prototype testing in real-world conditions, this study identified potential challenges and areas for improvement. This valuable insight is crucial for refining the mixed reality (MR) system for broader operational scales, ensuring its robustness and reliability in the manufacturing industry. The successful prototype testing confirmed the theoretical promise and practical applicability of implementing this MR system on a larger scale.

Furthermore, the study provides a practical roadmap for adopting smart glasses in industrial quality control. The detailed analysis of the development, implementation, and evaluation process serves as a valuable resource for organizations considering the integration of MR technology into their quality control processes. As a result, this study contributes to the advancement of MR technology and its applications in industrial settings.

The advantages presented by smart glasses in quality control, such as real-time feedback, increased accuracy, reduced human error, and improved overall efficiency, make the development of an MR-based quality control system an attractive proposition for industries seeking to innovate and optimize their production processes.

5.1. Comparison with Traditional Inspection Methods

The revolutionary MR-based quality control system, seamlessly incorporating smart glasses and advanced image recognition, represents a paradigmatic shift from traditional inspection methodologies. In this comparative analysis, we accentuate the advantages that position the proposed system as a pivotal advancement in industrial quality control.

Regarding speed and efficiency, traditional manual inspections reliant on human visual assessments are time-intensive and susceptible to inconsistencies. By contrast, the proposed MR-based system, armed with real-time object detection, expedites inspections. Digital overlays on smart glasses facilitate swift identification, eliminating the need for laborious manual cross-referencing against plans or physical checklists.

Accuracy and consistency, paramount in quality control, are often compromised by human error, fatigue, and subjectivity in traditional methods. Integrating cutting-edge technologies, specifically YOLOv8 and OpenCV, ensures precise and consistent inspections. Advanced computer vision algorithms detect defects with accuracy that surpass human assessments, eliminating the variability associated with traditional approaches.

The early detection and prevention capabilities of the MR-based system further distinguish it from conventional methods. This system mitigates downstream complications by proactively identifying and rectifying issues in the early stages of the manufacturing process, minimizing the risk of rework, reducing recalls, and fortifying the brand's image. The user-friendly interface presented through smart glasses and the system's adaptability to varied manufacturing scenarios underscores its transformative impact on operational efficiency. Real-time tracking and data logging provide actionable insights, allowing managers to optimize workflows, identify bottlenecks, and enhance productivity. In summary, the proposed MR-based quality control system transcends the limitations of traditional methods, ushering in a new era of technology-driven, proactive quality control aligned with the demands of modern industrial settings.

5.2. Future Work

In future work, it is imperative to address and overcome the challenges and limitations identified in the proposed MR-based quality control framework. Real-world deployment poses challenges such as varying lighting conditions, environmental noise, and hardware compatibilities, demanding meticulous consideration for seamless integration into diverse industrial settings. Research efforts should be directed toward devising solutions to these challenges to enhance the system's effectiveness in practical manufacturing facilities.

Moreover, the identification and validation of the assembly sequence of components are important topics that need to be explored to enhance the quality of equipment assembly.

Another avenue for future research involves addressing integration challenges with existing manufacturing systems. Strategies for seamless integration should be explored, taking into account the diverse hardware and software configurations prevalent in industrial setups. Collaborative efforts with industry partners can facilitate real-world testing and validation, ensuring that the proposed MR-based quality control system aligns with the practical needs of manufacturing environments.

Additionally, future research endeavors can explore the potential of enhancing system efficiency and usability through the analysis of NASA LTX data. By actively engaging with these considerations, the proposed framework and methodology can be refined, making substantial contributions to advancing MR-based quality control in the realm of industrial manufacturing.

Recognizing the importance of user well-being and its potential impact on overall comfort, it is vital in future work to further analyze the effects of using smart glasses on visual fatigue and cybersickness. This exploration aims to comprehensively understand the user experience and address potential concerns associated with the prolonged use of MR devices in industrial environments. Continuous refinement based on user feedback can optimize the user experience, promoting increased comfort and widespread system adoption.

Author Contributions: J.S. conceptualized the study, developed the methodology, and wrote the paper; P.C. and L.S. conceptualized the study and developed the methodology; P.V. provided validation and reviewed the manuscript; P.M. provided validation and reviewed the manuscript; A.L.-R. supervised the study and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by National Funds through the FCT—Foundation for Science and Technology—I.P., within the scope of the project Ref. UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD) and the Instituto Politécnico de Viseu for their support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Juran, J.M.; Defeo, J.A. *Juran's Quality Handbook: The Complete Guide to Performance Excellence*; McGraw-Hill: New York, NY, USA, 2010.
2. Thoben, K.; Wiesner, S.; Wuest, T. Industrie 4.0 and Smart Manufacturing—A Review of Research Issues and Application Examples. *Int. J. Autom. Technol.* **2017**, *11*, 4–16. [CrossRef]
3. Jayaram, A. Lean Six Sigma Approach for Global Supply Chain Management Using Industry 4.0 and IIoT. In Proceedings of the 2nd International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 14–17 December 2016; pp. 89–94.
4. Pacana, A.; Czerwińska, K.; Dwornicka, R. Analysis of Quality Control Efficiency in The Automotive Industry. *Transp. Res. Procedia* **2021**, *55*, 691–698. [CrossRef]
5. Kim, H.; Lin, Y.; Tseng, T.L.B. A Review on Quality Control in Additive Manufacturing. *Rapid Prototyp. J.* **2018**, *24*, 645–669. [CrossRef]
6. Rusell, J.P. *The ASQ Auditing Handbook*, 4th ed.; ASQ Quality Press: Milwaukee, WI, USA, 2012.
7. Kujawińska, A.; Vogt, K. Human Factors in Visual Quality Control. *Manag. Prod. Eng. Rev.* **2015**, *6*, 25–31. [CrossRef]
8. Drury, C.G. Human Factors and Automation in Test and Inspection. In *Handbook of Industrial Engineering*, 3rd ed.; Salvendy, G., Ed.; John Wiley and Sons: Hoboken, NJ, USA, 2001; pp. 1887–1920.
9. Gallwey, T.J. Selection Tests for Visual Inspection on a Multiple Fault Type Task. *Ergonomics* **1982**, *25*, 1077–1092. [CrossRef]
10. See, J.E. Visual Inspection Reliability for Precision Manufactured Parts. *Hum. Factors* **2015**, *57*, 1427–1442. [CrossRef]
11. Shahin, M.; Chen, F.F.; Bouzary, H.; Krishnaiyer, K. Integration of Lean Practices and Industry 4.0 Technologies: Smart Manufacturing for Next-Generation Enterprises. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 2927–2936. [CrossRef]
12. Silva, J.; Nogueira, P.; Martins, P.; Vaz, P.; Abrantes, J. Exploring the Potential of Mixed Reality as a Support Tool for Industrial Training and Production Processes: A Case Study Utilizing Microsoft HoloLens. In *Advances in Intelligent Systems and Computing*; Iglesia, D., Santana, J., Rivero, A., Eds.; Springer: Cham, Switzerland, 2023; pp. 187–196.
13. Li, W.; Nee, A.Y.C.; Ong, S.K. A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. *Multimodal Technol. Interact.* **2017**, *1*, 17. [CrossRef]
14. Segovia, D.; Mendoza, M.; Mendoza, E.; González, E. Augmented Reality as a Tool for Production and Quality Monitoring. *Procedia Comput. Sci.* **2015**, *75*, 291–300. [CrossRef]

15. Baroroh, D.K.; Chu, C.H.; Wang, L. Systematic Literature Review on Augmented Reality in Smart Manufacturing: Collaboration Between Human and Computational Intelligence. *J. Manuf. Syst.* **2021**, *61*, 696–711. [CrossRef]
16. Muñoz, A.; Martí, A.; Mahiques, X.; Gracia, L.; Solanes, J.E.; Tornero, J. Camera 3D Positioning Mixed Reality-Based Interface to Improve Worker Safety, Ergonomics and Productivity. *CIRP J. Manuf. Sci. Technol.* **2020**, *28*, 24–37. [CrossRef]
17. Rokhsaritalemi, S.; Sadeghi-Niaraki, A.; Choi, S.M. A Review on Mixed Reality: Current Trends, Challenges and Prospects. *Appl. Sci.* **2020**, *10*, 636. [CrossRef]
18. Flavián, C.; Ibáñez-Sánchez, S.; Orús, C. The Impact of Virtual, Augmented and Mixed Reality Technologies on the Customer Experience. *J. Bus. Res.* **2019**, *100*, 547–560. [CrossRef]
19. Runji, J.M.; Lin, C.-Y. Markerless Cooperative Augmented Reality-Based Smart Manufacturing Double-Check System: Case of Safe PCBA Inspection Following Automatic Optical Inspection. *Robot. Comput. Integr. Manuf.* **2020**, *64*, 101957. [CrossRef]
20. Diwan, T.; Anirudh, G.; Tembhurne, J.V. Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applications. *Multimed. Tools Appl.* **2023**, *82*, 9243–9275. [CrossRef] [PubMed]
21. Hussain, M. YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature Toward Digital Manufacturing and Industrial Defect Detection. *Machines* **2023**, *11*, 677. [CrossRef]
22. Ren, Y.; Zhu, C.; Xiao, S. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. *Appl. Sci.* **2018**, *8*, 813. [CrossRef]
23. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
24. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Washington, DC, USA, 11–18 December 2015; pp. 1440–1448.
25. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]
26. Shao, F.; Wang, X.; Meng, F.; Zhu, J.; Wang, D.; Dai, J. Improved Faster R-CNN Traffic Sign Detection Based on a Second Region of Interest and Highly Possible Regions Proposal Network. *Sensors* **2019**, *19*, 2288. [CrossRef]
27. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
28. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-time Object Detection. In Proceedings of the IEEE Conference on Computer Vision, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.
29. Ravi, N.; El-Sharkawy, M. Real-Time Embedded Implementation of Improved Object Detector for Resource-Constrained Devices. *J. Low Power Electron. Appl.* **2022**, *12*, 21. [CrossRef]
30. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo Algorithm Developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [CrossRef]
31. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37.
32. Li, Y.; Fan, Q.; Huang, H.; Han, Z.; Gu, Q. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. *Drones* **2023**, *7*, 304. [CrossRef]
33. Kim, J.H.; Kim, N.; Won, C.S. High-Speed Drone Detection Based on Yolo-V8. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–11 June 2023; pp. 1–2.
34. Lou, H.; Duan, X.; Guo, J.; Liu, H.; Gu, J.; Bi, L.; Chen, H. DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. *Electronics* **2023**, *12*, 2323. [CrossRef]
35. Liu, G.; Nouaze, J.C.; Touko, P.L.; Kim, J.H. YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3. *Sensors* **2020**, *20*, 2145. [CrossRef] [PubMed]
36. Yu, J.; Zhang, W. Face Mask Wearing Detection Algorithm Based on Improved YOLO-v4. *Sensors* **2021**, *21*, 3263. [CrossRef] [PubMed]
37. Gai, R.; Chen, N.; Yuan, H. A Detection Algorithm for Cherry Fruits Based on the Improved YOLO-v4 Model. *Neural Comput. Applic.* **2023**, *35*, 13895–13906. [CrossRef]
38. Ruiz-Ponce, P.A.; Ortiz-Perez, D.; Garcia-Rodriguez, J.; Kiefer, B. POSEIDON: A Data Augmentation Tool for Small Object Detection Datasets in Maritime Environments. *Sensors* **2023**, *23*, 3691. [CrossRef] [PubMed]
39. Ma, M.; Pang, H. SP-YOLOv8s: An Improved YOLOv8s Model for Remote Sensing Image Tiny Object Detection. *Appl. Sci.* **2023**, *13*, 8161. [CrossRef]
40. Yandouzi, M.; Grari, M.; Berrahal, M.; Idrissi, I.; Moussaoui, O.; Azizi, M.; Ghoumid, K.; Elmiad, A.K. Investigation of Combining Deep Learning Object Recognition with Drones for Forest Fire Detection and Monitoring. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 377–384. [CrossRef]
41. Kumar, N.; Stephi, S.; Meenakshi, R.; Sarath, G. Wheat Head Detection using YOLO: A Comparative Study. In Proceedings of the International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 13–15 December 2022; pp. 182–186.
42. Gašparović, B.; Mauša, G.; Rukavina, J.; Lerga, J. Evaluating YOLOV5, YOLOV6, YOLOV7, and YOLOV8 in Underwater Environment: Is There Real Improvement? In Proceedings of the 8th International Conference on Smart and Sustainable Technologies (SpliTech), Bol, Croatia, 20–23 June 2023; pp. 1–4.

43. Padilla, R.; Passos, W.L.; Dias, T.L.B.; Netto, S.L.; da Silva, E.A.B. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics* **2021**, *10*, 279. [CrossRef]
44. Liu, H.; Duan, X.; Chen, H.; Lou, H.; Deng, L. DBF-YOLO: UAV Small Targets Detection Based on Shallow Feature Fusion. *IEEE J. Trans. Electr. Electron. Eng.* **2023**, *18*, 605–612. [CrossRef]
45. Pu, Y.; Wang, Y.; Xia, Z.; Han, Y.; Wang, Y.; Gan, W.; Wang, Z.; Song, S.; Huang, G. Adaptive Rotated Convolution for Rotated Object Detection. *arXiv* **2023**, arXiv:2303.07820.
46. Yang, L.; Zheng, Z.; Wang, J.; Song, S.; Huang, G.; Li, F. An Adaptive Object Detection System based on Early-exit Neural Networks. *IEEE Trans. Cogn. Dev. Syst.* **2023**, *16*, 332–345. [CrossRef]
47. Pu, Y.; Liang, W.; Hao, Y.; Yuan, Y.; Yang, Y.; Zhang, C.; Hu, H.; Huang, G. Rank-DETR for high quality object detection. In Proceedings of the 36th Conference on Neural Information Processing Systems, New Orleans, LO, USA, 28 November–9 December 2022. Available online: https://proceedings.neurips.cc/paper_files/paper/2023/hash/34074479ee2186a9f236b8fd03635372-Abstract-Conference.html (accessed on 7 January 2024).
48. Wu, D.; Wang, Y.; Ma, H.; Ai, L.; Yang, J.; Zhang, S.; Li, W. Adaptive feature extraction method for capsule endoscopy images. *Vis. Comput. Ind. Biomed. Art* **2023**, *6*, 24.
49. Zhao, W.; Kang, Y.; Chen, H.; Zhao, Z.; Zhao, Z.; Zhai, Y. Adaptively Attentional Feature Fusion Oriented to Multiscale Object Detection in Remote Sensing Images. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–11. [CrossRef]
50. Cheng, S.; Li, F.Y.; Qiao, S.S.; Shang, D.L.; Zhou, Y.M. AMA-Det: Enhancing Shared Head of One-Stage Object Detection with Adaptation, Merging, and Alignment. *IEEE Access* **2023**, *11*, 11377–11389. [CrossRef]
51. Zhao, S.; Wang, J.; Shi, T.; Huang, K. Contrastive and transfer learning-based visual small component inspection in assembly. *Adv. Eng. Inform.* **2024**, *59*, 102308. [CrossRef]
52. Zhao, S.W.; Wang, J.F.; Li, W.; Lu, L.F. Online assembly inspection integrating lightweight hybrid neural network with positioning box matching. *IEEE Access* **2023**, *11*, 139223–139235. [CrossRef]
53. Deshpande, H.; Singh, A.; Herunde, H. Comparative Analysis on YOLO Object Detection with OpenCV. *Int. J. Res. Ind. Eng.* **2020**, *9*, 46–64.
54. Sharma, A.; Pathak, J.; Prakash, M.; Singh, J.N. Object Detection Using OpenCV and Python. In Proceedings of the 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 17–18 December 2021; pp. 501–505.
55. Howse, J.; Minichino, J. *Learning OpenCV 4 Computer Vision with Python 3: Get to Grips with Tools, Techniques, and Algorithms for Computer Vision and Machine Learning*, 3rd ed.; Packt Publishing: Birmingham, UK, 2020.
56. Samini, A.; Palmerius, K.L.; Ljung, P. A Review of Current, Complete Augmented Reality Solutions. In Proceedings of the International Conference on Cyberworlds (CW), Caen, France, 28–30 September 2021; pp. 49–56.
57. Dontschewa, M.; Stamatov, D.; Marinov, M.B. Mixed Reality Smart Glass Application for Interactive Working. In Proceedings of the XXVI International Scientific Conference Electronics, Sozopol, Bulgaria, 13–15 September 2017; pp. 1–4.
58. Anjum, T.; Lawrence, S.; Shabani, A. Augmented Reality and Affective Computing on the Edge Makes Social Robots Better Companions for Older Adults. In Proceedings of the 2nd International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS), Electronics Network, Virtual, 27–28 October 2021; pp. 196–204.
59. Yoon, Y.-S.; Kim, D.-M.; Suh, J.-W. Augmented Reality Services Using Smart Glasses for Great Gilt-bronze Incense Burner of Baekje. In Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC), Jeju, Republic of Korea, 6–9 February 2022; pp. 1–4.
60. Park, S.; Bokijonov, S.; Choi, Y. Review of Microsoft HoloLens Applications Over the Past Five Years. *Appl. Sci.* **2021**, *11*, 7259. [CrossRef]
61. Protik, A.A.; Rafi, A.H.; Siddique, S. Real-time Personal Protective Equipment (PPE) Detection Using YOLOv4 and TensorFlow. In Proceedings of the 2021 IEEE Region 10 Symposium (TENSYMP), Jeju, Republic of Korea, 23–25 August 2021; pp. 1–6.
62. Vovk, A.; Wild, F.; Guest, W.; Kuula, T. Simulator Sickness in Augmented Reality Training Using the Microsoft HoloLens. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montréal, QC, Canada, 21–26 April 2018; pp. 1–9.
63. Kaufeld, M.; Mundt, M.; Forst, S.; Hecht, H. Optical see-through augmented reality can induce severe motion sickness. *Displays* **2022**, *74*, 102283. [CrossRef]