MYU Tokyo

S & M 3750

Enhancing Intelligent Processing System with Generative Adversarial Networks

Yu-Ching Lin,[1] Cheng-Hsing Wu,[2] and Meng-Hua Yen[1*]

[1]Department of Electronic Engineering, National Chin-Yi University of Technology, Taichung 41170, Taiwan
[2]Weimao Precision Corporation Limited, Taichung 420, Taiwan

(Received November 1, 2023; accepted July 18, 2024)

Keywords: automatic feeding system, graphical user interface, automatic optical inspection

The integration of automation equipment with intelligence has undoubtedly become the trend in the automation industry. In this regard, automated optical inspection (AOI) application, with advantages such as not being limited by working hours, rapid inspection, and low labor costs, are applied in the automation field. For mass production scenarios, AOI can address the allocation of operators, learning costs, and deficiencies in production efficiency. In this study, we adopted the You Only Look Once version 7 (YOLOv7) architecture for the AOI system to conduct sample inspections. At the same time, generative adversarial networks (GANs) were used to expand the sample database. The Ethernet communication architecture was also employed to integrate equipment such as a six-axis robotic arm, programmable logic controller (PLC), and an industrial computer. This integration gave the system intelligent functions such as automatic scheduling, production report statistics, and real-time monitoring. A graphical user interface was also designed to reduce personnel learning costs, simplify operations, and enhance equipment uptime. Ultimately, through the training of the YOLOv7 detection model, we achieved excellent detection results. The detection precision reached 91%, and the mean average precision (mAP@0.5) reached 83%. This confirms the value of using GAN technology to expand the AOI sample database, especially when there is a shortage of samples in the early stages of production. This high accuracy not only helps improve the accuracy of AOI detection but also enhances the production efficiency of the intelligent processing system.

1. Introduction

The automation industry can be traced back to the Industrial Revolution. During that time, the invention and utilization of machinery significantly increased factory productivity but also brought challenges such as labor shortages and inefficiencies. To address these issues, researchers began exploring how to automate machines, aiming to enhance productivity and efficiency while ensuring the stability and safety of the production process.(1)

With the advancement of electronic technology, computer technology, and control theory, research in the automation industry has made significant progress. Modern automation now

*Corresponding author: e-mail: emh1989@ncut.edu.tw
https://doi.org/10.18494/SAM4746

encompasses a wide range of fields, from industrial control systems to artificial intelligence. Among these, computer control systems are a crucial component of the automation industry. They enable automated control and monitoring of equipment and sensors. Computer control systems also include a graphical user interface (GUI), which improves the interaction between users and computers, making industrial control systems less cumbersome and more user-friendly.(2)

Furthermore, the application of technologies such as machine vision, machine learning, and artificial intelligence has diversified research in the automation industry.(3) The development of machine vision technologies, such as charge coupled device (CCD) cameras, optical techniques, and image processing technologies, has enabled the realization of automated optical inspection (AOI) technology. AOI is a technique that uses optical systems to rapidly, efficiently, and accurately inspect products. Compared with traditional inspection methods, it is more capable of detecting minute defects and issues.(4)

In 2022, Wang and others published You Only Look Once version 7 (YOLOv7),(5) a real-time object detection model. We provide a comparison chart of detection speed and accuracy below. The chart shows that within the range of 5 to 160 frames per second (FPS), YOLOv7's performance surpassed those of many known object detection models, as shown in Fig. 1. Compared with the YOLOv4 detection model published by the same authors in 2020, YOLOv7 also demonstrated superior results. On the basis of the accuracy of the YOLOv7 detection model, we anticipate achieving even more exceptional object detection performance.

1.1 Research motivation and objectives

In the manufacturing industry, product inspection is a critical step, as it ensures product quality while also enhancing production efficiency and product competitiveness.(6) However, traditional product inspection methods often rely on manual operations, which are not only time-

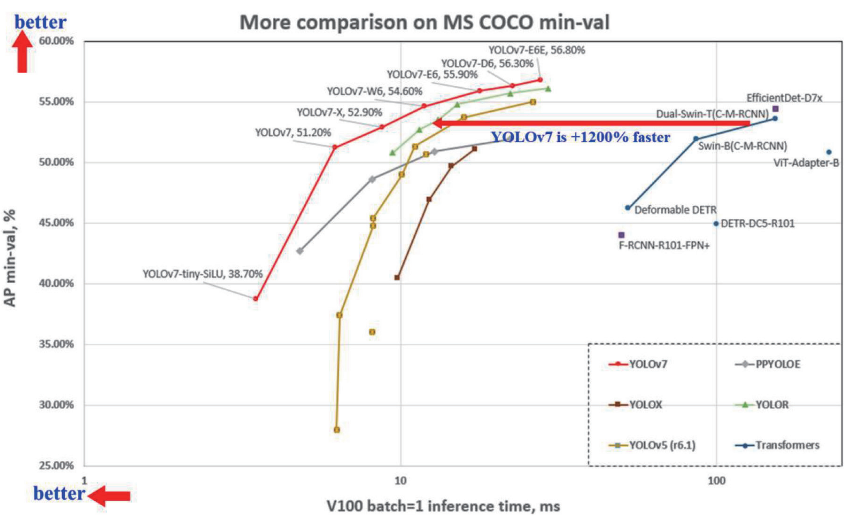

Fig. 1. (Color online) Comparison chart of various real-time object detection models. Source of information: "Wang et al.(5)"

consuming and labor-intensive but also susceptible to human factors. As a result, there is some rate of misinspection and missed inspection,(7) making it challenging to meet the inspection requirements of mass-produced products.

In recent years, the generative adversarial network (GAN)(8) has emerged as a promising tool in the field of image generation. By introducing GANs, we can improve various aspects of intelligent processing systems. This includes optimizing the production process, quality control, fault detection, and real-time monitoring. Such applications are expected to offer a higher degree of automation, reduce the need for human intervention, and simultaneously enhance production efficiency and product quality.

To improve the accuracy and efficiency of product inspection, we investigated how to enhance intelligent processing systems using GANs. By integrating computer technology, automated operations of robotic arms, and AI-driven product inspection, efficient product inspection and rapid feeding can be achieved. This not only reduces labor costs and the impact of human factors but also promotes the intelligent upgrade and technological advancement of the manufacturing industry.

2. Materials and Methods

2.1 Research methods

The methodology of this study is divided into six steps. First, sample collection is conducted by sampling multiple processed parts. Next, an automated loading and unloading inspection system is designed. On the basis of the collected data, a product inspection system is designed to enhance processing efficiency and includes robotic arms, sensors, and an image processing system. In the third step, system integration is carried out. Control software is developed for the real-time monitoring of signals from equipment such as robotic arms and programmable logic controllers (PLCs), as well as for the operation of the loading and unloading systems. Fourth, system testing is conducted. The automated loading system is tested, including the simulation of different product processing techniques, and the testing of the precision and stability of the robotic arm and the reliability of the control system. In the fifth step, modifications and optimizations are made referring to the test results to improve the performance and reliability of the automated loading system. Lastly, statistical research data are collected. Data from actual applications are gathered for feedback and optimization, which continuously enhance the system's performance and efficiency. When combined, these steps can effectively facilitate the research and development of the automated loading and unloading inspection system, as shown in Fig. 2.

2.2 Hardware architecture

Figure 3 provides a schematic diagram of the hardware developed in this study. The automated feeding system is specifically designed for production equipment. To address product diversity, this system incorporates a six-axis robotic arm. The high flexibility of the robotic arm

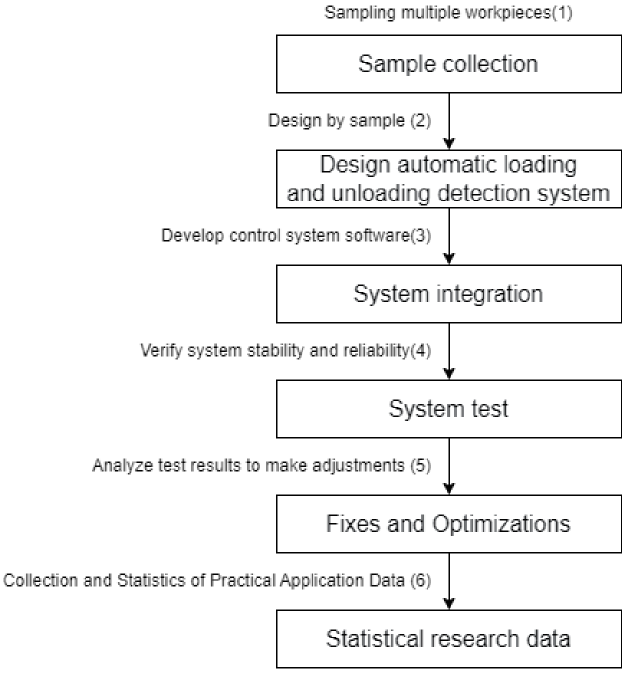

Fig. 2. Flowchart of research methodology.

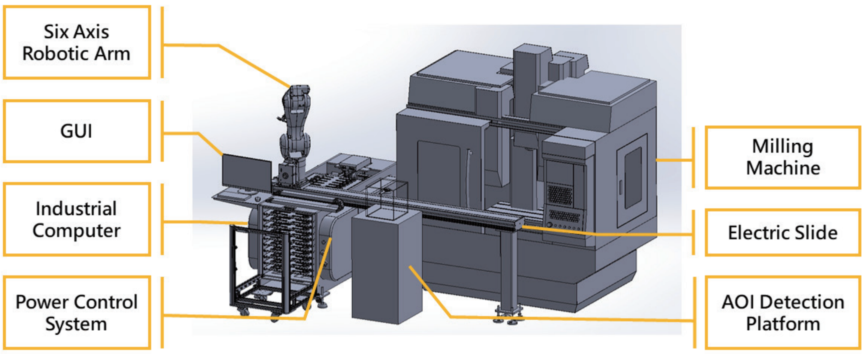

Fig. 3. (Color online) Schematic of hardware devices. The six-axis robotic arm is erected on the electric slide for product clamping.

makes it suitable for the task of picking up objects. For commonly used computer numerical control (CNC) milling machines in product processing, we designed the seventh-axis electric slide to expand the range of motion of the robotic arm in this study. The PLC provides power to the equipment and collects data from sensors. Finally, the system is integrated with an industrial computer, simplifying the operation process and improving equipment uptime. In summary, with the above hardware design, research on using GANs to enhance intelligent processing systems can be carried out.

Operators manually place the workpieces to be processed on a tray cart and push them into the tray storage area. For this area, in this framework, we plan the use of mechanical tensions and positioning sensors to help users confirm whether the tray is set correctly. The maximum sensing error of the positioning sensor, as specified in the sensor specifications, is 4 mm. Subsequently, the electronic control system controls the stepper motor to move the tray to a designated position for the robotic arm to pick up the workpieces. Additionally, we have constructed an AOI inspection platform, as shown in Fig. 4. The darkroom structure of this mechanism isolates most of the external light noise, and a ring light is added to enhance the effect of the industrial camera capturing samples. Simultaneously, the mechanism ensures consistent positioning of the workpiece during each capture, thus minimizing glare from the workpiece.

2.3 Communication architecture

Ethernet/IP is an industrial communication protocol based on Ethernet. It is an extension of the Ethernet protocol enabling real-time communication and remote equipment monitoring. The main advantages of Ethernet/IP lie in its speed, stability, and scalability. It can achieve high-speed data transmission and reliable remote control.(9) Given the above, Ethernet/IP meets the communication architecture requirements of this study. Therefore, our system architecture adopts the Ethernet/IP communication protocol to establish communication between devices. This communication protocol is mainly used in process control and other automation industries. Ethernet/IP uses the physical layer of Ethernet and connects to it via the RJ45 network port to achieve data transmission. Network switches are responsible for network bridging, addressing the shortage of PC network ports. Apart from the network switches, all other equipment is physically connected through, for example, physical I/O lines, USB 3.0, HDMI, and other transmission ports, to achieve data transmission and hardware equipment integration, as shown in Fig. 5.

2.4 GUI

In mass production scenarios, there are significant deficiencies in the allocation of operators, learning costs, and production efficiency. We achieve automation system integration by designing a GUI, enhancing the production efficiency of the robotic arm's loading system in this

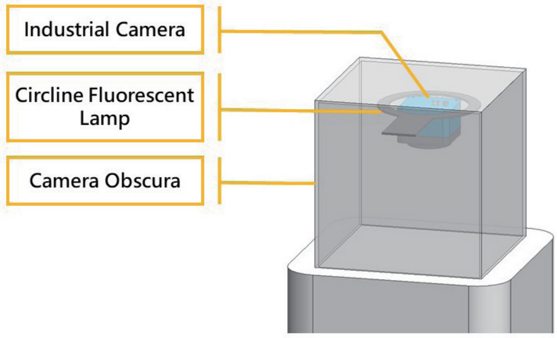

Fig. 4. (Color online) AOI detection platform. This institutional diagram is a perspective view.

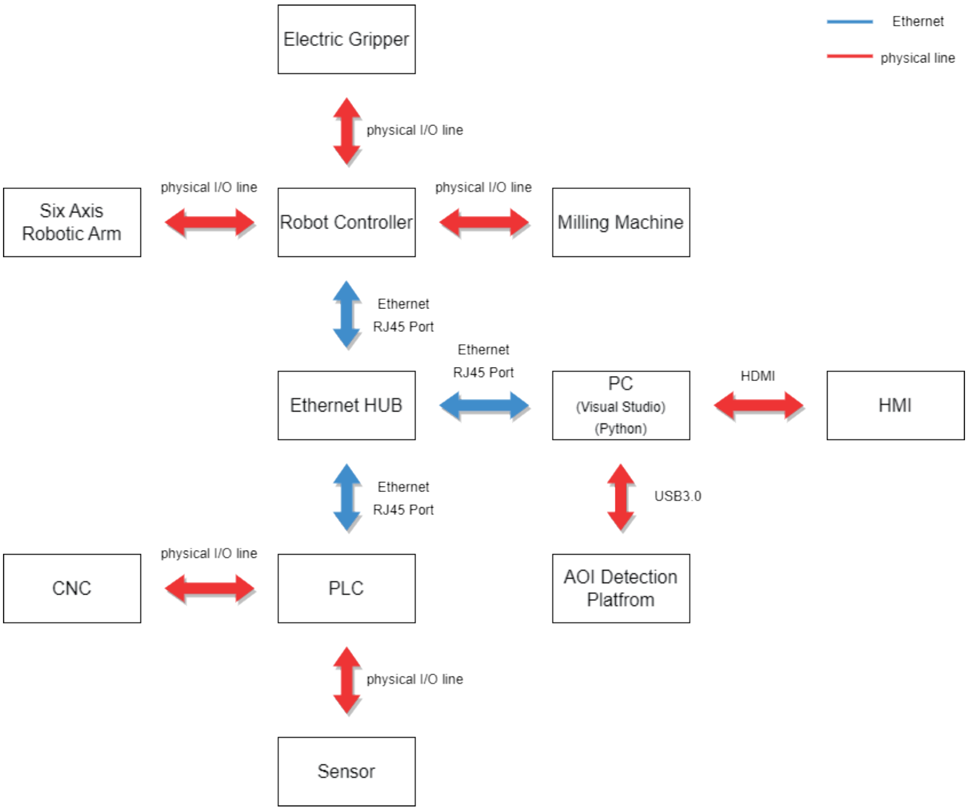

Fig. 5. (Color online) Communication architecture.

study. It also monitors signals from equipment such as the robotic arm and PLC in real time, ensuring the safe operation of the equipment. The designed GUI is shown in Fig. 6. The left side of the interface is the product count area, the center is for hardware equipment status monitoring, and the right side is the system operation area. Real-time monitoring of equipment signals is shown in Figs. 7 and 8.

2.5 Verifying the feasibility of GANs

To verify the effectiveness of GANs in generating sample images, we collected samples using the previously designed AOI platform. The samples are categorized into three types: defect-free, internal defects, and external defects. Most of the external defects are caused by the milling machine path, while internal defects are due to misalignment of the fixtures. All defect locations are marked with red boxes for easier reading by the user. However, during the subsequent experiments, these red boxes were not present. The sample images are shown in Fig. 9.

On the basis of the previous research results of this experiment,[10] two different types of GAN were used for image generation: CycleGAN[11] and Pix2Pix.[12] Additionally, we introduced WGAN-GP,[13] which is known for its generation stability, to further optimize the method of generating sample images for the intelligent processing system. The generated results are shown in Figs. 10 to 12. From these figures, it can be inferred that, under different sample

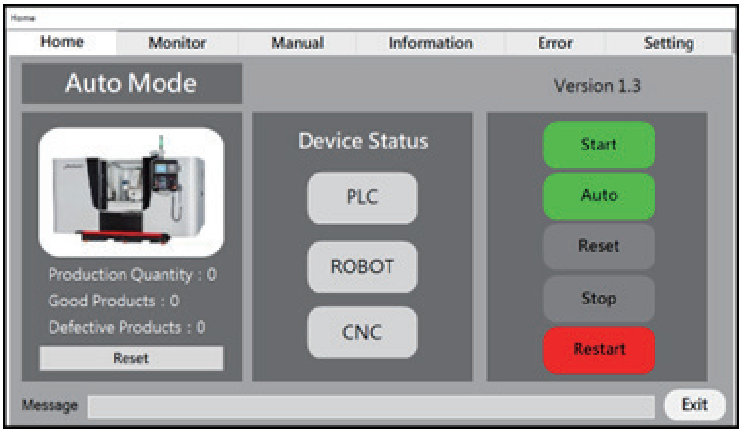

Fig. 6. (Color online) User interface.

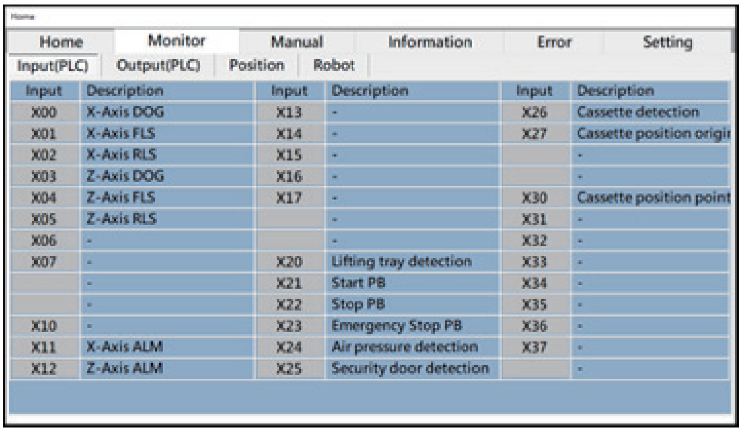

Fig. 7. (Color online) PLC real-time monitoring interface.

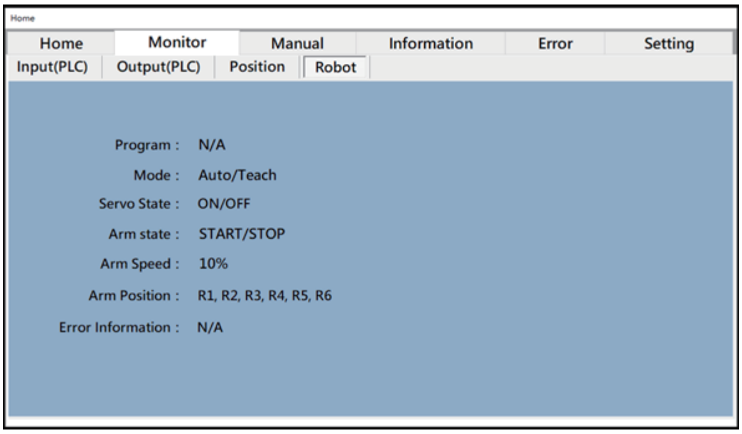

Fig. 8. (Color online) Robotic arm information real-time monitoring interface.

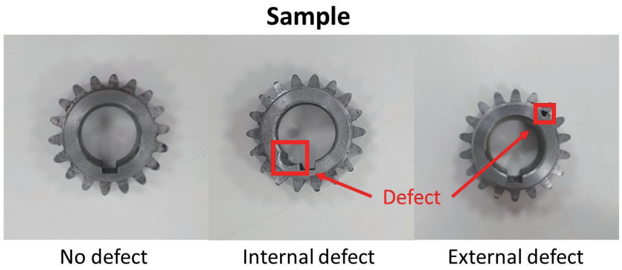

Fig. 9. (Color online) Sample images.

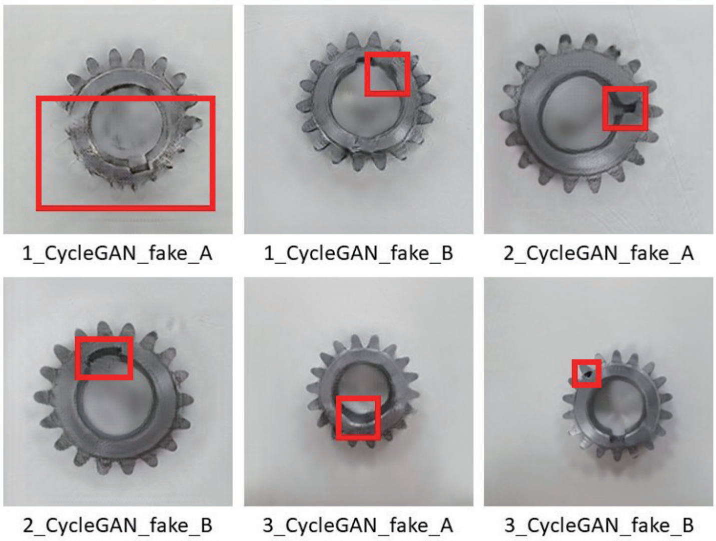

Fig. 10. (Color online) CycleGAN-generated images.

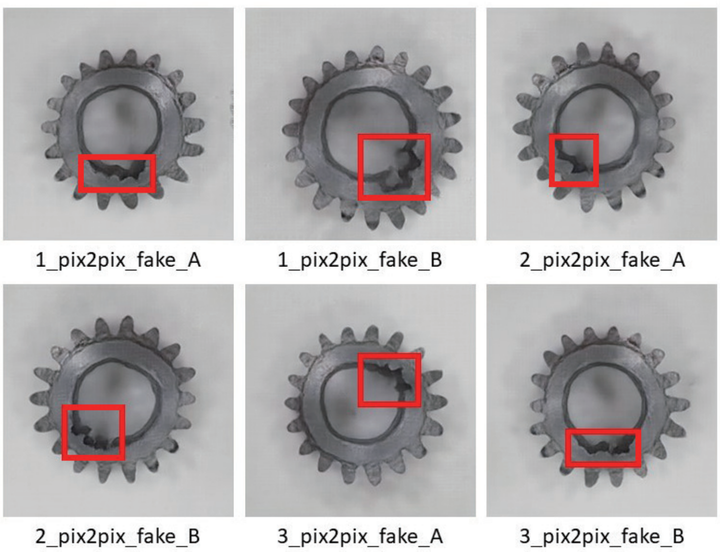

Fig. 11. (Color online) Pix2Pix-generated images.

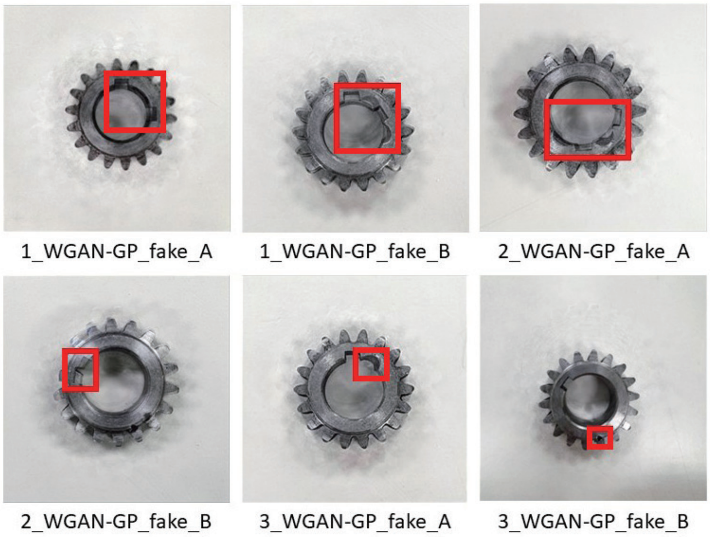

Fig. 12. (Color online) WGAN-GP-generated images.

conditions, such GANs possess a certain degree of repeatability and feasibility. The purpose of generating these images is to augment the sample dataset in quantity and diversity, thereby reducing the possibility of misinspection in AOI and enhancing the accuracy of defect detection.

2.6 Defect detection experiment

In this inspection experiment, we use the YOLOv7 model for defect detection. The objective is to apply the well-trained model to real-world production environments referring to the experimental results, facilitating automated defect detection. YOLOv7 is an object detection and identification model. It employs re-parameterized techniques to replace the original modules and uses a dynamic label assignment strategy. It has a function that can more efficiently allocate labels to different output layers. The image dataset for this inspection experiment consists of the collected samples described in Sect. 2.4 and the generated images. The quantity of samples is shown in Table 1.

To accurately analyze the results of the defect detection experiment, we use a confusion matrix to measure the classification accuracy of the model. The confusion matrix is an $N \times N$ matrix, where N is the number of categories. Each row of the matrix represents the actual category, and each column represents the category predicted by the model. The values on the diagonal of the matrix represent the number of correct predictions, while the values in other positions indicate incorrect predictions, as shown in Fig. 13. Using the confusion matrix, various evaluation metrics can be calculated, such as precision (1), recall (2), and F1 score (3). Precision is the ratio of samples correctly predicted for a particular category to the total number of samples predicted for that category. Recall is the ratio of samples correctly predicted for a particular category to the total number of actual samples in that category. The F1 score is the harmonic mean of precision and recall.

Table 1
Number of samples in image dataset.

Sample source	Number of samples
Defect-free images	50
Defective images	50
CycleGAN	20
Pix2Pix	20
WGAN-GP	20

Confusion Matrix

	Predicting Defects	Predicting Non-Defects
Defects	True Positive(TP)	False Negative(FN)
Non-Defects	False Positive(FP)	True Negative(TN)

Fig. 13. Confusion matrix.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1\,score = \frac{2TP}{2TP + FN + FP} \tag{3}$$

The metrics above can help us understand the model's performance across different categories, as well as its overall accuracy and bias. By thoroughly analyzing these metrics, we can obtain information about the model's efficiency and reliability for defect detection tasks.

3. Results

To ensure the safety and stability of the automated feeding system, we meticulously break down the system's operational process, as shown in Fig. 14. After undergoing system operation tests and physical disconnections (forcibly removing the RJ45 network cable), the automated feeding system showed no adverse reactions. Moreover, when any equipment was forcibly interrupted, it followed the preset safety procedures to halt operations. Additionally, the user

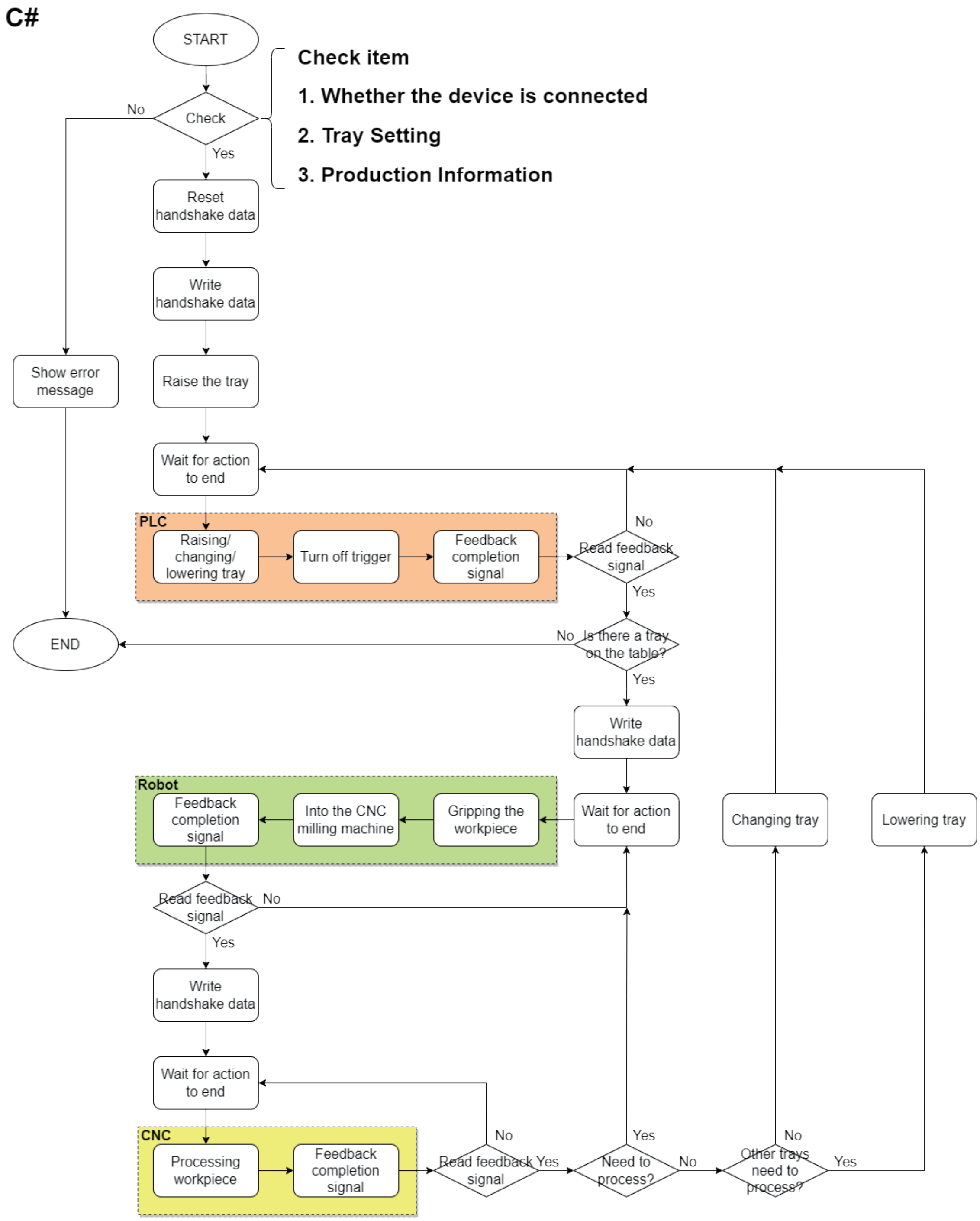

Fig. 14. (Color online) System test flow chart.

interface adopted in this study includes an anomaly information module. This module autonomously logs any abnormal conditions of the automated feeding system, as depicted in Fig. 15.

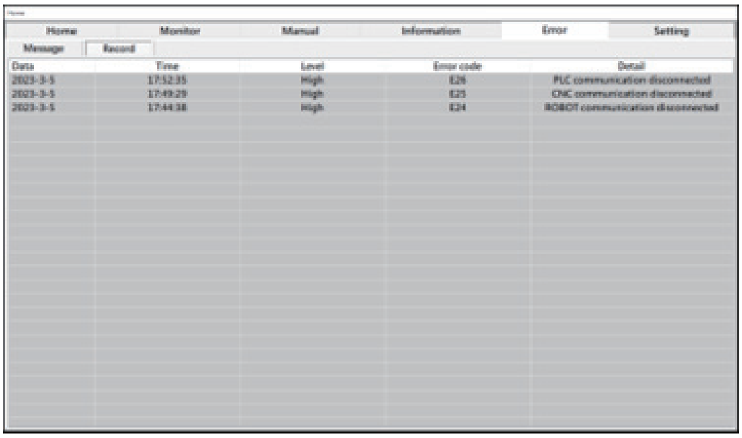

Fig. 15. System test result.

As described in Sect. 2.5, after the training, the YOLOv7 detection model made predictions on the images in the test set. The prediction results were evaluated on the basis of the actual labels, and a confusion matrix was generated to analyze the training results. The confusion matrix from this experiment is shown in Fig. 16. In this figure, it can be seen that the two types of defect category were correctly distinguished and predicted. The F1 score calculated using the confusion matrix is shown in Fig. 17. The results indicate that a better F1 score can be achieved within the interval of 0.2 to 0.6, and the highest F1 score is achieved at a confidence interval of 0.24. The distribution of the F1 score also shows that the precision and recall of this experiment are well balanced.

As shown in Fig. 18, which is derived from the confusion matrix, the detection precision reached 91% when not considering positioning. The Intersection over Union (IoU) threshold, used to distinguish true positives from false positives, commonly has a threshold of 0.5 in practical applications. This means that the IoU between the predicted box and the ground truth box must be at least 0.5 to be considered a true positive detection. In this experimental result, the average precision at the IoU threshold 0.5 (mAP@0.5) reaches 83%. These two precision metrics indicate that the defect detection capability of the intelligent processing system in this study has been significantly improved.

4. Discussion

In this paper, we propose the use of three GANs to generate sample images, aiming to enhance the detection capability of the intelligent processing system. As discussed in the analysis of the results in Sect. 3, it is evident that applying GANs to augment the training database for the YOLOv7 detection model can improve the AOI detection accuracy and training stability. In this study, we did not use BicycleGAN,[14] a subsequent model by the same authors of Pix2Pix, because BicycleGAN incorporates components of the variational autoencoder (VAE). Although VAE makes the training process easier, the generated results tend to be blurry. Therefore, we use Pix2Pix, the predecessor of BicycleGAN, as one of the image generation

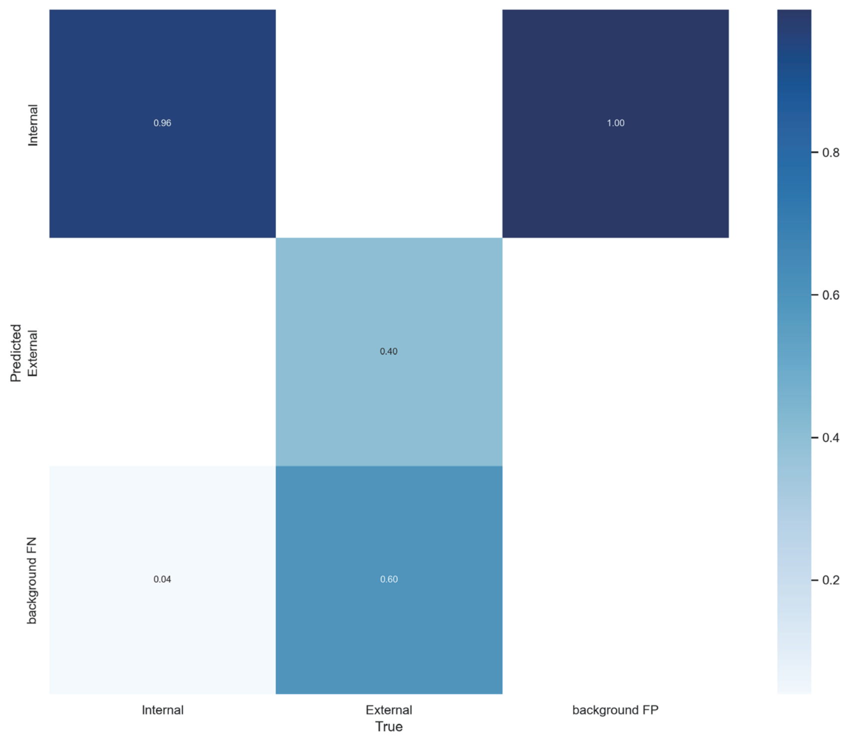

Fig. 16. (Color online) Confusion matrix result.

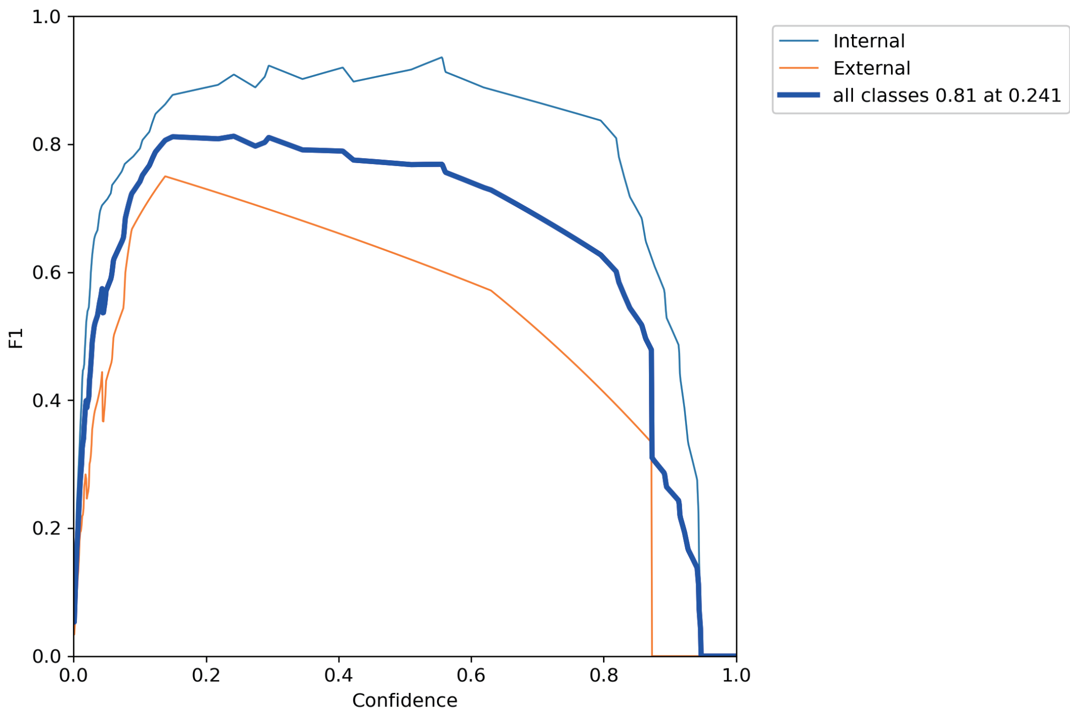

Fig. 17. (Color online) F1 score result.

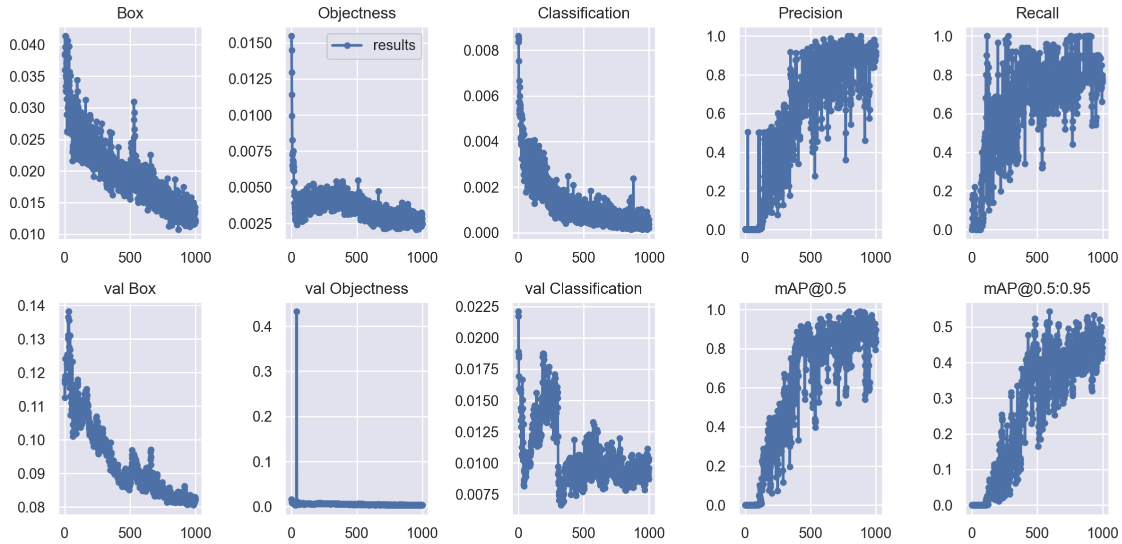

Fig. 18. (Color online) YOLOv7 training results.

models. Both Pix2Pix and BicycleGAN require paired training data. Unlike BicycleGAN, Pix2Pix is not a VAE training structure but uses a GAN training structure for image mapping. As a result, there are only minor variations in resolution, and the clarity of the generated images is enhanced, making them more suitable as samples in this study.

Furthermore, in our hardware design, a seventh axis of a motorized slide table was added to the six-axis robotic arm. While this expands the movement range of the robotic arm, it might increase the risk of collisions with personnel or other objects within the working area. Results of research by Fryman and Bjoern(15) indicate that no robotic arm currently offers a completely safe solution. However, external solutions can address such safety concerns. In our system, we employed a safety monitoring approach to prevent hazardous situations and ensure personnel safety. In conclusion, when adding a motorized slide table as the seventh axis to a six-axis robotic arm, safety and space occupancy must be considered. It is essential to set up an appropriate safety monitoring system and re-evaluate the workspace layout and space utilization efficiency to ensure the safe operation of the robotic arm and maximize production efficiency.

5. Conclusions

We created an intelligent processing system based on practical design in this study. This system uses the Ethernet communication architecture to integrate equipment such as a six-axis robotic arm, PLC, and industrial computer, enabling it to have intelligent functions such as automatic scheduling, production report statistics, and real-time monitoring. Additionally, a GUI was designed to reduce personnel learning costs, operational difficulty, and equipment uptime. Simultaneously, the AOI inspection platform was used for product inspection, endowing the automation system with intelligent inspection capabilities. The results verified that this automated system possesses a satisfactory level of safety and stability.

In this study, we utilized three different GANs and a limited number of images to generate sample images with various defect positions and types. The results showed that the generated images retained their original features, thereby expanding the AOI sample database. By generating new types of defect using GANs, we further reduced the problem of missed detection that arises in the detection model that had not learned about new defects. Ultimately, the training results of the YOLOv7 detection model showed a detection precision of 91%, and the mean average precision (mAP@0.5) reached 83%. This validates the application of GANs in expanding the AOI sample database, addressing the challenge of having too few sample images during the initial production phase, and enhancing both the AOI detection accuracy and the production efficiency of the intelligent processing system.

Author Contributions

Conceptualization, C.-H. W. and M.-H. Y.; methodology, Y.-C. L., M.-H. Y., and C.-H. W.; software, Y.-C. L. and M.-H. Y.; validation, Y.-C. L., M.-H. Y., and C.-H. W.; formal analysis, Y.-C. L., C.-H. W., and M.-H. Y.; investigation, Y.-C. L. and M.-H. Y.; resources, C.-H. W. and M.-H. Y.; data curation, Y.-C. L. and M.-H. Y.; writing—original draft preparation, Y.-C. L. and M.-H. Y.; writing—review and editing, Y.-C. L. and M.-H. Y.; visualization, Y.-C. L. and M.-H. Y.; supervision, M.-H. Y.; project administration, M.-H. Y.; funding acquisition, C.-H. W. and M.-H. Y. All authors have read and agreed to the published version of the manuscript.

Funding

The authors acknowledge the support provided for this study by the Ministry of Science and Technology (MOST 111-2221-E167-025-MY2) of Taiwan.

Conflicts of Interest

The authors declare no conflict of interest.

References

1 M. Rüßmann, M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, and M. Harnisch: Industry 4.0 The Future of Productivity and Growth in Manufacturing Industries. Boston Consulting Group 9 (Scientific Research Publishing, 2015) 54–89. https://www.scirp.org/reference/referencespapers?referenceid=2839429 (accessed 2 May 2024).
2 V. Villani, F. Pini, F. Leali, and C. Secchi: Mechatronics **55** (2018) 248. https://doi.org/10.1016/j.mechatronics.2018.02.009
3 Y. J. Qu, X. G. Ming, Z. W. Liu, X. Y. Zhang, and Z. T. Hou: Int. J. Adv. Manuf. Technol. **103** (2019) 3751. https://doi.org/10.1007/S00170-019-03754-7
4 A. A. R. M. A. Ebayyeh and A. Mousavi: IEEE Access **8** (2020) 183192. https://doi.org/10.1109/ACCESS.2020.3029127
5 C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao: IEEE/CVF Conf. Computer Vision and Pattern Recognition (2023) 7464–7475. https://doi.org/10.1109/CVPR52729.2023.00721
6 Y. Yin, K. E. Stecke, and D. Li: Int. J. Prod. Res. **56** (2018) 848. https://doi.org/10.1080/00207543.2017.1403664

7 Y. Chen, Y. Ding, F. Zhao, E. Zhang, Z. Wu, and L. Shao: Appl. Sci. **11** (2021) 7657. https://doi.org/10.3390/app11167657
8 I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio: Commun. ACM **63** (2020) 139. https://doi.org/10.1145/3422622
9 J.-D. Decotignie: Proc. IEEE **93** (2005) 1102. https://doi.org/10.1109/JPROC.2005.849721
10 Y.-C. Lin, T.-J. Cai, M.-H. Yen, C.-C. Chen, G.-S. Huang, and C.-Y. Peng: 2023 IEEE Int. Conf. Omni-Layer Intelligent Systems (COINS) (2023) 1–5. https://doi.org/10.1109/COINS57856.2023.10189198
11 J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros: Proc. IEEE Int. Conf. Computer Vision (2017) 2242–2251. https://doi.org/10.1109/ICCV.2017.244
12 P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros: Proc. 30th IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2017) 5967–5976. https://doi.org/10.1109/CVPR.2017.632
13 I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville: Adv. Neural Inf. Process. Syst. (2017). https://doi.org/10.48550/arXiv.1704.00028
14 J. Y. Zhu, R. Zhang, D. Pathak, T. Darrell, A. A. Efros, O. Wang, and E. Shechtman: Adv. Neural Inf. Process. Syst. (2017). https://doi.org/10.48550/arXiv.1711.11586
15 J. Fryman and M. Bjoern: ROBOTIK 2012 7th German Conf. Robotics (2012). https://ieeexplore.ieee.org/document/6309480