Automated Robotic Arm System for Handling Plastic Connectors Utilizing Deep Learning and Advanced Lighting Control*

Y.-C. Chen and S.-H. Chen

Abstract— **The manufacturing and assembly of small-batch electronic connectors involve significant labor, but industries are currently facing severe shortages. This study presents an automated grasping system utilizing a six-axis robotic arm that handles plastic components. It integrates vision technology, light source control, and deep learning. The system operates on Linux Ubuntu 18.04 and utilizes the Robot Operating System (ROS) to manage robotic operations.**

Using a distributed ROS framework, the system integrates the robotic arm, industrial camera, adaptive gripper, and machine vision based on deep learning. It accurately recognizes plastic components' front and reverse faces and their orientation and position, achieving a grasping and alignment success rate of 94%. This solution enhances automation in small-batch production, effectively addressing labor shortages.

I. Introduction

With technological advancements and industrialization, numerous companies invest in smart factories to enhance automation, digitalization, and intelligent manufacturing. In contemporary automated assembly lines, robotic arms increasingly replace manual labor, reducing labor costs and heightening production efficiency. These robotic arms perform repetitive tasks, providing stability and high efficiency. Additionally, machine vision and deep learning technologies are widely implemented in industrial applications, facilitating rapid object recognition, shape analysis, and dimensional measurement. Moreover, proper illumination is essential, as lighting conditions significantly influence image quality and object recognition accuracy.

Kim et al. [1] proposed an optimization approach to determine the optimal light intensity in images by adjusting the input voltage. This method achieves ideal illumination conditions for high-contrast printed circuit board (PCB) images. Their optimized algorithm identifies the sharpest regions within the camera's field of view, exploring the feasibility of automated lighting in industrial machine vision applications.

However, even under well-controlled lighting, electronic components can still show variations in appearance. To tackle this issue, Guo et al. [2] combined traditional image processing techniques with deep learning to enable real-time detection of electronic components. Using an industrial camera to capture images from a conveyor belt, they modified the loss function and network structure of the YOLOv4-Tiny object detection model. This modification allowed for the successful classification of components and accurate determination of their coordinates in real-time.

This study is conducted with an industry partner to address real-world assembly line challenges, focusing on handling plastic components, as shown in Figure 1. Due to labor shortages, manufacturers must invest in costly vibratory feeders for component assembly. However, frequent design changes make vibratory feeders less practical. We propose a system integrating light source control and deep learning for flat-laying plastic component grasping to overcome this. Initially, 3D point cloud recognition was explored for stacked components, but recognition accuracy was insufficient due to their small size. Therefore, a 2D recognition approach was adopted, with real-time lighting adjustments in each experiment to enhance accuracy. The system employs machine vision and deep learning to determine component orientation and grasping points, utilizing a six-axis collaborative robotic arm to pick and place flat-laying components onto an assembly jig. This solution effectively addresses the challenges of handling diverse small-batch or prototype components without the need for costly vibratory feeders and automated feeding mechanisms.

II. Research Methodology

A. Object Localization and Orientation Estimation Using Machine Vision

In this study, the robotic arm must determine the world coordinates of scattered plastic components to grip them successfully. Since the positions of the elements are unknown, machine vision technology is employed to locate them in the captured images. The system maps image coordinates to world coordinates through hand-to-eye calibration, ensuring accurate positioning.

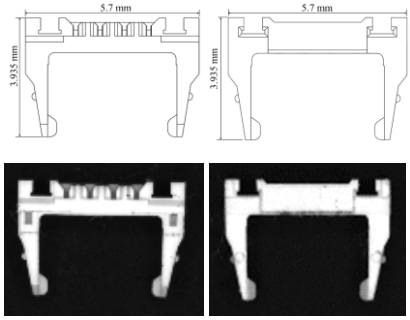

Figure 1. Plastic component (Left) front face (Right) reverse face

The robotic arm (TM5-900) and industrial camera are configured in an Eye-to-Hand setup, where the camera is mounted on a post to capture workspace images. The camera model is Basler acA2500-14uc color camera with 2590x1942

*Resrach supported by NSTC of Taiwan.

Y.-C. Chen is with Department of Mechanical Engineering of National Central University, Taoyuan, 320, Taiwan (corresponding author phone: +886-3-426-7313; fax: 886-3-425-4501; e-mail: ethan@ncu.edu.tw).

S.-H. Chen is with Department of Mechanical Engineering of National Central University, Taoyuan, 320, Taiwan (e-mail: q9906013@gmail.com).

(i.e. 5 Megapixel) pixels along with MK1614-C lens. The images undergo preprocessing, which includes converting to grayscale, binarization, median filtering, and contour edge detection [4, 5, 6].

Then, the centroid method is applied to identify the plastic components' image coordinates. This method calculates the number of enclosed pixels and the total area based on the contour points of the detected object. The centroid coordinates are derived using specific equations, which consider the total number of pixels within the enclosed region and the coordinates of each pixel inside that region, as illustrated in Figure 2.

Once the image coordinates of the plastic component are obtained, the rotation angle is determined using the minimum bounding rectangle method. This technique identifies the contour features from the processed image. A bounding rectangle is formed by extracting the contour's four vertex points and connecting them. The rotation angle of the plastic component is then calculated based on the angle between the longer side of the minimum bounding rectangle and the horizontal axis, as illustrated in Figure 3.

B. Lighting Parameter Adjustment and Control

After identifying the plastic component's position through image processing, the contrast of the detected region is calculated. Since this study focuses solely on the detailed surface features of the plastic component's front and reverse sides, the overall image contrast is not considered. The contrast calculation for the selected image region is defined by

$$C = \frac{I_{\max} - I_{\min}}{\sigma}, \tag{1}$$

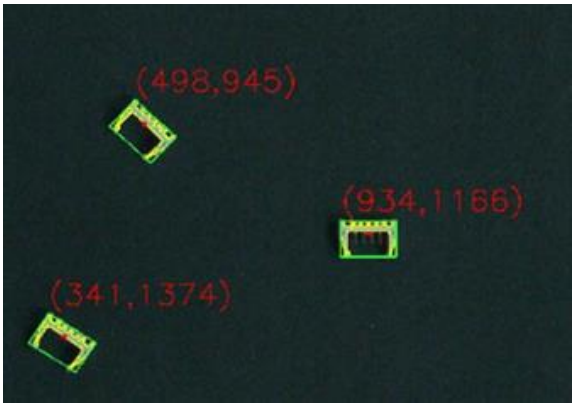

Figure 2. The calculated centroid for the objects.

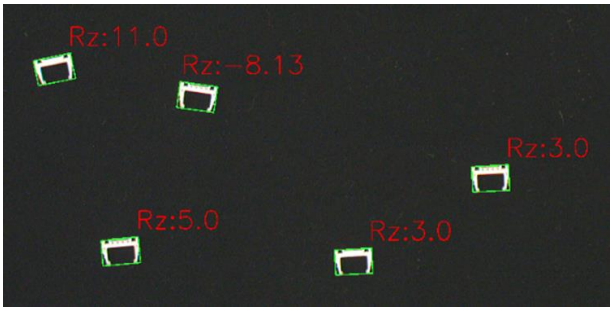

Figure 3. The minimum bounding rectangles for the objects.

where C represents the image contrast of the plastic component; $I_{\max}$ and $I_{\min}$ are the maximum and minimum pixel intensity values in the selected region; and σ is the standard deviation of the pixel intensities. This method effectively quantifies the degree of brightness variation within the area.

The Genetic Algorithm (GA), proposed by John Holland in the 1970s, is inspired by Darwin's theory of evolution. It applies the concept of natural selection, where the fittest individuals survive and reproduce while weaker candidates are eliminated. In GA, variables are represented as chromosomes, which compete and evolve over multiple generations. Through crossover (recombination), mutation, and reproduction, unsuitable chromosomes are discarded, ultimately leading to the optimal solution selection. GA is widely used as an optimization algorithm.

This study applies GA to adjust the brightness levels of the ring lighting with four light channels (Red, Green, Blue, and White). The algorithm optimizes light source parameters based on image contrast to enhance the visibility of plastic component features. Each chromosome represents a set of light source parameters, with four genes corresponding to Red, Green, Blue, and White intensity values. These values range from 0 to 255, forming a discrete search space. The fitness function evaluates the quality of each solution based on image contrast.

Experimental results indicate that higher contrast does not always yield better image quality. The optimal contrast for distinguishing the front and reverse faces of the plastic component falls within the range of 2.8 to 3.2. Excessive or insufficient contrast results in feature loss, affecting recognition accuracy. Therefore, the fitness function is designed as defined in the following to ensure optimal lighting conditions.

$$f(x) = \frac{1}{|C - 3|} \tag{2}$$

C. YOLO-Based Deep Learning Model for Object Recognition

The YOLO (You Only Look Once) deep learning model utilizes supervised learning for training. Introduced by Joseph Redmon et al. [3] in 2016, YOLO was developed as a fast and efficient alternative to RCNN (Region Convolutional Neural Network) and Fast-RCNN (Fast Region Convolutional Neural Network). The YOLO algorithm divides an image into an N×N grid, where each grid cell predicts the bounding box and class label of objects within it, allowing the model to detect object size, location, and category in a single forward pass. This study employs YOLOv4, an improved version proposed by Alex Bochkovskiy et al. [7] in 2020.

This study's target objects are plastic components for electronic connectors, including two appearances, front and reverse faces, when placed in the working area. One thousand images of plastic components were captured using an industrial camera under different orientations and placements. Among these, 800 images were used for training and 200 for testing. The trained YOLOv4 weights (yolov4.conv.137) were used for model initialization, and training was conducted for

4,000 iterations. Figure 4 shows that the loss function curve and the mean Average Precision (mAP) curve converge at 4,000 iterations, indicating that training iterations are adequate.

The Intersection over Union (IoU) threshold of 0.5 was used to evaluate the model's performance across different training iterations. After 2,000 iterations, the model achieved 88% accuracy, which improved to 96% after 4,000 iterations. Considering accuracy and other evaluation metrics, this study selects the model trained for 4,000 iterations as the final object detection model. The test results of the trained model are illustrated in Figure 5.

C. Hand-Eye Calibration

This study adopts the nine-point calibration method referenced in [8] to determine the transformation relationship between the camera's and robot's coordinate systems. The principle is based on calculating rotation and translation in a 2D plane using at least three non-collinear points. Herein, expanding to nine points reduces calibration errors, improving accuracy. The homogeneous coordinate transformation matrix between the camera and robot base is expressed as follows:

$$\begin{bmatrix} X_{robot} \\ Y_{robot} \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & X_t \\ r_{21} & r_{22} & Y_t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{camera} \\ Y_{camera} \\ 1 \end{bmatrix}, \quad (3)$$

where r_{11}, r_{12}, r_{21}, and r_{22} are components of the rotating matrix; X_t and Y_t denote the X-axis and Y-axis components of the camera's origin represented in the robot's coordinate system. The calibration process involves the following steps: (a) Capturing Calibration Board Images: The camera captures images of a calibration board. Hough Circle Detection is applied to identify the center coordinates of the nine circles on the calibration board, respectively, as shown in Figure 6. Each detected image coordinates are recorded as (X_{camera}, Y_{camera}). (b) Aligning the Robot's End Effector: The robotic arm is actuated to align its end-effector point precisely with the nine circle centers on the calibration board, respectively. Each of the corresponding coordinates of the robot shown on the controller are recorded as (X_{robot}, Y_{robot}). (c) Coordinate Transformation Matrix Computation: The nine sets of image coordinates and the robot's end effector coordinates are formulated using transformation matrix representations, as shown in (4) and (5).

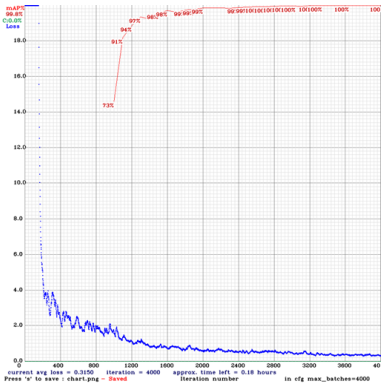

Figure 4. The loss function curve and the mAP curve.

Figure 5. Detection results for the front and reverse faces of the components

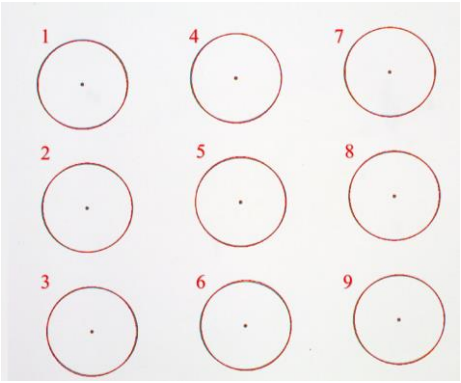

Figure 6. Identified circle centers for hand-eye calibration.

$$\begin{bmatrix} X_{robot1} \\ X_{robot2} \\ \vdots \\ X_{robot9} \end{bmatrix} = \begin{bmatrix} X_{camera1} & Y_{camera1} & 1 \\ X_{camera2} & Y_{camera1} & 1 \\ \vdots & \vdots & 1 \\ X_{camera9} & Y_{camera9} & 1 \end{bmatrix} \begin{bmatrix} r_{11} \\ r_{12} \\ X_t \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} Y_{robot1} \\ Y_{robot2} \\ \vdots \\ Y_{robot9} \end{bmatrix} = \begin{bmatrix} X_{camera1} & Y_{camera1} & 1 \\ X_{camera2} & Y_{camera1} & 1 \\ \vdots & \vdots & 1 \\ X_{camera9} & Y_{camera9} & 1 \end{bmatrix} \begin{bmatrix} r_{21} \\ r_{22} \\ Y_t \end{bmatrix} \quad (5)$$

Using the least squares method, the six unknown parameters are solved from the nine sets of coordinates. The final transformation matrix is computed using MATLAB, and the obtained transformation matrix in this study is presented in (6).

$$\overset{Base}{Camera}\mathbf{T} = \begin{bmatrix} r_{11} & r_{12} & X_t \\ r_{21} & r_{22} & Y_t \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.0667 & -0.0029 & -97.4352 \\ -0.0028 & -0.0663 & -689.0179 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

III. EXPERIMENTAL RESULTS

A. Experimental Setup and System Functionality

This study includes two system functionalities tested in separate experiments. The first experiment examines the adjustment of light source parameters to analyze the relationship between image contrast and a deep learning model. It also compares various parameter settings for the Genetic Algorithm (GA) to optimize lighting conditions. The second experiment integrates a robotic arm with the YOLOv4 model and machine vision for object detection and grasping tasks. This part ultimately evaluates the success rate of object recognition and pick-and-place operations.

The experimental framework, depicted in Figure 7, features a robotic arm mounted on a workbench measuring 83 cm in length, 110 cm in width, and 80 cm in height. An industrial camera is attached to the optical bench, positioned at a working distance of 53 cm from the objects.

B. *Effect of Image Contrast on Deep Learning Model Performance*

This study evaluates the detection capabilities of a deep learning model under various contrast levels and examines how image contrast affects recognition performance. The YOLOv4 model was trained to compare object detection in images with differing contrast levels. We created two test datasets: one comprised of images with a contrast level between 2.8 and 3.2, and another with images outside this contrast range. The exact original grayscale images were used to minimize differences between the two datasets, with only the light source parameters adjusted to control contrast. This ensured that the only variation was in the contrast levels of the plastic components.

Both datasets contained 150 test images, and the recognition results were analyzed and presented in Table I. Observations from Table I indicate that images with contrast values between 2.8 and 3.2 achieved higher precision, recall, and F1-score in distinguishing the front and reverse sides of the plastic component using the YOLOv4 model. Based on these results, the optimal light source parameters were set to maintain contrast levels within this range to improve recognition accuracy.

C. *Genetic Algorithm Parameter Optimization*

This study aims to identify optimal light source parameters that can produce suitable image contrast under various environmental conditions while minimizing Computational time. A Genetic Algorithm (GA) addresses this issue, focusing on finding the optimal balance between achieving the desired image contrast of plastic components and reducing the number of iterations needed for convergence.

Experiments were conducted to investigate the effect of initial population size while keeping other genetic algorithm (GA) parameters constant. Initial population sizes of 5, 8, and 10 were tested under two lighting conditions: well-light and low-light indoor environments. Each experiment was repeated five times for each population size, and the average fitness score, average computation time, and average image contrast were recorded.

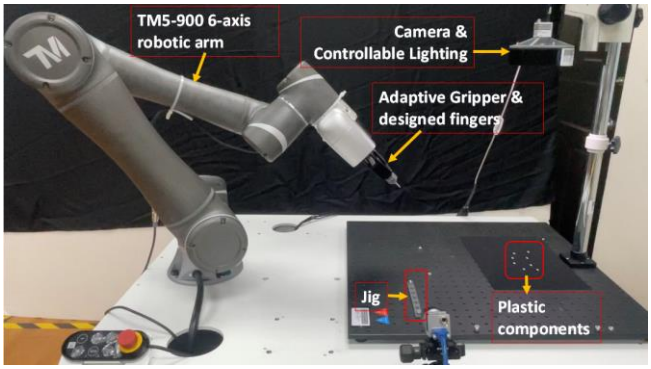

Figure 7. Experimental setup

TABLE I. IMAGE CONTRAST INFLUENCES ON THE MODEL PERFORMANCE

Testing Results	Image contrast	
	2.8< Contrast <3.2	*Contrast < 2.8 or Contrast > 3.2*
TP	346	330
FP	16	37
FN	4	13
Precision	96%	90%
Recall	99%	96%
F1-score	97%	93%
mAP@0.5	98.87%	99.14%

The findings indicate that computation time and average image contrast improved as the initial population size increased from 5 to 8 to 10. A larger initial population provided a broader solution space, reducing the risk of converging to a local optimum and yielding better overall results. However, increasing the population size also led to higher computational complexity. Additionally, the environmental lighting conditions affected both solution accuracy and computational efficiency. Based on these results, an initial population size of 8 emerged as the most suitable choice for this study, striking a balance between efficiency and solution quality.

D. *Plastic Component Placement Experiment*

When the codes developed for the Robot Operating System (ROS) receive information about the grasping point and object classification, the robotic arm begins the process of grasping and placement. To align with the upcoming assembly line production, scattered plastic components need to be picked from the workspace and placed onto a custom-designed jig, as shown in Figure 8. Additionally, it is essential that all objects are positioned in the same orientation. Since the plastic components are initially laid flat in the workspace, they must be reoriented to a vertical position for proper placement in the jig.

During the grasping process, the robotic arm maintains a consistent grasping posture, regardless of whether the component is classified as front-facing or reverse-facing, as illustrated in Figures 8 and 9. Based on the classification results, the robotic arm adjusts its pose and movement trajectory to ensure that all objects are inserted into the jig uniformly. Specifically, if the component is classified as front-facing, the robotic arm rotates 180 degrees around its Z-axis, as shown in Figure 10. If the component is classified as reverse-facing, the robotic arm tilts to a 45-degree angle relative to the work surface, as illustrated in Figure 11. This method ensures the precise and consistent placement of the components for further assembly operations. Figure 12 shows the plastic components inserted into the jig by the robotic arm.

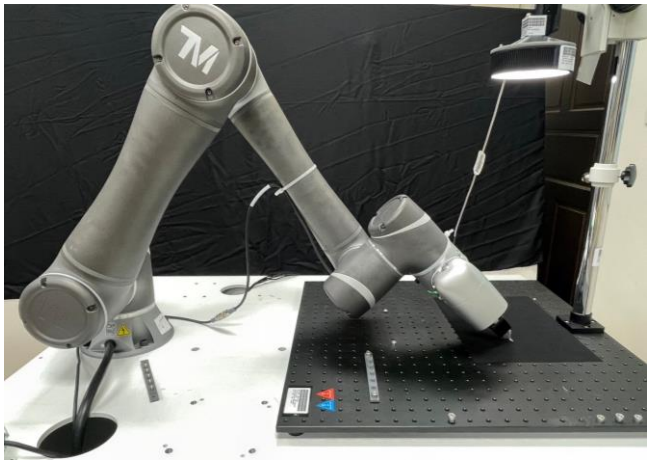

Figure 8. Grasping posture of the robotic arm

Figure 9. Grasping the plastic component

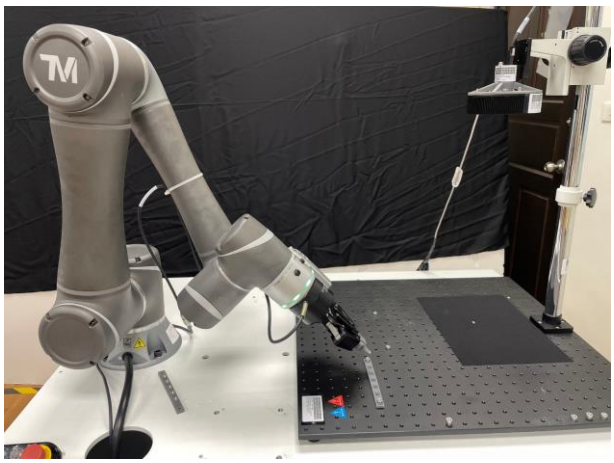

Figure 10. Placing posture of the robotic arm when grasping the component with the front face

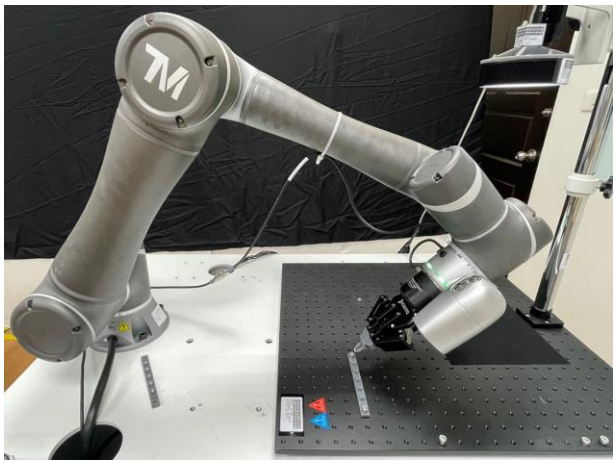

Figure 11. Placing posture of the robotic arm when grasping the component with the reverse face

Figure 12. Inserted plastic components on the jig.

IV. DISCUSSION: EVALUATION AND SUCCESS RATE

This study evaluates the success rate of a robotic arm's recognition and pick-and-place system for handling multiple plastic components. The success rate is defined as the proportion of times the robotic arm successfully grips a plastic component and correctly inserts it into the jig while maintaining a predefined orientation. A failure is recorded when a component is inserted into the jig but at an incorrect orientation due to misclassification. The system's success rate is calculated as the ratio of successful placements to the total number of trials.

For this experiment, ten trials were conducted with varying quantities of scattered plastic components in the workspace, and each trail contains picking-and-placing 5 components. The components were randomly placed, as shown in Figure 12, and the results were summarized in Table II. Throughout 50 trials, the system achieved a success rate of 94%, with an average pick-and-place cycle time of 23 seconds per component.

Failures mostly occurred when components slipped on the fingers during insertion into the jig, leading to misalignment with the jig and unsuccessful insertions. Additionally, excessive time was spent on the robotic arm's movement, particularly when it returned to the starting position after each pick-and-place cycle. Optimizing the robot's trajectory and minimizing unnecessary repositioning can further decrease the cycle time and improve the system's efficiency.

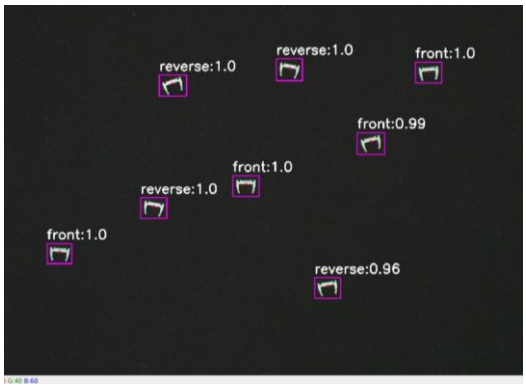

Figure 13. Object detection with front and reverse faces identified.

TABLE II. ROBOTIC GRASPING AND PLACING RESULTS

Experiments	Results		
	Trails	***Success***	***Success Rate***
	50	47	94%

V. CONCLUSION

This research presents a ROS-based distributed framework that integrates an Eye-to-Hand industrial camera, an adaptive gripper, and a collaborative 6-axis robotic arm, along with image recognition capabilities leveraging the YOLOv4 algorithm. The system is designed with an external ring light and a lighting controller that employs a Genetic Algorithm to optimize lighting parameters across various illumination conditions. This optimization process enhances image quality, thereby enabling YOLOv4 to attain high object detection accuracy in both dimly lit and well-lit indoor environments.

Furthermore, the study introduces an automated pick-and-place system specifically for plastic components, utilizing YOLOv4 in conjunction with machine vision. This system proficiently classifies components as either front-facing or reverse-facing while accurately determining their positions. The robotic arm is capable of precisely picking and placing these small components onto an assembly jig, thereby eliminating the necessity for vibratory feeders, which are frequently expensive and impractical for small-batch or prototype production.

In summary, this research offers an efficient, adaptive, and cost-effective solution for automated material handling within electronic connector assembly lines, effectively addressing the challenges encountered by manufacturers operating in flexible production environments.

ACKNOWLEDGMENT

National Science and Technology Council funded this research under grants NSTC 113-2218-E-008 014 and NSTC 113-2221-E-008-072.

REFERENCES

[1] H. Kim, S. Kim, and Y. Cho, “Quick and efficient light control for conventional automatic optical inspection (AOI) systems,” Int. J. Precis. Eng. Manuf., vol. 16, no. 2, pp. 46–54, Feb. 2015.

[2] C. Guo, X. L. Lv, Y. Zhang, and M. L. Zhang, “Improved YOLOv4 tiny network for real-time electronic component detection,” Sci. Rep., vol. 11, no. 1, Nov. 2021.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[4] J. Uusitalo and R. Tuokko, “Setting up task-optimal illumination automatically for inspection purposes,” Proc. SPIE, vol. 6503, 2007.

[5] C. K. Cheng and H. Y. Tsai, “Enhanced detection of diverse defects by developing lighting strategies using multiple light sources based on reinforcement learning,” J. Intell. Manuf., Jun. 2021.

[6] H. Wang, J. Wang, W. Chen, and L. Xu, “Automatic illumination planning for robot vision inspection system,” Neurocomputing, Feb. 2017.

[7] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, “YOLOv4: Optimal speed and accuracy of object detection,” arXiv Preprint, arXiv:2004.10934, Apr. 2020. [Online]. Available: https://arxiv.org/abs/2004.10934

[8] C. Joochim, S. Kaewkorn, and A. Kunapinun, “The 9 points calibration using SCARA robot,” in Proc. Res., Invention, Innov. Congr. (RI2C), Bangkok, Thailand, Dec. 2019, pp. 1–5.