Thiết kế và phát triển hệ thống tay máy phân loại các đối tượng dạng phương tiện dựa trên thị giác máy tính

TÓM TẮT

Việc ứng dụng thị giác máy tính trong các hệ thống phân loại đối tượng dạng phương tiện góp phần nâng cao năng suất, độ chính xác và giảm thiểu sai sót do con người trong các dây chuyền sản xuất công nghiệp. Thông qua việc khai thác các đặc trưng thị giác và kết hợp các kỹ thuật xử lý ảnh, học sâu cho phép nhận dạng tự động và đáng tin cậy các đối tượng phức tạp dưới các điều kiện chiếu sáng, góc quan sát và khoảng cách làm việc khác nhau. Bài báo này trình bày thiết kế và triển khai một hệ thống phân loại tự động tích hợp tay máy robot với mô hình học sâu YOLO nhằm phát hiện và phân loại theo thời gian thực ba nhóm phương tiện gồm ô tô con, xe tải và xe buýt, đồng thời mở rộng khả năng phân loại thông qua nhận dạng màu sắc và mã QR. Hệ thống hỗ trợ ba chế độ nhận dạng gồm YOLO kết hợp màu sắc, nhận dạng dựa trên mã QR và chế độ lai. Kết quả nhận dạng được truyền tới PLC Siemens S7-1200 để điều khiển tay máy robot, trong khi việc giám sát và vận hành được thực hiện thông qua giao diện SCADA. Kết quả thực nghiệm cho thấy hệ thống hoạt động ổn định theo thời gian thực và đạt độ chính xác cao dưới các điều kiện làm việc khác nhau, qua đó khẳng định tính khả thi và hiệu quả của việc tích hợp thị giác máy tính dựa trên học sâu với điều khiển PLC cho các hệ thống tự động hóa công nghiệp.

Từ khóa: *Thị giác máy tính, YOLO, mã QR, Hệ thống tay máy, SCADA.*

Design and Development of a Computer-Vision-Based Robotic Arm System for Sorting of Vehicle-Like Objects

ABSTRACT

The application of computer vision in vehicle-like object sorting systems contributes to improving productivity, accuracy, and reducing human-induced errors in industrial production lines. By exploiting visual features and combining image-processing techniques, deep learning enables reliable and automated recognition of complex objects under varying lighting conditions, viewing angles, and working distances. This paper presents the design and implementation of an automated sorting system integrating a robotic arm with a YOLO-based deep learning model for real-time detection and classification of three vehicle categories—cars, trucks, and buses—while extending classification capability through color recognition and QR code identification. The system supports three recognition modes: YOLO combined with color detection, QR code–based recognition, and a hybrid approach. Recognition results are transmitted to a Siemens S7-1200 PLC to control the robotic arm, while monitoring and operation are performed via a SCADA interface. Experimental results demonstrate that the proposed system operates stably in real time and achieves high classification accuracy under different working conditions, confirming the feasibility and effectiveness of integrating deep learning–based computer vision with PLC control for industrial automation systems.

Keywords: *Computer vision, YOLO, QR code,* Robotic arm*, SCADA.*

1. INTRODUCTION

In the context of the rapidly advancing Industry 4.0 revolution, intelligent automation systems have become essential for enhancing productivity, reducing operational errors, and optimizing resource utilization.[1] One of the most prominent technological trends is the integration of industrial robots with computer vision, enabling machines not only to perform mechanical actions but also to "see" and make autonomous decisions based on real-time visual data.[2]

In product-sorting applications, especially for vehicle-like objects such as model cars, trucks, and buses, accurate classification becomes challenging due to varying illumination, background noise, inconsistent object sizes, and dynamic conveyor movement.[3]

Among deep-learning-based object detectors, the YOLO (You Only Look Once) family has gained significant attention due to its ability to perform real-time object detection while maintaining competitive accuracy.[4] However, many existing studies rely solely on object shape recognition and do not incorporate additional attributes such as color classification or QR code decoding. Furthermore, only a few works have addressed the integration of deep-learning vision systems with industrial PLCs.[5]

To address these limitations, this paper presents the design and implementation of an automatic sorting system that integrates a robotic arm, a YOLO-based computer vision module, and a Siemens S7-1200 PLC. The proposed system detects and classifies three vehicle categories—cars, trucks, and buses—and further analyzes their colors using HSV color-space processing. In addition, a QR code recognition module is incorporated to enhance flexibility. The overall architecture is designed to ensure real-time operation and high classification accuracy. Moreover, the system's modular design allows easy scalability and adaptation to different industrial environments, supporting future integration of additional sensors or sorting criteria.

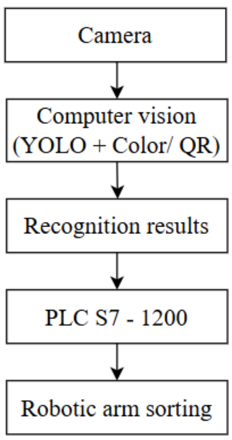

Figure 1. Overview System

Unlike conventional vision-based recognition approaches that focus primarily on algorithmic improvements, this work emphasizes a system-level contribution. The proposed solution integrates a hybrid YOLO–QR recognition architecture with a priority-based decision logic and real-time PLC–SCADA integration, targeting practical deployment in industrial environments. The system supports three recognition modes—YOLO with color detection, QR-based recognition, and a hybrid mode—allowing adaptive operation under different operational scenarios. Recognition results are transmitted to the PLC for real-time robotic arm control, while a SCADA interface enables continuous monitoring and supervision of system behavior.

It should be emphasized that this study does not aim to propose new detection algorithms or modify the internal structure of the YOLO model. Instead, the primary technical contributions of this work lie at the system level and can be summarized as follows:

1) A hybrid computer-vision architecture that combines YOLO-based visual detection and QR-code-based identification for real-time industrial sorting applications;
2) A priority-based recognition strategy in which QR decoding is treated as the primary identification method, while YOLO detection is automatically activated as a fallback mechanism when QR labels are missing, damaged, or unreadable, thereby improving robustness without increasing computational complexity;
3) A complete real-time integration framework connecting the vision system, a Siemens S7-1200 PLC, a robotic arm, and a SCADA interface, enabling closed-loop control and synchronized operation between perception, decision-making, and actuation;
4) An An experimental system-level evaluation demonstrating consistent and stable operation under varying lighting conditions and continuous conveyor motion in real industrial operating environments.

These contributions position the proposed work as a practical and scalable system-level solution for intelligent industrial sorting rather than an algorithm-level enhancement.

The remainder of this paper is organized as follows: Section 2 reviews related works, Section 3 presents the system design and methodology, Section 4 presents the experimental results, Section 5 provides discussion, and Section 6 provides conclusion.

2. RELATED WORKS

In recent years, the integration of computer vision and deep learning technologies into industrial automation has received significant attention from researchers and engineers. Traditional sensing-based classification methods, such as RGB color sensors, proximity sensors, and photoelectric detectors, are generally suitable only for simple objects and stable operating conditions. However, these methods often fail when facing challenges such as variable illumination, complex backgrounds, or objects with diverse shapes and colors, leading to reduced accuracy and system instability.[6] These limitations have motivated the use of image-processing and deep-learning approaches to enhance recognition performance in real-world environments.[7]

In object detection, the YOLO (You Only Look Once) family has emerged as one of the most widely adopted deep-learning models due to its ability to perform fast and accurate real-time detection. Models such as YOLOv3, YOLOv4, YOLOv5, and the more recent YOLOv8 have been applied to various industrial tasks, including defect detection,[8] agricultural sorting, traffic monitoring, and vehicle recognition. Existing studies show that YOLO-based systems can achieve more than 30 FPS while maintaining high detection accuracy even when objects are partially occluded or captured under low-light conditions.[9] Some research also integrates YOLO with color-space analysis to extend classification capabilities when objects share similar shapes but differ in color.

Additionally, QR code recognition has been increasingly used in manufacturing due to its high storage capacity, robustness to noise, and ease of integration with modern vision systems. Libraries such as ZBar and Pyzbar enable real-time decoding and are widely used in inventory management, traceability systems, and automatic product routing. The combination of visual recognition and QR decoding enhances system flexibility by allowing classification based not only on visual appearance but also on encoded product information.

In industrial control, Programmable Logic Controllers (PLCs) play a critical role in bridging computer vision modules with hardware equipment such as robotic arms and conveyor belts. Several studies have explored integrating computer vision with PLCs—such as Siemens or Mitsubishi—through communication protocols like Modbus/TCP, OPC-UA, or Snap7. These works highlight the importance of reliable and low-latency communication in practical automation systems.[10] However, many existing studies focus on isolated components, such as visual detection alone or robotic control alone, without developing a fully integrated system that combines deep learning, multi-modal recognition, and PLC-based control.

From the reviewed literature, it is evident that a gap remains in designing a comprehensive solution that integrates:

(1) YOLO-based object detection,

(2) Color recognition,

(3) QR code decoding, and

(4) Real-time communication with a PLC to perform robotic sorting.

The present paper aims to address this gap by developing a unified, practical product-sorting system suitable for small and medium-sized automation environments.

3. SYSTEM DESIGN AND METHODOLOGY

It should be emphasized that the novelty of the proposed approach does not stem from individual vision algorithms, but from the way multiple recognition modalities are organized and coordinated within a unified real-time system architecture. The design focuses on reliability, priority-based decision logic, and seamless integration between perception and industrial control.

3.1. System Architecture

The proposed system is designed as a unified real-time industrial sorting framework that integrates perception, decision-making, and actuation within a closed-loop control architecture. The system combines computer vision, industrial control, and robotic manipulation to enable reliable vehicle classification and sorting under dynamic conveyor conditions. A hybrid recognition strategy is employed to enhance robustness by leveraging complementary identification modalities. Real-time communication between system components ensures synchronized operation and continuous monitoring through a supervisory interface. The overall system architecture is illustrated in Figure 2. This integrated design facilitates stable, efficient, and scalable deployment in practical industrial environments.

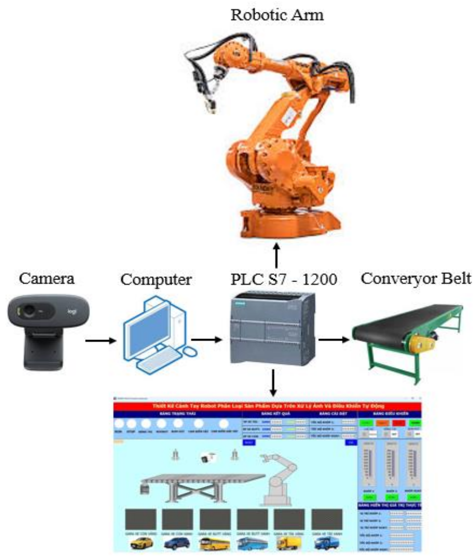

Figure 2. System Architecture

The system, as shown in Figure 2, comprises a camera, a computer running the vision modules, a Siemens PLC S7-1200, a robotic arm, a conveyor belt, and an SCADA interface. The camera captures images of vehicles moving on the conveyor belt and transmits them to the computer for processing.

The vision system adopts a hybrid YOLO–QR recognition architecture with a priority-based decision strategy. The system first attempts to decode the QR code attached to the vehicle due to its high reliability and direct data encoding capability. When QR decoding is successful, the extracted identification information is immediately forwarded to the PLC for real-time control execution.

In cases where the QR code is missing, damaged, or unreadable, the system automatically activates the YOLOv8-based visual recognition module as a fallback mechanism to identify the vehicle type. After vehicle identification, the color information is determined using the HSV color space. This conditional hybrid strategy ensures continuous system operation without manual intervention while maintaining robustness under real-world industrial conditions.

Once the classification results are obtained, the recognized vehicle type and color information are transmitted to the PLC via the Snap7 communication protocol. Based on the received data, the PLC controls the robotic arm to sort vehicles into designated locations on the conveyor system.

The PLC functions as the central real-time controller, coordinating data exchange between the vision system, robotic arm, and SCADA interface. The SCADA system enables real-time monitoring, data visualization, and operator interaction, forming a closed-loop industrial control architecture. This integrated PLC–SCADA framework ensures synchronized operation across all system components, thereby enhancing reliability and practical deployability of the proposed system.

3.2. Image Acquisition and Processing

The image acquisition and processing module serves as the foundation of the entire vision pipeline, ensuring stable input data and enabling downstream recognition algorithms to achieve optimal performance. A camera mounted above the conveyor continuously captures frames of objects as they move through the inspection area. These frames are immediately transferred to the vision-processing program running on the computer, where they are handled sequentially with minimal latency to maintain synchronization between the conveyor motion and the robotic arm's pick-and-place cycle.

To ensure reliable operation, the camera position and capture parameters are carefully configured to minimize motion blur and illumination variations. Continuous frame acquisition allows the system to maintain temporal consistency between consecutive observations, which is essential for stable real-time control. In addition, buffering and frame-handling mechanisms are employed to prevent data loss during high-throughput operation. This design choice contributes to maintaining deterministic system behavior and smooth interaction between perception and actuation components. Overall, the image acquisition stage establishes a reliable data foundation for subsequent recognition and decision-making processes.

Upon receiving a frame, the system first performs a preprocessing step by converting the image from the BGR format into the HSV color space. HSV is selected due to its strong separation between chromatic and luminance components,[13] providing robust color recognition under varying lighting conditions.[11] After conversion, the frame is routed into one of three processing pipelines depending on the selected operating mode.

In the YOLO-based detection mode, the frame is fed into the YOLOv8 model to identify predefined vehicle categories,[12] including cars, trucks, and buses. The detector outputs bounding boxes, class labels, and confidence scores. The bounding box also provides a precise region of interest for subsequent color analysis, effectively reducing the influence of background noise and irrelevant pixels.

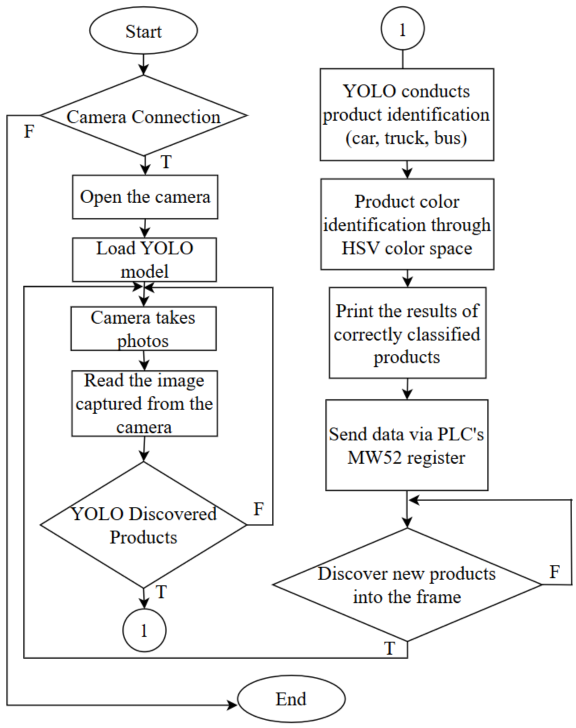

Figure 3. Flowchart of image processing algorithm using YOLO

In the QR-recognition mode, the system utilizes the Pyzbar library to perform comprehensive scanning and decoding of QR patterns within each processed frame.[13] Pyzbar is selected due to its stable recognition capability even under challenging imaging conditions such as drastic lighting variations, partial occlusions, motion blur, or low-resolution inputs. During operation, the system analyzes the shape and position of regions suspected to contain QR codes, verifies the structural validity of the detected patterns, and then applies Pyzbar's decoding algorithm to extract the embedded data. Once a valid QR code is identified, the information encoded within it is immediately adopted as the final classification output. This mechanism allows the system to bypass the YOLO-based processing steps, thereby reducing latency, improving processing speed, and enhancing overall accuracy. Thanks to these advantages, the QR-recognition mode is particularly suitable for automated production lines, inventory inspection, warehouse management, traceability systems, and applications that require encoded data with higher reliability compared to conventional visual recognition. Moreover, this mode contributes to consistent real-time performance while maintaining system stability.

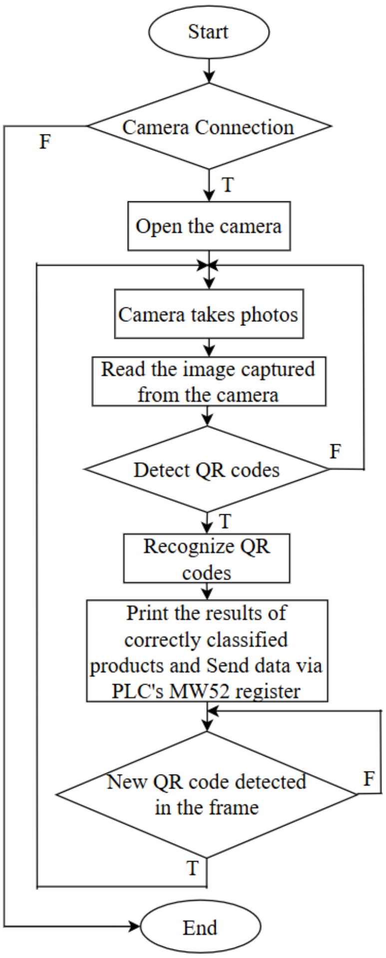

Figure 4. Flowchart of image processing algorithm using QR code

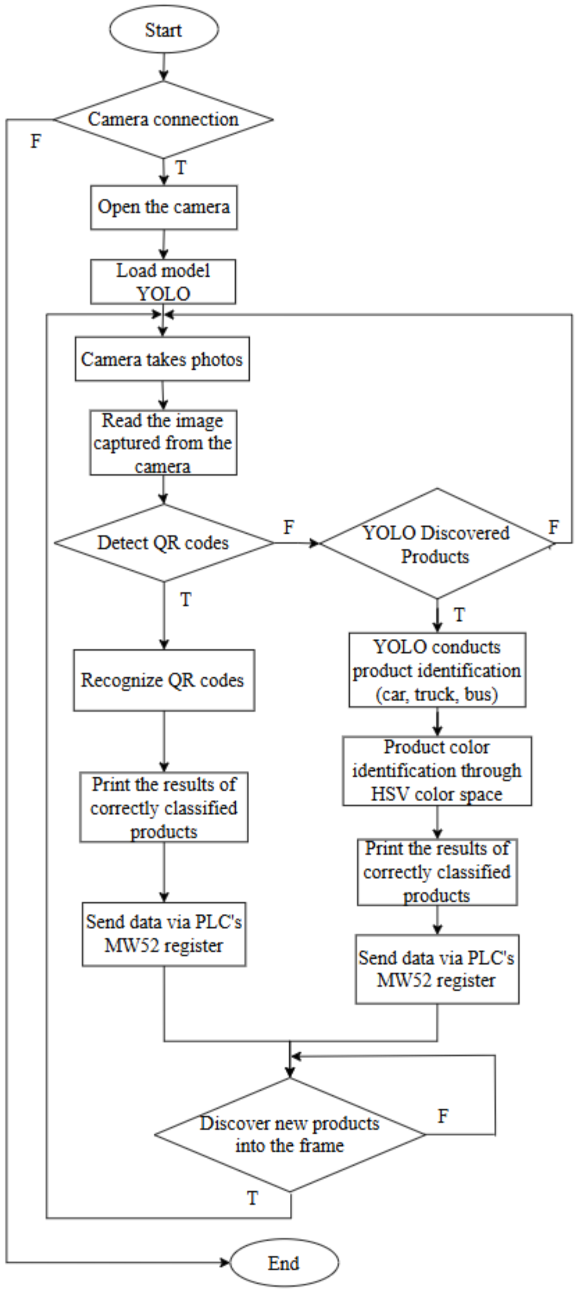

Figure 5. Flowchart of image processing algorithm using YOLO and QR code

When YOLO successfully detects an object, the color-recognition module extracts the central region of the bounding box to reduce the effects of shadows, object edges, and background gradients. The selected pixels are converted into HSV values and compared with predefined threshold ranges, while a voting-based algorithm determines the dominant color with greater stability. To prevent misclassification caused by motion blur or temporary lighting variations, a temporal consistency filter ensures that recognition results remain stable across multiple consecutive frames before being accepted. This combination of spatial filtering and temporal validation significantly improves the reliability of the final color output and ensures robust recognition under fluctuating environmental conditions. In addition, the system applies a normalization step to mitigate brightness fluctuations, helping the HSV values remain consistent during operation. The processing pipeline is further optimized to minimize latency, ensuring that color classification is synchronized with the motion of objects on the conveyor. By integrating these enhancements, the color-recognition subsystem achieves higher resilience against noise, reflections, and rapid illumination changes.

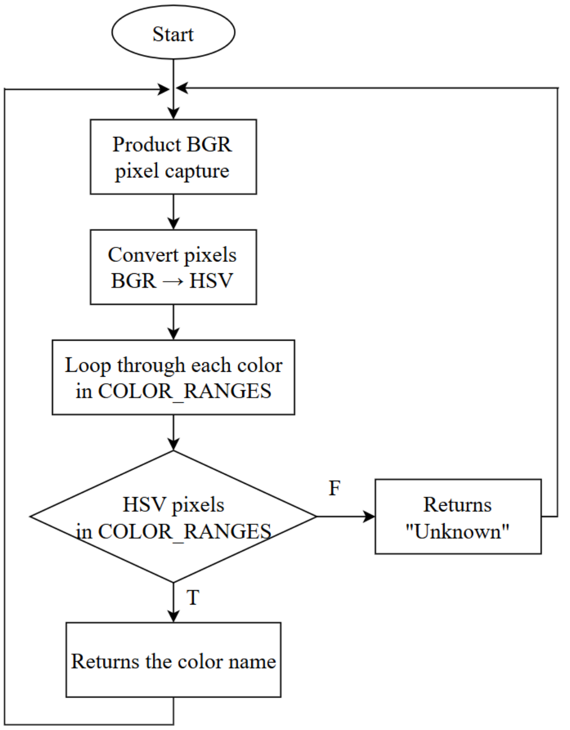

Figure 6. Flowchart of color recognition algorithm using HSV color space

After the object type and color are confirmed, the final classification result is packaged and transmitted to the Siemens S7-1200 PLC via the Snap7 protocol.[15] The PLC interprets the received data and triggers the corresponding robotic actions according to the predefined workflow. This streamlined communication and control pipeline enhances responsiveness, operational robustness, and fault tolerance. As a result, the system maintains stable performance in industrial automation environments, meeting strict requirements for accuracy, speed, and continuous operation. Furthermore, the PLC continuously monitors the communication status and the operational condition of the robotic arm to prevent unexpected failures. In the event of abnormal data or communication delays, the system automatically activates predefined safety responses, ensuring uninterrupted workflow and reliable high-throughput operation. This design approach supports long-term autonomous operation while minimizing downtime and maintenance overhead.

3.3 YOLOv8 Model Training

The YOLOv8 model used in this study was trained using a custom dataset collected directly from the experimental setup. The dataset consists of vehicle-like objects belonging to three categories: car, truck, and bus. All images were captured using a high-resolution camera under adequate lighting conditions, resulting in clear and sharp image quality.

A total of approximately 300 images were collected, including variations in object orientation, position, color, and illumination conditions. The dataset was designed to represent practical industrial scenarios where objects move continuously on the conveyor belt and may appear at different viewing angles and distances. This diversity helps improve the generalization capability of the trained model.

Figure 7. Sample images from the training dataset.

All images were manually annotated using bounding boxes to label the object categories. The annotation process was performed using standard labeling tools, ensuring consistent and accurate ground-truth data for training and evaluation. To improve robustness and reduce overfitting, data augmentation techniques were applied during training. These techniques included image rotation, brightness variation, scaling, and horizontal flipping. Data augmentation increased the effective dataset diversity and improved detection stability under varying environmental conditions.

The YOLOv8s model was selected due to its favorable balance between detection accuracy and computational efficiency for real-time industrial applications. Training was performed for 100 epochs with a batch size of 16 and an input image resolution of 640 × 640 pixels. The training process was executed on a computer equipped with an NVIDIA RTX 3050 GPU, ensuring efficient convergence and stable model performance.

After training, the best-performing model was selected based on validation accuracy and used in the experimental evaluation presented in Section 4.

3.4 PLC Integration and Control Logic

Once the objects have been fully identified and recognized, the resulting classification data are transmitted to the Siemens S7-1200 PLC through the Snap7 communication protocol. This protocol ensures fast, stable, and reliable data transfer between the vision-processing unit and the control system, even under continuous high-frequency operation.[16] Upon receiving the data, the PLC analyzes the classification information and determines the corresponding control commands that the robotic arm must execute according to the predefined sorting logic. Each classification type is mapped to a specific container, storage cell, or designated physical location on the conveyor system, ensuring that every object is delivered to the correct destination with high accuracy. Moreover, the system can dynamically adjust sorting parameters in real-time based on varying production rates or unexpected changes in object flow, further enhancing operational efficiency.

The control logic of the system is developed in Siemens TIA Portal using the ladder diagram programming language, which provides a clear and intuitive structure for handling sequential control operations.[17] Within this environment, additional safety checks, fault-handling routines, and motion-coordination mechanisms are implemented to guarantee stable and collision-free operation of the robotic arm. The machine-vision module communicates with the PLC through an optimized data-exchange routine, enabling minimal latency and ensuring that classification results are synchronized with the conveyor's movement and the robotic arm's pick-and-place cycle. In addition, data logging and automated alerts are integrated to notify operators of abnormal conditions immediately, allowing preventive measures to be taken without halting the production line.

To support supervision and operator interaction, a SCADA interface is integrated into the control architecture. This interface provides real-time visualization of the entire sorting workflow, including the detected object type, current operating mode, communication status, and the live position of the robotic arm.[17] Operators can monitor system performance, adjust parameters, and intervene when necessary, thereby improving operational transparency and enabling efficient system troubleshooting. The combination of PLC-based control and SCADA-based supervision ensures a highly reliable, safe, and easy-to-manage automation system. Additionally, the SCADA interface records historical operational data, allowing engineers to review trends and diagnose recurring issues more effectively. These features collectively enhance the overall maintainability of the system and contribute to long-term operational stability. Additionally, the SCADA interface records historical operational data, allowing review trends, analyze throughput efficiency, and diagnose recurring issues more effectively. The system also supports predictive maintenance by generating performance reports and forecasting potential equipment failures before they occur.

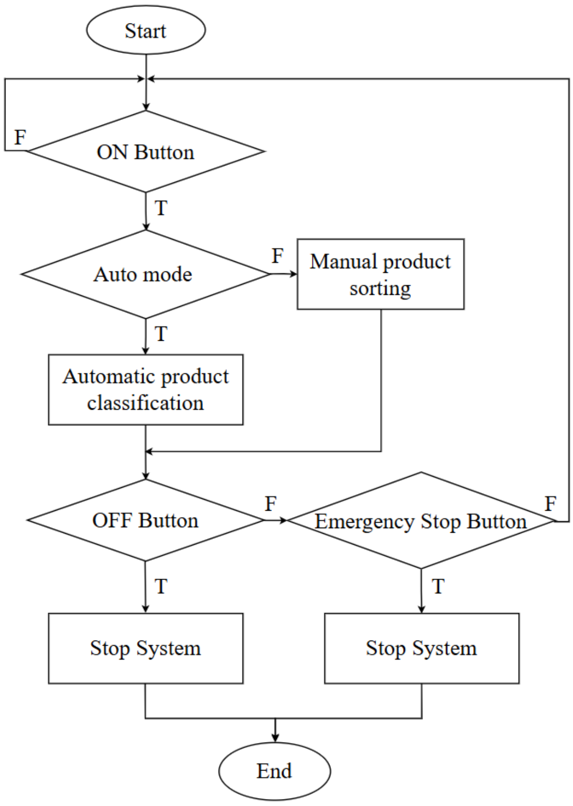

Figure 8. General product classification algorithm flowchart

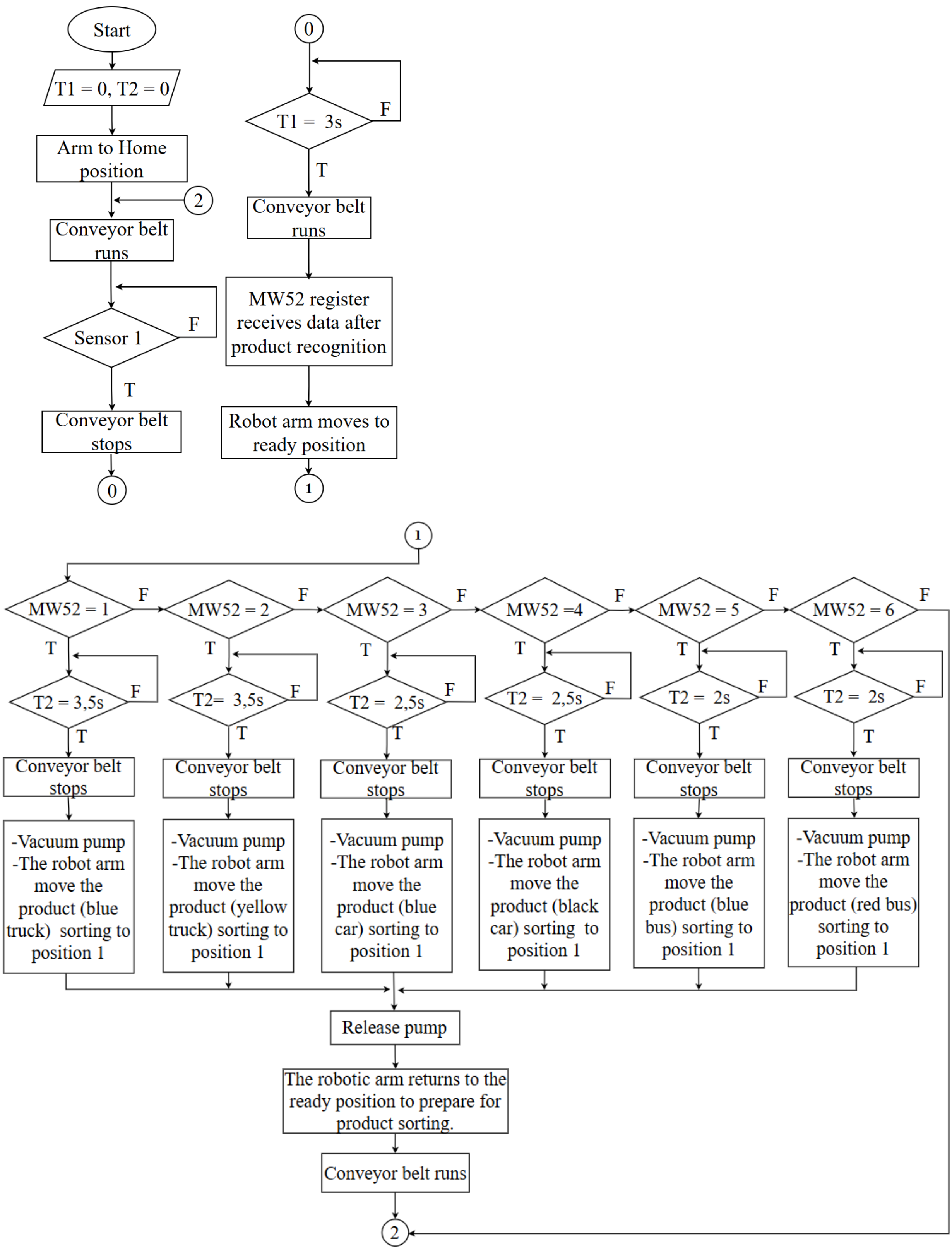

Figure 9. Flowchart of product classification algorithm in Auto mode

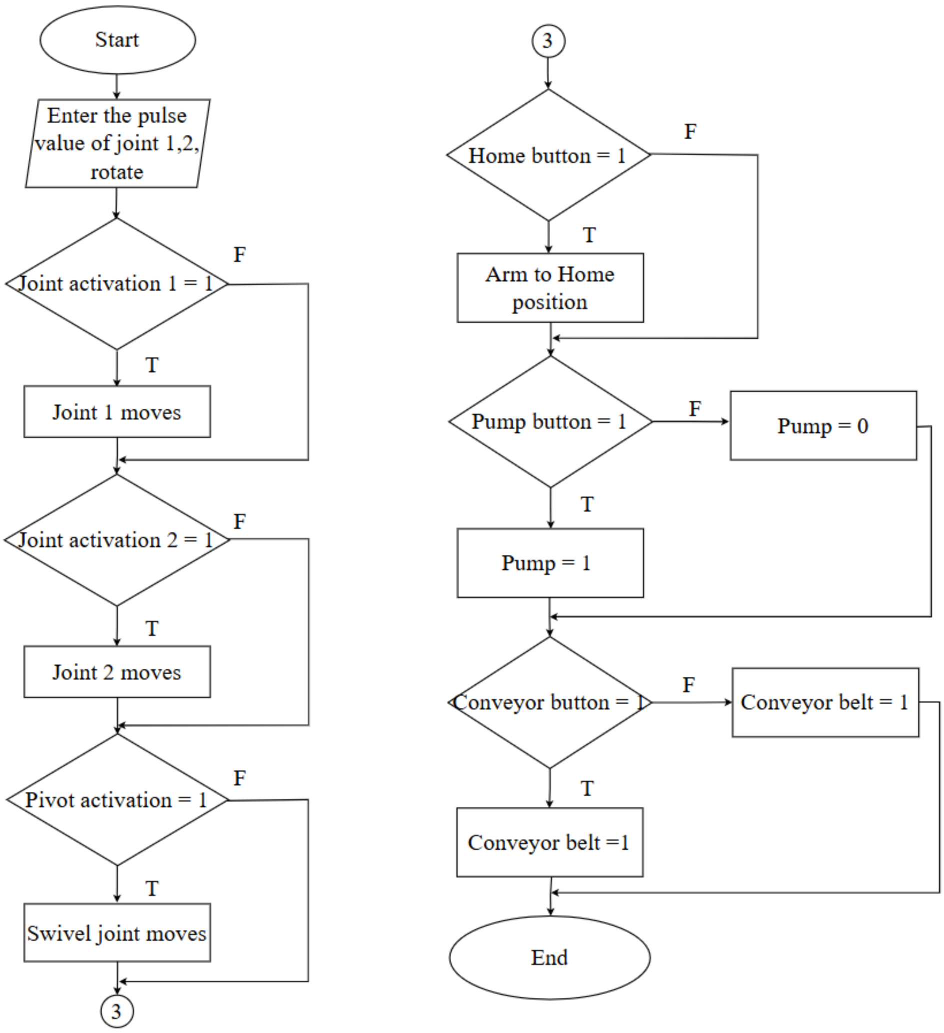

Figure 10. Flowchart of product classification algorithm in Manual mode

3.5. Formulas

3.4.1. YOLO-Based Image Processing Formulations

In the proposed system, YOLOv8 is employed as the object detection module due to its proven balance between detection accuracy and real-time processing capability.[18] YOLO follows a single-stage detection paradigm, in which object localization and classification are performed simultaneously within a single forward pass of the network. This design makes it particularly suitable for real-time industrial applications where low latency is required.[19]

In practice, YOLOv8 processes each input frame captured from the conveyor-mounted camera and outputs bounding boxes, class labels, and confidence scores for detected vehicle-like objects. These outputs provide reliable spatial information that is subsequently used by higher-level system logic, including color recognition, QR prioritization, and PLC-based decision making.

It is important to emphasize that this work does not aim to modify or improve the internal architecture or loss formulation of YOLOv8. Instead, the model is adopted as a mature and stable detection component within a larger system-level framework focused on robustness and real-time integration.

3.4.2. HSV Color Space–Based Image Processing Formulations

For color recognition, the system adopts the HSV (Hue–Saturation–Value) color space, which is widely used in industrial vision applications due to its ability to separate chromatic information from illumination intensity. Compared with the RGB representation, HSV provides improved robustness when lighting conditions vary, which is common in conveyor-based environments.[20]

After an object is detected by the YOLO module, the corresponding region of interest is extracted and converted into the HSV color space. Color classification is then performed by comparing pixel values within this region against predefined threshold ranges for each target color. A voting-based mechanism is applied to determine the dominant color, reducing the influence of noise, shadows, and minor illumination fluctuations.[21]

The HSV-based approach is computationally efficient and well suited for real-time operation.[22] In this study, it serves as a complementary module that extends object classification by incorporating color attributes without increasing system complexity.

3.4.3. Kinematic Equations of the Robotic Arm

The forward kinematic model is introduced to establish a mathematical relationship between joint variables and the end-effector pose, which is essential for defining reachable pick-and-place positions in the sorting workspace. Although real-time trajectory generation is executed by the PLC using predefined motion commands, the kinematic formulation provides the theoretical foundation for workspace analysis, joint limit verification, and collision-free motion planning during system design and validation.

The robotic manipulator under consideration is a three-link revolute configuration (R–R–R), as illustrated in Figure 11. The coordinate frames are assigned following the Denavit–Hartenberg (D–H) convention,[23] and the corresponding parameters are summarized in Table 1. This formulation enables a systematic derivation of the forward kinematic model describing the end-effector pose as a function of joint variables.[24]

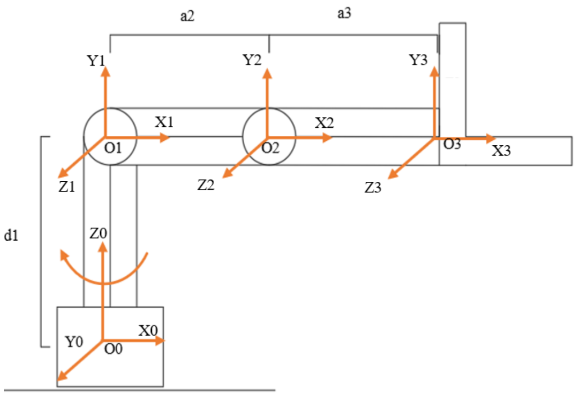

Figure 11. 3-DOF R-R-R robotic arm

Using the standard Denavit–Hartenberg (D–H) convention, each joint of the robotic manipulator is described by a homogeneous transformation matrix A_i, which relates coordinate frame $i$to frame $i-1$. The general form of the D–H transformation matrix is expressed as:[25]

$$A_i = \begin{bmatrix} c_{\theta_i} & -s_{\theta_i} \cdot c\,\alpha_i & s_{\theta_i} \cdot s_{\alpha_i} & a_i \cdot c_{\theta_i} \\ s_{\theta_i} & c_{\theta_i} \cdot c\,\alpha_i & -c_{\theta_i} \cdot s_{\alpha_i} & a_i \cdot s_{\theta_i} \\ 0 & s_{\alpha_i} & c_{\alpha_i} & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where: c is cos and s is sin, θ_i the joint angle, $a_i$is the link length, d_i is the link offset, and $\alpha_i$is the link twist angle.

Based on the mechanical structure of the proposed 3-DOF robotic arm, the corresponding D–H parameters are summarized in Table 1.

Table 1. Denavit–Hartenberg kinematic parameters

Joint	θ_i	α_i	a_i	d_i
1	θ_1	90°	0	L_1
2	θ_2	0°	L_2	0
3	θ_3	0°	L_3	0

Using these parameters, the individual transformation matrices A_1, A_2, and A_3 are obtained. The overall forward kinematics of the manipulator is computed by multiplying the transformations sequentially:

$$T_3 = A_1.A_2.A_3 \quad (2)$$

where:

$$A_1 = \begin{bmatrix} c_{\theta_1} & 0 & s_{\theta_1} & 0 \\ s_{\theta_1} & 0 & -c_{\theta_1} & 0 \\ 0 & 1 & 0 & L_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3a)$$

$$A_2 = \begin{bmatrix} c_{\theta_2} & -s_{\theta_2} & 0 & L_2 c_{\theta_2} \\ s_{\theta_2} & c_{\theta_2} & 0 & L_2 s_{\theta_2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3b)$$

$$A_3 = \begin{bmatrix} c_{\theta_3} & -s_{\theta_3} & 0 & L_3 c_{\theta_3} \\ s_{\theta_3} & c_{\theta_3} & 0 & L_3 s_{\theta_3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3c)$$

After substituting the D–H parameters and simplifying, the resulting homogeneous transformation matrix T_3 can be expressed in the following form:

$$T_3 = \begin{bmatrix} n_x & O_x & a_x & P_x \\ n_y & O_y & a_y & P_y \\ n_z & O_z & a_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

The position of the end-effector with respect to the base frame is given by:

$$Px = c_{\theta_1}.\left(L_2.c_{\theta_{1+}\theta_2} + L_2.c_{\theta_2}\right) \quad (5a)$$

$$Py = s_{\theta_1}.\left(L_3.c_{\theta_{2+}\theta_3} + L_2.c_{\theta_2}\right) \quad (5b)$$

$$P_z = L_1 + L_3.s_{\theta_{2+}\theta_3} + L_2.s_{\theta_2} \quad (5c)$$

These equations explicitly define the Cartesian position of the end-effector as a function of the joint angles and link dimensions.

The orientation of the end-effector is represented by the rotation submatrix of T_3, whose column vectors n, O, and a describe the directions of the local coordinate axes attached to the end-effector. These components are defined as:

$$n_x = c_{\theta_2+\theta_3} . c_{\theta_1} \tag{6a}$$

$$n_y = c_{\theta_2+\theta_3} . s_{\theta_1} \tag{6b}$$

$$n_z = s_{\theta_2+\theta_3} . n_y = c_{\theta_2+\theta_3} . s_{\theta_1} \tag{6c}$$

$$O_x = -s_{\theta_2+\theta_3} . c_{\theta_1} \tag{7a}$$

$$O_y = -s_{\theta_2+\theta_3} . s_{\theta_1} \tag{7b}$$

$$Oz = c_{\theta_2+\theta_3} \tag{7c}$$

$$ax = s_{\theta_1} \tag{8a}$$

$$ay = -c_{\theta_1} \tag{8b}$$

$$az = 0 \tag{8c}$$

The derived forward kinematic equations provide a direct mapping between joint space and Cartesian space. In the proposed system, the computed end-effector position $(P_x, P_y, P_z)$is used for trajectory planning and object positioning during pick-and-place operations, while the orientation components ensure consistent alignment of the gripper when handling vehicle-like objects.

The above expressions describe the motion of the robotic manipulator through the forward kinematic equations, using the D–H parameters derived from the joints of the system.

To ensure safe and collision-free motion during pick-and-place tasks, each joint is restricted to a predefined angular range. These constraints are implemented in the PLC control logic to prevent the manipulator from exceeding its mechanical limits.

Joint limits:

- Joint 1: $\theta_1 \in [0^o, 360^o]$;
- Joint 2: $\theta_2 \in [0^o, 78^o]$;

Joint 3: $\theta_3 \in [0^o, 102^o]$.

3.6. Hardware Setup

The hardware system consists of five main components: an image acquisition camera, a YOLOv8 processing computer, a Siemens S7-1200 PLC, a robotic arm, and a conveyor belt responsible for vehicle-like objects sorting. Figure 12 illustrates the overall layout of the experimental setup.

A Logitech C270 camera is mounted above the conveyor to capture input images for vehicle recognition, color classification, and QR code detection. The captured frames are then process on the computer to ensure real-time inference performance. The recognition results are transmitted to the Siemens S7-1200 PLC via the Snap7 communication protocol. The PLC processes the signals and controls the robotic arm to pick and sort the products according to the detected results. This hardware configuration ensures stable operation and meets the requirements of the experimental tests.

In addition, the system is designed to maintain robust synchronization between visual processing and mechanical motion, minimizing latency during sorting. All hardware components are connected through a structured wiring system to enhance reliability and reduce electromagnetic interference. The robotic arm operates within a predefined workspace optimized for the conveyor layout, ensuring smooth and collision-free movements. Furthermore, the modular structure of the hardware setup allows for easy expansion and modifications for future experiments or industrial applications.

To support continuous real-time operation, the processing computer is configured to handle concurrent image acquisition and inference tasks without interrupting PLC communication. The communication latency between hardware components is kept minimal to preserve synchronization during high-speed conveyor movement. This coordinated hardware design enables reliable closed-loop control throughout the sorting process.

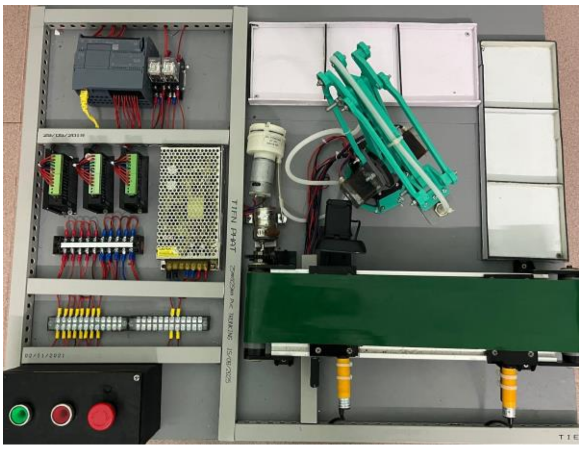

Figure 12. Experimental model

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed system, a series of experiments were conducted on a product classification model consisting of three vehicle types (car, truck, and bus), multiple color variations, and corresponding QR codes. The experiments were carried out under realistic operating conditions, with the conveyor running continuously, the robotic arm performing pick-and-place actions, and a Logitech C270 camera fixed above the observation area. The vision processing module was executed on a computer equipped with an RTX 3050 GPU, ensuring real-time inference throughout the experiments.

4.1 Accuracy of YOLOv8 and Color Recognition

To evaluate the performance of the trained YOLOv8s model[8] in vehicle classification, the system was tested on a dataset consisting of three object categories: car, truck, and bus. The YOLOv8s model was trained using the dataset described in Section 3.3 and then validated under realistic operating conditions. Each category included multiple samples with different sizes, shapes, and colors to ensure diversity in the input data. The recognition results were collected under realistic operating conditions with varying ambient lighting and continuously moving objects on the conveyor belt.

Experimental results show that the trained YOLOv8s model achieves high accuracy in identifying the three vehicle types, with an overall accuracy of approximately 94%. The detected vehicles exhibit relatively consistent accuracy across all categories, indicating that the trained model generalizes well to practical operating conditions.

In addition to vehicle recognition, the HSV-based color analysis algorithm was evaluated using a dataset containing multiple color samples. The experiments indicate that the method provides stable performance, particularly when ambient lighting does not change rapidly. Examples illustrating the vehicle detection and color classification processes are presented in Figure 13.

Furthermore, the combined use of the trained YOLOv8s model and HSV color analysis enables the system to classify both the type and color of vehicles simultaneously, providing a comprehensive solution for automated sorting. This integration also allows real-time monitoring and evaluation, which is crucial for practical industrial applications. Additionally, the system can adapt to minor variations in vehicle orientation and position, maintaining high accuracy under typical conveyor conditions. The results demonstrate the robustness of the proposed approach, making it suitable for continuous operation in real-world manufacturing environments.

Figure 13. Vehicle and color recognition using YOLO and HSV

To evaluate the recognition accuracy, the system wastested by performing 10 consecutiverecognition trials for each vehicle type with its corresponding colors. Table 2 presents the recorded values for all 10

measurements

Table 2. Short-term repeatability results over 10 recognition trials for each vehicle type and color

Vehicle Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) ± Std
Blue Truck	95.5%	94.2%	94.2%	95.8%	95.5%	95.3%	95.5%	95.0%	95.3%	95.3%	95.16 ± 0.55 %
Yellow Truck	95.8%	95.1%	94.2%	94.0%	95.6%	94.0%	95.1%	95.7%	95.7%	95.2%	95.04 ± 0.72 %

Blue Car	96.7%	97.2%	96.1%	97.0%	96.4%	97.1%	96.5%	97.0%	97.2%	96.9%	96.81 ± 0.37 %
Black Car	97.0%	98.1%	97.2%	97.0%	98.0%	98.2%	97.5%	98.0%	97.2%	97.3%	97.55 ± 0.48 %
Blue Bus	96.9%	96.3%	95.9%	96.7%	95.7%	96.5%	96.1%	96.2%	96.0%	96.0%	96.23 ± 0.37 %
Red Bus	96.2%	96.5%	96.1%	95.8%	96.3%	96.1%	96.0%	96.2%	95.5%	96.0%	96.07 ± 0.28 %

Each experiment was repeated ten times under identical operating conditions to evaluate short-term repeatability and operational consistency of the proposed system. The reported mean and standard deviation values describe variability across repeated trials and are not intended for large-scale statistical inference.

The accuracy of vehicle detection and color classification was evaluated over ten repeated recognition trials for each vehicle type and color. The results are illustrated in a line chart (Figure 14), which presents the accuracy obtained in each trial. As observed, the recognition accuracy consistently exceeds 92% across all tested scenarios.

In addition to the mean accuracy, the standard deviation remains low, typically below ±1.0%, indicating that the observed variations between trials are minimal. Such a small deviation suggests that the system's performance is highly stable and that the reported accuracy values are experimentally consistent rather than incidental.

From an engineering perspective, accuracy fluctuations within ±1.0% are generally considered acceptable for industrial vision-based sorting applications, particularly under varying illumination and motion conditions. Therefore, the obtained accuracy levels can be regarded as both valid and practically reliable, confirming that the proposed system maintains high recognition performance and robustness throughout repeated experiments.

Overall, the results demonstrate that the system not only achieves high accuracy but also exhibits strong repeatability, making it suitable for real-time automated vehicle-like object

sorting tasks.

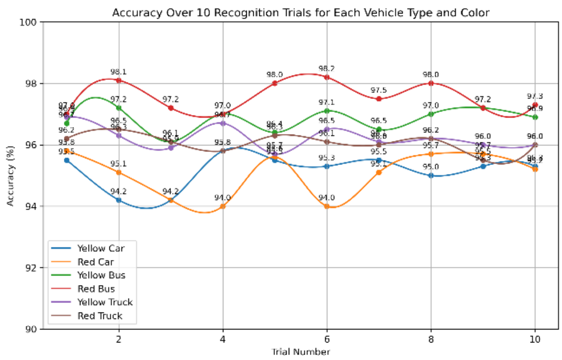

Figure 14. Short-term recognition repeatability over 10 trials for each vehicle type and color

4.2. QR Decoding Performance

The QR recognition module achieved a 98% decoding success rate for 3.2 × 3.2 cm QR labels under lighting-controlled conditions. Most decoding failures occurred when QR codes were bent, blurred, or partially occluded, highlighting the sensitivity of the module to physical distortions. Under varying lighting levels ranging from 300 to 700 lux, the decoding success remained consistently above 96.4%, demonstrating the robustness of the system. The overall classification accuracy was preserved due to the hybrid detection pipeline, which combines YOLO object detection and QR recognition.

To provide real-time feedback during video monitoring, each detected QR code is displayed on the output frame along with a confidence percentage. Since the QR decoding module only outputs successfully decoded codes, the displayed confidence is set to 100% for visual clarity. This visual representation aids operators in quickly identifying and tracking QR codes within the video feed, reducing the risk of oversight. Additionally, real-time highlighting of QR codes allows operators to immediately notice any missing or unreadable codes, improving workflow efficiency and ensuring accurate data collection. Furthermore, this visualization mechanism supports rapid system validation and debugging by enabling intuitive inspection of recognition outcomes during continuous operation.

Figure 15. Recognize QR codes

Furthermore, the actual decoding performance is quantitatively analyzed offline, ensuring that the reported success rates accurately reflect real operational conditions. This combined approach enhances both usability and reliability of the recognition system, particularly in dynamic industrial or laboratory environments. In addition, the system maintains stable decoding performance even when the conveyor speed increases moderately, confirming its suitability for real-time applications. The robustness of this QR module ensures that it can be integrated seamlessly into larger automation pipelines without requiring additional calibration or hardware adjustments.

To evaluate the success rate of QR code decoding under varying lighting conditions, an experiment was conducted in which each QR code was scanned 10 times across different illumination levels, as detailed in the table below:

Table 3. Short-term QR code decoding repeatability over 10 trials under different lighting conditions

Lighting Condition	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) ± Std
300 lux	97.5%	97.8%	98.1%	97.0%	97.2%	97.0%	97.6%	97.9%	97.3%	97.7%	97.51 ± 0.38%
500 lux	98.0%	98.2%	98.0%	97.9%	98.3%	98.0%	98.4%	98.1%	98.2%	98.0%	98.11 ± 0.16%
700 lux	96.5%	96.8%	97.0%	96.4%	96.7%	96.9%	96.6%	96.5%	96.8%	96.7%	96.69 ± 0.19%

Each lighting condition was evaluated over ten repeated trials to assess short-term decoding repeatability and operational consistency. The reported mean and standard deviation values describe variability across repeated trials and are not intended for statistical generalization beyond the tested conditions.

The chart illustrates the QR code decoding success rate over 10 trials under three different lighting conditions: 300 lux, 500 lux, and 700 lux. Each curve represents the decoding accuracy variation across trials, with individual data points labeled by their corresponding percentage values, allowing direct visual comparisonof performance stability under varying illumination levels.

As summarized in Table 3, the mean decoding accuracies are 97.51% ± 0.38%, 98.11% ± 0.16%, and 96.69% ± 0.19% for 300 lux, 500 lux, and 700 lux, respectively. The low standard deviation values (all below 0.4%) indicate that the QR decoding performance is highly stable and exhibits minimal variability across repeated experiments.

The decoding success rates consistently remain within a narrow range from 96.4% to 98.4%, demonstrating that no significant performance degradation occurs under different illumination levels. Minor fluctuations observed between individual trials are primarily attributed to slight variations in QR code alignment, camera focus, or surface reflections rather than inherent system instability.

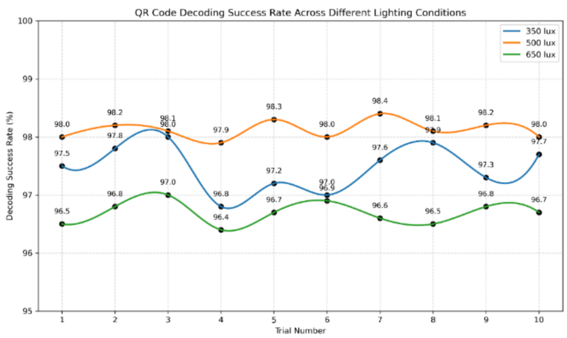

Figure 16. Short-term QR code decoding repeatability over 10 trials under different lighting conditions

Overall, the small standard deviation confirms the robustness and repeatability of the QR recognition module. These results demonstrate that the proposed system can maintain reliable and consistent decoding performance under varying lighting conditions, making it suitable for real-world industrial applications where ambient illumination may fluctuate. Furthermore, this stability enhances the effectiveness of the hybrid YOLO and QR recognition pipeline by ensuring dependable identification during continuous operation.

4.3. Accuracy of the YOLO + QR Mode

In this operating mode, QR decoding is prioritized as the primary method. When a valid QR code is detected, the system immediately adopts the decoded information as the final classification result. If the QR label is missing, damaged, blurred, or occluded, the system automatically switches to the YOLO-based visual detection module, ensuring accuracy comparable to QR-only operation and maintaining high reliability under adverse conditions. This conditional switching mechanism avoids unnecessary visual processing when QR information is available, while guaranteeing uninterrupted recognition in cases where QR decoding fails. The decision logic operates automatically without operator intervention, contributing to stable and autonomous system behavior.

The hybrid priority-based structure achieved an overall accuracy of 97.6%, demonstrating performance equivalent to QR recognition while offering greater robustness. This indicates that the system remains effective even when QR labels fail due to distortion, inconsistent lighting, or rapid object motion, significantly enhancing practical usability in real-world environments. Compared to single-mode recognition systems, the proposed approach dynamically adapts to input quality variations and mitigates performance degradation under challenging conditions.

To further assess this mode, ten consecutive trials were conducted for each vehicle type. Recognition rates were recorded and summarized in Table 4, with a corresponding figure illustrating performance across vehicle categories. All six vehicle types achieved high accuracy ranging from 93.8% to 97.4%, with larger vehicles benefiting from more distinctive shapes and surface areas, while smaller vehicles showed slightly higher variation but still maintained strong performance. These results confirm the scalability of the hybrid strategy across different object sizes and visual characteristics.

These findings reveal patterns related to size, geometry, and color variations, confirming stable performance across changes in illumination, orientation, or conveyor speed. The combined table and figure support a clearer interpretation of recognition behaviors, reinforcing the effectiveness of the QR-first + YOLO-fallback strategy for industrial automation, smart transportation, and real-time inspection applications. The observed consistency further supports the suitability of the proposed system for continuous production-line deployment.

Moreover, the results highlight that integrating multiple recognition strategies not only preserves accuracy but also improves system adaptability, allowing it to handle unexpected variations in product appearance or positioning. This approach ensures seamless operation in dynamic production environments and can be extended to other automated sorting or inspection tasks. Furthermore, the system demonstrates consistent performance over prolonged operation periods, indicating reliability for continuous industrial deployment. The combination of QR and YOLO methods also reduces the need for manual intervention, making the system more autonomous and suitable for large-scale implementation. Overall, the experimental results validate the proposed hybrid framework as a practical and robust system-level solution. The proposed architecture effectively balances recognition accuracy, operational reliability, and real-time industrial deployment requirements.

Table 4. Short-term vehicle recognition repeatability over 10 trials using the hybrid YOLO–QR system

Vehicle Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) + std
Blue Truck	95.8%	95.9%	95.7%	96.0%	95.6%	95.9%	95.8%	96.0%	95.7%	95.9%	95.83 ± 0.13%
Yellow Truck	97.3%	97.1%	97.1%	97.2%	97.3%	97.3%	97.1%	97.4%	97.2%	97.3%	97.23 ± 0.11%
Blue Car	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.64 ± 0.12%

Black Car	96.5%	96.6%	96.6%	96.8%	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.63 ± 0.12%
Blue Bus	96.8%	96.9%	96.9%	96.1%	96.8%	96.0%	95.9%	96.1%	96.8%	96.0%	96.13 ± 0.44%
Yellow Bus	96.0%	96.2%	96.1%	95.3%	96.0%	96.2%	96.1%	96.3%	96.0%	96.2%	96.04 ± 0.28%

Each experiment was repeated ten times under identical operating conditions to evaluate short-term repeatability and operational consistency of the proposed system. The reported mean and standard deviation values describe variability across repeated trials and are not intended for large-scale statistical inference.

The chart shows that the YOLO + QR hybrid system maintains consistently high and stable recognition accuracy, ranging from approximately 94% to 97% across all tested vehicle categories. Among them, the yellow vehicle groups achieve the highest mean accuracy, reaching about 96–97%, while the remaining categories consistently maintain accuracy levels around 94–95%.

As summarized in the experimental results, the reported standard deviation values are extremely low, ranging from ±0.11% to ±0.13% across all vehicle types. Such small standard deviations indicate minimal performance variation over the 10 repeated trials, demonstrating that the system produces highly repeatable and reliable recognition results.

The narrow dispersion of accuracy values confirms that the hybrid YOLO + QR recognition approach is robust against minor disturbances, such as lighting fluctuations, surface color variations, and object positioning changes. No noticeable performance degradation is observed across different vehicle categories, highlighting the system's consistent performance capability.

Overall, the combination of high mean accuracy and very low standard deviation demonstrates that the proposed hybrid system operates stably and effectively, making it well suited for real-time industrial sorting applications, particularly in continuous operation environments where consistency and reliability are critical.

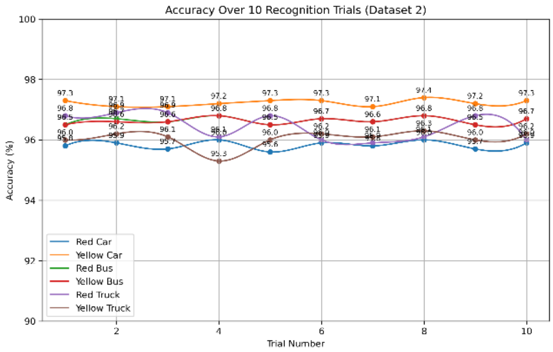

Figure 17. Short-term vehicle recognition repeatability over 10 trials using the hybrid YOLO–QR system

4.4. Processing Time and Real-Time Response

In automated sorting systems, processing time and real-time responsiveness are key factors that determine overall operational efficiency. This section presents the average processing time for each stage in the product-sorting cycle, from the moment an object enters the observation area until the robot completes the pick-and-place operation. This analysis provides an essential evaluation of the system's speed and effectiveness, ensuring that it meets the performance requirements of small- to medium-capacity sorting lines.

In addition, understanding these timing parameters supports the optimization of robot trajectories, minimizes system idle time, and reduces potential bottlenecks during operation. The results serve as an important foundation for improving overall throughput. Furthermore, analyzing the timing of each stage allows for the identification of specific areas where efficiency can be enhanced, such as adjusting conveyor speed or fine-tuning robotic motion. This data also facilitates predictive maintenance and proactive system adjustments, ensuring sustained high performance during continuous operation.

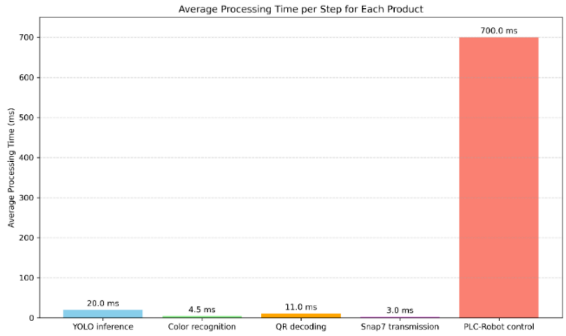

Figure 18. Average Processing Time per Step for Each Product

Based on experimental results, the chart illustrates the average processing time for each step in the vehicle-like objects sorting cycle. Specifically, YOLO inference takes about 20 ms, color recognition 4.5 ms, QR decoding 11 ms, Snap7 data transmission 3 ms, and PLC-Robot control 700 ms. These results meet the real-time requirements for small-to-medium capacity sorting lines while ensuring operational stability and reliability. Moreover, the measured timing breakdown highlights that mechanical actuation dominates the cycle time, indicating sufficient computational margin for future algorithmic or system-level extensions.

4.5. Stability Under Environmental Variations

The system was tested under low (300 lux), normal (500 lux), and high (700 lux) illumination.

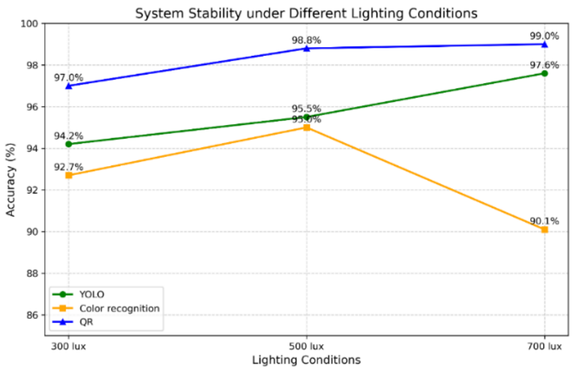

Figure 19. System Stability under Different Lighting Conditions

The figure 19 illustrates the system's recognition accuracy under three different lighting conditions: low (300 lux), normal (500 lux), and high (700 lux). YOLO maintained high accuracy across all lighting levels, with only a slight improvement from 94.2% at 300 lux to 97.6% at 700 lux, indicating minimal sensitivity to illumination changes. The color recognition module exhibited moderate fluctuations, achieving 92.7% at 300 lux, peaking at 95.0% under normal lighting, and dropping to 90.1% under high illumination, suggesting a moderate dependence on lighting conditions. In contrast, QR-based recognition remained highly stable, ranging from 97.0% to 99.0% across all conditions, reflecting its robustness to ambient light due to reliance on contrast rather than overall illumination. Overall, the combined analysis confirms that while YOLO and QR provide consistently reliable performance, the color recognition module may require additional compensation or preprocessing under extreme lighting conditions.

4.6. Quantitative Comparison Between Recognition Modes

To further evaluate the effectiveness of the proposed hybrid recognition approach, a quantitative comparison was conducted between three recognition modes: YOLO-only detection, QR-only recognition, and the proposed hybrid YOLO–QR method.

Table X summarizes the average recognition accuracy and processing characteristics of the three approaches under identical experimental conditions.

Table 5. Comparison of Object Identification Methods

Recognition Method	Accuracy	Processing Time	Robustness
YOLO only	94.0%	20ms	Medium
QR only	98.0%	11ms	High
Hybrid YOLO + QR	97.6%	15ms	Very high

The results indicate that QR-based recognition achieves the highest standalone accuracy due to its deterministic decoding mechanism. However, its performance depends on the presence and quality of QR labels.

The YOLO-based recognition method provides reliable visual detection but shows slightly lower accuracy due to sensitivity to lighting variations and object orientations.

The proposed hybrid YOLO–QR approach combines the advantages of both methods. By prioritizing QR decoding and activating YOLO detection as a fallback mechanism, the system achieves high accuracy while maintaining continuous operation even when QR codes are unavailable or unreadable.

These results demonstrate that the hybrid approach provides a balanced solution in terms of accuracy, robustness, and real-time performance, making it more suitable for practical industrial sorting applications.

4.7. Evaluation of accuracy and error reduction rate of three modes

In automated vehicle recognition systems, the choice of detection strategy directly impacts the accuracy, robustness, and operational reliability of the entire pipeline. To assess the effectiveness of the proposed Hybrid YOLO + QR mechanism, we compared its average recognition accuracy against two individual modes: YOLO-based detection only and QR-based decoding only. This comparison provides clear insight into how combining visual object detection and code-based identification enhances overall system performance, especially under industrial conditions with fluctuating illumination and diverse object appearances.

Figure 20 illustrates the recognition accuracy across the three operating modes. QR

decoding achieves the highest standalone accuracy at 98%, benefiting from its deterministic and noise-resistant nature. YOLO achieves an accuracy of 94%, offering strong visual-feature-based recognition but remaining sensitive to lighting variations, motion blur, and occlusion. The Hybrid YOLO + QR mechanism reaches 97.6%, demonstrating that prioritizing QR recognition while using YOLO as a fallback maintains nearly QR-level performance while significantly improving robustness in cases where QR labels are missing, damaged, or unreadable. These results highlight the practical advantages of integrating multiple recognition strategies, ensuring consistent identification under dynamic production conditions.

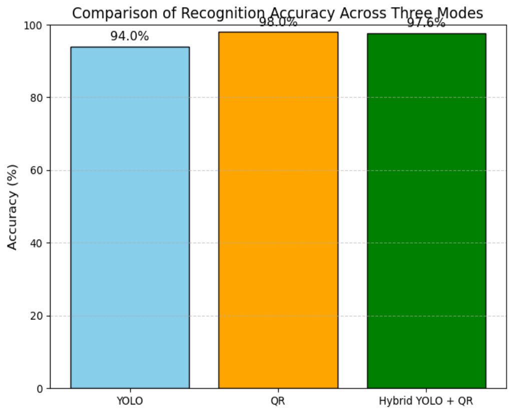

Figure 20. Comparison of Accuracy Across Three Modes

The chart clearly highlights the performance differences among the three approaches. While YOLO provides reliable baseline visual detection and QR yields the highest precision when labels are intact, the hybrid architecture delivers the most balanced and consistent results across varying environmental and operational conditions. Overall, this comparison confirms the hybrid mode as an effective and practical solution for real-world vehicle classification systems, offering both high accuracy and strong resilience to recognition challenges. Additionally, the hybrid approach allows the system to adapt dynamically to temporary failures or occlusions in QR labels, ensuring uninterrupted operation and reducing the need for manual intervention. This capability is particularly valuable in industrial settings where high throughput and minimal downtime are essential. Furthermore, the hybrid strategy improves overall system reliability by minimizing single-point failure risks and supporting continuous operation under non-ideal conditions. The experimental comparison demonstrates that combining complementary recognition mechanisms yields superior robustness compared to standalone methods.

5. DISCUSSION

5.1. Performance Analysis and Discussion

Based on the updated experimental results, the proposed recognition system demonstrates strong performance and reliability across a wide range of operating conditions. The YOLO-based detection module achieves an average accuracy of 94%, maintaining stable processing rates even under changes in illumination or viewing angles, confirming its ability to generalize well to industrial environments. This consistency indicates that the vision pipeline is sufficiently robust for continuous operation on conveyor-based sorting systems.

The color-based recognition component shows slightly lower accuracy due to its sensitivity to lighting variations, but the reduction remains acceptable and does not significantly affect the workflow when combined with YOLO or QR verification. In practice, color information mainly serves as a supplementary attribute rather than a standalone decision factor.

QR decoding proves to be the most robust method, achieving 98% accuracy and maintaining stability thanks to its high-contrast binary structure, which is largely unaffected by environmental changes. This reliability makes QR decoding particularly suitable as the primary identification mechanism in industrial settings.

Across the three operating modes—YOLO, YOLO + Color, and YOLO + QR—the benefits of multi-source fusion become evident. YOLO + QR reaches 97.6% accuracy, outperforming YOLO alone and providing consistently high performance, while YOLO + Color remains helpful for distinguishing visually similar vehicles. In addition, prioritizing QR recognition reduces unnecessary visual inference when label conditions are favorable, thereby improving overall system efficiency. This reduction in redundant computation contributes to more effective utilization of processing resources in continuous operation scenarios.

Overall, the results confirm that a QR-first + YOLO fallback configuration is optimal for industrial applications requiring high reliability and low error tolerance. Remaining limitations include the narrow dataset and lack of outdoor testing; future work should expand data diversity, enhance color recognition, and evaluate deployment on low-power embedded hardware. Such extensions would further improve the system's adaptability and broaden its applicability.

A limitation of this study is the relatively small number of experimental repetitions for each test scenario. Although the results demonstrate consistent system behavior during short-term operation, a larger number of trials would be required to support statistically generalized conclusions. This aspect will be addressed in future work. Long-term testing under continuous industrial workloads is also planned to further validate system robustness.

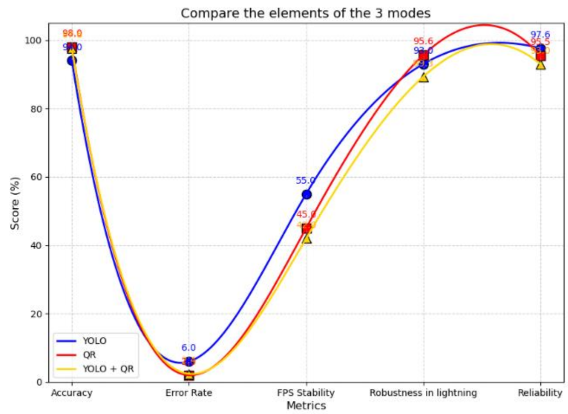

Figure 21: Compare the elements of the 3 modes

5.2. System Scalability and Deployment Considerations

The proposed computer-vision-based robotic arm system is designed with scalability in mind, allowing future expansion to more object categories and higher throughput. The YOLOv8-based detection model can be retrained with additional datasets to support new object types without significant modification to the hardware structure. Furthermore, the modular design of the robotic arm and control system enables flexible integration with other industrial automation components such as conveyor systems and programmable controllers.

Regarding long-term reliability, the system employs stable mechanical structures and industrial-grade electronic components to ensure continuous operation. The vision-based identification approach reduces mechanical contact and minimizes wear, which contributes to improved durability. Regular calibration of the camera and periodic inspection of the robotic arm joints are recommended to maintain stable performance over extended operating periods.

In practical deployment, several constraints must be considered. The system can be implemented on embedded computing platforms such as industrial PCs or edge devices to reduce power consumption and improve portability. Environmental factors such as lighting variation, dust, and vibration may affect detection performance; therefore, controlled illumination and protective enclosures are recommended. In addition, routine maintenance procedures such as cleaning the camera lens and checking mechanical connections are necessary to ensure reliable operation in industrial environments.

6. CONCLUSION

This study presented a system-level solution for automated vehicle-like object sorting by integrating deep-learning-based vision, QR-code identification, and PLC-controlled robotic manipulation. Rather than proposing a new detection algorithm, the contribution of this work lies in the design of a robust hybrid recognition architecture and its real-time integration into an industrial automation pipeline.

This paper presented the design and implementation of an automated product-sorting system that integrates computer vision, deep learning, QR code recognition, and PLC-based robotic control. By combining YOLOv8 for vehicle detection, HSV-based color classification, and a QR decoding module, the system achieves flexible and multi-modal recognition suitable for a wide range of industrial applications. Experimental results demonstrated that the YOLOv8 model provides high accuracy and stable performance under varying lighting conditions, while the QR module maintains exceptional robustness, ensuring reliable classification even in challenging scenarios. The hybrid YOLO + QR mechanism further enhances system reliability, achieving the highest overall accuracy and significantly reducing misclassification compared to individual recognition methods.

The integration of the vision system with the Siemens S7-1200 PLC and a robotic arm enabled real-time sorting operations with minimal latency, fulfilling the requirements of small-to-medium-capacity production lines. The system also maintained stable operation under environmental variations, confirming its suitability for practical deployment.

Although the results are promising, the current system still has limitations, particularly in color recognition under rapidly changing illumination and the restricted scope of tested object categories. Future research may focus on expanding the dataset, enhancing illumination-invariant color-processing techniques, developing adaptive parameter-tuning mechanisms, and deploying the system on embedded platforms to reduce cost and increase portability. Additionally, integrating more advanced multimodal sensing or implementing predictive control for the robotic arm may further improve performance in real-world industrial environments.

Overall, the proposed system demonstrates the effectiveness of combining computer vision and PLC-based control for automated vehicle-like objects sorting, offering a scalable and practical solution for modern intelligent manufacturing systems

REFERENCES

1. A. Nayyar, M. Naved, R. Rameshwar (eds.). *New Horizons for Industry 4.0 in Modern Business*, Springer, Cham, 2023.

2. J. Blasco. Robotic weed control using machine vision, *Biosystems Engineering*, **2002**, *83(2)*, 149–157.

3. Q.-C. Hsu, N.-V. Ngo, R.-H. Ni. Development of a faster classification system for metal parts using machine vision under different lighting environments, *The International Journal of Advanced Manufacturing Technology*, **2019**, *100(9)*, 3219–3235.

4. D.-L. Pham, T.-W. Chang. *A YOLO-based real-time packaging defect detection system*, 5th International Conference on Industry 4.0 and Smart Manufacturing (ISM 2023), Lisbon, Portugal, 2023.

5. M. Maślanka, D. Jancarczyk, J. Rysinski. Integration of Machine Vision and PLC-Based Control for Scalable Quality Inspection in Industry 4.0, *Sensors*, **2025**, *25(20)*, 6383.

6. A. E. Elwakeel. Designing, optimizing, and validating a low-cost, multi-purpose, automatic system-based RGB color sensor for sorting fruits, *Agriculture*, **2023**, *13(9)*, 1824.

7. D. K. Moru, D. Borro. A machine vision algorithm for quality control inspection of gears, *The International Journal of Advanced Manufacturing Technology*, **2020**, *106(1)*, 105–123.

8. A. M. Mezher, A. E. Marble. Computer vision defect detection on unseen backgrounds for manufacturing inspection, *Expert Systems with Applications*, **2024**, *243*, 122749.

9. J. Lee, K.-I. Hwang. YOLO with adaptive frame control for real-time object detection applications, *Multimedia Tools and Applications*, **2022**, *81(25)*, 36375–36396.

10. E. Tapia. Assessing industrial communication protocols to bridge the gap between machine tools and software monitoring, *Sensors*, **2023**, *23(12)*, 5694.

11. S. Sural, G. Qian, S. Pramanik. *Segmentation and histogram generation using the HSV color space for image retrieval*, Proceedings of the 2002 IEEE International Conference on Image Processing (ICIP 2002), Rochester, NY, USA, 2002.

12. M. Bakirci. Enhancing vehicle detection in intelligent transportation systems via autonomous UAV platform and YOLOv8 integration, *Applied Soft Computing*, **2024**, *164*, 112015.

13. N. N. Bokhani, N. K. Modi, H. N. Patel. *Efficiency analysis of barcode and QR code in context to pharmaceutical drug label: an implementation approach with open-source libraries and PCB technology*, International Conference on Deep Learning and Visual Artificial Intelligence (ICDLAI 2024**)**, Singapore, 2024.

14. J. Li, M. Zhao, Z. Qin, R. Yuan, A. Huang, M. Li. Detection and recognition of multiple QR codes based on YOLO_CBAM algorithm, *International Journal of Bio-Inspired Computation*, **2024**, *23(3)*, 179–188.

15. A. S. Ozer, I. Cinar. Real-Time and Fully Automated Robotic Stacking System with Deep Learning-Based Visual Perception, *Sensors*, **2025**, *25(22)*, 6960.

16. S. Şahin, Y. İşler. Microcontroller-based robotics and SCADA experiments, *IEEE Transactions on Education*, **2013**, *56(4)*, 424–429.

17. A. Lobov, J. L. M. Lastra, R. Tuokko, V. Vyatkin. *Modelling and verification of PLC-based systems programmed with ladder diagrams*, 18th IFAC World Congress, IFAC Proceedings Volumes, Milano, Italy, 2004.

18. K. Shizhao, Z. Hu, L. Liu, K. Zhang, Z. Cao. Object detection YOLO algorithms and their industrial applications: Overview and comparative analysis, *Electronics*, **2025**, *14(6)*, 1104.

19. D.-S. Bacea, F. Oniga. Single-stage architecture for improved accuracy real-time object detection on mobile devices, *Image and Vision Computing*, **2023**, *130*, 104613.

20. H.-C. Kang, S. Kim, J. Lee. HSV color-space-based automated object localization for robot grasping without prior knowledge, *Applied Sciences*, **2021**, *11(16)*, 7593.

21. M. Roser, F. Moosmann. *Classification of weather situations on single colour images*, 2008 IEEE Intelligent Vehicles Symposium, Eindhoven, Netherlands, 2008.

22. S. Li, G. Guo. *The application of improved HSV color space model in image processing*, 2nd International Conference on Future Computer and Communication (FCC 2010), Wuhan, China, 2010.

23. M. M. U. Atique, M. A. R. Ahad. *Inverse kinematics solution for a 3DOF robotic structure using Denavit–Hartenberg Convention*, 2014 International Conference on Informatics, Electronics & Vision (ICIEV 2014), Dhaka, Bangladesh, 2014.

24. F. Zhong, G. Liu, Z. Lu, Y. Han, F. Liu, T. Ye. Inverse Kinematics Analysis of Humanoid Robot Arm by Fusing Denavit–Hartenberg and Screw Theory to Imitate Human Motion With Kinect, *IEEE Access*, **2023**, *11*, 67126–67139.

25. J. J. Craig. *Introduction to Robotics: Mechanics and Control*, Third Edition, Pearson Prentice Hall, Upper Saddle River, New Jersey, 2005.