

Design of a Visual Welding Defect Inspection Robot on The Outer Piping Line System based on The CNN Algorithm

Noorman Rinanto¹, Ari Shinta², Ryan Yudha Adhitya³, Mahmoud Elsis⁴

^{1,2,3}Department of Marine Electrical Engineering, Shipbuilding Institute of Polytechnic Surabaya, Surabaya, Indonesia

⁴Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

Article Info

Article history:

Received Dec 15, 2025

Revised Mar 30, 2026

Accepted Apr 16, 2026

Keywords:

Welding Flaw Detection

Deep Learning

Pipeline Inspection

Automated Visual Inspection

ABSTRACT

The investigation focuses on using deep learning techniques to design a testing camera system and identify welding-related issues in outer pipes. The system uses an RGB camera, an ESP32 microcontroller, a DC motor with a rotary encoder, an LED strip, and a small computer that has a graphical user interface. This prototype was made to help with automating the inspection of carbon steel pipe welding, focusing on identifying and grouping different types of defects like porosity, undercut, and too much reinforcement. The results show that the system works in real time, achieving an average of 10.35 frames per second, taking 0.21 seconds to detect something, and having a confidence level of 97.13%. The built-in GUI shows the original image, the edge detection results, and the cropped ROI area, along with motor control and an LED strip. Mechanically, the prototype demonstrates sufficient stability for laboratory research; however, for industrial-scale research purposes, adjustments to materials and dimensions are required. This further develops an automated inspection system based on digital image processing and is designed to be safe, efficient, and compliant with safety standards.

Copyright © 2026 Institute of Advanced Engineering and Science.

All rights reserved.

Corresponding Author:

Noorman Rinanto

Department of Marine Electrical Engineering, Shipbuilding Institute of Polytechnic Surabaya

Jl.Teknik Kimia, Kampus ITS, Sukolilo, Surabaya, Indonesia

Email: noorman.rinanto@ppns.ac.id

1. INTRODUCTION

The pipeline construction industry heavily relies on the welding sector of the manufacturing and energy fields, which are critical for both the development and maintenance of oil and gas pipelines [1, 2]. In economies that export natural resources, these pipeline systems are crucial for the nation's energy exports and greatly improve the country's foreign exchange earnings [3–5]. The safety and reliability of the framework, particularly for the carbon steel pipes used in the oil and gas sectors, depend heavily on the integrity of the welded connections [6]. Currently, the process of visual inspection hinges primarily on the subjective focus and variable reliability of the welding inspectors who are limited by the visual environment and other factors [7]. This necessitates an inspection system that is more accurate, faster, and more reliable in order to maintain compliance with safety regulations.

Advancements in technology have recently allowed deep learning techniques to transform the field of image processing and make automated visual inspections possible. In this area, the Convolutional Neural Networks (CNNs) have created the most impact because of their capacity to hierarchically identify and extract features for optimum recognition and classification of images. These characteristics are of great importance in non-destructive testing (NDT) because of the pressing need to identify and locate defects that compromise the

structural integrity of the materials used.

In welding inspection, deep learning models have been used to detect typical defects, including cracks, porosity, lack of penetration, and other discontinuities. A case in point is a recently carried out study on the RIAWELC dataset, who implemented a ResNet-50-based CNN and obtained results of classification accuracy of 98.75%, 90.25% for GDXray, and 75.83% for a set of low quality private images [8]. On the other hand, Zhu et al. [9] achieved an accuracy of 0.9875 for the detection of surface defects (overlap, spatter, and pores) by proposing a method that used CNN for feature extraction coupled with Random Forest classifier. Moreover, a machine-vision system that uses CNN (ResNet-50) has been developed to achieve about 96.06% average accuracy in the classification of weld defect from radiographic images (crack, porosity, lack of penetration) [10].

Researchers have begun to address deployment beyond just classification. Spruck et al. [11] developed an optical inspection system using laser triangulation imaging combined with a deep neural network to detect weld seam faults; the system was designed to be compatible with Industry 4.0 environments and achieved up to 96.9% classification accuracy. Also, in real industrial radiography applications, semantic-segmentation networks have been used to locate weld defects; one study compared the system's performance against expert human inspectors, showing promising practical utility [12].

Several recent studies have developed robot prototypes for joint weld inspection using computer vision technology, generally using the YOLO method [13–15] and some based on CNN models [16]. These prototypes have been applied to inspect weld joints in the shipping industry [13], civil construction [14], industrial tanks [16], and laboratory-scale weld samples [15]. The subject of inspecting welded joints on pipelines has not been explored by researchers. Most studies on pipeline inspection using robotic technology focus on material corrosion inspection [17, 18], not welded joints. Furthermore, their prototypes are only used to inspect the inside of pipelines. The task of inspecting welded joints on the outside of pipelines has not been widely discussed.

Following previously identified gaps in research, we offer a holistic framework for this study: designing an automated inspection system that integrates a webcam and a CNN-based classification algorithm. While most research focuses on the development of an algorithm, the main focus of our research is the practical design and construction of the inspection device, which involves the arrangement of sensors, the procedures for data collection, and the building of a prototype. With the goal of providing an inspection system that is accurate, reliable, inexpensive, and portable enough to be used in field inspections, we have placed equal emphasis on both the hardware and software components. In considering the CNN model for this study, it was noted that the majority of prior research on inspection robots was based on YOLO [13–15] methods. In addition, CNN detectors such as the Faster R-CNN are considered to be superior to YOLO whenever higher accuracy and more precise object localization are critical, as their region proposal mechanism enables greater feature extraction for small, complex objects. In contrast, YOLO may even be less accurate. This is because it relies on single-stage regression and prioritizes speed over all [19, 20]. Thus, for applications where accuracy is critical, CNN methods are preferred.

This research has several significant contributions, including:

- The design and development of a prototype inspection rig that adheres to iron pipes using magnetic wheels, and allows for automated inspection using an easily mounted standard web camera which is low-cost.
- The integration of a CNN-based classification pipeline into this hardware setup for balancing classification accuracy and inference speed.
- The incorporation of the CNN methodology with the CBAM attention module to improve the system's accuracy in detecting the presence of welding defects such as porosity, excessive reinforcement, undercut, and normal welding.
- The evaluation of the system using a dataset collected in lab- and field-like scenarios to illustrate the real-world applicability of deep-learning-based weld inspection systems for NDT.
- A discussion of the practical challenges (e.g. lighting variation, data noise) and offered suggestions for overcoming these challenges for future improvements and more industrial usability.

The rest of this document is organized as follows: The methods describe the design of the mechanical, electronic, and software hardware systems. In the results and discussion, the prototype testing and evaluation is discussed in detail. Lastly, the conclusion presents the study's findings.

2. METHOD

This section offers an overview of the system, details the hardware and software design, and concludes with the assembly of datasets for welding images.

2.1. System overview

There are three main categories for the components of the system: input, processing, and output. The input devices include a camera, rotary encoder, and push button that allow the system to gather information from both the environment and the user. The data processing components include a tiny PC and an ESP32 microcontroller. The tiny PC will process data coming from the camera, while the ESP32 will process data coming from the rotary encoder and push button. The tiny PC and ESP32 then work together to generate outputs which may be through GUI software, LEDs, LCDs, and DC motors. The proposed system diagram is shown in Figure 1.

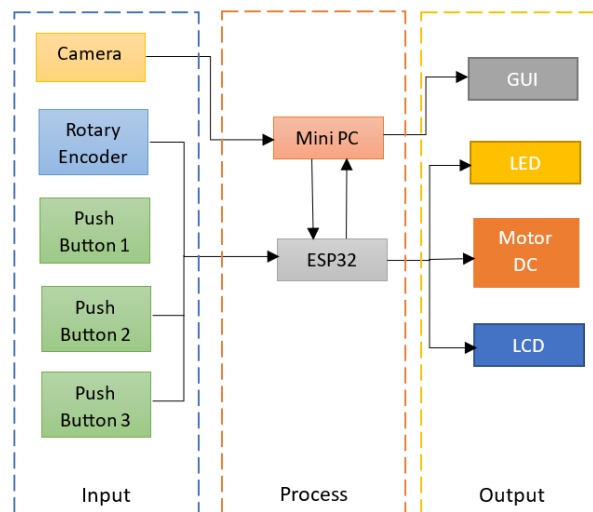


Figure 1. Block Diagram of the Proposed System

The system obtains data from several key components: a webcam, a rotary encoder, and push buttons. The webcam takes pictures of the weld on the steel pipe. The rotary encoder detects the rotation of the motor, which allows for precise control over the positioning of the pipe. Push buttons 1 to 3 are added as extra manual control buttons for specific functions, such as turning on or off the motor, turning on or off the light, or stopping the entire system.

During the processing stage, data from the camera is sent to a tiny PC for processing. The tiny PC acts as a processing unit where it runs the detection and classification algorithms. Meanwhile, the ESP32 microcontroller controls the operation of the actuators (i.e. DC motors, LEDs and LCDs) according to the commands received from the tiny PC and the push buttons. Communication between the ESP32 and the tiny PC can be done either via remote desktop or a serial connection.

The output module displays the inspection results based on the processed data, which is presented through a GUI on the tiny PC. The GUI displays the information extracted from the images, the results of the detection, and the electronic report. The ESP32 also controls an LED strip for additional lighting, a DC motor for the movement of the pipe, and an LCD for the display of information, such as the status of the system or the operating mode. All subsystems work together seamlessly to enable the automated inspection process.

2.2. Hardware Design

This subsection describes the mechanical and hardware design of the prototype pipe welding inspection robot as well as the parts used.

For the system's design, let us focus on the isometric view in Figure 2, which also shows the internal details of the configuration. A camera, which is attached to the front, provides a clear view of the welding zone. This camera takes pictures of pipe welding and sends the data to the model on the tiny PC for further processing. We use the Logitech C270 webcam in this study, which has a maximum resolution of 720 pixels and a frame rate of 30 fps. The LED strip's next positional adjacency to the camera guarantees an even spread of supplementary lighting when the image of pipe welding is taken. A compact PC with the ACER Veriton N4640G brand, equipped with the following specifications: A 6th generation Intel i5 processor, 4GB RAM, and Windows 10 operating system are horizontally arranged on top to facilitate efficient air circulation and cable connectivity.

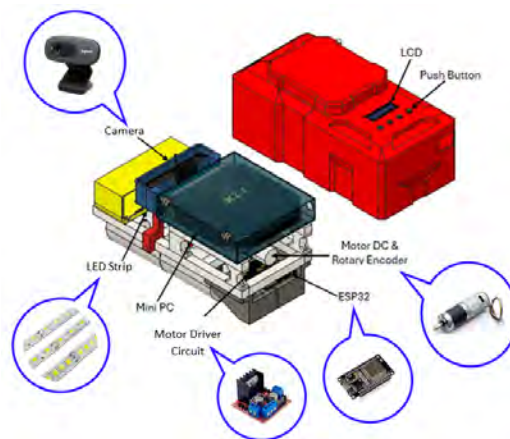


Figure 2. Design of a Mechanical Prototype Robot for Visual Inspection of Pipe Welding and Its Supporting Components.

The ESP32 board and motor driver module are positioned at the base adjacent to the DC motor, facilitating direct connectivity to the actuator. The ESP32 is employed to regulate actuators and receive command data from a tiny PC using serial communication. In the motor driver portion, this study employs the L298N module to control the PG36 DC motor, which is powered by 12 volts and attains a maximum speed of 1500 RPM. The LCD panel and push buttons are affixed to the upper section of the frame to facilitate user operation.

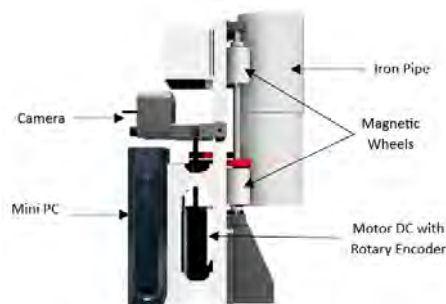


Figure 3. Prototype visual inspection robot for welding, depicted in lateral perspective.

The primary elements of the welding visual inspection robot prototype are depicted in a lateral view as illustrated in Figure 3. A cylindrical motor with a rotary encoder is mounted on the right side of the frame as the main actuator. A camera is located at the front to take images of the area being inspected and a small PC is mounted horizontally on top of the frame as the controller. This design provides a fair distribution of the load and improves access to the electrical parts.

Figure 4 shows the bottom view of the prototype robot used in the study. The figure shows a cylindrical motor located in the center of the system, with frame support on both sides for stability. The front of the frame has the web camera and LED strip, while the back has the cable connections and tiny PC ventilation.

Figure 5 presents an isometric top view of the camera oriented directly towards the inspection region. A black mini PC is situated adjacent to the camera, and an LED strip is aligned parallel to the camera to illuminate the weld joint. This configuration guarantees optimal illumination for the welding zone and enhanced image clarity.

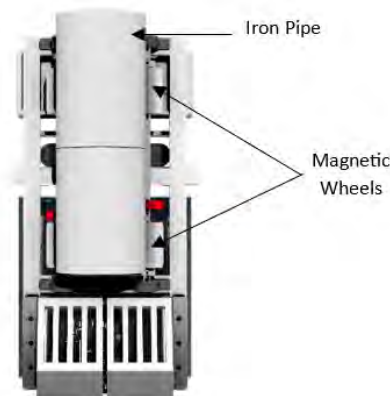


Figure 4. Inferior perspective of the suggested robotic system for pipe weld inspection.

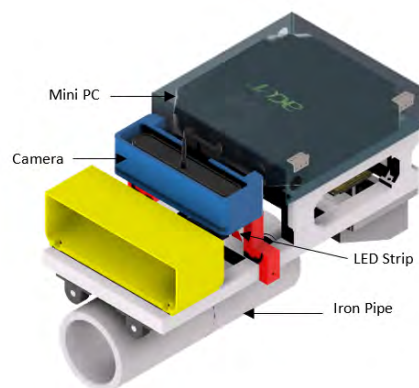


Figure 5. Proposed prototype top view during pipe welding flaw detection.

2.3. Software Design

The graphical user interface (GUI) design on the pipe welding defect detection robot prototype can be visually shown in Figure 6.

The GUI was developed in Python to present the results of welding flaw identification. This application was developed utilizing the PyQt5 library to present classification results in real-time, with three primary views: the original image, the edge-detected image, and the cropped region of interest image. This interface additionally presents critical information including defect kind, classification results, model accuracy, FPS value, and test result photographs. This graphical user interface is linked with the motor and LED control system via serial communication, enabling control of the motor and LED through the GUI buttons. Table 1 delineates the functionalities of all buttons within the GUI.

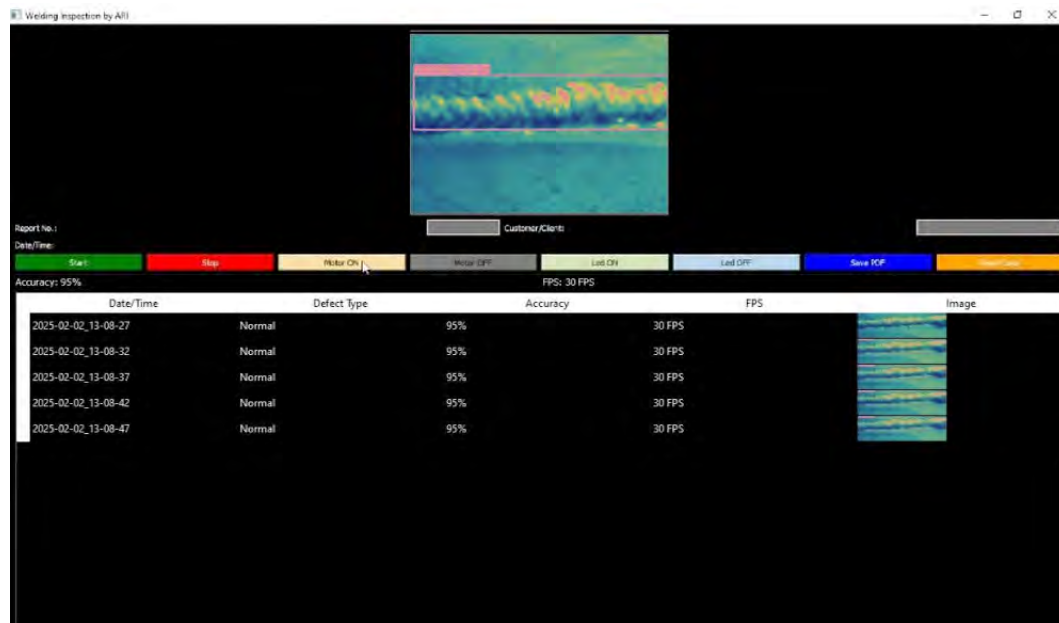


Figure 6. GUI view of the proposed automated visual welding inspection robot.

Table 1. Types of buttons and their functions in the proposed GUI.

No.	Button Colour	Button Label	Description
1	Green	START	Start button to activate the camera and start the classification system
2	Red	STOP	Stop button to deactivate the camera and stop the classification system
3	Cream	MOTOR ON	The button utilized to initiate the motor using the graphical user interface
4	Grey	MOTOR OFF	The button utilized to deactivate the engine using the graphical user interface
5	Light Green	LED ON	The GUI button that activates the LED strip
6	Light Blue	LED OFF	The GUI button for turning off the LED strip
7	Blue	SAVE PDF	The button to save the information displayed in the GUI monitoring table in PDF format
8	Golden Yellow	RESET DATA	The button to reset the classification result information in the monitoring table in the GUI

2.4. Welding Image Dataset

The welding defect images in this investigation were obtained using an RGB webcam affixed to the suggested robot prototype, which was operated manually by an operator. Figure 7 illustrates the procedure for obtaining weld pictures on a steel pipe specimen.

The images of every type of weld defect were captured and cataloged as they appeared under various lighting conditions and from different angles. These images were then incorporated into a training and test dataset. For each image in the dataset, the collection was cropped to 320 x 320 pixels so as to control file size and to match the required input size of the training model. The dataset was divided into four classes corresponding to the defined classes with respect to welding defects: Normal, Undercut, Excessive Reinforcement, and Porosity. Each of the four volume classes contained 500 images.



Figure 7. The procedure for capturing welding photos on sample pipes.

Figure 8 shows an example of a normal pipe weld. It appears that each weld joint on the pipe is neatly arranged with a consistent pattern from start to finish. Next, an example of a welding defect type with the Undercut class is shown in Figure 9. Undercuts are indentations along the weld edge that can reduce the strength of the weld joint [21]. Chen et al. [22] indicate that undercut defects can be caused by excessively high arc currents. Furthermore, these defects can facilitate corrosion.



Figure 8. An example of a normal welding pattern on an iron pipe.



Figure 9. An example of welding defects classified as Undercut.

Figure 10 exemplifies a welding fault image characterised by excessive reinforcement, which will serve as a dataset for this study. Excessive reinforcement is a welding outcome characterised by an excessive height that mimics a mould [23]. This fault is typically attributable to the comparatively low electric current employed during capping. The tolerance limit, as specified by the AWS D1.1 standard, is 3 mm; any deviation exceeding this is classified as an excessive reinforcement defect [24].



Figure 10. An example of welding defects classified as excessive reinforcement.

Meanwhile, Figure 11 is an example of a welding defect image with the type of porosity that will be used in this study. Porosity defects are welding defects in the form of small cavities in the weld metal, which

can be on the surface or inside it [25]. According to the American Welding Society Standard, porosity can be caused by damp or wet field conditions during the welding process. However, sometimes the condition of the metal used is rusty or oily, which can also produce porosity in the welding results [26].

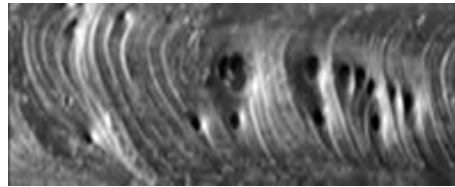


Figure 11. An example of welding defects classified as porosity.

The total number of pipe weld images from this process is 2000 images for the four classes (normal, undercut, excessive reinforcement, and porosity). The concluding stage is manually annotating each image based on their respective categories with the web program Roboflow. The labelled image set is partitioned into three segments: 80% for training, 15% for testing, and 5% for validation. The dataset is now prepared for training a convolutional neural network (CNN) model.

2.5. CNN with Attention Layer

In this section, an explanation of the CNN model architecture with the Convolutional Block Attention Module (CBAM) layer applied in this study is reviewed in depth and comprehensively. The complete CNN and CBAM architecture is detailed in Figure 12. To assist with the training and recognition accuracy, additional steps in the image preprocessing phase are incorporated, which involves normalization and image augmentation. Within image augmentation, modifications such as image flipping, brightness modification, and an adjustment of the grayscale can occur. The procedure commences with an input image and traverses through the convolutional layers and ReLU activation to generate an initial feature map. This feature map is then sent to the Channel Attention module, which identifies the most salient feature channels and sent to the Spatial Attention module to highlight important areas in the image. The feature improvements from both modules are combined and sent to a flattening operation to be processed by a fully connected layer. In the last step, the Softmax function is used to classify the image into one of the defect categories, which can be either porosity, undercut, excessive, or normal. Faster, better model training can occur when the model is able to focus on more important features, thus better improving classification accuracy. Therefore, CNN and CBAM provide the best equilibrium of deep feature extracting and the most efficient focus on the best areas. Comprehensive data are available on each architectural layer's size and proportions in Table 2.

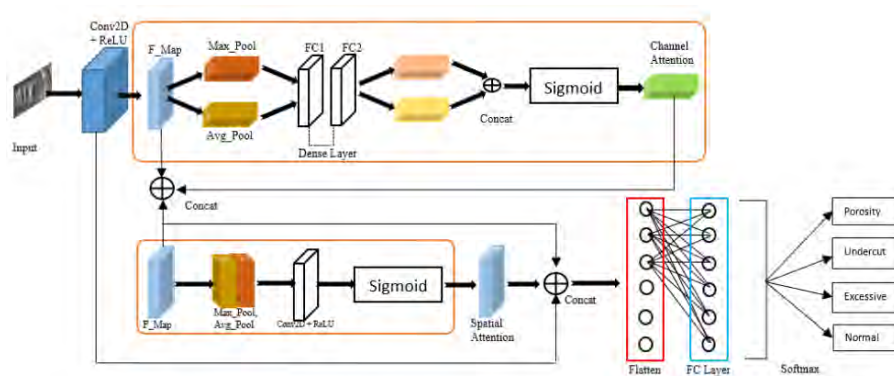


Figure 12. CNN architecture with CBAM attention module in this study.

Table 2. Detailed information on the proposed architecture.

Module	Process Description	Size of Kernel/Pooling	Input Dimension	Output Dimension
Input	Processing input images	-	128x128x3	128x128x3
Conv2D + ReLU	Initial feature extraction from input image	Kernel 3x3, stride 1	128x128x3	128x128x64
Max Pool	Dimension reduction dominant features	Pooling 2x2	128x128x64	128x128x64
Channel Attention	Selecting relevant feature channels	Dense layer (FC1:1x1, FC2:1x1)	64x64x64	64x64x64
Spatial Attention	Highlighting important areas of the image	Kernel 7x7	64x64x64	64x64x64
Flatten	Converting the feature matrix to a 1D vector	-	64x64x64	262,144
Fully Connected Layer	Combining and processing features	-	262,144	128
Softmax	Generates final probabilities for each category	-	128	4

The channel attention approach highlights the importance of the feature representations generated by the CNN. Features have dimensions of $CxHxW$, where C stands for the channel count, H signifies height, and W represents width. This procedure computes average pooling and max pooling along the spatial dimension to provide two representations ($Cx1x1$) that are subsequently processed by a fully connected layer. The outcomes are linked via sigmoid activation to produce attention weights for each channel. The output is a feature with size $Cx1x1$, subsequently multiplied by the initial feature. The computation of the attention channel employs equations (1) and (2).

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

$$F' = M_c(F).F \quad (2)$$

Where F is an input feature tensor with size $CxHxW$, the average pooling of spatial dimensions H and W is denoted by $AvgPool(F)$, $MaxPool(F)$ is maximum pooling spatial dimension, MLP represents the multi-layer perceptron. The sigmoid function starts from the range between 0 and 1 and is symbolized by σ , $M_c(F)$ is the attention mask on channel dimension $Cx1x1$ and the output feature denoted by F' .

The spatial attention mechanism concentrates solely on the spatial position of the feature. This procedure computes average pooling and max pooling to yield two $1xHxW$ feature maps, which are subsequently amalgamated via a $7x7$ convolutional layer utilising features derived from the attention channel. This approach generates a spatial attention weight of $1xHxW$, which is then applied to the starting features. Before the spatial attention step, the attention channel step occurs; this integration dynamically refines key information and helps the model to better grasp pertinent qualities. Equations (3) and (4) are the formulas employed in spatial attention.

$$M_s(F') = \sigma(Conv2D([AvgPool(F'), MaxPool(F')])) \quad (3)$$

$$F'' = M_s(F').F' \quad (4)$$

F' is input tensor from output of channel attention, Pooling the average channel dimension produces a tensor of size $1xHxW$ denoted by $AvgPool(F')$, $MaxPool(F')$ is the maximum pooling channel dimension with the same size as the average pooling output. The 2D convolution operation with a $7x7$ kernel to capture spatial information is represented by $Conv2D$, $M_s(F')$ is the attention mask in the spatial dimension, and the output feature of spatial attention is denoted by F'' .

3. RESULTS AND DISCUSSION

This section discusses the evaluation results of the prototype visual inspection robot for pipe welding defects, starting from the results of design implementation, model training, model testing, and full prototype testing.

3.1. Design Implementation Results

The design of a weld flaw detection system is brought to life in the shape of a prototype put together from mechanical and electrical parts. Equipped with a rotary encoder, the mechanical frame acts as the principal support for the webcam, LED strip, mini PC, and DC motor actuator.

An elevated angle of the model is shown in Figure 13. This attitude shows where the battery is housed and provides the system its main source of energy. There are also push buttons and an LCD screen to manage the LED lights and DC motor. The motor speed may be changed as required during the inspection process using push buttons.



Figure 13. Top View Prototype Design Implementation Results.

Figure 14 shows the front view of the suggested prototype. From this perspective, the camera aimed right at the subject of study, the features are rather evident. The side of the camera has a light-emitting diode strip fastened to give the welding area enough light. A lux meter was used to assess the LED strip's brightness against the outcomes from the Canny edge detection technique. The results show that the Canny edge detection method works consistently and produces clear defect outlines with little noise over the 170–270 lux illumination spectrum. Using four LED strips at once will produce about 250 lux of light, which is ideal for this prototype. At the rear, a direct current motor featuring a rotary encoder is placed to be the motion actuator. This configuration guarantees the prototype performs well with a small but strong design.



Figure 14. Front View Prototype Design Implementation Results.

3.2. Model Training Results

The graph depicting training loss in Figure 15 illustrates a reduction in loss throughout the training stages. At the start of the first epoch, the loss was significantly elevated, approximately 100. Nevertheless, with each passing epoch, there was a substantial decline in the loss value, nearing zero. This reduction signifies that the model is becoming better at grasping data trends and minimizing prediction inaccuracies. While there are minor variations during the middle of the epoch, they remain within acceptable bounds and reflect an ongoing optimization effort.

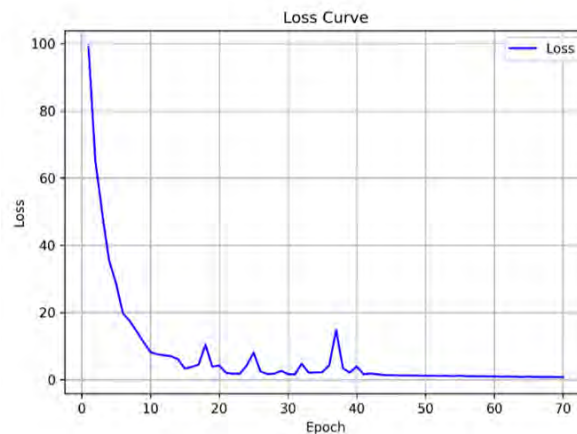


Figure 15. The training loss graph of the proposed model.

Figure 16 shows that, especially within the first 10 to 20 epochs, the values for precision, recall, and F1 score have increased sharply since training started. Later on, all three metrics usually reach plateaus near 1.00. While recall gauges the share of data the model properly identifies from the entire data set, precision mirrors the model's correctness in forecasting the correct class. Combining precision and recall, the F1 score represents the balance between prediction coverage and accuracy. Starting in the 30th epoch, the F1 score started to reach exceptional levels, so demonstrating the model's ability to identify items both correctly and consistently. This graph shows how well the CNN-CBAM model performs in the classification task after training.

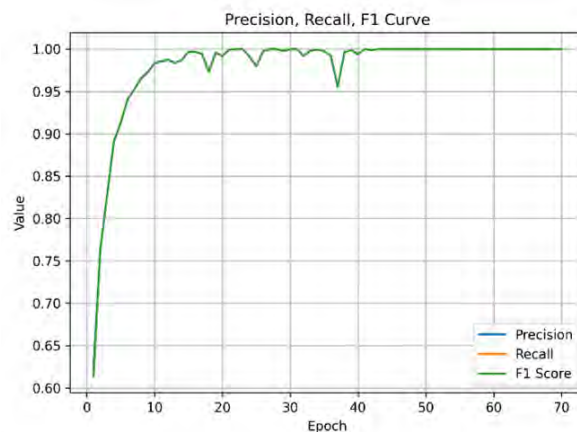


Figure 16. Precision, Recall, and F1-Score graphs of the proposed system.

3.3. Model Testing Results

This part describes the results of assessing the model that was trained using the pipe welding defect information. The assessment was conducted with real pipe welding test data. The suggested approach was then contrasted with the popular object detection algorithm known as YOLOv8. This version of YOLO was chosen for its reliability in terms of both quickness and precision when used on a handheld computer [27], like the mini PC employed in this study.

The proposed method in Figure 17(a) properly found porosity defects with a 99.86% confidence. The defect area is marked with a green box, showing small holes that are characteristic of porosity. The detection process took place in 0.09 seconds with a processing speed of 10.6 FPS, indicating that the system is able to work quickly and accurately on this type of defect compared to the detection results of the YOLOv8 model in 17(b), which only had a confidence level of 94.64% with a detection time and processing speed of each image frame of 2.22 seconds and 0.5 FPS, respectively.

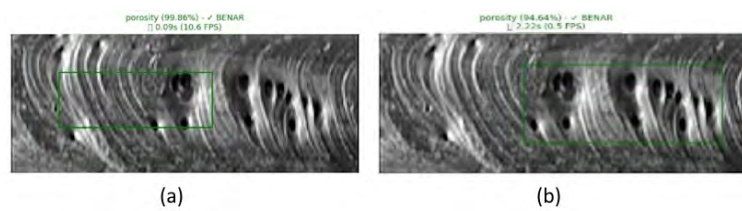


Figure 17. Comparison of the detection results of the two models on porosity-type welding defect test data.
(a) CNN-CBAM model. (b) YOLOv8 model

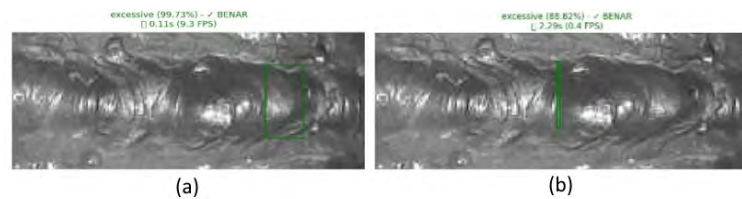


Figure 18. Comparison of the detection results of the two models on excessive class welding defect test data.
(a) CNN-CBAM model. (b) YOLOv8 model

Figure 18(a) shows the results of the CNN-CBAM model detection of excessive weld defects with a confidence level of 99.73%. The system successfully recognized the presence of excessive protrusions on the weld surface marked by a green detection line. The system processing time was measured at 0.11 seconds with a speed of 9.3 FPS, which is better than the identification output with the YOLOv8 algorithm in Figure 18(b) with a confidence level of only 88.82%, a long detection time of around 2.29 seconds, and a low FPS value of 0.4 FPS.

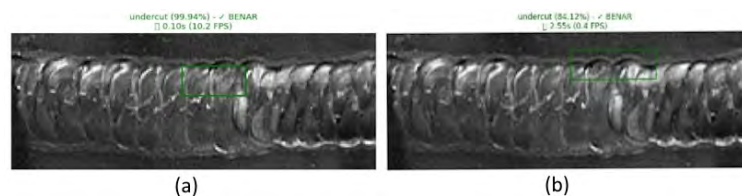


Figure 19. Comparison of the detection results of the two models on undercut-type welding defect test data.
(a) CNN-CBAM model. (b) YOLOv8 model

Figure 19(a) illustrates the detection outcomes of the suggested approach for undercut faults, achieving a confidence level of 99.94%. The green box highlights the empty space along the weld's edge, indicating this kind of flaw. Operating at 10.2 frames per second, the system takes 0.10 seconds to identify it. The results demonstrate that this model is very reliable in identifying undercut flaws with significantly improved accuracy over the YOLOv8 approach presented in Figure 19(b), which has an 84.12% confidence rate, takes 2.44 seconds to run, and operates at around 0.4 FPS.

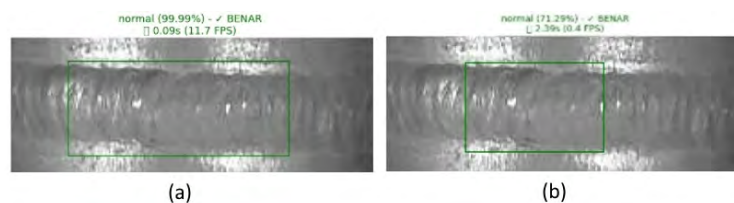


Figure 20. Comparison of the detection results of the two models on normal class welding test data. (a) CNN-CBAM model. (b) YOLOv8 model

Finally, the detection results of the CNN-CBAM model on a normal weld test sample are shown in Figure 20(a). The system can classify with a very high confidence level of 99.99%. The analyzed area is marked with a green box that highlights the weld joint surface without any indication of defects. The classification process is fast with a time of 0.09 seconds and a speed of 11.7 FPS, making it one of the best results in terms of accuracy and responsive FPS performance compared to the YOLOv8 approach in Figure 20(b) which has a 71.29% confidence level, a detection time of around 2.39 seconds, and an average frame rate of 0.4 FPS.

Table 3 presents a summary of the results of welding defect detection on iron pipes using the proposed approach (CNN-CBAM) and the YOLOv8 benchmark method with 10 tests in each class using the test dataset. The test results indicate that the CNN-CBAM approach achieves a high degree of accuracy, with a total confidence of 97.13%, ranging from 94.79% for normal situations to 99.61% for porosity cases. Using the same data, the YOLOv8 approach has a total confidence of 92.21%. For normal situations, its minimum is 90.62%, and for porosity problems, its maximum is 93.83%. The CNN-CBAM technique outperforms the YOLOv8 approach in terms of picture processing speed (FPS), reaching an average rate of 10.35 FPS. The average speed at which the YOLOv8 model processes images is 0.65 FPS, indicating it is insufficient for applications requiring real-time processing.

Table 3. Comparison of the proposed model to the YOLOv8 method.

Welding Defects	YOLOv8 Conf. (%)	YOLOv8 (s)	YOLOv8 (FPS)	CNN-CBAM Conf. (%)	CNN-CBAM (s)	CNN-CBAM (FPS)
Normal	92.43	2.15	0.8	94.79	0.22	10.2
Excessive	90.62	1.98	0.6	96.72	0.35	9.7
Undercut	91.94	2.33	0.5	97.45	0.17	10.9
Porosity	93.83	2.24	0.7	99.61	0.11	10.6
Average	92.21	2.18	0.65	97.13	0.21	10.35

3.4. Full Prototype Testing

The camera arrangement is parallel to the surface of the test pipe, allowing it to image the welded section directly free from any angle distortion; see Figure 21. To maximize brightness and guarantee excellent photographs, LED lights are positioned around the camera. From this perspective, the mechanical frame is also apparent since it can sustain the weight of the carbon steel pipe while the system is operating. The frame's robust and well-balanced construction ensures the prototype remains stable even as the DC motor with a rotary encoder turns to move the pipe. This indicates that the frame's design is not only compact but also has sufficient strength to manage the continuous testing weight.



Figure 21. Bird's eye view of the proposed model in operation.

From this angle, one can also see a little computer positioned atop the structure with a camera in front and a DC motor with a rotary encoder at the rear that helps to move the pipe. This perspective emphasizes the small character of the frame's construction, therefore guaranteeing that every component has a unique location and does not impede the others. This view also suggests that the balanced weight distribution of the components helps the system remain stable as the motor operates.

The picture plainly shows the inside layout, notably the placement of the battery and the main power supply, placed in the centre of the frame to maintain equilibrium. The wires are arranged neatly and don't obstruct the operating area of the camera. The top view also shows the mechanical arrangement, therefore simplifying access for users to the control components including the LCD, buttons, and connectors, so accelerating maintenance and troubleshooting any problems.

By oiling the prototype's surface, it underwent an additional testing scenario on a lubricated pipe. The findings revealed the prototype slid because of the slick surface. To enhance the grip of the prototype, a circular strap was inserted. This study sought to pinpoint the flaws of the prototype for real-world application and considered all restrictions for design enhancements in future research, carried out on a modest basis in the lab.

The picture in Figure 22 illustrates the interaction of hardware and software within the inspection system. Live images of the weld junction are captured by a depth camera linked to a tiny computer, which then handles this data using a classification approach depending on CNN-CBAM. As shown in Figure 23, the findings are shown in a graphical user interface created with Python (PyQt5) that functions as a hub for control and monitoring. Three kinds of pictures are shown by this graphical user interface: the original image, the edge-detected image, and the image cut from the region of interest (ROI). Moreover, interactive buttons on the graphical user interface let users operate actuators such DC motors and LED strips, therefore enabling the seamless cooperation of all the system components.

This picture shows how the graphical user interface looks when the system is being used to find weld flaws. It shows the classification outcomes, including the kind of flaw and the degree of accuracy. Monitoring data is also shown by the graphical user interface in a tabular format, including test findings such detection time, FPS value, and model confidence level. Eliminating the necessity for human inspections, this interface lets consumers digitally examine the state of the weld junction.

This evaluation demonstrated that the system can run concurrently, completing image capture, data processing, classification, and displaying real-time results all within one platform. A major advantage of the technology is this degree of integration, since it simplifies user inspections and automatically creates digital reports.

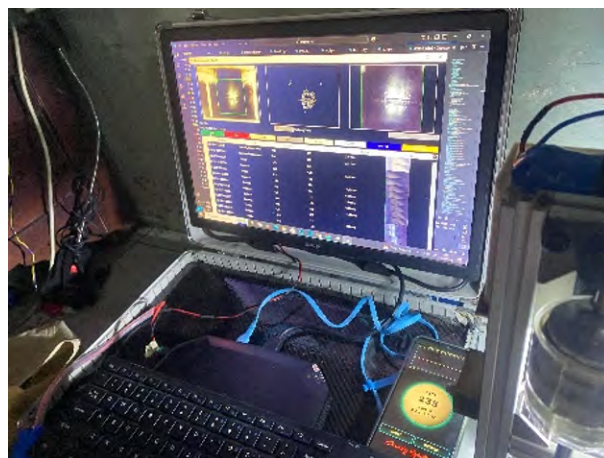


Figure 22. Combination of software and hardware for the proposed prototype.

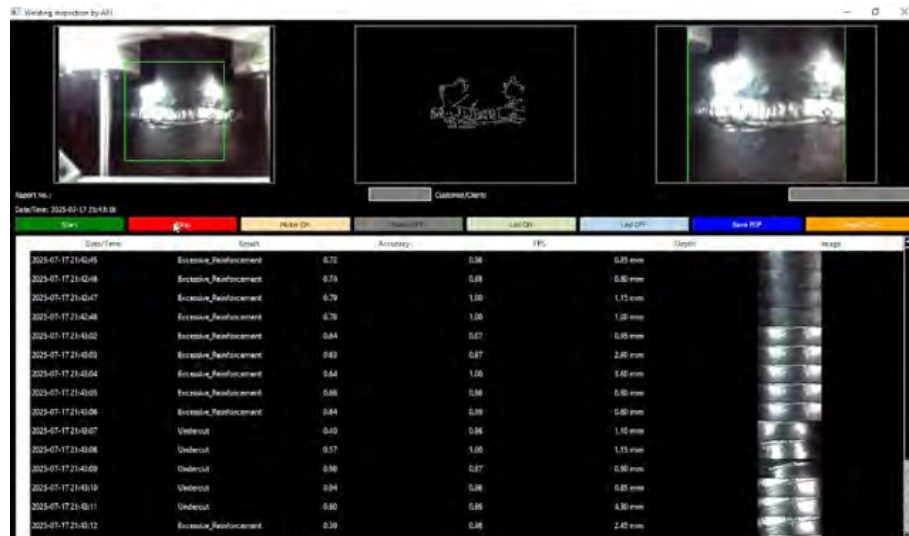


Figure 23. Graphical view of the proposed system's interface.

Along with the range of pipe diameters it can evaluate, Table 4 shows the features of the prototype robot meant to visually search the outer pipeline for welding flaws. With dimensions of 42 by 24 by 22 cm and a weight of 7.5 kg, the IVIS-02 prototype is highly portable. It can find pipes with diameters between 6 and 30 inches, so it can be used in many industries.

Table 4. The proposed prototype's technical specifics.

Prototype Name	Length (cm)	Width (cm)	Height (cm)	Weights (kg)	Pipe inspection capability diameter (inch)
IVIS-02	42	24	22	7.5	6 - 30

Using the model on a mini-PC (ACER Veriton N4640G) and a DELL Latitude E7270 laptop, both with Intel i5 CPUs and 4GB of RAM, we conducted tests to check the effectiveness of the CNN-CBAM approach with this prototype. The two gadgets differ just in weight and size. Weighing around 1.5 kg, this little PC is 204 mm long, 192 mm broad, and 33.2 mm high. On the other hand, the laptop's dimensions are 310.5 mm in length, 215.15 mm in width, and 18.85 mm in height; it weighs around 1.26 kg. Table 5 shows the results of evaluating the CNN-CBAM model on the two devices. To guarantee a fair evaluation, five samples for every kind of welding defect were tested with the same CNN-CBAM models on both devices. With the laptop's mean average precision at 0.84 against the mini-PC's 0.83, the table shows that both gadgets have somewhat equivalent performance; the laptop somewhat outperforms the small PC in average accuracy. The laptop also led in terms of frames per second (FPS) by roughly 0.15 FPS. This testing shows that the mini-PC works well inside the intended design, with a particular focus on how the mini-PC's smaller form factor is a major advantage even while it matches the laptop in welding fault-detection capacity.

Table 5. Performance evaluation results for the CNN-CBAM on laptops and small PCs.

Welding defects	Mini-PC		Laptop	
	mAP50	FPS	mAP50	FPS
Normal	0.79	11.5	0.8	10.9
Excessive	0.83	9.3	0.85	10.5
Undercut	0.87	10.2	0.84	10.1
Porosity	0.81	10.6	0.86	10.7
Average	0.83	10.4	0.84	10.55

4. CONCLUSION

Using an RGB camera and a CNN approach, this study successfully created and implemented a paradigm for detecting welding flaws, with enhanced accuracy attained by adding a CBAM module. The setup includes a webcam, a DC motor with a rotary encoder, an ESP32, an LED strip, and a small PC, which

together allow for capturing images, classifying them, and showing the inspection results in a live graphical user interface. The test results showed that the system has an average accuracy of 97.13%, takes 0.21 seconds to detect issues, and is capable of identifying problems such as porosity, undercut, and excessive reinforcement. It operates at an average frame rate of 10.35 frames per second. The prototype's mechanical design is strong enough for testing in a lab environment, but it requires advancements in materials and dimensions to be fit for application on a bigger industrial scale. To move the project forward, it is important to upgrade the hardware specifications to include GPU support or move to a smaller model (like TensorRT/ONNX) for a more steady frame rate. Furthermore improving the mechanical design with proper materials and measurements matching industry criteria can improve the prototype's durability and readiness for use. To increase flexibility and precision of detection in real-world settings, it is also recommended to include wireless communication and depth-based segmentation methods. In the future, prototype testing will not only be for pipeline inspection but will also be tested in shipping environments such as on ship hulls; manufacturing industries, for example, in industrial tank inspections; and perhaps also in civil construction environments such as bridges or buildings. In addition, improvements to the mechanical design are absolutely necessary so that the robot can grip firmly and be more stable when implemented on slippery inspection objects due to the presence of oil, water, or other materials.

REFERENCES

- [1] T. Yin, J. Wang, H. Zhao, L. Zhou, Z. Xue, and H. Wang, "Research on Filling Strategy of Pipeline Multi-Layer Welding for Compound Narrow Gap Groove," *Materials*, vol. 15, no. 17, p. 5967, 2022.
- [2] S. K. Sharma and S. Maheshwari, "A review on welding of high strength oil and gas pipeline steels," *J. Nat. Gas Sci. Eng.*, vol. 38, pp. 203–217, 2017.
- [3] G. Molnar, "Economics of gas transportation by pipeline and LNG," in *The Palgrave Handbook of International Energy Economics*, M. Hafner and G. Luciani, Eds. Cham: Springer International Publishing, 2022, pp. 23–57.
- [4] J. B. Geng, Q. Ji, and Y. Fan, "A dynamic analysis on global natural gas trade network," *Appl. Energy*, vol. 132, pp. 23–33, 2014.
- [5] V. Arora and Y. Cai, "US natural gas exports and their global impacts," *Appl. Energy*, vol. 120, pp. 95–103, 2014.
- [6] M. P. Singh, D. K. Shukla, R. Kumar, and K. S. Arora, "The structural integrity of high-strength welded pipeline steels: A review," *International Journal of Structural Integrity*, vol. 12, no. 3, pp. 470–496, 2021.
- [7] W. Hou, X. Zhang, J. Gao, et al., "Review on Computer Aided Weld Defect Detection from Radiography Images," *Appl. Sci.*, vol. 10, no. 5, pp. 1878, 2020.
- [8] D. Palma-Ramírez, B. D. Ross-Veitia, P. Font-Arriosa, A. Espinel-Hernández, A. Sanchez-Roca, H. Carvajal-Fals, J. R. Nuñez-Alvarez, and H. Hernández-Herrera, "Deep convolutional neural network for weld defect classification in radiographic images," *Heliyon*, vol. 10, no. 9, Art. no. e30590, 2024, doi: 10.1016/j.heliyon.2024.e30590.
- [9] H. Zhu, W. Ge, and Z. Liu, "Deep Learning-Based Classification of Weld Surface Defects," *Applied Sciences*, vol. 9, no. 16, Art. no. 3312, 2019, doi: 10.3390/app9163312.
- [10] G. A. Elhendawy and Y. El-Taybany, "Machine Vision-Assisted Welding Defect Detection System with Convolutional Neural Networks," *Int. J. Precis. Eng. Manuf.*, pp. 1–10, 2025, doi: 10.1007/s12541-025-01281-y.
- [11] A. Spruck, J. Seiler, M. Roll, T. Dudziak, J. Eckstein, and A. Kaup, "Quality assurance of weld seams using laser triangulation imaging and deep neural networks," in 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT, June 2020, pp. 407–412.
- [12] T. Tyystjärvi, I. Virkkunen, P. Fridolf, A. Rosell, and Z. Barsoum, "Automated defect detection in digital radiography of aerospace welds using deep learning," *Welding in the World*, vol. 66, no. 4, pp. 643–671, 2022.
- [13] Y. Du, M. Liu, K. Wang, X. Liu, K. Wang, Z. Liu, Q. Dong, J. Yao, D. Lu, and Y. Su, "A wall climbing robot based on machine vision for automatic welding seam inspection," *Ocean Engineering*, vol. 310, p. 118825, 2024.
- [14] M. Zhao, X. Liu, K. Wang, Z. Liu, Q. Dong, P. Wang, and Y. Su, "Welding seam tracking and inspection robot based on improved YOLOv8s-Seg model," *Sensors*, vol. 24, no. 14, p. 4690, 2024.
- [15] P. Zhang, J. Wang, F. Zhang, P. Xu, L. Li, and B. Li, "Design and analysis of welding inspection robot," *Scientific Reports*, vol. 12, no. 1, p. 22651, 2022.
- [16] J. Li, B. Li, L. Dong, X. Wang, and M. Tian, "Weld seam identification and tracking of inspection robot based on deep learning network," *Drones*, vol. 6, no. 8, p. 216, 2022.
- [17] R. S. Elankavi, D. Dinakaran, A. S. A. Doss, R. M. K. Chetty, and M. M. Ramya, "Design and motion planning of a wheeled type pipeline inspection robot," *Journal of Robotics and Control (JRC)*, vol. 3, no. 4, pp. 415–430, 2022.
- [18] C. Xu, W. Jiang, H. Ma, Y. Liu, H. Men, X. Xu, and Z. Li, "A review: research and application of pipeline robots in the oil and gas industry," *Journal of Pipeline Science and Engineering*, p. 100356, 2025.
- [19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.
- [20] 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)*, 2016, pp. 779–788.
- [21] A. Niraula, H. Remes, and A. Nussbaumer, "Strain-based analysis on the influence of local undercut geometry on fatigue crack initiation life," *International Journal of Fatigue*, vol. 186, p. 108392, 2024.
- [22] J. Chen, Y. Zhang, C. Wu, and G. K. Padhy, "Suppression of undercut defects in high-speed GMAW through a compound magnetic field," *J. Mater. Process. Technol.*, 2019, doi: 10.1016/j.jmatprotec.2019.116288.
- [23] KOBE STEEL, LTD., *Weld Imperfections and Preventive Measures*, Kobe Steel, Tokyo, Japan, 2015. [online] Available: https://www.kobe-steel.com/en/technology/welding/weld_imperfections_and_preventive_measures

- //www.kobelco-welding.jp/images/education-center/pdf/Defect_5Ed.pdf (accessed August 14, 2025)
- [24] American Welding Society, *Structural Welding Code — Steel, AWS D1.1/D1.1M:2020*, Miami, FL: American Welding Society, 2020. [online] Available: https://pubs.aws.org/Download_PDFS/D1_1_D1_1M_2020_index_PV.pdf (accessed September 16, 2025)
- [25] S. Shin and S. Rhee, "Porosity Characteristics and Effect on Tensile Shear Strength of High-Strength Galvanized Steel Sheets after the Gas Metal Arc Welding Process," *Metals*, vol. 8, no. 12, art. 1077, Dec. 2018, doi: 10.3390/met8121077.
- [26] American Welding Society, *A3.0M/A3.0 — Standard Welding Terms and Definitions*, Miami, FL: American Welding Society, 2010. [online] Available: <https://dl.gasplus.ir/standard-ha/Standard-AWS/AWS%20A3.0%202020.pdf> (accessed October 2, 2025)
- [27] R. Reveles-Martínez, H. Gamboa-Rosales, E. Sánchez-Femat, J. Saldívar-Pérez, T. Ibarra-Pérez, L. C. Reveles-Gómez, O. A. Guirrette-Barbosa, J. I. Galván-Tejada, C. E. Galván-Tejada, H. Luna-García, and J. M. Celaya-Padilla, "Benchmarking YOLOv8 to YOLOv11 Architectures for Real-Time Traffic Sign Recognition in Embedded 1:10 Scale Autonomous Vehicle," *Technologies*, vol. 13, no. 11, p. 531, 2025.

BIOGRAPHIES OF AUTHORS



Noorman Rinanto    was born in Surabaya, Indonesia in 1976. He received the bachelor degree in 2006 with the department of computer system engineering and the master degree in 2012 with the department of intelligent network and multimedia, both from the Sepuluh Nopember Institute of Technology (ITS), Surabaya. He also received the Ph.D. degree from National Taiwan University of Science and Technology, Taiwan in 2023 with the department of electrical engineering. Currently, he is a lecturer in Shipbuilding Institute of Polytechnic Surabaya (SHIPS), Surabaya, Indonesia. His research interests include Artificial Intelligence, Image and Signal Processing, Machine Learning, Automation and Robotic Control System, IoT, and Optimization Algorithm. He can be contacted at email: noorman.rinanto@ppns.ac.id



Ari Shinta was born in Sidoarjo in 2003. She received the bachelor of applied science from the Shipbuilding Institute of Polytechnic Surabaya (SHIPS), Surabaya, in 2025 with the automation engineering program study. Her research interests include IoT, Automation and Robotic Control Systems, and Artificial Intelligence. She can be contacted at email: arishin767@gmail.com.



Ryan Yudha Adhitya   was born in Kediri, Indonesia in 1991. He received the bachelor of applied science from the Shipbuilding Institute of Polytechnic Surabaya (SHIPS) - ITS, Surabaya in 2013 with the department of automation engineering and the master's degree from Sepuluh Nopember Institute of Technology (ITS), Surabaya in 2015 with the department of physics engineering. His research interests include Softcomputing, Artificial Intelligence, Digital Image Processing, Underwater Acoustic, and Spectrum Analysis. He has a certificate of competency in automation system engineer given by SIEMENS. He can be contacted at email: ryanyudhaadhitya@ppns.ac.id



Mahmoud Elsisy (Senior Member, IEEE)    received the B.Sc., M.Sc., and Ph.D. degrees from the Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt, in 2011, 2014, and 2017, respectively. He worked as an Assistant Professor at the Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University. From August 2019 to July 2022, he worked as an Assistant Professor with the Industry 4.0 Implementation Center, Center for Cyber-Physical System Innovation, National Taiwan University of Science and Technology, Taipei, Taiwan. He is currently an Associate Professor with the Electrical Engineering Department, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan. His research interests include studying the machine learning, deep learning, the Internet of Things (IoT), cybersecurity, model predictive control, neural networks, fuzzy logic, Kalman filter, observers, decentralized control of large scale systems, robotic control, autonomous vehicle control, renewable energy, power system dynamics: stability and control, nuclear power plant control, and wind energy conversion systems, faults diagnosis, power transformers, maritime microgrids.. He can be contacted at email: mahmoudelsisy@nkust.edu.tw