e-ISSN: 2146 - 9067

International Journal of Automotive Engineering and Technologies

journal homepage:
https://dergipark.org.tr/en/pub/ijaet

Original Research Article

Simulation-based spot welding inspection on automotive chassis using YOLO-powered image processing

Adem Dilbaz [1, *], İlker Ali Özkan [2]

[1, *] Nükte Automotive, Konya, Türkiye.
[2] Department of Computer Engineering, Faculty of Technology, Selçuk University, Konya, 42250, Türkiye.

ARTICLE INFO

Orcid Numbers
1. 0000-0002-3135-7032
2. 0000-0002-5715-1040

Doi: 10.18245/ijaet.1729908

* Corresponding author
ademdilbaz25@gmail.com

Received: Jun 29, 2025
Accepted: Jul 25, 2025

Published: 30 Sep 2025

Published by Editorial Board Members of IJAET

To cite this paper: Dilbaz, A., Özkan İ., A., Simulation-based spot welding inspection on automotive chassis using yolo-powered image processing, International Journal of Automotive Engineering and Technologies. 2025, 14 (3), 170 – 180. http://dx.doi.org/10.18245/ijaet.1729908

ABSTRACT

This study examines a simulation-based testing platform designed to enhance the quality control processes of Resistance Spot Welding (RSW), a technology widely used in the automotive industry. A virtual testing environment was developed to eliminate the need for physical prototypes. The platform was assembled by placing ESP32-CAM-based virtual cameras on a vehicle chassis obtained from the RoboDK library within the simulation environment. A dataset of approximately 1,000 real RSW images from Kaggle was labeled using Roboflow and converted into a format compatible with YOLO(You Only Look Once) architecture. During image processing and object recognition, YOLOv3-s and YOLOv5-m models were utilized. The models' classification performance was evaluated using metrics such as F1 score, precision, recall, mean average precision (mAP), and Confidence Score (CS). Both models required low hardware requirements; however, YOLOv5-m displayed overall superior performance. Notably, the YOLOv5-m model achieved higher confidence scores in detecting critical welding defects classified as Class 2 (explosion weld); an approximate increase of 8–9% was observed in experimental results, reaching a CS of around 0.58. In addition, the F1 score for Class 2 (explosion weld) improved by approximately 5–6%, reaching a value of around 0.85. This simulation-based method has made RSW quality control faster, more cost-effective, and reliable. Consequently, robotic welding systems can be thoroughly tested for accuracy and safety in a virtual environment before being integrated into the production line.

Keywords: Automotive Industry, Deep Learning, Image Processing, Resistance Spot Welding, RoboDK, YOLO

1. Introduction

Resistance spot welding (RSW) has become one of the most widely used joining methods within the framework of Industry 4.0 within the automotive industry, due to its advantages in high-speed operation, energy efficiency, and structural integrity [1]. In this process, two metal sheets are fused by a short-duration, high-intensity electric current. However, the quality of this process is directly influenced by numerous parameters such as the applied

current, duration, electrode pressure, and mechanical stability [2]. Especially in mass production lines, maintaining consistent and repeatable weld quality is essential for both product safety and manufacturing efficiency [3].

In typical spot welding systems, the lower jaw (lower electrode) remains stationary, while the upper jaw (upper electrode) is equipped with pneumatic, servo-motor-driven, or—less commonly—hydraulic actuators to allow mobility. During the welding process, these electrode structures may adversely affect system performance under certain unfavorable conditions; therefore, the detection and monitoring of such conditions are of utmost importance for ensuring process reliability and maintaining weld quality. In particular, malfunctions in the control of the upper electrode may lead to a non-uniform pressure distribution at the weld zone, thereby compromising the overall quality of the weld joint [4][5].

In this context, Figure 1 illustrates the working principle of a spot welding machine. In the diagram, the movable upper electrode functions as the anode (positive pole), while the fixed lower electrode serves as the cathode (negative pole). The material to be welded is positioned between these two electrodes, held in place by the applied pressure, and an electric current is passed through the contact point. The resistance generated at the contact region produces heat, which in turn facilitates the fusion of the materials. The stylized representation of multi-layered lines in spot welds symbolically depicts both the magnitude of the applied force and the path through which the current flows.

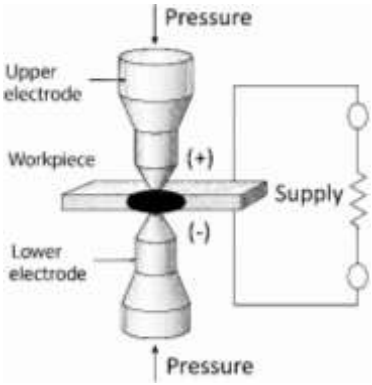

Figure 1. Working principle of spot welding

The evaluation and analysis of resistance spot welding (RSW) quality is a crucial step in enhancing process efficiency [6]. Today, a variety of CAD-based welding simulation software tools are employed as powerful instruments for generating digital twins of materials used in manufacturing processes [7]. In this context, RoboDK stands out by enabling the seamless integration of spot welding machines from various brands into the simulation environment, thanks to its extensive libraries of robots and equipment.

Furthermore, the platform allows for the realistic visualization of weld seam appearances on a digital chassis following the welding operation. Utilizing RoboDK's camera interface, welding programs can be developed without the need for actual field tests. In addition, advanced rendering settings available within the simulation environment enhance the visual quality of the data, enabling virtual welding defects to be represented with greater realism.

RoboDK, a simulation program that offers free access for academic research, is compatible with various software platforms. It enables seamless integration with applications based particularly on SolidWorks and OpenCV, facilitating the transfer of design data into robotic applications [8]. Through its API support, accessible via the Python programming language, RoboDK allows users to develop customized scenarios and dynamically control robot movements. Thanks to its versatile capabilities, RoboDK proves to be an effective tool not only for creating realistic robot motion simulations but also for modeling and analyzing welding applications based on camera integration and image processing.

Among artificial intelligence–based object detection algorithms, YOLO (You Only Look Once) stands out in image processing applications within the automotive industry, particularly due to its detection capabilities when implemented with Python [9]. This algorithm performs both classification of weld images and prediction of bounding boxes using a single neural network. It operates by dividing the image into S × S grids, where each grid cell predicts parameters such as the center coordinates (x, y), width (w), height (h), and

confidence score of the bounding boxes, along with an F1 score that ranges between 0 and 1 [10].

This deep learning–based approach enables the detection and analysis of various weld defects that may occur during the resistance spot welding (RSW) process. Such defects can be so minor that they are undetectable by manual inspection methods in automated production lines, potentially posing serious risks to production reliability [11]. In this study, the objective is to perform classification tasks, identify potential welding defects, and visually present the results using pre-existing datasets of the RSW process on a realistic simulation platform.

2. Material and Methods

In this study, the imaging system employed within both the simulation and practical application environments is based on a camera module suitable for low-cost and embedded systems projects. This module was modeled in three dimensions using SolidWorks software and subsequently integrated into the RoboDK platform for use within the simulation environment. The representative hardware used in the simulation—shown in Figure 2—is a compact, cost-effective ESP32-based camera module with approximate dimensions of 70×30×25 mm, including its protective enclosure.

Figure 2. ESP 32-CAM module used in the simulation tests

Equipped with an integrated OV2640 image sensor, the module supports wireless communication via built-in Wi-Fi and Bluetooth functionalities. These features enable effective use of the system even in real-time visual monitoring applications [12]. The hardware architecture is particularly well-suited for simulation applications with spatial constraints. Owing to its wireless connectivity and capability to operate with an external battery, this camera does not require any cable carrier system when mounted on an RSW machine. As a result, it minimizes restrictions on robotic movements and eliminates the risk of cable entanglement, thereby significantly enhancing the system's overall reliability and operational usability. For the RoboDK simulation experiments presented in this article, two ESP32-CAM camera units were utilized.

In deep learning–based object detection models such as YOLO, the datasets used during training play a critical role in determining the model's accuracy, generalization capability, and real-world performance [13]. Factors such as dataset diversity, image quality, class balance, and accurate labeling are essential for enabling the model to correctly identify objects across varying scenarios [14]. In this study, a pre-existing dataset consisting of approximately 1,000 real RSW (Resistance Spot Welding) images was utilized.

Figure 3. Kaggle dataset prepared for RSW applications [15]

The visual datasets presented in Figure 3 are employed in image processing and deep learning–based classification studies. Since these datasets—comprising images from various angles—are not sufficient on their own for accurate classification, proper labeling of the data is required. In this context, the data were annotated according to three distinct class

IDs, and corresponding label files (label ID files) were created accordingly.

To carry out the labeling process, all images in the dataset were uploaded to the Roboflow platform, where each real image was manually labeled. As illustrated in Figure 4, this process was completed step by step for each image on the platform. Thus, objects related to resistance spot welding were identified and enclosed within appropriate bounding boxes. Subsequently, the labeled dataset was exported in YOLO-compatible format for model training.

Within the scope of this study, visual data related to RSW found in the literature were subjected to preliminary image processing steps. The resulting quality classifications were grouped under three main categories: *Good*, *Bad*, and *Explosion class.*

- The **Good** class (*Class 1*) refers to cases where the welding process has been carried out under optimal parameters, resulting in a properly formed weld nugget and a mechanically sound joint between the parts.
- The **Explosion** class (*Class 2*) typically results from parameter-related errors such as excessive heat, high pressure, or overly strong welding current. In this case, molten metal abruptly and uncontrollably ejects from the weld zone. Such welds often result in severe spatter, surface deformation, or even hole formation on the welded parts.
- The **Bad** class (*Class 3*) represents substandard outcomes in which the nugget formation is insufficient or the joint quality is low; however, some level of fusion between the parts is still present.

A study in the literature reported that composite plates exhibiting explosion welding defects showed the presence of micro-voids and micro-cracks at the weld interface, emphasizing the severe implications of such defects for weld quality [16]. In another similar study, the applicability of the RSW method in the automotive and related industries was evaluated. It was noted that the explosion welding defects encountered in this context led to various geographical limitations and raised serious safety concerns [17]. Accordingly, welds classified under the *Explosion* category should not pass through quality control processes. Given the critical importance of this class for both operational safety and weld integrity in the automotive industry, the data quantity pertaining to this category may be increased using various data augmentation techniques, depending on the current availability of test samples.

Figure 4. Dataset Creation for Model Training in Roboflow Software

In YOLO-based image processing applications, computer hardware plays a critical role in terms of model training duration and real-time performance [18]. The use of GPUs in particular enables faster and more efficient training of deep learning models, while components such as sufficient RAM and SSDs help optimize data processing and access times. Therefore, proper hardware selection is taken into serious consideration, especially regarding its impact on model performance with large datasets. In this study, an RTX 2050 graphics card—offering advantages for image processing via CUDA technology—and a 12th-generation Intel i5-12500H processor were utilized.

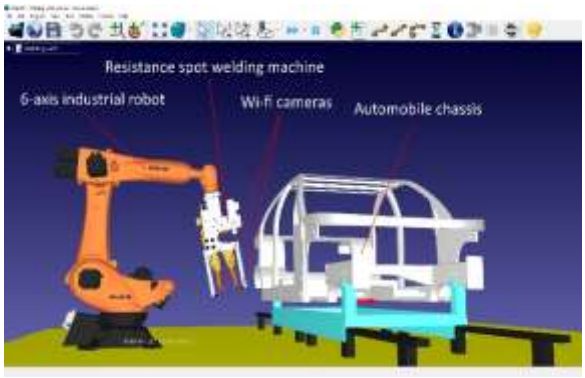

Figure 5. Integration of realistic materials from the RoboDK library into the simulation platform

As shown in Figure 5, the KUKA KR 210 R2700 extra B industrial robot and the servo-motor controlled C-type OBARA spot welding machine were selected from the RoboDK equipment library and virtually modeled within the simulation environment. These

components were integrated into the RoboDK platform to ensure that the robot was mechanically prepared in terms of mobility within the welding cell and accessibility to welding points. As a result, it becomes possible to effectively analyze and observe operational validation and cycle time optimization before any physical implementation.

The welding torches and clamping mechanisms used in spot welding applications typically possess substantial mass. In this context, the payload capacity of the KR 210 robot enables the secure handling of heavy welding equipment, making it suitable for integration with the OBARA spot welding machine employed in the simulation study. Additionally, its 2700 mm reach provides easy and effective access to weld points, particularly in the large-scale chassis structures commonly used in the automotive industry. This capability enhances the system's flexibility and overall operational efficiency [19].

The YOLO algorithm offers significant improvements in detection accuracy and speed, particularly in complex object detection scenarios where overlaps and small-scale defects must be handled effectively [20]. YOLOv3 is a lightweight and fast deep learning model optimized for object detection tasks on systems with limited hardware resources [21]. Built on the Darknet framework, this architecture is simplified to include fewer layers and parameters, thereby achieving high processing speed even under low computational power. Compared to the standard YOLOv3 architecture, it features a more compact design.

However, this simplicity also introduces certain drawbacks. The model may yield a high number of false negatives—objects present in the scene but not correctly detected—and may produce relatively low confidence scores [22]. Therefore, in applications requiring higher accuracy and reliability, the use of YOLOv5 or more recent versions is recommended.

Although the advanced versions of YOLO offer significant advantages in terms of accuracy and reliability, they also introduce certain limitations in specific cases. Due to increased model complexity, training times tend to be longer, and YOLO models trained on large datasets require more computational power to achieve optimal performance. As the quality and scope of the training data improve, the accuracy of object detection correspondingly increases.

Table 1. Class-wise distribution of Kaggle RSW data

Class ID	Test	Validation	Train
Class 1	63	76	320
Class 2	49	57	238
Class 3	32	38	117
All classes	144	171	675

In the literature, optimizing key hyperparameters is considered critical for enhancing the performance of the YOLOv3 model [23]. The learning rate (LR), which determines the step size for updating model weights, must be carefully tuned; an excessively high learning rate can lead to instability, whereas a very low learning rate may slow down the training process. Additionally, activation functions such as Leaky ReLU are commonly employed in YOLO architectures, as they mitigate the "dying neuron" problem by allowing a small gradient when inputs are negative, thereby facilitating more effective learning. Furthermore, in this study, the weld images were resized to 480×480 pixels to ensure compatibility with the YOLO architectures and to enhance processing efficiency. Higher-resolution images, while containing more detail, impose additional load on GPU memory, leading to extended training times and less efficient utilization of hardware resources.

In both model variants—YOLOv3-s and YOLOv5-m—the dataset was divided into three subsets: 68% for training, 17% for validation, and 15% for testing. The numerical distribution of samples for each class in these subsets is presented in Table 1. During the training phase, conducted in a Python-based development environment, each model underwent 50 epochs with a learning rate (LR) set to 0.005. Adam optimizer was selected to update the model weights, utilizing its adaptive learning rate and momentum features to ensure stable and efficient training. The optimizer is also favored by practitioners new to deep learning, especially when training models like

YOLOv3, for its ease of use and stability. [24]. Since the application involved real-time data processing, a batch size of 1 was selected. As a result, two distinct weight files were generated: v3sbest.pt and v5mbest.pt. These weight files enabled the test data to be processed at the millisecond level, and performance metrics such as recall and precision could be calculated rapidly following the classification task.

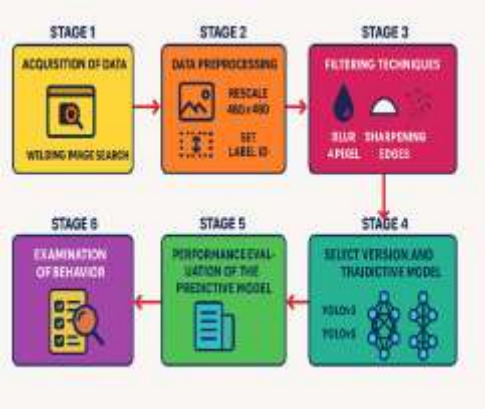

Figure 6. The six fundamental stages of the YOLO algorithm

In real-world conditions, imaging systems can be significantly affected by factors such as variations in ambient lighting, sensor noise, reflections, and shadows [25]. These external influences may reduce the accuracy of deep learning–based object detection algorithms, potentially leading to misclassification or missed detection of small or low-contrast welding defects. To minimize such adverse effects, various filtering and preprocessing techniques are applied to the images, ensuring that YOLO models receive more consistent and higher-quality input data. The YOLO-based image processing approach, illustrated in Figure 6, involves multiple stages and utilizes a variety of filtering techniques. One such technique, blurring, is used in preprocessing to soften the image and reduce detail. This filter is applied to the background areas, allowing foreground objects to become more prominent. Specifically, blurring smooths the background using the Gaussian method, reducing distractions caused by noise, texture, or complex patterns. The sharpening technique emphasizes edges and details within the image, which helps make object boundaries clearer, particularly in low-quality images. Meanwhile, noise reduction techniques help minimize unwanted random distortions. In this study, the bilateral filter effectively reduced noise while preserving object boundaries, and the Non-Local Means filter, despite increasing computational load, contributed to improved detection accuracy.

Intersection over Union (IoU), frequently used in object detection tasks, is calculated as the ratio of the area of overlap between the predicted bounding box and the ground truth box to the area of their union [26]. The **F1 score** is the harmonic mean of precision and recall, representing the balance between false positives and false negatives [27]. **mAP**, Mean Average Precision, provides a comprehensive measure of overall detection performance by averaging the **Average Precision (AP)** scores obtained under different threshold values.

Accordingly, the values used in calculating the F1 score are as follows:

- **True Positive (TP):** Refers to instances where the model correctly detects the presence of an actual object. In other words, the model makes a positive prediction that aligns with reality.
- **True Negative (TN):** Represents cases in which the model correctly identifies the absence of an object. The model predicts a negative outcome, and this matches the actual condition.
- **False Positive (FP):** Occurs when the model incorrectly predicts the presence of an object where none exists; commonly referred to as a "false detection."

False Negative (FN): Denotes cases where the model fails to detect an existing object—i.e., it classifies the instance as negative, even though the object is present.

Recommendation 2: Correction and Formatting of Equations

For these variables, the first two equation models are applied, as shown in Eqs. (1) and (2).

$$F_1\ score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

$$CS = P(nv)\ x\ \text{IoU}_{\text{truth,pred}} \quad (2)$$

Based on these formulas, precision is determined by the values of true positives (TP) and false positives (FP), whereas recall relies on true positives (TP) and false negatives (FN). In the literature, the objectness score, denoted

as P(nv), refers to the probability score assigned by the neural network for the presence of a visual object in the corresponding region [28][29]. The P(nv) plays a critical role in determining these classification outcomes, especially for TP; it is a key parameter used in decision making [30]. The CS builds upon this concept by representing the overall likelihood that a predicted bounding box both contains an object and correctly identifies its class. In deep learning–based object detection models, this score is typically calculated by multiplying the objectness score with the Intersection over Union (IoU) between the predicted and ground truth bounding boxes. The numerical computation of these related metrics utilized the class distribution data presented in Table 2. Here, A_{gt} represents the ground truth bounding box, a reference area defined by an expert that reflects the actual location of the object. A_{pred} denotes the area of the predicted bounding box, calculated using its width and height as estimated by the model. N_C refers to the total number of classes (categories) the model is designed to learn. In this study, three distinct classes were defined for the analysis of the RSW dataset obtained from the Kaggle platform. **AP_i** or Average Precision for the i-th class, indicates how effectively the model detects that class. It is computed as the area under the precision-recall curve for the corresponding class.

For these parameters, the last three equation models are applied, as shown in Eqs. (3), (4), and (5).

$$mAP = \frac{1}{N_C} x \sum_{i=1}^{N_C} AP_i \quad (3)$$

$$A_{int} = \max(0,\ X_{right} - X_{left})\, x\, max(0, y_{bottom} - y_{top}) \quad (4)$$

$$IoU = \frac{A_{int}}{A_{gt} + A_{pred} - A_{int}} \quad (5)$$

In this study, the RoboDK simulation software was utilized to integrate real Resistance Spot Welding (RSW) images onto a digital vehicle chassis model. This approach leverages the capability of processing images data to analyze post-welding visuals captured from multiple camera modules placed at various positions. Within this framework, various RSW images prepared for testing were evaluated using pre-trained deep learning weight files.

Table 2. Calculation of YOLO metrics: Confidence score (CS) and F1 score

Object Detection	Actual condition	P(nv) Value	Result
Object Detected	Object Detected	0.95	TP
Object Detected	No Object Detected	0.85	FP
No Object Detected	Object Detected	0.10	FN
No Object Detected	No Object Detected	0.05	TN

As shown in Figure 7, the weld images generated by the pressure of the moving upper electrode in the RSW process can be monitored in real time through the camera window located at the upper left corner of the simulation interface.

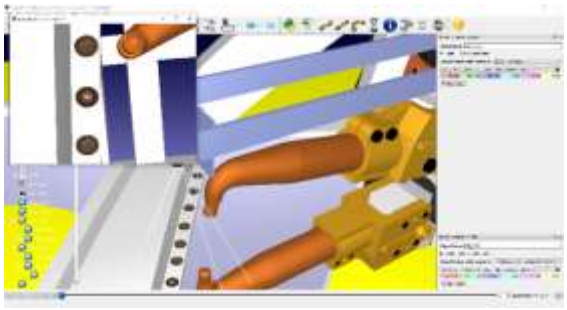

Figure 7. Real-Time image monitoring in RoboDK software

Within the RoboDK simulation environment, it is possible to calculate the robot's field of view, optimize the camera's viewing angle, and pre-simulate various image processing applications by configuring essential camera parameters such as **focal length**, **pixel size**, **field of view**, and **working distance**. As shown in Figure 8, through the camera settings window located on the right side of the interface, the camera was configured to operate within a viewing range of up to 220 mm.

Similarly, the white pyramid-shaped area displayed on the simulation screen represents the physical region detectable by the camera, based on the defined parameters. On the left side of the interface, Python-based YOLO algorithm scripts were inputted, initiating the image processing operations.

3. Test Results and Discussion

In this study, real-time image processing tests were conducted using the YOLOv3-s

(v3sbest.pt) and YOLOv5-m (v5mbest.pt) model weight files, trained on a total of 675 images obtained from the Kaggle dataset within the RoboDK simulation environment. During the test phase, after determining the appropriate camera field of view and working distance, real RSW (Resistance Spot Welding) images placed along the metal plate of a virtual car chassis were sequentially analyzed.

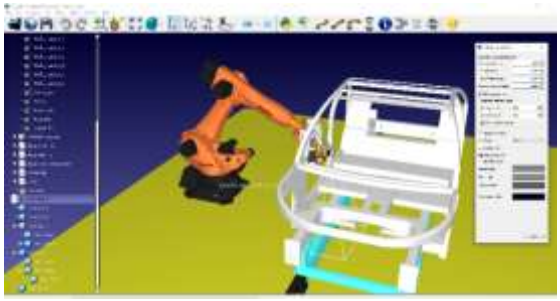

Figure 8. Camera Settings in RoboDK Software

As illustrated in Figures 9 and 10, the weld zones in these images were classified by the respective deep learning models. The detected objects were visualized using bounding boxes along with the classification results. These boxes were highlighted in red frames within the upper and lower camera windows located on the right-hand side of the images.

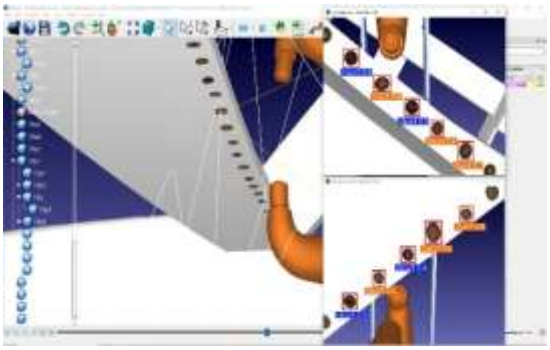

Figure 9. Bottom View and Camera Perspectives of the Test Study with YOLOv5-m

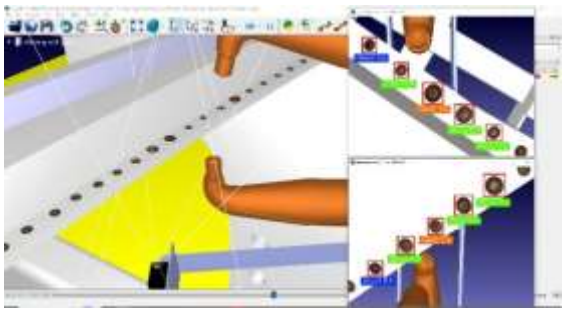

Figure 10. Top View and Camera Window Perspectives of the Test Study with YOLOv3-s

In these figures, Class 1 (Good Weld) is represented in orange, Class 2 (Explosion Weld) in navy blue, and Class 3 (Bad Weld) in green. Although Class 1 and Class 2 exhibit structural similarities, their training data were respectively 173% and 102% more abundant than that of Class 3. Similarly, their test data were 54% and 95% more numerous compared to Class 3. As a result, the YOLO models produced higher confidence scores for these two classes. The corresponding score tables are presented for both YOLOv3-s and YOLOv5-m models.

In this study, classification of RSW data was performed using YOLO-based metrics through both the YOLOv3-s model, known for its low hardware requirements, and the YOLOv5-m model, known for its superior accuracy. It was observed that both models performed effectively and delivered satisfactory results even on low-budget hardware systems. The image processing result graphs for 144 test images processed by each model are presented in Figures 11 and 12, respectively.

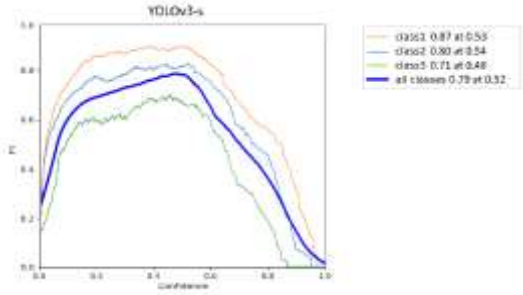

Figure 11. Class-Based Graph Using the YOLOv3-s Model

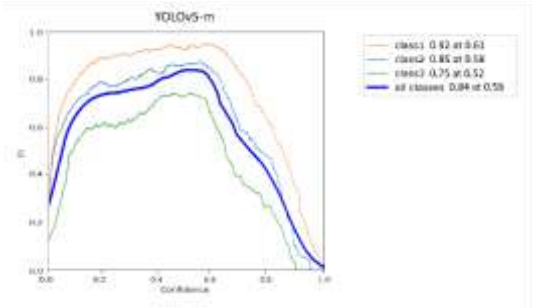

Figure 12. Class-Based Graph Using the YOLOv5-m Model

Both graphs present a comparative analysis of the three different classes (Class 1: Good Weld, Class 2: Explosion Weld, Class 3: BadWeld) in terms of F1 and confidence scores. The YOLOv5-m model demonstrated superior performance across all classes, producing higher scores in both accuracy (F1) and confidence. The particularly high performance in Class 1 is directly related to the

larger quantity of data available for this class. Notably, the model also achieved above-average classification success and reliable outputs for Class 2 (Explosion Weld), which is of critical importance for quality control.

Table 3. Class-Wise Comparison of Image Processing F1 and Confidence Scores for YOLO Models

Model	Class ID	F1 score	CS	IT
YOLOv3-s	1	0.866	0.534	55.3 ms
YOLOv5-m	1	0.924	0.605	72.1 ms
YOLOv3-s	2	0.804	0.536	43.0 ms
YOLOv5-m	2	0.851	0.584	61.2 ms
YOLOv3-s	3	0.714	0.484	63.7 ms
YOLOv5-m	3	0.746	0.520	78.5 ms
YOLOv3-s	all	0.792	0.516	54.1 ms
YOLOv5-m	all	0.837	0.575	69.3 ms

Table 4. Class-Wise Comparison of Additional Image Processing Performance Metrics for YOLO Models

Model	Class ID	Precision	Recall	mAP
YOLOv3-s	1	0.851	0.892	0.76
YOLOv5-m	1	0.910	0.930	0.79
YOLOv3-s	2	0.803	0.826	0.73
YOLOv5-m	2	0.847	0.868	0.77
YOLOv3-s	3	0.691	0.749	0.62
YOLOv5-m	3	0.725	0.772	0.67
YOLOv3-s	all	0.750	0.827	0.68
YOLOv5-m	all	0.813	0.870	0.74

These graphs further highlight a key factor influencing inter-class discrimination ability: data diversity and class imbalance. During training, imbalances in the dataset can limit the model's ability to generalize—particularly for underrepresented classes—resulting in misclassifications. Therefore, the relatively lower performance observed in Class 3 (Bad Weld), which had fewer samples, is a direct outcome of this data imbalance. The data supporting these results are presented comparatively in Tables 3 and 4, along with inference times (IT).

In this study, the mAP@0.5 criterion was adopted as the threshold for evaluating confidence scores. Accordingly, the YOLOv3-s model achieved a value of 0.68, while the YOLOv5-m model reached 0.74, both successfully surpassing the defined threshold. Based on the obtained tables, while the YOLOv5-m model delivered higher accuracy and confidence scores, the YOLOv3-s model demonstrated faster inference times.

4. Conclusion

In this study, image processing operations using YOLO-based deep learning models were conducted within a virtual test platform developed in the RoboDK simulation environment, aiming to detect welding defects occurring during the Resistance Spot Welding (RSW) process. A comparative evaluation of the YOLOv3-s and YOLOv5-m models revealed that YOLOv5-m outperformed in terms of both classification accuracy and confidence scores. Specifically, the F1 score improved by 6.7% for Class 1, 5.8% for Class 2, 4.5% for Class 3, and 5.7% overall. Regarding the confidence score, enhancements of 13.2% for Class 1, 9.0% for Class 2, 7.4% for Class 3, and 10.3% overall were achieved. Additionally, with YOLOv5-m, an improvement of approximately 7–8% in IoU was observed across all classes, reaching a value of around 0.70.

Both YOLO models demonstrated the capability to operate on low-end hardware while delivering reliable and rapid predictions in the simulation environment. This enables quality assessments of welding processes without requiring physical hardware investments, thereby offering significant advantages for process engineering and predictive maintenance planning.

In conclusion, the virtual test platform developed through this study provides a comprehensive solution for the pre-evaluation of robotic spot welding applications, accessibility analysis, and validation of image processing algorithms prior to field deployment. The outputs of this research—including open-access weight files and Python-based software—are well-suited to contribute to the literature and future studies. The findings lay the groundwork for the development of decision-support mechanisms aimed at enhancing weld quality in both academic research and industrial automation projects. For future work, expanding the dataset with a larger number of annotated images, particularly targeting underrepresented defect classes, will be essential to improve model robustness and generalization. Additionally, incorporating advanced synthetic data augmentation techniques such as Generative Adversarial

Networks (GANs) can help to diversify the training samples and alleviate class imbalance issues. Exploring real-world deployment scenarios by integrating the virtual test platform with actual robotic welding systems will further validate the effectiveness of the approach. Moreover, investigating adaptive learning strategies and real-time feedback mechanisms can enhance the system's ability to cope with varying environmental conditions and weld quality fluctuations, ultimately contributing to more reliable and automated quality control in industrial settings.

5. References

1. Capezza, C., Centofanti, F., Lepore, A., & Palumbo, B., Functional clustering methods for resistance spot welding process data in the automotive industry. Applied Stochastic Models in Business and Industry, 37 (5), 908-925, 2021. https://doi.org/10.1002/asmb.2648

2. Li, D., Yang, P., & Zou, Y., Optimizing insulator defect detection with improved DETR models. Mathematics, 12 (10), 1507 2024. https://doi.org/10.3390/math12101507

3. Dai, W., Li, D., Zheng, Y., Wang, D., Tang, D., Wang, H., & Peng, Y. Online quality inspection of resistance spot welding for automotive production lines. Journal of Manufacturing Systems, 63, 354-369, 2022. https://doi.org/10.1016/j.jmsy.2022.04.008

4. Mathiszik, C., Köberlin, D., Heilmann, S., Zschetzsche, J., & Füssel, U., General approach for inline electrode wear monitoring at resistance spot welding. Processes, 9 (4), 685, 2021. https://doi.org/10.3390/pr9040685

5. Liu, W., Hu, J., & Qi, J., Resistance Spot Welding Defect Detection Based on Visual Inspection: Improved Faster R-CNN Model. Machines, 13 (1), 33, 2025. https://doi.org/10.3390/machines13010033

6. Wang, X. J., Zhou, J. H., Yan, H. C., & Pang, C. K. Quality monitoring of spot welding with advanced signal processing and data-driven techniques. Transactions of the Institute of Measurement and Control, 40 (7), 2291-2302, 2018. https://doi.org/10.1177/0142331217700703

7. Yu, X., Sun, X., & Ou, L., Graphics-based modular digital twin software framework for production lines. Computers & Industrial Engineering, 193, 110308, 2024. https://doi.org/10.1016/j.cie.2024.110308

8. Wang, Z., Zhang, M., & Xu, Y., Development of a robotic arm control platform for ultrasonic testing inspection in remanufacturing. In 2022 27th International Conference on Automation and Computing (ICAC), (pp. 1-6). IEEE, 2022. https://doi.org/10.1109/ICAC55051.2022.9911174

9. Gheorghe, C., Duguleana, M., Boboc, R. G., & Postelnicu, C. C., Analyzing Real-Time Object Detection with YOLO Algorithm in Automotive Applications: A Review. CMES-Computer Modeling in Engineering & Sciences, 141 (3), 2024. https://doi.org/10.32604/cmes.2024.054735

10. Singh, A., Kalaichelvi, V., DSouza, A., & Karthikeyan, R., GAN-Based image dehazing for intelligent weld shape classification and tracing using deep learning. Applied Sciences, 12 (14), 6860, 2022. https://doi.org/10.3390/app12146860

11. Lang, X., Ren, Z., Wan, D., Zhang, Y., & Shu, S. (MR-YOLO: An improved YOLOv5 network for detecting magnetic ring surface defects. Sensors, 22 (24), 9897, 2024. https://doi.org/10.3390/s22249897

12. Elhattab, K., Abouelmehdi, K., & Elatar, S, New model to monitor plant growth remotely using esp32-cam and mobile application. In 2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM) IEEE.,(pp. 1-6), 2023. https://doi.org/10.1109/WINCOM59760.2023.10322939

13. Ragab, M. G., Abdulkader, S. J., Muneer, A., Alqushaibi, A., Sumiea, E. H., Qureshi, R., ... & Alhussian, H., A comprehensive systematic review of YOLO for medical object detection (2018 to 2023). IEEE Access., 2024. https://doi.org/10.1109/ACCESS.2024.3386826

14. Ramchandani, M., Sahu, S. P., & Dewangan, D. K., A comparative study in pedestrian detection for autonomous driving systems. In 2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON) IEEE. (pp. 1-6), 2023.

https://doi.org/10.1109/OTCON56053.2023.10113992
15. Domínguez, L., Rivas-Araiza, E. A., Jáuregui-Correa, J. C., González-Córdoba, J. L., Pedraza-Ortega, J. C., & Takács, A, Resistance spot welding insights: A dataset integrating process parameters, infrared, and surface imaging. Data in Brief, 59, 111373, 2025. https://doi.org/10.1016/j.dib.2025.111373
16. Karaman, A., Karaboga, D., Pacal, I., Akay, B., Basturk, A., Nalbantoglu, U.,& Sahin, O., Hyper-parameter optimization of deep learning architectures using artificial bee colony (ABC) algorithm for high performance real-time automatic colorectal cancer (CRC) polyp detection. Applied Intelligence, 53 (12), 15603-15620, 2023. https://doi.org/10.1007/s10489-022-04299-1
17. Gullino, A., Matteis, P., & D'Aiuto, F., Review of aluminum-to-steel welding technologies for car-body applications. Metals, 9 (3), 315, 2019. https://doi.org/10.3390/met9030315
18. Guo, K., Sui, L., Qiu, J., Yu, J., Wang, J., Yao, S., ... & Yang, H., Angel-eye: A complete design flow for mapping CNN onto embedded FPGA. IEEE transactions on computer-aided design of integrated circuits and systems, 37 (1), 35-47, 2017. https://doi.org/10.1109/TCAD.2017.2705069
19. Kulikov, A. A., Sidorova, A. V., & Balanovskii, A. E. Programming industrial robots for wire arc additive manufacturing. In International Conference on Industrial Engineering Cham: Springer International Publishing, (pp. 566-576), 2021. https://doi.org/10.1007/978-3-030-54817-9_66
20. Zhou, S., Ao, S., Yang, Z., & Liu, H., Surface defect detection of steel plate based on SKS-YOLO. IEEE Access, 2024. https://doi.org/10.1109/ACCESS.2024.3422244
21. Singh, A., Raj, K., Kumar, T., Verma, S., & Roy, A. M., Deep learning-based cost-effective and responsive robot for autism treatment. Drones, 7 (2), 81, 2023. https://doi.org/10.3390/drones7020081
22. Sozzi, M., Cantalamessa, S., Cogato, A., Kayad, A., & Marinello, F., Automatic bunch detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms. Agronomy, 12 (2), 319, 2022. https://doi.org/10.3390/agronomy12020319
23. Wu, J., Shen, T., Wang, Q., Tao, Z., Zeng, K., & Song, J. Local adaptive illumination-driven input-level fusion for infrared and visible object detection. Remote Sensing, 15 (3), 660, 2023. https://doi.org/10.3390/rs15030660
24. Swain, S., & Tripathy, A. K., Automatic detection of potholes using VGG-16 pre-trained network and Convolutional Neural Network. Heliyon, 10 (10), 2024. https://doi.org/10.1016/j.heliyon.2024.e30957
25. Mohammadrezaei, E., Ghasemi, S., Dongre, P., Gračanin, D., & Zhang, H. Systematic review of extended reality for smart built environments lighting design simulations. IEEE Access, 2024; 12, 17058-17089, 2024. https://doi.org/10.1109/ACCESS.2024.3359167
26. Kshirsagar, V., Bhalerao, R. H., & Chaturvedi, M. Modified yolo module for efficient object tracking in a video. IEEE Latin America Transactions, 21 (3), 389-398, 2023. https://doi.org/10.1109/TLA.2023.10068842
27. Glučina, M., Anđelić, N., Lorencin, I., & Car, Z., Detection and classification of printed circuit boards using YOLO algorithm. Electronics, 12 (3), 667, 2023. https://doi.org/10.3390/electronics12030667
28. Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., & Wen, S., PP-YOLO: An effective and efficient implementation of object detector arXiv preprint arXiv:2007.12099, 2020. https://doi.org/10.48550/arXiv.2007.12099
29. Wu, S., Li, X., & Wang, X., IoU-aware single-stage object detector for accurate localization. Image and Vision Computing, 97, 103911, 2020. https://doi.org/10.1016/j.imavis.2020.103911
30. Sotres, J., Boyd, H., & Gonzalez-Martinez, J. F., Enabling autonomous scanning probe microscopy imaging of single molecules with deep learning. Nanoscale, 13 (20), 9193-9203, 2021. https://doi.org/10.1039/D1NR01109J