

Article

An AI-Based Vision System for Detecting Defects in Resistance Spot Welding and Predicting Maintenance in Robotic Cells

Alfonso Alejo-Ramirez ¹, Rogelio Cedeño-Moreno ¹, Luis A. Morales Hernandez ¹, Juan C. Jauregui-Correa ² and Irving A. Cruz-Albarran ^{1,*}

¹ Laboratory of Artificial Vision and Thermography/Post-Graduated in Mechatronics, Faculty of Engineering, Autonomous University of Queretaro, Campus San Juan del Rio, San Juan del Rio 76807, Mexico

² Faculty of Engineering, Autonomous University of Queretaro, Queretaro 76010, Mexico

* Correspondence: irving.cruz@uaq.mx

Abstract

The automotive industry is adopting the Industry 4.0 model to reduce failures in electrical resistance spot welding by using automated non-destructive testing systems. This study presents a real-time machine vision system that has been implemented on an automotive cabin assembly line to detect weld defects in door frames. The system extracts the physical parameters of each spot weld, including nugget diameter and heat-affected zone, to identify failing robots and welding points and prioritise maintenance actions (maintenance-free, predictive, preventive or corrective). Machine learning models were trained using images from 674 cabins collected over 30 continuous hours. YOLOv8 was used for spot weld detection and feature extraction. A neural network with linear discriminant analysis was applied for defect classification, achieving 95.80% accuracy, 95.61% precision, 96.00% recall and 95.80% F1-score. Additionally, a convolutional neural network was developed for maintenance prediction, achieving 94.72% precision, 95.87% accuracy, 93.97% F1-score and 94.42% recall. The results demonstrate the effectiveness of real-time defect detection and reliable maintenance prediction, supporting informed decision-making and efficient resource management.

Keywords: computer vision; defect detection; machine learning; maintenance; resistance spot welding



Academic Editors: Victor García,
David Zorrilla and
Jesús Sánchez-Márquez

Received: 10 April 2026

Revised: 21 July 2026

Accepted: 27 July 2026

Published: 31 July 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Resistance spot welding (RSW) is a common method of joining metal parts. It involves applying pressure to the parts to be welded and passing an electric current through them. The resulting resistance heat is used as an energy source to melt the metal and form a weld nugget [1]. It is widely used in the automotive industry to join thin sheets of coated and uncoated materials [2,3]. Robots are currently being used to apply RSW to build vehicle cabins [4]. The main advantages of RSW are ease of automation, speed and high quality [5]. However, to achieve the desired quality of RSW and avoid defects, certain conditions must be considered, such as the type of material used, the amount of current applied, the duration of the welding process and the power used for welding [6]. To assess the quality of spot welds, destructive and non-destructive techniques are used [7]; some are made during the welding process, and others are made after the welding process [8]. Of the many methods of weld defect detection, visual inspection stands out for its speed, non-contact nature, rapid results, and cost-effectiveness [9].

Vision systems can capture an image of a specific area, extract information from it, analyse it, and provide a diagnosis. Regarding inspecting RSW, the following research has been conducted. Choi et al. [10] qualitatively predicted RSW quality using a convolutional neural network (CNN) and information from the heat trace image, focusing specifically on the surface appearance of the nugget. Alghannam et al. [11] implemented a vision system with a six-axis robotic arm to capture images at various points after welding and estimate RSW quality using image processing and a fuzzy inference system. Younes et al. [12] developed a vision system to detect defects and estimate the quality of RSW (e.g., lateral surface ejection, deformed metal, pitting, blowing, cracking and slippage) using image processing and a fuzzy support vector machine to analyse the surface of a nugget on the underbody of a car. Xiao et al. [13] proposed RswNet, a method based on a CNN, to evaluate RSW quality in an automotive production line. Stavropoulos et al. [14] evaluated the feasibility of integrating an RSW quality assurance system incorporating an infrared (IR) camera, machine learning (ML) models and the RSW robotic process into an industrial process. They concluded that it is feasible under various conditions. Lee et al. [15] proposed an explainable failure diagnosis method based on a CNN and gradient-weighted class activation mapping (Grad-CAM) to infer the causal relationship of the failure. This method was tested on a robotic spot welding process. Li et al. [16] proposed a flexible, fine-grained graph convolution network (FFGCN), which enables feature extraction in point space for the visual inspection of RSW, by combining computer vision with natural language processing. Dai et al. [17] proposed a Balancing Generative Adversarial Network and Gradient Penalty (BGANGP) algorithm to generate defective spot images. They then ran a classifier based on transfer learning to assess spot defects. Finally, Yang et al. [18] developed an RSW weld nugget classification method under laboratory conditions based on a deep analysis of the appearance and shape of point cloud data, named the offset-transformer hierarchical model (OFTFHC). While there is research into assessing the quality of RSW nuggets, to the authors' knowledge, there is no research that can predict the level of maintenance required to ensure good quality based on the quality of spots. Likewise, many works are offline, i.e., not within the production line.

Due to their powerful feature extraction and pattern recognition capabilities, CNNs have been widely applied in various industrial applications. They have been successfully used for automated quality inspection, predictive maintenance and fault detection in manufacturing systems, for example. Chian y Tian [19] proposed a method for inspecting surface defects that combines convolutional neural networks with statistical patch fusion techniques. This approach extracts deep features from images and analyses multiple local regions to improve the detection of small, noisy defects on industrial surfaces. Experimental results showed that integrating deep learning with statistical image analysis significantly improves defect detection performance compared with traditional vision-based inspection methods. Similarly, [20] analysed the application of CNN-based models for fault detection in photovoltaic systems, highlighting potential limitations related to model interpretability and generalisation. Their study investigated the presence of shortcut learning in deep learning models by applying explainable artificial intelligence techniques to evaluate how CNNs make classification decisions. The results showed that, while CNN architectures can achieve high accuracy in defect detection tasks, their predictions sometimes rely on spurious patterns in the dataset. These findings emphasise the importance of interpretability and robust dataset design when implementing deep learning models in industrial inspection systems.

This study proposes the design and implementation of a computer vision system combined with artificial intelligence (AI) algorithms for detecting and classifying RSW spots and predicting the maintenance required for the welding robot. The system will

monitor and classify welding points as either good or bad. It will also show how often failures occur in the welding robots and determine the maintenance forecast through identifying trends. This will make it possible to identify the robot and the failing point, as well as the level of priority with which the problem should be addressed. The system will consider the categories of maintenance: predictive, preventive, corrective, or when maintenance is not necessary. One advantage of the system is that NDT tests can be performed based on the physical characteristics of the spot welds in real time without interfering with the production line. In summary, the main contribution is the development of an artificial vision system based on three fundamental points: (1) automatic detection of RSW spots; (2) classification of each RSW spot as 'good' or 'bad' with an accuracy of 95.80%, a precision of 95.61%, an F1-score of 95.80% and a recall of 96.00%; and (3) prediction of maintenance required for the welding robot based on the RSW spots identified, with a precision of 94.72%, an accuracy of 95.87%, an F1-score of 93.97% and a recall of 94.42%.

2. Materials and Methods

The artificial vision system has been designed and implemented in accordance with the general diagram shown in Figure 1. This diagram illustrates the image acquisition process along the line, as well as the configuration of the vision system. These images are used to create a database over three days, from which spot welds are detected and segmented. The physical parameters of the spots are extracted from these images, such as nugget diameter and heat-affected zone (HAZ). These parameters are analysed to classify the spots as good or bad based on the presence of defects. A CNN model then processes the extracted information to determine the robots' tendency to work on defective points and predict the priority of maintenance. The model considers corrective, preventive, predictive or free maintenance.

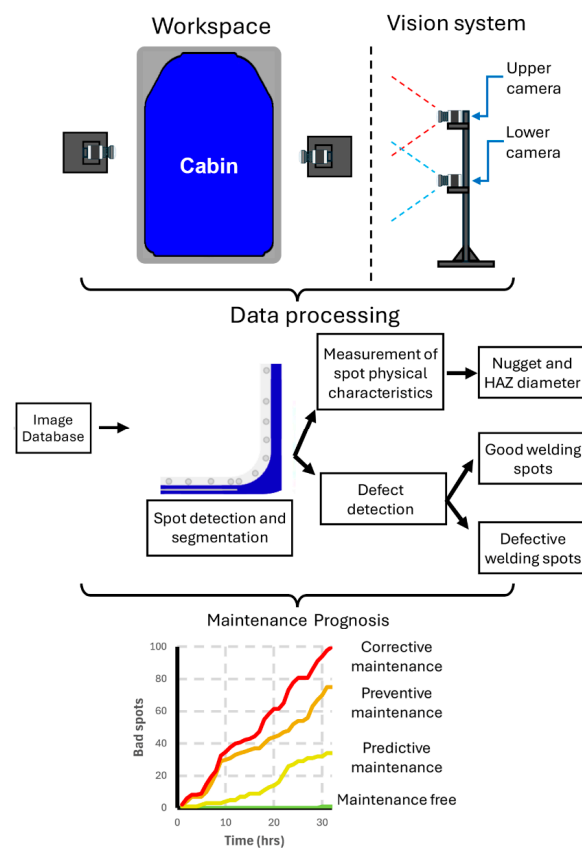


Figure 1. General methodology diagram.

2.1. Workspace

The work area is located on the production line immediately after cabin assembly. As shown in Figure 2, which illustrates the general layout of the components and their locations, the vision system is installed in this area. Inside the workspace, an array of cameras is mounted on each side of the cabin at 1.3 m. These are complemented by an illumination system designed to enhance image quality. All this equipment is connected and managed from a workstation located outside the workspace.

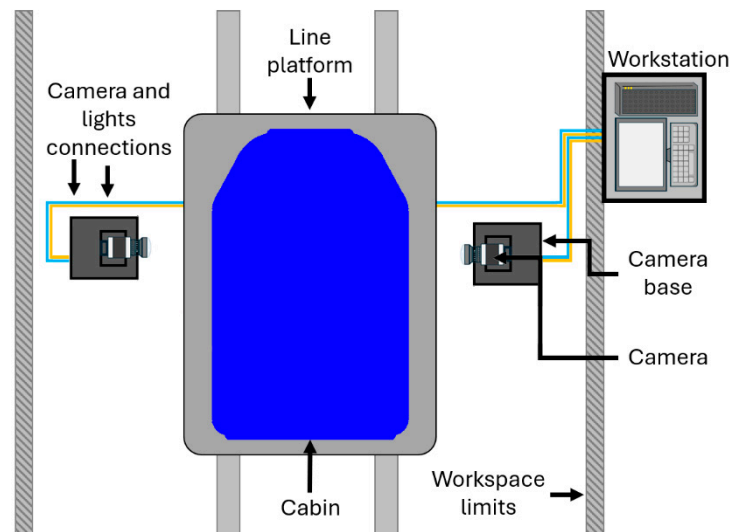


Figure 2. In-line vision system workspace.

2.2. Vision System

To implement the vision system on the production line, it was determined that the assembly and processing speeds would not interfere with the production processes [21,22]. The vision system consists of four monochrome cameras with a resolution of 12 MP, complemented by a high-performance chip and real-time GigE communication for transmitting uncompressed images. The industrial-grade cameras can operate in temperatures of up to 50 °C and humidity of up to 80% and are resistant to dust and dirt particles. Each camera is complemented by a 12.5 mm lens with a 625 nm red light filter and a set of LED reflectors with the same wavelength. A high-end industrial workstation with graphics cards is used to manage the vision system in real time.

Two cameras are positioned to cover the cabin, with one focused on the upper section and the other on the lower section. Based on the characteristics of the equipment used and the space available, the maximum vertical field of vision (FOV) of the cameras is 0.79 m. The following equation is therefore used:

$$FOV = \frac{WD \cdot H}{f} \quad (1)$$

where WD is the working distance, H is the sensor size, and f is the focal length of the optics [23]. The mounting arrangement for the vision system consists of a base on which two cameras are placed at different heights: 0.8 m and 1.5 m. This enables us to visualise the entire 1.56 m cabin while generating a 10 cm overlap between the cameras' FOV.

2.3. Database

The image database covers three consecutive days of uninterrupted operation of the vision system. It consists of 8092 RGB images, each measuring 4096×3000 pixels, belonging to 674 cabins. There are 63,073 RSWs, which are segmented into single images of

30×30 pixels, as described in Section D. Samples are taken from the database to train the algorithms in the subsequent sections, but the entire database is used to evaluate the results.

2.4. Spot Detection and Segmentation

Although taking individual snapshots of the spot welds provides a high level of detail for defect detection, the time required makes it difficult to implement online in real time [10]. Therefore, images containing multiple spot welds that can be detected and segmented into 30×30 -pixel images are used instead, as this significantly reduces the computational load. To accomplish this, YOLOv8 in Python 3.8 was employed, enabling the development of deep learning (DL) models for tasks such as detection, segmentation, pose estimation, tracking, and classification that demand speed and precision [24].

For segmentation purposes, it is assumed that the spot welds of interest are located within the car door frame. Therefore, a DL model is generated to identify the frame in each new image. The CVAT platform, an online tool focused on supporting image annotation, is used to build the model. A sample of 100 cabin images is used to identify the frame, generating a database in YOLO 1.1 format. This contains information about the position, morphology, and dimensions of the labels. This database is then exported to Python to train a new DL algorithm using YOLOv8. This algorithm analyses new cabin images (Figure 3A), identifies the position of the edge (Figure 3B) and generates a dynamic mask that separates the edge from the background (Figure 3C).

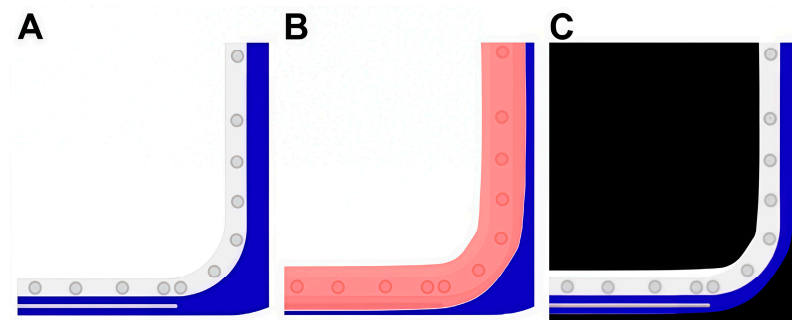


Figure 3. Segmentation of cabin edge using dynamic masking. (A) Original image. (B) Image with labelled edge. (C) Segmentation of area of interest using dynamic masking.

The same process is then used to generate a DL model for identifying spot welds. One hundred cabin images on which the edges have already been isolated are taken, and each spot in each image is labelled manually, as shown in Figure 4. The resulting model can identify the centroid of the spot welds and uses this information to segment them into individual 30×30 -pixel RSW images.

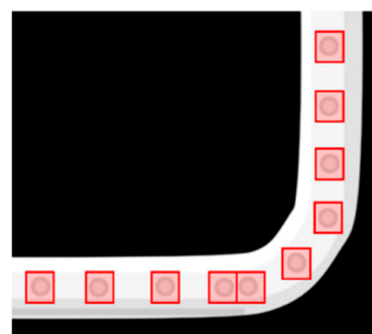


Figure 4. Annotation of the spot weld present on the edge.

2.5. Measurements of Spots' Physical Characteristics

Previous research has shown that nugget diameter and the size of the heat-affected zone (HAZ) are physical parameters that vary according to weld strength, welding current, electrode force, welding time, and other process conditions [25–27]. In RSW, the nugget corresponds to the region of the joint between the sheets that are internally fused and solidified, whereas the HAZ is the surrounding area that is affected by heat without melting and which is often visible on the surface through changes in colour and texture [28]. In this study, image annotations were generated to identify regions associated with visible surface manifestations of the weld and classify them as HAZ or spot weld features. Using the procedure described in Section D, a DL model was trained to detect these regions and estimate their centroids and enclosing polygons (see Figure 5A,B). The characteristic size of each detected region was estimated from the minimum side length of the corresponding polygon and expressed in pixel units.

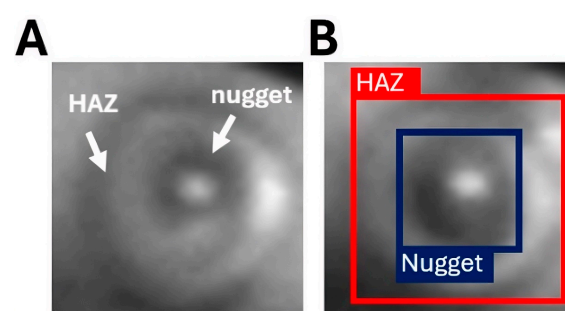


Figure 5. Identification of components of spot weld. (A) Nugget. (B) HAZ.

2.6. Defect Detection

Under normal conditions, a quality weld is characterised by inner and outer elliptical contours and a fusion zone, as well as a smooth, rounded, fusion-free appearance and a diameter of 2.0–3.0 mm [12]. An acceptable spot weld exhibiting these characteristics is shown in Figure 6.

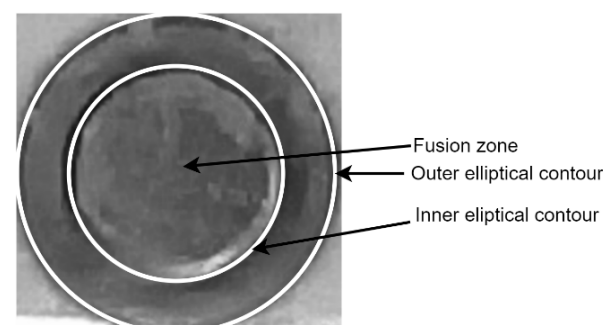


Figure 6. Components of the welding spot.

The fusion zone of the weld nugget is the area inside the inner elliptical circle, while the heat-affected zone is the area inside the outer elliptical circle. Defects include cold spots and poor welds, which occur when the physical variables of the welding machine are not correctly set. Defects also occur when there is no contact between the two welding surfaces and when there is material expulsion in the fusion zone. Figure 7 shows the various defects in the 'bad spot' class. These include deformed metal (excessive electrode current), ejection (burr in the fusion zone), cold spots (lack of current and fusion) and, finally, pitting (prolonged electrode pressure time).

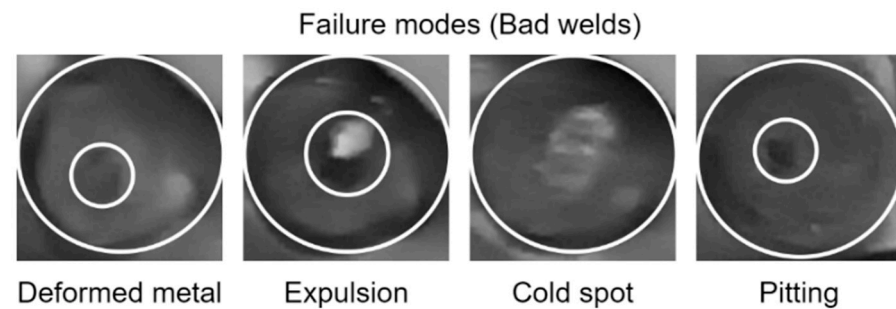


Figure 7. Types of spot weld defects.

In this study, labels of ‘good spot’ or ‘bad spot’ are assigned based on visible surface characteristics of the weld and the presence or absence of surface defects [29]. Therefore, the proposed classification is limited to an assessment of surface appearance and should not be interpreted as a direct evaluation of the mechanical integrity of the weld, since destructive testing was not performed as part of this study.

The process shown in Figure 8 is implemented for defect detection, using a sample from the welding spot database obtained in Section D for training, testing and validating the model for class labelling (good or bad spot). The RSW images selected for the classification process were taken using different welding robots on different days to account for variations in position, illumination and distribution, among other variables. Linear discriminant analysis (LDA) is then used to reduce the dimensionality of the data and define a two-dimensional projection. The data obtained using the LDA technique are then used in a learning model consisting of cross-validation with five iterations and a neural network with one neuron in the input layer, fifteen neurons in the hidden layer and two neurons in the output layer.

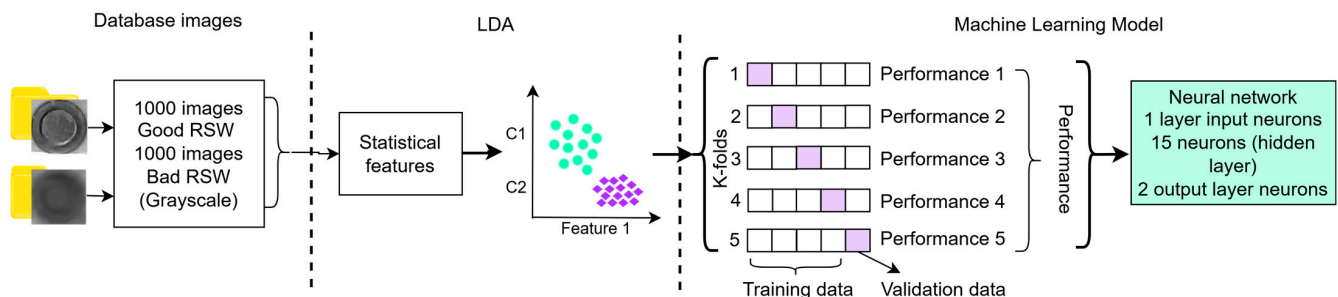


Figure 8. Steps of spot classification model.

The features used are shown in Table 1, where n, k is the number of data points, σ is the standard deviation, x is the data vector, $\bar{x}$ is the mean of the data, μ is the ordinary moment of order 1, $s(k)$ is the sample value, and g is the sample mean [30].

The selected statistical features characterise the intensity distribution within the spot weld image, providing an indirect representation of the welding conditions. As the visual appearance of the weld and heat-affected zone (HAZ) is affected by factors such as heat input, nugget formation, electrode condition and process stability, changes to these physical phenomena are reflected in the spatial distribution of pixel intensities and texture patterns. Specifically, mean intensity and energy describe the brightness of the weld region, which is associated with thermal input and surface appearance. The standard deviation quantifies intensity variability and can reveal non-uniform heat distribution. Meanwhile, skewness and kurtosis characterise asymmetries and concentration effects in the intensity values, which may be related to irregular weld formation, variations in nugget development or modifications to the HAZ. Additionally, higher-order moments provide complementary

information regarding the distribution and concentration of intensity values within the weld region. Therefore, these descriptors capture visual characteristics that are relevant to weld quality and process stability, making them suitable for subsequent dimensionality reduction and classification.

Table 1. Statistical features.

No.	Feature	Equation	Description
1	Mean intensity	$f_m = \frac{1}{n} * \sum_{k=1}^n (x_k - \bar{x})^2$	Measures the average brightness of the weld area, which can indicate uniformity in heat distribution.
2	Standard deviation	$f = \sqrt{\frac{\sum_{x=1}^m \sum_{y=1}^n [I(x,y) - \mu]^2}{m*n}}$	Quantifies the variation in brightness within the weld, helping to detect irregularities.
3	Skewness	$f = \frac{\sum_{x=1}^m \sum_{y=1}^n [I(x,y) - \mu]^3}{m*n*\sigma^2}$	Evaluates the asymmetry in the intensity distribution of the weld. An asymmetrical pattern may indicate uneven heating.
4	Kurtosis	$f = \frac{\sum_{x=1}^m \sum_{y=1}^n [I(x,y) - \mu]^4}{m*n*\sigma^4}$	Identifies the degree of concentration of light energy (brightness or contrast), which is associated with the definition of the weld nugget and the uniformity of the thermal process.
5	Energy	$f = \frac{\sum_{x=1}^m \sum_{y=1}^n I(x,y)^2}{m*n}$	Provides information on image intensity, highlighting regions where the electrode heat is strongest.
6	Third moment	$ft = \frac{1}{k(\sqrt{fm^2})} \sum_{k=1}^k (s(k) - g)^3$	Quantifies information related to the shape and distribution of intensity values.
7	Fourth moment	$fc = \frac{1}{k(\sqrt{fm^2})} \sum_{k=1}^k (s(k) - g)^4$	Quantifies information related to the shape and distribution of intensity values.

The LDA technique (Figure 9) is used to maximize the separation between classes (minimising S_w) while keeping the classes as compact as possible (maximising S_b) [31].

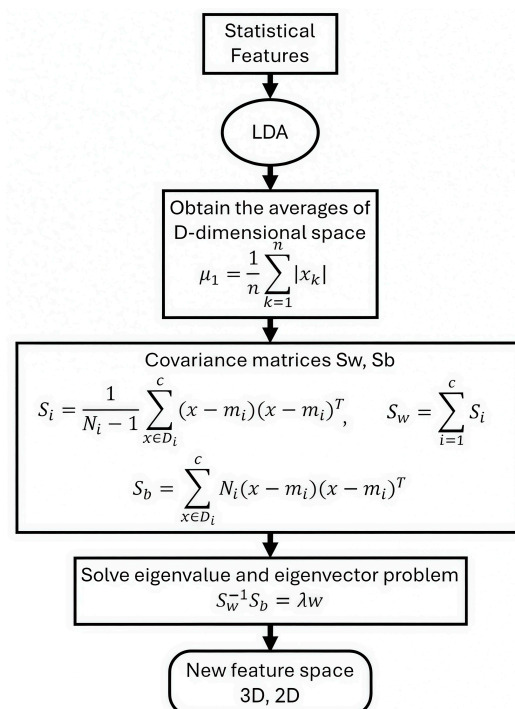


Figure 9. LDA process.

2.7. Maintenance Prognosis

Using the failure classification from the previous section, it is possible to determine how frequently failures occurred at specific robots and welding points. Based on this, it can be assumed that there may be trends in failure occurrence that reflect the need for maintenance. The application of maintenance is based on ISO 18278-1:2022 [32], which states that failures in a vehicle chassis should be kept below 1% for the system to be considered efficient and not require maintenance. However, when this limit is exceeded, it is necessary to determine whether the failure was spontaneous and caused by an external factor, or whether it is frequent and characterised by a trend. To predict trends, a CNN is generated using a Keras API that takes as input the number of good and bad spots, the average nugget diameter and HAZ diameter. This considers changes over time for each spot weld, averaged over a cumulative period of 30 production hours collected on different days under different operating conditions. It naturally incorporates variations such as changes in illumination, minor camera vibrations, changes in weld surface reflectivity, and tolerances in workpiece positioning.

To train the network, the maintenance categories were labelled according to the defect rate thresholds presented in Table 2. Based on the quality criteria established for the welding process, four maintenance levels were defined: ‘no maintenance’ for defect rates below 0.1%, ‘preventive maintenance’ for defect rates between 0.2% and 0.4%, ‘predictive maintenance’ for defect rates between 0.6% and 0.9%, and ‘corrective maintenance’ for defect rates equal to or greater than 1.0%. These categories were then used to label the historical trends employed for network training. Figure 10 shows examples of each maintenance category and how the defect rate evolves over time.

Table 2. Quantitative criteria used for maintenance category labelling.

Maintenance Category Label	Defect Rate
No Maintenance	<0.1%
Preventive Maintenance	0.2–0.4%
Predictive Maintenance	0.6–0.9%
Corrective Maintenance	≥1.0%

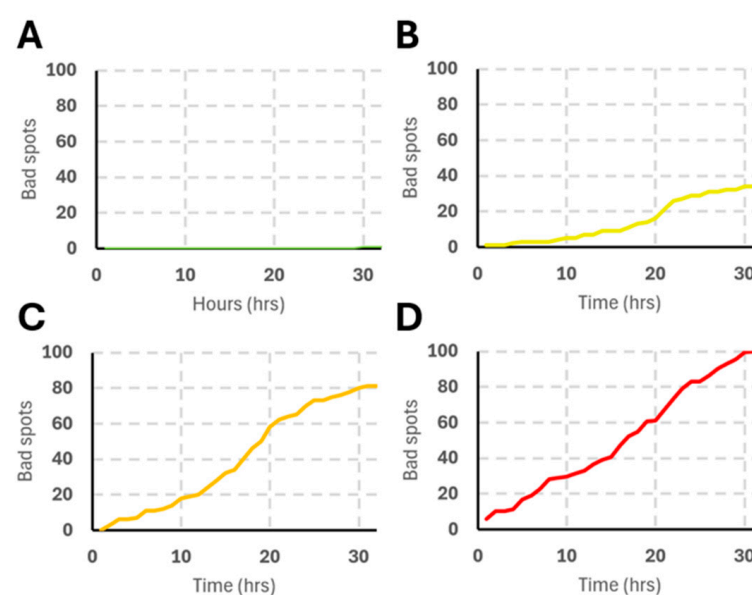


Figure 10. Types of failure frequency trends. (A) No maintenance. (B) Preventive maintenance. (C) Predictive maintenance. (D) Corrective maintenance.

3. Results

3.1. Defect Classification Performance

To identify the size of the nugget and HAZ, independent models were trained using the same methodology. One hundred RSW images were used for training and thirty for validation, with the boundaries of the nugget and HAZ labelled manually. These models were then used to evaluate new RSW images, determining the centroids of the nugget and HAZ and generating a polygon adapted to their morphology. The diameter is determined by taking the smaller of these two dimensions, resulting in average errors of 5% and 3% for the nugget and HAZ diameters, respectively.

The number of RSW images per class was limited to 1000 to train and validate the neural network, as further increases in image numbers were found to no longer yield significant performance gains. For each class, 800 images were used for training and 200 for validation. The statistical features described in Table 1 were subsequently transformed using linear discriminant analysis (LDA) to maximize class separability. Figure 11 presents the resulting feature space, where the projected samples from the ‘good’ and ‘bad’ classes exhibit a clear separation. This distribution indicates that the extracted descriptors contain sufficient discriminative information to distinguish weld quality conditions and can therefore be used as inputs for the classification stage.

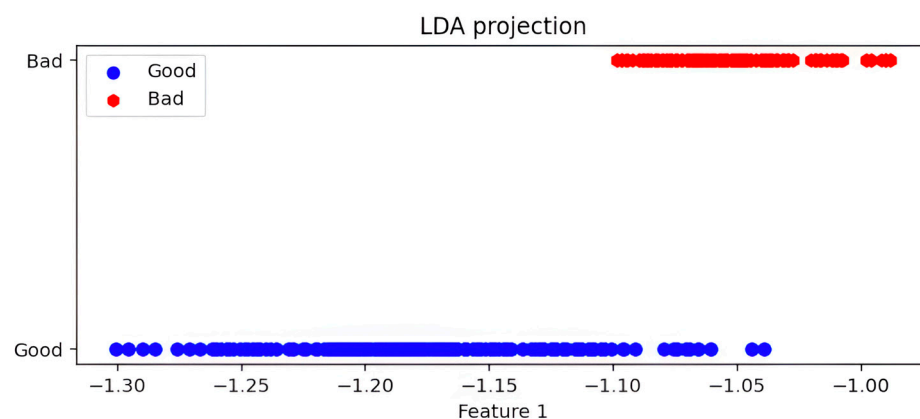


Figure 11. Dispersion of data resulting from model with LDA.

As outlined in the methodology, a neural network consisting of one neuron in the input layer, fifteen neurons in the hidden layer and two neurons in the output layer was trained using the two discriminant features obtained from the LDA projection. Although LDA provides linear class separation, some overlap remains due to the dispersion of the projected samples. Direct comparison showed that classification based solely on the LDA linear discriminant boundary was 1.8% less accurate than classification using the neural network. Therefore, the hidden layer was introduced to capture non-linear decision boundaries within the reduced feature space, thereby improving the separation of samples located near class boundaries.

Figure 12 shows the confusion matrix obtained with the proposed classifier. The resulting model achieved an accuracy of 95.80%, a precision of 95.61%, a recall of 96.00%, and an F1-score of 95.80%, while maintaining a classification time of approximately 0.5 s per cabin. The trained classifier was subsequently used to evaluate, label, and compile all RSW images into a database containing the robot identifier, timestamp, and weld point ID for use in Section 3.2.

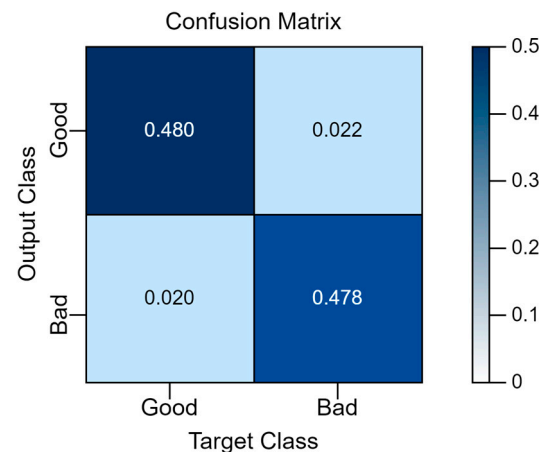


Figure 12. Confusion matrix NN.

3.2. Performance of Maintenance Prognosis

The YOLOv8 model was chosen for edge segmentation and spot detection because, at the time of the experiment, it was the first version of the YOLO family to offer robust instance segmentation based on polygonal masks, as well as a variety of model sizes suitable for deployment under different computational constraints [33,34]. To justify this selection, YOLOv8 was quantitatively compared with YOLOv9 and YOLOv10 using the same dataset. As shown in Table 3, YOLOv8 achieved a segmentation precision of 99.78%, a recall of 100% and an F1-score of 99.88%. This performance is comparable to that of YOLOv9.

Table 3. Comparison of YOLO-based models for RSW detection and segmentation.

Model	Segmentation			Detection		
	Precision	Recall	F1-Score	Precision	F1-Score	Precision
YOLOv8s	99.78%	100.00%	99.88%	89.81%	90.63%	90.21%
YOLOv9c	99.79%	100.00%	99.89%	90.92%	89.31%	90.11%
YOLOv10s	N/A ^a	N/A ^a	N/A ^a	93.24%	87.94%	90.51%

^a N/A = not applicable.

Although YOLOv9 obtained slightly higher segmentation metrics ($\pm 0.01\%$), these differences were negligible, yet YOLOv9 required larger computational resources and longer training and inference times due to the limited availability of model size configurations. YOLOv10 achieved the highest object detection precision (93.24%) but does not provide native instance segmentation capabilities. This makes it necessary to employ an additional segmentation model. This would increase system complexity without significantly improving overall performance. Therefore, YOLOv8 was adopted as the final solution because it offers an optimal balance between segmentation accuracy, object detection performance, computational efficiency and implementation simplicity.

The trained YOLOv8 model was subsequently used to isolate individual RSW points from cabin images. A training dataset of 100 cabin images and a validation dataset of 30 images were collected, including images of different cabin sections, illumination conditions and defect scenarios. The model was trained for 500 epochs to identify cabin boundaries and detect the centroid of each weld spot. These centroids were then processed using OpenCV to segment and extract individual spot images with a resolution of 30×30 pixels, achieving an average processing time of 5.88 s per cabin.

The segmentation process was applied to all acquired images to identify and extract the corresponding weld regions, resulting in a database of 63,073 RSW images. This database

was subsequently reduced in size to optimize computational requirements and analysis time for the subsequent classification and maintenance prediction stages.

To identify defect trends, the information collected from the 63,073 spot welds is considered. The information presented in the database includes the number of the robot that placed the point, the ID of the point checked, the day it was checked, the time it was checked, how many good and bad points were obtained, and the average of the nugget diameter and the HAZ for both cases. This was compiled in 1 h periods for a total span of 30 h, as shown in the database extract in Table 4.

Table 4. Extract from welding spot database.

Robot No.	Day	Hour	No. of Good Spots	Good Spot Mean Diameter	No. of Bad Spots	Bad Spot Mean Diameter
1	3	9	17	26.30	0	0
1	3	10	3	26.30	0	0
1	3	11	2	26.30	0	0
1	4	23	21	26.32	0	0
1	4	0	17	26.28	3	26.39
1	4	1	20	26.31	3	26.37

This database is used to train a convolutional neural network (CNN) model that takes time-series data as input to identify defect evolution trends over a three-day period. To determine the most suitable CNN architecture, an ablation study was performed by modifying several hyperparameters, including the number of hidden layers, filter dimensions, activation functions, learning rate, and dropout configuration. Model performance was evaluated using accuracy as the primary optimization metric. The results of the ablation study are presented in Table 5, which summarizes the evaluated architectures and their corresponding performance.

Table 5. Results of ablation test.

Test No.		1	2	3	4	5
First layer	Filter Activation function	24 ReLu	64 Sigmoid	24 Sigmoid	64 Sigmoid	44 ReLu
Second layer	Filter Activation function	24 Sigmoid	64 ReLu	4 ReLu	--- ---	4 ReLu
Third layer	Filter Activation function	44 ReLu	24 ReLu	4 ReLu	--- ---	4 ReLu
Fourth layer	Filter Activation function	--- ---	4 Sigmoid	4 ReLu	--- ---	--- ---
Output layer	Dropout Learning rate	True 0.0005	True 0.0057	True 0.0023	True 0.0016	True 0.0002
Accuracy %		94.44	88.88	87.50	59.6	59.72

To assess the suitability of the CNN architecture for time-series prognosis, additional experiments were conducted using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, which are widely employed for modelling temporal dependencies. All models were trained and evaluated using the same dataset and preprocessing procedure to ensure a fair comparison. The quantitative results are presented in Table 6.

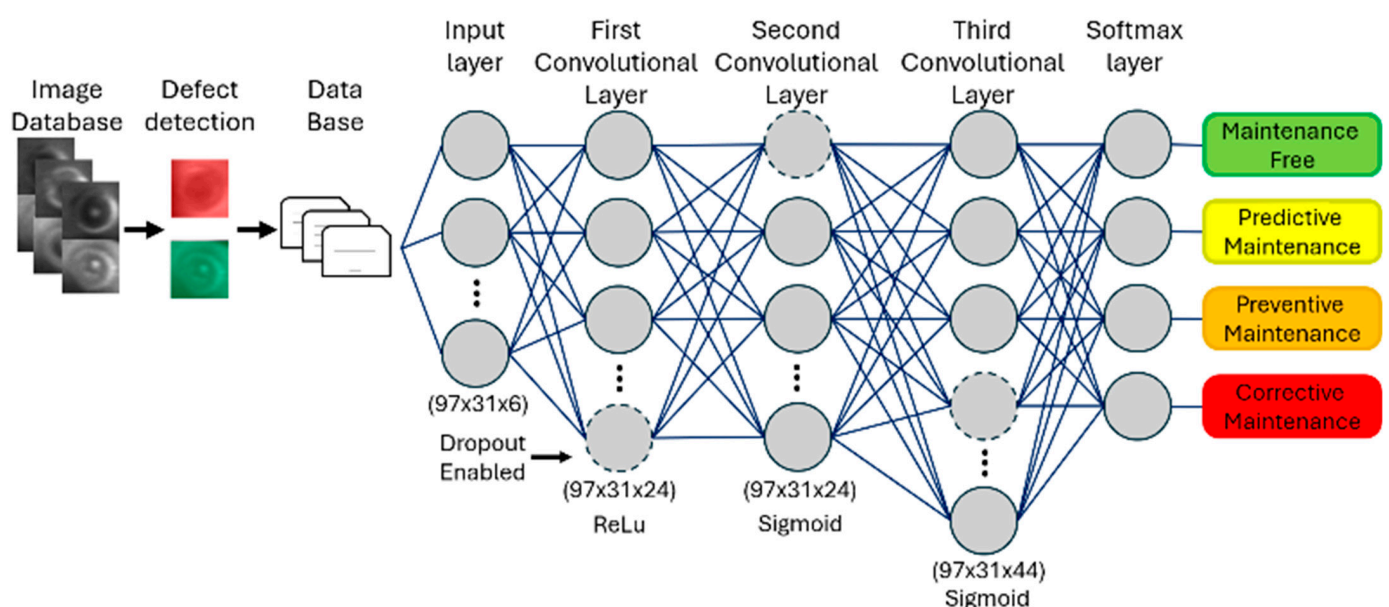
Table 6. Performance comparison models for defect trend prediction.

Model	Precision	Accuracy	Recall	F1-Score
CNN	94.72%	95.87%	94.42%	93.97%
LSTM	96.94%	98.97%	97.52%	97.19%
GRU	96.94%	98.97%	97.75%	97.32%

Although the recurrent architectures produced slightly higher classification metrics, the proposed maintenance prognosis is based on a short observation horizon of three days, where local temporal patterns are more relevant than long-term dependencies. Under these conditions, CNNs effectively capture discriminative temporal features while maintaining a lower computational complexity and reduced training requirements compared with recurrent architectures. Furthermore, the selected CNN model achieved an average processing time of 0.5 s for maintenance prognosis, making it suitable for online industrial monitoring applications where fast response times are required.

The results indicate that LSTM and GRU networks are advantageous when long-term temporal dependencies must be modelled; however, for the proposed application, the CNN architecture provides a favourable trade-off between prediction performance [35], computational efficiency, implementation simplicity, and real-time deployment capability. Therefore, the CNN model was selected as the final solution for maintenance prognosis.

Based on the results obtained from the ablation study, the hyperparameters corresponding to Test 1 were selected as the final CNN architecture. The resulting network is shown in Figure 13, while its corresponding confusion matrix is presented in Figure 14. The confusion matrix demonstrates the model's ability to correctly classify the maintenance categories, exhibiting a low level of confusion among classes and confirming the effectiveness of the selected architecture for maintenance prognosis.

**Figure 13.** Structure of algorithm for maintenance order prognosis.

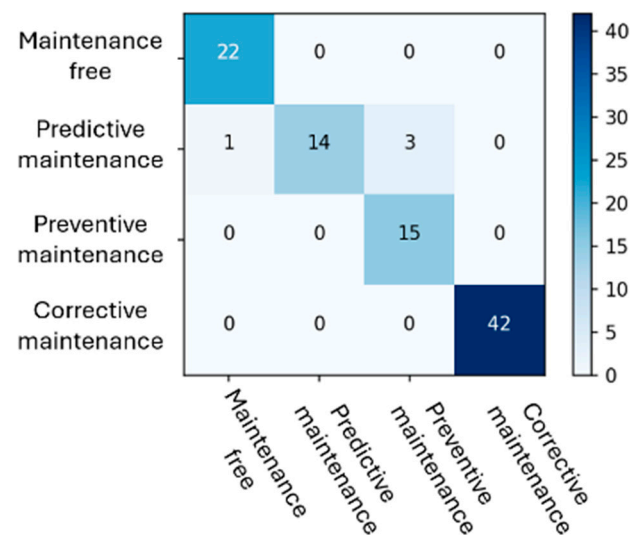


Figure 14. Confusion matrix for maintenance prognosis.

4. Discussion

An online, real-time artificial vision system has been implemented to detect defects in welding points, identify trends, and predict maintenance tasks, optimizing time and resources. The results presented in Section 3 confirm that the system satisfies the required production cycle times by performing image acquisition and analysis in real time. A comparison with previous studies regarding performance and capabilities is presented in Table 7. Although direct comparisons should be interpreted carefully due to differences in sensors, datasets, and application objectives, the proposed system demonstrates competitive performance while providing the additional advantage of online operation in an industrial environment.

Compared with vision-based approaches using digital cameras [11,17,36], the proposed method achieves comparable or improved classification performance while maintaining real-time processing capabilities. Furthermore, while the work of Boriwal et al. [37] achieves higher accuracy, it relies on ultrasonic sensor data and performs offline analysis, which makes direct comparison of the two approaches difficult. The main contribution of the proposed system lies not only in its classification performance, but also in its ability to integrate defect detection, weld quality assessment, trend analysis and maintenance prediction within a single online framework that is suitable for industrial deployment. The combination of real-time operation and predictive capabilities underscores the system's practical value for monitoring RSW and managing maintenance.

The methodology used by each author to classify the welds has been considered a comparative factor, since no methods have been implemented to predict the required maintenance. These methods focus on the use of fuzzy support machines and artificial intelligence models, which require a high level of time and processing power for training and validation. They are also systems that are implemented in laboratories and offline processes or use other methods such as ultrasound. Although the accuracy of the proposed system is slightly lower in some cases, it should be emphasised that this system is evaluated from image acquisition to maintenance prediction and that its main advantages are that it is fully implemented within an automotive production line and runs online. Nevertheless, it demonstrates an impressive ability to accurately identify positive cases without compromising precision. It shows the potential of deep learning-based approaches for online applications, offering a more robust and reliable solution than traditional techniques.

Table 7. Comparison of classification system performance.

Author	Sensor	Algorithms	Implementation	Precision (%)	Recall (%)	Accuracy (%)	F1-Score (%)
Boriwal et al. [37]	Ultrasonic sensor	Manual analysis	Outside of line	95.00	---	---	---
Alghannam et al. [11]	Camera	Support Vector Machine	Outside of line	90.00	---	---	---
Younes et al. [12]	Ultrasonic sensor	Fuzzy Support Vector Machine	In line	93.00	---	---	---
Dai et al. [17]	Camera	Faster R-CNN	Laboratory	92.04	---	---	---
Dai et al. [36]	Camera	BAGAN, GP Y ResNet-50	Laboratory	87.60	83.40	87.60	83.90
Implemented system	Camera	LDA-NN, YOLOv8-CNN	In line	94.72	94.42	95.87	93.97

The presented system has some limitations, such as the fact that the database only classifies spots into two categories, which is why there are some defects in it. Therefore, continuing the project would involve increasing the number of spot welds in the database and focusing on specific defects. Additionally, intermittent use of the vision system makes it challenging to compile a comprehensive database over extended periods. The aim would therefore be to create a structured database with fixed time periods for analysing the cabin. Furthermore, the exact impact on the performance of the production line is unknown, so integration is a priority to monitor the project and estimate its benefits to the company. To achieve this, it is important to measure changes in the spot welding correction tasks over time and the frequency with which robot maintenance is applied.

Another limitation of the present study is that the mechanical integrity of the spot welds classified by the vision system was not verified by destructive testing. Consequently, the classification results are based exclusively on the visual appearance of the weld surface. Future work will involve subjecting representative samples of welds from each quality category to destructive tests to validate the correlation between visual classification and actual mechanical strength. This additional validation will enhance confidence in the system's ability to assess weld quality, contributing to the establishment of a robust relationship between surface characteristics and weld integrity. Additionally, although the proposed system was evaluated using data acquired under real production conditions, a systematic robustness assessment under controlled perturbations was beyond the scope of the present study. Factors such as artificial illumination variations, camera micro displacements, and surface contamination were not independently introduced and analysed during the experimental phase. While the system demonstrated stable performance in the operational environment, future work will include a comprehensive robustness evaluation under controlled interference conditions. Such an assessment will provide a deeper understanding of the system's resilience to environmental and operational variations and will further support its long-term deployment in industrial production lines.

5. Conclusions

The proposed system uses machine vision to analyse spot weld and the frequency with which defects occur on different robots in real production conditions. Research has identified this as a limitation, as it requires fast processing and a robust system that can handle variations in the mass production process. This issue was resolved by utilising several open-source machine learning (ML) and deep learning (DL) tools. YOLOv8 was implemented to provide robust identification, measurement and segmentation of spot

welds with an error margin of 5%. These are then processed by an optimised NN with dimensionality reduction techniques to classify them as good or bad spots according to their physical characteristics, achieving an accuracy of 95.8%. This enables the frequency of occurrence to be monitored and trends to be identified through the NN, generating a forecast of maintenance orders according to the recurrence of failure with an accuracy of 95.87%, a precision of 94.72%, an F1-score of 93.97% and a recall of 94.42%. It is expected that the developed system will be fully integrated into the line inspection process with an average total processing time of 7 s per cabin. This could reduce production downtime and economic losses while improving the management of human resources and consumables used in the cabin assembly and robot maintenance processes.

Author Contributions: Conceptualization, L.A.M.H., J.C.J.-C. and I.A.C.-A.; methodology, all authors; software, A.A.-R. and R.C.-M.; validation, L.A.M.H., J.C.J.-C. and I.A.C.-A.; formal analysis, A.A.-R. and R.C.-M.; investigation, all authors; resources, L.A.M.H., J.C.J.-C. and I.A.C.-A.; data curation, A.A.-R. and R.C.-M.; writing—original draft preparation, all authors; writing—review and editing, all authors; visualization, L.A.M.H., J.C.J.-C. and I.A.C.-A.; supervision, L.A.M.H., J.C.J.-C. and I.A.C.-A.; project administration, L.A.M.H. and J.C.J.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The first author would like to express his sincere gratitude to the Secretary of Science, Humanities, Technology and Innovation (SECIHTI) for the scholarship awarded (CVU 1241430).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, B. A study on spot welding quality judgment based on improved generative adversarial network and auto-encoder. *Mech. Syst. Signal Process.* **2022**, *165*, 108318. [\[CrossRef\]](#)
2. Johnson, N.N.; Madhavadas, V.; Asati, B.; Giri, A.; Hanumant, S.A.; Shajan, N.; Arora, K.S.; Selvaraj, S.K. Implementation of Machine Learning Algorithms for Weld Quality Prediction and Optimization in Resistance Spot Welding. *J. Mater. Eng. Perform.* **2023**, *33*, 6561–6585. [\[CrossRef\]](#)
3. Soomro, I.A.; Pedapati, S.R.; Awang, M. A review of advances in resistance spot welding of automotive sheet steels: Emerging methods to improve joint mechanical performance. *Int. J. Adv. Manuf. Technol.* **2021**, *118*, 1335–1366. [\[CrossRef\]](#)
4. Banga, H.K.; Kalra, P.; Kumar, R.; Singh, S.; Pruncu, C.I. Optimization of the cycle time of robotics resistance spot welding for automotive applications. *J. Adv. Manuf. Process* **2021**, *3*, e10084. [\[CrossRef\]](#)
5. Bousnina, K.; Hamza, A.; Ben Yahia, N. Effect of vibration and welding parameters on spot weld resistance: Modeling integrating PSO-ANN and GA algorithm. *Multiscale Multidiscip. Model. Exp. Des.* **2023**, *7*, 1755–1767. [\[CrossRef\]](#)
6. Zaharuddin, M.F.A.; Kim, D.; Rhee, S. An ANFIS based approach for predicting the weld strength of resistance spot welding in artificial intelligence development. *J. Mech. Sci. Technol.* **2017**, *31*, 5467–5476. [\[CrossRef\]](#)
7. Summerville, C.; Compston, P.; Doolan, M. A comparison of resistance spot weld quality assessment techniques. *Procedia Manuf.* **2019**, *29*, 305–312. [\[CrossRef\]](#)
8. Martin, O.; Ahedo, V.; Santos, J.I.; Galan, J.M. Comparative Study of Classification Algorithms for Quality Assessment of Resistance Spot Welding Joints From Pre- and Post-Welding Inputs. *IEEE Access* **2022**, *10*, 6518–6527. [\[CrossRef\]](#)
9. Guo, Q.; Yang, Z.; Xu, J.; Jiang, Y.; Wang, W.; Liu, Z.; Zhao, W.; Sun, Y. Progress, challenges and trends on vision sensing technologies in automatic/intelligent robotic welding: State-of-the-art review. *Robot. Comput.-Integr. Manuf.* **2024**, *89*, 102767. [\[CrossRef\]](#)
10. Choi, S.G.; Hwang, I.; Kim, Y.M.; Kang, B.; Kang, M. Prediction of the weld qualities using surface appearance image in resistance spot welding. *Metals* **2019**, *9*, 831. [\[CrossRef\]](#)
11. Alghannam, E.; Lu, H.; Ma, M.; Cheng, Q.; Gonzalez, A.A.; Zang, Y.; Li, S. A novel method of using vision system and fuzzy logic for quality estimation of resistance spot welding. *Symmetry* **2019**, *11*, 990. [\[CrossRef\]](#)

12. Younes, D.; Alghannam, E.; Tan, Y.; Lu, H. Enhancement in quality estimation of resistance spot welding using vision system and fuzzy support vector machine. *Symmetry* **2020**, *12*, 1380. [\[CrossRef\]](#)
13. Xiao, M.; Yang, B.; Wang, S.; Chang, Y.; Li, S.; Yi, G. Research on recognition methods of spot-welding surface appearances based on transfer learning and a lightweight high-precision convolutional neural network. *J. Intell. Manuf.* **2022**, *34*, 2153–2170. [\[CrossRef\]](#)
14. Stavropoulos, P.; Sabatakakis, K.; Papacharalampopoulos, A.; Mourtzis, D. Infrared (IR) quality assessment of robotized resistance spot welding based on machine learning. *Int. J. Adv. Manuf. Technol.* **2021**, *119*, 1785–1806. [\[CrossRef\]](#)
15. Lee, J.; Noh, I.; Lee, J.; Lee, S.W. Development of an Explainable Fault Diagnosis Framework Based on Sensor Data Imagification: A Case Study of the Robotic Spot-Welding Process. *IEEE Trans. Ind. Inf.* **2021**, *18*, 6895–6904. [\[CrossRef\]](#)
16. Li, Q.; Yang, B.; Wang, S.; Zhang, Z.; Tang, X.; Zhao, C. A fine-grained flexible graph convolution network for visual inspection of resistance spot welds using cross-domain features. *J. Manuf. Process* **2022**, *78*, 319–329. [\[CrossRef\]](#)
17. Dai, W.; Li, D.; Tang, D.; Jiang, Q.; Wang, D.; Wang, H.; Peng, Y. Deep learning assisted vision inspection of resistance spot welds. *J. Manuf. Process* **2021**, *62*, 262–274. [\[CrossRef\]](#)
18. Yang, B.; Peng, Q.; Zhang, Z.; Zhang, Y.; Li, Y.; Xi, Z. An offset-transformer hierarchical model for point cloud-based resistance spot welding quality classification. *Comput. Ind.* **2024**, *161*, 104134. [\[CrossRef\]](#)
19. Chian, Y.T.E.; Tian, J. Surface Defect Inspection in Images Using Statistical Patches Fusion and Deeply Learned Features. *AI* **2021**, *2*, 17–31. [\[CrossRef\]](#)
20. Diaconu, B.M. Diagnosing Shortcut Learning in CNN-Based Photovoltaic Fault Recognition from RGB Images: A Multi-Method Explainability Audit. *AI* **2026**, *7*, 94. [\[CrossRef\]](#)
21. Simončič, S.; Podržaj, P. Resistance spot weld strength estimation based on electrode tip displacement/velocity curve obtained by image processing. *Sci. Technol. Weld. Join.* **2014**, *19*, 468–475. [\[CrossRef\]](#)
22. Ruisz, J.; Biber, J.; Loipetsberger, M. Quality evaluation in resistance spot welding by analysing the weld fingerprint on metal bands by computer vision. *Int. J. Adv. Manuf. Technol.* **2006**, *33*, 952–960. [\[CrossRef\]](#)
23. Pajares, G.; de la Cruz, J.M. *Visión Por Computador, Imágenes Digitales y Aplicaciones*; Alfaomega: Mexico City, Mexico, 2007.
24. Valaee-Tale, M.; Sheikhi, M.; Mazaheri, Y.; Malek Ghaini, F.; Usefifar, G.R. Criterion for predicting expulsion in resistance spot welding of steel sheets. *J. Mater. Process. Technol.* **2020**, *275*, 116329. [\[CrossRef\]](#)
25. Kim, S.; Hwang, I.; Kim, D.Y.; Kim, Y.M.; Kang, M.; Yu, J. Weld-quality prediction algorithm based on multiple models using process signals in resistance spot welding. *Metals* **2021**, *11*, 1459. [\[CrossRef\]](#)
26. Dejans, A.; Kurtov, O.; Van Rymenant, P. Acoustic emission as a tool for prediction of nugget diameter in resistance spot welding. *J. Manuf. Process.* **2020**, *62*, 7–17. [\[CrossRef\]](#)
27. Farson, D.F.; Chen, J.Z.; And, K.E.; Frech, T. Monitoring Resistance Spot Nugget Size by Electrode Displacement. *J. Manuf. Sci. Eng.* **2004**, *126*, 391–394. [\[CrossRef\]](#)
28. Muhammad, N.; Manurung, Y.H.P.; Hafidzi, M.; Abas, S.K.; Tham, G.; Haruman, E. Optimization and modeling of spot welding parameters with simultaneous multiple response consideration using multi-objective Taguchi method and RSM. *J. Mech. Sci. Technol.* **2012**, *26*, 2365–2370. [\[CrossRef\]](#)
29. Liu, W.; Hu, J.; Qi, J. Resistance spot welding defect detection based on visual inspection: Improved faster R-CNN model. *Machines* **2025**, *13*, 33. [\[CrossRef\]](#)
30. Zhang, H.; Senkara, J. *Resistance Welding: Fundamentals and Applications*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2011.
31. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2000.
32. ISO 18278-1:2022; Resistance Welding—Weldability—Part 1: General Requirements for Evaluation of Resistance Spot, Seam and Projection Welding. International Organization for Standardization: Geneva, Switzerland, 2022.
33. Hussain, M. YOLOv1 to v8: Unveiling each variant—A comprehensive review of YOLO. *IEEE Access Pract. Innov. Open Solut.* **2024**, *12*, 42816–42833. [\[CrossRef\]](#)
34. Varghese, R.; M., S. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. In Proceedings of the 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 18–19 April 2024; pp. 1–6. [\[CrossRef\]](#)
35. Lee, J.; Hwang, H.; Jeong, T.; Kim, D.; Ahn, J.; Lee, G.; Lee, S.H. Review on welding process monitoring based on deep learning using time-series data. *J. Weld. Join.* **2024**, *42*, 333–344. [\[CrossRef\]](#)
36. Dai, W.; Li, D.; Tang, D.; Wang, H.; Peng, Y. Deep learning approach for defective spot welds classification using small and class-imbalanced datasets. *Neurocomputing* **2022**, *477*, 46–60. [\[CrossRef\]](#)
37. Boriwal, L.; Sarviya, R.M.; Mahapatra, M.M. Failure modes of spot welds in quasi—Static tensile—Shear loading of coated steel sheets. Shear I. In *Materials Today: Proceedings*; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; pp. 3672–3677. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.