COMPUTER AIDED IDENTIFICATION AND CLASSIFICATION OF WELDING DEFECTS

Dominic O. Cagadas* and Marjun S. Sequera
University of Science and Technology of Southern Philippines-Cagayan de Oro
Cagayan de Oro City, 9000, Philippines
* dominic.cagadas@ustp.edu.ph

Date of publication: 25 July, 2026

***ABSTRACT** - Visual inspection of welding connections is traditionally subjective, labor-intensive, and susceptible to human error. This study introduces an automated computer vision system utilizing a deep Convolutional Neural Network (CNN) based on the VGG19 architecture for identifying and classifying welding defects. A dataset of surface weld images was collected and preprocessed to train and evaluate the model across four target categories: Crack, Good Weld, Porosity, and Spatter. The fine-tuned VGG19 model achieved an overall classification accuracy of 88.76% on an independent test dataset of 89 samples. Specific class performances demonstrated high precision and recall, with the Crack class reaching a 100% recall rate and the Good Weld class achieving 95% precision. The results indicate that the proposed automated vision framework offers a consistent, objective, and reliable solution for quality control and inspection in industrial manufacturing processes.*

Keywords: *Computer vision, Convolutional Neural Network, quality assurance, VGG19, welding defect detection.*

1. INTRODUCTION

In fields like industrial, automotive, aerospace, and construction, welding is a common joining technique. But while welding, a number of defects might appear that affect the durability and reliability of the weld connection. For welded components to be structurally sound and reliable, it is essential to find these faults.

The main technique for locating defects in welding has been visual inspection by human inspectors. Manual inspection, however, is subjective, labor-intensive, and open to human mistake. Automating the detection and classification of defects in welding has been done in order to get around these restrictions [1].

Welding defect detection using image processing is a study that aims to develop a system for classifying welding defects. This study can significantly benefit human inspectors because it automates the identification of welding errors. This [2] presents a method for detecting and classifying defects in weld radiographs using a Neural Network. The method detects discontinuities in weld images, such as false alarms or defects like worm holes, porosity, linear slag inclusion, gas pores, lack of fusion, or crack. A set of 43 descriptors corresponding to texture measurements and geometrical features is extracted for each segmented object, which is then trained to classify them into defect classes or non-defect classes. Three-fold cross validation was used, and experimental results were reported for three different classifiers (Support Vector Machine, Neural Network, k-NN). Around industries, image processing technology is frequently used in quality control procedures to enable automated defect identification and accurate measurement of goods and components. In addition, it is essential for medical imaging, helping with the diagnosis and planning of treatments for a variety of illnesses and ailments. The method presents a novel defect detection method using laser vision and model-based segmentation for laser brazing welds on car body surfaces. This fully automatic solution identifies defects and accurately characterizes their nature, without requiring prior knowledge of the specific defects being inspected. Laser brazing technology has made significant advancements in automotive manufacturing, improving inspection processes and reducing defects [3].

External factors, such as tool tilt angle and material conditions, can induce imperfections in friction stir welding connections. The mistakes can be conventional or unique, with the cause of these defects remaining unknown. Using optical microscopy, energy-dispersive X-ray spectroscopy, and scanning electron microscope, the study indicates that tilt angle alters plastic material flow patterns in the stir zone, altering weld qualities. Additionally, the oxide layer from the original butt surface during friction stir welding is a primary cause of joint failure. Through the utilization of advanced techniques such as optical microscopy, energy-dispersive X-ray spectroscopy, and scanning electron microscopy, researchers have gained valuable insights into the intricate mechanisms underlying the impact of tool tilt angle on the plastic material flow patterns within the stir zone. The stir zone, which is the region subjected to strong plastic deformation and temperature cycles during friction stir welding, plays a critical role in defining the overall quality of the weld. The interplay between the tool tilt angle and the material flow patterns within the stir zone has been discovered to have a direct influence on the mechanical and structural qualities of the weld. Studies have showed that differences in the tilt angle of the tool can disrupt the desired flow patterns within the stir zone. This disruption can result in irregular mixing of the components, poor consolidation of the weld, and the creation of defects such as voids, cracks, or inclusions. These faults weaken the strength, fatigue resistance, and overall structural integrity of the joint, making it more susceptible to early failure [4].

For quality control in welding, a crucial step in automated manufacturing, the car industry primarily relies on machine learning (ML) methods. The adoption of ML approaches is made possible by IoT technology, however there are substantial obstacles due to the lack of equal access to data scientists, engineers, process experts, and managers. SemML, a system for ontology-enhanced ML pipeline creation that uses ontologies and ontology templates for job negotiation and data annotation, is offered as a solution to these problems. In the Bosch use-case of electric resistance welding, the system has shown encouraging results, highlighting the potential of ML-based quality monitoring techniques [5]. Machine vision technology called Automatic

X-ray Inspection (AXI) is used to inspect flaws in electronics and mechanical engineering. It takes the place of the Automated Optical Inspection (AOI) as the last step in the front end process at Continental Temic Automotive Phils Inc. AXIs are better suited for a variety of applications than AOIs because they have a faster cycle time and more flexible inspection restrictions [6].

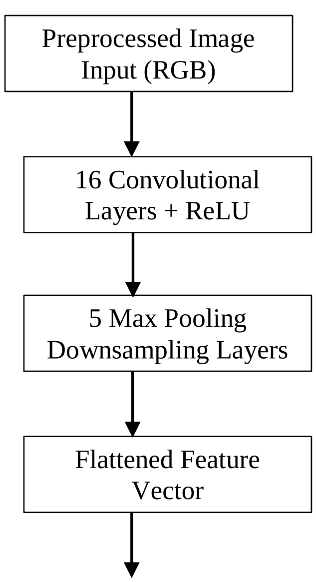
2. METHODOLOGY

The development of the proposed computer vision-based welding defect detection system involves three primary stages: data collection and preprocessing, network architecture and configuration, and model training.

2. Data Collection and Preprocessing

To establish objective visual acceptance criteria for dataset annotation, the four defect categories must be defined according to international welding standards. Under these standards, a Crack (Class)) is defined as a linear fracture resulting from localized stress and carries a strict zero-tolerance policy across all quality levels due to its high structural risk. Porosity (Class 2) refers to gas cavities trapped during solidification, which are identified by their spherical or clustered visual geometry and evaluated based on allowable percentage surface area and pore diameter. Spatter (Class 3) consists of metal droplets ejected onto adjacent surfaces; it is classified as a surface imperfection rather than an internal structural defect, though it must be rejected under strict quality levels if it interferes with non-destructive testing or protective coatings. A Good Weld (Class 1) is defined by complete fusion, a uniform bead profile, and the absence of unacceptable discontinuities, adhering strictly to dimensional limits for reinforcement and toe blending. Finally, image dataset labeling must follow standardized visual testing protocols such as ISO 17637 [7] and AWS B1.11 [8], which dictate minimum light intensity of 350 lux and viewing angles of 30 degrees to eliminate subjective bias during manual image annotation [9].

All collected images were acquired using standard camera sensors under variable lighting conditions. The raw images were preprocessed, resized to uniform spatial dimensions, and normalized to facilitate efficient feature extraction during network backpropagation. An independent test set consisting of 89 samples was withheld specifically for unbiased model evaluation.



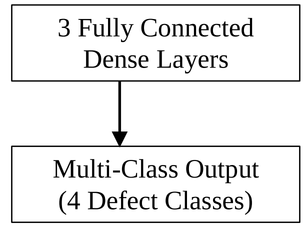


Figure 2. Operational Block Diagram of the VGG19

Network Architecture Configured for Weld Classification.

2.2 Network Architecture and Configuration

The system incorporates the VGG19 architecture, a deep 19-layer Convolutional Neural Network known for strong feature hierarchy representation [10][11] as shown in figure 1. The functional layers are configured as follows: Convolutional Layers: Small 3×3 receptive field filters operate across spatial dimensions to compute feature maps. The mathematical operation for a transformed pixel output using a kernel matrix is defined as:

$$S(i,i) = (I * K)(i,j) = I(i-m,j-n)K(m,n)$$

where I represents the input image matrix and K denotes the convolution kernel.

1. Activation Function: The Rectified Linear Unit ($ReLU$) activation function is applied across all hidden convolutional layers to introduce non-linearity and prevent vanishing gradient issues:

$$f(x) = \max(0,x)$$

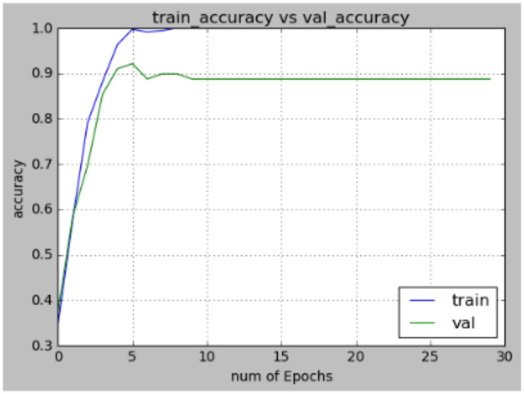
2. Pooling Layers: Max pooling layers with 2×2 spatial windows downsample the activation feature maps, summarizing spatial regions while preserving spatial invariance.
3. Dense and Softmax Layers: Multi-dimensional feature maps are flattened into a dense vector, which passes into fully connected layers. The terminal fully connected layer comprises four output nodes matching the target class count.

To fine-tune a VGG19 model for welding defect identification, transfer learning is applied by initializing the convolutional backbone with pre-trained ImageNet weights to leverage its learned visual primitives like edges and textures. To prevent overfitting and accelerate training, the lower and mid-level convolutional layers are frozen, while the high-level feature extraction block is unfrozen and retrained to adapt specifically to complex defect geometries such as cracks and porosity. The original 1,000-class fully connected head is completely replaced with a streamlined classifier comprising pooling layer, a dense layer with ReLU activation, and a final output layer matched to the specific number of weld defect categories [12].

2.3 Model Training

Training was conducted over 30 epochs. Cross-entropy loss was minimized using an adaptive optimization algorithm. The validation split was continuously monitored across training iterations to track convergence behavior and guard against overfitting.

3. RESULTS AND DISCUSSION



3.1 Overall Model Performance

The performance of the fine-tuned VGG19 model was verified against the 89-sample independent test dataset. The system achieved an overall classification accuracy of 88.76% across all combined classes. Detailed precision, recall, and F1-score breakdown metrics per class are provided in Table 1 below.

Table 1 Classification Performance Metrics of the Fine-Tuned VGG19 Model

Class Identifier	Class Description	Precision	Recall	F1-Score	Support (Samples)
Class 0	Crack	0.81	1.00	0.89	21
Class 1	Good Weld	0.95	0.75	0.84	24
Class 2	Porosity	0.88	0.83	0.86	18
Class 3	Spatter	0.93	0.96	0.94	26
Macro Average	Overall System	0.89	0.89	0.89	89

In an image-based welding defect identification system, evaluation metrics derived from the confusion matrix are essential for measuring model reliability across safety-critical defect categories. Accuracy measures overall correct predictions but can be misleading on imbalanced industrial datasets; therefore, Precision is evaluated to quantify the system's ability to avoid false alarms that cause unnecessary inspection rework costs, while Recall is prioritized as the most critical metric to ensure severe flaws like cracks are not missed, preventing catastrophic structural failures. The F1-Score provides a unified harmonic mean that balances these trade-offs, and together with confusion matrix analysis, validates that the deep learning model achieves both high detection sensitivity and low false-alarm rates [13].

As presented in Table 1, the model demonstrated notable performance in detecting Class 0 (Crack), securing a 100% recall rate. This indicates that all 21 actual crack instances were correctly identified without any false negative omissions—a key requirement for structural safety applications. Class 1 (Good Weld) achieved the highest precision at 0.95, demonstrating high reliability when classifying non-defective welds. Class 3 (Spatter) achieved balanced metrics with an F1-score of 0.94.

3.2 Defect Detection Performance

Figure 2 and Figure 3 depict the system's training and validation accuracy and loss dynamics over 30 epochs.

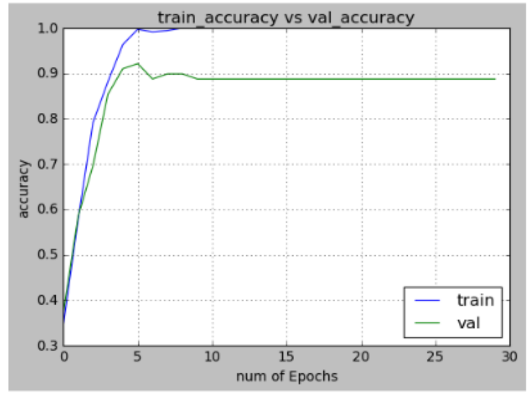


Figure 2. CNN Accuracy
Figure 3. CNN Loss

Training accuracy increased rapidly during the initial 5 epochs, settling near 100%, while validation accuracy stabilized at approximately 88.76%. Correspondingly, validation loss decreased rapidly during initial epochs and reached a steady state around epoch 8. The close alignment between training and validation trends demonstrates stable learning behavior.

3.3 Confusion Matrix Analysis

In an image-based weld defect identification system, a confusion matrix acts as a performance scorecard by tracking how often the system correctly predicts defect types versus confusing them with others such as mistaking a crack for porosity [14]. To assess specific misclassifications across categories, a confusion matrix was compiled for the test set (Table 2).

Table 2 Confusion Matrix Mapping for Test Dataset Predictions

	Predicted: Crack (Class 0)	Predicted: Good Weld (Class 1)	Predicted: Porosity (Class 2)	Predicted: Spatter (Class 3)	Total Actual
Actual: Crack (Class 0)	21	0	0	0	21
Actual: Good Weld (Class 1)	4	18	2	0	24
Actual: Porosity (Class 2)	1	0	15	2	18
Actual: Spatter (Class 3)	0	1	0	25	26

An analysis of the confusion matrix reveals key performance insights across the evaluated classes. Regarding crack reliability, all 21 actual Crack instances were correctly classified, yielding 21 true positives and maintaining a 100% recall rate for Class 0. However, five samples—consisting of four Good Welds and one Porosity instance—were incorrectly categorized as Cracks, which resulted in a lower precision of 0.81 for this class. For spatter detection, the model demonstrated high accuracy by correctly identifying 25 out of 26 true Spatter samples, with only a single sample misclassified as a Good Weld. Finally, minor discrepancies and mutual misclassifications occurred between Good Weld and Porosity samples, primarily due to their overlapping surface visual characteristics.

4. Conclusions and Recommendation

This study demonstrated an automated image processing system using the fine-tuned VGG19 CNN model for detecting and classifying welding surface defects. The system achieved an overall accuracy of 88.76% across four defect categories: Crack, Good Weld, Porosity, and Spatter. The model demonstrated 100% recall in detecting structural cracks, confirming its ability to identify critical joint defects. Deep learning-based visual classification offers a practical, reliable alternative to manual visual inspection in industrial quality assurance workflows.

To build upon the findings of this study, several future directions are recommended. First, expanding the dataset size and variety to include diverse lighting conditions, various surface finishes, and additional defect classes—such as lack of penetration and undercut—would significantly enhance model robustness. Second, deploying the trained network onto portable mobile devices or edge-computing hardware would enable real-time, on-site quality inspection for field welding inspectors. Finally, establishing standardized lighting and camera positioning setups during image capture would effectively minimize environmental noise and boost overall classification accuracy.

When comparing these findings with prior welding-defect studies, the VGG19 visual inspection framework demonstrates notable operational advantages and strong alignment with established literature. Traditional radiographic defect detection methodologies, such as those utilizing multi-descriptor feature extractions paired with machine learning classifiers like Support Vector Machines or k-NN, or deep learning applied to radiographic images, rely heavily on specialized subsurface imaging data. In contrast, the fine-tuned VGG19 approach achieves high visual feature extraction directly from standard surface RGB images acquired under variable lighting conditions, eliminating the need for manual extraction of complex geometric descriptors or costly X-ray equipment. Furthermore, while advanced laser vision systems and real-time object detection models like YOLOv8 require specialized hardware or prioritize detection speed, the proposed model balances precision and recall at 89% while strictly addressing zero-tolerance safety requirements for crack detection under standard visual testing protocols. These results confirm that deep learning-based visual classification offers a practical, consistent, and objective alternative to manual visual inspection in industrial quality assurance workflows.

REFERENCES

[1] Jayawardena, A., Sul, J. H., Moratuwage, D., Wijayaraja, J. L., and Piyathilaka, L. State of the Art in AI-Based Visual Inspection for Industrial Quality Control: Methods, Benchmarks, Challenges, and Autonomous Systems. *Electronics*, Vol. 15, Issue 12, Art. no. 2727, 2026.

[2] Palma-Ramírez, D., Ross-Veitía, B. D., Font-Ariosa, P., Espinel-Hernández, A., Sanchez-Roca, A., Carvajal-Fals, H., Nuñez-Alvarez, J. R. and Hernández-Herrera, H. Deep convolutional neural network for weld defect classification in radiographic images. *Heliyon*, Vol. 10, Issue 9, Art. no. e30590, 2024.

[3] Xue-Wu, Z., Yan-Qiong, D., Yan-Yun, L., Ai-Ye, S., and Rui-Yu, L. A vision inspection system for the surface defects of strongly reflected metal based on multi-class SVM. *Expert Systems with Applications*, Vol. 38, Issue 5, pp. 5930-5939, 2011.

[4] Singh, V. P., Patel, S. K., and Kuriachen, B. Mechanical and microstructural properties evolutions of various alloys welded through cooling assisted friction-stir welding: A review. *Intermetallics*, Vol. 133, Art. no. 107122, 2021.

[5] Chuenmee, N., Phothi, N., Chamniprasart, K., Khaengkarn, S., and Srisertpol, J. Machine learning for predicting resistance spot weld quality in automotive manufacturing. *Results in Engineering*, Vol. 25, Art. no. 103570, 2025.

[6] Masuch, S., Gümbel, P., Kaden, N., and Dröder, K. Applications and development of X-ray inspection techniques in battery cell production. *Processes*, Vol. 11, Issue 1, Art. No. 10, 2023.

[7] Aranđelović, M., Milovanović, N., Đorđević, B., Sedmak, S., and Martić, I. Reparation, inspection and damage analysis of steam boiler. *Welding & Material Testing*, Vol. 3, pp. 9-12, 2020.

[8] Luciano, H. G., Costa Vicente, L. F., Daleffe, A., Casagrande, H. C., Crozetta, N. D., and de March, G. Comparative analysis of mig/mag and laser welding processes on low-carbon sheet metal. *Environmental and Social Management Journal/Revista de Gestão Social e Ambiental*, Vol. 19, Issue 12, Art. no. 1, 2025.

[9] Rappe, N. A., Kirda, A. W., Yassin, H., Bartoszuk, M., and Caesarendra, W. Application of YOLOv8 in fusion welding defect detection on carbon steel for potential remote visual inspection. *Advances in Science and Technology Research Journal*, Vol. 19, Issue 12, pp. 436-453, 2025

[10] Cagadas, D. O., Ekpanyapong, M., and Madrid, C. M. A leaf-based rice disease recognition system using convolutional neural network. *Science International (Lahore)*. Vol. 35, Issue 4, pp. 421-424, 2023.

[11] Cagadas, D. O., and Labajan, R. A. Leaf-based cacao diseases classification using image processing. *Science International (Lahore).* Vol. 35, Issue 4, 2024

[12] López, G., Ramírez, P. D., Vega, E., Pizarro, F., Toro, J., and Parra, C. WeldVGG: a VGG-inspired deep learning model for weld defect classification from radiographic images with visual interpretability. *Sensors*, Vol. 25, Issue 19, Art. no. 6183, 2025.

[13] Nikam, D., Nikam, S., Bhosale, T., Harkin, D., Sawant, M., and McGarrigle, C. Deep CNN-based multi-class TIG welding defect classification using HDR images with explainable AI. *Journal of Manufacturing and Materials Processing*, Vol. 10, Issue 6, Art. no. 193, 2026.

[14] Simonyan, K., and Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv*, Art. no. 1409.1556, 2014.