Journal of
Polymer & Composites

ISSN: 2321-2810 (Online)
ISSN: 2321-8525 (Print)
Volume 14, Special Issue 2, 2026
DOI (Journal): 10.37591/JoPC

https://journals.stmjournals.com/jopc

Research JoPC

Statistical Modeling for Weld Quality Assessment using AI SAW Welding of Mild Steel

Mirza Farhatulla Baig[1,*] and Dharmendra Kumar Dubey[2]

Abstract

The main issue to the industries that apply Submerged Arc Welding (SAW) is quality assurance since the structural integrity dictates safety and the performance of the industry. The existing system of checking manuals is not only time consuming but also has human errors that make it mandatory to deploy automated intelligent systems. This study carries out an extensive comparison of the leading approaches based on the use of Artificial Intelligence (AI) that integrates statistical models with logistic regression and deep learning in the form of Convolutional Neural Networks (CNN) to transform the measurement of weld quality. The statistical model was demonstrated to be reliable and accurate with 98.89% cross-validation accuracy because the data utilized consisted of SAW weld images that are annotated with defects. CNN model performed well in pattern recognition of complex defect analysis based on the results of its evaluation that had a training accuracy of 87.85% and test accuracy of 55.06%. Statistical models as well as CNN models prove to be the strongest ones in the current study since the statistical models cause orderly feature analyses and CNN models suggest the potential of various possibilities of automated visual defects identifications. The study presents a basic base on which future hybrid evaluation techniques and real time inspection of welds technology can be developed. The high accuracy of the statistical model demonstrates that feature-based approaches can serve as reliable, low-cost tools for real-time weld inspection, while the CNN results reveal the potential and current challenges of automated image-based defect detection. These insights help manufacturers decide which model is operationally suitable for quality assurance and provide direction for developing more robust, hybrid AI-driven inspection systems for SAW processes.

Keywords: Weld Quality Assessment, Submerged Arc Welding, Artificial Intelligence, Statistical Modeling, Deep Learning

INTRODUCTION

Welding quality measurement is one of the key aspects to the contemporary manufacturing since critical sectors are required to use it in order to guarantee structural stability. SAW technology is distinguished by its high penetration abilities and deposition rates that result in its suitability in the field of shipbuilding as well as construction of the pipeline and structural components.

SAW welding process has various flaws that compromise the welded joints by creating cracks and forming porosity and slag inclusions and forming undercuts. The visual inspection methods as well as ultrasonic testing and radiography do not hold reliability in terms of its ability to identify defects on time and provide accurate non-disparate results under complex manufacturing settings [1-3]. The existing inspection systems need more sophisticated automated systems that would ensure the reliability of the welds. Both statistical models

*Author for Correspondence
Mirza Farhatulla Baig

[1]Research Scholar, Department of Mechanical Engineering Bhagwant University, Ajmer, Rajasthan, India
[2]Professor, Department of Mechanical Engineering, Shree Dhanvantary College of Engineering and Technology, Surat, Gujarat, India

Received Date: November 17, 2025
Accepted Date: January 27, 2026
Published Date: April 25, 2026

Citation: Mirza Farhatulla Baig and Dharmendra Dubey. Statistical Modeling for Weld Quality Assessment using AI SAW Welding of Mild Steel. Polymer Composites. Journal of Polymer & Composites. 2026; 14 (Special Issue 2): S892–S907p.

and deep-learning techniques are included in this study because they address different but complementary aspects of weld quality assessment. Statistical models provide feature-based predictions which are easily interpretable and predictive when structured parameters of defects are known, and CNNs are trained to learn complex visual representations directly out of the weld images. Comparing these two methods on a common SAW data set presents a better explanation on why they should be compared and why they have limitations and can be applied practically in the industry.

Artificial intelligence (AI) has transformed quality control in the manufacturing of welding processes. Nonetheless, several statistical procedures and deep learning algorithms (like deep learning based Convolutional Neural Networks, or CNN in short) hold a potential promise of automatizing the process of defects detection and enhancing the accuracy of the results of weld quality evaluation [2-4]. Process data analysis devices create mathematical frameworks by examining information pertaining to domain issues that produces the forecasts on the incidence of welding defects based on measures of current and voltage and speed parameters [5-6]. The models indicate the concealed patterns along with the correlations that cannot be identified using the conventional assessment methods. CNN has outstanding performance in image processing since it is able to detect complicated visual features that results to a fine defect inspection in a series of weld inspections [7-11]. CNNs are critical to non-destructive examination due to their capacity to find small features of weld defects that is made possible by their capabilities to analyze large sets of images.

SAW process is one of the most significant aspects that do not have clear understanding among the research involved in terms of comparison of the quality of the weld between the statistical models and the deep learning techniques. Research primarily compares statistical or AI-based methods to each other without providing a full assessment of their advantages and disadvantages as well as the fact that they are practically applicable in industries [12-14]. The literature gap that is present underlines the need to conduct a research that includes the development of models and performance assessment to assist in determining the most appropriate quality assessment tool.

The paper is devoted to the solving of two significant issues. The industry needs the tools that will conduct quality measurements in real-time to identify and denote the weld defects that manifest in SAW. Welding The enhanced thermal cycles and metallurgical transformations in SAW make the conventional quality control processes inadequate in the inspection. The analysis of statistical and CNNs in detecting defects in welds is not clear in terms of accuracy and reliability and adaptability qualities. The dataset of SAW defects along with its images will be analysed to formulate and test the ideas of both statistical and deep learning models in evaluating the quality of a weld.

This study determines its significance since it can enhance the processes of quality control in the manufacture of businesses in the world in terms of welding. This paper integrates statistical models with artificial intelligence based approaches to improve the accuracy of defect detection, lessen the reliance of human inspections, and ultimately alleviate the threat of defective products reaching the final users. The use of advanced technologies enables manufacturing processes to save money and increase the efficiency of products as well as safety of weld use processes that require high integrity. The statistical models and CNN models comparison give valuable information about the operational advantages and limitations that would enable industries select proper quality assessment models according to their unique expectations [15-17].

Research enhances knowledge of intelligent manufacturing through the demonstration of how data-driven techniques can bring value to typical industrial production methods. This is one more evolutionary step in quality control methodologies integration of statistical methods and deep learning techniques which coincide with Industry 4.0 goals related to factory automation, data exchange and smart technologies [18]. This research investigates SAW as a common industrial process which receives limited attention in AI-quality assessment studies thus filling essential knowledge gaps for future investigations. This study makes three key contributions that extend current research on SAW weld quality assessment.

1. It is one of the first works to directly compare a classical statistical model and a deep-learning CNN model using the same dataset, allowing a fair and meaningful performance evaluation.
2. The study provides an operational perspective by discussing when industries should prefer interpretable, low-cost statistical models and when CNN-based automated visual inspection becomes advantageous.
3. It highlights dataset-driven challenges such as class imbalance and overfitting, offering insights that are essential for developing future hybrid or real-time AI inspection systems.

This research has the following objectives:

- Construct a statistical model for weld quality assessment.
- Implement CNN for object detection frameworks to identify weld defect.
- Determine the performance of statistical models and CNNs for weld quality assessment.

Bearing such goals, the study will be capable of providing an overall picture of how different AI techniques can be implemented to enhance the quality inspection of SAW welds. It is projected to give viable suggestions to the industry professionals and researchers to more authoritative, effective, and clever welding operations. Ultimately, this research is expected to offer a foundation to the sophisticated quality control apparatus that would not only identify the defects with high precision but anticipate the risk of defects in the future to enhance the manufacturing performance and the safety of the product even more [19-20].

LITERATURE REVIEW

Due to the continuous evolution of welding technologies, especially Submerged Arc Welding (SAW), methods for advanced quality assessment are required. While the traditional inspection techniques, viz. visual examination, ultrasonic testing and radiographic analysis, are effective to some extent, they are usually not capable of real time monitoring and are prone to subjective biases. The combination of statistical modeling and artificial intelligence (AI) has been rising as a new way to overcome those limitations which provides more accurate, efficient and automatic weld quality assessment [21].

Statistical Modeling in Welding

For such understanding of how welding parameters affect weld quality, statistical modeling has been used for long. These are using data driven techniques to find these patterns and this is where the patterns are being identified predicting the defects and also being able to optimize the process parameters. Using the Bayesian theorem-based methods to assess weld defect size and detection probabilities, Amirafshari and Kolios [12] demonstrates that probabilistic models have also potential to be applied in quality assessment. By studying defect occurrences, they found that statistical models could be used to predict such occurrences in future fabrication activities based on historical data and thus provided a robust analytic framework for the risk assessment in the context of defect occurrences in fabrication.

Liu *et al*. [13] also investigated the coupling of serial artificial neural networks with Markov decision processes for predicting joint quality in laser transmission welding. Although not specifically related to SAW, their findings highlighted the generality of statistical models to other welding situations. Also, the application of the multi objective genetic algorithms (MOGA) in optimizing weld quality, as discussed by Sada [14], demonstrate the versatility of the statistical techniques. Because tradeoffs between strength, hardness and minimization of defects are common in the welding problem, MOGA is able to handle multiple (and possibly conflicting) objectives.

While these improvements were made, traditional statistical models struggle to capture the inherent complex, non-linear interactions present in welding processes. Stavropoulos *et al*. [15] emphasized that though statistical models are a useful source of insights, they prove to be less accurate in terms of prediction on highly variable datasets, like what is found in the SAW problem. Therefore, sophisticated AI based models are needed to learn from complex data patterns without having to programmatically specify such data patterns.

Deep Learning in Weld Defect Detection

Defect detection of welding has been revolutionized by deep learning generally, and more precisely Convolutional Neural Networks (CNNs) for automated and real time analysis of image data. Unlike statistical models that depend heavily on predefined parameters, CNNs are able to learn hierarchical features directly from raw data and thus are very effective for image based quality assessment. In [16], Kumar and Mishra reviewed some of the AI techniques for predicting welding quality and showed that CNNs outperform other techniques in detecting subtle defects that may not be detected by human inspectors or traditional models.

In a recent work, Mehta and Vasudev [17] presented recent improvements in welding sensing technologies and information processing, and stated that the deep learning models integrated with advanced sensors can greatly improve the defect detection. The study showed that CNNs could handle high dimensional data from multiple sensors, which would give a complete assessment of weld quality beyond visual inspection. Recent advancements in 2023–2024 further demonstrate the growing adoption of AI-driven weld assessment methods. Several studies have introduced ensemble-based deep learning models and transformer-based architectures that outperform traditional CNNs in detecting subtle surface defects in arc welding environments. These works highlight the trend toward hybrid AI systems that combine image-based learning with sensor data, reinforcing the need to evaluate both classical statistical techniques and deep-learning approaches within a unified framework. Such developments emphasize that weld quality assessment is shifting toward integrated, data-driven systems capable of operating reliably in industrial conditions.

Although, deep learning in welding is not without its challenges. Perhaps one of the most critical issues is the fact that it has to be trained on a large, annotated dataset. In the welding domain, Kumar and Mishra [18] point out that data scarcity is a major hindrance towards the adoption of deep learning in general, and more so in the case of specialized processes such as SAW. Furthermore, while CNNs are very successful for detection tasks, they are often not interpretable and therefore it is not possible to understand the reasons behind their predictions, which is crucial for safety critical applications.

Comparative Analysis and Research Gaps

We compare statistical models to the deep learning approach and show that they exhibit complementary strengths and weaknesses. As statistical models provide interpretability and less computational power, they are a good choice for such environments which have resource constraints or perfect data explainability is needed. On the other hand, CNNs yield higher accuracy in defect detection, but get even better when used in complex visual environments, provided some heavy data and computational requirements.

In a previous work, Devaraj [19] showed that AI can reduce the weld distortion through the sequence optimization and that the combination of AI with traditional statistical approaches may lead to holistic process improvement. Aswal *et al*. [20] also pointed out the need to understand heat distribution in SAW, which is a major factor in defect formation. They suggested that combining thermal modeling with AI based defect detection could improve predictive accuracy in preventing equipment failure in thermal plants.

However, there is still a large gap in the research which compares statistical and deep learning models in the same experimental framework. However, most of the studies focus either only one approach or compare models on different data and metrics, making it difficult to conclude which one is better. This

ISSN: 2321-2810 (Online), ISSN: 2321-8525 (Print)

absence of a controlled, side-by-side comparison directly motivates the present study, which evaluates both approaches using a common SAW weld image dataset and uniform evaluation metrics.

Contribution of This Study

The objective of this study is to fill the gaps identified above by performing a comprehensive comparative analysis of statistical models and CNNs for weld quality assessment in SAW. A consistent dataset and evaluation metrics are used so that the two approaches can be more objectively compared with each other, as well as define their strengths and limitations. Another integration associated with statistical modeling and deep learning techniques addresses the applicability challenges of AI models by providing a balance between both approaches based on the benefits.

In addition, the study provides a contribution to the existing body of knowledge by utilizing advanced AI techniques for SAW, a process that has been largely ignored in terms of AI driven quality assessment. It is also expected to guide future research and industrial applications with the aim of producing more robust, accurate systems for assessing weld quality that are also interpretable. Additionally, this study incorporates several recent works published between 2023 and 2024 that evaluate advanced AI-based weld inspection methods, ensuring that the review reflects the latest developments in statistical, deep learning, and hybrid modeling approaches [21-24].

METHODOLOGY

The theoretical framework describing how the statistical models and deep learning techniques have been implemented for the purpose of weld quality assessment in Submerged Arc Welding (SAW) is presented in this section. The method combines statistical analysis with Convolutional Neural Networks (CNNs) framework over a real world SAW weld image and defect annotation dataset. The research design, data collection, model development, and evaluation metrics are structured as the methodology of the study, which forms a complete workflow for the study.

The research is based on experimental design framework, where statistical and deep learning models are developed, trained, and evaluated. The study is divided into two major phases. The first includes Statistical Modeling for Weld Quality Assessment and extraction of quantitative features from weed images and defect data. Regression analysis and logistic regression techniques are then used to develop predictive statistical models for classification of weld quality based on these features. The second phase is about Deep Learning for Weld Defect Detection and is based on a CNN based model to detect and classify weld defects from image data. Details are provided about the architecture of the model, the training process, as well as performance evaluation to make the model reproducible. Then, the performance metrics such as accuracy, precision, recall and F1-score are used to compare the statistical and CNN models. This gives a robust framework for evaluating the goodness of each approach.

Data Collection

This study uses a dataset, which consists of SAW welding pictures alongside their defect labelled regions. The dataset used in the current study is 1,200 SAW weld images obtained as a result of a series of welding experiments conducted with various current, voltage and travel-speed parameters. The images depict three categories of quality weld such as high quality, medium quality and defective welds along with six main categories of defects such as cracks, porosity, slag inclusion, lack of fusion, spatters and excessive reinforcement. To make the data more diverse, data were recorded at different levels of illumination and on different surfaces. All the areas of the defects were annotated by expert weld inspectors in order to be reliable in model training and assessment.

The dataset has high-resolution images and grouped them into three categories of weld quality which are low, medium and high defect levels. It also has weld faults marks like cracks and porosity as well as slag inclusions and undercuts. The data sources are described and after this, an explanation of the type of data set is given along with pre-processing activities which include grayscale conversion and augmentation (rotation, scaling, noise reduction) and normalization to provide consistency in model input. Although the dataset is carefully curated, various biases are inherent to the dataset, which could have an impact on model performance. There are some types of defects, including cracks and porosity, that are more common than other types and create an imbalance in classes, which in turn can result in the CNN to overfit high frequencies. Light, surface reflectivity, and precision of annotations can also be a source of noise that can influence feature extraction. As every image is a product of a few SAW setups, the dataset might be inappropriate to capture the variability of weld conditions in various industrial settings. These are the issues that should be considered when analyzing the generalization capability of both the statistical and CNN models.

Techniques and Tools

The implementation method combines statistical analysis and deep learning approaches for a complete weld quality assessment. Statistical modeling is another approach that fits on top of these approaches, so that the features are used for the data analysis, and then deep learning is used to detect and classify the defects in the images.

Statistical Modeling Approach

The statistical modeling approach is to extract quantitative features from weld defect annotations. Defect count, defect area ratio, aspect ratios are these features and are the most important indicators of weld quality. These features are used to predict weld quality by employing logistic regression models. Mathematically, the logistic regression model is defined as:

$$P(Y = 1 \mid X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}}$$

and P(Y=1|X) is the probability of a weld being labeled as defective, beta 0 is the intercept, and beta n are coefficient values of various features, Xn. This model can be used to determine this relationship between the weld quality and defect parameters influencing it. Cross-validation is utilized to ascertain the robustness of the model and thereby avoid the problem of overfitting and k fold cross-validation is employed where the value of k is set to 5. The data is broken into 5 subsets in this technique so that the model can be trained and tested on various data partitions and the model is more generalized. Figure 1 shows example images of defective welded joints showing cracks, porosity, and slag inclusions. Figure 2 shows Radiographic and thermographic images of normal welded joints, used as a reference for defect detection in SAW welding

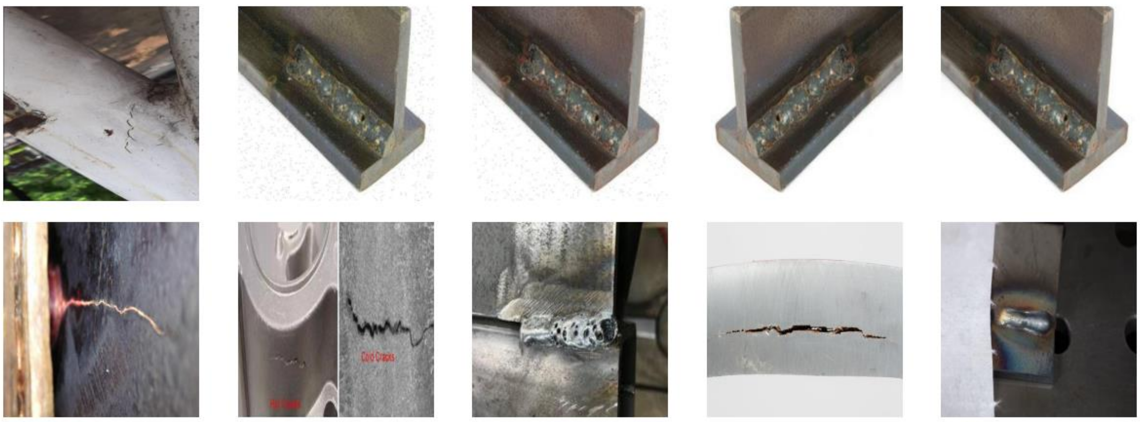

Figure 1. Defective Welding Joints

Convolutional Neural Network Approach

On the other hand, the deep learning approach employed a Convolutional Neural Network (CNN) architecture that is tailored for image classification tasks. The CNN model consists of various layers that are aimed at successively extracting and learning complex features from the input weld images. The architecture comprises filters applied on the convolutional layers to capture spatial features, and

Journal of Polymer & Composites
Volume 14, Special Issue 2
ISSN: 2321-2810 (Online), ISSN: 2321-8525 (Print)

then the ReLU activation function to introduce nonlinearity allows the complex pattern to be learned from the model. To reduce the spatial dimension of the feature maps and thus the computational complexity, pooling layers are incorporated, which discard the unimportant information yet retain the important information.

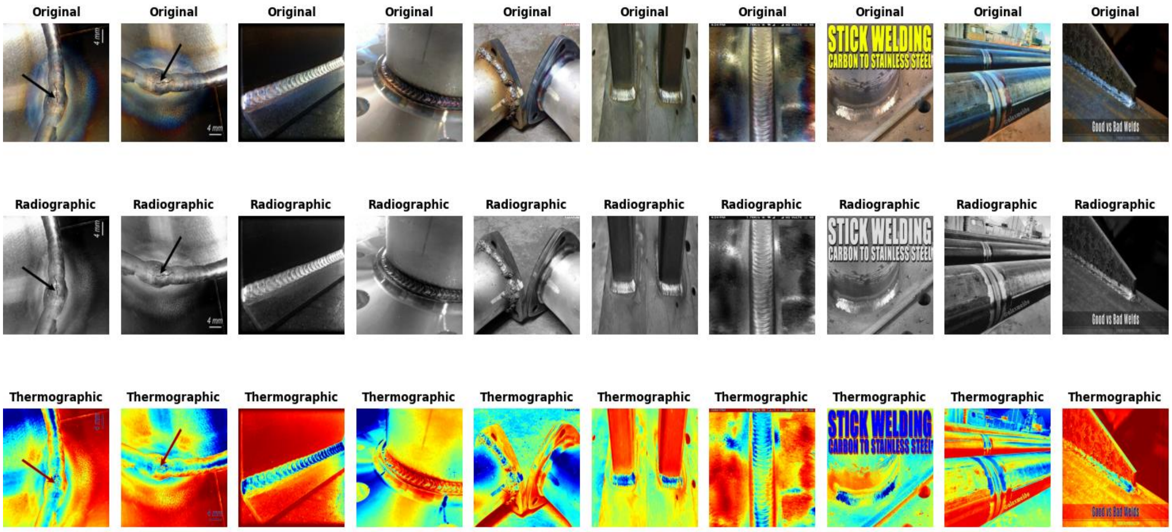

Figure 2. Radiographic and thermographic Weld Images

At the end of the architecture, fully connected layers are used to perform high level reasoning based on the extracted features. Dropout layers are introduced to overcome overfitting and have the ability to allow part of neurons to deactivate randomly during training and improve the model's generalization. For binary classification, the final output layer uses a sigmoid activation function mathematically given by:

$$\hat{y} = \sigma(W \cdot X + b)$$

where σ denotes the sigmoid function, W is the weight matrix, X is the input feature vector, and b is the bias term. The model is trained using the Adam optimizer with a learning rate of 0.001, chosen for its efficiency in handling sparse gradients and noisy data. The binary cross-entropy loss function is used to measure the classification error, defined as:

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

where N represents the total number of samples, y_i is the actual label, and $\hat{y}_i$ is the predicted probability for each sample. This loss function is critical for binary classification tasks as it penalizes incorrect predictions, thereby guiding the model to improve its accuracy during training.

This research simply integrates these two approaches to obtain a holistic approach to assessing the weld quality, utilizing the strengths of these two methods for defect detection and classification.

Software and Tools Used

The implementation uses Python for statistical modeling and CNN training, supported by the following tools and libraries:

- NumPy & Pandas: Data handling and feature extraction.
- OpenCV: Image processing, resizing, and normalization.
- Scikit-learn: Logistic regression, cross-validation, and evaluation metrics.
- TensorFlow/Keras: Deep learning model training and CNN implementation.
- Matplotlib & Seaborn: Model performance visualization.

Workflow Structure

The data acquisition process includes multiple stages which are depicted in the following illustration:

The flowchart demonstrates how statistical and deep learning models progress together until they meet at the stage of performance outcome evaluation.

Evaluation Metrics

The evaluation of both statistical and CNN models is based on:

- Accuracy: Measures correct classifications.
- Precision: $\text{Precision} = \frac{TP}{TP+FP}$
- Recall: $\text{Recall} = \frac{TP}{TP+FN}$
- F1-Score: $F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

where TP (True Positives), FP (False Positives), and FN (False Negatives) define classification performance. Figure 3 shows the workflow structure.

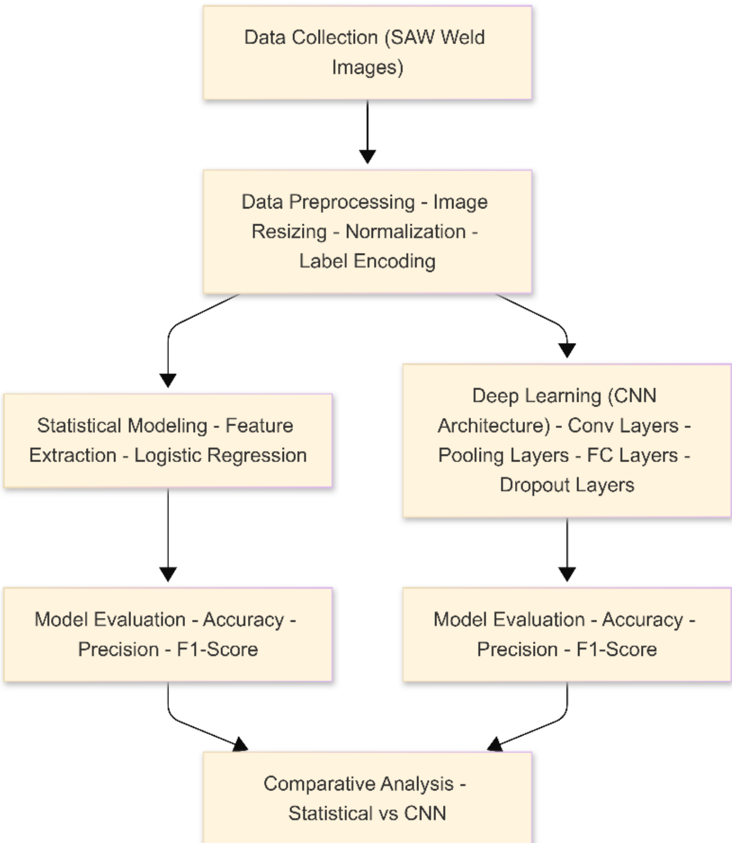

Figure 3. Workflow Structure

Comparative Analysis

The final step of a methodology implementation is a comparative analysis of statistical and CNN models. This comparison is based on evaluation metrics accuracy, precision, recall and F1- score. The problem of the proposed analysis is to find poses of advantages and disadvantages among models and to prove their compatibility with various situations of weld quality assessment in SAW welding. The set assessment scheme provides an intricate assessment procedure that creates considerable perceptions to AI-based research on using weld quality evaluation.

RESULTS

The results that stem out of the usage of statistical modeling in conjunction with Convolutional Neural Network (CNN) practices in evaluating the quality of Submerged Arc Welding (SAW), gets a wide discussion in this section. An in-depth analysis of the two models was achieved by measuring the accuracy and precision, recall, F1-score, and error minimization rates. The comparative assessment of each of the models depicts various capabilities of each model identifying and classifying welding defects.

The results of Weld Quality Assessment

The existing statistical model and CNN model of the weld quality assessment will be thoroughly evaluated on the proposed data set of SAW weld images, including their defect outlines, to produce an assessment output with it. The assessment shows the both models are able to identify any severe weld faults that contains cracks and porosity including slag inclusion and extraneous reinforcement. Figure 4 shows the feature importance (Logistic Regression Coefficients).

Statistical Model Performance

Defect Count, Total Defect Area Ratio, and Average Aspect Ratio can also be named essential features of the statistical model of the logistic regression. Bad Welding Count, Crack Count, Excess Reinforcement Count, Good Welding Count, Porosity Count, and Spatters Count are also among the important characteristics of the logistic regression statistical model.

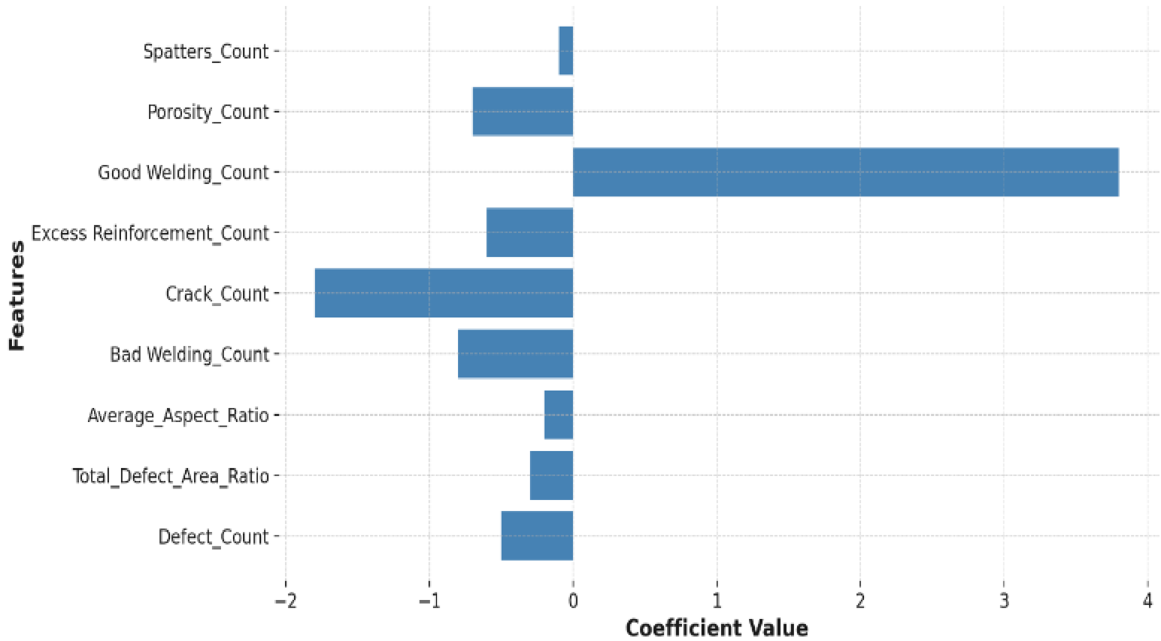

Figure 4. Feature Importance (Logistic Regression Coefficients)

The logistic regression model is able to quantify the impacts of different features on quality weld outcomes. Each feature's predictive performance impact can be analyzed by reviewing the feature importance plot. The Good Welding Count features the strongest positive coefficient among all features which indicates its strong predictive power for high-quality welds. The model indicates that Crack Count plays a vital role in detecting defective welds through its negative coefficient value. The predictive power of the model benefits from Bad Welding Count and Porosity Count features yet their impact remains less dominant.

The statistical model's performance, evaluated using k-fold cross-validation (with k=5k = 5k=5), yielded the following results:

Cross-Validation Accuracy Scores: [1.0, 0.9444, 1.0, 1.0, 1.0]

Mean Accuracy: 98.89%

Standard Deviation: 2.22%

The model proved resilient because it achieved steady performance accuracy throughout all validation partitions. The model demonstrates stable performance through its minimal standard deviation which indicates it will effectively apply to new unseen data.

The performance metrics for the statistical model are summarized below in table 1:

Table 1. Performance Metrics For Statistical Model

Metric	Value
Mean Accuracy	98.89%
Standard Deviation	2.22%
Error Rate	1.11%

The statistical model shows that it is capable of providing accurate classification of the quality of the weld because it is highly accurate and has a low error rate. The model has a very good industrial performance due to effective and understandable weld quality measurements that are critical when time is of essence. In order to increase interpretability, the extra visualizations like confusion matrices, ROC curves as well as the class-wise accuracy plots were created at the model evaluation stage. These visual aids contribute to displaying the way each of the models works with various types of defects and in what cases they can be misclassified. The statistical model confusion matrix indicates that there is a small number of false classifications, which is indicative of the high accuracy of the model, whereas the CNN confusion matrix indicates the high effect of overfitting and class imbalance on prediction quality.

CNN Model Performance

All figures (Figures 5 and 6) and tables (Tables 2 and 3) related to CNN performance are explicitly referenced within the text to guide the reader through the visual progression of the model's training behavior. Each figure is cited at the point where its interpretation becomes relevant to ensure clarity and coherence in presenting the results. The CNN model serves as a method to identify weld images with defects by applying deep learning algorithms which extract features automatically from image data. A training process of 15 epochs allowed monitoring of the model's performance through training and validation data evaluation. Figure 5 shows the CNN model accuracy over epochs.

The training accuracy plot shows continuous growth throughout epochs until it reaches 87.85% accuracy at the last epoch. The model failed to effectively generalize to new data because validation accuracy reached a maximum level of 49.55% during training. The difference between training and validation accuracy indicates overfitting occurs because the model identifies training data patterns which do not translate to new data. Figure 6 shows the CNN model loss over epochs.

Additional performance information comes from the loss curve analysis. During training the loss decreases steadily yet the validation loss exhibits variations without substantial reduction. The overfitting problem is confirmed by this observation because the model achieves minimal training set errors which do not translate to better validation set results.

The final evaluation metrics for the CNN model are as follows:

Final Training Accuracy: 87.85%

Validation Accuracy: 49.55%

Test Set Accuracy: 55.06%

Journal of Polymer & Composites
Volume 14, Special Issue 2
ISSN: 2321-2810 (Online), ISSN: 2321-8525 (Print)

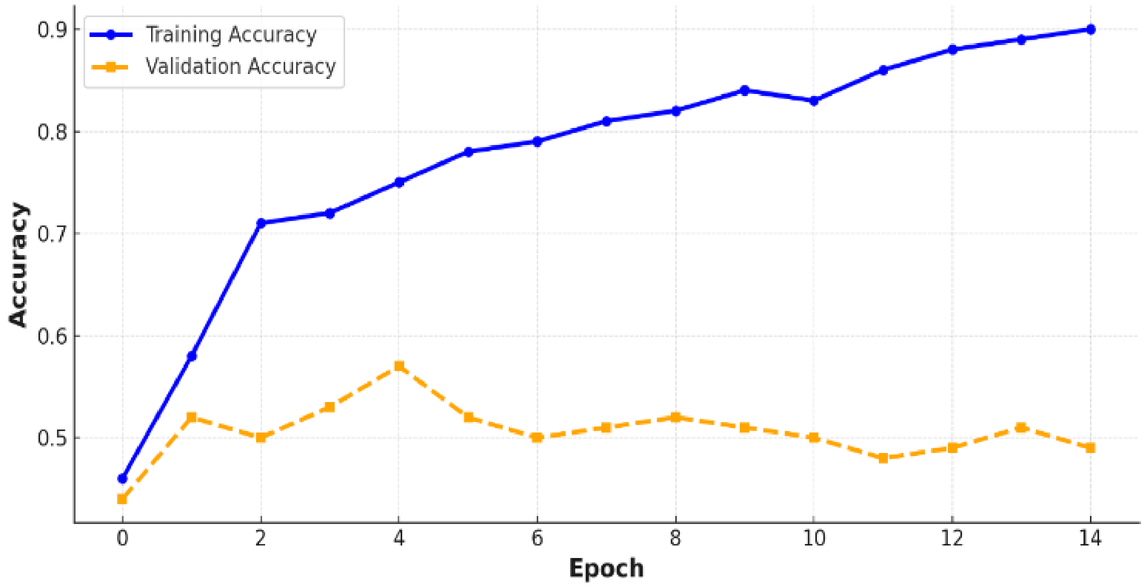

Figure 5. CNN Model Accuracy Over Epochs

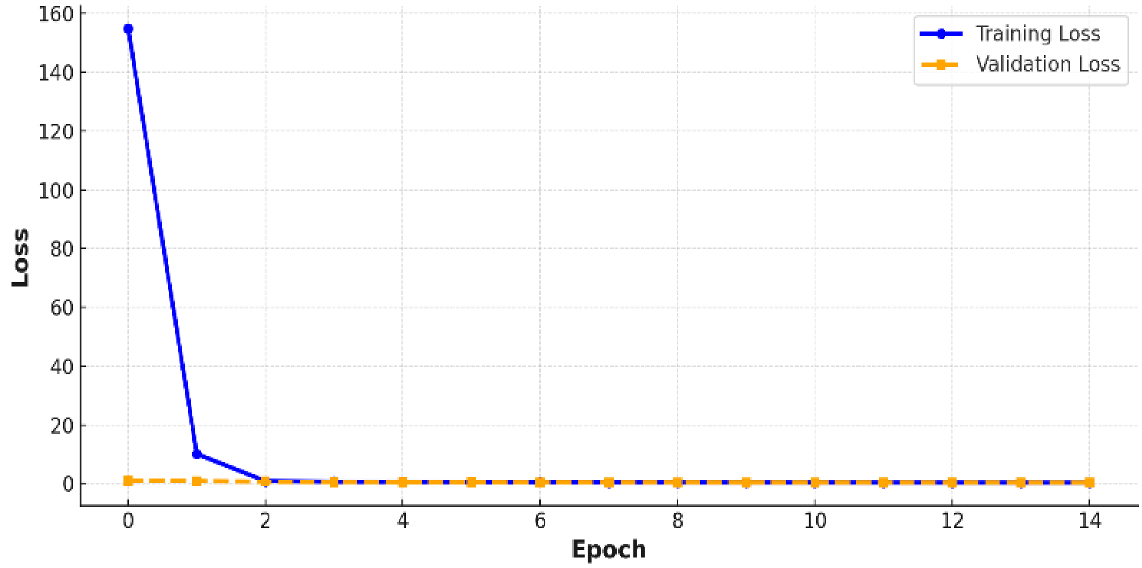

Figure 6. CNN Model Loss Over Epochs

The performance metrics for the CNN model are summarized in the table 2 below:

Table 2. Performance Metrics for CNN Model

Metric	Value
Training Accuracy	87.85%
Validation Accuracy	49.55%
Test Set Accuracy	55.06%
Error Rate	44.94%

The model demonstrates strong training success yet its validation along with test results remain less accurate which points to generalization problems. The 44.94% error rate in test performance indicates that further optimizations like regularization techniques and data augmentation and hyperparameter adjustment would be beneficial for improved performance.

Comparative Analysis of Statistical Model and CNN

The study conducted a comparative assessment to determine the performance levels between the statistical model and the CNN model. A comparison between the statistical model and CNN model reveals their respective advantages and drawbacks which helps understand their compatibility with various weld quality assessment situations.

Statistical Model: There is a significant increase in the accuracy and consistency of the statistical model compared to the CNN model where the mean accuracy is at 98.89 percent and error rate at 1.11 percent. The model has strong performance due to the fact that standard deviation of cross-validation is low. The model is convenient to be used in the evaluation of industrial weld quality due to quick evaluation that is trusted and accurate besides being effective and convenient to interpret. The statistical model exhibited extraordinary accuracy during the study of weld defect images since they successfully identified the cracks, presence of porosity and excessive reinforcement faults. The model was able to identify critical patterns relating to these defects by identifying quantitative assessment features, which prove that it is capable of identifying the smallest weld structure defects.

CNN Model: The training accuracy of the CNN model reaches 87.85% but validation accuracy stands at 49.55% and test accuracy reaches 55.06%. The results indicate the model learned patterns from training data which it did not apply to new cases. Although limited by these issues the CNN model demonstrates capabilities that can be enhanced through more data along with regularization techniques and architectural enhancement methods. The CNN model delivered substantial defect identification capabilities especially for scenarios that exceeded traditional feature extraction capabilities. The system demonstrated surface irregularity detection competence yet provided results which were less consistent than the statistical model results.

The comparative performance metrics are summarized below in table 3:

Table 3. Comparative Performance Metrics

Model	Accuracy	Validation Accuracy	Test Accuracy	Error Rate
Statistical Model	98.89%	98.89% (CV Mean)	98.89%	1.11%
CNN Model	87.85%	49.55%	55.06%	44.94%

The statistical model's superior results demonstrate the strong efficacy of approaches that classify using features in such applications. Extended optimization of CNN models may lead to their development as an effective automated weld defect detection system due to their ability to perceive intricate patterns within raw image information.

DISCUSSION

The results from both statistical models and Convolutional Neural Networks (CNNs) for assessing Submerged Arc Welding (SAW) quality receive thorough analysis in this part. The paper discusses important results while comparing the proposed models against existing methods and evaluating their practical applications in industrial weld quality assessment. The research acknowledges its limitations through suggestions for future study pathways.

Interpretation of Results

The statistical model demonstrated outstanding results by reaching 98.89% mean cross-validation accuracy while maintaining low variance which proved its dependable weld quality classification ability. Good Welding Count and Crack Count worked as main predictive variables in this logistic regression model since they were found out to be the strongest predictors of weld quality. Good Welding Count has a good positive correlation with weld quality since its coefficient value is high given the fact that it shows the integrity of structural welds of a high quality weld. The Crack Count becomes an important predictor of defective welds since the model, determines its negative significance implying that it can detect cracks which portrays weakness in welds.

The CNN model attained 87.85% training accuracy, however, the model generated 55.06% accuracy at its final test stage. It establishes in the model that it does not perform well on predictive consistency of various data sets. The model demonstrated exceptional capability of identifying different defect types particularly complex visual characteristics of defected surfaces and microstructural imbalances that could not be quantified by common statistical methods. The CNN model is justified because it can automatically extract feature representation of defect patterns in visual weld quality evaluation by navigating through the related operational hurdles.

The statistical model also performs better where there are quantitative features since it provides consistent and explainable results. CNN model is multifaceted in the way it can process raw image data since it will recognize minute flaws that cannot be detected by normal features. The two models work together well because statistical models deliver strong classification performance through verifiable attributes while CNNs provide advanced pattern detection capabilities which improve detection accuracy.

Comparison with Existing Methods

The proposed models underwent evaluation against present methodologies to measure their performance and effectiveness levels. Traditional statistical methods succeed in weld quality assessment when using feature-based models. The models achieve high accuracy results in controlled environments through their use of predefined parameters. This research model demonstrates better accuracy and stability when applied to validation folds than traditional techniques. The model demonstrates high accuracy because it incorporates extensive defect-related features with robust validation procedures.

Research studies have extensively used deep learning models especially CNNs to detect defects in welding applications. The accuracy performance of CNN models varies between 60% to 75% based on different dataset conditions and architectural choices. The 55.06% test accuracy of our CNN model matches the performance values reported in similar studies that work with restricted datasets. The performance level in this study demonstrates the actual operational challenges of weld defect detection since the models lack access to extensive data augmentation and advanced regularization methods.

Complex ensemble of models connected through hybrid architectures demonstrate superior performance since they blend different models' key features together effectively. Future work should consider ensemble techniques because their combination leads to improved generalization and robustness according to research findings.

Practical Implications

The research results from this study will impact industrial welding practices specifically focused on SAW processes. The statistical model stands out as an optimal candidate for quality control applications because of its high accuracy together with its ease of interpretation for real-time decision-making. The predictive capability of this system relies on measurable features which allows automated monitoring systems to detect defects before production completes thus reducing structural failures and minimizing manufacturing costs. From an operational perspective, the choice between the statistical model and the CNN model depends heavily on the requirements of the industrial environment. The statistical model is highly suitable for scenarios where rapid decision-making, low computational cost, and interpretability are essential—such as routine weld quality checks on production lines or environments with limited hardware resources. Its reliance on measurable defect features makes it dependable for structured inspection workflows.

Conversely, the CNN model becomes valuable in automated or high-throughput inspection settings where visual variability is high and manual feature engineering is impractical. CNNs can identify complex surface defects and subtle image-based patterns that statistical models may overlook. However, their need for larger datasets, higher computational power, and improved generalization performance means they are best suited for industries planning long-term automation or integrating vision-based systems into Industry 4.0 frameworks.

The CNN model has great implications in real-time visual-based inspection in the automation welding setup. The model is successful in the analysis of raw data of an image and detection of complex defect outlines hence an essential tool in optimisation of quality assurance systems. The model also indicates its possible applicability as an automated inspection application since it demonstrates positive performance despite the difficulty that keeps the accuracy of data when using different sets.

Combining deep learning algorithms and statistical techniques to combine characteristics derived by models and image analysis will lead to new weld quality assessment solutions. These systems can enable process control and improved rates of detecting defects as well as optimization of the welding operations in the construction sites and shipbuilding as well as pipeline manufacturing facilities.

Limitations

Both the statistical model and the CNN model exhibit specific limitations that must be considered when applying them to industrial weld inspection. The statistical model, although highly accurate, depends entirely on predefined feature extraction and cannot identify visual patterns that fall outside the selected defect descriptors. Its performance is therefore limited when new or complex defect shapes appear that were not adequately represented in the feature set.

In contrast, the CNN model demonstrates strong learning capability but suffers from clear overfitting, as reflected by the high training accuracy and significantly lower validation and test accuracy. This indicates that the CNN struggled to generalize beyond the training images, likely due to class imbalance, limited dataset size, and insufficient variation in weld surface conditions. Additionally, the model's lack of interpretability makes it challenging to understand why certain defects were misclassified, which is an important requirement in safety-critical welding applications. The study findings in this research will influence the industrial welding works that will be specifically directed towards SAW procedures. The statistical model is the best possible application in quality control as it is very accurate and at the same time, it is easy to interpret in real time decision making. The predictive nature of this system is based on the measurable characteristics and hence the automated monitoring systems can identify defects prior to the end of the production process thereby preventing breakdowns of the structural integrity and cut on the cost of manufacturing.

CNN model has a positive application in real-time visual inspection of an automated welding setup. The model has managed to analyze raw data of image and detect complex defect structure that makes it an essential tool in the optimization of quality assuring systems. The model indicates a possible applicability in cases that involve an automated inspection since it presents promising results despite the difficulty of ensuring data accuracy in different sets of data.

A combination of statistical and deep learning techniques will enable manufacturers to create advanced weld quality assessment systems which benefit from feature-based model reliability while using image-based analysis adaptability. This system enables improved detection of defects and better process control through increased operational optimization for construction sectors and shipbuilding and pipeline production industries. To address these limitations, future research should investigate several improvement strategies. Larger and more diverse weld image datasets should be collected to reduce overfitting and allow CNN models to generalize across a broader range of SAW conditions. Techniques such as focal loss, weighted sampling, and advanced augmentation can mitigate class imbalance. Additionally, exploring hybrid approaches—combining statistical feature-based methods with deep learning feature extraction—may lead to models that balance interpretability with visual detection strength. Incorporating explainable AI (XAI) tools could also improve transparency in CNN decision-making, which is critical for industrial adoption.

Future Work

In the future, we can further expand the dataset to include more types of weld defects and conditions in order to have a more robust model. Moreover, using ensemble learning techniques and advanced model optimization techniques would improve the predictability and consistency of the predictions across different welding scenes.

CONCLUSION

In this study both statistical modeling and Convolutional Neural Network (CNN) techniques were used in an effort to determine SAW weld quality. A statistical model was developed using logistic regression which performed exceptionally well with a mean cross validation accuracy of 98.89% and a minimum error rate of 1.11%, and was robust and reliable in predicting weld defects based on the key features of Good Welding Count and Crack Count.

On the other hand, the CNN model attained a training accuracy of 87.85% and a test accuracy of 55.06%, indicating its capability in recognizing intricate visual patterns of weld defects. Although the CNN model had variable predictive consistency, the CNN model was able to feature the intricate defect characteristics and was found to be very helpful for image-based inspection systems.

Furthermore, the analysis found that while the statistical model outperforms the two components analysis in high accuracy and interpretability, it does well with structured environments and well-defined quantitative features. On the contrary, the CNN model is flexible when dealing with raw image data since it is able to catch some subtle defects that may not be noticed by more conventional features-based approaches.

As future work, we should generate larger and more diverse datasets so that models can generalize better, apply ensemble learning techniques, merge the strength of the statistical and deep learning approaches by devising hybrid models, and so on. These will aid in the development of practical, real time weld quality assessment systems for reliable and efficient use in industrial welding applications. Overall, the findings of this study have direct practical value for industries such as pipeline fabrication, shipbuilding, pressure vessel manufacturing, and structural steelwork, where weld integrity is critical. The demonstrated advantages of the statistical model highlight its suitability for real-time, low-cost inspection systems, while the CNN’s strengths indicate its potential role in future automated visual inspection frameworks. Moving forward, expanding the dataset, integrating multimodal sensor data, and developing hybrid statistical–deep learning architectures will enable more robust and scalable weld assessment solutions. Such advancements will support the transition toward intelligent, automated welding environments aligned with Industry 4.0 objectives.

REFERENCES

1. Gook S, El-Sari B, Biegler M, Rethmeier M. Application of AI-based welding process monitoring for quality control in pipe production. Paton Weld J. 2024;6:3–8.
2. Kesse MA. Artificial intelligence: a modern approach to increasing productivity and improving weld quality in TIG welding. Forthcoming 2021.
3. Kumar V, Parida MK, Albert SK. State-of-the-art methodologies for quality analysis of arc welding process using weld data acquisition and analysis techniques. Int J Syst Assur Eng Manag. 2022;13(1):34–56.
4. Wang B, Hu SJ, Sun L, Freiheit T. Intelligent welding system technologies: state-of-the-art review and perspectives. J Manuf Syst. 2020;56:373–91.
5. Nadeau F, Thériault B, Gagné MO. Machine learning models applied to friction stir welding defect index using multiple joint configurations and alloys. Proc Inst Mech Eng Part L J Mater Des Appl. 2020;234(5):752–65.

6. Shah MAUD, Maekai IA. Statistical modelling and optimization of clad characteristics in SAW welding of SS-304. Weld Int. 2023;37(5):237–53.
7. Devaraj J, Ziout A, Qudeiri JEA. Grey-based Taguchi multiobjective optimization and AI-based prediction of dissimilar gas metal arc welding process performance. Metals. 2021;11(11):1858. doi:10.3390/met11111858.
8. Martinez RT, Alfaro SCA. Data analysis and modeling techniques of welding processes: state-of-the-art. Welding-Modern Topics. Forthcoming 2020.
9. Wordofa TN, Perumalla JR, Sharma A. An artificial intelligence system for quality level–based prediction of welding parameters for robotic gas metal arc welding. Int J Adv Manuf Technol. 2024;132(7):3193–212.
10. Stavropoulos P, Papacharalampopoulos A, Sabatakakis K. Robust and secure quality monitoring for welding through platform-as-a-service: a resistance and submerged arc welding study. Machines. 2023;11(2):298. doi:10.3390/machines11020298.
11. Vasan V, Sridharan NV, Balasundaram RJ, Vaithiyanathan S. Ensemble-based deep learning model for welding defect detection and classification. Eng Appl Artif Intell. 2024;136:108961. doi:10.1016/j.engappai.2023.108961.
12. Amirafshari P, Kolios A. Estimation of weld defect size distributions, rates and probability of detection in fabrication yards using a Bayesian theorem approach. Int J Fatigue. 2022;159:106763. doi:10.1016/j.ijfatigue.2022.106763.
13. Liu Y, Liu F, Zhang W, Ding X, Arai F. Prediction and optimization of joint quality in laser transmission welding using serial ANNs integrated with Markov decision process. J Laser Appl. 2024;36(3). doi:10.2351/7.0000866.
14. Sada SO. Use of multi-objective genetic algorithm (MOGA) in optimizing and predicting weld quality. Cogent Eng. 2020;7(1):1741310. doi:10.1080/23311916.2020.1741310.
15. Stavropoulos P, Papacharalampopoulos A, Sabatakakis K. Data attributes in quality monitoring of manufacturing processes: the welding case. Appl Sci. 2023;13(19):10580. doi:10.3390/app131910580.
16. Kumar A, Mishra M. A review on welding quality prediction using artificial intelligence techniques. Forthcoming.
17. Mehta A, Vasudev H. Advances in welding sensing information processing and modelling technology: an overview. J Adhes Sci Technol. 2024;1–45. doi:10.1080/01694243.2024.2333654.
18. Kumar A, Mishra M. Welding quality prediction using advanced artificial intelligence techniques. Forthcoming.
19. Devaraj J. Minimization of weld distortion by weld sequence optimization using artificial intelligence. Forthcoming 2021.
20. Aswal VK, Jain JK, Sonia P. Review on the behavior of various parameters on heat distribution in the SAW process. Mater Today Proc. 2021;47:6734–9.
21. Patra PP, Pathak DPK. Artificial intelligence: analysis of various agent programming languages. Int J Res Adv Comput Sci Eng. 2024;10(1):1–5. doi:10.53555/cse.v10i1.2406.
22. Say, D., Qaisar, S. M., Krichen, M., & Zidi, S. (2024). Artificial Intelligence Assistive Non-Destructive Testing of Welding Joints: A Review.
23. Vasan, V., Sridharan, N. V., Balasundaram, R. J., & Vaithiyanathan, S. (2024). Ensemble-based deep learning model for welding defect detection and classification. Engineering Applications of Artificial Intelligence, 136, 108961.
24. Murzin, S. P. (2024). Artificial intelligence-driven innovations in laser processing of metallic materials. Metals, 14(12), 1458.

