Middle East Journal of Science

https://dergipark.org.tr/mejs

MEJS

e-ISSN:2618-6136

Review

APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES FOR DEFECT PREVENTION AND QUALITY CONTROL IN ARC WELDING PROCESSES: A COMPREHENSIVE REVIEW

Turgay Tugay BILGIN[1*] Musa Selman KUNDURACI [2] Ahmet METİN[3] Merve PAYAR[4] Erdal NAYIR[5]

[1] Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Türkiye
[2]Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Bursa, Turkiye
[3]Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Bursa, Turkiye
[4]Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Bursa, Turkiye
[5]Bursa Technical University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Bursa, Turkiye
*Corresponding author: turgay.bilgin@btu.edu.tr

Abstract: *This study presents a comprehensive review of research applying artificial intelligence (AI) techniques to prevent defects in arc welding processes. Arc welding is essential across various industries, but numerous issues can arise, impacting weld quality and production efficiency. The review systematically analyzes relevant studies published since 2018, focusing on three key aspects: datasets used, methodologies and approaches adopted, and performance metrics reported. The findings reveal significant adoption of both machine learning and deep learning techniques, with the choice depending on factors like input data nature, welding process dynamics, and computational requirements. Deep learning models, particularly convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have demonstrated superior performance in image-based defect detection and time-series analysis for quality prediction. However, traditional machine learning algorithms have also been utilized, often coupled with dimensionality reduction or feature selection techniques. The review highlights the diverse range of performance metrics employed, such as accuracy, precision, recall, F1-score, mean squared error (MSE), and root mean squared error (RMSE). Metric selection depends on the specific task (classification or regression) and the desired trade-off between different performance aspects. While many studies reported promising results with accuracy rates frequently exceeding 90%, challenges remain in real-world industrial settings due to factors like noise, occlusions, and rapidly changing welding conditions. This review serves as a comprehensive guide for researchers and practitioners in AI-assisted defect prevention and quality control for arc welding processes, highlighting current trends, methodologies, and future research directions.*

Keywords: *arc welding, Deep Learning, Machine Learning, welding defect prevention*

Received: June 8, 2024 *Accepted: November 3, 2024*

1. Introduction

Welding techniques are among the most fundamental requirements of the industry. They are employed in a wide range of applications, from automotive to construction, and from aircraft and aviation technologies to the defense industry.

Arc welding encompasses various techniques that use an electric arc to join metals, each suited to different materials and applications.

1.1. Primary Types of Arc Welding

1.1.1 MIG/MAG welding

MIG/MAG welding techniques are gas metal arc welding methods. In gas metal arc welding, the necessary heat for welding is generated by an arc created between a continuously fed melting wire electrode and the workpiece, and by the resistance heating of the welding current passing through the electrode. Gas metal arc welding (GMAW), also called metal inert gas (MIG) welding, is an arc welding process in which the heat for melting the metal is generated by an electric arc between a consumable electrode and the metal. When welding is done under an active gas shield, it is called MAG welding [1].

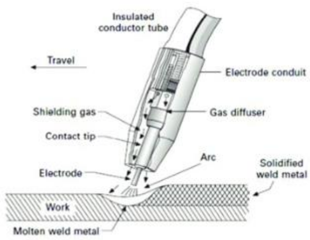

Figure 1. MIG/MAG Welding Technique [2].

For MIG welding, commonly used gases are pure argon and helium. In MAG welding, pure CO_2, argon, and O_2 are used. Schematic representation of MIG/MAG welding technique. General gases used for MIG welding include pure argon and helium, whereas pure CO_2, argon, and O_2 are used for MAG welding. Figure 1 is a schematic representation of the MIG/MAG welding technique. Gas metal arc welding, such as MIG/MAG, provides high welding speed and uses less filler metal. It is suitable for robotic welding applications [3].

There are different metal transfer modes in MIG/MAG welding:

- Short Circuit: Short circuit is the coolest form of MIG welding, utilizing low voltage. In this transfer method, the consumable electrode wire arcs and momentarily contacts the base material, causing a short circuit. This produces a small weld puddle that solidifies rapidly, often described as "fast freezing," as it drips into the joint to fuse the materials. It's ideal for welding thinner materials, though there is a risk of "cold lapping" when used on thicker metals. Additionally, this method tends to generate more spatter [4].
- Globular: Globular transfer method is very similar to the short circuit transfer method, which the consumable electrode wire arcs and touches the base material and shorts. The difference comes in how long the consumable electrode melts. In Globular method, the wire is heated longer and creates a large volume of weld metal that drips into the weld joint. It uses a high heat input and also risks less fusion because of large amounts of spatter disrupting the weld puddle. You are limited to flat and horizontal fillet welds with this method [4].
- Spray Arc: In the Spray Arc transfer method, small droplets of molten metal from the consumable electrode are sprayed into the weld joint. This is a pure CV (constant-voltage) process that sends a constant stream of weld metal across the arc to the base material. This method uses a high heat input and you risk burn-through on thinner materials and only allows for limited to flat and horizontal weld positions [4].
- Pulsed MIG: This Pulsed MIG transfer method is a modified form of the Spray Arc method, taking the best parts of all the transfer methods and minimizing their disadvantages. Pulse MIG

welding does require a special power source, which pulses the voltage many times per second. This allows one droplet of molten metal to form at the end of the consumable wire and the current, and then pushes the droplet across the arc into the weld puddle. A droplet is formed every pulse. Since the voltage drops on every pulse, this creates a longer cooling off period and may reduce the HAZ from the weld. Pulse MIG transfer minimizes spatter or the risk of cold lapping, and weld positioning is not as limited as the Globular and Spray methods [4].

During the welding process using MIG/MAG welding machines, adherence to certain parameters is imperative. These parameters constitute crucial factors that determine the quality of the weld being executed. As the thickness of the material subjected to the welding process increases, it becomes necessary to augment the intensity of the electrical current being applied.[5]. Analogously, the current value must be calibrated in accordance with the diameter of the wire being utilized. Voltage affects the arc length and the fluidity and width of the weld bead. Wire feed speed controls the amount of filler material supplied, while penetration is influenced by the current.

Defects can occur at excessive speeds, and if the wire feed speed is too low, it can result in a narrow, low-penetration, and sometimes convex weld area. Welding speed: A high welding speed results in a narrow weld area and insufficient penetration. Conversely, a slow welding speed causes an excessively wide weld area.

Additionally, the presence of shielding gas, gas flow rate, torch angle and ambient temperature also affect the quality of the weld. For optimal MIG/MAG welding, it should be performed in a closed environment to maintain the quality, as the shielding gas direction can change quickly in open air, reducing weld quality. This welding type is also sensitive to moisture and dust, so the environment must be considered carefully[6].

Aside from these factors, there are problems that can occur during welding[7]. Porosity occurs when gases present in the atmosphere are absorbed while the weld pool is still in liquid form. These gases become trapped as the metal solidifies, creating a weld filled with holes. The main causes are insufficient shielding gas or using the wrong type of shielding gas.

Undercut occurs when the weld metal expands too much and the base metal collapses towards the edges. Generally, this happens when the torch angle is too excessive in one direction, there is insufficient filler metal, or the current is too high. Burn-through, also known as a burn hole, is when the filler metal penetrates through the other side of the metal, creating a hole. The main causes are generally slow welding speed, high current, and heat being focused on one spot.

When the metals being welded do not sufficiently fuse with the weld, this is known as lack of fusion. Sometimes gaps around the edges of the weld can be observed as indicators.. The primary causes are welding too quickly, insufficient heat, an incorrect torch angle, or using electrodes or filler metal that are too small. Spatter refers to small metal droplets that surround the weld after welding. These metal particles are expelled from the torch or electrode during welding and accumulate around the weld once they solidify. While it does not damage the weld itself, it increases the time required for cleaning. Causes include low voltage, high wire feed speed, high current (amperage), and an excessively steep torch angle.

At the end of welding, craters are empty spaces. They may lead to cracks occasionally. This happens when there is not enough filler metal built up at the weld bead's end. These problems are closely related to MIG/MAG welding parameters. Therefore, before starting the welding process, attention should be paid to the welding parameters, and the correct parameters should be used.

1.1.2 TIG (Tungsten Inert Gas Welding) welding

TIG Welding (Tungsten Inert Gas Welding), also known as Gas Tungsten Arc Welding (GTAW), is a highly precise welding process that uses a non-consumable tungsten electrode to create an electric arc between the electrode and the workpiece. An inert shielding gas, typically argon or helium, protects the weld from atmospheric contamination, ensuring a clean and high-quality weld. This process can be used with or without filler metal, depending on the requirements of the job[8].

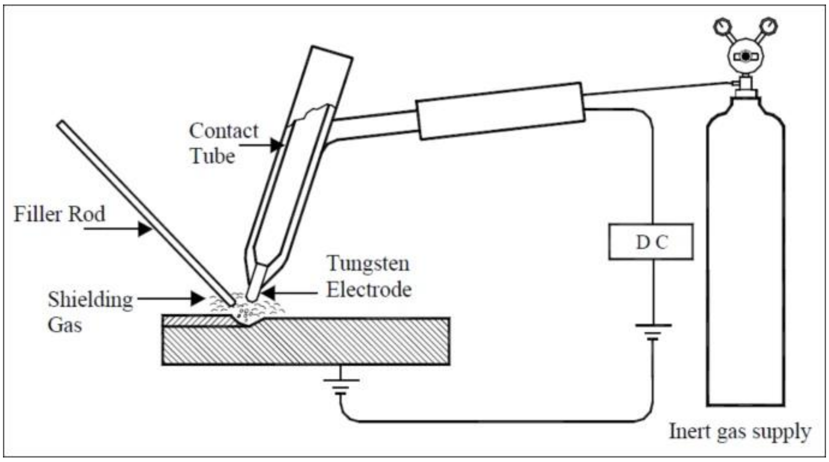

Figure 2. Schematic diagram of a TIG welding process [9].

One of the key advantages of TIG welding is the level of control it offers. The welder can manually adjust the heat input and filler material, making it ideal for welding thin materials and delicate applications. The result is a weld that is both strong and visually smooth, with minimal spatter or defects. TIG welding is versatile and can be applied to a wide range of metals, including stainless steel, aluminum, and titanium. This precision makes it widely used in industries like aerospace, automotive, and art fabrication[10].

However, TIG welding also has its challenges. It is a slower process compared to methods like MIG welding, and it requires a high level of skill and experience to master. Additionally, the equipment and labor costs are generally higher due to the manual nature of the process.

In summary, TIG welding is a method that excels in producing high-quality, clean welds with superior control over the welding process, making it an essential tool in industries requiring precision and reliability. Despite its slower pace and higher cost, the benefits it offers in terms of weld integrity and versatility make it a preferred choice for critical applications.

1.1.3 MMA (Manual Metal Arc Welding)

Manual Metal Arc Welding MMA welding, also known as Manual Metal Arc Welding or Shielded Metal Arc Welding (SMAW), is a manual welding process that uses a consumable electrode coated with flux to create the weld. In this method, an electric arc is generated between the electrode and the workpiece, producing the heat needed to melt both the electrode and the base metal. As the electrode melts, it forms a weld pool and deposits filler material, while the flux coating burns off to create a protective gas shield around the arc. This gas shield, along with the slag that forms on top of the weld, prevents the molten metal from being contaminated by the atmosphere [11].

One of the key advantages of MMA welding is its simplicity and portability. Since it doesn’t require external shielding gas or complicated equipment, it is easy to transport and set up, making it ideal for outdoor and fieldwork applications. This welding method can be used on a wide range of metals, including carbon steel, stainless steel, cast iron, and some non-ferrous metals, making it highly versatile. Additionally, MMA welding can be performed in various positions, such as horizontal, vertical, overhead, or flat, which further enhances its adaptability to different welding scenarios.

However, the process does require some skill and practice to master, as the welder must maintain the correct arc length and electrode position throughout the weld. Another challenge of MMA welding is the slag that forms from the flux coating, which must be removed after the weld cools. This adds an extra step to the process and requires more time for post-weld cleaning compared to other welding methods like MIG or TIG welding. Despite these limitations, MMA welding remains a widely used method in industries like construction, maintenance, shipbuilding, and general fabrication due to its reliability, ease of use, and flexibility in various working conditions.

Overall, while MMA welding may not be the fastest or most automated welding method, its ability to deliver strong, durable welds in a variety of settings ensures its continued relevance in many industrial applications.

1.1.4 Other Types of Arc Welding

Submerged Arc Welding (SAW) is a process where the weld is protected by a layer of granular flux. The flux not only shields the weld from contamination but also improves weld quality and allows for deep penetration. SAW is often used for welding thick materials in a horizontal position, and it is highly efficient for large-scale projects like shipbuilding or pipeline construction. The process can be automated, making it ideal for industries that require high productivity and consistency. However, SAW is typically limited to flat or horizontal welding positions due to the flow of the molten flux[11].

Flux-Cored Arc Welding (FCAW) is similar to MIG welding but uses a special flux-cored wire that provides its own shielding gas as it burns. This makes FCAW suitable for outdoor environments where wind may disrupt traditional gas shielding. It is particularly effective for welding thicker materials and for use in construction, shipbuilding, and heavy machinery fabrication. While FCAW can generate more spatter and requires post-weld cleaning, it offers the advantage of being able to weld in a variety of conditions [12].

Plasma Arc Welding (PAW) is a more advanced form of TIG welding that uses a highly concentrated plasma arc to achieve a precise and high-temperature weld. Plasma arc welding is known for its ability to produce very narrow and deep welds, making it ideal for precision work on thin materials or in applications that require high-quality, defect-free welds. The process is often used in aerospace, electronics, and medical device manufacturing due to its accuracy and control[13].

2. Approaches and techniques used in arc welding processes

In this section, the techniques and approaches used in the studies on arc welding processes that we examined are examined. Techniques and approaches other than those used in the studies are not examined.

2.1. Machine Learning

Machine Learning is defined as a way to use mathematical models to help a computer learn on its own without intervention. It is a subfield of Artificial Intelligence. Therefore, sometimes artificial intelligence and machine learning are used interchangeably. However, it is important to note that while all machine learning solutions are a form of artificial intelligence, not all artificial intelligence solutions involve machine learning.

Machine Learning is divided into two categories: Supervised and Unsupervised Learning. Supervised Learning is a type of machine learning where a model is built by training labeled data to make predictions. In supervised learning, the goal is for the model, trained with labeled input data, to correctly predict the outputs for new test data. Algorithms such as XGBoost, which is based on decision trees and provide faster and more effective solutions by reducing the loss function [14], Naive Bayes for probabilistic classification, and SVM used for regression, classification, and outlier detection are examples of supervised learning algorithms [15].

Random Forest is a solution to the overfitting problem of decision trees by combining multiple decision trees. The dataset is randomly divided into smaller parts to create decision trees. The performance of the entire model is determined by averaging the results of each decision tree [16].

Unsupervised Learning, on the other hand, involves training a model to find similarities or patterns within unlabeled data. Since there is no labeled data in unsupervised learning, it learns on its own and is used to find relationships or clustering within data. Clustering algorithms like K-means, anomaly detection, and principal component analysis are examples of unsupervised learning [15].

2.2. Artificial Neural Network (ANN)

ANN (Artificial Neural Networks), namely Artificial Neural Networks, is an artificial intelligence model inspired by biological nervous systems. These networks perform tasks such as making predictions, classification, and decision-making by learning from data. ANNs are modeled with nerve cells (neurons) and the connections (weights) between them. ANN basic components; Input Layer, Hidden Layers, Output Layer, Neurons, Weights, and Activation Function. Data is transferred to the network in the Input Layer. Hidden Layers are processing layers inside the network. Inputs are processed in these layers and converted into more abstract representations with weights and activation functions (such as ReLU, sigmoid, and tanh). The layer where the final prediction or output of the model is formed is the output layer. ANNs are usually trained using the backpropagation algorithm to learn from data. This process calculates how much the model output differs from the expected result (error) and updates the weights to minimize this error. Errors are minimized using the gradient descent method together with the backpropagation algorithm [17].

2.2.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS (Adaptive Neuro-Fuzzy Inference System) is a hybrid learning model that combines the strengths of artificial neural networks (ANN) and fuzzy logic systems. ANFIS integrates the adaptive learning capability of neural networks with the inference mechanism of fuzzy logic to create a powerful tool for complex decision-making and modeling non-linear systems.

ANFIS consists of 5 layers; Input Layer, Fuzzification Layer, Rule Layer, Output Layer, Summation Layer [18].

2.2.2 BP (Backpropagation) Neural Networks

Backpropagation Neural Networks are a type of Artificial Neural Network (ANN) that uses the backpropagation algorithm to train the network. The backpropagation algorithm is the foundation for learning in many neural networks, especially in supervised learning tasks.

Backpropagation Neural Networks (BP Neural Networks) are multilayered neural networks that adjust their weights through the backpropagation process. This algorithm works by propagating the error from the output layer back to the input layer, allowing the network to fine-tune its weights and improve performance on a given task [19].

2.3. Deep Learning

Deep Learning, a subset of Machine Learning, operates on artificial neural networks designed based on neural networks in the human brain. Typically trained with labeled data, deep learning can create models capable of recognizing complex tasks and patterns. The created model can predict outputs or perform classification.

2.3.1 Convolutional Neural Network (CNN)

DNN (Deep Neural Network) is a type of artificial neural network with multiple layers between the input and output layers. The "deep" in DNN refers to the presence of multiple hidden layers, which enables the network to learn complex patterns and representations in data. Each layer in a DNN consists of neurons, or nodes, that are connected to neurons in the adjacent layers, and these connections are associated with adjustable weights.

2.3.2 Convolutional Neural Network (CNN)

Convolutional neural networks are a commonly used deep learning algorithm for analyzing images from large datasets. They use linear algebra, specifically matrix multiplication, to recognize images, bringing them into a format suitable for computer recognition and processing. The matrix format provides scalability in image classification and object detection applications.

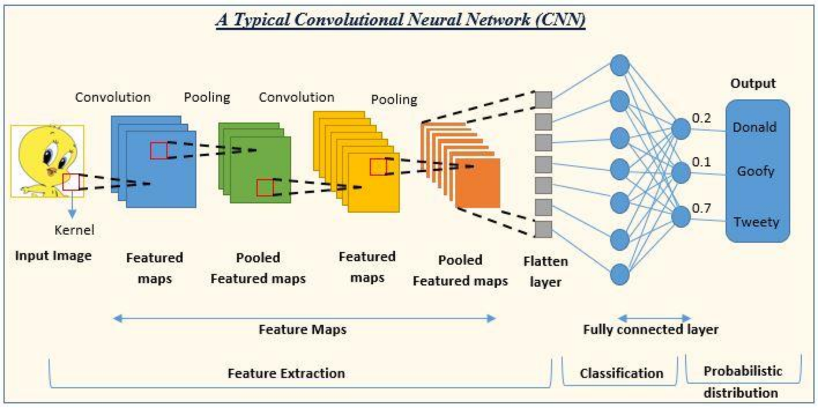

Figure 3. Convolutional Neural Network Layers [20]

Convolutional Neural Networks consist of several layers. Convolutional layer applies filters to the input images to produce feature outputs. Pooling layer reduces the data size by taking subsamples from the input data. Flatten Layer flattens the data coming in matrix form for transmission to the fully connected layer. The input data for the fully connected layer is created in this layer.

R-CNN, a region-based CNN, is used to identify the classes of objects and their bounding boxes in images. Mask R-CNN, an advanced version of R-CNN, can also predict masks for each object based on pixels, in addition to object detection or classification. It is commonly used in areas such as facial recognition and autonomous driving [21].

ResNET, one of the frequently used architectures of CNN, is designed to address the degradation problem and the vanishing gradient issue caused by the increased number of layers in deep neural networks. It facilitates training, enhances model performance, and provides higher accuracy rates [22].

2.3.3 Recurrent Neural Networks (RNN)

A Recurrent Neural Network (RNN) is a deep learning technology designed to handle time-series data. RNNs are applied and used in many research areas, including translation, document summarization, speech recognition, image recognition, disease prediction, click-through rate prediction, stock prediction, synthetic music, and e-commerce fraud detection.

In an RNN, feedback from the output variable is provided to the input. The feedback variable contains a time-delayed network. RNNs use directed cycles to solve problems in the context of input nodes. They overcome the connection between the traditional neural network structure layer and the hidden layer. The transition between each layer node is no longer an input to a hidden layer. An RNN is a sequence-to-sequence model that can appropriately handle sequential data of any length.

The idea of RNN structure is to fully utilize the information from the previous sequence, which is common in traditional neural networks. It assumes that all inputs or outputs are independent of each other. An RNN is directly transformed into a convolutional neural network because different inputs pass through the same neural network, and the difference is hidden by the past state information of the hidden layer.

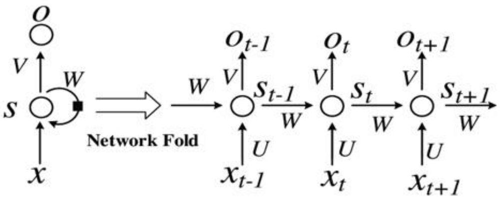

Figure 4. Basic Structure of a Recurrent Neural Network[23]

As seen in Figure 4, input data in an RNN structure is labeled as {x0, x1, x2, ..., xt+1}, while output units and hidden units are respectively labeled as {o0, o1, o2, ..., ot+1} and {s0, s1, s2, ..., st+1}. Hidden units do the most critical work. Information flows unidirectionally from inputs to hidden units and output units. Each output unit is sent to the hidden unit, performing backpropagation. Thus, the input data going to the hidden layers also include the past state information in the hidden layer.

2.3.4 Long Short-Term Memory (LSTM)

LSTM, or Long Short-Term Memory, is a neural network used in recurrent neural networks. It is particularly successful in learning long-term dependencies in sequence prediction problems.

A traditional RNN has a single hidden state transmitted over time, making it difficult for the network to learn long-term dependencies. LSTMs solve this problem by using a memory cell that can retain information for a long period. The memory cell is controlled by three gates: the input gate, the forget gate, and the output gate. These gates determine which information should be added to the memory cell, which should be removed, and which should be output from the memory cell [24].

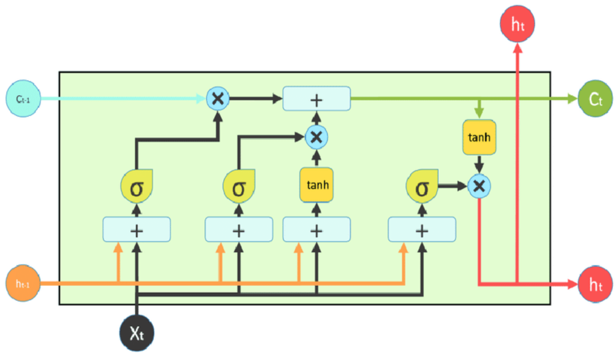

Figure 5. Long Short-Term Memory Neural Network (LSTM) Structure [24]

The input gate controls what information is added to the memory cell. The forget gate controls what information is removed from the memory cell. The output gate controls what information is output from the memory cell. This allows LSTM networks to selectively retain or discard information as it flows through the network, enabling them to learn long-term dependencies.

In LSTMs, the gates are composed of the sigmoid function. The sigmoid function produces values between 0 and 1. Due to its ability to produce positive values and definitive outputs (0 or 1), the sigmoid function is used in LSTM gates. If the value is zero, it means all information is blocked. Conversely, if the value is one, it means the information is allowed to pass.

BiLSTM-CTC is highly effective in sequence-to-sequence tasks, particularly in scenarios where alignment between inputs and outputs is uncertain, such as speech and handwriting recognition. The BiLSTM's ability to learn from both past and future context, combined with the CTC loss function's flexibility, makes this a powerful architecture for time-series and sequence problems.

2.3.5 1D Convolutional Neural Networks (1DCNN)

Deep learning models have made significant progress in automatically extracting distinctive features from time series data for classification tasks. The 1DCNN model utilizes channel attention mechanisms and residual connections to boost performance in detecting errors in signals. It shares core components with traditional CNNs like convolutional layers, ReLU activations, batch normalization, and global average pooling. A key innovation is the Squeeze and Excitation (SE) module, which models channel dependencies to emphasize important feature maps and suppress less useful ones. The residual connections enable effective reuse of features, reducing redundant computation. Collectively, the channel attention via SE and residual connections in 1DCNN allow for improved feature extraction and classification accuracy on time series data.

2.3.6 Genetic Algorithms

GAN (Generative Adversarial Networks) has provided an enhanced related dataset by using an architecture that employs two neural networks: a generative model and a discriminative model, which compete with each other to improve their predictions [25].

2.3.7 Hybrid Fuzzy Deep Learning (HFDL)

Hybrid Fuzzy Deep Learning (HFDL) refers to an advanced machine learning approach that combines elements of deep learning with fuzzy logic. This hybrid system is designed to leverage the strengths of both methodologies deep learning's capacity to model complex, hierarchical data and fuzzy logic's ability to handle uncertainty and approximate reasoning. By integrating these two, HFDL can effectively tackle problems involving ambiguous or imprecise data while still benefiting from deep learning's feature extraction and prediction capabilities[26].

3. Research Methodology

First, information on arc welding processes was collected. Subsequently, the defects and problems encountered in these processes were identified. After that, a literature review performed to explore the information technology solutions to fix these issues. Determining the right keywords was a crucial part of this research. The keywords used for the literature review are listed in Table 1.

Table 1. Keywords used in literature review

Keywords	Synonyms
Arc welding processes	Welding processes, mig/mag, tig
Deep Learning	CNN, LSTM, RNN
Machine Learning	kNN,
Welding deep learning	Tig welding
automotive	Auto, vehicle
Artificial intelligence	AI

The databases searched included Google Scholar, ACM Digital Library, Scopus, IEEE Xplore Digital Library, and arXiv.org. The search was filtered to include studies published from 2018 onwards containing the keywords listed in Table 1.

3.1. Purpose of the Review

Every review seeks to answer specific questions. The questions addressed in this review study are as follows:

- **Q1.** What approaches and techniques have been used in studies related to defect detection in arc welding processes?
- **Q2.** What datasets have been used in these studies?
- **Q3.** What performance metrics have been used in these studies?
- **Q4.** What are the performances of the studies according to these metrics?

Based on these research questions, a systematic literature review has been performed.

4. Review and Findings

In this section, a summary of the reviewed articles, dataset analysis, techniques and approach analysis, and performance analyses are provided.

4.1. Literature Summary

Recent studies on artificial intelligence in solving defects in arc welding processes, particularly in the automotive sector, have been reviewed. These studies employ different methods and approaches, resulting in varying performance metrics and datasets. This review analyzes both the differences and similarities in these studies. Additionally, the commonalities among the reviewed studies are identified to show the direction of the research.

Welding evaluation processes are generally performed post-welding, identifying defects after the welding is completed, which often leads to the disposal of expensive materials or lengthy repair processes. Even for automatic welding machines, quality inspections are typically conducted manually by humans. The studies reviewed aim to enable control by an auxiliary support mechanism. These support mechanisms demonstrate that deep learning methods perform better than general machine learning methods for time series quality classification prediction and diagnosis in various processes and systems. To address this issue in machine learning, a parallel strategy has been adopted where the overall dataset is classified, and each classified data subset is further divided into different subgroups, with a machine learning model applied to each subgroup. However, deep learning models using the LSTM structure have now dominated these efforts.

In the study by Kaiser Asif et al., an acoustic emission monitoring method for gas metal arc welding in real-time is introduced. The acoustic emission system is designed to cover a wide frequency range from 5 to 400 kHz. Various features extracted from the acoustic emission and welding parameters are fed into a machine learning algorithm. Additionally, welding parameters (current, voltage, gas flow rate, and heat input) are recorded simultaneously with the acoustic emission. Adversarial Sequence Tagging and Logistic Regression methods are applied to predict the presence of four welding conditions: good, over-penetration, burn-through, porosity, and spatter-over-penetration[27].

Due to the complexity of the resistance spot welding process, accurately knowing the operational state of the welding robot under current parameter settings and evaluating the quality of electrode caps under different sheet metal types in real-time remains a challenge. To address this issue, Geng Chen et al. (Chen et al., 2022) propose a parallel strategy for predicting the quality of welding connections using machine learning for subsets of data with different distribution models. First, PCA (Principal Component Analysis) dimensionality reduction model was used to reduce the dimensionality of welding process feature values and the classification difficulty of data subgroups. Based on the dimension reduction data, the k-means model was applied to complete the classification of sub-data sets. The elbow method was used to adjust the number of clustering centers. Finally, the feature parameters of each sub-data set were used as inputs for machine learning, and a parallel prediction strategy for welding connection quality was developed based on the data distribution characteristics of each sub-data set. The parallel strategy adapted to the characteristics of the data population works well with more complexly distributed large welding data [28].

In the welding process of zinc-coated steel, zinc vapor causes serious porosity defects. Porosity is a significant indicator of weld quality and reduces the durability and productivity of the weld. Therefore, Seungmin Shin et al. propose a method based on a deep neural network (DNN) that can detect and predict porosity defects in real-time without the need for an additional device based on the welding voltage signal in the GMAW process. The welding current and arc voltage signals produced in the GMAW process are measured in real-time. Then, data preprocessing is performed at 0.1-second intervals to extract feature variables from the welding voltage signal. Meaningful feature variables were selected through correlation analysis between the feature variables and porosity defects, and classification models based on artificial neural networks and deep neural networks were developed. The proposed DNN-based model showed a 15% improvement in performance over the artificial neural network model. [29]

Monitoring the quality of welded joints is an important issue in cold metal transfer (CMT) overlay welding. In the CMT welding process, there are usually some abnormal problems in actual production. For example, welding gas leak, caused by excessive assembly gaps and insufficient shielding gas directly affects weld quality. In this study, Liang Liu et al. conducted an experiment on the CMT overlay welding of low-carbon steel sheets and examined the sound characteristics of CMT. In different operating modes of CMT, short-term comparisons were made of welding voltage, welding current, and arc sound. It was found that the faster the arc changes, the higher the corresponding sound pressure value. Additionally, under two abnormal welding conditions, insufficient gas supply and welding wear defect, the analysis of welding electrical parameters and welding sound signals using feature extraction and fusion methods was examined. The BiLSTM-CTC model was proposed to detect insufficient gas supply and welding wear in the welding process [30].

Kevin Meyer and Vladimir Mahalec, in their 2022 paper "Anomaly detection methods for infrequent failures in resistive steel welding" [31], attempted to detect defects in the seam resistance welding machine using a single-class neural network autoencoder and PCA (principal component analysis). The models were trained with only normal welding process data without faulty welding data, and defect detection was performed with these trained models [31].

In the paper entitled "Development of anomaly detection model for welding classification using arc sound," Phongsin Jirapipattanaporn and colleagues tried to detect the type of welding process result by using the arc sound of gas metal arc welding, commonly used in the industry, with artificial intelligence. Recurrent neural networks (RNN), long short-term memory (LSTM), and support vector machines (SVM) were used as artificial intelligence algorithms for defect detection in this project [32] . In a similar study titled "Detection of defective welds made by gas metal arc welding through analysis of sound signals," the sound signals of gas metal arc welding and the current signals drawn by the machine were used to detect problematic welds. After collecting the necessary signals and sounds, signals in the time domain were analyzed using the principal component analysis (PCA) method. These analyzed data were used with artificial neural networks, support vector machines, decision trees, and nearest neighbor approaches to try to detect defects [33].

The University of Birmingham's Metallurgy and Materials Engineering department designed a system to monitor the TIG welding process using an HDR camera capturing images of the weld pool and surrounding area and a training system based on neural networks to identify welding defects. Unlike previous studies, the camera filters the strong light emitted by the arc, and the system adapts without needing to identify the welding direction [34]. In a similar study, seam inspection based on image processing and deep learning was conducted. This study, which aimed to achieve a 97% accuracy rate on defective reference parts, compared a standard deep learning algorithm applied to raw data with data augmentation approaches [35].

In a study conducted in China, four parameters were identified as having a greater impact on welding quality, and two main welding parameters that could determine other parameters were obtained based on correlation analysis. The study concluded that welding current and welding speed could impact welding quality by 85%. ELM-based algorithms were found to be more effective and efficient in predicting the optimal welding voltage range [36].

Another study conducted in China investigated factors affecting welding quality. Lin et al. developed a deep learning-based model for welding quality analysis and prediction to address the challenges posed by the high nonlinearity of the welding process and the complex interaction of multiple factors. Utilizing data accumulated from an automotive production line, the study proposed a Residual Neural Network (ResNet) model to predict welding quality. The model incorporated key input parameters such as the characteristic point at the end of the metal bonding process, maximum voltage, and the wear state of the electrode cap. Their results demonstrated an accuracy rate of 88.9%, offering a foundation for the automated control of welding quality [37].

Ma et al. developed an intelligent welding robot system utilizing deep learning and machine vision. Their approach involved a calibrated binocular vision system, where 1000 images were captured, and the weld positions were labeled using LabelImg software. The dataset was split into training and validation sets in a 3:1 ratio, and a convolutional neural network was trained until convergence. The TensorFlow model was optimized and deployed to a Raspberry Pi 3B+ using Intel's NCS2 edge computing stick to identify weld locations in images. By matching left and right images, the system calculated weld depth through parallax. Finally, the UR3 robot was controlled via Ethernet for automated welding [38].

Jin et al. proposed a welding seam recognition model based on the Mask R-CNN network to address challenges in vision-guided welding systems, particularly in recognizing complex welding seams. Traditionally, stereo vision systems are used to correct welding seam position defects caused by local overheating or positional deviations, but they struggle with accurately guiding laser sensors along curves with gradually changing curvatures. By employing transfer learning, Jin et al.'s model effectively identifies welding seams and segments instances in images, improving the accuracy of solder joint positioning. This approach enhances the precision of industrial robots in welding processes [39].

Charbel El Hachem worked on the automation of quality control and the reduction of nonconformity using machine learning techniques at Faurecia Clean Mobility. The thesis mentions five different contributions to the literature. The first contribution is automated quality control. The second focuses on the automation of quality control for weld seams that cover external aspects of weld defects and cannot be reached by leak tests, using deep learning methods. The third contribution combines deep learning model explain ability with weld seam classification accuracy. A hybrid approach of CNN-Machine Learning classifier was proposed to increase the accuracy achieved in the second contribution. This study presents a new model-driven optimization that achieved over 98% accuracy when applied to a weld seam dataset. The fourth contribution examines the location and relative fractures of the ceramic monolith. Quality control of these monoliths should be carried out during the production of exhaust pipes. A comparison of image processing filters for straight line detection is presented. Tests were performed by rotating the ceramic in 5-degree increments. In 2021, Hachem et al. conducted a paper study based on the doctoral thesis. In this study, ResNet-50 and MobileNet were used. Classification of weld seams was performed using XGBoost, Decision Tree, SVM Linear, and SVM Poly5, and accuracy rates were measured [40]. In another study, Hachem et al. worked on the classification of weld seams in the automotive industry using deep learning algorithms. The welds created in this study were classified into six categories. MobileNet was used as the CNN algorithm. Data augmentation methods were used to increase the number of images used in training [35].

Martin Appiah Kesse's doctoral dissertation addresses this topic. In his study, Kesse presented a modern approach using artificial intelligence to improve efficiency and welding quality in TIG welding. This doctoral thesis aims to contribute to the state of the art in terms of the applicability of AI in welding technology by developing an AI framework using an ANFIS and a fuzzy deep neural network. The study used an ANFIS fuzzy logic model and a DNN model. After taking the necessary features for the welding process as input data, the required data is provided to the automatic welding system using a fuzzy-driven deep network model [41].

In another study that forms the basis for Kesse's doctoral thesis, Kesse et al. modeled an artificial intelligence system to predict the structural integrity in robotic GMAW of UHSS corner-welded joints [42].

Nogay and Akıncı conducted a study focused on the classification and determination of operating zones in MAG electric arc welding machines using deep learning techniques. The study analyzed the welding current graphs over a 5-second operation period. To classify the operating zones, they utilized five deep convolutional neural networks, four of which were pre-trained models through the concept of transfer learning. Their results showed that the designed model achieved 93.5% accuracy in estimating the operating range of the welding machine, while the pre-trained models achieved accuracy rates between 95% and 100% [43].

Wang et al. proposed a hybrid model that utilizes multi-source information fusion to enhance defect recognition in MIG lap welding processes. To address the limitations of single-sensor recognition methods, the model integrates molten pool images and voltage signals to detect defects such as burn-through, weld bias, and uneven width. Using a synchronized trigger device, molten pool images and weld voltage data were collected, and the Short-Time Fourier Transform (STFT) was applied to analyze the frequency variations in the voltage data. The molten pool images and time spectrograms were then fed into a hybrid CNN model for real-time defect detection. The results demonstrated that the hybrid model achieved an accuracy of 98.87%, outperforming other single-sensor-based models. The confusion matrix and Grad-CAM visualizations showed that the hybrid model effectively detects transition states in welding, offering a more accurate and balanced recognition across the entire process[44].

El Houd et al. focused on addressing the limitations of manual visual inspection in welding seam classification, which remains subjective and costly in many industries. They explored the integration of deep learning methods to enhance the accuracy and reliability of welding seam classification, particularly in the automotive sector. The study proposed a novel hybrid approach, combining model prediction scores with visual explanation heatmaps to improve defect classification. Their results demonstrated that the hybrid model outperformed target metrics, increasing classification accuracy by at least 18%, highlighting the importance of deep learning model explainability and interpretability in industrial applications[45].

Pernambuco et al. proposed a low-cost system for monitoring the stability and transfer mode of the MIG/MAG welding process by analyzing the sound signal produced by the electric arc. Recognizing the challenges of instability in welding caused by electrical variables, the authors introduced a nonintrusive, real-time solution using an Artificial Neural Network (ANN) to detect discontinuities in the welding process. A sound signal dataset was developed through experiments simulating real-world welding conditions, including processes with adequate welds and those with two types of discontinuities. The methodology was validated using classification accuracy and a confusion matrix, demonstrating that discontinuities can be identified solely by analyzing the sound generated during welding [46].

Weld defect detection is an important task in the welding process. Although there are many excellent weld defect detection models, there is still much room for improvement in stability and accuracy. In this study, a lightweight deep learning model called WeldNet is proposed to improve the existing weld defect recognition network for its poor generalization performance, overfitting, and large memory occupation, using a design with a small number of parameters but with better performance. We also proposed an ensemble-distillation strategy in the training process, which effectively improved the accuracy rate and proposed an improved model ensemble scheme. The experimental results show that the final designed WeldNet model performs well in detecting weld defects and achieves state-of-the-art performance. Its number of parameters is only 26.8% of that of ResNet18, but the accuracy is 8.9% higher, while achieving a 24.2 ms inference time on CPU to meet the demand of real-time operation. The study is of guiding significance for solving practical problems in weld defect detection, and provides new ideas for the application of deep learning in industry[47].

Li et al. developed a vision-based monitoring system for automatic tungsten inert gas (TIG) welding of thin plates with a reserved gap, focusing on weld pool shape and fusion hole size as key factors for weld quality. To accurately detect the unstable location of the fusion hole, they proposed an improved YOLOv3 network with ResNet-d as the backbone. The optimized YOLOv3 achieved 95.03% mAP50 accuracy and a prediction speed of 49.43 frames per second (FPS). Additionally, a region-growing image processing algorithm was designed to extract clear edges and the width of the fusion hole, with the algorithm's accuracy verified through comparison with simultaneous backside images[48].

Wang et al. developed a visual sensing system for tungsten inert gas (TIG) welding, focusing on molten pool contour extraction, a key aspect of online welding quality monitoring. They designed a multi-scale feature fusion semantic segmentation network, Res-Seg, based on a residual network to improve target segmentation. To enhance the generalization ability of the model, deep convolutional generative adversarial networks (DCGAN) were used to augment the molten pool dataset, followed by color and morphological data enhancement. The proposed method was compared with traditional edge detection algorithms and other segmentation networks, demonstrating superior accuracy and robustness in real welding environments. Additionally, a back propagation (BP) neural network was used to predict weld width, with the system achieving an average defect of less than 0.2 mm, meeting welding accuracy requirements [49].

Zhang et al. proposed a deep learning-based approach for real-time defect detection in keyhole tungsten inert gas (TIG) welding, leveraging a multi-layer deep neural network trained on a large welding image dataset. Unlike support vector machines (SVMs), which are limited by manual kernel selection and perform poorly in recognizing complex weld defects such as burn-through, the neural network excels in capturing deep feature maps of molten pools and distinguishing between various weld states. The study employed a four-class classification task, identifying good welds, burn-through, partial penetration, and undercut with high accuracy and real-time performance. The method's effectiveness was supported by a comprehensive dataset, enhanced through preprocessing and data augmentation, offering a robust solution for quality control and defect prevention in keyhole TIG welding [50].

Xia et al. developed a visual monitoring system for keyhole Tungsten Inert Gas (TIG) welding, aimed at improving manufacturing quality and automation through real-time process monitoring. Using an HDR welding camera, the system monitored the weld pool and keyhole during the welding process.

A ResNet-based convolutional neural network was implemented to classify various welding states, including good weld, incomplete penetration, burn-through, misalignment, and undercut. To enhance the training dataset's diversity, image augmentation was applied, and the training process was optimized using a center loss metric learning strategy. Visualization techniques such as guided Grad-CAM, feature maps, and t-SNE were used to explain the deep learning process, providing insights into the network's effectiveness. This research lays the groundwork for developing an online monitoring system for keyhole TIG welding [51].

4.2. Input-Output Characteristics and Data Analysis

In this section, we aim to address the answer to research question Q2. During the data analysis of the studies, input data, storage method, dataset used, and output data were examined. The status of each study in these areas is shown in Table 2.

The principal objective of evaluating the methodologies and approaches employed in the research is to comprehend the ways in which artificial intelligence addresses issues in Arc welding procedures. Table 2 provides a summary of our analysis. Table 2 makes it clear that deep learning and machine learning techniques are used most often. Studies using machine learning have applied techniques such as classification, feature selection, prediction, correlation analysis, and regression. Algorithms such as XGBoost, K-means, and BP Neural Network have been utilized in these studies. A common characteristic of studies employing machine learning is that the incoming data is generally numerical.

Table 2. Data analysis of studies

Reference	Input Data	Storage Method	Dataset Used	Output Data
[27]	Welding Parameters (welding current, voltage, gas flow rate and heat input), acoustic emission	Local	Own dataset	Weld Quality (good (G), excessive penetration (Pn), burn-through (B), porosity (Pr) and porosity-excessive penetration (Pr + Pn))
[28]	Welding Parameters	Local	Own dataset	Welding quality prediction
[29]	Standard deviation of voltage, Standard deviation of instantaneous short circuit voltage, Average electrical arc (Arc) time, Number of short circuit periods, Standard deviation of short circuit duration, Standard deviation of voltage during electrical arc (Arc) period	Local	Own dataset	Porosity detection and estimation in welds
[30]	Welding Current, Welding voltage, Electrical arc sound pressure, Welding heat input electrode voltage, welding current	Local	Own dataset	insufficient gas supply and weld wear defect detection

Table 2. Continued.

Reference	Input Data	Storage Method	Dataset Used	Output Data
[31]	Welding Parameters (Temperature, current, voltage, clamping forces, carrier position, centering position)	Local	Own Dataset	Anomaly Detection
[32]	Welding Sound	Local with Raspberry Pi		weld bead type
[33]	Current and Audio Signals	WaveSCAN 2.0	Own Dataset	Online Weld Fault Detection with current signals
[34]	Welding images taken from the camera reading 55 fps are 1280x1024 pixels	Local	Own dataset	Welding quality prediction
[35]	Welding images	Local	Own dataset	Classification of source image
[36]	Oxygen Pressure, Pulse Width, Pulse Frequency, Cutting Speed, Welding Speed, Welding Current, Gas Pressure	No information	Own dataset	Welding Quality
[37]	maximum voltage, Wear status of the electrode cover	Local	Own dataset	Weld quality estimation
[39]	Welding Images	Local	Weld seam images	Irregularity of Weld Seam
[40]	Images taken from 18Mp camera	Local	Images of Faurecia Clean Mobility company	Weld seams classification
[41]	Flow, Electrical jump (Arc) length, welding speed, Weld bead width, penetration width, Depth of the heat affected zone, Width of heat affected zone	Local	Own dataset	Tensile Strength and yield strength estimation
[42]	Electrical surge (Arc) current, Electrical surge (Arc) voltage, welding speed, gas flow rate, electrode extension, Electrical jump (Arc) length, wire feeding speed, wire diameter, welding torch position, Welding torch movement angle	Local	Own dataset	Estimation of resource properties
[43]	Welding current graph,	Local	Own dataset	Classification of operation cases in electric arc welding machine

Table 2. Continued.

Reference	Input Data	Storage Method	Dataset Used	Output Data
[45]	Welding Images	Local	Own dataset	Classification of Weld Defective and Perfect
[46]	Welding audio signals	Local	Own dataset	Welding defect detection
[47]	Welding Images	Local	TIG Al5083 [34]	Weld defect detection
[48]	Welding images	Local	Own dataset	Detect fusion hole state and size during TIG welding
[49]	Welding Images	Local	Own dataset	Extract the contour of the molten pool in TIG stainless steel welding
[50]	Welding images	Local	Own dataset	Classification of weld quality
[51]	Weld images, Grad-CAM Images Guided Backprop	Local	Own dataset	Classification of weld quality

Deep learning studies have used techniques such as classification and object detection. Studies focused on classification have categorized the quality of MIG/MAG welds based on images, classifying them as either defective or non-defective. Object detection techniques have been used alongside classification to identify defects in welds. Deep learning studies have employed network models and algorithms such as ResNet, LSTM, and MobileNet. These models and algorithms may have been preferred due to their speed and low error rates.

In summary, upon reviewing recent studies in this field, it is observed that both machine learning and deep learning are utilized in nearly all studies. In this domain, only image processing or signal processing techniques are not solely used. Researchers should determine their approach and method based on the type of incoming data (image, numerical data), and choose algorithms and models according to time, speed, and hardware requirements.

4.3. Analysis of Methods and Techniques Used

In this section, the approaches and techniques used in the studies are analyzed. Additionally, for approaches utilizing deep learning, any CNN models and libraries employed are identified.

Table 3. Method and Technical Analysis of Studies.

Reference	Approach	Techniques Used	Network Model or Algorithm
[27]	Machine Learning NLP	Classification Attribute Selection	Logistic Regression - Adversarial Sequence Tagging (AST)
[28]	Machine Learning	Classification Classification	BP Neural Networks / PKB (PCA-K-means-BP)
[29]	Machine Learning	Classification, Correlation analysis	Deep neural networks (DNN), Artificial neural network (ANN)
[30]	Deep Learning, Audio and Signal Processing, Time series	Classification, normalization	BiLSTM-CTC LSTM
[31]	Machine Learning Signal Processing	Classification	PCA, Autoencoder, SUM-PCA, DUM-autoencoder
[32]	Machine Learning	Classification	RNN, SVM, LSTM
[33]	Machine Learning Signal Processing	Classification	PCA, Naive Bayes C. SVM, Decision Trees, Boosted Trees, Random Forest, ANN, Nearest Neighbor, KNN

Table 3 Continued.

Reference	Approach	Techniques Used	Network Model or Algorithm
[34]	Deep Learning	Binary image classification, Multi image classification	Conv6, Conv2, FullyCon6, FullyCon2
[35]	Deep Learning, Image processing,	Binary classification, Data augmentation,	CNN, MobileNet
[36]	Machine Learning	Regression, Correlation analysis,	Back Propagation Neural Network, Extreme Learning Machine, Support vector machine
[37]	Deep Learning	Normalization, Regression, Residual connection	ResNet18
[39]	Deep Learning Image processing	Classification Object Detection – object identification	Mask R-CNN / ResNet - RoIAlign(a method used to extract a small feature map from each RoI)
[40]	Image processing, Deep Learning,	Data augmentation, Image classification, Imbalanced data,	ResNet50, MobileNet-XGboost classifier, MobileNet-SVM Poly5 kernel, GradCAM, Support Vector Machine Classifier, Decision Tree Classifier, canny filter, Laplacian edge detector, Hough lines
[41]	Machine Learning	Regression, Data augmentation	Adaptive Neuro-fuzzy Inference System (ANFIS), DNN
[42]	Machine Learning	Regression	Artificial neural network with back propagation, Artificial neural networks with Levenberg-Marquardt
[43]	Deep Learning	Classification	DCNN
[45]	Deep Learning Machine Learning	Detection Classification	ResNet MobileNet – XGBoost MobileNet-SVM
[46]	Machine Learning	Prediction, Classification	Neural Network(ANN)
[47]	Deep Learning	Classification	CNN, FCN, Own model similar Resnet18
[48]	Deep Learning	Classification	CNN, YOLOv3
[49]	Deep Learning	Segmentation, Data augmentation	Res-Seg, DCGAN
[50]	Deep Learning	Classification	CNN, ResNet
[51]	Deep Learning Machine Learning	Classification	CNN, Center-loss ResNet, ResNet, SVM,

4.4. Performance Metrics

Performance metrics are used in academic research to measure the performance of an approach. Evaluating the efficiency and accuracy of machine learning models, particularly deep learning models, relies heavily on performance measurements. There are several performance metrics designed to capture different aspects of model performance for each algorithm. This performance is typically measured by comparing the model's predictions or classifications with the ground-truth values obtained from labeled data.

4.4.1 Classification Metrics

Confusion Matrix is a comparison of model predictions and ground-truth labels in tabular form. When the output can include two or more different types of classes, it is the simplest approach to measure how well a classification task performs. The confusion matrix serves as a basis for evaluating results rather than serving as a performance metric on its own. Accuracy is the percentage of examples that the algorithm correctly classifies out of all examples. It is useful when all classes are equally important.

Precision represents the ratio of true positive results to the number of all positive results predicted by the model. Precision indicates the probability of a positive result being correct. When dealing with imbalanced distributions of classes, classification accuracy is not an ideal measure to evaluate how well a model works. We need a precision measure that is produced by dividing true positives by the sum of true positives and false positives to address a specific issue for a certain class [52]. Precision indicates the percentage of positive predictions that are actually correct. Recall, also known as Sensitivity, is a measure of how many of the actual positive instances were correctly predicted. Specificity measures the performance in categorizing negative examples. It represents the percentage of true negatives correctly identified among all negative examples. F1 Score is the harmonic mean of precision and recall and is used to evaluate the accuracy of a test. It can range from a minimum of 0 to a maximum of 1. The model's robustness and sensitivity are generally measured with it. It is difficult to evaluate two models with high recall but low precision. To compare them, we use the F1 score. The F1 score is the harmonic mean of two numbers, one small and one large, both of which are precision and recall.

The performance of a classification problem at various threshold values is measured using the AUC-ROC curve. ROC is a probability curve, the area under ROC is known as AUC, which is calculated by taking the integral. It speaks of the ability to distinguish different model types. The higher the AUC, the more accurate the model is expected to predict. Thresholds are values that need to be determined depending on the problem.

In computer vision and object detection tasks, Intersection over Union (IoU) metrics are often used to evaluate the accuracy of an object's position. IoU is a frequently used evaluation metric that assesses how accurately an object's location is predicted. The actual location of the object is indicated by the ground truth bounding box. The first step in calculating IoU is to find the intersection area of the predicted bounding box and the ground truth bounding box. If there is no overlap, there is no intersection area. The area created by both boxes together, including the intersection area, is called the union area. IoU measures how much of the predicted bounding box overlaps with the ground truth bounding box. To calculate IoU, the intersection area is divided by the union area:

$$IoU \; = \; Intersection\ Area \;/\; Union\ Area \tag{1}$$

An Intersection over Union (IoU) value of 1 indicates a perfect match, where the predicted and ground truth bounding boxes completely overlap; whereas a value of 0 signifies no overlap. It offers a single scalar value indicating the accuracy in determining object locations, facilitating comparison across various detection techniques or models.

Another popular evaluation statistic in the field of computer vision is the mean Average Precision (mAP), which provides a comprehensive assessment of the performance and accuracy of an algorithm or model across multiple object categories. The Average Precision (AP) score for each object category is determined individually, and the average of these individual AP scores is then used to form the mAP. The general method for calculating mAP is as follows:

- Generate a Precision-Recall (PR) curve. The algorithm or model ranks discovered objects based on confidence scores for each object category. Plotting Precision (y-axis) against Recall (x-axis) as the detection threshold varies results in the PR curve. Precision represents the ratio of true positives to all positive detections, while Recall represents the ratio of true positives to all ground truth objects.

- Calculate the area under the PR curve to obtain the Average Precision (AP) score. Computing the integral of precision values at various recall levels yields this area. For a specific object category, the model's precision-recall performance is summarized by the AP score.
- Compute the mAP: After determining the AP scores for all object categories, the mAP is calculated by taking the average of each category's AP scores. (2).

The Mean Average Precision (MAP) metric is useful because it provides a comprehensive evaluation of the model's detection capabilities by assessing its performance across multiple object categories. By offering a single scalar value that represents the model's average performance, MAP facilitates the comparison of various algorithms or models.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \qquad (2)$$

The False Discovery Rate (FDR) is a measure of accuracy used in multiple hypothesis testing when several measurements or changes are observed in a single experiment, which accounts for multiple hypotheses being simultaneously investigated. The FDR regulates the percentage of false discoveries (also known as false positives) among the rejected hypotheses. As the number of tests conducted increases, so does the probability of false positive findings. The estimated ratio of false discoveries (false positives) to all discoveries (rejected hypotheses) is known as the False Discovery Rate. Since it is dependent on unforeseen outcomes, controlling the FDR is more challenging.

4.4.2 Regression Metrics

Mean Squared Error (MSE) is the average of the squared differences between the target value and the value predicted by the regression model. Essentially, it measures what has been discovered. MSE provides a measure of model accuracy by considering both the size and direction of errors.

Root Mean Squared Error (RMSE) is the square root of the average of the squared differences between the value predicted by the regression model and the target value. A larger RMSE indicates larger prediction errors, while a lower RMSE indicates better prediction accuracy. Mean Absolute Error (MAE) is the average of the absolute differences between the actual values and the predicted values across all observations in the dataset. Better accuracy is indicated by a lower value. This metric does not consider the direction of errors because taking the absolute value guarantees the differences are positive. Unlike MSE, which is differentiable, MAE is not differentiable. It provides a measure of how far the model's predictions are from the actual results. Unlike MSE and RMSE, MAE does not square the errors, making it more robust to outliers. The R-squared metric gives a measure of how well a set of predicted output values matches the actual output values and is often used for explanatory purposes. R-squared is not a reliable indicator of how well a regression model fits your data. A good model might have a low R-squared value, and conversely, a biased model might have a high R-squared value. The model's performance can be evaluated by comparing the R-squared metric against a fixed baseline. The fixed baseline is selected by taking the average of the data and placing the line at the mean.

A model performing at the fixed baseline will have an R-squared value equal to 0. Higher R-squared values indicate better models, with the best model giving an R-squared of 1, indicating all correct predictions. The R-squared value increases as new features are added to the model. We can calculate R-squared as follows:

$$R^2 = 1 - \frac{MSE(Model)}{MSE(Baseline)} \quad (3)$$

$$\frac{MSE(Model)}{MSE(Baseline)} = \sum_{i=1}^{N}(y_i - \hat{y})^2 / (\overline{y_i} - \hat{y}_i)^2 \quad (4)$$

- MSE (Model): Average squared error of predictions compared to the actual values.
- MSE (Baseline): Average squared error of average predictions compared to the actual values.

As more independent variables are included, the R-squared value of a model always increases. As independent variables come into play, the model becomes more complex, and as the model becomes more complex, it tends to overfit. Thus, when features that do not improve the model are added, the R-squared value increases, and R-squared does not penalize them. To address these issues, it is necessary to use adjusted R-squared. Each independent variable is considered by adjusted R-squared only when it significantly improves the variable model. Adjusted R-squared is used to address the problem of R-squared, but it will always show a lower number than R-squared.

4.4.3 Other Metrics

Mean Square Reconstruction Error (MSRE) is commonly used to assess the accuracy of reconstruction in various signal processing and data compression applications, especially when it comes to autoencoders and other unsupervised learning methods. MSRE calculates the mean squared difference between the original and reconstructed forms of input data. The goal when expressing and reconstructing data is typically to minimize distortion or reconstruction error. MSRE provides a quantitative evaluation of a model's or algorithm's ability to reconstruct the original input data. By squaring the inconsistencies between original and reconstructed data, it penalizes larger errors more severely than smaller ones, emphasizing the value of capturing the features and patterns of input data accurately. Lower MSRE levels indicate more accurate reconstruction, as they suggest that the recovered data closely resembles the original input. Conversely, higher values indicate greater discrepancies between original and reconstructed data.

When alignment between inputs and target labels is unknown, Connectionist Temporal Classification (CTC) is used to solve sequence labeling problems, especially for temporal classification tasks. Machine learning, particularly deep learning, employs the CTC method for sequence labeling problems. First proposed by Alex Graves in 2006, it has since been widely used for various applications, including speech recognition, handwriting recognition, and machine translation.

The choice of metric depends on the task at hand, and often requires customization to meet specific requirements and prioritize different objectives. Performance metrics contribute to improving machine learning models by providing useful information about the advantages and disadvantages of various models.

Prediction Interval (PI) provides a measure of uncertainty for predictions in regression problems. For example, a 95% prediction interval indicates that out of 100 samples, the actual value will be between the lower and upper bounds of the interval 95 times. There are several performance evaluation metrics that can be used accordingly.

Coverage Ratio (CR) represents the degree to which actual data is covered by the prediction interval. As this value increases, the accuracy of the model in classifying actual data increases.

Performance analyses of the studies have been conducted and summarized in Table 4. As seen in Table 4, in many studies, since classification is involved, Accuracy has been used as the performance metric to measure classification success. The choice of performance metric varies depending on the desired outcome and the technique used. When making a classification using Deep Learning or Machine Learning, metrics such as Recall, Precision, and F1 score are used. However, when performing object detection with Deep Learning, the IoU metric is used. If the desired output is regression, metrics such as MSE, RMSE, MAE have been used [28].

Table 4. Method and Technical Analysis of Studies

Reference	Performance Metrics	Results
[27]	Accuracy	91.18% accuracy for Adverserial Sequence Tagging
[28]	R^2 , MSE, MAE, MRE	BP: 0.4-0.5 / PBP: 0.81-0.96 Best results in order: 18.4223, 3.7143, 0.0500, 0.9600(96% success)
[29]	Prediction Accuracy	ANN: 0.659(dataset 1), 0.790(dataset 2) DNN: 0.858(dataset 1), 0.895(dataset 2)
[30]	Classification Error Rate	0.295
[31]	MSRE	Mean 0.0267
[32]	Accuracy	RNN Model Test Accuracy: 89.67% LSTM Model Test Accuracy: 78.57%
[33]	Accuracy	Decision Tree model accuracy: 80%
[34]	Accuracy, Precision, Recall, F-score	Fully-conv6: Accuracy: 69% F-score: 0.56 Conv6: Accuracy: 93.4% F-score: 0.78 Fully-conv2 Accuracy: 89.5% F-score: 0.89 Conv2: Accuracy: 75.5% F-score: 0.75
[35]	Accuracy	Overall accuracy: 97 %
[36]	CR, AIW	BPNN Model - CR Average: 0.8516 AIW Average: 1.5084 ELM Model - CR Average: 0.9297 AIW Average: 1.4536 SVM Model - CR Average: 0.9041 AIW Average: 1.4358
[37]	Accuracy	Accuracy: 88.9%
[39]	mAP - Recall	mAP:95% – Recall: 76%
[40]	R^2 Accuracy per Classifier	R^2 Laplacian-Hough lines: 0,8291 R^2 Canny-Hough lines: 0,9969 XG-Boost Accuracy: 99.3% (Weld2) , 98.7% (Weld3) Decision Tree Accuracy: 98.1% (Weld2) , 97.2% (Weld3) SVM Linear Accuracy: 98.8% (Weld2) , 98.1% (Weld3) SVM Poly5 Accuracy: 98.8% (Weld2) , 98.7% (Weld3)
[41]	RMSE	Tensile strength: 2.392 Yield strength: 12,546
[42]	RMSE	0.0000845
[43]	Accuracy	Accuracy : 93.5 %
[45]	Accuracy	MobileNet Model Average Accuracy: 98% ResNet50 Model Average Accuracy: 97%
[46]	Accuracy	Accuracy: 80.61%
[47]	Accuracy Precision Recall F1-score	Accuracy: 83.5% Precision: 75.1% Recall: 78.4% F1-score: 74.9%
[48]	mAP50, FPS	mAP50 Mean: 95.11% FPS: 49.71

Table 4 Continued.

Reference	Performance Metrics	Results
[49]	Segmentation accuracy	Molten pool(M),Background(B) Res-Seg (Based on ResNet-34) M: 91.94% B: 99.33% Res-Seg (Based on ResNet-50) M: 93.71% B: 99.28% Res-Seg (Based on ResNet-101) M: 93.80% B: 99.21% Res-Seg (Based on ResNet-50) + DCGAN M: 94.77% B: 99.32%
[50]	Accuracy, Precision	SVM based model accuracy:
[51]	Accuracy, Precision	Center-loss Resnet Model: Average Precision value: 0.984 Average Accuracy value: 0.984 Resnet: Average Precision value: 0.978 Average Accuracy value: 0. 978 SVM: Average Precision value: 0.90 Average Accuracy value: 0.846

5. Discussion

This study examines the use of artificial intelligence (AI) techniques for defect prevention and quality control in Arc welding processes. The reviewed studies generally cover AI applications such as image processing and time-series analysis, with deep learning methods demonstrating superior performance, particularly in image-based defect detection and time-series analysis. Although deep learning algorithms have achieved higher accuracy rates compared to traditional machine learning methods, these techniques face challenges in industrial environments due to noise, obstructions, and varying welding conditions.

While many studies reported accuracy rates exceeding 90%, it was concluded that further research is needed to improve the performance of these systems in real-world industrial applications. Additionally, AI-assisted error prevention systems are expected to significantly contribute to reducing costs and speeding up quality control processes in manufacturing.

In this context, future research should focus on improving the real-time performance of these systems and enhancing the generalizability of AI models. This review provides a broad overview of the current state of AI techniques for error prevention and quality control in Arc welding processes and aims to guide future research in this field.

Method, dataset and performance analyses of artificial intelligence studies to detect errors in Arc welding processes have been conducted. Many different data types, methods and performance metrics were used in the studies.

The techniques and approaches used in the study vary according to the problems and objectives of the studies. For example, while Deep Learning approaches and techniques are used in a study that classifies the quality of the source from the source image, machine learning techniques are used in a study that classifies the source from the data obtained from various sensors during the process.

Each technique has its specific strengths depending on the task (e.g., classification, time series prediction, anomaly detection). Deep learning models like DNNs and CNNs offer powerful pattern recognition but require significant data and computational power. Traditional machine learning models like SVM and decision trees are easier to implement and interpret but may not perform as well on highly complex data. Hybrid methods like PCA-K-means-BP and BiLSTM-CTC combine multiple approaches to improve performance in specific use cases.

This review has some limitations. These are as follows:

The use of different and non-standard data collection forms in studies reviewed, experiencing difficulties with interpretation of data arc welding processes.

The limitations of the study are as follows:

The lack of standardization in data collection methods across the reviewed studies has created challenges for comparative analysis. Some studies utilized custom datasets, but the lack of access to these datasets limits the generalizability of the findings.

The reviewed studies employed different types of data (images, audio signals, current/voltage data) and data sizes, making it difficult to compare techniques. Some studies focused solely on image data, while others relied on sensor data, leading to inconsistencies in the comparison of results.

While AI models have achieved high accuracy rates in controlled environments, their performance in real-world industrial settings is limited by factors such as noise, environmental variability, and rapidly changing conditions. Many challenges remain for real-time applications in these environments.

Most studies provided solutions tailored to specific problems using customized datasets. However, there is limited information on whether these solutions can be generalized to different welding processes, materials, or industrial environments.

Many studies relied on proprietary datasets created by researchers. The lack of publicly available datasets makes it difficult for other researchers to replicate experiments or compare results.

AI techniques, particularly deep learning models, require large datasets and significant computational power to achieve high accuracy. This can lead to delays and performance issues in certain applications. In welding processes that require real-time decision-making, hardware limitations may hinder the models' effectiveness.

These limitations affect the scope of the study and the generalizability of the findings.

6. Conclusions

This review paper has presented a comprehensive analysis of recent studies employing artificial intelligence techniques to prevent defects in Arc welding processes. The analysis covered three key aspects: datasets utilized, methodologies and approaches adopted, and performance metrics reported.

Several notable conclusions can be drawn from this study:

Most researchers have relied on their own custom datasets tailored to the specific welding defects or applications under investigation, given the unique challenges and solutions involved in each study.

For classification tasks, deep learning models like convolutional neural networks (CNNs) and long short-term memory (LSTM) networks have demonstrated superior performance compared to traditional machine learning algorithms, particularly when dealing with image or time-series data.

A diverse range of performance metrics have been employed, including accuracy, precision, recall, F1-score, mean squared error (MSE), and root mean squared error (RMSE), with the selection of metric contingent upon the specific task (classification or regression) and the desired trade-off between different performance aspects.

While many studies have reported promising results, with accuracy rates frequently exceeding 90%, there remains room for improvement, especially in real-world industrial settings where factors such as noise, occlusions, and rapidly changing welding conditions can pose significant challenges.

Overall, this review highlights the growing adoption of artificial intelligence techniques, particularly deep learning, in the domain of error prevention and quality control for Arc welding processes. As the field continues to evolve, further research is warranted to enhance the robustness, generalizability, and real-time applicability of these methods across.

Ethical statement

The authors declare that this document does not require ethics committee approval or any special permission. This review does not cause any harm to the environment and does not involve the use of animal or human subjects.

Acknowledgment

This review study was supported by TUBITAK 1711 Artificial Intelligence Ecosystem Call with the project number "3227006" under the name "Artificial Intelligence Supported Error Preventive and Predictive Smart Production System Development in Robotic MIG/MAG Welding Processes".

Conflict of interest

All authors have none to declare.

Authors' Contributions

All authors mentioned in the paper must have significantly contributed to the research. The contributions of the authors are listed below.

TTB: Conceptualization, Methodology, Formal analysis, Investigation
MSK: Formal analysis, Resources, Writing - Original draft preparation
AM: Formal analysis, Resources, Writing
MP: Resources, Investigation
EN: Investigation

References

[1] K. Weman and G. Lindén, *MIG welding guide*. Woodhead Publishing, 2006.

[2] J. Norrish, *Advanced Welding Processes*. Institute of physics Publishing, 1992.

[3] "What is MIG/MAG Welding?" [Online]. Available: https://www.fronius.com/en/welding-technology/world-of-welding/mig-mag-welding

[4] D. Young, "MIG Welding Transfer Methods - A.E.D. Motorsport Products," *A.E.D. Motorsport Products*. [Online]. Available: https://www.aedmetals.com/news/mig-welding-transfer-methods

[5] Miller, "MIG Welding: Setting the Correct Parameters," Miller.

[6] S. C. A. Alfaro and P. Drews, "Intelligent Systems for Welding Process Automation," *J. of the Braz. Soc. of Mech. Sci. & Eng.*, vol. XXVIII, no. 1, pp. 25–29, 2006.

[7] Unimig, "Troubleshooting Your Weld – The 12 Most Common Problems & How to Fix Them," Unimig.

[8] R. Singh, *Arc welding processes handbook*. John Wiley & Sons, 2021.

[9] A. B. Short, "Gas tungsten arc welding of α + β titanium alloys: A review," *Materials Science and Technology*, vol. 25, no. 3, pp. 309–324, Mar. 2009, doi: 10.1179/174328408X389463.

[10] K. Weman, *Welding processes handbook*. Elsevier, 2011.

[11] L. F. Jeffus, H. V Johnson, and A. Lesnewich, *Welding: principles and applications*. Delmar Publishers New York, 1999.

[12] P. Kah, R. Suoranta, and J. Martikainen, "Advanced gas metal arc welding processes," *The International Journal of Advanced Manufacturing Technology*, vol. 67, pp. 655–674, 2013.

[13] F. Khoshnaw, I. Krivtsun, and V. Korzhyk, "Arc welding methods," in *Welding of Metallic Materials*, Elsevier, 2023, pp. 37–71.

[14] T. Chen *et al.*, "Xgboost: extreme gradient boosting," *R package version 0.4-2*, vol. 1, no. 4, pp. 1–4, 2015.

[15] T. M. Mitchell and T. M. Mitchell, *Machine learning*, vol. 1, no. 9. McGraw-hill New York, 1997.

[16] L. Breiman, "Random forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001.

[17] B. Yegnanarayana, *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

[18] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans Syst Man Cybern*, vol. 23, no. 3, pp. 665–685, 1993.

[19] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.

[20] E. Eze and J. Eze, "Artificial Intelligence Support For 5g/6g-Enabled Internet of Vehicles Networks: An Overview." [Online]. Available: https://www.itu.int/en/journal/j-fet/Pages/default.aspx

[21] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," Mar. 2017, [Online]. Available: http://arxiv.org/abs/1703.06870

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03385

[23] L. R. Medsker and L. C. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, pp. 64–67, 2001.

[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997.

[25] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.

[26] Y. Zheng, Z. Xu, and X. Wang, "The fusion of deep learning and fuzzy systems: A state-of-the-art survey," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 8, pp. 2783–2799, 2021.

[27] K. Asif, L. Zhang, S. Derrible, J. E. Indacochea, D. Ozevin, and B. Ziebart, "Machine learning model to predict welding quality using air-coupled acoustic emission and weld inputs," *J Intell Manuf*, vol. 33, no. 3, pp. 881–895, Mar. 2022, doi: 10.1007/s10845-020-01667-x.

[28] G. Chen, B. Sheng, R. Luo, and P. Jia, "A parallel strategy for predicting the quality of welded joints in automotive bodies based on machine learning," *J Manuf Syst*, vol. 62, pp. 636–649, Jan. 2022, doi: 10.1016/j.jmsy.2022.01.011.

[29] S. Shin, C. Jin, J. Yu, and S. Rhee, "Real-time detection of weld defects for automated welding process base on deep neural network," *Metals (Basel)*, vol. 10, no. 3, Mar. 2020, doi: 10.3390/met10030389.

[30] L. Liu, H. Chen, and S. Chen, “Quality analysis of CMT lap welding based on welding electronic parameters and welding sound,” *J Manuf Process*, vol. 74, pp. 1–13, Feb. 2022, doi: 10.1016/j.jmapro.2021.11.055.

[31] K. Meyer and V. Mahalec, “Anomaly detection methods for infrequent failures in resistive steel welding,” *J Manuf Process*, vol. 75, pp. 497–513, Mar. 2022, doi: 10.1016/j.jmapro.2021.12.003.

[32] P. Jirapipattanaporn and W. Lawanont, “Development of Anomaly Detection Model for Welding Classification Using Arc Sound,” in *KST 2022 - 2022 14th International Conference on Knowledge and Smart Technology*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 57–62. doi: 10.1109/KST53302.2022.9729058.

[33] S. F. Laving, “Gas Metal Arc Welding Defect Detection using Sound signals,” Master Thesis, Seljuk University Institute of Science, Konya, 2019.

[34] D. Bacioiu, G. Melton, M. Papaelias, and R. Shaw, “Automated defect classification of SS304 TIG welding process using visible spectrum camera and machine learning,” *NDT and E International*, vol. 107, Oct. 2019, doi: 10.1016/j.ndteint.2019.102139.

[35] C. El Hachem, G. Perrot, L. Painvin, J. B. Ernst-Desmulier, and R. Couturier, “Welding Seam Classification in the Automotive Industry using Deep Learning Algorithms,” in *Proceedings - 2021 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, IAICT 2021*, Institute of Electrical and Electronics Engineers Inc., Jul. 2021, pp. 235–240. doi: 10.1109/IAICT52856.2021.9532569.

[36] H. Guo, L. Lin, Y. Lv, J. Liu, and C. Tong, “Machine Learning for Determining Key Parameters in Welding Process of Underground Engineering Equipment,” in *Proceedings of 2021 IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 33–41. doi: 10.1109/SDPC52933.2021.9563365.

[37] J. Lin, J. Lu, J. Xu, and D. Li, “Welding quality analysis and prediction based on deep learning,” in *Proceedings - 2021 4th World Conference on Mechanical Engineering and Intelligent Manufacturing, WCMEIM 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 173–177. doi: 10.1109/WCMEIM54377.2021.00045.

[38] X. Ma, S. Pan, Y. Li, C. Feng, and A. Wang, “Intelligent welding robot system based on deep learning,” *Proceedings - 2019 Chinese Automation Congress, CAC 2019*, pp. 2944–2949, 2019, doi: 10.1109/CAC48633.2019.8997310.

[39] X. Jin, L. Lv, C. Chen, F. Yang, and T. Chen, “A New Welding Seam Recognition Methodology Based on Deep Learning Model MRCNN,” in *2020 7th International Conference on Information, Cybernetics, and Computational Social Systems, ICCSS 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 767–771. doi: 10.1109/ICCSS52145.2020.9336927.

[40] C. El Hachem, “Automation of quality control and reduction of non-compliance using machine learning techniques at Faurecia Clean Mobility.” [Online]. Available: https://theses.hal.science/tel-03775370

[41] M. A. Kesse, “Artificial intelligence: A modern approach to increasing productivity and improving weld quality in TIG welding,” Doctoral Thesis, Lappeenranta-Lahti University of Technology LUT, Lappeenranta, 2021.

[42] E. A. Gyasi, P. Kah, H. Wu, and M. A. Kesse, “Modeling of an artificial intelligence system to predict structural integrity in robotic GMAW of UHSS fillet welded joints,” *International Journal of Advanced Manufacturing Technology*, vol. 93, no. 1–4, pp. 1139–1155, Oct. 2017, doi: 10.1007/s00170-017-0554-0.

[43] H. S. Nogay and T. C. Akinci, “Classification of operation cases in electric arc welding wachine by using deep convolutional neural networks,” *Neural Comput Appl*, vol. 33, no. 12, pp. 6657–6670, Jun. 2021, doi: 10.1007/s00521-020-05436-y.

[44] J. Wang *et al.*, “On-line defect recognition of MIG lap welding for stainless steel sheet based on weld image and CMT voltage: Feature fusion and attention weights visualization,” *J Manuf Process*, vol. 108, pp. 430–444, Dec. 2023, doi: 10.1016/j.jmapro.2023.10.081.

[45] A. El Houd, C. El Hachem, and L. Painvin, “Deep Learning Model Explainability for Inspection Accuracy Improvement in the Automotive Industry,” Oct. 2021, [Online]. Available: http://arxiv.org/abs/2110.03384

[46] B. S. G. Pernambuco, C. R. Steffens, J. R. Pereira, A. V. Werhli, R. Z. Azzolin, and E. Da Silva Diaz Estrada, “Online sound based arc-welding defect detection using artificial neural networks,” in *Proceedings - 2019 Latin American Robotics Symposium, 2019 Brazilian Symposium on Robotics and 2019 Workshop on Robotics in Education, LARS/SBR/WRE 2019*, Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 263–268. doi: 10.1109/LARS-SBR-WRE48964.2019.00053.

[47] R. Wang, H. Wang, Z. He, J. Zhu, and H. Zuo, “WeldNet: a lightweight deep learning model for welding defect recognition,” *Welding in the World*, 2024, doi: 10.1007/s40194-024-01759-9.

[48] S. Li, J. Gao, E. Zhou, Q. Pan, and X. Wang, “Deep learning-based fusion hole state recognition and width extraction for thin plate TIG welding,” *Welding in the World*, vol. 66, no. 7, pp. 1329–1347, Jul. 2022, doi: 10.1007/s40194-022-01287-4.

[49] Y. Wang, J. Han, J. Lu, L. Bai, and Z. Zhao, “TIG stainless steel molten pool contour detection and weld width prediction based on Res-Seg,” *Metals (Basel)*, vol. 10, no. 11, pp. 1–15, Nov. 2020, doi: 10.3390/met10111495.

[50] X. Zhang, S. Zhao, and M. Wang, “Deep Learning-Based Defects Detection in Keyhole TIG Welding with Enhanced Vision,” *Materials*, vol. 17, no. 15, Aug. 2024, doi: 10.3390/ma17153871.

[51] C. Xia, Z. Pan, Z. Fei, S. Zhang, and H. Li, “Vision based defects detection for Keyhole TIG welding using deep learning T with visual explanation,” *J Manuf Process*, vol. 56, pp. 845–855, Aug. 2020, doi: 10.1016/j.jmapro.2020.05.033.

[52] Y. Liu, Y. Zhou, S. Wen, and C. Tang, “A strategy on selecting performance metrics for classifier evaluation,” *International Journal of Mobile Computing and Multimedia Communications (IJMCMC)*, vol. 6, no. 4, pp. 20–35, 2014.