Dicle University
Journal of Engineering
https://dergipark.org.tr/tr/pub/dumf
duje.dicle.edu.tr

Research Article

Advancing Welding Quality through Intelligent TIG Welding: A Hybrid Deep Learning Approach for Defect Detection and Quality Monitoring

Emre GORGUN[1*]

[1] Sivas Cumhuriyet University, *Sivas Technical Sciences Vocational School, Railway Systems Department*, emregorgun@cumhuriyet.edu.tr, Orcid No: 0000-0002-1971-456X

ARTICLE INFO

Article history:

Received 19 February 2025
Received in revised form 13 March 2025
Accepted 20 March 2025
Available online 30 September 2025

Keywords:

Artificial Intelligence, Convolutional Neural Networks, Deep Learning, Tungsten Inert Gas, Welding)

Doi: 10.24012/dumf.1642978

* Corresponding author

ABSTRACT

Modern welding procedures are intricate, requiring a variety of variables and occasionally lacking a complete understanding of their underlying mechanics. Despite the adoption of intelligent welding processes in a few applications, there are still several obstacles. By combining advanced search, combinatorial optimisation, geometric reasoning techniques, and comprehensive Artificial Intelligence (AI) modelling cognitive capabilities, the proposed research aims to build intelligent welding. The three main scientific foci of the research are feature correlation to forecast process performance and facilitate corrective actions, feature extraction utilising intense signal analysis, and the use of simulated or supplied data for analysis. Previous research led to the development of an intelligent Tungsten Inert Gas (TIG) welding platform for materials made of aluminium. On the other hand, TIG welding is susceptible to fluctuations in the root gap, which affect the quality of the weld and could result in electrode contamination. Common welding errors include excessive heat-affected zone width, fusion width, bead height, and inadequate penetration. These errors directly affect the strength and load-bearing capacity of the joint while also making it more susceptible to stress and fracture propagation. The proposed AI-powered welding tool is made to overcome common weld imperfections. Therefore, the research's objective is to develop a hybrid deep learning-powered platform for TIG welding. Convolutional Neural Networks (CNNs) will be employed to extract discrete visual characteristics linked to each type of weld defect, establish correlations between these features, and give weld images to identify the types of defects or their absence. The objective of the research is to create a neural network model that can determine whether a given weld image is good or bad due to contamination, burn-through, or lack of fusion. These findings will make precise weld quality monitoring and process improvement possible.

Introduction

Tungsten Inert Gas (TIG) welding is a crucial process in the industrial sector that calls for qualified and experienced welders, which entails high costs and potential health concerns [1]. However, with the introduction of big data and Industry 4.0, efforts have been made to discover weld flaws in several other welding procedures in addition to TIG welding [2]. Therefore, this research has the potential to ease problems related to using human labour in the TIG welding process if weld automation when successfully implemented. The ability of a worker to make intelligent decisions based on cursory investigations is crucial for welding. Making these choices is essential for figuring out and altering process parameters to meet requirements. The choice of these parameters directly affects product quality, underscoring the significance of this decision.

However, the introduction of machine learning algorithms that can mimic human cognitive capacities offers a potential substitute by reproducing abilities like learning, image classification, and feature identification. The ResNet-50 and ResNet-18 networks are examined and compared in this research as they have shown potential in the past such as [3], [4], [5]. These neural networks are trained to correlate characteristics from a verification dataset after extracting them from a training dataset [6]. The objective is to categorise a given weld image as either a good weld or a particular type of defective weld, taking into account common TIG weld faults [7] [8] [9]. Therefore, the main goal is to create a neural network that can categorise post-weld images of TIG welds into different groups, such as weld quality, burn-through, contamination, lack of fusion, misalignment, and lack of penetration. However, data accessibility determines this research's scope. The TIG welding process on 5083 aluminium will be represented in the ResNet model exclusively as post-weld images taken with High Dynamic Range (HDR) camera. It's crucial to remember that numerous weld flaws can exist in a single piece in real-world applications. However, occurrences containing many flaws in a single sample will not be taken into consideration in this research due to the limits of the training data [7]. Therefore, the main contributions of this paper are as follows:

Development of an Intelligent TIG Welding Platform

The research contributes to the field of welding technology by developing an intelligent TIG welding platform. This platform is designed to enhance the welding process for materials made of aluminum, which is known to be particularly challenging due to its susceptibility to common welding errors [9] [10] [11]. The development of this platform showcases the application of advanced search, combinatorial optimization, geometric reasoning techniques, and comprehensive AI modeling cognitive capabilities to address real-world welding challenges.

Hybrid Deep Learning-Powered Welding Tool

The research introduces a novel approach by leveraging CNNs to extract discrete visual characteristics associated with various types of weld defects. By establishing correlations between these features and using weld images as input, a hybrid deep learning-powered platform is developed. This platform has the ability to identify common weld imperfections such as contamination, burn-through, or lack of fusion. This contribution holds significant promise for precise weld quality monitoring and process improvement, ultimately leading to enhanced weld quality and reliability.

Advancements in Welding Defect Detection

The research significantly advances the capabilities of defect detection in welding processes [12]. By training a neural network model to analyze weld images and determine whether a given weld is of good or bad quality, the study introduces a new level of automation and accuracy to the inspection and assessment of welds. This contribution not only enables real-time quality control but also facilitates rapid corrective actions in response to welding errors, thereby improving the overall efficiency and reliability of welding operations in various industries.

These contributions collectively represent a substantial step forward in the integration of AI and deep learning techniques into the welding domain, addressing critical issues related to weld quality and performance prediction while paving the way for more intelligent and efficient welding processes.

Background

The common welding method known as TIG creates welds of excellent quality. This process employs a non-consumable tungsten electrode, Argon shielding gas, and an electric arc as the heat source to melt the weld pool. The accuracy and superior quality of the welds produced are enhanced by the intense and small size of the arc. The quality of the welded connection is significantly influenced by the operator's control of crucial welding parameters such as arc current, arc voltage, travel speed, and gas flow rate [8]. Numerous weld faults may develop when any of these variables depart from their ideal ranges, endangering the connection's structural integrity. One such issue that develops when there is insufficient penetration into the base metal is “burn-through”. This issue is frequently brought on by an excessively high arc current or voltage and/or a slow enough travel speed, which produces too much extra heat in comparison to what is needed for melting and fusing. A key factor in understanding this phenomenon is the specific heat input, which may be calculated using Equation (1) [9]. These factors highlight the complexity and importance of precise control over welding settings in preserving the quality and integrity of TIG welding.

Specific Heat Input =(Voltage * Amperage)/(Travel speed*mass) (1)

In welding, “contamination” is the presence of foreign particles in or close to the weld zone. On the surface of base metals, these particles—which can be anything from dust to oil to metal shavings—are frequently observed. Contamination can be efficiently reduced during pre-welding procedures by wiping, chemical cleaning, or mechanical cleaning techniques. Whereas, “lack of fusion” in welding describes an inadequate bond between the weld and the base metals. This occurrence is commonly attributed to insufficient heat input, which happens when there is not enough energy given to allow for proper melting and fusing. When welding, low voltage or current, as well as an incredibly fast travel speed, can all result in insufficient heat input. The link between these parameters is seen in Equation (1). Fusion failure can also happen in circumstances when the surface isn't properly prepared before welding. When two base metals are not precisely aligned in their appropriate dimensions (“misalignment”), the necessary degree of interaction between them cannot be maintained. However, when the welding process fails to fill the target connection, this is known as a “lack of penetration”. This results from improper management or inadequate control of welding parameters including arc current, voltage, or travel speed. As a result, the joint is weaker than necessary, raising the possibility of structural issues.

Related Works

The goal of this research is to integrate AI into welding systems. Although numerous studies examine AI and welding sciences individually such as [13], [14], [15], there is a noteworthy dearth of research on the development of intelligent welding systems. Robots have achieved significant gains in several production processes [18], [19], including machining [16], [17], thanks to the quick development of intelligent systems. The welding industry, particularly in the field of process automation, has been hesitant to adopt new technological advancements. Critical TIG welding process variables including current, voltage, welding speed, and gas flow rate directly affect the shape of the weld bead and, consequently, the joint's Ultimate Tensile Strength (UTS). The UTS of AISI 4340 low alloy steel underwent a considerable 57MPa change as a result of changing production conditions [12], [20]. Even though there has been advancement in automating process parameter determination, it is still limited to online programming and large batch sizes. Robotic welding arms' actuation and control systems have been shown to work such as[21], [22], [23], but sophisticated sensing and decision-making systems still need to be improved.

The ability of AI to make decisions is still significantly lacking, despite attempts to replicate human brain cognitive processes. Humans are better than machines at making the appropriate decisions at the right time while exhibiting ingenuity and resourcefulness in problem-solving [24], [25]. The strength of AI, on the other hand, is in its capacity to perform lower-level human judgements with outstanding speed, efficiency, and cost-effectiveness. Each of the divisions of AI has its collection of algorithms created to solve particular problem sets. Notably, GoogleNet and ResNet-50 have proven effective at classifying images, identifying objects, recognising faces, and classifying objects [26]. Layers may be disregarded if they do not bring value to the network thanks to ResNet-50's addition of residual blocks. Utilising inception modules, GoogleNet's convolutional operations are affordable and free from the risk of overfitting. Despite the dearth of research on AI-based welding systems, significant advancements have been made. Notably, this research gathers visual data with an emphasis on either weld feature extraction or weld problem identification for intelligent systems to handle. A weld geometric feature monitoring system is developed in [27] to assist welders in selecting the best welding settings by calculating the bead width that will be produced based on input data. Another study i.e. [28] predicts welding parameters based on desired weld penetration and vice versa, showing prediction accuracies reached with different approaches. Another critical step in the welding process is non-destructive testing for weld defect detection and image-based defect sensing devices that have been proposed and put into use by [29]. For instance, researchers employed ResNet with 18 convolutional layers for image classification to accurately identify a variety of weld flaws [30]. Another study in the same spirit emphasises the potential benefit of CNNs for image classification [31].

Emiprical study

Dataset

The enormous amount of data required for successfully training a CNN is a key factor in this endeavour. The difficulty, though, is that substantial testing and data collection are difficult due to the inherent limitations of the available internal resources. To address this, it was determined that the necessary data would be obtained from outside sources, mainly online archives, ensuring a large and varied dataset suitable for training and validation. The requirement for a dataset that precisely replicates the wide range of real-world welding circumstances, a prerequisite for training and verifying the CNN model, led to the choice to obtain data externally. The goal was to gather a dataset that would include various welding circumstances and faults, which made access to a variety of images necessary. Leveraging other sources became a practical strategy to alleviate resource shortages given the enormous volume required. Given its depth and clear connection with the research's goals, the TIG Aluminium 5083 dataset from Kaggle was specifically chosen as the fundamental corpus for training the CNN. This effort was greatly helped by the dataset, a treasure trove of 33,254 images painstakingly categorised into separate classifications. The collection is intelligently divided into several groups that correspond to important welding circumstances and flaws. The following subcategories are included in this list: "Good Weld", "Burn Through", "Contamination", "Lack of Fusion", "Misalignment", and "Lack of Penetration". This dataset's extensiveness offered a solid foundation for training the CNN and attaining the research's objectives. The provenance of the dataset adds to its legitimacy. The Department of Metallurgy and Materials at the University of Birmingham in the UK and TWI Ltd in Abington, Cambridge, were the organisations responsible for taking the images in this collection. These organisations captured images using cutting-edge technology, particularly HDR cameras. The dataset's diversity and quality were greatly increased by the HDR cameras' high-fidelity capture of welding settings.

The dataset is then organised into designated subfolders, which is an essential step for speeding up data processing. Then, image augmentation is carried out to increase dataset diversity and efficiency. The ResNet-50 and ResNet-18 models are then imported and adjusted as necessary to meet the research's specifications. The main stage in the process is to train these modified ResNet models using the collected TIG Aluminium 5083 dataset, which is essential for the models. The models' ability to recognise welding conditions and flaws accurately depends on this training step. Following training, the effectiveness of the models in identifying and categorising welding instances is evaluated in detail by utilising confusion matrices and Grad-CAM to examine network performance. With the use of these analytical tools, models may be thoroughly evaluated, directing future model improvements and optimisations for better performance.

Dataset Preprocessing

As they pass through several convolutional layers, filters, and the neural network in this research, the images undergo significant processing. Red, Green, or Blue (RGB) channels correspond to the dimensions of each image's three-dimensional array [13]. Each component of a dimension corresponds to a pixel in the image and has a value between 0 and 255 that represents the brightness of the corresponding pixel. Neural network training is a computationally demanding process that uses a lot of memory, especially on the GPU. As a result, the processing time varies greatly depending on the image resolution. Image sizes are minimised to the greatest extent possible to speed up processing and lower computational load. The images in the dataset are initially cropped from their original size of 1280*1024 pixels to 800*974 pixels to remove extraneous black pixels. The images are then changed to grayscale, which further shrinks them to an array of 800*974*1 pixels each. The dataset is, however, extended and altered to meet this criterion for compatibility with the ResNet-50 model, which analyses RGB images with 224 pixels in height and 224 pixels in width. For the network to perform better with unseen images, gained by augmentation, the dataset must be diverse. Each image is given a 50% probability of being either vertically mirrored or rotated by 20 degrees in either a clockwise or anticlockwise direction during the augmentation process. The data must be arranged in a specified folder structure with a subfolder for each class of image to enable effective training with the MATLAB network training tool. The initial dataset, though, did not follow this structure. A Python script was used to properly reorganise the dataset. By comparing image names to a JSON file that comprised image names and their corresponding classifications, this script iterated over each file in the folder, transferred images to the appropriate subfolder, and assured accurate labelling for supervised learning. By carefully processing the data, it is ensured that the dataset is formatted and enhanced to fulfil the needs for effective training and validation of the AI-based welding system.

Model architecture

The layer count and its ramifications are the most important factors to take into account when choosing a suitable CNN architecture. In this decision-making process, we place special emphasis on the convolutional layers because they extract features from the input images. In the TIG dataset, these layers are crucial for identifying the digital properties connected to each unique weld class. The CNN's earliest layers are tasked with identifying basic features including the borders of the weld zone and colour gradients coming from the Heat-Affected Zone (HAZ). The network can distinguish increasingly complicated traits as the information moves through succeeding layers. It has been demonstrated that an architecture with more layers can find complex and sophisticated features in the dataset. While improving feature extraction capabilities, this increase in layer count does present some questions. When dealing with deep architectures that have a high layer count, the overfitting issue must be taken into account. When a model is overturned to the complexities of the training data, it captures highly specific features, even subtleties like scratches on the base metal used in the training dataset. As a result, the model predicts the training data with remarkable accuracy but fails when attempting to predict the unknown data. Finding the right number of layers to avoid overfitting and guaranteeing the network's generalizability presents a complex task in this case. The risk of overfitting in large networks reduces the model's capacity to generalise effectively to new data.

Finding the ideal balance between the number of layers is a difficult undertaking that calls for a practical strategy incorporating testing. To find the ideal configuration that strikes the ideal balance between complexity and generalisation ability, it is essential to experiment and assess the model's performance with a range of layer counts. Given this, we incorporate the ResNet-18 architecture into our experimental framework to compare its performance to that of the ResNet-50 model. This methodical comparison will show how an 18-layer deep network compares to its 50-layer counterpart in terms of effectiveness and efficiency for our particular goals. As we move from the domain of network parameters to the domain of training the ResNet-50 model on the dataset, it becomes clear that careful changes and fine-tuning are required. These variables have a significant impact on the network's learning dynamics, successfully guiding it in the direction of optimal convergence and reliable predictions. The output layer's number of neurons is a crucial parameter to configure, to start. The model's ability to accurately describe the classes found in the dataset depends on this parameter.

To guarantee that the model can recognise the varied welding circumstances and faults in the dataset, the proper balance must be struck here. The learning rate, which affects how big of a step to take throughout the optimisation process, is another key component. The network's capacity to settle into an ideal solution and the speed at which it converges are both greatly impacted by the learning rate. To promote effective training and avoid problems like overshooting or slow convergence, a well-calibrated learning rate is essential. Another important component of the network parameters is momentum, which can be written as a constant. During optimisation, it affects how the weights are updated. Smoother convergence and escape from local minima can result from proper tuning of the momentum constant. Another crucial setting to choose is the maximum number of iterations, which determines how many times the algorithm will run through the full dataset during training. This parameter affects how long the model is trained for and how much of the dataset it is exposed to as a result. The bias value, which affects the model's adaptability and capacity to precisely capture the underlying patterns in the dataset, is the final crucial parameter to adjust. Table I lists the hyperparameter that has been used.

Table 1. Hyperparameters value.

Hyperparameter	Value
Momentum	0.9
Initial learning rate	0.001
Learning rate drop factor	0.1
Learning rate drop period	10
Gradient Threshold method	'12norm'
Gradient threshold	Inf
Maximum epochs	30
Mini batch size	32
Verbose	1
Verbose frequency	50
Validation frequency	50
Sequence padding direction	'right'
Sequence padding value	0
Batch normalisation statistics	'Population'

Experiments

Fig. 1 shows a large dataset sample and highlights the notable differences in appearance between images within each class. It should be noted that the "contamination" class exhibits substantial variability, making it difficult to pinpoint a typical image for feature extraction. As a result, an individual technique is required, in which unique features are extracted from each image within a class. This guarantees that each class is accurately characterised by the CNN by capturing all pertinent information related to that class. The goal of network training is to create a network capable of using convolutional layers to extract distinctive properties for each class. The CNN should then compare these characteristics to any given image and determine which class it belongs to. Although deeper networks are capable of extracting more complex characteristics, they run a higher risk of overfitting the training set.

The performance of the ResNet-50 model is contrasted with that of the shallower ResNet-18 model to establish the ideal depth of the network. Both models go through training on 1000 images from each class, followed by validation on 200 images from each class. The ResNet-50 model was trained for 257 minutes and 30 seconds, with a validation accuracy of 99.75% and a validation loss of 0.0144 at the end. While the ResNet-18 model took 106 minutes and 49 seconds to train, it finished with a validation accuracy of 99.67% and a loss of 0.0081. ResNet-18 was more effective in terms of training time and loss value, despite the ResNet-50 model showing a 0.8% greater prediction accuracy. Because of its 18 layers and quicker training time, the ResNet-18 model is recommended for training. It also has a lower loss value. It is clear from examining Figs. 2 and 3's training progression that both prediction accuracy and loss have peaked. The training procedure should have been finished after 1000 iterations, but it took 5610 instead. Both networks are ready to classify the test dataset after the training is finished, and a variety of analytical methods are used to assess their performance. Figs. 4 and 5 shows that ResNet-50 almost precisely predicts the classes of images in the test dataset, except for one misclassification in which an image tagged as "good weld" is wrongly predicted as "lack of penetration". Despite the excellent forecast accuracy, this misclassification calls for further inquiry.

The Grad-CAM algorithm is applied to investigate the prediction process of the network and maybe to comprehend the misclassification. Fig. 6 shows the Grad-CAM visualisation for each class of images. There aren't many distinguishing characteristics in the images from the "good weld" and "lack of fusion" classes, which could confuse these classes. Grad-CAM heat maps show that the network is highlighting features on the left and right sides instead of the "good weld" class, which is an improper focus for the network. Other classes like "lack of fusion" and "misalignment" also show this pattern. Instead of concentrating on the basic weld aspects that are essential for classification, the majority of images from each class come from the same experiment and share some experiment-specific data that the model employs for identification. As a result, the model is outstanding at categorising images from this particular dataset but less effective for more general industrial applications. This shows overfitting to the training data. Surprisingly, the ResNet-18 model's prediction accuracy with the test dataset matches that of ResNet-50 with only one misclassification: an image tagged "contamination" was projected mistakenly to be "lack of penetration". The Grad-CAM visualisations for each image class using ResNet-18 are shown in Fig. 6. It becomes clear that due to their visual resemblance, classifications like "contamination" and "lack of fusion" could be mistaken for one another. According to the Grad-CAM explanation, the network concentrates on comparable areas of the images (upper right corner), which have nothing to do with the weld quality itself. The ResNet-18 model also demonstrates a propensity to focus on characteristics unrelated to weld quality, as observed in the ResNet-50 model. This indicates that overfitting is still an issue in both 50-layer and 18-layer deep CNNs. The models, ResNet-50 and ResNet-18, exhibit great prediction accuracy but also show evidence of overfitting, making them most useful on this particular dataset.

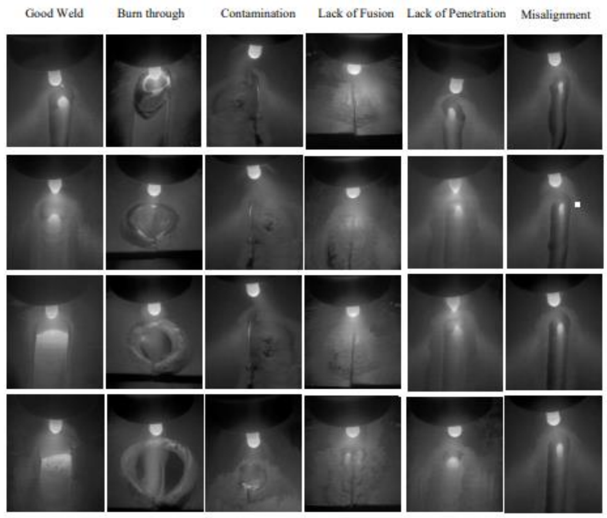

Figure 1. Dataset sample. [15, 18]

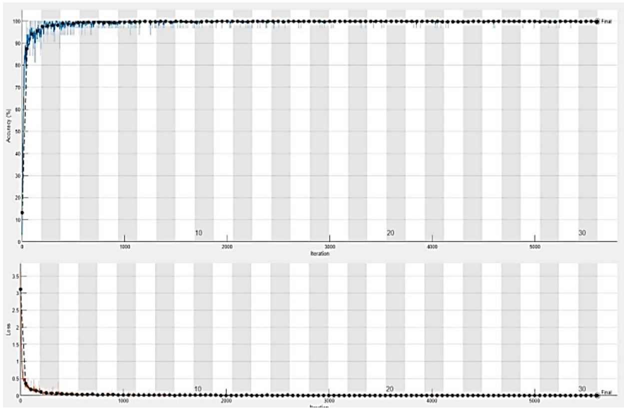

Figure 2. ResNet-18 accuracy vs loss.

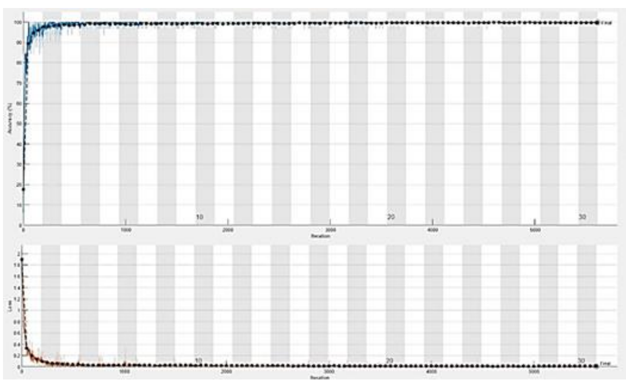

Figure 3. ResNet-50 accuracy vs loss.

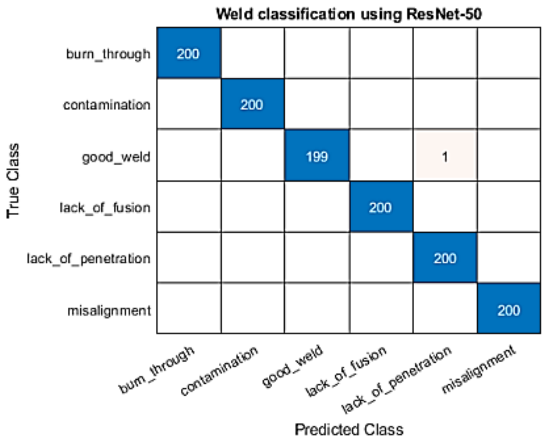

Figure 4. Confusion matrix chart- ResNet-50.

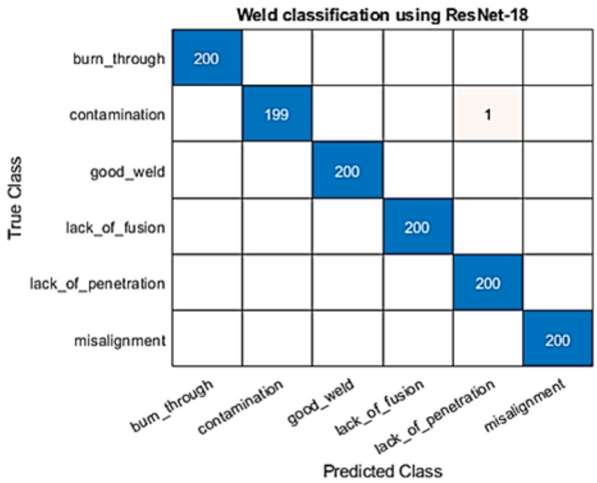

Figure 5. Confusion matrix chart- ResNet-18.

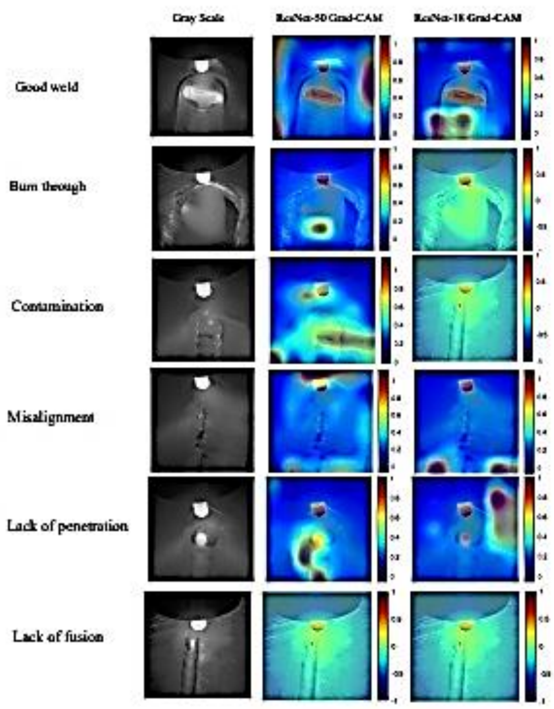

Figure 6. Grad-CAM ResNet-50 and ResNet-18. [15, 18]

Result analysis

The effectiveness of both models in extracting distinctive properties for each class is evaluated through a thorough examination. A comparison research is also done to see how these features may be defined digitally manually as opposed to automatically using a neural network. On a scale of 0 to 1, the heat map serves as a crucial tool for highlighting the importance of the features that CNN used to digitally designate each class. A score of 1 denotes the most prominence, whereas a score of 0 denotes the least significance in terms of feature representation.

Good weld

An important finding from the heat map is that the ResNet-50 model's emphasis is too narrow. It shows that the ResNet-50 model is highlighting aspects of the image that have nothing to do with the weldment. These components may include base metal-specific patterns or gradients that the deeper network wrongly classified as defining characteristics of the "good weld" class. On the other hand, the heat map's prominent focus on the weldment area shows that the ResNet-18 model has acquired characteristics associated with a good weld. In real-world terms, this means that the ResNet-50 model is expected to produce extremely precise predictions only for the photos in this dataset. In contrast, the ResNet-18 model is anticipated to produce precise predictions for both images within and outside of this dataset. If we were to describe these features digitally, we would use a 5*5 filter to flatten colour gradients by averaging the pixel values. The image would therefore be labelled as having a "good weld" if there is a rectangular area of pixels that is lighter (pixels with lower values) between two darker (base metal-representing) areas.

Burn through

A noteworthy finding is highlighted by the heat map analysis of the burn-through class: neither the ResNet-50 model nor the ResNet-18 model was able to successfully learn the characteristics of this class. The heat map shows that contrary to what was initially thought, the ResNet-50 model emphasises a specific area at the burn-through's middle rather than its outskirts. On the other hand, the ResNet-18 model did not learn any specific aspects of the image, as seen by the green mask over its heat map (green corresponds to a value of 0). Although the models learned false features for this class, their burn-through prediction accuracy topped 98%. This implies that the models perform exceptionally well in predicting images of this class inside this dataset but may fail to predict images of this class precisely from additional datasets. If this class were to be defined digitally, the procedure would entail averaging the pixel values with a 5*5 filter to balance out the colour gradients. A darker rectangular section of pixels (pixels with higher values) of the same value is present between two lighter portions (indicating the base metal), and the image is afterwards categorised as "burn through" if this is the case.

Contamination

The green mask on the heat map shows that, like the "burn through" class, the ResNet-18 model has trouble distinguishing the distinctive characteristics of the "contamination" class. On the other hand, the heat map for the ResNet-50 model provides information: the light red zones covering the weld area show that specific areas of the actual weld are receiving focus. However, a dark red area in the heat map's corner implies that the deeper model is picking up elements that are both connected to and unconnected from the actual weld. This suggests that the ResNet-50 model might outperform the shallower 18-layer model in terms of image prediction from outside the dataset. Identifying foreign particles based on their pixel values may be necessary to define this class. Contamination is when a group of pixels considerably differs (by more than 50%) from the pixels around it. Additionally, if more than 20% of the image's pixels comprise foreign particles, the image may be labelled as "contaminated".

Misalignment

The heatmaps unmistakably show that both models are emphasising the bottom half of the image, which has little to do with the weld's specifics. Surprisingly, this is the only situation in which the shallower and deeper networks concentrate on the same image regions. Despite both models' remarkable prediction accuracy within this dataset, they are unlikely to correctly categorise images belonging to this class outside of it. A procedure utilising pixel value averaging with a 2*2 filter and finding vertical lines of higher pixel values might be used to manually establish this image class. The image might be correctly categorised as "misalignment" if it has several misaligned lines running vertically across the middle of it.

Lack of penetration

The deeper design of the ResNet-50 model demonstrates that it is more effective at extracting the distinguishing characteristics of this particular class. The heatmap demonstrates that although the 18-layer CNN tends to emphasise information around the weld, the 50-layer CNN nearly entirely concentrates on the fine weld details. This distinction strongly implies that the ResNet-50 model is likely to show a better prediction accuracy for this particular image class than it did on the training dataset. A method that uses a 5*5 filter size to average all pixel values can be used for digital classification. The core of the image, where the weld bead should be, should show a gradient shift; if it doesn't, the image can be correctly categorised as having a "lack of penetration".

Lack of Fusion

Both models fail to accurately capture particular characteristics of the "lack of fusion" class, as seen by the green mask overlaid on the images. Therefore, it seems that CNN's performance in this situation may not be primarily influenced by the depth of the network. To ensure that both networks can correctly categorise images of this class outside the boundaries of this dataset, changes to other CNN parameters are likely needed. In a digital setting, defining the characteristics of this class requires a vertical scan through the image. If a vertical matrix of pixels in the centre of the image is discovered to be 50% darker than the surrounding pixels, the weld can be safely categorised as having a "lack of fusion".

Discussion

Carefully separating the dataset into training and validation images is a crucial step in starting the model training and evaluation phase. This division is necessary to ensure a complete assessment of the model's performance. Additionally, it is crucial to estimate the volume of data used for training with caution; given the size of the TIG Aluminium 5083 dataset. When the model is too suited to the training data, overfitting can lead to overfitting, which weakens the model's generalisation skills. During the training phase, the network is intensively trained and features are retrieved using the designated training images. The performance of the model is then scrutinised using the validation images. This evaluation consists of guessing the class to which each validation image belongs and contrasting those predictions with the actual labels attached to each image. Such a comparison research offers valuable insight into the accuracy and effectiveness of the model. The training and validation data are meticulously plotted in various ways to thoroughly examine the network's performance as shown in Section 6. Critical performance measures are represented graphically, which aids in understanding the behaviour of the model. This thorough analysis points up potential areas for enhancement and optimisation, opening the door to well-informed decisions on how to improve the model's performance and accuracy in locating and classifying welding errors in the dataset. To achieve the goals of using AI for better welding defect identification and categorization, this ongoing process of training, evaluating, and refining is essential. The prediction accuracy is evaluated graphically and quantitatively (as a percentage) to give a thorough analysis of the neural network's performance as shown in Section 6. This comprehensive graphic allows for a quick assessment of incorrect classifications and offers important insights into both correct and incorrect image classifications. If a particular class is regularly misclassified, a thorough investigation is necessary to ascertain the underlying causes. However, it takes a lot of work to manually go through the output of each convolutional layer and find the crucial traits that the model thought were important. In this case, automated algorithms save the day by providing a clear and comprehensive explanation of the model's decision-making process. One such incredible algorithm is the directed Gradient-Class Activation Mapping (Grad-CAM). Grad-CAM generates a heat map that successfully highlights the portions of the image that the network considers crucial for prediction. By shedding light on the model's inner workings, assisting in the identification of crucial components, and permitting potential model alterations, this representation enhances understanding. Grad-CAM is employed in this way to add a crucial layer of interpretability and transparency to the model, improving the evaluation procedure and ultimately increasing the model's expected accuracy and efficiency in identifying and categorising welding defects in the dataset.

Conclusion and future works

To create the framework for the later development of a specific AI deep learning system, this research started with an examination of AI applications in the welding industry. The major objective was to develop an algorithm capable of correctly determining if a weld image represented a "good weld", "burn through", "contamination", "lack of fusion", "lack of penetration", or "misalignment". The successful installation and alterations of the ResNet-50 and ResNet-18 CNN models were among this research's accomplishments. The acquisition and improvement of a TIG on the Aluminium 5083 dataset, which was necessary for effectively training the models, was a significant component. Through the training process, both models showed exceptional prediction accuracy, reaching 98%. Despite the excellent forecast accuracy, a further investigation utilising the Grad-CAM approach identified a significant issue. It was shown that, in most situations, neither model was able to adequately extract the critical characteristics that distinguished each type of weld. More study is needed to address the risk of overfitting, which was cited as the cause of this disadvantage. Potential methods for reducing overfitting include using a more varied dataset to expose the models to a wider variety of features, modifying the learning rate to encourage efficient convergence during training, and investigating different CNN architectures that might enhance feature extraction and generalisation. Future iterations of this AI deep learning algorithm are anticipated to produce predictions that are even more accurate by resolving these shortcomings, and as a result, will be more widely used in the sector.

References

[1] A. W. Fande, R. V. Taiwade, and L. Raut, ‘Development of activated tungsten inert gas welding and its current status: A review’, Jun. 11, 2022, *Taylor & Francis*. doi: 10.1080/10426914.2022.2039695.

[2] E. A. Gyasi, H. Handroos, and P. Kah, ‘Survey on artificial intelligence (AI) applied in welding: A future scenario of the influence of AI on technological, economic, educational and social changes’, in *Procedia Manufacturing*, Elsevier, Jan. 2019, pp. 702–714. doi: 10.1016/j.promfg.2020.01.095.

[3] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, ‘Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer’, in *Procedia Computer Science*, Elsevier, Jan. 2021, pp. 423–431. doi: 10.1016/j.procs.2021.01.025.

[4] A. Sirco, A. Almisreb, N. M. Tahir, and J. Bakri, ‘Liver Tumour Segmentation based on ResNet Technique’, in *ICCSCE 2022 - Proceedings: 2022 12th IEEE International Conference on Control System, Computing and Engineering*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 203–208. doi: 10.1109/ICCSCE54767.2022.9935636.

[5] N. Saleem, J. Gao, M. Irfan, E. Verdu, and J. P. Fuente, ‘E2E-V2SResNet: Deep residual convolutional neural networks for end-to-end video driven speech synthesis’, *Image and Vision Computing*, vol. 119, p. 104389, Mar. 2022, doi: 10.1016/j.imavis.2022.104389.

[6] M. H. R. Sobuz, M. K. I. Kabbo, T. S. Alahmari, J. Ashraf, E. Gorgun, and M. M. H. Khan, *‘Microstructural behavior and explainable machine learning aided mechanical strength prediction and optimization of recycled glass-based solid waste concrete’*, Case Studies in Construction Materials, p. e04305, 2025.

[7] A. Mayr, M. Weigelt, M. Masuch, M. Meiners, F. Hüttel, and J. Franke, ‘Application Scenarios of Artificial Intelligence in Electric Drives Production’, in *Procedia Manufacturing*, Elsevier, Jan. 2018, pp. 40–47. doi: 10.1016/j.promfg.2018.06.006.

[8] E. Gorgun, *‘Numerical analysis of inflow turbulence intensity impact on the stress and fatigue life of vertical axis hydrokinetic turbine’*, Physics of Fluids, vol. 36, no. 1, 2024, Accessed: Mar. 06, 2024. [Online]. Available: https://pubs.aip.org/aip/pof/article/36/1/015111/2932752.

[9] E. Görgün, *‘Çoklu Yükleme Koşulları Altında Yük Vagonu Şasisinin Topoloji Optimizasyonu’*, Karadeniz Fen Bilimleri Dergisi, vol. 12, no. 2, pp. 593–604, Dec. 2022, doi: 10.31466/kfbd.1078425.

[10] E. Görgün, *‘Ultrasonik Muayene Prob Çaplarının Darbe Yankı Değerine Etkisinin Araştırılması’*, Karadeniz Fen Bilimleri Dergisi, vol. 12, no. 1, pp. 381–389, 2022.

[11] E. Görgün, *‘Investigation of The Effect of SMAW Parameters On Properties of AH36 Joints and The Chemical Composition of Seawater’*, International Journal of Innovative Engineering Applications, vol. 8, no. 1, pp. 28–36, 2024.

[12] E. Gorgun, *‘Ultrasonic testing and surface conditioning techniques for enhanced thermoplastic adhesive bonds’*, J Mech Sci Technol, vol. 38, no. 3, pp. 1227–1236, Mar. 2024, doi: 10.1007/s12206-024-0218-6.

[13] E. Gorgun, A. Ali, and Md. S. Islam, *‘Biocomposites of Poly(Lactic Acid) and Microcrystalline Cellulose: Influence of the Coupling Agent on Thermomechanical and Absorption Characteristics’*, ACS Omega, vol. 9, no. 10, pp. 11523–11533, Mar. 2024, doi: 10.1021/acsomega.3c08448..

[14] D. Vijayan and V. Seshagiri Rao, ‘Process Parameter Optimization in TIG Welding of AISI 4340 Low Alloy Steel Welds by Genetic Algorithm’, in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, Jul. 2018, p. 012066. doi: 10.1088/1757-899X/390/1/012066.

[15] TWI, ‘What is Tungsten Inert Gas (GTAW or TIG) Welding?’, Job Knowledge 6. Accessed: Sep. 28, 2023. [Online]. Available: https://www.twi-global.com/technical-knowledge/job-knowledge/tungsten-inert-gas-tig-or-gta-welding-006

[16] Z. Abbasi *et al.*, 'The Detection of Burn-Through Weld Defects Using Noncontact Ultrasonics', *Materials 2018, Vol. 11, Page 128*, vol. 11, no. 1, p. 128, Jan. 2018, doi: 10.3390/MA11010128.
[17] B. Wang, S. J. Hu, L. Sun, and T. Freiheit, 'Intelligent welding system technologies: State-of-the-art review and perspectives', Jul. 01, 2020, *Elsevier*. doi: 10.1016/j.jmsy.2020.06.020.
[18] R. Tsuzuki, 'Development of automation and artificial intelligence technology for welding and inspection process in aircraft industry', Jan. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s40194-021-01210-3.
[19] M. A. Kesse, E. Buah, H. Handroos, and G. K. Ayetor, 'Development of an artificial intelligence powered tig welding algorithm for the prediction of bead geometry for tig welding processes using hybrid deep learning', *Metals*, vol. 10, no. 4, p. 451, Mar. 2020, doi: 10.3390/met10040451.
[20] S. Wazir, G. S. Kashyap, and P. Saxena, 'MLOps: A Review', Aug. 2023.
[21] N. Marwah, V. K. Singh, G. S. Kashyap, and S. Wazir, 'An analysis of the robustness of UAV agriculture field coverage using multi-agent reinforcement learning', *International Journal of Information Technology (Singapore)*, vol. 15, no. 4, pp. 2317–2327, May 2023, doi: 10.1007/s41870-023-01264-0.
[22] W. Ji and L. Wang, 'Industrial robotic machining: a review', *International Journal of Advanced Manufacturing Technology*, vol. 103, no. 1–4, pp. 1239–1255, Apr. 2019, doi: 10.1007/s00170-019-03403-z.
[23] S. H. Kim *et al.*, 'Robotic Machining: A Review of Recent Progress', *International Journal of Precision Engineering and Manufacturing*, vol. 20, no. 9, pp. 1629–1642, Sep. 2019, doi: 10.1007/S12541-019-00187-W/FIGURES/12.
[24] M. H. M. Ali and M. R. Atia, 'A lead through approach for programming a welding arm robot using machine vision', *Robotica*, vol. 40, no. 3, pp. 464–474, Mar. 2022, doi: 10.1017/S026357472100059X.
[25] N. Oh and H. Rodrigue, 'Toward the Development of Large-Scale Inflatable Robotic Arms Using Hot Air Welding', *Soft Robotics*, vol. 10, no. 1, pp. 88–96, Feb. 2023, doi: 10.1089/soro.2021.0134.
[26] A. Ghosh, D. Chakraborty, and A. Law, 'Artificial intelligence in Internet of things', Dec. 01, 2018, *The Institution of Engineering and Technology*. doi: 10.1049/trit.2018.1008.
[27] G. S. Kashyap, K. Malik, S. Wazir, and R. Khan, 'Using Machine Learning to Quantify the Multimedia Risk Due to Fuzzing', *Multimedia Tools and Applications*, vol. 81, no. 25, pp. 36685–36698, Oct. 2022, doi: 10.1007/s11042-021-11558-9.
[28] Z. Guo, Y. Sun, M. Jian, and X. Zhang, 'Deep residual network with sparse feedback for image restoration', *Applied Sciences (Switzerland)*, vol. 8, no. 12, p. 2417, Nov. 2018, doi: 10.3390/app8122417.
[29] D. Das, D. K. Pratihar, G. G. Roy, and A. R. Pal, 'Phenomenological model-based study on electron beam welding process, and input-output modeling using neural networks trained by back-propagation algorithm, genetic algorithms, particle swarm optimization algorithm and bat algorithm', *Applied Intelligence*, vol. 48, no. 9, pp. 2698–2718, Sep. 2018, doi: 10.1007/s10489-017-1101-2.
[30] C. Xia, Z. Pan, Z. Fei, S. Zhang, and H. Li, 'Vision based defects detection for Keyhole TIG welding using deep learning with visual explanation', *Journal of Manufacturing Processes*, vol. 56, pp. 845–855, Aug. 2020, doi: 10.1016/j.jmapro.2020.05.033.
[31] D. Bacioiu, G. Melton, M. Papaelias, and R. Shaw, 'Automated defect classification of Aluminium 5083 TIG welding using HDR camera and neural networks', *Journal of Manufacturing Processes*, vol. 45, pp. 603–613, Sep. 2019, doi: 10.1016/j.jmapro.2019.07.020.