# MobileNetV3 for Pipeline Welding Defect Detection: A Computer Vision Approach on Noisy RGB Images

*Ata Jahangir* Moshayedi[1,2], *Amir Sohail* Khan[3], *Ming Jun* Zhang[1,4], *Shuxin* Yang[5], and *David* Bassir[1,5,*]

[1]Smart Structural Health Monitoring and Control Laboratory, DGUT-CNAM Institute, Dongguan University of Technology, D1, Daxue Rd., Songshan Lake, Dongguan, Guangdong Province, P.R. China
[2]DGUT-CNAM Institute, Dongguan University of Technology, D1, Daxue Rd., Songshan Lake, Dongguan, Guangdong Province, P.R. China
[3]School of Computer Science and AI, Wuhan Textile University ,Wuhan,430200
[4]CEDRIC/Laetitia, CNAM, Paris, France
[5]School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi, China
[6]ENS -Paris-Saclay University, Centre Borelli, UMR CNRS 9010, 91190 Gif-sur-Yvette,france

**Abstract.** Weld defect detection plays a central role in pipeline inspection because it directly influences structural reliability and operational safety.Conventional manual inspection methods suffer from several limitations, including high labor demands, subjective judgment, and inconsistent results, which restrict their effectiveness in large-scale industrial applications. To overcome these challenges, this study evaluates the use of the lightweight convolutional neural network MobileNetV3, considering both its Small and Large configurations, for automatic multi-class classification of pipeline weld defects.The proposed approach integrates image preprocessing and data augmentation with a lightweight deep learning architecture to extract discriminative visual features from weld images. Experimental validation is carried out using two datasets acquired under practical inspection conditions. The AI5083 dataset contains 4,927 images captured using HDR cameras, while the PWDID dataset includes 78 images collected with a mobile phone camera in uncontrolled environments.The experimental results demonstrate strong classification performance, achieving an accuracy of 99.99% on the AI5083 dataset and improved results on the PWDID dataset. Compared with previously reported methods, the proposed approach yields an accuracy improvement of approximately 3%, indicating its robustness across different acquisition settings. In addition to its high accuracy, MobileNetV3 offers fast inference, low memory consumption, and reduced power requirements. Owing to these characteristics and ias compatibility with deployment frameworks such as TensorFlow Lite, ONNX, and CoreML, MobileNetV3 is well suited for mobile and handheld inspection tools. Based on these findings, the use of MobileNetV3 and other modern lightweight architectures along with denoised function are recommended for practical, on device pipeline weld defect inspection systems.

**Keywords:** Computer vision (CV), weld defect detection, pipeline inspection, deep learning, MobileNetV3, convolutional neural networks.

## 1 Introduction

Pipeline welding is a critical process in industrial sectors such as oil and gas, petrochemical plants, and infrastructure development[1]. Operational safety, system reliability, and maintenance costs of welded joints are directly dependent on structural integrity of the welded joints. The welding defects include cracks, absence of fusion, porosity, corrosion, and misalignment which will cause great weaknesses to the pipeline structures which may cause leakages, failure or a disastrous accident. Thus, ensuring quality control and maintenance prevention requires correct and timely identification of defects in welds aids in their prevention [2].Visual inspection of the welds, the ultrasonic measurements and radiographic analysis are some of the wide used methods of inspection of the welds in the industry toolkit. These procedures are also usually limited to expert inspectors, equipment sales, and time-consuming wastes of inspection [3]. In addition, manual inspection, to a great extent, is prone to human subjectivity and inconsistency, especially when inspection measures involve a large scale of inspection.

These constraints have encouraged the establishment of automated inspection to offer a reliable and speedy and objective way of detecting defects. These shortcomings have inspired the formulation of automated inspection systems, which can deliver credible, rapid and objective way of fading away the defects. [4].The recent developments in computer vision (CV) and deep learning have enabled a great improvement regarding automated image-based systems of inspection[5]. Moshayedi et al. (2019) introduced the design and construction of a flexible pipe inspection robot, which is fitted with a visual inspection system based on a camera to inspect the inside of the pipes.Image processing was applied in the form of a MATLAB based GUI

*e-mail: david.bassir@ens-paris-saclay.fr

with classical edge detectors, i.e., Sobel, Roberts, Canny, and Laplacian used to visualize the defects as opposed to automated use of the classification[6]. In order to use the deeplearning methods, the convolutional neural networks (CNNs) have shown incredible abilities in visual pattern recognition, in which a complex of hierarchy feature representation is automatically learned using unprocessed image data as input to the network [7]. in comparison to traditional hand developed feature-based techniques, deep learning-based CV programmes do not require hand-driven feature engineering, and are better resistant to changes in light, texture, and even defect appearance. Consequently, there has been an influx of the use of CNN-based techniques to industrial inspection of surfaces and classification of defects taking place in the industry as well as in defect classification of the products involved in the process of classification [8].Even some sucessful project shown the implementation of Handy Pipe Defect (HPD) recognition system which incorporates a Personal Image Classifier into a mobile app to detect weld defects.The training sample consisted of over 150 images and this was tested on 28 actual samples where the system had high accuracy of detecting defect-names and thus confirmed the feasibility of lightweight, vision-related solutions to on-site industrial inspection [9]. The review on available resources shows that YOLO method have more demand and used for the detection of defects in a pipeline. As the recent work shows, Zhang et al. (2024) proposed YOLOv8-CM with CBAM and U-net segmentation for tunnels, achieving 0.908 mAP detection and 0.890 segmentation accuracy [10].Chen et al., 2025 focused on detecting surface defects in long-distance oil and gas pipelines using Magnetic Flux Leakage (MFL) inspection data combined with deep learning. The authors proposed a cascaded model that integrates the YOLOv11 detection framework for defect recognition with a physics-informed, data-driven model for defect size estimation. Their method achieved an mAP50 of 92.1%, a precision of 100%, a recall of 84.29%, and an F1-score of 91.47%, proving that the model can effectively identify and quantify defects even with limited data and offers high reliability for pipeline safety assessment [11].Khan et al. (2025) presented a pipe line surface defect detection framework improve with attention mechanisms which was built upon the basis of YOLOv11. Based on a dataset of 1,500 images, the proposed YOLOv11+SelfAttention model has a highest accuracy of 98.95 percent, it is higher than other attention modules like LKA and CBAM. [12].

Zhang et al. (2025) came up with an object-detection framework which uses deep learning to detect the presence of welding defects in pressure-pipeline radiographs. The use of CNN-based detection models on X-ray images significantly improved the accuracy of the defect-detection compared with manual check, and the system showing great results in detecting defects like porosity, cracks, and absence of fused material, however, per-class performance differed among defects [13].Xiao et al (2025) added MSDA and SCConv modules to YOLOv8,attaining 94.4% mAP with improved multi-scale feature extraction [14].Sha et al. (2025) developed a lightweight cross-scale YOLOv8 model with HS-BiFPN and DySample, improving mAP by 3.8% on 1,952 sewer pipe images while reducing parameters and computational cost [15]. As the study shows most of the successful deep learning models work with large models with a high computational complexity, thereby reducing their use in a real-time system or machine with limited resources [16] which shows the need of lightweight architectures in this work domain. Even though there have been significant improvements in automated detectors of weld defects, a number of limitations still permeate through the literature that remains. First of all, the majority of the previous research focuses on enhancing the accuracy of classification and it does not consider the construction of a consistent computer-vision (CV) algorithm that would analyze the primary functions of preprocessing, data augmentation, feature extraction, and model optimization in a single lightweight model [17]. Additionally, many deep learning approaches rely on computationally heavy architectures, limiting their practical deployment in real-time or resource-constrained industrial environments and few work reported the implimention on handy device [18].

Another weakness comes as a result of performance of the models on individual datasets that provides limited information about the performance with regard to the robustness and generalizability when switching the datasets with different sizes and complexities.Besides working on the models tools are essential which have a unfiltered images and captured by non advanced camera.Furthermore currently used methods usually favor large sets of people-based corpora or small experiment-based sets without developing a consistent cross-analysis between the two in a single experimental pipeline. This is because the impact of lightweight convolutional neural networks architecture in concurrently attaining high accuracy and computational efficiency has not been adequately studied. Also, the efficacy of the models has not been well addressed in situations where data are scarce, and this reflects real-life industrial conditions [19].

To address this challenge, in this work the perfoamce of lightweight CNN architectures named MobileNet have been introduced, offering an effective trade-off between accuracy and computational efficiency which have the possibility to run on the handy device with respect to its Lightweight architecture,Low memory footprint, and Low latency [20, 21]. beside most of the used image data sets are captured under controlled environment which are far from real conditions and this study tried to use the of real image captured by athours mobile phone and labeled with expert (data set PWDID) [22] under uncontrolled environment. The main objectives of this study are summarized as follows:**1)**Test and report the performance of a light computer vision framework for multi-class weld defect classification applicable for the implementation of a handy device.**2)** Extensive evaluation the model on large and small size of two diffrent data based captured with different environment.**3)** Validation on a small self-collected dataset using MobileNetV3-Small and large, demonstrating strong generalization despite limited training data.**4)** compare the model performance with the pervious reported result for

one of the data sets.**5)**Test and check the model performance over an uncontrolled environment.
The remainder of this paper is organized as follows. Section II describes the datasets and preprocessing methods. Section III presents the proposed computer vision framework and model architecture. Section IV discusses the experimental setup and evaluation metrics. Section V reports the results and comparative analysis, followed by conclusions and future work in Section VI.

## 2 Methodology

### 2.1 Dataset

In this work two set of data set named aI5083 and PWDID .Figure 1 summarizes the defect categories included in both the aI5083 dataset and the PWDID dataset [22], along with representative sample images and dataset sources described as follows:

#### 2.1.1 Dataset 1:Large data set:al5083

The publicly available aI5083 dataset contains 4,927 weld images acquired with HDR camera, spanning six defect categories and is widely used for benchmarking weld defect classification. The images were captured using an industrial high-resolution camera under controlled laboratory conditions, ensuring consistent illumination and imaging geometry.

#### 2.1.2 Dataset 2:Small dataset:PDWID

As The proposed dataset comprises 78 real-world weld images covering seven defect types, collected under practical inspection conditions. The images were acquired using a standard RGB camera of mobile phone in unconstrained environments, introducing variations in lighting and surface appearance. Due to its limited size, all samples are utilized for both training and testing to assess model adaptability in data-scarce industrial scenarios.

### 2.2 MobileNetV3:

The main core of the proposed model is the MobileNetV3 framework, which is specifically created to offer the best balance between the accuracy of the classification and the computing power[23]. The model enables the framework to have good classification rate whilst being appropriate in industrial inspection scenarios with real-time and resource limitations. MobileNetV3 uses depthwise separable convolutions, inverted residual blocks and squeeze-and-excitation to build lean yet powerful models with great feature representation potential as it shown in algorthem 1. In this paper, two versions of MobileNetV3, including MobileNetV3-Large and MobileNetV3-Small, are used to fit various data availability and deployment requirements (see Figure 2).The two variants are feature extractors in the postulated computer vision framework to convert image inputs of welders into non-dimensional feature vectors that are then fed to fully connected layers

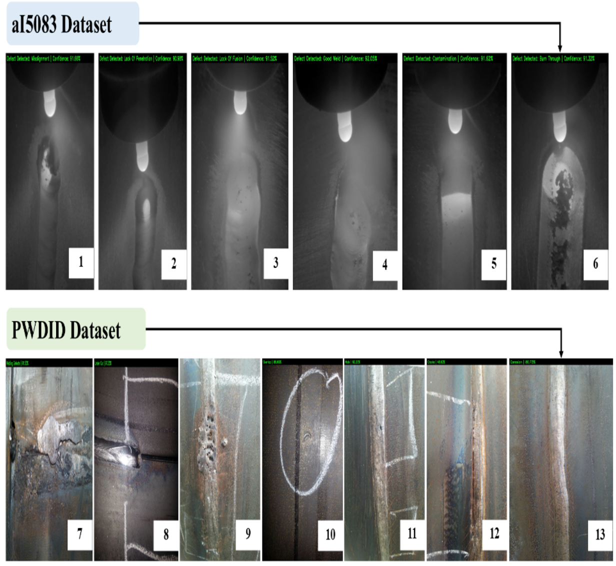


**Figure 1.** DEFECT CATEGORIES AND REPRESENTATIVE SAMPLES FROM THE PUBLIC AI5083 & PWDID DATASET, aI5083 Dataset: 1= Misalignment, 2= Lack of Penetration, 3= Lack of Fusion, 4= Good Weld, 5= Contamination, 6= Burn Through, 7= Welding Defects, 8= Under Cut, 9= Porosity, 10= Overlap, 11= Hole, 12= Cracks, 13= Corosion

where multi-class defects are classified. MobileNetV3-Large is used to perform experiments on the large-scale open dataset, wherein large training samples makes the model learn varied and discriminative features of the visuals[24].Conversely, MobileNetV3-Small is used to train a ,self-collected dataset, its small size has minimized chances of overfitting and it can be successfully trained with the limited data situation.In the present work, to experiment with the large public dataset, MobileNetV3-Large is employed to obtain more detailed feature representations, and MobileNetV3-Small is considered on the self-collected dataset because of the fewer number of parameters and better generalization in the case of the limited amount of data. As shownn in Figure 2 MobileNetV3 is the featured base extractor in the computer vision structure proposed in both scenarios, and the output feature vectors of the feature extractor are subjected to fully connected layers to classify defects in welds into multiple categories.
MobileNetV3 as it shown in algorthem 1 uses depthwise separable convolutions which break down regular convolution operations into depthwise and pointwise convolutions, which greatly reduce the computational cost. The inverse of blocked residual blocks are used to carry out the expansion of feature dimensions in the intermediate layers and then put them so they are passed back to the compact form to facilitate effective flow of information.Also, squeeze-and-excitation blocks can adaptively recalibrate channel-wise feature responses to inter-channel dependencies, and thus improve significant visual features with regard to weld defects [25].
For the aI5083 dataset, MobileNetV3-Large [26] is used

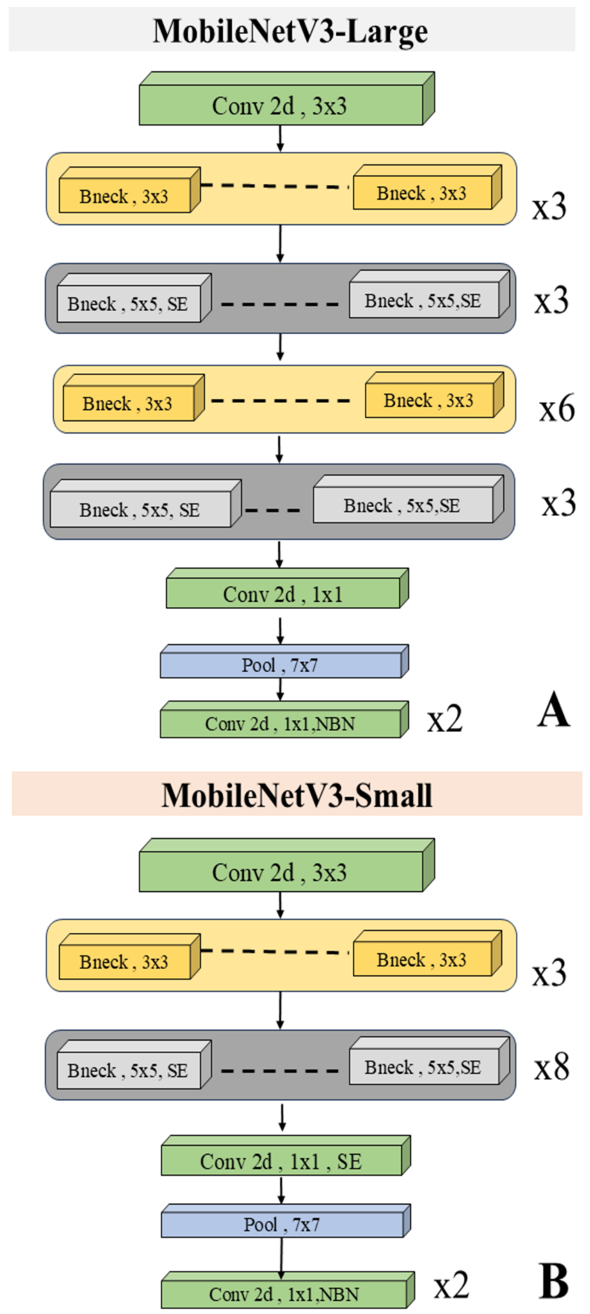


**Figure 2.** MobileNetv3-Large and MobileNetv3-Small feature extraction network.

due to its enhanced representational capacity and ability to capture complex defect patterns. For the PWDID dataset, MobileNetV3-Small [27] is employed to reduce model complexity while maintaining strong performance. Both models utilize depthwise separable convolutions and optimized network design, making them well-suited for real-time and resource-limited industrial applications. The final classification layer of each model is modified to match the number of defect classes in the respective datasets.The detail paramter of models are shown in Table 1
For classification and prediction of these images the ex-

---

**Algorithm 1** Pipe Defect Classification Using MobileNetV3

---

**Require:**
Training images $\mathcal{D}_{tr}$,
validation images $\mathcal{D}_{val}$,
number of classes $C$, learning rate $\alpha$,
batch size $B$, epochs $E$
**Ensure:** Trained model, prediction results, confusion matrix
Initialize random seed and computation device
Define defect label mapping
Create result directories
Define image preprocessing: Resize to 224×224, normalize using ImageNet statistics
Load training dataset $\mathcal{D}_{tr}$ and validation dataset $\mathcal{D}_{val}$
Create DataLoaders with batch size $B$
Load pre-trained MobileNetV3-Small
Replace final classifier layer with $C$ output neurons
Initialize Cross-Entropy loss and Adam optimizer
**for** epoch = 1 to $E$ **do**
    Set model to training mode
    **for** each batch $(x, y)$ in $\mathcal{D}_{tr}$ **do**
        Forward pass: $\hat{y} \leftarrow f(x)$
        Compute loss $\mathcal{L}(\hat{y}, y)$
        Backpropagate loss
        Update model parameters
    **end for**
**end for**
Set model to evaluation mode
Initialize prediction records
**for** each sample $x_i$ in $\mathcal{D}_{val}$ **do**
    Predict class probabilities using softmax
    Obtain predicted label and confidence
    Save annotated output image
    Store true and predicted labels
**end for**
Save results to Excel file
Compute and save confusion matrix **return** Trained model and evaluation results

---

tracted feature vector is passed to a fully connected classification layer to obtain class scores:

$$\mathbf{z} = \mathbf{W}\mathbf{f} + \mathbf{b}, \tag{1}$$

where $\mathbf{W}$ and $\mathbf{b}$ denote the weight matrix and bias vector of the classifier. The predicted class probabilities are computed using the softmax function:

$$P(y = k \mid \mathbf{I}) = \frac{\exp(z_k)}{\sum_{j=1}^{C} \exp(z_j)}, \tag{2}$$

where $C$ is the total number of defect classes.

Due to the limited size of this dataset, all images are used for both training and testing to analyze the framework's performance under constrained data conditions.

## 2.3 Evaluation Metrics

To make a numerical analysis of how well the proposed computer vision-based weld defect classification framework functions, are accuracy, precision, recall, and F1

**Table 1.** Training Settings Comparison: MobileNetV3-Small vs MobileNetV3-Large

| **Parameter** | **MobileNetV3-Small** | **MobileNetV3-Large** |
|---|---|---|
| Classes | 7 | 6 |
| Batch Size | 32 | |
| Epochs | 40 | |
| Learning Rate | 0.0003 | |
| Image Size | 224 × 224 | |
| Device | CUDA/CPU | |
| Random Seed | 42 | |
| Loss Function | CrossEntropy | (label smoothing=0.1) |
| Optimizer | AdamW | AdamW (weight decay=1e-4) |
| Scheduler | CosineAnnealingLR (T_max=40) | |
| Weighted Sampler | Yes | |
| Data Augmentation | Resize, RandomCrop, Flip, Rotation, ColorJitter, GaussianBlur, Normalize | Random Resized Crop, HorizontalFlip, Rotation, ColorJitter, GaussianBlur, Normalize |
| Train/Test Split | 100-100 | 70-30 |
| model.classifier[3] | nn.Linear(1024, 7) | nn.Linear(1280, 6) |
| Computational Complexity | Lightweight | Higher parameter count |

score, commonly applied to the multi-class classification tasks in order to have a complete and interpretable measure of model efficacy as it shown in Table (2). Besides, confusion matrices are created to depict the behavior in the shape of classes and the possible misclassifications steps.these metrics provide a robust evaluation framework for assessing both overall performance and class-wise prediction quality of the proposed model.

## 3 workflow and experimental result

The general grasp of the workflow of the suggested computer vision-based system of detecting weld defects is shown in Figure 3.As it shown,the computer vision (CV) model adheres to the systematic end-to-end visual processing pipeline which includes image retrieval starting with the aI5083 dataset and PWDID dataset. The obtained images are systematically preprocessed, i.e. they are resized and normalized to guarantee that the spatial and color consistency is consistent across datasets.In order to promote robustness, generalization, data augmentation models like rotation, flipping, color jittering, and blurring are used in the process of training. An impartial three-step feature extraction process is subsequently employed based on a lightweight MobileNetV3 backbone which trains on the extracted images in an attempt to discover discriminative

**Table 2.** Evaluation Metrics,$TP$ (True Positive) represents the number of correctly classified positive samples, $TN$ (True Negative) denotes the number of correctly classified negative samples, $FP$ (False Positive) corresponds to negative samples incorrectly classified as positive, and $FN$ (False Negative) refers to positive samples incorrectly classified as negative.

| **Metric** | **Formulation** |
|---|---|
| Accuracy | $$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$ |
| Precision | $$P = \frac{TP}{TP + FP} \quad (2)$$ |
| Recall | $$R = \frac{TP}{TP + FN} \quad (3)$$ |
| F1-score | $$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (4)$$ |

visual representations of the images. The detached feature vectors are then subjected to dense layers in multi-class defect classification.Lastly, the framework measures framework performance via metrics.In the pipeline, the first step is the image acquisition of both the public and self-collected datasets or data, which is then proceeded to image preprocessing and augmentation to improve the quality and generalization of the data. The models are then fed with the images preprocessed and features extracted thereof, trained. Standard classification measures are used to assess model performance and the ultimate result is in expected classes of weld defects with the corresponding confidence score.

### *3.0.1 Image Preprocessing and Data Augmentation*

Weld defect detection involves image processing of features to enhance the learning process of the imaging system efficiency of detecting features in an image of the weld line [28]. This is done to standardize all input images to a common resolution of $224 \times 224 pixels$. ImageNet mean and standard deviation values are used to compute color normalization and stabilize training and hasten convergence. Large-scale data augmentation strategies are used in the training so as to increase the robustness of a model and minimize overfitting; these methods include resized cropping and horizontal flipping, rotation, color jittering and Gaussian blurring. These additions are a simulation of actual variations in the real world like variations in lighting, change of perspective, surface texture which are often found in the industrial inspection circumstances. The pre-processing of image is used to equalize the intensity of pixels and enhance the stability of training. The normalized image $\mathbf{I}_n$ is computed as:

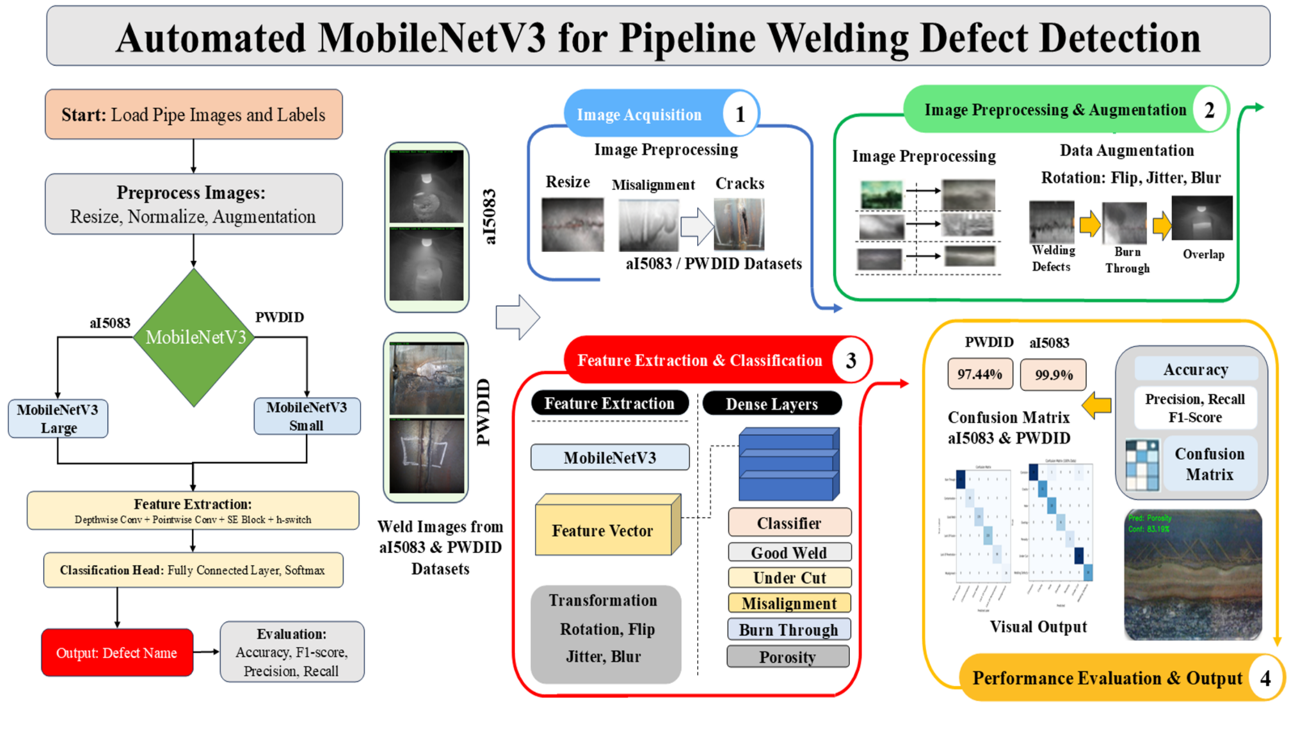


**Figure 3.** Overall computer vision (CV) framework proposed for automated pipeline weld defect detection.The MObilenetV3 large for I5083 and the PDWID data set for small model.

$$\mathbf{I}_n = \frac{\mathbf{I} - \boldsymbol{\mu}}{\boldsymbol{\sigma}}, \tag{3}$$

where $\boldsymbol{\mu} = [0.485, 0.456, 0.406]$ and $\boldsymbol{\sigma} = [0.229, 0.224, 0.225]$ is the channel mean and standard deviation.

### 3.1 Training Strategy

AdamW optimizer is used to train the models and the learning rate is 0.0003 and a batch size of 32 and 40 epochs. A cross-entropy loss is used to tackle label smoothing as a mitigation measure in generalizing and decrease the overconfidence of prediction. The random seeds are fixed in all experiments in order to make them reproducible.

The network is trained using the cross-entropy loss function:

$$\mathcal{L} = -\sum_{k=1}^{C} y_k \log\left(P(y = k \mid \mathbf{I})\right), \tag{4}$$

where $y_k$ is the ground-truth label encoded in one-hot format. The loss is minimized using the AdamW optimizer with cosine annealing learning rate scheduling.

## 4 Results and Discussion

The experimental evaluation of the proposed framework on both dataset are shown in this section. The first experiment was conducted on the publicly available dataset consisting of 4,927 weld images categorized into six defect classes. The dataset was divided into 3,695 training images (75%) and 1,232 testing images (25%). MobileNetV3-Large is employed to evaluate the effectiveness of the proposed framework under large-scale conditions. Table 3 presents the classification performance achieved on the test set. The proposed framework achieves perfect performance across all evaluation metrics, indicating robust feature learning and excellent class separability.

The figure 4 A & B shows the graph for the test and train loss for the aI5083 dataset. The corresponding confusion matrix shown in figure 4 C, demonstrates a perfectly diagonal structure, confirming zero misclassification across all defect categories in handling large and imbalanced datasets. To further validate the robustness of the proposed framework under limited data conditions, experiments are conducted on a self-collected dataset consisting of 78 images representing seven weld defect categories. Due to the small dataset size, all images are used for both training and testing. MobileNetV3-Small is selected to reduce model complexity while maintaining reliable performance.

Table 3 summarizes the classification results obtained on the self-collected dataset. The framework achieves an overall accuracy of 97.44%, demonstrating strong generalization capability despite the limited number of training samples.The confusion matrix for the self-collected dataset in figure 4D, shows a predominantly diagonal structure with minor misclassification observed in the corrosion class. These results indicate that the proposed

**Table 3.** Performance Metrics on aI5083 [19] and PWDID Datasets[22] Using MobileNetV3

| **Dataset: aI5083 (MobileNetV3-Large)** | | | | |
|---|---|---|---|---|
| **Class** | **Precision** | **Recall** | **F1-score** | **Support** |
| Burn Through | 0.99 | 0.99 | 0.99 | 615 |
| Contamination | 0.99 | 0.99 | 0.99 | 89 |
| Good Weld | 0.99 | 0.99 | 0.99 | 170 |
| Lack of Fusion | 0.99 | 0.99 | 0.99 | 233 |
| Lack of Penetration | 0.99 | 0.99 | 0.99 | 99 |
| Misalignment | 0.99 | 0.99 | 0.99 | 26 |
| **Accuracy** | | 0.99 | | |
| Macro Avg | 0.99 | 0.99 | 0.99 | 1232 |
| Weighted Avg | 0.99 | 0.99 | 0.99 | 1232 |

| **Dataset: PWDID (MobileNetV3-Small)** | | | | |
|---|---|---|---|---|
| **Class** | **Precision** | **Recall** | **F1-score** | **Support** |
| Corrosion | 1.00 | 0.90 | 0.9474 | 20 |
| Cracks | 1.00 | 1.00 | 1.00 | 11 |
| Hole | 0.91 | 1.00 | 0.9524 | 10 |
| Overlap | 1.00 | 1.00 | 1.00 | 6 |
| Porosity | 1.00 | 1.00 | 1.00 | 3 |
| Under Cut | 0.95 | 1.00 | 0.9730 | 18 |
| Welding Defects | 1.00 | 1.00 | 1.00 | 10 |
| **Accuracy** | | 0.9744 | | |
| Macro Avg | 0.9795 | 0.9857 | 0.9818 | 78 |
| Weighted Avg | 0.9762 | 0.9744 | 0.9742 | 78 |

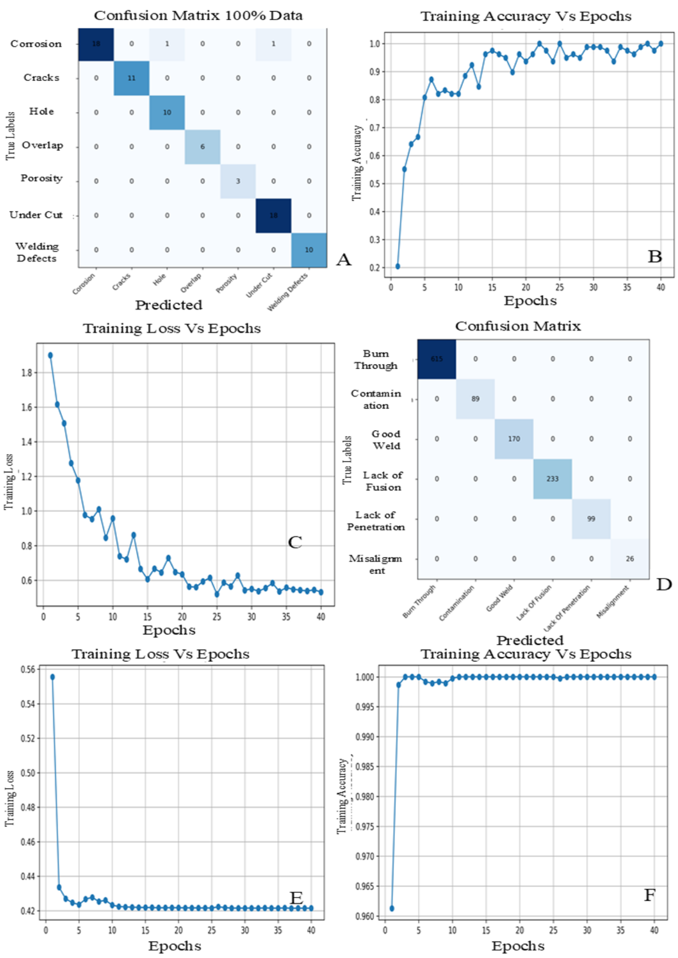


**Figure 4.** A: Confusion Matrix (PWDID), B: Train accuracy vs Epochs (PWDID), C: Training Loss vs Epochs (PWDID), D: Confusion Matrix (ai5083 dataset), E: Training Loss vs Epochs (ai5083 dataset), F: Training Accuracy vs Epochs (ai5083 dataset)

CV framework remains effective even under constrained data availability, although limited sample size introduces slight performance variations. While the figure 4 E & F, shows the graph for the test and train loss for our PWDID dataset. The experiment findings show that the suggested deep learning framework that is based on computer vision, can reach superior performance in both datasets. On the aI5083 , the ideal classification shows the appropriateness of MobileNetV3-Large to massive weld defects inspection. On the PWDID data, MobileNetV3-small is more accurate but minimizes its computing power, which evidences flexibility of light CNNs architecture and the results shows that uncontolled envirment data will not limited the model performance. The integrated systematic preprocessing, data enhancement, balanced sampling, and effective model construction allows effective defect detection regardless of large-scale and small-data conditions, which means that the framework can be deployed to practice in the industry. In contrast to most of the literature, who have dealt mostly with model-centric improvements or accuracy of detection or used YOLO versions with different attention mechanisms, the tested model is explicitly structured by formulating the entire CV pipeline, that is, it includes image acquisition, standardized preprocessing, data augmentation, feature extraction, supervised classification, global multi-metric evaluation, and visual decision support. This two-dataset analysis indicates the strength and flexibility of the framework in data-abundant as well as data-sparse situations, both of which are scarcely tackled in the former literature.

## 5 Conclusion

This paper presented a robust computer vision-based deep learning framework, MobileNetV3 for automated weld defect detection in pipeline inspection systems. The tested approach integrates systematic image preprocessing, data augmentation, and lightweight convolutional neural networks to achieve highly accurate and computationally efficient multi-class defect classification. Two datasets were used to comprehensively evaluate the framework with 4,927 images that captured with the HDR camera and 78 image which used the mobile phone to acquire are used.for bothe the image the denoise process didnt applay and specially the second data set have some unctrolled enviroment effect which didnt remove to evaluate the framework. Experimental results on first data set(aI5083 datase) demonstrated exceptional 99.9% accuracy and the second data set PWDID have the accuracy of 97.43 . along with limited number of data in second data set considering the first data set performance with the perior result(6-class: 71%, 4-class: 89%, 2-class: 95% ) and obtained result in this student with out denoise steps and controlling the environment effect, the model shown the promising performance which can be counted as the potentinal of MobileNetV3 model to used as the baseline model in welding pipe detection activity.Despite the constrained dataset size and reuse of samples for training and testing, the framework demonstrated reliable performance, highlighting its adaptability to small-scale and practical industrial scenarios. This makes the proposed approach suitable for real-time and resource-constrained industrial inspection systems, such as on-site pipeline monitoring and automated quality assurance. Unlike the earlier work, which relied on high-resolution HDR imagery and task-specific architectures.These results highlight the robustness, scalability, and superior discriminative capability of the proposed method for automated weld defect detection.Future work will focus on expanding the dataset and add the denoise step along with the imporvment on MobileNetV3 model for the acquired image on real-time welidng pipeline defect industrial applications.

## References


[1] Li, D., Xie, Q., Yu, Z., Wu, Q., Zhou, J., & Wang, J. (2021). Sewer pipe defect detection via deep learning with local and global feature fusion. Automation in Construction, 129, 103823.

[2] Madhvacharyula, A. S., Pavan, A. V. S., Gorthi, S., Chitral, S., Venkaiah, N., & Kiran, D. V. (2022). In situ detection of welding defects: A review. Welding in the World, 66(4), 611-628.

[3] Zhang, J., Liu, X., Zhang, X., Xi, Z., & Wang, S. (2023). Automatic detection method of sewer pipe defects using deep learning techniques. Applied Sciences, 13(7), 4589.

[4] Goharinezhad, S., Hadi, A., & Mirzaei, S. (2025). Automated defect classification and localization in sewer pipelines using hybrid ResNet50–Swin transformer and modified YOLOv8 on CCTV inspection images. Scientific Reports, 15(1), 41755.

[5] Moshayedi, A. J., Khan, A. S., Khan, Z. H., & Zhang, X. (2025). Safeguarding smart infrastructure: A review of deep learning techniques for automatic pipeline defect detection. Empowering AI Applications in Smart Life and Environment, 67-92.

[6] Mohammed, A., & Hussain, M. (2025). Advances and challenges in deep learning for automated welding defect detection: A technical survey. IEEE Access.

[7] Matusz-Kalász, D., Bodnár, I., & Jobbágy, M. (2025). An Overview of CNN-Based Image Analysis in Solar Cells, Photovoltaic Modules, and Power Plants. Applied Sciences, 15(10), 5511.

[8] Reddy, B. S. H. (2025). Deep learning-based detection of hair and scalp diseases using CNN and image processing. Milestone Transactions on Medical Technometrics, 3(1), 145-155.

[9] Zhang, M., Feng, M., Chen, C., Yu, X., & Lian, G. (2026). Weld defect detection: deep learning-based image processing and the mechanisms of defect formation. Archives of Computational Methods in Engineering, 33(3), 3525-3563.

[10] Zhang, C., Chen, X., Liu, P., He, B., Li, W., & Song, T. (2024). Automated detection and segmentation of tunnel defects and objects using YOLOv8-CM. Tunnelling and Underground Space Technology, 150, 105857.

[11] Chen, X. M., Fu, M. K., Shao, J., & Liu, X. B. (2025). A cascaded pipeline defect detection and size estimation method based on the YOLOv11 and physics-informed network using the MFL data. Petroleum Science.

[12] Khan, A. S., Rahman, A., & Moshayedi, A. J. (2025). Pipeline Surface Defect Detection Using YOLOv11 with Attention Mechanisms: A Comparative Study of SA, LKA, and CBAM Approaches. Journal of Robotics Research (JRR), 2(2).

[13] Zhang, W., Liu, W., Yu, X., Kang, D., Xiong, Z., Lv, X., ... & Li, Y. (2025). Deep learning-based automated detection of welding defects in pressure pipeline radiograph. Coatings, 15(7), 808.

[14] Xiao, L., & Zhang, H. (2025, January). Improved drainage pipe defect detection algorithm based on YOLOv8. In The International Conference Optoelectronic Information and Optical Engineering (OIOE2024) (Vol. 13513, pp. 1025-1030). SPIE

[15] Sha, R., Zhang, Z., Cui, X., & Mu, Q. (2025). A lightweight cross-scale feature fusion model based on YOLOv8 for defect detection in sewer pipeline. PLoS One, 20(8), e0330677.

[16] Guyon, R., Newson, M., Fisher, C., Miorelli, R., & Roué, D. (2025, July). Towards embedded AI models for welding defect detection in pipes. In 2025 IEEE Sensors Applications Symposium (SAS) (pp. 1-6). IEEE.

[17] Mumuni, A., Mumuni, F., & Gerrar, N. K. (2024). A survey of synthetic data augmentation methods in computer vision. arXiv preprint arXiv:2403.10075.

[18] Khamparia, A., & Singh, K. M. (2019). A systematic review on deep learning architectures and applications. Expert Systems, 36(3), e12400.

[19] Bacioiu, D., Melton, G., Papaelias, M., & Shaw, R. (2019). Automated defect classification of Aluminium 5083 TIG welding using HDR camera and neural net-

works. Journal of manufacturing processes, 45, 603-613.

[20] Oulahyane, A., Kodad, M., Addou, E. H., Ourarhi, S., & Chafik, H. (2025). Hybrid Approach for Early Road Defect Detection: Integrating Edge Detection with Attention-Enhanced MobileNetV3 for Superior Classification. International Journal of Advanced Computer Science & Applications, 16(4).

[21] Jahangir Moshayedi, A., Nowzari, R., Taherinezhad, M., Esmaeili Najafabadi, H., Khan, A. S., Ghadiri Nejad, M., & Ghanbari, N. (2025). Investigation and Characterization of Pipe Defects and Techniques, and Challenges Toward the Protection of Environmental Protection. Iranian Journal of Chemistry and Chemical Engineering, 44(6), 1724-1756.

[22] Moshayedi, A. J. (2025). PWDID (Pipe Welding Defect Image Dataset) (Version V01). Zenodo. https://doi.org/10.5281/zenodo.16935930

[23] Qian, S., Ning, C., & Hu, Y. (2021, March). MobileNetV3 for image classification. In 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) (pp. 490-497). IEEE.

[24] Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 3(1), 91-99.

[25] Chen, L., Yao, H., Gan, K., Huang, Z., Zhang, J., Ng, C. T., & Fu, J. (2025). Evaluation of MobileNetV3-Large for crack classification across low-and high-resolution images. Frontiers in Built Environment, 11, 1724879.

[26] Koonce, B. (2021). MobileNetV3. In Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization (pp. 125-144). Berkeley, CA: Apress.

[27] Zhao, Y., Huang, H., Li, Z., Yiwang, H., & Lu, M. (2022). Intelligent garbage classification system based on improve MobileNetV3-Large. Connection Science, 34(1), 1299-1321.

[28] Cao, Z., Li, J., Fang, L., Li, Z., Yang, H., & Dong, G. (2025). Research on efficient classification algorithm for coal and gangue based on improved MobilenetV3-small. International Journal of Coal Preparation and Utilization, 45(2), 437-462.