



ARTICLE

A Deep Learning Approach to Industrial Corrosion Detection

Mehwash Farooqui¹, Atta Rahman^{2,*}, Latifa Alsuliman¹, Zainab Alsaif¹, Fatimah Albaik¹,
Cadi Alshammari¹, Razan Sharaf¹, Sunday Olatunji¹, Sara Waslallah Althubaiti¹ and Hina Gull³

¹Department of Computer Engineering (CE), College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, Dammam, 31441, Saudi Arabia

²Department of Computer Science (CS), College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, Dammam, 31441, Saudi Arabia

³Department of Computer Information Systems (CIS), College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University, Dammam, 31441, Saudi Arabia

*Corresponding Author: Atta Rahman. Email: aaurahman@iau.edu.sa

Received: 21 June 2024 Accepted: 06 September 2024 Published: 18 November 2024

ABSTRACT

The proposed study focuses on the critical issue of corrosion, which leads to significant economic losses and safety risks worldwide. A key area of emphasis is the accuracy of corrosion detection methods. While recent studies have made progress, a common challenge is the low accuracy of existing detection models. These models often struggle to reliably identify corrosion tendencies, which are crucial for minimizing industrial risks and optimizing resource use. The proposed study introduces an innovative approach that significantly improves the accuracy of corrosion detection using a convolutional neural network (CNN), as well as two pretrained models, namely YOLOv8 and EfficientNetB0. By leveraging advanced technologies and methodologies, we have achieved high accuracies in identifying and managing the hazards associated with corrosion across various industrial settings. This advancement not only supports the overarching goals of enhancing safety and efficiency, but also sets a new benchmark for future research in the field. The results demonstrate a significant improvement in the ability to detect and mitigate corrosion-related concerns, providing a more accurate and comprehensive solution for industries facing these challenges. Both CNN and EfficientNetB0 exhibited 100% accuracy, precision, recall, and F1-score, followed by YOLOv8 with respective metrics of 95%, 100%, 90%, and 94.74%. Our approach outperformed state-of-the-art with similar datasets and methodologies.

KEYWORDS

Deep learning; YOLOv8; EfficientNetB0; CNN; corrosion detection; Industry 4.0; sustainability

1 Introduction

In Saudi Arabia, combating corrosion in critical sectors such as oil and gas is an immensely challenging task. Corrosion predominantly affects metals under specific environmental conditions: high humidity, wetness, elevated temperatures, exposure to hazardous chemicals, galvanic reactions, ultraviolet (UV) light, and pollution [1]. These elements lead to rapid structural degradation, resulting in significant economic and environmental impacts, as observed in pipeline failures and infrastructure



deterioration. To counter this issue, a study introduces an innovative application of deep learning, a form of artificial intelligence that replicates the human brain's ability to recognize patterns. This approach represents a marked departure from the slow, unsafe, and often inaccurate traditional manual methods of corrosion inspection [1]. Through the application of advanced image recognition algorithms in deep learning, the research aims to revolutionize the early detection of corrosion, improving safety, accuracy, and efficiency. This research stands out for its automation of a labor-intensive process, addressing a critical gap in current methodologies with a more reliable and efficient AI-driven approach. The serious impact of corrosion, which includes frequent structural failures and economic losses, emphasizes the urgent need for enhanced detection methods. Existing studies have limitations as they mostly concentrate on specific areas for corrosion detection, such as bolts, pipelines, and selected industrial parts, rather than covering the general types of corrosion. This new approach targets the detection of any kind of corrosion from industrial images. This study aims to reduce incidents of undetected corrosion, enhancing safety, preserving the environment, and providing economic benefits. Based on the literature, YOLOv8 and EfficientNet are identified as promising techniques for image processing and corrosion detection. The study contributes to existing knowledge by conducting a comprehensive literature review to identify research gaps. Three candidate techniques, including a convolutional neural network (CNN) developed from scratch and two pre-trained models (YOLOv8 and EfficientNetB0), are shortlisted. An augmented dataset comprising publicly available and locally obtained industrial data is used to train the models. The proposed models are comprehensively analyzed against state-of-the-art techniques in the literature, and the results are described.

The remaining part of the paper is structured as follows: [Section 2](#) covers related work in the literature on corrosion detection. [Section 3](#) details the materials and methods utilized in the study. [Section 4](#) presents the results and discussion, while [Section 5](#) concludes the paper.

2 Literature Review

In their research, Bondada et al. [2] used an automated technique to detect corrosion in industrial pipes. This approach uses machine vision principles to not only identify corrosion, but also measure the resulting damage. The process starts by identifying scene constraints to guide effective vision processing. Next, the image of the corroded parts is captured, and noise is removed using appropriate filtering methods. Converting the color space to HSI enhances corrosion identification. Corrosion is then detected by calculating the mean saturation value, and measures are applied to evaluate the extent of corrosion. The experimental results showed that this method accurately identifies and measures the level of corrosion on the pipeline surface.

Yao et al. [3] proposed a new method using convolutional neural networks (CNNs) to detect and recognize corrosion damage in marine structures, to prevent catastrophic accidents. They utilized a dataset of 330 high-resolution hull plate images, and before training, the data was processed to have a zero mean to improve the model's convergence properties. Their CNN model was based on the Alexnet architecture, along with an overlap-scanning sliding window algorithm to recognize and locate the position of corrosion damage. The model achieved an accuracy of 98.9% in recognizing and labeling corrosion damage, with the best recognition results obtained from the highest light intensity images. Bastian et al. [4] published an interesting work on corrosion detection as well. They developed a computer vision-based system for detecting corrosion in pipelines. To aid in this task, they gathered a comprehensive collection of over 140,000 optical pictures of pipelines displaying various phases of corrosion development. Their method included the deployment of a custom CNN specifically created

for the categorization of pipeline photos based on their corrosion levels. Remarkably, the model achieved a classification accuracy rate of 98.8%, demonstrating the ability to distinguish between photos of corroded pipelines and those of rust-free pipelines with patterns resembling corrosion features. Notably, the suggested CNN model outperformed most state-of-the-art classifiers in this domain. Additionally, they presented a localization technique based on a recursive region-based methodology, with the objective of locating corroded spots within a given image with greater precision.

In their study, the authors of [5] proposed a novel approach to address the challenge of detecting pitting corrosion. They integrated history-based adaptive differential evolution with linear population size reduction (LSHADE), image processing techniques, and the support vector machine (SVM). The study involved capturing a total of 213 images representing both corroded and non-corroded metal structures in Danang City, Vietnam. These images underwent a rigorous region of interest (ROI) extraction process. The researchers used various techniques to optimize and preprocess the dataset images, including multilevel image thresholding, morphological operations, and the LSHADE metaheuristic. Subsequently, texture computation was carried out using statistical measurements from color channels, GLCM, and LBP descriptors. The evaluation results were impressive, with the LSHADE-SVC-PCD method demonstrating a classification accuracy rate of 91.80%, precision of 91%, recall of 94%, negative predictive value of 93%, and an F1-score of 92%. However, the authors noted the limited amount of data in the dataset and the computational expense of the LSHADE metaheuristic. They also suggested utilizing deep learning models with a larger number of images.

The article [6] introduced a new deep learning structure called Densely Connected Cascade Forest-Weighted K Nearest Neighbors (DCCF-WKNNs) to model data related to corrosion and extract knowledge about corrosion. The study begins by collecting corrosion samples of low-alloy steels and describes the methods used, including the combination of RF-WKNNs, random forests-K, and DCCF-WKNNs. The results of this approach show that it outperforms commonly used methods such as artificial neural networks (ANN), support vector regression (SVR), random forests (RF), and cascade forests (cForest) in predicting corrosion rates. Additionally, the method can predict corrosion rates based on variations in individual environmental variables, allowing for the identification of specific values for each variable beyond which the rate of corrosion could undergo significant changes. Another related study [7] explored corrosion detection on metals with a focus on quick and safe inspections. According to Katsamenis et al., the major limitation of image analysis approaches that involve bounding boxes to identify corroded areas is that they are insufficient for in-depth structural analysis and innovative construction techniques such as prefabrication. After comparing the deep learning models for corrosion detection—FCN, U-Net, and Mask R-CNN—the study found that while these models were not adequate for structural analysis, they exhibited superior accuracy and efficiency compared to traditional approaches. To enhance accuracy, the study proposes a data projection technique that combines contour segmentation with processed deep learning masks [7].

Matthaiou et al. [8] conducted a study on automated corrosion detection solutions, focusing on utilizing visual attributes of corrosion and artificial intelligence techniques for image analysis. They employed three approaches for corrosion detection: texture-based using convolutional neural networks, color-based using color algorithms, and a single shot detector using transfer learning on real-world images. The study revealed that deep learning approaches, specifically the use of single shot detectors, produced better results compared to other methods, making them more suitable for real-world applications. Zuchniak et al. [9] carried out a study to demonstrate the effectiveness of proposed CNN architectures in non-destructive aircraft fuselage inspection methods. The images used in the study were obtained through the DAIS (Digital Automated Inspection System) imaging system, which is a non-invasive imaging approach used in technical inspections in Polish aircraft

systems maintenance. The study's methodology involved using a multi-teacher/single-student strategy combined with an ensemble learning scheme and a new version of knowledge distillation. The proposed project utilized an ensemble of CNN models to enhance the system's classification accuracy. The study found that the geometric student distillation model performed the best, detecting 100% of the "moderate corrosion" class and outperforming other metrics such as accuracy, recall, precision, and F1-score. However, the study had a relatively small dataset, which may limit the generalizability of the findings. Additionally, evaluating the performance of the proposed methodology in comparison to others presented a challenge.

A study by Brandoli et al. [10] aimed to identify corrosion in aircraft fuselage images. The dataset includes 210 fuselage joint images from Boeing 727 and Airbus 300 aircraft models. The images of corroded and non-corroded sections were captured using a DAIS sensor. The experiment used five-fold cross-validation, image cropping, resizing, and various pre-trained architectures such as Inceptionv3, ResNet-101, NasNet, SqueezeNet-201, and DenseNet-121. Data augmentation techniques like shearing and scaling were also used. According to the findings, DenseNet achieved an accuracy of 92.2%, while SqueezeNet had an accuracy of 91.6%.

Another study [11] introduced a deep learning model, AMCD, for real-time corrosion detection on Micro Aerial Vehicles (MAVs). The model is based on the YOLOv3-tiny architecture with advanced DSConv layers for efficiency. The dataset comprised 5625 images categorized into four types of corrosion: nubby, bar, exfoliation, and fastener. Several enhancements were implemented, such as the convolutional block attention module (CBAM), three-scale object detection, focal loss, and an improved spatial pyramid pooling (SPP) module. The model achieved a mean average precision (mAP) of 84.96% and outperformed other models like YOLOv2-tiny, YOLOv3-tiny, YOLOv4-tiny, Single-shot object detection (SSD) network, and RetinaNet in terms of accuracy rate, model size, and speed of detection. Ahuja et al. proposed an optimized deep-learning framework for corrosion detection using image segmentation in their study [12]. Their research primarily aims to predict the course of corrosion based on the relationship between the shape and size of corrosion pits and their conditions. They employed a Residual UNet model together with a Bidirectional Convolution LSTM (BCLSTM) for the detection and classification of pitting corrosion. Furthermore, image processing techniques were used to explore the corrosion's varied color and textural properties. The study was implemented in the MATLAB platform and utilized SEM image-based datasets. The experimental findings demonstrated an accuracy rate of 94%, showcasing the suitability of the proposed strategy for corrosion detection. In a separate study, Forkan et al. [13] introduced CorrDetector, a structural corrosion detection framework developed for drone imagery. Their framework is based on CNN models and is aimed at detecting structural corrosion in civil engineering. Their dataset comprised 573 high-resolution photos capturing unique views and lighting situations relevant to telecommunication towers. The results of their model are impressive, with an accuracy of 92.5% and an F1-score of 98%, showing significant improvements over existing approaches. Comparative studies highlighted the enhanced performance features of the proposed model compared to contemporaneous state-of-the-art.

In their study [14], the researchers aimed to develop a robust system for accurately detecting and assessing damage and corrosion in civil infrastructure using machine learning (ML) and image processing techniques. The final dataset consisted of 1300 images, each with dimensions of 4864×3648 pixels. The process involved initial image processing followed by data augmentation, which was applied sixteen times to improve model generation. The authors utilized a modified deep hierarchical CNN (HCNN), CycleGAN, and U-Net architecture, as well as convolutional neural network, U-Net, along with various tools and functions for training the data. Additionally, they used conditional random field (CRF)-based Fusion and Multi-Level Feature Aggregation (MLFA) to enhance predictions and

learning during training. The model was then compared against industry standards such as PSPNet, DeepLab, and SegNet. The results were promising, with the model achieving a global accuracy of 98.9%, class average accuracy of 93.1%, Mean IOU of 87.8%, Precision of 84.9%, Recall of 81.8%, and an F-score of 83.3%. Despite these positive outcomes, the authors encountered challenges including the limited diversity in training datasets, environmental conditions and camera angles affecting model accuracy, the requirement for specialized knowledge in ML and image processing, as well as the high costs associated with acquiring and maintaining the necessary equipment for ML-based damage detection.

Nash et al. [15] conducted a study on corrosion detection using deep learning. In their work, they proposed a method based on deep learning for identifying corrosion regions through pixel-level segmentation. They also introduced three Bayesian variations to provide uncertainty estimates and confidence levels for each pixel, aiming to enhance decision-making. The study involved a dataset of 225 photographs from an industrial site, captured in visible light using a consumer digital SLR camera and saved as compressed JPEGs. The dataset was manually labeled by experts with per-pixel “corrosion” and “background” labels. Due to the limited size of the dataset, the researchers used 10-fold cross-validation for training. According to the results, all three model variations achieved higher accuracy in corrosion detection compared to human-level performance, with the maximum F1-score reaching 93%.

The study referenced in [16] presented an automated solution for addressing corrosion detection challenges. They developed RustSEG, a corrosion segmenter, using a CNN model and a “Classify, localize, refine” approach, along with heatmaps from Grad-CAM++ to pinpoint corrosion locations. However, Grad-CAM++ produced false positive regions for classification, which the authors addressed by integrating heatmaps with final pixel-level segmentation and refining the mask using a CRF. RustSEG achieved promising results without the need for per-pixel labeled datasets. To prevent overfitting, data augmentation techniques were used. The training dataset consisted of 7000 corrosion and 9000 non-corrosion images from publicly available photos on [flickr.com](https://www.flickr.com) (accessed on 10 July 2024), and an additional balanced set of 600 images per category for testing, along with 1200 expertly labeled images from various platforms. The model achieved an accuracy of 77.83%, an Area Under the Curve (AUC) of 86.17%, a precision rate of 81.27%, a recall rate of 72.33%, and an F1-score of 76.54%. For the Flickr Test Dataset, the model’s accuracy was 86.81%, with an AUC of 94.82%, a precision of 89.38%, a recall of 83.86%, and an F1-score of 86.53%. Additionally, the researchers provided an interactive platform to ensure accessibility for all users.

A study referenced in [17] explored the potential of secure coatings based on two-dimensional materials. These coatings are known for their strong bonding with metals and their ability to be functionalized. The study highlights the success of graphene coatings and mentions the exploration of other two-dimensional materials such as hexagonal boron nitride and molybdenum disulfide as promising alternatives. However, limited experimental data makes training classifiers challenging and often leads to overfitting. To tackle this issue, the article investigates the use of deep learning data augmentation methods, specifically variation autoencoder (VAE) and generative adversarial network (GAN), to create synthetic electrochemical data resembling the training data. These techniques were applied to a few-layered graphene on copper surfaces. The results show that the synthetic data generated by GAN outperformed VAE in terms of neural network system performance, achieving an accuracy of 83%–85% compared to VAE’s 78%–80%. However, when using XGBoost, VAE-generated data performed better with a 90% accuracy compared to GAN’s 84%–85%. In conclusion, the study illustrates that synthetic data generated using VAE and GAN models can be valuable for training machine learning models focused on developing corrosion-resistant two-dimensional coatings.

A study by Lemos et al. [18] used unmanned aerial vehicles (UAVs) and artificial intelligence to automatically detect corrosion on metal sandwich panels of large industrial buildings. The dataset includes 8400 high-resolution aerial images of metallic roofing systems of industrial facilities in northern Portugal. The images were captured at high sunlight exposure from a height of 12–15 m using a drone. The data was augmented using techniques like flipping and rotating before being divided into three parts for training, validation, and testing, respectively. The study used a deep learning algorithm called Mask R-CNN for instance segmentation. After conducting a sensitivity analysis, the results showed a precision of 85.8% and recall of 84.0% for all identified corrosion instances. In another research by Guzmán-Torres et al. [19], machine learning and computer vision techniques were used for concrete visual inspection and damage detection in real-time. The study aimed to demonstrate how machine learning and computer vision can be used to identify and analyze complex patterns of structural damage, with a focus on early detection of corrosion in concrete structures. The initial dataset included 159 images and was augmented to a total of 790 images. The training process involved using transfer learning with medium- and high-resolution training images, implementing the YOLOv3 algorithm based on Darknet-53, and applying data augmentation techniques to generate a robust model. By experimenting with 30 hyperparameter combinations, the authors identified optimal values for learning rate, momentum, patience, and weight decay to address overfitting. With the adjusted model, the study achieved a high precision of 82.12% in detecting corrosion damage.

A similar study in [20] addressed the continued dominance of fossil fuels in the global energy supply, despite significant investments in renewable energy. It specifically focuses on the corrosion-related failures, particularly pitting corrosion, in offshore oil and gas pipelines. To predict the maximum depth of pitting corrosion in these pipelines, the paper briefly covers the use of Generalization and Generalization-Memorization models. These models are trained using various variables, including soil properties and types of protective coating on the pipes. Deep neural networks are utilized, and the mean squared error of prediction for training data is 0.0055, while for test data it is 0.0037. These results indicate that deep learning models outperform empirical, and hybrid models used in earlier experiments with the same dataset. In summary, the deep learning model proposed in this study has the potential to predict pipeline failures caused by external corrosion, thereby enhancing the security and reliability of oil and gas infrastructure.

In the study referenced in [21], researchers tackled the challenge of detecting corrosion, a task that is risky and difficult for humans and often yields results with poor accuracy. Moreover, it is costly. They proposed an alternative approach for automated corrosion detection on steel structures. The methodology included data preprocessing, training, and hyperparameter tuning. By leveraging drone technology and CNN methodologies, they employed the MobileNet model for setting CNN hyperparameters. The relatively small dataset comprised 200 corrosion images collected from Google Images and ITS, which were split into 85% for training and 15% for testing. Data augmentation was used to increase the number of images, and data labeling was applied to assist CNN in recognizing corrosion objects. The training process updated the CNN model parameters until optimal results with a slight loss were achieved. The study successfully reached its goals with a total loss of 1.673 and an accuracy of 84.66%. The authors made several recommendations for future work to enhance the detection accuracy of the CNN model. These encompass increasing the number and diversity in the dataset by incorporating various colors, texture variations, and environmental conditions such as wind speed, humidity, and temperature. Moreover, optimizing crucial parameters such as batch size, drone speed, learning rate, and observational distance can substantially improve corrosion detection accuracy. Tan et al. [22] proposed a multitask deep learning approach called detection and segmentation network (DSNet) for corrosion detection and segmentation. The approach achieved the

highest detection precision, recall, and F1-score, with values of 94.3%, 98.1%, and 96%, respectively. Additionally, it achieved the best segmentation results with a mean Average Precision (mAP) of 0.786. The dataset used in the study was obtained from a public repository, and the focus of the approach was on corrosion detection from bolts.

[Table 1](#) provides a summary of the related work, including the methods used, dataset sizes, and results.

Table 1: Summary of the literature review

Study	Technique/s	Dataset size	Results
[2] 2018	Color space transformation to HSI	–	Identify and measure the corrosion level on the pipeline surface with high accuracy.
[3] 2019	CNN model based on Alexnet alongside an overlap-scanning sliding window algorithm	A total of 330 raw images (4032 × 3024) of hull structural plate corrosion damage are collected under different conditions.	Accuracy of 98.9%.
[4] 2019	A custom CNN	60,000 images	The model achieved a classification accuracy rate of 98.8%.
[5] 2020	Support Vector Machine (SVM), linear population size reduction (LSHADE) Metaheuristic	213 image	Accuracy rate of 91.80%, a precision of 91%, a recall of 94%, a negative predictive value of 93%, and an F1-score of 92%.
[6] 2020	ANN, support vector regression (SVR), random forests (RF), and cascade forests (cForest)	409 LAS corrosion samples	The weighted combination model assigns different values to the base tree models to weaken the influences of low accuracy tree.
[7] 2020	Comparing the deep learning models for corrosion detection: FCN, U-Net, and Mask R-CNN.	Contains 116 images of various resolutions	Proposed models exhibited superior accuracy and efficiency as compared to traditional approaches
[8] 2021	Texture-based, using CNN; color-based, using color algorithms; and the use of a single shot detector using transfer learning	Real world images	The use of single-shot detectors yielded better results as compared to other methods, making them more apt for real-world applications.

(Continued)

Table 1 (continued)

Study	Technique/s	Dataset size	Results
[9] 2021	Used a multi-teacher/single-student strategy combined with an ensemble learning and a new knowledge distillation version	Dataset consisting of $1.3 \cdot 10^4$ 320×240 images of various fuselage parts from several aircraft types, brands, and service life	The geometric student distillation model performed best, detecting 100% of the “moderate corrosion” class and outperforming other metrics like accuracy, recall, precision, and F1-score. An accuracy of 92.2%.
[10] 2021	DenseNet a pretrained neural network architecture	210 images	mAP 84.96%, AMCD model outperformed other models like YOLOv2-tiny, YOLOv3-tiny, YOLOv4-tiny, SSD, and RetinaNet. An accuracy rate of 94%.
[11] 2021	Deep Learning-based Metallic Corrosion Detector (AMCD)	5625 images	Accuracy of 92.5% and F1-score of 98%. Global Accuracy of 98.9%, Class Average Accuracy of 93.1%, Mean IOU of 87.8%, Precision of 84.9%, Recall of 81.8%, and an F-score of 83.3%.
[12] 2021	Residual UNet model with a Bidirectional Convolution LSTM	–	The maximum F1-score of 93%.
[13] 2021	CNN	573 high-resolution images	Model’s accuracy at 86.81%, with an AUC of 94.82%, a precision of 89.38%, a recall of 83.86%, and an F1-score of 86.53%.
[14] 2022	Deep Hierarchical CNN, CycleGAN, U-Net Architecture.	1300 images	
[15] 2022	Three deep learning Bayesian variations	225 images	
[16] 2022	CNN model	Flicker dataset: 1600 images, corrosiondetector dataset: 1200 images	

(Continued)

Table 1 (continued)

Study	Technique/s	Dataset size	Results
[17] 2022	Namely variation autoencoder (VAE) and generative adversarial network (GAN)	The size of the dataset is 49. Synthetic data generated by VAE and GAN	GAN data performance (83%–85% accuracy), VAE data performance (78%–80% accuracy). But, VAE data performed better (90%) than GAN data (84%–85%) with XGBoost.
[18] 2023	Deep learning algorithm called Mask R-CNN	The database contains 8400 images, with a 2000×2000 dimension	Precision of 85.8% and recall of 84.0%.
[19] 2023	Implemented the YOLOv3 algorithm based on Darknet-53	The initial dataset with 159 images. Augmented dataset with 790 images	Precision of 82.121%.
[20] 2023	Generalization and Memorization models	–	Mean squared error of prediction of 0.0055 in training data and 0.0037 in test data.
[21] 2023	Deep learning: CNN with the MobileNetV1 SSD	200 corrosion images	Total loss of 1.673 and an accuracy of 84.66%.
[22] 2024	DSNet a deep learning approach	1441 images of size 640×640	Precision, recall and F1-score of 94.3%, 98.1% and 96%. mAP score 0.786.

3 Proposed Techniques

The diagram in Fig. 1 illustrates the methodological steps that were utilized in the proposed study.

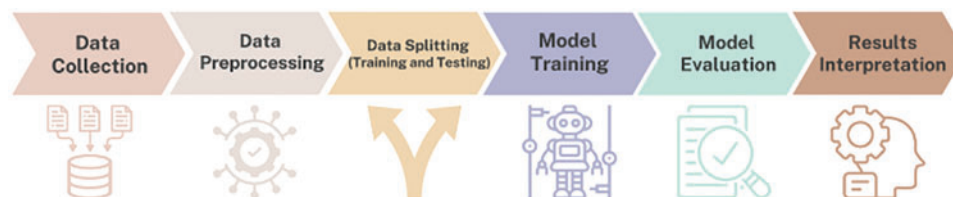


Figure 1: Methodological steps

3.1 You Only Look Once–Version 8 (YOLOv8)

YOLOv8 is the latest version of the YOLO model, known for its accuracy in important computer vision tasks such as object detection, image classification, and instance segmentation. Developed by Ultralytics, YOLOv8 is an improvement on YOLOv5, featuring enhancements to the model's structure and usability. YOLO was first introduced in 2015 by Joseph Redmond, and it quickly became known for its ability to balance accuracy and computational efficiency, making advanced computer vision accessible to many developers. YOLOv8 continues this tradition of high accuracy, as confirmed by

metrics from sources like Microsoft COCO, while still being trainable on a single GPU. YOLOv8 stands out with its unique architecture, consisting of two main components: the Backbone and Head. The model uses an anchor-free approach and separates the head to handle detection, classification, and regression tasks independently, allowing for specialized optimization for each function and improving the model's accuracy [23].

3.2 Convolutional Neural Networks (CNNs)

CNNs are a type of deep learning model specifically designed to analyze structured grid data, such as images and videos. They are widely used in computer vision tasks. The fundamental building blocks of CNNs are convolutional layers, which apply filters or kernels to input data and extract relevant information like edges, patterns, and formations in images. The resulting feature maps are then down sampled using pooling layers to minimize computing complexity, reduce spatial dimensions, and improve the network's resilience to input changes. CNN models create connections between neurons in successive layers and typically include one or more fully connected layers for handling classification or regression problems. Activation functions, such as Rectified Linear Units (ReLUs), introduce non-linearity into the network and aid the learning process by enabling the network to recognize complex patterns. CNNs are trained using optimization techniques like gradient descent and back-propagation, in which network weights and biases are adjusted to minimize a specified loss function. Normalization and augmentation are common data preprocessing strategies used to maximize training and improve the network's generalization capacity. CNNs are also used in transfer-learning paradigms, which reduce training time and data requirements by optimizing pre-trained models for specific tasks. CNNs represent a revolutionary force in computer vision and have significantly advanced the discipline. They are still the focus of ongoing research and development, with their basic function in contemporary image recognition systems serving as evidence of their lasting significance [24].

3.3 EfficientNetB0

EfficientNetB0 is a series of CNN models designed to efficiently utilize computational resources while delivering high accuracy. What makes it unique is its use of a compound coefficient to effectively scale up CNNs across depth, width, and resolution in a balanced manner. The initial model for this family of models is the EfficientNetB0 model, which combines squeeze-and-excitation blocks with inverted bottleneck residual blocks from MobileNetV2. The architecture's scaling technique allows for uneven scaling of network dimensions, setting it apart from existing CNNs and improving efficiency and speed. One of the key concepts of EfficientNetB0 is the use of compound scaling, employing a coefficient to systematically scale the model's dimensions rather than arbitrarily scaling these factors. This results in improved performance without increasing computational cost. EfficientNetB0 can be used practically as a feature extractor in transfer learning processes. For instance, custom layers tailored to the number of classes in a new dataset can replace the top layer of a model pre-trained on datasets like ImageNet [25].

3.4 Dataset Description

The dataset for the corrosion detection project was created using two main sources: an open-source website [26] and contributions from an author of a relevant study [4]. It consists of a total of 1000 images. These images were carefully split to meet the requirements for a robust deep learning project: 70% of the images are for training the models, 20% for validation to fine-tune model parameters, and the remaining 10% for testing to evaluate the models' performance. This structured division aims to enhance the accuracy of the models (CNN, YOLOv8, and EfficientNet0) in distinguishing

between corroded and non-corroded surfaces. Fig. 2 shows a sample of random images selected from the dataset.

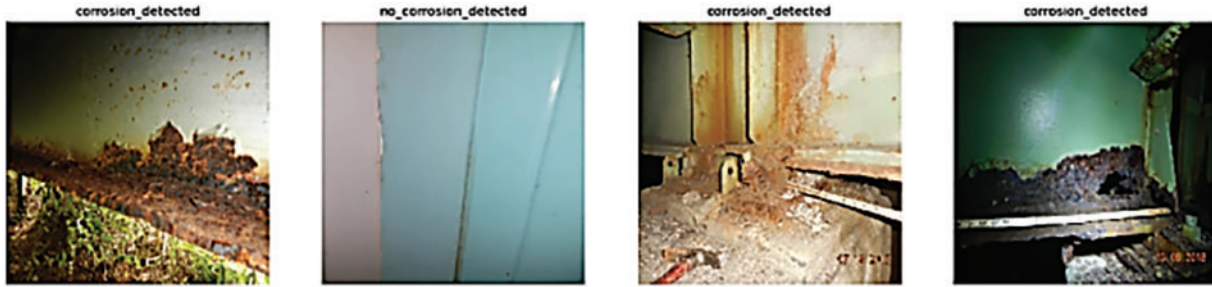


Figure 2: Sample of the dataset

3.5 Performance Evaluation

The model's predictive accuracy will be rigorously assessed using a suite of key performance indicators. Accuracy measures the fraction of predictions that the model made correctly out of the total number of predictions. The Loss Value will also be carefully examined to capture how frequently the model makes errors when tested against the dataset. Precision is important for determining the model's reliability in identifying images that show corrosion. Additionally, the recall metric will be used to evaluate how well the model identifies instances of corrosion in the images. Furthermore, the F1-score, which is a harmonic mean of precision and recall, provides a balanced performance evaluation considering both types of prediction errors. These evaluation criteria offer a detailed and stringent appraisal of the model's performance. The Eqs. (1)–(4) use TP (true positive), TN (true negative), FP (false positive), and FN (false negative) as of [27–30].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F1 - score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

4 Results and Discussion

4.1 CNN

The lines labeled 'Accuracy' and 'val_accuracy' in Fig. 3 represent the model's accuracy on the training and validation sets (y -axis), respectively in terms of number of epochs given on the x -axis. These lines rise quickly and level off close to 100%, indicating high accuracy across both datasets. This suggests that the model is not overfitting and can effectively generalize to new data, as seen by the convergence of training and validation accuracy near 100%.

In Fig. 4, the lines 'loss' and 'val_loss' represent the training loss and validation loss (y -axis), respectively in terms of number of epochs given on the x -axis. The validation loss is calculated on a separate set of data that the model did not see during training, while the training loss represents the model's error on the training dataset. Both loss metrics sharply fall and converge to a low value

near zero, indicating efficient learning and effective generalization without overfitting. Examining the hyperparameters, 'adam' is used as the optimizer, 'binarycrossentropy' as the loss function, and a batch size of 22. As shown in Table 2, the CNN model performed flawlessly in every metric. It accurately classified each instance with 100% accuracy. Additionally, precision reached 100%, indicating no false positives, and all predictions of the target class were accurate. Similarly, recall was 100%, meaning no instances of the target class were missed. The F1-score is 100%, showing ideal balance between recall and precision and demonstrating the model's promising performance.

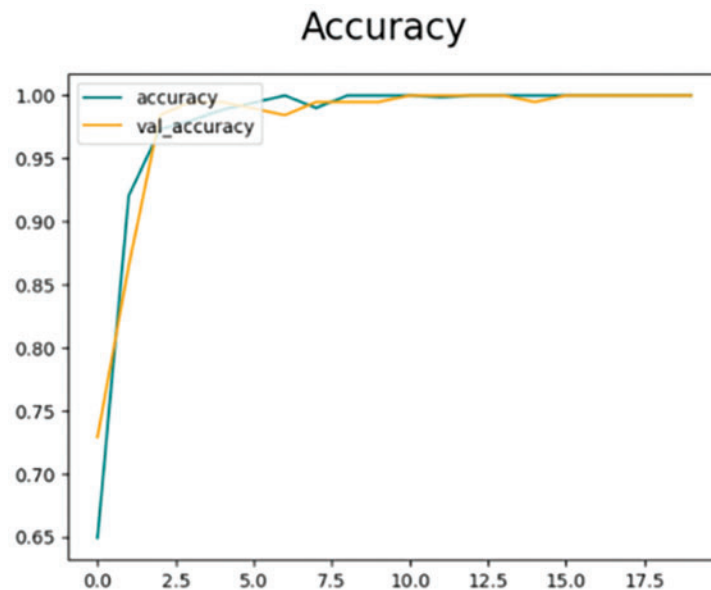


Figure 3: CNN accuracy curve

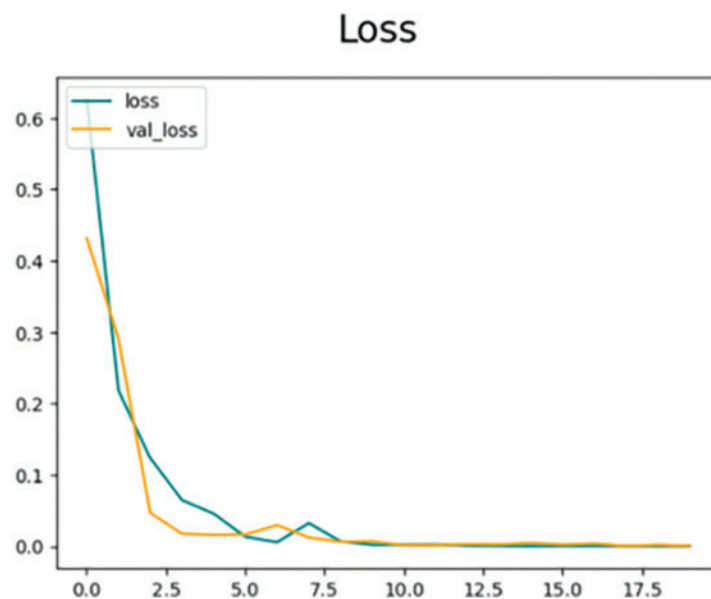


Figure 4: CNN loss curve

Table 2: CNN evaluation results

	Accuracy	Precision	Recall	F1-score
CNN	100%	100%	100%	100%

4.2 YOLOv8

The left graph shows the training loss, indicated by the orange dotted line, which smooths the trend. The actual loss values throughout epochs are represented by the blue line. The model seems to be reaching a stable learning point from the training data. This is evidenced by the first large drop, indicating quick learning, followed by a steadier descent and then a leveling out. The top right graph displays the validation loss, which shows a significant initial decline leveling out as epochs increase, like the training loss. Regarding hyperparameters, the optimizer function used is ‘adamw’ with a batch size of 44. Although the actual loss values (blue) vary, the smoothing trend (orange) remains generally constant, indicating consistent performance on the validation set. In Fig. 5, the graph on the bottom left shows the top 1 accuracy measure on the training data. The blue line represents the real data, exhibiting significant changes as training progresses. The model may have reached a point during training where it no longer performs well in generalizing, as seen by the slight decrease in top 1 accuracy shown by the smoothed orange line. The graph on the bottom right displays the top 5 accuracy metrics, with the blue line demonstrating that the true class is almost always within the top 5 predictions made by the model. This shows that the model consistently achieves near-perfect top 5 accuracy with minimal fluctuation across the epochs (Fig. 5). Table 3 illustrates the YOLOv8 model’s exceptional performance in detecting corrosion, with an accuracy of 95%, ensuring that almost all its predictions are correct. Its precision stands at an impeccable 100%, meaning it has zero false positives; no non-corroded instances are wrongly identified as corrosion. While it misses 10% of actual corroded instances, as indicated by the 90% recall, the model still captures most corrosion cases. The F1-score of 94.74% confirms the model’s effectiveness in balancing precision and recall, making it a reliable tool for corrosion detection.

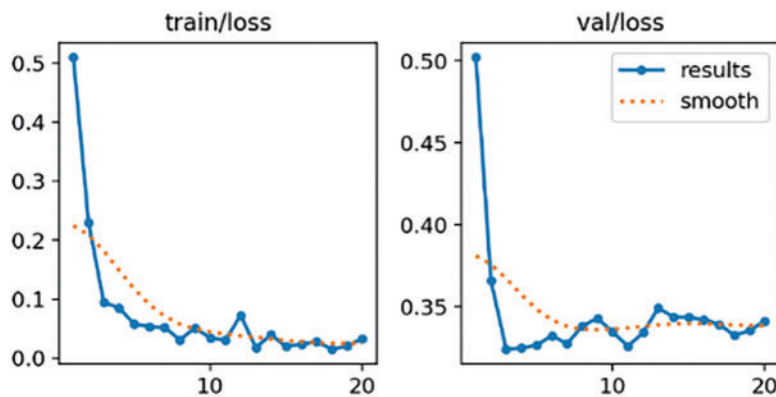


Figure 5: (Continued)

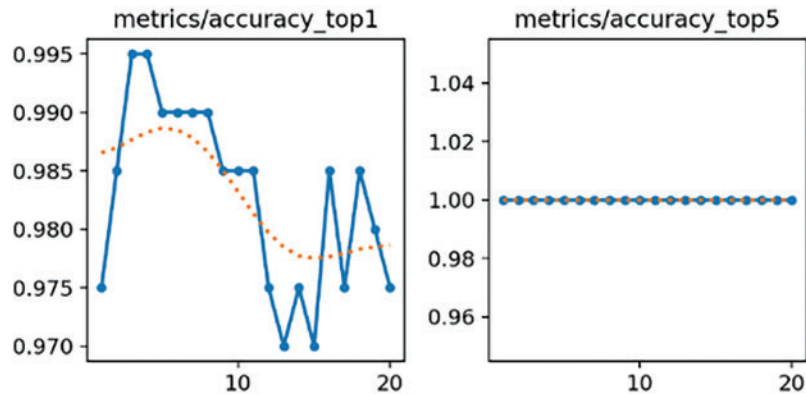


Figure 5: YOLOv8 performance curves

Table 3: YOLOv8 evaluation results

	Accuracy	Precision	Recall	F1-score
YOLOv8	95%	100%	90%	94.74%

4.3 EfficientnetB0

In Fig. 6, we can observe the training progress of the EfficientNetB0 model over 30 epochs. It shows consistent improvement and eventual perfection. Initially, the model was trained for 20 epochs, reaching a certain level of accuracy. To further enhance its performance, it underwent an additional 10 epochs of training (Fig. 7). The graphs demonstrate a clear contrast between the training and validation accuracy and loss, which eventually converge. The hyperparameters used include the ‘adam’ optimizer and ‘categoricalcrossentropy’ loss function, with a batch size of 22. This convergence indicates the model’s increasing generalizability and stability over time. As shown in Table 4, the model ultimately achieved a flawless 100% score across accuracy, precision, recall, and F1-score. This remarkable achievement underscores the model’s ability to classify every image in the test set correctly, with precision signifying no false positives and recall indicating that every positive instance was captured. The F1-score of 100% demonstrates a perfect balance between precision and recall, which is an ideal outcome for any classification model. These impressive results suggest that the additional epochs were highly beneficial, allowing the model to refine and adjust its parameters to better understand the complexities of the dataset.

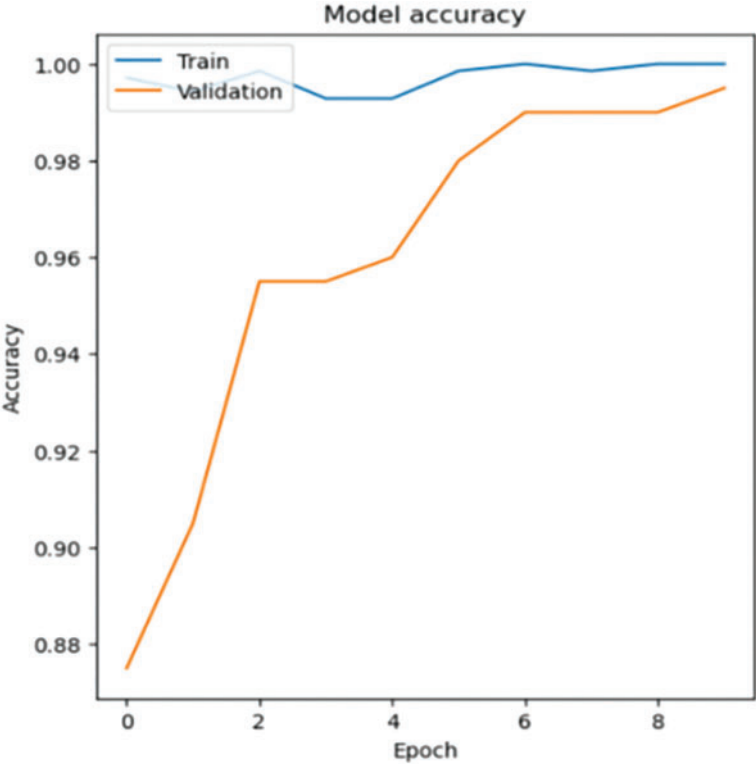


Figure 6: EffcientNetB0 accuracy curve

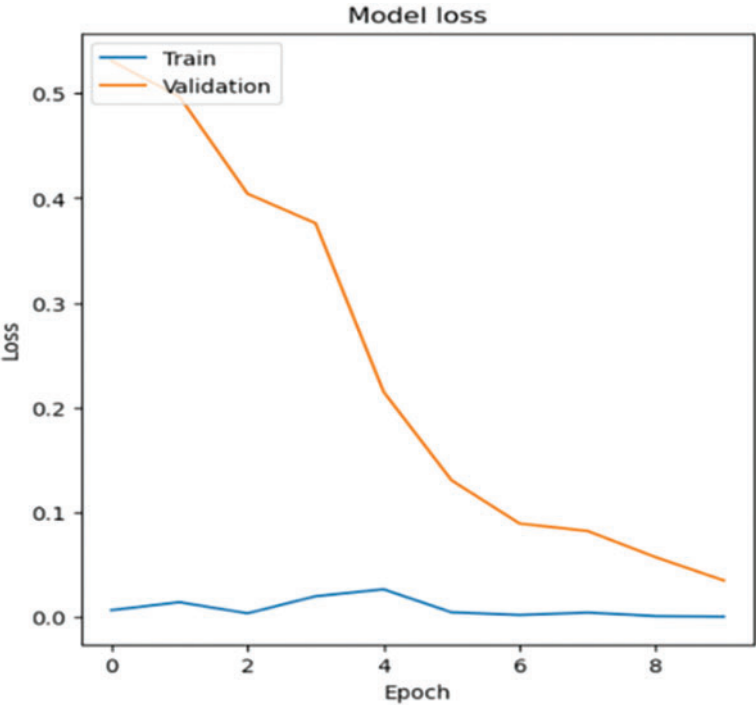


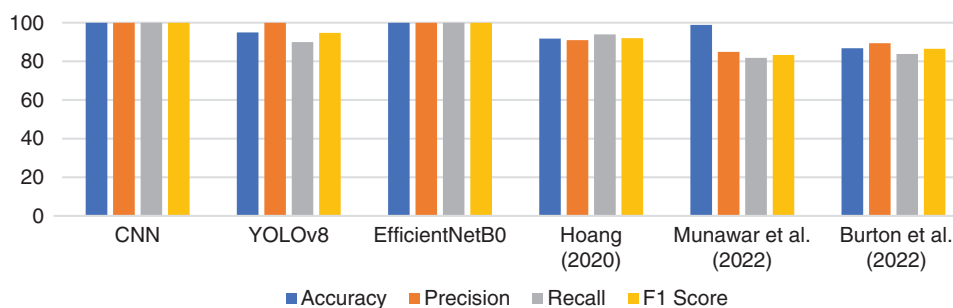
Figure 7: EffcientNetB0 loss curve

Table 4: EffcientNetB0 evaluation results

	Accuracy	Precision	Recall	F1-score
EfficientNetB0	100%	100%	100%	100%

4.4 Comparison with State-of-the-Art

The proposed scheme has been compared to similar studies and schemes in the literature from 2021 to 2023. For the comparison, the schemes were selected based on the algorithms used, dataset, and target industrial images. Fig. 8 illustrates the comparison.

**Figure 8:** Comparison with state-of-the-art [5,14,16]

It is evident that the proposed schemes CNN and EfficientNetB0 outperform the state-of-the-art in all four metrics. The studies chosen for comparison are SVM [5], Deep CNN with transformers [14], and CNN [16] for common grounds. The key differences lie in the set of hyperparameters, methodology (data augmentation), and feature selection used. The scheme in [16] exhibited relatively lower performance compared to all the proposed schemes, including CNN, YOLOv8, and EfficientNetB0. The scheme in [5] exhibited similar behavior, except it had 4% better recall compared to YOLOv8. The scheme in [14] underperformed in all metrics, except that the accuracy was comparable to the proposed schemes, just 1.1% lower compared to CNN and EfficientNetB0, while being 3.9% better than YOLOv8.

4.5 Discussion

The research focuses on utilizing deep-learning methods to identify corrosion in industrial images. The study uses an augmented dataset compiled from open sources and local images. The CNN and two pre-trained models, YOLOv8 and EfficientNetB0, show promise compared to existing studies. The learning rate is adjusted during training to manage the model's learning pace, controlling the number of steps taken in the learning process. Additionally, variable-sized batches are used during each iteration to improve weight updating. Techniques such as mini-batch gradient descent and stochastic gradient descent are used to achieve this. Intelligent tuning of the learning rate, typically set around 0.001, promotes progressive learning and helps to avoid overfitting. Adding a momentum parameter enhances the model's resistance to getting stuck in local minima, preventing stagnation and improving its learning dynamics. These measures contribute to the model's robust validation process and its ability to generalize effectively. The complexity of the deep learning models is primarily expressed in terms of learning rate, batch size, number of epochs, and training time, which are typically performed offline.

Testing, on the other hand, happens in real-time [28]. In this case, the method shows polynomial complexity in both time and space, with the size of the image (in 3 dimensions) being expressed as $O(x * y * z)$. As for the study's limitations, it primarily focuses on detecting corrosion in images from the oil and gas industry, encompassing a wide range. However, it does not consider other industries such as manufacturing. Additionally, environmental conditions and camera angles may impact the model's accuracy. In future studies, other sectors could also be joined. Furthermore, additional deep learning models such as YOLOv9 and Federated learning could be explored for further improvement and understanding [31,32].

5 Conclusions

The study aims to detect corrosion using deep learning and transfer learning models. Our comprehensive evaluation of CNN, YOLOv8, and EfficientNetB0 models in corrosion detection demonstrates impressive results. Both models achieved perfect scores in accuracy, precision, recall, and F1-score, showcasing their ability to accurately identify instances of corrosion without overfitting. This marks a significant technological advancement in detection, providing reliable solutions for industrial challenges. The YOLOv8 model achieved 95% accuracy and 100% precision, but its recall was slightly lower. However, it displayed a strong capability in minimizing false positives, which is essential for practical applications. These results underscore the model's reliability in detecting corrosion with high precision. These findings highlight the potential of deep learning techniques in transforming corrosion detection and setting new standards for accuracy and reliability. We also discovered that the strategic use of these models not only advances the field but also indicates their readiness for real-world application, promising to improve safety and efficiency in industries facing corrosion challenges. Moreover, it may be beneficial to explore hybrid intelligent systems with diverse datasets to further advance research in this area.

Acknowledgement: Authors like to acknowledge CCSIT for using the resources during the study.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Conceptualization, Atta Rahman and Sunday Olatunji; Data curation, Fatimah Albaik and Razan Sharaf; Formal analysis, Mehwash Farooqui, Sara Waslallah Althubaiti and Hina Gull; Investigation, Sunday Olatunji; Methodology, Atta Rahman, Latifa Alsuliman, Zainab Alsaif, Fatimah Albaik and Cadi Alshammari; Project administration, Hina Gull; Software, Latifa Alsuliman, Zainab Alsaif, Fatimah Albaik, Cadi Alshammari and Razan Sharaf; Supervision, Atta Rahman, Mehwash Farooqui and Sunday Olatunji; Validation, Sara Waslallah Althubaiti and Hina Gull; Visualization, Sara Waslallah Althubaiti; Writing—original draft, Latifa Alsuliman, Zainab Alsaif, Cadi Alshammari and Razan Sharaf; Writing—review & editing, Atta Rahman and Mehwash Farooqui. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data publicly available at GitHub as given in [26].

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

- [1] S. S. Aljameel *et al.*, “Oil and gas pipelines leakage detection approaches: A systematic review of literature,” *Int. J. Saf. Secur. Eng.*, vol. 14, no. 3, pp. 773–786, 2024. doi: [10.18280/ijssse.140310](https://doi.org/10.18280/ijssse.140310).
- [2] V. Bondada, D. K. Pratihar, and C. S. Kumar, “Detection and quantitative assessment of corrosion on pipelines through image analysis,” *Procedia Comput. Sci.*, vol. 133, pp. 804–811, Jan. 2018. doi: [10.1016/j.procs.2018.07.115](https://doi.org/10.1016/j.procs.2018.07.115).
- [3] Y. Yao, Y. Yang, Y. Wang, and X. Zhao, “Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network,” *Appl. Ocean Res.*, vol. 90, Sep. 2019, Art. no. 101823. doi: [10.1016/j.apor.2019.05.008](https://doi.org/10.1016/j.apor.2019.05.008).
- [4] B. T. Bastian, J. N. S. K. Ranjith, and C. V. Jiji, “Visual inspection and characterization of external corrosion in pipelines using deep neural network,” *NDT & E Int.*, vol. 107, Oct. 2019, Art. no. 102134. doi: [10.1016/j.ndteint.2019.102134](https://doi.org/10.1016/j.ndteint.2019.102134).
- [5] N. D. Hoang, “Image processing-based pitting corrosion detection using metaheuristic optimized multilevel image thresholding and machine-learning approaches,” *Math. Probl. Eng.*, vol. 2020, pp. 1–19, 2020. doi: [10.1155/2020/6765274](https://doi.org/10.1155/2020/6765274).
- [6] Y. Zhi, T. Yang, and D. Fu, “An improved deep forest model for forecast the outdoor atmospheric corrosion rate of low-alloy steels,” *J. Mater Sci Technol.*, vol. 49, pp. 202–210, Jul. 2020. doi: [10.1016/j.jmst.2020.01.044](https://doi.org/10.1016/j.jmst.2020.01.044).
- [7] I. Katsamenis, E. Protopapadakis, A. Doulamis, N. Doulamis, and A. Voulodimos, “Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation,” in *Lecture Notes in Computer Science*, Cham, San Diego, CA, USA: Springer, Nov. 5–7, 2020, vol. 12509, pp. 160–169. doi: [10.1007/978-3-030-64556-4_13](https://doi.org/10.1007/978-3-030-64556-4_13).
- [8] A. Matthaïou, G. Papalambrou, and M. S. Samuelides, “Corrosion detection with computer vision and deep learning,” in *Developments in the Analysis and Design of Marine Structures*, 1st ed. USA: CRC Press, Dec. 2021, pp. 289–296.
- [9] K. Zuchniak, W. Dzwiniel, E. Majerz, A. Pasternak, and K. Dragan, “Corrosion detection on aircraft fuselage with multi-teacher knowledge distillation,” in *Lecture Notes in Computer Science*, Springer, Cham, 2021, vol. 12747, pp. 318–332. doi: [10.1007/978-3-030-77980-1](https://doi.org/10.1007/978-3-030-77980-1).
- [10] B. Brandoli *et al.*, “Aircraft fuselage corrosion detection using Artificial Intelligence,” *Sensors*, vol. 21, no. 12, Jun. 2021, Art. no. 4026. doi: [10.3390/s21124026](https://doi.org/10.3390/s21124026).
- [11] L. Yu, E. Yang, C. Luo, and P. Ren, “AMCD: An accurate deep learning-based metallic corrosion detector for MAV-based real-time visual inspection,” *J. Ambient Intell. Humaniz Comput.*, vol. 14, no. 7, pp. 8087–8098, 2023. doi: [10.1007/s12652-021-03580-4](https://doi.org/10.1007/s12652-021-03580-4).
- [12] S. K. Ahuja, M. K. Shukla, and K. K. Ravulakollu, “Optimized deep learning framework for detecting pitting corrosion based on image segmentation,” *Int. J. Perform. Eng.*, vol. 17, no. 7, Jul. 2021, Art. no. 627.
- [13] A. R. M. Forkan *et al.*, “CorrDetector: A framework for structural corrosion detection from drone images using ensemble deep learning,” *Expert. Syst. Appl.*, vol. 193, Feb. 2021, Art. no. 116461. doi: [10.1016/j.eswa.2021.116461](https://doi.org/10.1016/j.eswa.2021.116461).
- [14] H. Munawar, F. Ullah, D. Shahzad, A. Heravi, S. Qayyum and J. Akram, “Civil infrastructure damage and corrosion detection: An application of machine learning,” *Buildings*, vol. 12, no. 2, Feb. 2022, Art. no. 156.
- [15] W. Nash, L. Zheng, and N. Birbilis, “Deep learning corrosion detection with confidence,” *npj Mat. Degrad.*, vol. 6, no. 1, pp. 1–13, Mar. 2022. doi: [10.1038/s41529-022-00232-6](https://doi.org/10.1038/s41529-022-00232-6).
- [16] B. Burton, W. T. Nash, and N. Birbilis, “RustSEG-Automated segmentation of corrosion using deep learning,” May 2022. doi: [10.48550/arXiv.2205.05426](https://doi.org/10.48550/arXiv.2205.05426).
- [17] C. Allen *et al.*, “Deep learning strategies for addressing issues with small datasets in 2D materials research: Microbial corrosion,” *Front. Microbiol.*, vol. 13, Dec. 2022, Art. no. 1059123.
- [18] R. Lemos, R. Cabral, D. Ribeiro, R. Santos, V. Alves and A. Dias, “Automatic detection of corrosion in large-scale industrial buildings based on artificial intelligence and unmanned aerial vehicles,” *Appl. Sci.*, vol. 13, no. 3, Jan. 2023, Art. no. 1386.

- [19] J. J. A. Guzmán-Torres, F. J. Domínguez-Mota, W. Martínez-Molina, M. Z. Naser, G. Tinoco-Guerrero and J. G. Tinoco-Ruiz, "Damage detection on steel-reinforced concrete produced by corrosion via YOLOv3: A detailed guide," *Front. Built Environ.*, vol. 9, no. 1, Mar. 2023, Art. no. 1144606.
- [20] B. Akhlaghi, H. Mesghali, M. Ehteshami, J. Mohammadpour, F. Salehi and R. Abbassi, "Predictive deep learning for pitting corrosion modeling in buried transmission pipelines," *Process Saf. Environ. Prot.*, vol. 174, pp. 320–327, Jun. 2023. doi: [10.1016/J.PSEP.2023.04.010](https://doi.org/10.1016/J.PSEP.2023.04.010).
- [21] M. K. Effendi, B. Atmaja, A. Wahjudi, and D. B. Purwanto, "Automated corrosion detection on steel structures using convolutional neural network," *Int. J. Mech. Eng. Sci.*, vol. 7, no. 1, Mar. 2023, Art. no. 36.
- [22] L. Tan, X. Chen, D. Yuan, and T. Tang, "DSNet: A computer vision-based detection and corrosion segmentation network for corroded bolt detection in tunnel," *Struct. Control Health Monit.*, vol. 2024, 2024. doi: [10.1155/2024/1898088](https://doi.org/10.1155/2024/1898088).
- [23] R. Varghese and M. Sambath, "YOLOv8: A novel object detection algorithm with enhanced performance and robustness," in *Proc ADICS*, Chennai, India, 2024, pp. 1–6. doi: [10.1109/ADICS58448.2024.10533619](https://doi.org/10.1109/ADICS58448.2024.10533619).
- [24] M. Talha *et al.*, "Voting-Based Deep Convolutional Neural Networks (VB-DCNNs) for M-QAM and M-PSK signals classification," *Electronics*, vol. 12, no. 8, 2023, Art. no. 1913. doi: [10.3390/electronics12081913](https://doi.org/10.3390/electronics12081913).
- [25] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. ICML 2019*, Arlington, VA, USA, May 2019, pp. 10691–10700.
- [26] P. Sun, "Deep learning for automated corrosion detection," 2012. Accessed: Jan. 30, 2024. [Online]. Available: https://github.com/pjsun2012/Phase5_Capstone-Project
- [27] A. Das, S. Dorafshan, and N. Kaabouch, "Autonomous image-based corrosion detection in steel structures using deep learning," *Sensors*, vol. 24, no. 11, 2024, Art. no. 3630. doi: [10.3390/s24113630](https://doi.org/10.3390/s24113630).
- [28] M. A. A. Khan *et al.*, "Road damages detection and classification using deep learning and UAVs," in *2022 2nd Asian Conf. Innov. Technol. (ASIANCON)*, Ravet, India, 2022, pp. 1–6. doi: [10.1109/ASIANCON55314.2022.9909043](https://doi.org/10.1109/ASIANCON55314.2022.9909043).
- [29] N. Daoudi, M. Y. Haouam, L. Laimeche, I. Bendib, and M. Amroune, "Applications of machine learning in corrosion detection," in *Proc. PAIS*, Algeria, 2024, pp. 1–8. doi: [10.1109/PAIS62114.2024.10541125](https://doi.org/10.1109/PAIS62114.2024.10541125).
- [30] K. Mundada, M. Kulkarni, P. Sagar, R. Asegaonkar, and N. Mulgir, "Corrosion detection using deep learning and custom object detection," in *2nd Asian Conf. Innov. Technol. (ASIANCON)*, Ravet, India, 2022, pp. 1–5. doi: [10.1109/ASIANCON55314.2022.9909026](https://doi.org/10.1109/ASIANCON55314.2022.9909026).
- [31] R. S. Yuvaneshwari, S. T. Prakash, and K. Sudha, "A YOLOv8-based model for precise corrosion segmentation in industrial imagery," in *2024 3rd Int. Conf. Artif. Intell. Internet of Things (AIIoT)*, Vellore, India, 2024, pp. 1–6. doi: [10.1109/AIIoT58432.2024.10574659](https://doi.org/10.1109/AIIoT58432.2024.10574659).
- [32] S. K. Fondevik, A. Stahl, A. A. Transeth, and O. Ø. Knudsen, "Image segmentation of corrosion damages in industrial inspections," in *2020 IEEE 32nd Int. Conf. Tools Artif. Intell. (ICTAI)*, Baltimore, MD, USA, 2020, pp. 787–792. doi: [10.1109/ICTAI50040.2020.00125](https://doi.org/10.1109/ICTAI50040.2020.00125).