

Article

Classification and Regression of Pinhole Corrosions on Pipelines Based on Magnetic Flux Leakage Signals Using Convolutional Neural Networks

Yufei Shen  and Wenxing Zhou * Department of Civil & Environmental Engineering, The University of Western Ontario,
London, ON N6A 5B9, Canada

* Correspondence: wzhou@eng.uwo.ca

Abstract: Pinhole corrosions on oil and gas pipelines are difficult to detect and size and, therefore, pose a significant challenge to the pipeline integrity management practice. This study develops two convolutional neural network (CNN) models to identify pinholes and predict the sizes and location of the pinhole corrosions according to the magnetic flux leakage signals generated using the magneto-static finite element analysis. Extensive three-dimensional parametric finite element analysis cases are generated to train and validate the two CNN models. Additionally, comprehensive algorithm analysis evaluates the model performance, providing insights into the practical application of CNN models in pipeline integrity management. The proposed classification CNN model is shown to be highly accurate in classifying pinholes and pinhole-in-general corrosion defects. The proposed regression CNN model is shown to be highly accurate in predicting the location of the pinhole and obtain a reasonably high accuracy in estimating the depth and diameter of the pinhole, even in the presence of measurement noises. This study indicates the effectiveness of employing deep learning algorithms to enhance the integrity management practice of corroded pipelines.

Keywords: deep learning algorithm; convolutional neural network; pipeline; corrosion; pinhole; finite element analysis; magnetic flux leakage signal



Citation: Shen, Y.; Zhou, W. Classification and Regression of Pinhole Corrosions on Pipelines Based on Magnetic Flux Leakage Signals Using Convolutional Neural Networks. *Algorithms* **2024**, *17*, 347. <https://doi.org/10.3390/a17080347>

Academic Editors: Chih-Lung Lin, Bor-Jiunn Hwang, Shaou-Gang Miaou and Chi-Hung Chuang

Received: 24 June 2024
Revised: 1 August 2024
Accepted: 6 August 2024
Published: 8 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Pipelines are economical and safe means for transporting and distributing large volumes of oil and gas products across great distances [1]. However, pipeline failures can lead to severe consequences such as fatalities, economic losses, and environmental damages [2]. Among the various failure mechanisms threatening the structural integrity of pipelines, corrosion is a leading threat [3]. The in-line inspection (ILI) tool is widely used for detecting and measuring corrosions to assess pipeline integrity [4], and magnetic flux leakage (MFL) is the predominant ILI technology for identifying internal and external corrosion defects in both liquid and gas pipelines [5]. The Pipeline Operators Forum (POF) suggests that a single corrosion anomaly on the internal or external surface of a pipeline can be categorized into one of seven classes based on the ILI-reported corrosion length and width as illustrated in Figure 1 [6], i.e., axial slotting, circumferential slotting, axial grooving, circumferential grooving, general, pinhole and pitting, to better quantify the accuracy of ILI tools and facilitate the selection of appropriate models to evaluate the burst capacity of pipelines containing corrosion defects [6,7].

The estimation of the corrosion anomaly sizes based on MFL signals falls into the inverse modeling realm. The pipeline industry frequently employs the iterative inverse model because of its remarkable precision [8–10]. However, the iterative inverse model is computationally costly. Inverse models based on machine learning provide promising alternatives to sizing corrosion defects from MFL signals with high accuracy but are more computationally efficient compared to the iterative inverse models [11–13]. For instance,

Kandroodi et al. [14] fed the algorithm-estimated defect width and length, as well as the extracted peak-to-peak values from the MFL signals, into a Gaussian radial basis function neural network to predict the corrosion depth. Feng et al. [15] proposed an error adjustment methodology for reconstructing the corrosion profiles from MFL signals based on a radial basis function neural network. However, these models have certain limitations: the training of these models relies on manually chosen features from the MFL signals, potentially omitting valuable data necessary for accurately predicting the corrosion profile. Additionally, the majority of these techniques focus solely on one MFL signal dimension, neglecting the critical information present in the other two dimensions.

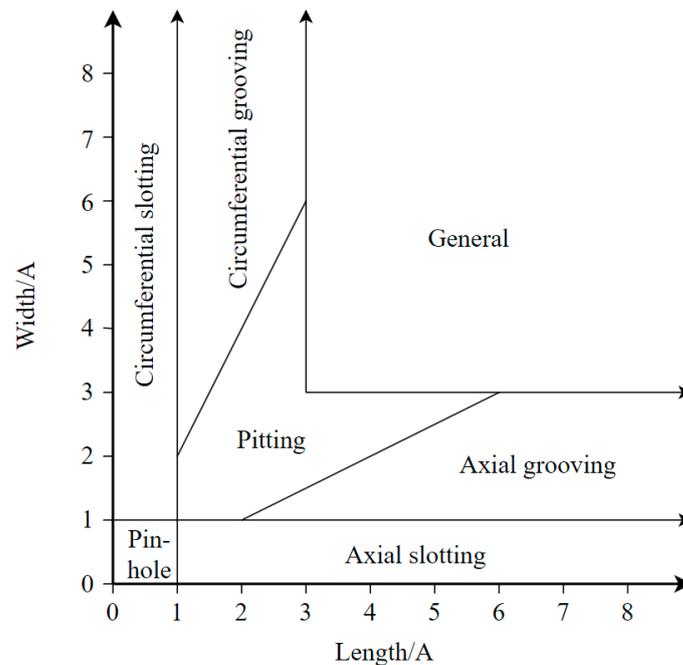


Figure 1. Seven corrosion anomaly categories are based on the anomaly length and width. Note: $A = \max\{10 \text{ mm, pipe wall thickness}\}$.

Over the past decade, convolutional neural networks (CNNs) have risen as a revolutionary deep learning algorithm, yielding remarkable results across various pattern recognition fields, including image and voice processing [16,17]. The CNN tool can automatically extract deep features from input data in place of manually designed feature extractors; therefore, it is well suited to solve the inverse modeling based on the MFL signals. The application of CNN in the inverse analysis of MFL signals has been reported in the literature. Lu et al. [18] introduced the visual transformation CNN, which predicts the length, width, and depth of anomalies by taking into account two components (i.e., circumferential and radial) of the MFL signal while disregarding the longitudinal component. Shen and Zhou [19] employed CNN to estimate the locations and dimensions of corrosions on steel pipelines based on three components of the MFL signals. Wang and Chen [20] proposed a CNN architecture comprising two main modules: one for anomaly classification and another for anomaly size regression. The input of the classification module consists of all three components of the MFL signal, and the outputs are seven anomaly categories, as illustrated in Figure 1. The regression module incorporates seven separate CNNs, each tailored to a specific anomaly type, predicting dimensions such as width, length, and depth of the defect.

According to a report from the European Gas Pipeline Incident Data Group (EGIG) [21], pinhole corrosion is identified as a major contributor to pipeline failures due to its potential to evolve from a small initial volume of metal loss into more severe metal loss, given the challenge in detection via pressure monitoring, particularly when the leaked products do not exceed ground level [22,23]. Despite the impact of pinholes on pipeline integrity, studies

that are specifically focused on detecting and sizing pinhole corrosions are limited in the literature, thus creating a knowledge gap. Furthermore, small corrosion anomalies such as pinholes may sometimes overlap with a relatively large corrosion area [23–25], creating a complex defect known in this study as a pinhole-in-general corrosion (PIC) defect. There are also few studies in the literature investigating the detection and sizing of PIC defects.

The objective of this study is to develop deep learning algorithms, i.e., CNN models, to classify and size pinholes and PIC defects based on a large set of MFL signals. The classification CNN model deals with six scenarios, i.e., pinholes on the internal or external pipe surface, general corrosions on the internal or external surface, and PIC defects on the internal or external surface. The signals for the pinholes and PIC defects identified by the classification model are then fed into the regression CNN model to predict the sizes and location of the pinhole and PIC defects. In practice, it is not necessary to apply the two CNN models sequentially. Depending on the specific application, data can be fed directly into the regression model if the classification has already been carried out through other means and the primary goal is to predict the size of the defect. Since real MFL signals are typically proprietary and not publicly accessible, the MFL signals used in this study are numerically generated by applying the three-dimensional (3D) magneto-static finite element analyses (FEA) using the commercial software COMSOL (Version 5.6). The three-dimensional (i.e., longitudinal, circumferential, and radial) MFL signal maps serve as inputs for training and validating CNN models to classify the defects and predict the sizes, as well as the circumferential and longitudinal positions, of pinholes on the pipeline. The impact of the measurement noises associated with the MFL signal on the accuracy of the regression CNN model is also investigated. This research fills a knowledge gap in the literature by applying CNN to detect, locate, and size pinholes and pinhole-in-corrosion defects. It is noted that accurately identifying the position of the pinhole within the general corrosion is crucial as the relative position of the pinhole within the general corrosion has a marked effect on the burst capacity of the PIC defect and, therefore, has strong implications for the accurate prediction of the burst capacity. This study demonstrates the effectiveness of CNN for the detection and sizing of pinhole corrosion and the viability of CNN for applications in pipeline integrity management practice.

The rest of the paper is structured as follows. Section 2 introduces the principles of the MFL technology and essential background information on the relationship between MFL signals and anomaly features. Section 3 introduces the proposed FEA model and the simulation parameters of MFL signals. Section 4 describes the structure of the developed CNN models and their application to the classification and regression of pinholes. The corresponding predictive accuracy of the CNN models is also discussed in Section 4. Section 5 summarizes the concluding remarks of the study.

2. Principles of MFL Technique

The magneto-statics FEA is employed to numerically generate the MFL signals in the study. The corresponding governing equation derived from Maxwell's equations is shown in Equation (1):

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{A} \right) = 0 \quad (1)$$

where $\nabla \times$ denotes the curl operation; μ represents the magnetic permeability of the ferromagnetic specimen (N/A^2), and \mathbf{A} is the magnetic potential vector. Equation (1) is solved under the specified boundary conditions to obtain \mathbf{A} , which is then used to obtain the magnetic flux density \mathbf{B} as follows [26,27]:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2)$$

The MFL technique relies on the measurement principle where a ferromagnetic material demonstrates MFL in the vicinity of a defect, as illustrated in Figure 2 [18]. The measurement procedure is as follows: during the ILI process, the MFL tool is propelled

through the pipeline, and magnets within the tool generate a magnetic field. In a perfect pipe, all magnetic flux remains within the pipe wall [28], resulting in a uniform magnetic field distribution [29]. However, in the presence of defects, some magnetic flux “leaks” outside the pipe, forming a magnetic leakage field close to the defect [30]. This leakage field is detected by Hall sensors, generating electrical signals [29]. The strength and distribution of the flux leakage vary depending on the geometry of the defect on the pipeline. By capturing the MFL signal, which typically consists of a three-dimensional magnetic flux density (\mathbf{B}) vector field, represented by B_x in the circumferential direction of the pipe, B_y in the radial direction, and B_z in the longitudinal (axial) direction, it is possible to estimate the size and location characteristics of the corrosion defect.

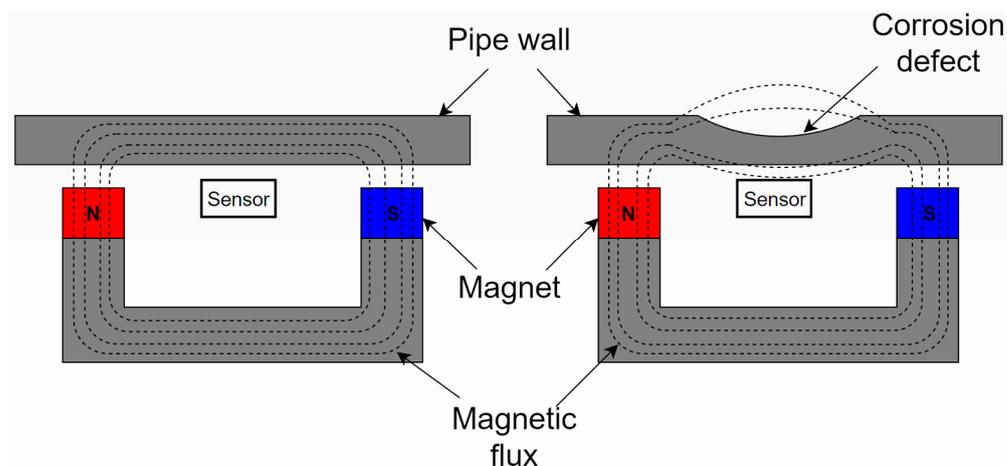


Figure 2. Principle of MFL technique.

Axial (AMFL) and circumferential (CMFL) MFL tools are the two types of MFL tools categorized by the arrangement of the magnets inside the MFL tool. In AMFL tools, the permanent magnets are oriented parallel to the longitudinal axis of the pipe. These tools excel in detecting and measuring defects in the circumferential direction of the pipe but are less effective in sizing features oriented axially [31]. Conversely, CMFL tools have magnets positioned in the circumferential direction of the pipeline. They are more precise in measuring features aligned along the longitudinal direction compared to AMFL tools. The CMFL tool is the main focus of this study because it is more accurate than the AMFL tool for sizing longitudinally-oriented corrosion anomalies, which are perpendicular to the hoop stress resulting from the pipe’s internal pressure and, therefore, can have a great influence on the pipeline burst capacity.

3. Simulating MFL Signals Using FEA

3.1. Simulation Parameters

The CMFL tool considered in this study includes three pairs of permanent magnets. NdFeB is selected as the magnetic material with a high coercive force of 895,000 A/m, stable magnetic properties, and high magnetic energy [32]. The steel used to simulate the pipe is assumed to be grade X52, characterized by a remanent flux density (B_r) of 1T, an electrical conductivity of 5.882×10^6 S/m, coercive field strength (H_c) of 415 A/m, and saturation flux density (B_s) of 1.9T. Figure 3 shows the schematics of a cross-section of the CMFL tool and pipe section, and the geometric and material properties of each component are summarized in Table 1. It is noted that the lift-off value is assumed to be 5 mm, which is the distance between the sensor and the pipe’s internal surface [33].

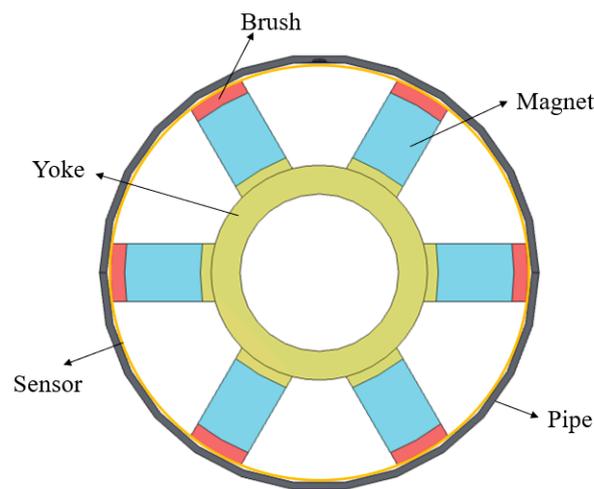


Figure 3. Illustration of CMFL tool.

Table 1. Geometry of every component within the proposed 3D CMFL model.

Elements	Pipe Section	Magnets	Brushes	Yoke
geometry (mm)	$600 \times 610 \times 10^a$	$300 \times 81 \times 105^b$	$300 \times 81 \times 20^b$	$300 \times 300 \times 40^a$

^a Length \times outside diameter \times wall thickness. ^b Length \times width \times thickness.

3.2. FEA Cases

A series of FEA models are created and analyzed with the same magnetization conditions and pipe attributes, but they vary in corrosion types, geometric parameters, and locations in this study. The simulation involves 9600 analysis cases, with 1600 cases allocated to each of the six scenarios as illustrated in Figure 4, namely internal general corrosion (G_{in}), external general corrosion (G_{ex}), internal pinhole (P_{in}), external pinhole (P_{ex}), internal PIC defect (PIC_{in}), and external PIC defect (PIC_{ex}). The general corrosion is idealized as cuboidal-shaped, defined by the width (w_g), depth (d_g), and length (l_g), and the pinhole is idealized as cylindrical-shaped, defined by the radius (r_p) and depth (d_p). Note the center of the pinhole in a PIC defect is assumed to be coincident with that of the general corrosion, which is defined by two coordinates, i.e., its longitudinal location (h) and circumferential location (ϕ), within a 160×160 mm designated area as shown in Figure 5. This area is marked in light blue, with its center specified by the coordinates ($r, \phi = 0, h = 0$), where r represents the inner radius of the pipe. This specific 160×160 mm square area on the inner pipe wall was selected because it encompasses the largest area, showing a consistent magnetic flux density profile between the neighboring magnets.

The analysis cases involving general corrosion only, i.e., G_{in} and G_{ex} , include four values of w_g , i.e., 30, 50, 70, and 90 mm, four values of d_g/t , i.e., 20, 40, 60, and 80%, where t denotes the pipe wall thickness, and four values of l_g , i.e., 30, 50, 70, and 90 mm. This results in 64 general corruptions in total with varying sizes. In terms of the corrosion location, five values of h (−50, −25, 0, 25, and 50 mm) and five values of ϕ (−10, −5, 0, 5, and 10 degrees) are considered. The permutations of 64 defects and 25 locations lead to 1600 cases for each of G_{in} and G_{ex} .

The analysis cases involving pinhole corrosion only, i.e., P_{in} and P_{ex} , include eight values of r_p , i.e., 1, 1.5, 2, 2.5, 3, 3.5, 4, and 4.5 mm, and eight values of d_p/t , i.e., 10, 20, 30, 40, 50, 60, 70, and 80%. This results in 64 pinhole corruptions in total with varying sizes. In terms of the corrosion location, five values of h (−50, −25, 0, 25, and 50 mm) and five values of ϕ (−10, −5, 0, 5, and 10 degrees) are considered. The permutation of 64 defects and 25 locations also leads to 1600 cases for each of P_{in} and P_{ex} .

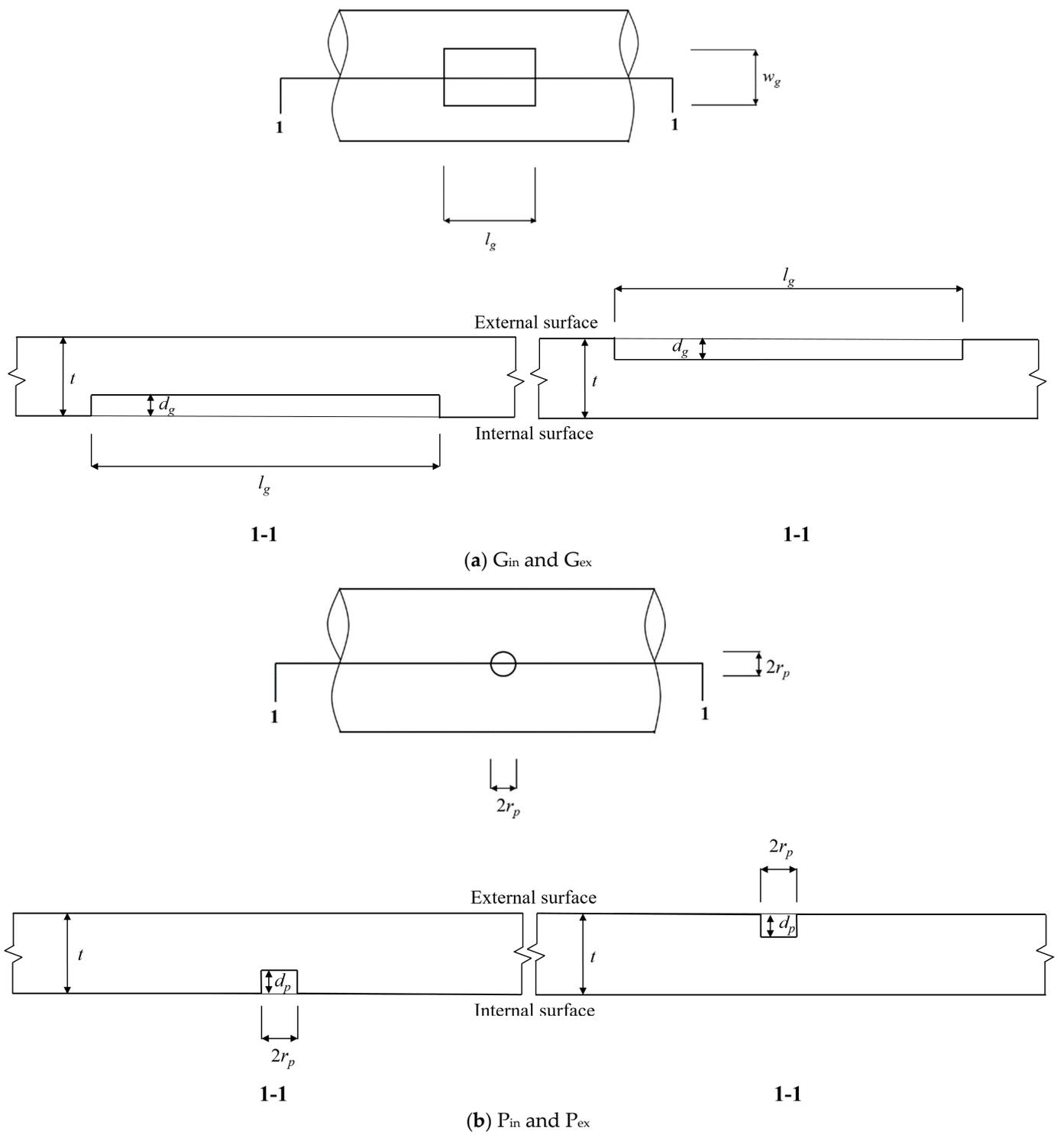


Figure 4. Cont.

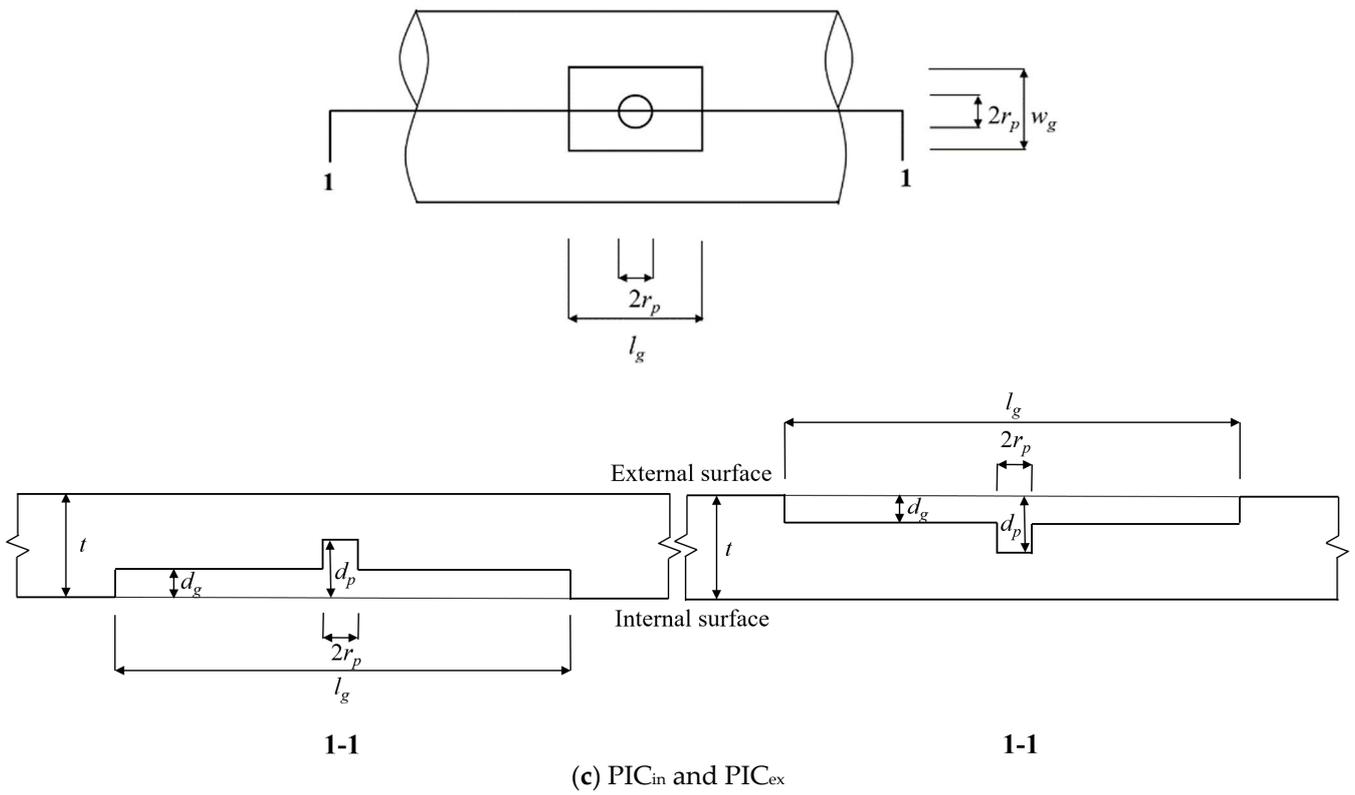


Figure 4. Six types of corrosion situations: (a) G_{in} and G_{ex} ; (b) P_{in} and P_{ex} ; (c) PIC_{in} and PIC_{ex}.

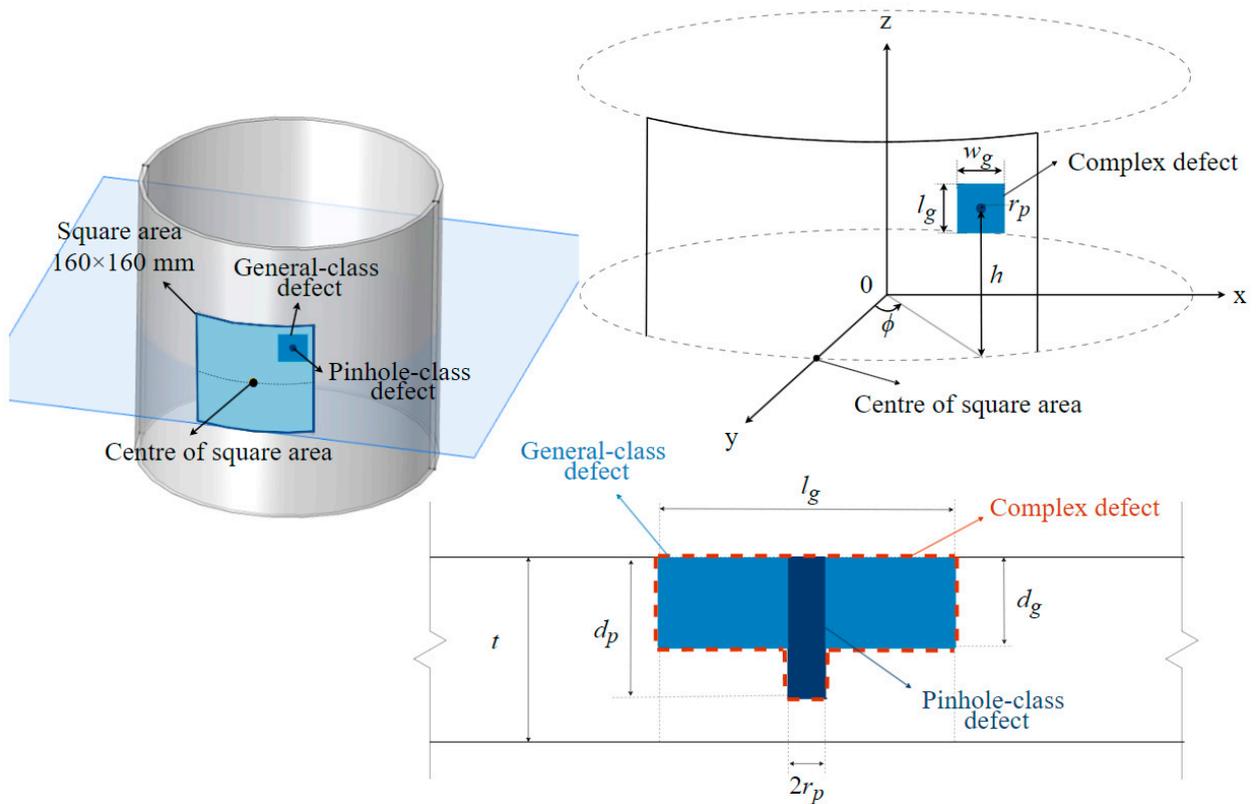


Figure 5. Cylindrical coordinate system defining location parameters h and ϕ .

Analysis cases involving PIC defects, i.e., PIC_{in} and PIC_{ex}, include $w_g = l_g = 50$ mm and four values of r_p , i.e., 1, 2, 3, and 4 mm. In terms of the sizes of general and pinhole corrosions, three groups are considered to capture a wider range of relative relationships between general and pinhole corrosions. The first includes d_g/t equal to 20% and d_p/t equal to 30, 40, 50, 60, 70, or 80%. The second group includes d_g/t equal to 30% and d_p/t equal to 40, 50, 55, 60, 70, or 80%. The last group includes d_g/t equal to 40% and d_p/t equal to 50, 60, 70, or 80%. It should be noted that d_p/t is assumed to be at least 10% larger than d_g/t in all three groups. This results in 64 PIC defects in total with varying sizes. In terms of the corrosion location, five values of h (−50, −25, 0, 25, and 50 mm) and five values of ϕ (−10, −5, 0, 5, and 10 degrees) are considered. The permutation of 64 defects and 25 locations leads to 1600 cases for each of PIC_{in} and PIC_{ex}. It follows that a dataset consisting of 9600 cases ($1600 \times 6 = 9600$) is generated in this study. This large dataset is intended to facilitate the training and validation of the CNN model, enabling comprehensive analysis and robust model development.

The developed FEA model meshes using 4-node tetrahedral elements, where the minimum size is 4 mm, and the maximum mesh size is 55 mm after a convergence study. Specifically, grids inside the 160 mm × 160 mm area are finely resolved at 1 mm intervals. At each grid point, the value of B is calculated via interpolation between two mesh nodes. The system of nonlinear equations is solved using the Newton-Raphson method. As an illustration, Figure 6 depicts the 3D magnetic flux density information (i.e., B_x , B_y , and B_z) within the 160 mm × 160 mm area of a selected FE model. The proposed FEA model was validated by comparing the simulation results with experimental data from a CMFL tool reported by Ireland and Torres [34] and an AMFL tool reported by Azizzadeh and Safizadeh [35]. Readers are referred to Shen and Zhou [19] for details of the validation.

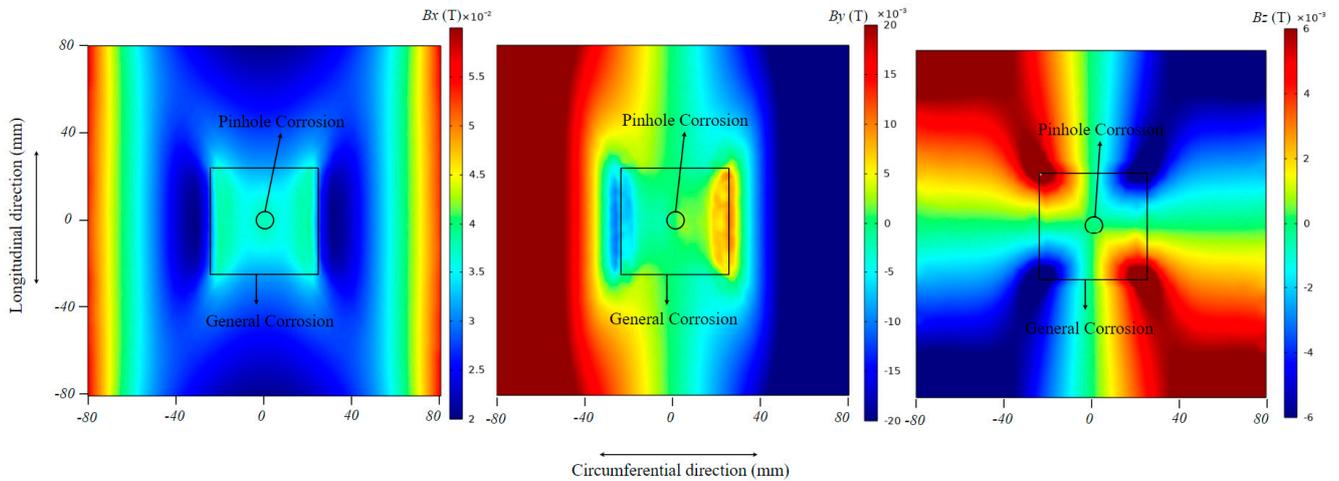


Figure 6. FEA-obtained B_x , B_y , and B_z corresponding to a PIC defect on the internal pipe surface ($w_g = 50$ mm, $l_g = 50$ mm, $d_g/t = 40\%$, $r_p = 4$ mm, $d_p/t = 80\%$, $h = 0$ mm and $\phi = 0$ degree).

MFL signals in practical applications inevitably include noise stemming from factors such as non-uniform pipe wall thickness and differences in sensor lift-off values [36]. These noises can impact the accuracy of the inverse model, necessitating the inclusion of noises in the FEA-simulated MFL signals. The signal-to-noise (SNR) ratio serves as an important parameter for quantifying the noise in the three-dimensional MFL signal. Equation (3) defines SNR in this study [37]:

$$SNR = 10 \log_{10} \left(\frac{\sum_{n=1}^{160} \sum_{m=1}^{160} [B_x(n, m)^2 + B_y(n, m)^2 + B_z(n, m)^2]}{\sum_{n=1}^{160} \sum_{m=1}^{160} [w_x(n, m)^2 + w_y(n, m)^2 + w_z(n, m)^2]} \right) \quad (3)$$

where w_x , w_y and w_z denote the noises along three directions of the pipe, namely circumferential, radial and longitudinal directions, respectively, and (n, m) (where $n, m = 1, 2, \dots, 160$) represents the coordinate of each grid (with 1 mm resolution) in the $160 \text{ mm} \times 160 \text{ mm}$ area, with a total of 25,600 grids. By considering a representative value of SNR equal to 20 [38], three 160×160 matrices of Gaussian-distributed white noises, which are assumed to be independent of each other, are simulated and then applied to B_x , B_y , and B_z , respectively. Each matrix assumes a mean value of zero, with standard deviations adjusted through trial and error until achieving the predefined signal-to-noise ratio. Let σ_n denote the standard deviation of w_x ; the standard deviations of w_y and w_z are assumed to be $0.75\sigma_n$ and $0.25\sigma_n$, respectively. The scaling of σ_n is based on the relative magnitudes of B_x , B_y , and B_z derived from the finite element analysis: for a given analysis case, B_x consistently has the highest magnitude, followed by B_y and B_z , respectively. The final noisy input to the CNN model consists of the three matrices B_x , B_y , and B_z , each combined with respective noisy matrices w_x , w_y , and w_z , resulting in noisy representations of the MFL signals.

It is emphasized that the noisy MFL signal matrices are only fed into the regression CNN model as an extended part in Section 4.2.4. Fan et al. [39] reported that the effect of Gaussian noise on the CNN classification model accuracy is minor. Acharya et al. [40] indicated that removing the noise is not necessary for image classification with deep learning algorithms [41]. Therefore, we only consider the impact of the noises on the accuracy of the regression model in the present study.

4. Convolutional Neural Network

4.1. Classification CNN

4.1.1. Input Information

The input to a CNN model is commonly represented as an RGB color image, including three color channels corresponding to green, red, and blue, respectively. This input is compatible with the three-dimensional MFL signals. For a given FE model, the 3D magnetic flux density information (i.e., B_x , B_y , and B_z) within the designated square area is exported from the COMSOL software at the $160 \text{ mm} \times 160 \text{ mm}$ area grids with a resolution of 1 mm. The three generated 160×160 matrices of noise-free MFL signals (one matrix for each dimension) are denoted as the input of the proposed CNN model, and the six types of corrosion (i.e., G_{in} , G_{ex} , P_{in} , P_{ex} , PIC_{in} , and PIC_{ex}) are the output.

4.1.2. Proposed Structure

The proposed classification CNN model in this study comprises a total of 20 layers: layers 1 to 12 are dedicated to feature extraction, while layers 13 to 20 are focused on classification. Layer 20 serves as the output layer, comprising six components that represent the six corrosion types. The detailed information on each layer is summarized in Table 2. As shown in Table 2, the size of the filter in the convolution layer is selected to be (5×5) , which is a typical size for large filters [42]. The size of the filter in the maxpooling layer is (2×2) , consistent with those commonly used in the existing literature [43,44]. Several dropout layers that randomly drop a certain percentage of neurons from the neural network during the training process are included in the model to avoid overfitting. The employed dropout rate is 0.2, i.e., 20% of the neurons in the neural network are randomly deactivated. Figure 7 depicts the architecture of the proposed classification model with the dimension of each layer, where the input, convolution layers, and max-pooling layers are depicted in pink, blue, and green, respectively. Note that the dropout layers are not shown in Figure 7 as they do not impact the data dimensionality. The thick white line represents the flattened layer, the fully connected layers are represented in orange, and the output layers are shown in yellow in Figure 7.

Table 2. CNN classification model layer information.

Part	Layer No.	Layer Name	Parameters
Feature extraction	1	Convolution	64 filters with size equals to (5×5)
	2	Maxpooling	Pooling size equals to (2×2)
	3	Dropout	Rate = 0.2
	4	Convolution	128 filters with size equals to (5×5)
	5	Maxpooling	Pooling size equals to (2×2)
	6	Dropout	Rate = 0.2
	7	Convolution	256 filters with size equals to (5×5)
	8	Maxpooling	Pooling size equals to (2×2)
	9	Dropout	Rate = 0.2
	10	Convolution	512 filters with size equals to (5×5)
	11	Maxpooling	Pooling size equals to (2×2)
	12	Dropout	Rate = 0.2
Classification	13	Flatten layer	Unit number equals to 51,200
	14	Fully connected layer	Unit number equals to 128
	15	Dropout	Rate = 0.2
	16	Fully connected layer	Unit number equals to 64
	17	Dropout	Rate = 0.2
	18	Fully connected layer	Unit number equals to 32
	19	Dropout	Rate = 0.2
	20	Output 1	Unit number equals to 1
		Output 2	Unit number equals to 1
		Output 3	Unit number equals to 1
Output 4		Unit number equals to 1	
Output 5		Unit number equals to 1	
Output 6		Unit number equals to 1	

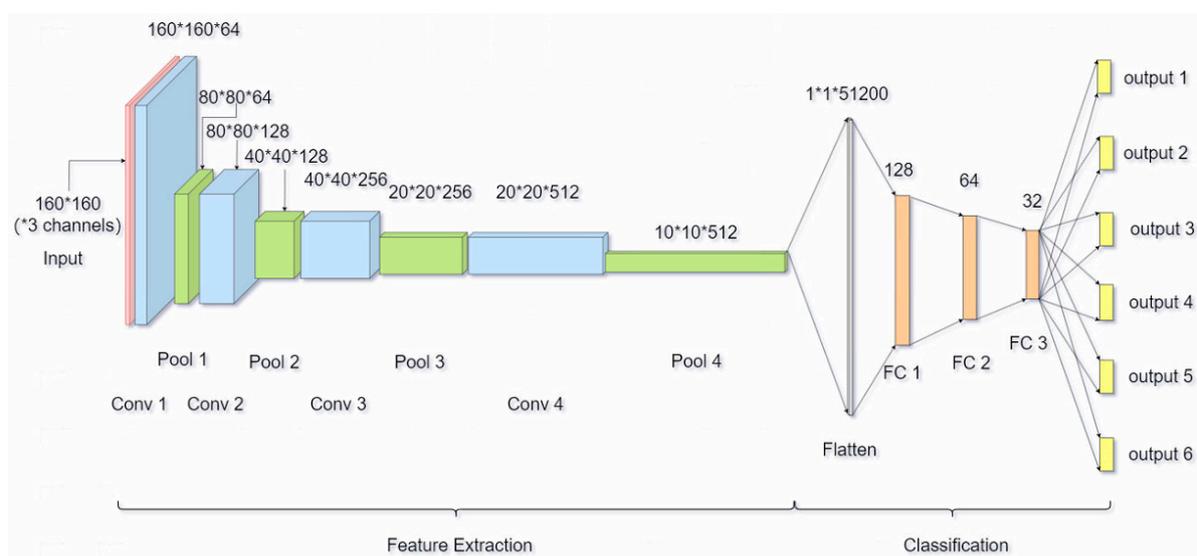


Figure 7. Proposed CNN classification model structure.

4.1.3. Results

By adopting an 80–20% training-test data ratio, there are 1280 training cases and 320 test cases for each of the six corrosion types. The classification results obtained from the proposed CNN model are summarized in Table 3. Table 3 demonstrates that the proposed model consistently achieves prediction accuracies higher than 0.93 in both the training and test datasets, with the lowest accuracy of 0.934 observed in the training dataset for G_{ex} . It is noted that the model performs better in the test dataset than the training dataset, with accuracy rates of each corrosion type in the test dataset higher than or equal to the rates in the training dataset. This may be due to the regularization techniques, such as dropout, adopted in the CNN model. Although these techniques can prevent overfitting, they may also introduce noise or uncertainty during training, which could temporarily degrade the model's performance on the training data while improving the generalization and performance on the test data [45].

Table 3. Results of the classification CNN model.

Category No.	Category	Total	Train	Test	Misprediction in Train	Misprediction in Test	Train Accuracy	Test Accuracy
0	G_{in}	1600	1280	320	7	0	0.978	1.000
1	G_{ex}	1600	1280	320	21	0	0.934	1.000
2	P_{in}	1600	1280	320	0	0	1.000	1.000
3	P_{ex}	1600	1280	320	2	0	0.994	1.000
4	PIC_{in}	1600	1280	320	1	0	0.997	1.000
5	PIC_{ex}	1600	1280	320	1	1	0.997	0.997

Notably, 32 cases in the training dataset and one case in the test dataset have been misclassified, resulting in a misprediction rate of 0.34% ($33/9600 = 0.34\%$), demonstrating the excellent accuracy of the developed model. Figure 8 illustrates the distribution of mispredictions in the training dataset across different categories. For instance, 65.6% (i.e., 21) of misclassified training cases belong to the G_{ex} category. Among these, four cases are misclassified as P_{in} , and 17 cases are misclassified as P_{ex} . Furthermore, six out of the 33 cases have misclassified locations (internal/external surface of the pipe), i.e., two P_{ex} cases misclassified as P_{in} , and four G_{ex} cases misclassified as P_{in} .

Note that the model achieves better classification accuracy for P_{in} , P_{ex} , PIC_{in} , and PIC_{ex} than for G_{in} and G_{ex} . The values of d_g/t included in the entire dataset for all six types of corruptions are 20, 30, 40, 60, and 80%, while the values of d_p/t included in the entire dataset are 20, 30, 40, 50, 55, 60, 70 and 80%. The wider range of the corrosion depth for cases involving pinholes (i.e., P_{in} , P_{ex} , PIC_{in} , and PIC_{ex}) compared to the cases involving general corrosion (G_{in} and G_{ex}) allows the model to learn a richer set of features and variations associated with pinhole corrosion, making the model more adept at distinguishing between different types and severity of pinhole corrosion. All of the 28 misclassified cases in the G_{in} and G_{ex} categories involve relatively shallow corruptions, among which 26 cases have $d_g/t = 20\%$ and two cases have $d_g/t = 40\%$. The distribution of the 28 misclassified cases by w_g/l_g is shown in Figure 9. Figure 9 indicates that the classification CNN model is less accurate for wide general corruptions with w_g/l_g larger than 1.67. It is also noteworthy that the developed model is less accurate for cases containing external corruptions (i.e., G_{ex} , P_{ex} , and PIC_{ex}) than for cases involving internal corruptions (i.e., G_{in} , P_{in} , and PIC_{in}). A potential explanation is that the model input (i.e., the MFL signal map) for external defects has less pronounced features because the variation in B due to the metal-loss defect on the external pipe surface is smaller than that caused by the internal metal-loss defect as the external defect is farther away from the sensor than the internal defect.

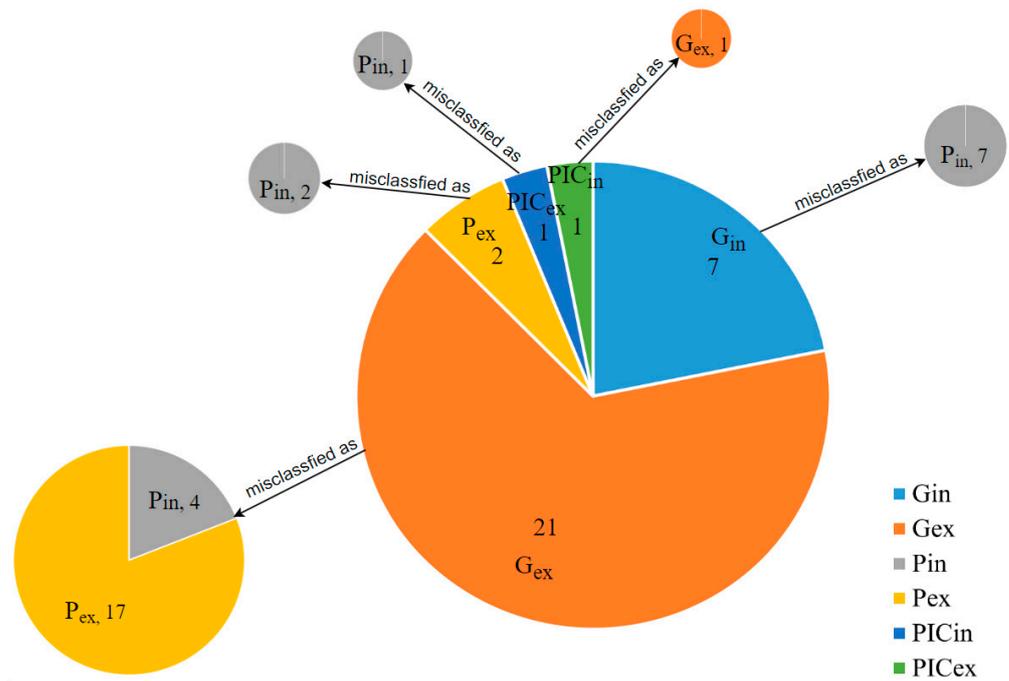


Figure 8. Distribution of misclassifications in the training dataset.

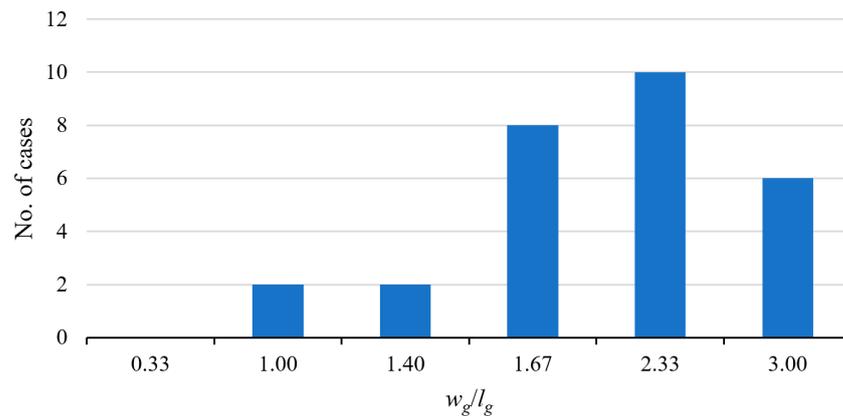


Figure 9. Distribution of misclassified cases in G_{in} and G_{ex} by w_g/l_g .

4.2. Regression CNN

4.2.1. Input Information

The results of the classification CNN model are utilized to extract correctly classified cases for P_{in}, P_{ex}, PIC_{in}, and PIC_{ex}, excluding four misclassifications in the training dataset and one misclassification in the test dataset, resulting in 1600, 1598, 1599, and 1598 input cases for P_{in}, P_{ex}, PIC_{in}, and PIC_{ex}, respectively, to train and validate the regression CNN model. The three generated 160 × 160 matrices of noise-free MFL signals for each of the 6395 (=1600 + 1598 + 1599 + 1598) cases are the input to the developed regression model, and the predicted r_p , d_p/t , h , and ϕ values for each case are the output. The correctly classified data are adopted to train the regression CNN model because including misclassified data (such as a general corrosion anomaly misclassified as a pinhole) in training is problematic, e.g., defining an error function that is applicable to all input data, including misclassified data. Alternatively, the training of the regression CNN model can be completely separated from the classification model; that is, the entire 6400 parametric finite element analysis cases involving pinholes can be employed to train the regression model. In this study, the correctly classified dataset, which consists of 6395 cases and is almost identical to the complete dataset, is employed to train and validate the regression CNN model.

4.2.2. Proposed Structure

The proposed regression CNN model in the present study comprises a total of 23 layers: layers 1 to 16 are dedicated to feature extraction, while layers 17 to 23 are focused on regression. Layer 23 is the output layer, which includes four components representing the predicted r_p , d_p/t , h , and ϕ values, respectively. The detailed CNN regression model layer information is presented in Table 4. It is noted that, in addition to the dropout layers introduced in Section 4.1.2, several batch normalization layers are incorporated within the typical CNN layers to prevent overfitting [46–48]. As shown in Table 4, the tuned hyperparameters in the feature extraction process of the regression model, such as the filter size of the convolution layer, the pooling size, and the dropout rate, are the same as those in the classification CNN model, while the employed dropout rate in the regression process of the regression CNN model is 0.5. Figure 10 depicts the architecture of the developed regression model with the dimensions of different layers. Note that the dropout and batch normalization layers are excluded from the diagram as they do not impact the data dimensionality. The color scheme utilized to distinguish various layers in Figure 10 is consistent with those utilized in Figure 7.

Table 4. CNN regression model layer information.

Part	Layer No.	Layer Name	Parameters
Feature extraction	1	Convolution	64 filters with size equals to (5×5)
	2	Batch normalization	
	3	Maxpooling	Pooling size equals to (2×2)
	4	Dropout	Rate = 0.2
	5	Convolution	128 filters with size equals to (5×5)
	6	Batch normalization	
	7	Maxpooling	Pooling size equals to (2×2)
	8	Dropout	Rate = 0.2
	9	Convolution	256 filters with size equals to (5×5)
	10	Batch normalization	
	11	Maxpooling	Pooling size equals to (2×2)
	12	Dropout	Rate = 0.2
	13	Convolution	512 filters with size equals to (5×5)
	14	Batch normalization	
	15	Maxpooling	Pooling size equals to (2×2)
	16	Dropout	Rate = 0.2
Regression	17	Flatten layer	Unit number equals to 51,200
	18	Dropout	Rate = 0.5
	19	Fully connected layer	Unit number equals to 64
	20	Fully connected layer	Unit number equals to 128
	21	Fully connected layer	Unit number equals to 256
	22	Fully connected layer	Unit number equals to 512
	23	Output 1	Unit number equals to 1
	Output 2	Unit number equals to 1	
	Output 3	Unit number equals to 1	
	Output 4	Unit number equals to 1	

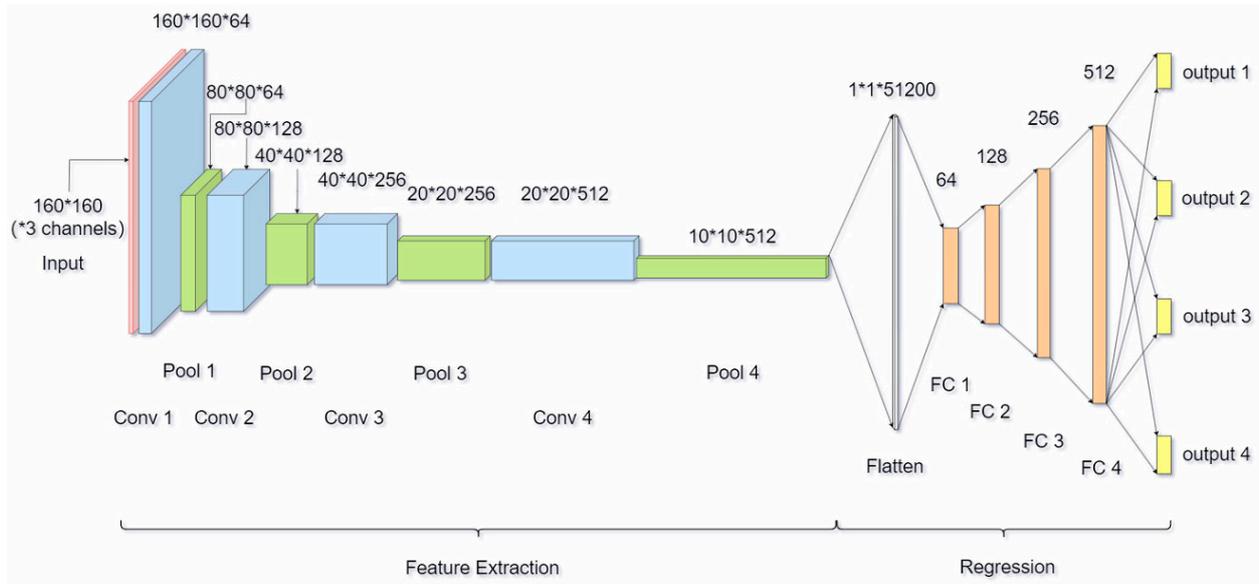


Figure 10. Proposed CNN regression model structure.

4.2.3. Results

The coefficient of determination (i.e., R^2) is used to measure the accuracy of each parameter prediction (i.e., r_p , d_p/t , h , and ϕ), defined by the following equation.

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_{true,i} - Y_{pred,i})^2}{\sum_{i=1}^N (Y_{true,i} - \bar{Y}_{true})^2} \tag{4}$$

where $Y_{true,i}$ and $Y_{pred,i}$ represent the true and predicted values, respectively, of each predicted parameter for the i th ($i = 1, 2, \dots, N$) point in a given dataset containing N data points, and \bar{Y}_{true} denotes the true mean value of N data points associated with a predicted parameter. Table 5 summarizes the R^2 values for the four parameters in the test dataset included in the regression model. The R^2 values are also plotted in a so-called radar chart in Figure 11. The true values of the four parameters r_p , d_p/t , h , and ϕ compared with the estimated values by the CNN for the corrosions in the test set are depicted in Figure 12. It is evident from Table 5 that the CNN model’s predictions of the corrosion location parameters (i.e., h and ϕ) show strong agreement with the actual values, as reflected by the R^2 values of 1.0 for all four corrosion categories. In terms of the size predictions (i.e., r_p and d_p/t), the model is more accurate for internal corrosions (P_{in} and PIC_{in}) than for external corrosions (P_{ex} and PIC_{ex}). These results also indicate that the performance of the model is better for cases containing pinholes only than for cases containing PIC. This is expected as PIC introduces additional complexity, including the presence of general corrosion, which may impact the model’s accuracy. Overall, the obtained R^2 values demonstrate the CNN model’s accuracy in predicting the size parameters, with a higher accuracy observed for corrosions on the internal surface and pinholes. However, note that although the predictions for r_p and d_p/t are, in general, accurate, Figure 12 reveals relatively large error bands in both r_p and d_p/t predictions. For instance, the predicted values of r_p range from 1.14 mm to 3.45 mm for $(r_p)_{true}$ equal to 2 mm. The large error bands may be explained by the small dimensions of the pinhole corrosion and the limited spatial resolution of the model. Furthermore, the presence of composite corrosion features can introduce additional complexities and uncertainties, leading to larger error bands in the predictions. Therefore, further investigation is required to explore the factors contributing to the reduction of regression errors for pinhole corrosions.

Table 5. Results of the regression CNN model without noise.

Category No.	Category	Cases			Test R^2			
		Total	Train	Test	r_p	d_p/t	h	ϕ
0	P_{in}	1600	1280	320	0.91	0.95	1.00	1.00
1	P_{ex}	1598	1277	321	0.88	0.92	1.00	1.00
2	PIC_{in}	1599	1280	319	0.89	0.85	1.00	1.00
3	PIC_{ex}	1598	1279	319	0.86	0.77	1.00	1.00
Overall		6395	5116	1279	0.89	0.91	1.00	1.00

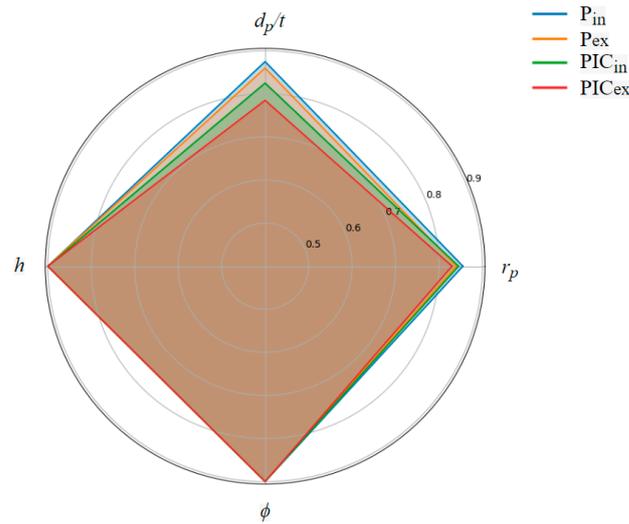


Figure 11. Radar chart for R^2 of metrics (r_p , d_p/t , h , and ϕ) for P_{in} , P_{ex} , PIC_{in} , and PIC_{ex} corrosions in the test dataset.

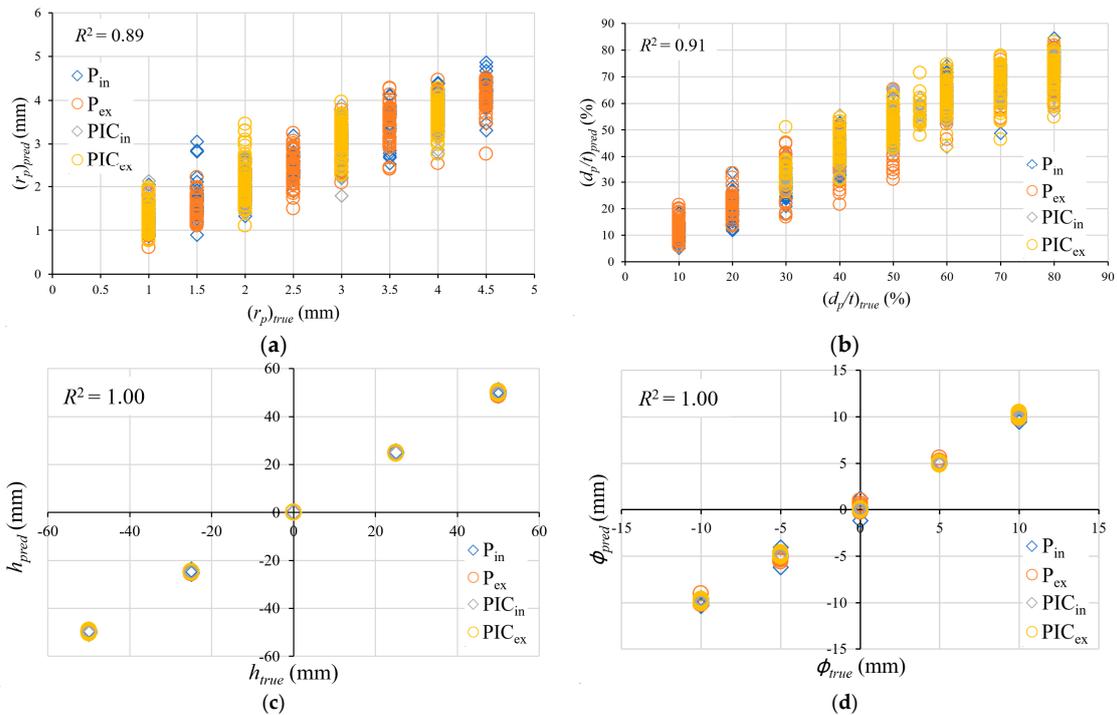


Figure 12. Comparison of true and predicted values of r_p , d_p/t , h , and ϕ for the anomalies in the regression test dataset without noise.

4.2.4. Influence of Noise

Figure 13 compares the results of the regression CNN model for noisy input (SNR = 20) with those of the noise-free scenario described in Section 4.2.3. With the presence of noise, the R^2 values for predicting r_p and d_p/t decrease somewhat compared to the noise-free scenario. The decrease in R^2 values is more prominent for P_{ex} and PIC_{ex} cases, indicating that the model is more affected by noise when predicting the sizes of external corrosion. Despite the decrease in the accuracy in predicting the corrosion sizes, the model maintains $R^2 = 1$ for the location parameters (h and ϕ) across all corrosion types except for P_{ex} , indicating its robustness in capturing the spatial information even in the presence of noise.

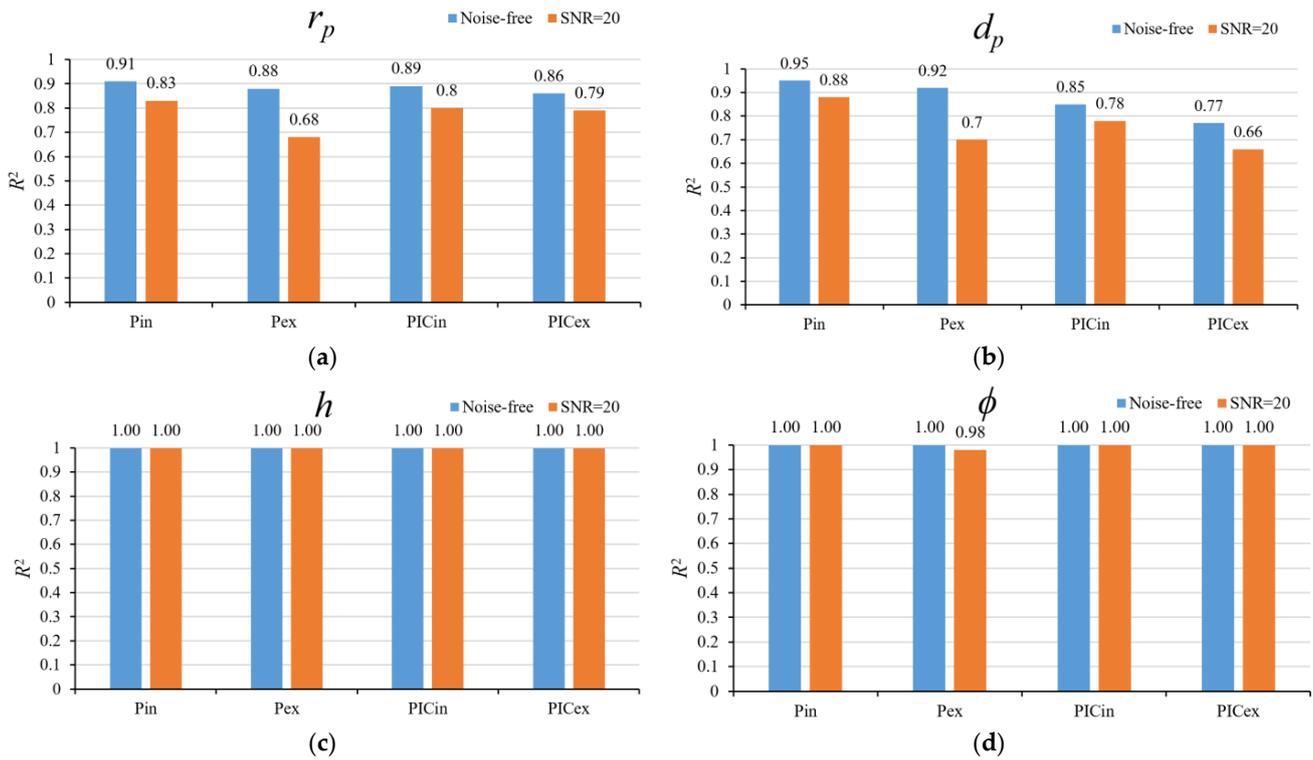


Figure 13. Comparison of R^2 between noise-free and SNR = 20 scenarios for different corrosion types.

Finally, it is important to point out that the CNN classification and regression models developed in this study are specifically tuned to certain pipeline and MFL tool attributes, such as the pipe wall thickness, magnet properties, lift-off values, and defect shapes. Therefore, further investigations are needed to determine how well the CNN model performs with various pipe wall thicknesses and different MFL tool parameters. Further studies are also needed to take into account the movement of the MFL tool into the prediction models.

5. Conclusions

This paper reports a novel study that focuses on the classification and sizing of pinhole corrosions on pipelines using deep learning algorithms. We propose a CNN classification model to classify six different types of corrosions (i.e., G_{in} , G_{ex} , P_{in} , P_{ex} , PIC_{in} , and PIC_{ex}) on pipelines and a CNN regression model to estimate the sizes and location of the pinhole defects based on MFL signals generated using magneto-static FEA. Extensive 3D parametric FEA cases involving box-shaped general corrosions and cylinder-shaped pinholes are used to simulate the 3D MFL signals by varying the defect depth, length, width, and longitudinal and circumferential locations. The CNN classification and regression models are then trained and validated using the simulated MFL signals.

The proposed classification model is shown to have excellent accuracy in classifying six types of corrosion defects with a misprediction rate of 0.34% ($33/9600 = 0.34\%$). The

proposed regression model is highly accurate in predicting the location of the pinhole and achieves good accuracy in predicting the depth and diameter of the pinhole. Through a comparative analysis of the CNN regression model with noise-free and noisy signals (SNR = 20) as input, it is observed that the noise impact on the predictive accuracy of the regression model is moderate. This study demonstrates the application of deep learning algorithms to facilitate the integrity management of pipelines containing complex-shaped corrosion defects.

Author Contributions: Y.S.: Investigation, Visualization, Formal analysis, Methodology, Writing—original draft. W.Z.: Funding acquisition, Supervision, Conceptualization, Methodology, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the financial support received from the Natural Sciences and Engineering Research Council of Canada (NSERC) (Grant No. RGPIN-2019-05160) and the Faculty of Engineering of the University of Western Ontario.

Data Availability Statement: Data are available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Zhou, Q.; Wu, W.; Liu, D.; Li, K.; Qiao, Q. Estimation of corrosion failure likelihood of oil and gas pipeline based on fuzzy logic approach. *Eng. Fail. Anal.* **2016**, *70*, 48–55. [[CrossRef](#)]
- Murphy, J.F. Nightmare pipeline failures, fantasy planning, black swans, and integrity management—A review. *Process Saf. Prog.* **2015**, *34*, 207. [[CrossRef](#)]
- Shen, Y.; Zhou, W. A comparison of onshore oil and gas transmission pipeline incident statistics in Canada and the United States. *Int. J. Crit. Infrastruct. Prot.* **2024**, *45*, 100679. [[CrossRef](#)]
- Vanaei, H.R.; Eslami, A.; Egbewande, A. A review on pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models. *Int. J. Press. Vessels Pip.* **2017**, *149*, 43–54. [[CrossRef](#)]
- Song, H.; Yang, L.; Liu, G.; Tian, G.; Ona, D.I.; Song, Y.; Li, S. Comparative analysis of in-line inspection equipments and technologies. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *382*, 032021. [[CrossRef](#)]
- Pipeline Operators Forum. Specifications and Requirements for In-Line Inspection of Pipelines. Version 2021. Available online: <https://pipelineoperators.org/documents> (accessed on 18 May 2022).
- Sutherland, J.; Bluck, M.; Pearce, J.; Quick, E. Validation of latest generation MFL in-line inspection technology leads to improved detection and sizing specification for pinholes, pitting, axial grooving and axial slotting. In Proceedings of the ASME 2010 8th International Pipeline Conference, Calgary, AB, Canada, 27 September–1 October 2010.
- Peng, X.; Anyaoha, U.; Liu, Z.; Tsukada, K. Analysis of magnetic-flux leakage (MFL) data for pipeline corrosion assessment. *IEEE Trans. Magn.* **2020**, *56*, 1–15. [[CrossRef](#)]
- Han, W.; Yang, P.; Xia, F.; Xue, Y. Magnetic flux leakage signal inversion of corrosive flaws based on modified genetic local search algorithm. *J. Shanghai Jiaotong Univ. (Sci.)* **2009**, *14*, 168–172. [[CrossRef](#)]
- Priewald, R.H.; Magele, C.; Ledger, P.D.; Pearson, N.R.; Mason, J.S.D. Fast magnetic flux leakage signal inversion for the reconstruction of arbitrary defect profiles in steel using finite elements. *IEEE Trans. Magn.* **2012**, *49*, 506–516. [[CrossRef](#)]
- Hwang, K.; Mandayam, S.; Udpa, S.S. Characterization of gas pipeline inspection signals using wavelet basis function neural networks. *NDT E Int.* **2000**, *33*, 531–545. [[CrossRef](#)]
- Ramuhalli, P.; Udpa, L.; Udpa, S.S. Neural network based inversion algorithms in magnetic flux leakage nondestructive evaluation. *J. Appl. Phys.* **2003**, *93*, 82748276. [[CrossRef](#)]
- Han, W.H.; Que, P.W. 2-D defect reconstruction from MFL signals based on genetic optimization algorithm. In Proceedings of the IEEE 2005 International Conference on Industrial Technology, Hong Kong, China, 14–17 December 2005.
- Kandroodi, M.R.; Araabi, B.N.; Bassiri, M.M.; Ahmadabadi, M.N. Estimation of depth and length of defects from magnetic flux leakage measurements: Verification with simulations, experiments, and pigging data. *IEEE Trans. Magn.* **2016**, *53*, 1–10. [[CrossRef](#)]
- Feng, J.; Li, F.; Lu, S.; Liu, J. Fast reconstruction of defect profiles from magnetic flux leakage measurements using a RBFNN based error adjustment methodology. *IET Sci. Meas. Technol.* **2017**, *11*, 262–269. [[CrossRef](#)]
- Yao, G.; Lei, T.; Zhong, J. A review of convolutional-neural-network-based action recognition. *Pattern Recognit. Lett.* **2019**, *118*, 14–22. [[CrossRef](#)]
- Taye, M.M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation* **2023**, *11*, 52. [[CrossRef](#)]
- Lu, S.; Feng, J.; Zhang, H.; Liu, J.; Wu, Z. An estimation method of defect size from MFL image using visual transformation convolutional neural network. *IEEE Trans. Ind. Inform.* **2018**, *15*, 213–224. [[CrossRef](#)]

19. Shen, Y.; Zhou, W. Development of a convolutional neural network model to predict the size and location of corrosion defects on pipelines based on magnetic flux leakage signals. *Int. J. Press. Vessels Pip.* **2023**, *207*, 105–123. [[CrossRef](#)]
20. Wang, H.A.; Chen, G. Defect size estimation method for magnetic flux leakage signals using convolutional neural networks. *Insight* **2020**, *62*, 86–91. [[CrossRef](#)]
21. EGIG. *10th Report of the European Gas Pipeline Incident Data Group (Period 1970–2016)*; EGIG: Groningen, The Netherlands, 2018; Doc. number VA 17.R.0395.
22. Zhang, H.; Sha, S.; Willis, C.; Qingshan, F.; Chen, P. Feasibility study of pinhole inspection via magnetic flux leakage and hydrostatic testing in oil & gas pipelines. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1043*, 022053.
23. Feng, Q.; Yan, B.; Chen, P.; Shirazi, S.A. Failure analysis and simulation model of pinhole corrosion of the refined oil pipeline. *Eng. Fail. Anal.* **2019**, *106*, 1–28. [[CrossRef](#)]
24. Subramanian, C. Localized pitting corrosion of API 5L grade A pipe used in industrial fire water piping applications. *Eng. Fail. Anal.* **2018**, *92*, 405–417. [[CrossRef](#)]
25. Askari, M.; Aliofkhaeaei, M.; Afroukhteh, S. A comprehensive review on internal corrosion and cracking of oil and gas pipelines. *J. Nat. Gas Sci. Eng.* **2019**, *71*, 102971. [[CrossRef](#)]
26. Kadhim, K.N.; Al-Rufaye, A.H.R. The effects of uniform transverse magnetic field on local flow and velocity profile. *Int. J. Civ. Eng. Technol.* **2016**, *7*, 140–151.
27. Ji, F.; Wang, C.; Sun, S.; Wang, W. Application of 3-D FEM in the simulation analysis for MFL signals. *Insight* **2009**, *51*, 32–35. [[CrossRef](#)]
28. Bubenik, T. *Electromagnetic Methods for Detecting Corrosion in Underground Pipelines: Magnetic Flux Leakage (MFL)*; Underground Pipeline Corrosion; Woodhead Publishing: Cambridgeshire, UK, 2014; pp. 215–226.
29. Shi, Y.; Zhang, C.; Li, R.; Cai, M.; Jia, G. Theory and application of magnetic flux leakage pipeline detection. *Sensors* **2015**, *15*, 31036–31055. [[CrossRef](#)]
30. Zhang, Y.; Ye, Z.; Wang, C. A fast method for rectangular crack sizes reconstruction in magnetic flux leakage testing. *NDT E Int.* **2009**, *42*, 369–375. [[CrossRef](#)]
31. Walker, J. *In-Line Inspection of Pipelines: Advanced Technologies for Economic and Safe Operation of Oil and Gas Pipelines*; Verlag Moderne Industrie: Landsberg am Lech, Germany, 2010.
32. Liu, Y.; Gao, X.; Wang, Y.; Yang, X. Sensitive parameters' optimization of the permanent magnet supporting mechanism. *J. Mech. Sci. Technol.* **2014**, *28*, 2707–2714. [[CrossRef](#)]
33. Yang, L.; Zhang, G.; Liu, G.; Gao, S. Effect of lift-off on pipeline magnetic flux leakage inspection. In Proceedings of the 2008 17th World Conference on Nondestructive Testing, Shanghai, China, 25–28 October 2008.
34. Ireland, R.C.; Torres, C.R. Finite element modelling of a circumferential magnetizer. *Sens. Actuators A Phys.* **2006**, *129*, 197–202. [[CrossRef](#)]
35. Azizzadeh, T.; Safizadeh, M.S. Three-dimensional finite element and experimental simulation of magnetic flux leakage-type NDT for detection of pitting corrosions. In Proceedings of the 2017 4th Iranian International NDT Conference (IRNNDT), Tehran, Iran, 21–22 February 2017.
36. Chen, L.; Li, X.; Qin, G.; Lu, Q. Signal processing of magnetic flux leakage surface flaw inspect in pipeline steel. *Russ. J. Nondestruct. Test.* **2008**, *44*, 859–867. [[CrossRef](#)]
37. Piao, G.; Guo, J.; Hu, T.; Leung, H. The effect of motion-induced eddy current on high-speed magnetic flux leakage (MFL) inspection for thick-wall steel pipe. *Res. Nondestruct. Eval.* **2020**, *31*, 48–67. [[CrossRef](#)]
38. Li, F.; Feng, J.; Zhang, H.; Liu, J.; Lu, S.; Ma, D. Quick reconstruction of arbitrary pipeline defect profiles from MFL measurements employing modified harmony search algorithm. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 9. [[CrossRef](#)]
39. Fan, X.; Dai, M.; Liu, C.; Wu, F.; Yan, X.; Feng, Y.; Feng, Y.; Su, B. Effect of image noise on the classification of skin lesions using deep convolutional neural networks. *Tsinghua Sci. Technol.* **2019**, *25*, 425–434. [[CrossRef](#)]
40. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adam, M.; Gertych, A.; San Tan, R. A deep convolutional neural network model to classify heartbeats. *Comput. Boil. Med.* **2017**, *89*, 389–396. [[CrossRef](#)] [[PubMed](#)]
41. Tesfai, H.; Saleh, H.; Al-Qutayri, M.; Mohammad, M.B.; Tekeste, T.; Khandoker, A.; Mohammad, B. Lightweight shufflenet based cnn for arrhythmia classification. *IEEE Access.* **2022**, *10*, 111842–111854. [[CrossRef](#)]
42. West, N.E.; O'shea, T. Deep architectures for modulation recognition. In Proceedings of the 2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Baltimore, MD, USA, 6–9 March 2017.
43. Baranwal, S.; Khandelwal, S.; Arora, A. Deep learning convolutional neural network for apple leaves disease detection. In Proceedings of the 2019 International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur, India, 26–28 February 2019.
44. Virupakshappa, K.; Marino, M.; Oruklu, E. A multi-resolution convolutional neural network architecture for ultrasonic flaw detection. In Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS), Kobe, Japan, 22–25 October 2018.
45. Bilmes, J. Underfitting and Overfitting in Machine Learning. UW ECE Course Notes. 2020. Available online: https://people.ece.uw.edu/bilmes/classes/ee511/ee511_spring_2020/overfitting_underfitting.pdf (accessed on 20 March 2023).
46. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 2015 International Conference on Machine Learning, Lille, France, 6–11 July 2015.

47. Thakkar, V.; Tewary, S.; Chakraborty, C. Batch normalization in convolutional neural networks—A comparative study with CIFAR-10 data. In Proceedings of the IEEE 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT), IEST Shibpur, West Bengal, India, 12–13 January 2018.
48. Bjorck, N.; Gomes, C.P.; Selman, B.; Weinberger, K.Q. Understanding batch normalization. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 1–12.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.