Image analysis and CNN-based crack depth estimation using eddy current data

Pham Van Dung, Cung Thanh Long*

School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, 1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam.

*Corresponding author: long.cungthanh@hust.edu.vn

Received 15 Mar. 2024; Revised 06 May 2024; Accepted 12 Jun. 2024; Published 25 Jun. 2024.

DOI: https://doi.org/10.54939/1859-1043.j.mst.96.2024.12-20

ABSTRACT

This study presents a comprehensive approach for crack depth estimation utilizing advanced image analysis techniques and a Convolutional Neural Network (CNN) model. The aim is to enhance accuracy and reliability in predicting crack depths, particularly for sub-millimeter cracks. The research addresses challenges arising from noise in images by employing a pre-processing technique and augmentation methods. The proposed method's effectiveness is showcased through its application to experimental crack data from diverse specimens. The outcomes exhibit a Mean Relative Error (MRE) of around 6%, indicating a high level of precision. These results affirm the potential of the methodology for real-world industrial applications. Additionally, the study explores the integration of eddy current image processing with CNN for Non-Destructive Evaluation (NDE) problems, offering a new approach for tiny surface-crack detection and characterization.

Keywords: Non-Destructive Evaluation; Eddy-current technique; Convolutional Neural Network; Data augmentation.

1. INTRODUCTION

In the aeronautical industry and other sectors that utilize laminated metal fuselage structures, there is a constant concern about the potential damage caused by minor cracks arising under challenging working conditions. To ensure aviation safety, early detection of such cracks in aircraft fuselages is critical. Similarly, in various industrial sectors, the detection and measurement of damages in metal parts during manufacturing, operation, and equipment maintenance are of utmost importance. In the past, cracks were manually detected by experts who examined the defective component visually and utilized specific tools to identify any deficiency in the components. However, this method was ineffective due to labor extensive and prone to human error [1]. Researchers now mainly focus on automatic methods, which utilize technologies to identify defects. Eddy current (EC) Non-Destructive Testing/Estimation (NDT/E) has emerged as a widely applied method [2–5] for this purpose due to its high sensitivity, mobility, and capability to detect defects on and below the surface without requiring complex surface preparation [6, 7].

Eddy current NDT/E has numerous advantages, making it versatile for qualitative and quantitative assessment of defects in metal structures. It finds applications in different areas, including evaluating surface corrosion in gas pipelines, inspecting aerospace structures, maintaining steel plates in construction projects, and examining steel pipes used in various fields [8–14].

Over time, researchers have made significant efforts to detect cracks in massive metal structures using different methods. Some studies have reported positive outcomes in this area. For instance, one study enhanced the reliability of eddy current testing by using a split-D reflection differential probe, effectively working with surface roughness and measurement frequency noise [15]. Another approach proposed a method to determine the length and depth of cracks in conductive metal plates by analyzing changes in magnetic density at measuring points and utilizing finite element method (FEM) modeling for depth estimation [16]. Additionally, a technique using traditional eddy current

(TEC) and pulse eddy current (PEC) simultaneously was introduced to estimate multiple crack chains in metal specimens [17].

Incorporating machine learning techniques has further improved crack inspection and evaluation in metal parts [18–21]. Previous works primarily focus on feature selection and feature extraction to improve estimation performance, in both the time domain and frequency domain [22–25]. Statistical approaches have been used to select optimal data and features from eddy current datasets affected by noise. Traditional image-based feature extraction techniques have been applied for crack angular detection, and a combination of wavelet transform and Multilayer Perceptron (MLP) network has been used to estimate crack depth with reduced error rates.

While research has focused on millimeter-sized cracks [11, 12, 15, 17, 19, 20], there is a lack of studies addressing crack depths smaller than 1 mm. Our previous works [13, 14] examined a problem of crack depth estimation in the range of 200 to 800 µm through processing S-scan EC signals. The results achieved mean relative errors of approximately 10% by applying various noise elimination techniques and carefully choosing manually extracted features. These features were subsequently utilized to train either a simple polynomial forward model or a MLP for the purpose of prediction.

With the same problem, this paper presents a comprehensive approach to crack depth estimation. Instead of manually calculating hand-crafted features for training prediction models as in previous studies, we proposed a Convolutional Neural Network (CNN) to automatically extracted features from experimental images obtained from impedance maps acquired by EC sensors. To enhance accuracy, a pre-processing procedure is employed to minimize the influence of noise in input images for the crack depth estimation problem. Data augmentation techniques for images are also applied with the aim of enhancing and diversifying small datasets, as well as improving the performance and generalization ability of machine learning models. The following sections provide a comprehensive breakdown of the proposed method in section 2, present the outcomes and subsequent discussions of the proposed method in section 3, and conclude with final remarks and future prospects in section 4.

2. PROPOSED METHOD

2.1. Data description

In this study, we evaluated our proposed method using the dataset examined in [13]. The dataset consists of C-scan eddy current images, where each pixel/point represents a complex value of the sensor at a corresponding position on the surface of various massive specimens. To collect the data, two EC sensors were utilized: an absolute sensor operating at a frequency of 500 kHz and a differential sensor operating at 400 kHz (figure 1).

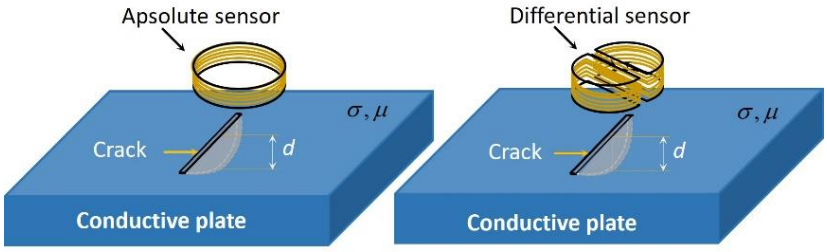

Figure 1. *Illustration of data acquisition for the problem of crack depth (d) estimation, using two kinds of EC sensor.*

The dataset comprises a total of 8 specimens subjected to mechanical fatigue to generate the necessary data. These specimens were evenly divided into two groups based on the type of force stimulation applied, resulting in the difference in shape of the cracks (figure 2).

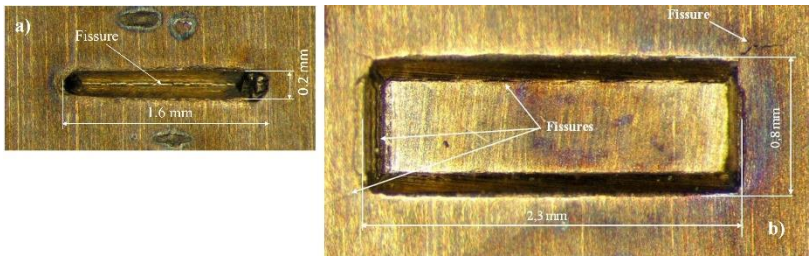

Figure 2. *Two groups of specimens with different types of force applied: (a) group one and (b) group two.*

The first group includes four specimens labeled as M11, M12, M13, and M14, while the second group consists of specimens M21, M22, M23, and M24. Each specimen in both groups underwent three cycles of impact force, resulting in typical crack depths of 200, 400, and 800 µm after each cycle. After each cycle, the EC measurements were conducted to obtain two C-scan images, representing the real and imaginary parts of the sensor's impedance. As a result, each group contains a total of 12 samples (four for each crack depth: 200 µm, 400 µm, and 800 µm).

2.2. Data preprocessing

In order to leverage the capabilities of CNN architecture for extracting features from image data, we decided to use the entire C-Scan image as input instead of manually extracting specific features to construct the model. As a result, we proposed a preprocessing method for the module image of the sensor. The complete process of preparing the module images is implemented as follows.

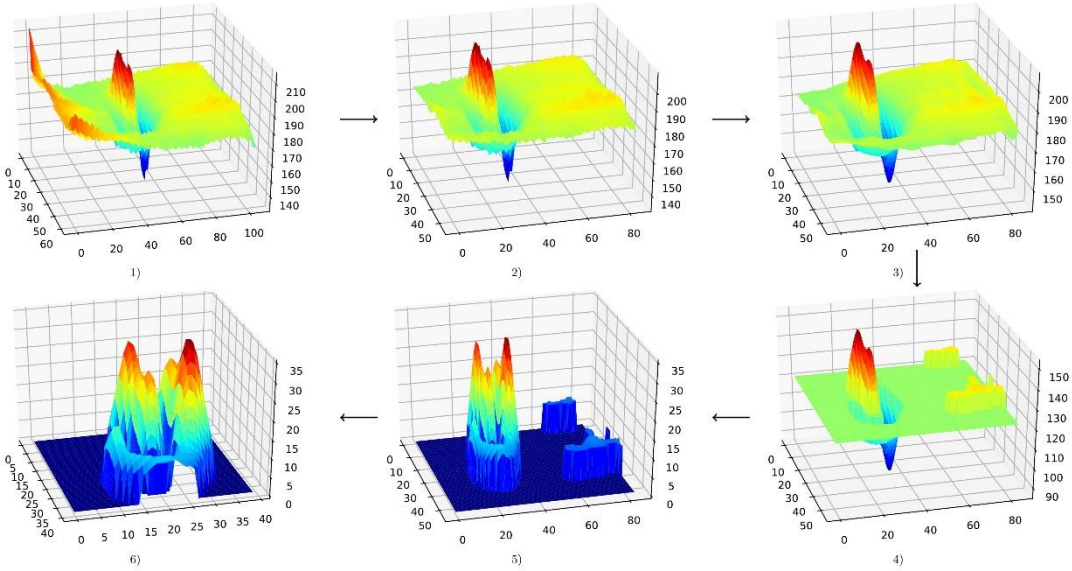

Figure 3. *Preprocess module images - (1): original module image, (2): remove edge noise, (3): remove high-frequency noise, (4) flatten points around background pixel, (5) subtract everything to background pixel and take the absolute, (6): crop or padding to the size of (40, 40).*

The real part and the imaginary part of the C-Scan from the PEC sensor were initially used to calculate the module, as we considered it contains essential information for estimating the crack depth, as demonstrated in [13]. During this stage, we observed that some values of the module images exceeded 255, the maximum value for a pixel in an RGB image. To address this,

we subtracted a value of 60 from all images in the dataset to ensure the module images could be preserved in the storage for training the model.

Next, we applied noise reduction methods to mitigate the impact of noise in the images on crack depth estimation. Two types of denoising techniques were utilized at this stage. The first technique, as adopted in [13], involved cutting out some rows and columns from the image outline to eliminate the effects of edge noises (figure 3. step (1) to step (2)). After removing edge noise, we implemented the second technique to eliminate high-frequency and background noises from the module image to smooth out the module images. The 2D Fourier transform was used to filter out high-frequency noises in the module images [26] (figure 3. Step (2) to step (3)). Regarding the background noises, we determined the frequency of appearance for each pixel in the module image. The pixel with the highest frequency was selected as the "background pixel", and all pixels within a range of ±7 around the background pixel were flattened to the ground pixel value. To comply with the CNN model's requirement of the same input size for all images, we resized all module images to a standardized size of (40, 40) by either cropping or padding them as necessary.

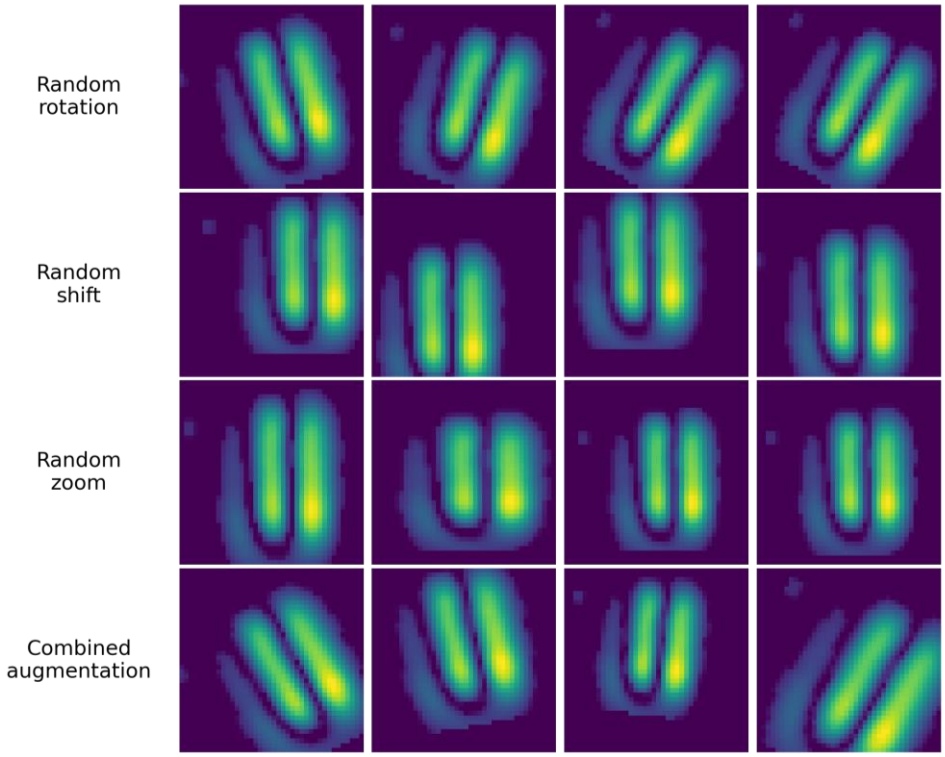

Figure 4. *Examples of image augmentation techniques used in training procedure.*

In NDE, data acquisition procedures can be complex, leading to limited data availability. Consequently, we adopted image augmentation techniques to increase the amount of training data for the CNN model, thus preventing overfitting and enhancing generalization. As we considered the pixel values in the module images to be the most significant feature related to crack depth, we applied augmentation methods that do not alter pixel values, such as random rotation and random shift. For random rotation, we set the rotation range to 45 degrees, considering that we also use horizontal and vertical flips. To account for the small size of the original images, we opted for a 25% range for random zoom and random shift. This choice ensures that most of the crack remains visible within the image frame. Figure 4 illustrates some examples of the images generated using these image augmentation techniques.

2.3. Models

CNN has emerged as a revolutionary network architecture, reshaping the landscape of Machine Learning in the field of Computer Vision in recent years. Specifically designed for image-based tasks, CNNs employ the mathematical operation of "convolution" to replace conventional matrix multiplication. This technique utilizes learnable filters known as kernels, which efficiently capture crucial information from images. By leveraging shared-weight matrices, CNNs self-adapt and extract essential features from the data, making them highly effective in various tasks. Typically, these kernels are of small size (e.g., 3x3, 5x5), reducing model complexity and facilitating the stacking of additional layers for enhanced learning capabilities. In essence, the CNN component serves as a powerful feature extractor, outperforming traditional machine learning methods that often rely on manually crafted features. This characteristic highlights the superiority of CNNs in tackling complex computer vision challenges.

Our model comprises two main parts. The base section consists of a fundamental CNN architecture, featuring four convolution blocks in a sequential order: Convolution, ReLU activation (to increase the training speed and reduce the phenomenon of vanishing gradient), and MaxPooling (figure 5). Each block is responsible for extracting critical features from the preprocessed images. The second part of our model constitutes a fully connected neural network with two layers. The entire model architecture is depicted in figure 5, showcasing the synergy of the base CNN and the fully connected neural network, which culminate in an effective and accurate approach for crack detection and depth estimation.

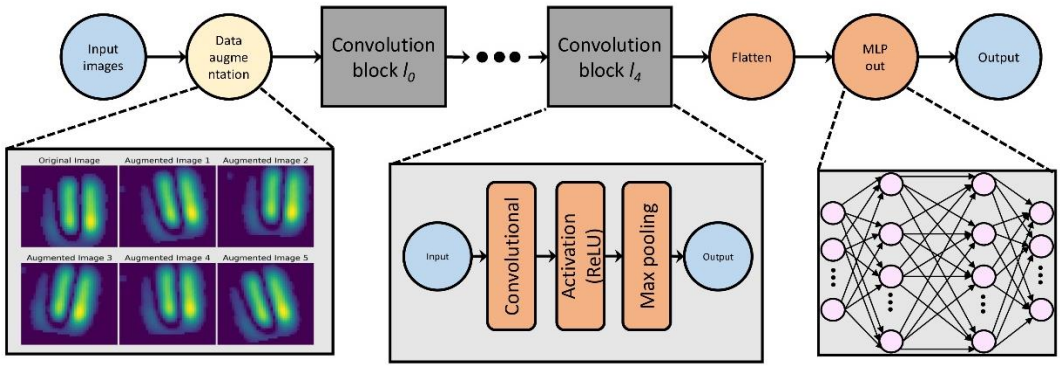

Figure 5. *Data flow of the network and the architecture of CNN model.*

We trained this base CNN separately on two types of cracked specimens. To optimize the model's performance, we employed a random search to tune the hyperparameters of the network within specific search boundaries. The hyperparameters include the number of convolution blocks (ranging from 2 to 5), the number of kernels in each layer (selected from [64, 128, 256]), the number of nodes in the dense layer (chosen from [64, 128]), and the learning rate of the Adam optimizer (ranging from 1e-5 to 1e-3). Our final model has (64, 128, 256, 256) kernels of size (3x3) for each convolution block, while the fully connected network has 256 nodes for the first layer and one node for the second layer, responsible for outputting the crack's depth value in the given image.

In our model's training process, K-fold cross-validation is implemented to ensure the robustness of the proposed method due to the limited availability of data. K-fold cross-validation offers several notable advantages in this context. It minimizes the dependence on a single train test split, thereby reducing the likelihood of results being skewed by a specific random split of data. Additionally, it provides a more comprehensive assessment of the model's generalization

capability, as it evaluates performance across various data permutations. By systematically rotating through the folds, the method ensures that each data point is eventually used for both training and validation, maximizing the utilization of the available information. These advantages collectively enhance the dependability and credibility of the network's outcomes in addressing challenges associated with limited datasets. By using K-fold cross validation, the dataset is divided into four folds for each type of fatigue, determined by the number of samples in each category. For every fold, a new model with the same architecture is trained using data from the remaining three folds, and then evaluated on the specific fold in question. The training procedure iterates through all folds, and the overall performance is computed as the average across all folds.

The development of our model's code was carried out using the Keras library [27] in the Python programming language. The pre-processing procedures, on the other hand, were undertaken using the Numpy [28] and Scipy [29] packages.

3. RESULTS AND DISCUSSION

We evaluate our methods using the following metrics: Mean Square Error (MSE), Relative Error (RE) and Mean Relative Error (MRE). Let $\hat{p}_i$ and p_i are the estimation results and actual values of the crack depth of sample i ($i=1, 2, ..., N$) where N is the number of samples, the used evaluation metrics are calculated as equations (1-3):

- Mean Square Error:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2 \quad (1)$$

- Relative Error:

$$RE_i = \left|\frac{\hat{p}_i - p_i}{p_i}\right| * 100\% \quad (2)$$

- Mean Relative Error:

$$MRE = \frac{1}{N}\sum_{i=1}^{N} RE_i \quad (3)$$

Our CNN model was trained separately for each type of crack data. The results, encompassing predicted values for individual samples, as well as their corresponding Relative Error, Mean Square Error, and the Mean Relative Error for each cross validation fold, have been tabulated below. Table 1 reports the results for the first type of data (samples M11, M12, M13, and M14), while table 2 presents the results for the second type of data (samples M21, M22, M23, and M24).

The MSE values of the CNN model over 4 testing folds exhibit a range from 165.64 to 2494 for the two models trained with distinct crack type data.

Concerning the first crack type data, most samples display predictions featuring low Relative Error values, predominantly under 10%. A couple of exceptions exist, notably involving instances of smaller crack depth, such as sample M13 (with RE values of 10.58% and 13.87%) and the minutest crack within sample M14 (12.86%). Nevertheless, the MRE values across all four cross-validation folds stand at 5.22%, 3.48%, 10.09%, and 7.37%, respectively. These MRE values are deemed acceptable within the context of industrial applications.

The second crack type data similarly yields promising outcomes, with only two instances exhibiting RE values exceeding 10% (sample M21 with a crack depth of 400 μm at 10.17% RE, and sample M24 with a crack depth of 400 μm at 20.94% RE). Importantly, all MRE values across the four folds within this crack type's results remain below 10%.

Table 1. *Result of CNN model on first crack type data.*

Fold	Real values (μm)	Predicted values (μm)	MSE	RE (%)	MRE (%)	
Fold 1 (M11)	200	183.05	385.04	8.47	5.22	6.54
	400	370.57		7.36		
	750	748.67		0.18		
Fold 2 (M12)	150	155.28	165.64	3.52	3.48	
	350	369.52		5.58		
	700	709.38		1.34		
Fold 3 (M13)	200	178.84	1491.51	10.58	10.09	
	350	301.46		13.87		
	700	740.87		5.84		
Fold 4 (M14)	150	130.71	494.42	12.86	7.37	
	300	285.02		4.99		
	700	670.22		4.25		

Furthermore, it's worth noting that the MRE values for both crack types, encompassing all four folds, hover around 6%, indicative of the consistency and reliability of the model's predictions. Especially, for cracks with a standard depth of 800 µm, the MRE calculated on both samples set is approximately 2.77%. The empirical findings from our testing affirm the efficacy of our method, even for cracks with diminutive depths. This robust performance underscores the considerable potential for practical application within industrial settings.

Table 2. *Result of CNN model on second crack type data.*

Fold	Original values (μm)	Predicted values (μm)	MSE	RE (%)	MRE (%)	
Fold 1 (M21)	200	183.23	1010.68	8.38	7.35	6.12
	400	359.34		10.17		
	950	916.87		3.49		
Fold 2 (M22)	200	204.64	242.20	2.32	3.09	
	400	426.34		6.58		
	900	896.64		0.37		
Fold 3 (M23)	200	196.19	1089.09	1.91	4.31	
	400	375.60		6.10		
	1050	998.45		4.91		
Fold 4 (M24)	200	187.09	2494.36	6.46	9.74	
	400	316.24		20.94		
	950	932.66		1.82		

4. CONCLUSIONS

This study has presented a comprehensive framework for crack depth estimation using advanced image analysis techniques and a CNN model. By effectively addressing the challenges posed by noise in the images, our proposed approach demonstrates its robustness in crack depth prediction across various crack types, including instances with small depths. Our method yielded a remarkably low mean relative error of approximately 6% across the four testing folds for both crack data types. The achieved results showcase a high level of accuracy and reliability, highlighting the potential of our methodology for real-world industrial applications.

The overall results demonstrate that techniques from the Computer Vision field can be applied to NDE and yield improved outcomes. In future studies, we intend to expand our focus beyond mere crack depth estimation. Our research endeavors will encompass the comprehensive detection of not

only crack depth but also its spatial position and geometric attributes. To facilitate this, we will curate purpose-specific datasets and leverage the strengths of advanced computer vision techniques.

REFERENCES

[1]. H. S. Munawar, A. W. Hammad, A. Haddad, C. A. P. Soares, and S. T. Waller, "*Image-based crack detection methods: A review*," Infrastructures, **vol. 6**, no. 8, pp. 115-135, (2021).

[2]. N. Yusa, H. Huang, and K. Miya, "*Numerical Evaluation of the Ill-posedness of Eddy Current Problems to Size Real Cracks*", NDT&E International, **vol. 40**, pp. 185-191, (2007).

[3]. Z. Xia, R. Huang, Y. Shao, X. Bai, and W. Yin, "*Estimation of defect depth on plates by eddy-current coil array*", Sensors and Actuators A: Physical, p. 115114, (2024).

[4]. Y. Liu *et al.*, "*Depth quantification of rolling contact fatigue crack using skewness of eddy current pulsed thermography in stationary and scanning modes*", NDT \& E International, **vol. 128**, p. 102630, (2022).

[5]. R. Gansel, H. J. Maier, and S. Barton, "*Detection and characterization of fatigue cracks in butt welds of offshore structures using the eddy current method*", Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, **vol. 6**, no. 2, p. 21001, (2023).

[6]. M. Zergoug, G. Kamel, and S. Lebailia, "*Characterization of the corrosion by eddy current*," EUROCORR, **vol. 36**, pp. 1-7, (2004).

[7]. D. C. Copley, "*Eddy-current imaging for defect characterization*", in: Review of Progress in Quantitative Nondestructive Evaluation: Volume 2A, Springer, pp. 1527–1540, (1983).

[8]. L. Xie, B. Gao, G. Tian, J. Tan, B. Feng, and Y. Yin, "*Coupling pulse eddy current sensor for deeper defects NDT*," Sens. Actuators Phys., **vol. 293**, pp. 189–199, (2019).

[9]. D. Kim, L. Udpa, and S. Udpa, "*Remote field eddy current testing for detection of stress corrosion cracks in gas transmission pipelines*," Mater. Lett., **vol. 58**, no. 15, pp. 2102–2104, (2004).

[10]. S. Xie, Z. Duan, J. Li, Z. Tong, M. Tian, and Z. Chen, "*A novel magnetic force transmission eddy current array probe and its application for nondestructive testing of defects in pipeline structures*," Sens. Actuators Phys., **vol. 309**, p. 112030, (2020).

[11]. Z. Chu, Z. Jiang, Z. Mao, Y. Shen, J. Gao, and S. Dong, "*Low-power eddy current detection with 1-1 type magnetoelectric sensor for pipeline cracks monitoring*," Sens. Actuators Phys., **vol. 318**, p. 112496, (2021).

[12]. Z. Tong *et al.*, "*Quantitative mapping of depth profile of fatigue cracks using eddy current pulsed thermography assisted by PCA and 2D wavelet transformation*," Mech. Syst. Signal Process., **vol. 175**, p. 109139, (2022).

[13]. L. T. Cung, T. D. Dao, P. C. Nguyen, and T. D. Bui, "*A model-based approach for estimation of the crack depth on a massive metal structure*," Meas. Control, **vol. 51**, no. 5–6, pp. 182–191, (2018).

[14]. T.-D. Bui, V.-D. Pham, and T.-L. Cung, "*Multilayer perceptron neural network and eddy current technique for estimation of the crack depth on massive metal structures*," J. Mil. Sci. Technol., no. 77, pp. 3–12, (2022).

[15]. E. Mohseni, D. R. França, M. Viens, W. F. Xie, and B. Xu, "*Finite element modelling of a reflection differential split-D eddy current probe scanning surface notches*," J. Nondestruct. Eval., **vol. 39**, pp. 1–14, (2020).

[16]. M. Jesenik, V. Gorican, A. Hamler, and M. Trlep, "*Finding a crack in a material and determining of depth*," IET 8th International Conference on Computation in Electromagnetics, 11-14 April, Wroclaw, pp. 1-2, (2011).

[17]. Z. Wang and Y. Yu, "*Traditional eddy current–pulsed eddy current fusion diagnostic technique for multiple micro-cracks in metals*," Sensors, **vol. 18**, no. 9, p. 2909, (2018).

[18]. L. Tian, Y. Cheng, C. Yin, X. Huang, B. Zhang, and L. Bai, "*Data-Driven Method for the Measurement of Thickness/Depth Using Pulsed Eddy Current.*," Sens. Mater., **vol. 29**, no. 9, pp. 1325-1338, (2017).

[19]. F. Nafiah, A. Sophian, M. R. Khan, S. B. Abdul Hamid, and I. M. Zainal Abidin, "*Image-based feature extraction technique for inclined crack quantification using pulsed eddy current*," Chin. J. Mech. Eng., **vol. 32**, no. 1, pp. 1–9, (2019).

[20]. M. Smetana, L. Behun, D. Gombarska, and L. Janousek, "*New Proposal for Inverse Algorithm Enhancing Noise Robust Eddy-Current Non-Destructive Evaluation*," Sensors, **vol. 20**, no. 19, p. 5548, (2020).

[21]. T. Meng *et al.*, "*Depth evaluation for metal surface defects by eddy current testing using deep residual convolutional neural networks*," IEEE transactions on instrumentation and measurement, **vol. 70**, pp. 1–13, (2021).
[22]. Z. Zeng, Y. Li, L. Huang, and M. Luo, "*Frequency-domain defect characterization in pulsed eddy current testing*," Int. J. Appl. Electromagn. Mech., **vol. 45**, no. 1–4, pp. 621–625, (2014).
[23]. T. Chen, G. Y. Tian, A. Sophian, and P. W. Que, "*Feature extraction and selection for defect classification of pulsed eddy current NDT,*" Ndt E Int., **vol. 41**, no. 6, pp. 467–476, (2008).
[24]. R. Edwards, A. Sophian, S. Dixon, G.-Y. Tian, and X. Jian, "*Dual EMAT and PEC non-contact probe: applications to defect testing,*" NDT E Int., **vol. 39**, no. 1, pp. 45–52, (2006).
[25]. G. Y. Tian and A. Sophian, "*Defect classification using a new feature for pulsed eddy current sensors,*" Ndt E Int., **vol. 38**, no. 1, pp. 77–82, (2005).
[26]. M. Elad and M. Aharon, "*Image denoising via sparse and redundant representations over learned dictionaries,*" IEEE Trans. Image Process., **vol. 15**, no. 12, pp. 3736–3745, (2006).
[27]. F. Chollet and others, "Keras." 2015, [Online]. Available: https://keras.io.
[28]. C. R. Harris *et al.*, "*Array programming with NumPy*," Nature, **vol. 585**, no. 7825, pp. 357–362, (2020).
[29]. P. Virtanen *et al.*, "*SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*," Nat. Methods, **vol. 17**, pp. 261–272, (2020).

TÓM TẮT

Ước lượng độ sâu vết nứt với mạng nơ ron tích chập và phân tích hình ảnh sử dụng dữ liệu dòng điện xoáy

Nghiên cứu này trình bày một phương pháp ước lượng độ sâu vết nứt, bằng cách kết hợp các kỹ thuật phân tích hình ảnh và mô hình Mạng Nơ-ron Tích chập (CNN). Mục tiêu của nghiên cứu là cải thiện độ chính xác và mức độ tin cậy trong việc dự đoán độ sâu các vết nứt nhỏ, đặc biệt đối với các vết nứt có độ sâu dưới 1 milimet. Nghiên cứu đã sử dụng loạt các kỹ thuật tiền xử lý và tăng cường dữ liệu nhằm giảm ảnh hưởng của nhiễu trong dữ liệu đo. Phương pháp đề xuất đã được thử nghiệm và đánh giá trên dữ liệu thực tế. Kết quả cho thấy, sai số tương đối (Mean Relative Error - MRE) đạt khoảng 6%, thể hiện tính hiệu quả và tiềm năng ứng dụng tốt của phương pháp đề xuất trong thực tế công nghiệp. Thêm nữa, việc áp dụng kỹ thuật xử lý hình ảnh, kết hợp sử dụng mạng nơ ron tích chập cho bài toán Đánh giá Không Phá Hủy (NDE), mở ra một cách tiếp cận mới cho bài toán phát hiện và đặc tính hóa các vết nứt cực nhỏ trên bề mặt, sử dụng phương pháp dòng điện xoáy truyền thống.

Từ khoá: Đánh giá không phá hủy; Kỹ thuật dòng điện xoáy; Mạng nơ ron tích chập; Tăng cường dữ liệu.