A Mixed Samples-Driven Methodology Based on Denoising Diffusion Probabilistic Model for Identifying Damage in Carbon Fiber Composite Structures

Peng Chen, Chaojun Xu, Zhigang Ma, and Yaqiang Jin

***Abstract*— X-ray imaging is a common nondestructive detection method for carbon fiber composite structures (CFCSs) that is useful in identifying damage in CFCS-cored wires. In recent years, deep learning models that incorporate classification and objection detection have become frequently utilized by the nondestructive testing industry. These models typically rely on the assumption that there are sufficient annotated failure samples from history that have been measured and can be used for training. Unfortunately, in real-world measurements, it is often challenging to obtain these types of samples. To address the issue of small sample size in such scenarios of real-world field testing, this article propose a mixed samples-driven methodology based on the denoising diffusion probabilistic model (DDPM) for identifying damage in CFCS. First, new samples are synthesized through DDPM module to improve the robustness of a small sample size. Then, the synthesized sample, along with a small number of authentic samples measured from real-world testing, are then integrated and fed into a DenseNet-based module. Lastly, the mixed samples-driven architecture is then constructed and employed to diagnose the damage of CFCS. The effectiveness of this approach is demonstrated through experiments in real-world field testing.**

***Index Terms*— Carbon fiber composite structures (CFCSs), damage diagnosis, data augmentation, denoising diffusion probabilistic model (DDPM), nondestructive testing.**

I. Introduction

CARBON fiber composite structures (CFCSs) have been extensively utilized in a variety of industries, including wind turbine blade manufacturing, automotive production, and the production of power grid equipment [1]. CFCS-cored aluminum conductor composite core (ACCC) wire has gained popularity for use in power grid equipment due to its excellent conductivity, lightweight properties, high tensile strength, and corrosion resistance. However, CFCS-cored ACCC wire is often subjected to harsh conditions such as high temperatures, corrosion, fatigue, creep, and prolonged stress, which can lead to damage and various types of failures. These failures can inevitably result in unexpected maintenance expenses. An effective damage diagnosis approach for overhead ACCC wire can help prevent failures by providing early warning and enabling better maintenance strategies, ultimately reducing the costs of operation and maintenance.

Historically, the assessment of damage to ACCC wires has been carried out by human inspection, which can be costly, risky, and time-consuming. In contrast to manual detection, several nondestructive technologies have been developed for overhead ACCC wires. These includes sensors attached to the wires, such as ultrasonic gilded wave (UGW) [2], power line carrier [3], as well as sensors integrated into robotics, such as infrared ray [4], electromagnetic induction [5], eddy current [6], stress [7], sound wave [8], and others. However, some of these technologies may be subject to factors such as ambient illumination that can impact their accuracy, and certain materials used in ACCC wire construction may not be compatible with certain technologies, such as eddy current-based methods. In some cases, the use of a magnetic cover on the composite core of ACCC wires may be required to address these issues when using nondestructive technologies for detection in the grid [9].

Despite their potential, the previously mentioned techniques have proven inadequate for reliable damage detection [10]. According to recent research [11], [12], X-ray imaging has demonstrated potential in identifying internal ACCC wire damage using nondestructive detection techniques. One of its advantages is that it can transmit X-rays through thick composite components, making it possible to inspect thicker materials. Moreover, X-ray imaging can detect both surface and subsurface defects and provide a record of the inspection. Deep learning algorithms, particularly those involving classification and object detection, have been widely used in the field of nondestructive testing [12], [13], [14], [15], [16], with

Manuscript received 15 January 2023; revised 18 March 2023; accepted 30 March 2023. Date of publication 17 April 2023; date of current version 1 May 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 52105111, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515010859, in part by the Guangdong Provincial Science and Technology Special Fund Project under Grant STKJ2021171, and in part by the STU Scientific Research Initiation Grant under Grant NTF21029. The Associate Editor coordinating the review process was Dr. Bruno Albuquerque De Castro. *(Peng Chen and Chaojun Xu contributed equally to this work.) (Corresponding author: Peng Chen.)*

Peng Chen is with the College of Engineering and the Key Laboratory of Intelligent Manufacturing Technology, Ministry of Education of China, Shantou University, Shantou, Guangdong 515063, China (e-mail: dr.pengchen@foxmail.com).

Chaojun Xu and Zhigang Ma are with the College of Engineering, Shantou University, Shantou, Guangdong 515063, China.

Yaqiang Jin is with the Center for Structural Acoustics and Machine Fault Diagnosis, Qingdao University of Technology, Qingdao 266520, China, and also with Qingdao Mingserve Technology Company Ltd., Qingdao 266041, China.

Digital Object Identifier 10.1109/TIM.2023.3267522

researchers developing robots equipped with portable X-ray generators and deep learning-based defect diagnostic systems to detect ACCC wire defects. Gao et al. [17] developed a robot that includes a portable X-ray generator, a DR detecting pane for gathering X-ray images, and a deep learning-based defect diagnostic system combined with manual evaluation to discover defects to ACCC wires. Wei et al. [18] provide an automatic defect detection approach based on deep convolution network (DCN) for defect detection of the aluminum conductor carbon fiber composite core through X-ray images. However, these approaches, particularly the supervised deep learning algorithms mentioned above, require a large number of labeled failure data for training. In practice, obtaining sufficient amounts of labeled failure samples can be difficult, making these approaches somewhat challenging to implement in real-world industrial applications.

In real-world testing, the quality of captured X-ray images is often compromised by noise, spurious edges, and other environmental interference. This can be especially problematic when examining CFCSs, as the low contrast of the carbon core and the high energy absorption of the outer steel core can make it difficult to identify cracks. This can result in data imbalance, where the number of samples for a specific type of defect may be limited, leading to potentially misleading deep learning performance estimates [19]. To address this issue, researchers [20] have explored ways to design deep learning models to avoid such inaccurate estimates, particularly in cases of small sample sizes. Generative adversarial networks (GANs) [21] have frequently been regarded as state-of-the-art generative models in terms of sample quality, but they are known for being difficult to train and prone to mode collapse [22]. On the other hand, diffusion models [23], which are likelihood-based, provide more diversity and a steady training process, may be a more practical solution for dealing with these issues.

As discussed above, few investigations for addressing the challenge of small sample sizes in nondestructive testing of CFCS-cored ACCC wires. In this article, a mixed samples-driven methodology based on the denoising diffusion probabilistic model (DDPM) for identifying damage in CFCSs is proposed. This approach leverages the parameterization of diffusion models to synthesize a certain number of high-quality samples through equivalency with denoising score matching and annealed Langevin dynamics. These synthesized samples are then combined with a limited number of real samples and fed into a dense-based diagnosis module to construct the mixed samples-driven model. Testing on a dataset from actual field testing shows that the proposed approach is effective at recognizing defects in CFCS-cored ACCC wires, such as fracture, sawing, shifting, and splitting.

The structures of this article is outlined as follows. In Section II, the DDPM and the primary topology of DenseNet are discussed. The proposed mixed samples-driven methodology is then presented in Section III. A case study from real-world field testing is presented in Section IV, and the fundamental perspective of the proposed methodology and a summary of the results are provided in the concluding section, Section V.

II. Related Theory

A. Denoising Diffusion Probabilistic Model

In this section, the derivation of the DDPM is examined. It is noted that several simplifying assumptions were made in this derivation, including the use of a fixed noising process q that adds diagonal Gaussian noise at each time-step, as reported in [20].

Let $q(x_0)$ be the density of input data x_0, with the index 0 representing the fact that the data is uncorrupted (raw). To construct a forward noising process q that yields underlying data x_1 through x_T from this set of input data, the Markovian process can be utilized. This process involves introducing Gaussian noise at each time step t with variance β_t, where β_t is a value in the range $(0, 1)$. Specifically, the process can be expressed as follows:

$$q(x_1, \ldots, x_T \mid x_0) := \prod_{t=1}^{T} q(x_t \mid x_{t-1}) \tag{1}$$

$$q(x_t \mid x_{t-1}) := \mathcal{N}\left(x_t; \sqrt{1-\beta_t} \cdot x_{t-1}, \beta_t \cdot \mathbf{I}\right) \quad \forall t \in \{1, \ldots, T\} \tag{2}$$

where T is the number of diffusion steps, $\beta_1, \ldots, \beta_T \in [0, 1)$ are hyper-parameters that denote the variance across diffusion steps, and $\mathbf{I}$ is the identify matrix with dimensions equal to those of the input data x_0. The function $\mathcal{N}(x; \mu, \sigma)$ represents the normal distribution with mean μ and covariance σ that produces x. This recursive formulation appears to have the unique advantage of allowing direct sampling of x_t when t is randomly selected from a uniform distribution, as expressed in the following statement: $\forall t \sim \mathcal{U}(1, \ldots, T)$:

$$q(x_t \mid x_0) = \mathcal{N}\left(x_t; \bar{\alpha}_t \cdot x_0, \left(1-\bar{\alpha}_t\right) \cdot \mathbf{I}\right) \Longrightarrow x_t = \sqrt{\bar{\alpha}_t} \cdot x_0 + z_t \cdot \sqrt{1-\bar{\alpha}_t} \tag{3}$$

where $z_t \sim \mathcal{N}(0, \mathbf{I})$, $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$.

The posterior $p(x_{t-1} \mid x_t, x_0)$ can be determined using Bayes theorem in term of $\tilde{\beta}_t$ and $\tilde{\mu}_t(x_t, x_0)$, which are specified as follows:

$$\tilde{\beta}_t := \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t. \tag{4}$$

$$\tilde{\mu}_t(x_t, x_0) := \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}x_0 + \frac{\sqrt{\alpha_t}\left(1-\bar{\alpha}_{t-1}\right)}{1-\bar{\alpha}_t}x_t. \tag{5}$$

$$p(x_{t-1} \mid x_t, x_0) = \mathcal{N}\left(x_{t-1}; \tilde{\mu}(x_t, x_0), \tilde{\beta}_t \mathbf{I}\right). \tag{6}$$

A new sample can be generated by starting with Gaussian noise at $x_T \sim \mathcal{N}(0, \mathbf{I})$ and then proceeding backwards using the reverse distribution $p(x_{t-1} \mid x_t)$, which is defined as $\mathcal{N}(x_{t-1}; \mu(x_t, t), \Sigma(x_t, t))$. Once this reverse distribution has been determined, it can be approximated using a neural network that takes as input the noisy data x_t and the time-step embedding t, and uses this information to predict the mean $\mu_\theta(x_t, t)$ and covariance $\Sigma_\theta(x_t, t)$

$$p_\theta(x_1, \ldots, x_T) := p(x_T) \cdot \prod_{t=1}^{T} p_\theta(x_{t-1} \mid x_t) \tag{7}$$

$$p_\theta(x_{t-1} \mid x_t) := \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)). \tag{8}$$

A variational autoencoder [24] is created by combining p and q, and the variational lower bound (VLB) can be computed as follows:

$$\mathcal{L}_{\text{vlb}} = \mathbb{E}\Bigg[-\log p_\theta(x_0 \mid x_1) + \text{KL}(p(x_T \mid x_0) \| q(x_T)) + \sum_{t>1} \text{KL}(p_\theta(x_{t-1} \mid x_t, x_0) \| q_\theta(x_{t-1} \mid x_t)) \Bigg]. \quad (9)$$

In (9), the Kullback–Leibler (KL) divergence is used to measure the difference between two probability distributions. With the exception of the first term, each term in the equation is a KL divergence between two Gaussian distributions. The second term does not depend on the parameter θ, and will be close to zero if the forward noising process effectively destroys the data distribution so that $q(x_T \mid x_0) \approx \mathcal{N}(0, \mathbf{I})$. The last term indicates that the neural network is trained to make the posterior $p_\theta(x_{t-1} \mid x_t)$ at each time step t as close as possible to the true posterior of the forward process when conditioned on the original sample. Additionally, it can be shown that the posterior $p_\theta(x_{t-1} \mid x_t)$ is a Gaussian distribution, which allows for closed-form expressions for the KL divergences.

To improve the quality of the output samples, an alternate objective was introduced in [23]. This objective focuses on training a neural network to measure noise from arbitrary instances generated using (3), as follows:

$$\mathcal{L}_{\text{simple}} = \mathbb{E}_{t\sim[1,T]} \mathbb{E}_{x_0\sim p(x_0)} \mathbb{E}_{z_t\sim\mathcal{N}(0,\mathbf{I})} \| z_t - z_\theta(x_t, t) \|^2 \quad (10)$$

where $\mathbb{E}$ is the expected value, and $z_\theta(x_t, t)$ is a network predicting the noise in x_t. In [23], the covariance $\Sigma_\theta(x_t, t)$ to a constant value and the $\mu\theta(x_t, t)$ can be rewrite as follows:

$$\mu_\theta = \frac{1}{\sqrt{\alpha_t}} \cdot \left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \cdot z_\theta(x_t, t) \right). \quad (11)$$

The reverse process is defined by $p_\theta(x_{t-1} \mid x_t)$ [as shown in (8)], but the neural network is not trained to directly estimate the mean and covariance of this distribution. Instead, it is trained to measure noise in the input, and the mean is calculated using (11), with the covariance set to a constant value.

B. Primary Topology of DenseNet

The DenseNet topology has a structure in which every layer of the feed-forward neural network is closely connected to all the layers that follow it [25]. This design allows for maximum information flow throughout the network. The DenseNet topology consists of three dense blocks connected by transition blocks, and each of these blocks is made up of basic blocks.

1) Basic Blocks: The DenseNet architecture is built on a low-level structure called basic blocks. Each basic block is made up of four layers that are used to concatenate the input: a batch normalization layer, a rectified linear unit (ReLU) activation layer, and a 3×3 convolutional layer with a stride of 1. Fig. 1 shows these blocks, which include a convolutional layer with k filters.

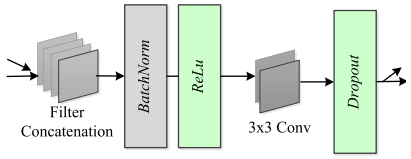

Fig. 1. Basic blocks.

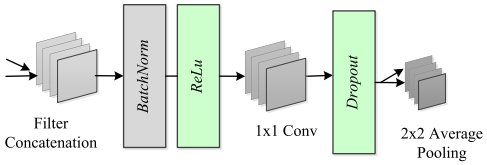

Fig. 2. Transition blocks.

2) Transition Blocks: Transition blocks, which connect the dense blocks in a DenseNet architecture, are constructed using the following sequence of layers: a batch normalization layer, a ReLU activation layer, a 1×1 convolutional layer with a stride of 1 and k output filters, where k is the number of basic blocks per dense block multiplied by the size of the input to the previous dense block. After the convolution operation, dropout is applied at a rate of 0.2, followed by a 2×2 average pooling layer with a stride of 2. Fig. 2 shows the transition layers.

3) Dense Blocks: The DenseNet architecture used in this work contains three dense blocks, which are made up of an initial input and several basic blocks, and feature permissible skip connections within the dense blocks. The depth of each dense block is determined by dividing the overall depth by the number of trainable layers that are not located in the dense blocks.

III. Proposed Architecture

In this section, a mixed samples-driven architecture for detecting damage in CFCSs is proposed. The architecture, shown in Fig. 3, consists of a DDPM module, a DenseNet-based module [25], and a mixed samples-driven module. The DDPM module is used to synthesize new samples to compensate for the small sample size commonly encountered in real-world field testing. These synthesized samples, along with a small number of authentic samples collected from real-world testing, are then integrated and fed into the DenseNet-based module. The mixed samples-driven architecture is then constructed and employed to diagnose damage in the CFCSs of ACCC wires.

As described in (7) and (8), a new sample $\tilde{x}_0$ can be generated by starting at $\tilde{x}_T$ and then moving backwards using $p_\theta(x_{t-1} \mid x_t)$

$$\tilde{x}_0 = \tilde{x}_T \cdot \prod_{t=1}^{T} p_\theta(x_{t-1} \mid x_t) = \tilde{x}_T \cdot \prod_{t=1}^{T} \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \quad (12)$$

where $\tilde{x}_T$ is drawn from a normal distribution with mean 0 and identity covariance matrix, $\mathcal{N}(0, \mathbf{I})$. The covariance $\Sigma_\theta(x_t, t)$

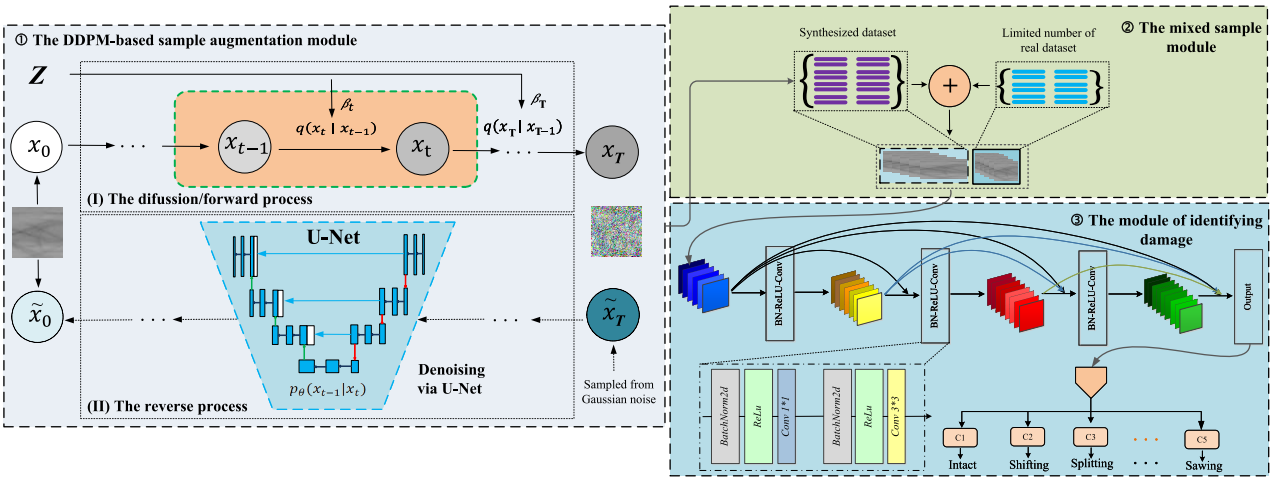

Fig. 3. Proposed mixed samples-driven architecture for damage identifying in CFCS-cored ACCC wires.

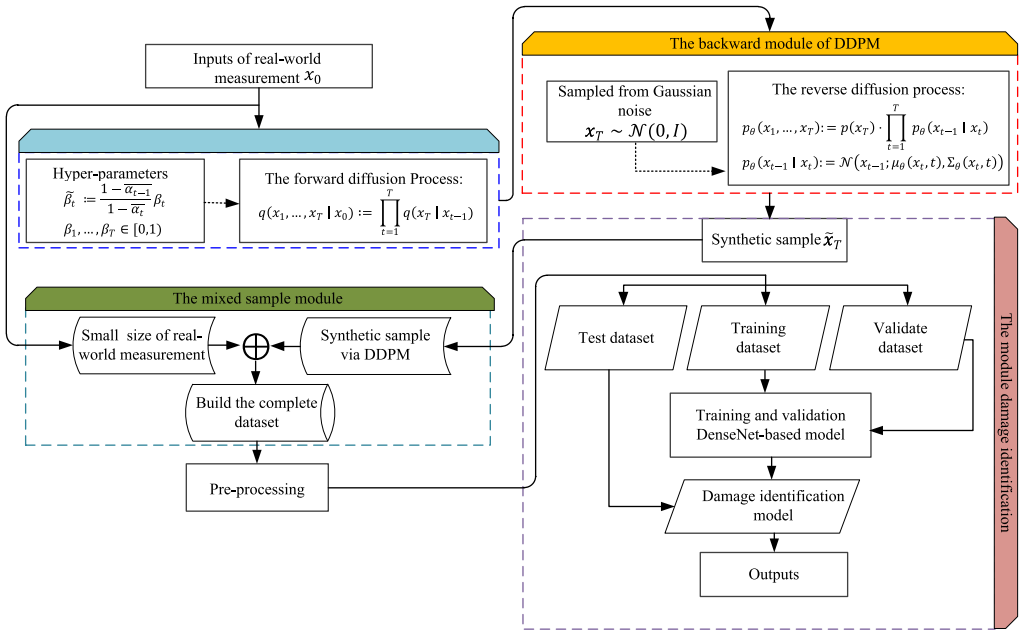

Fig. 4. Flowchart of proposed model.

can be approximated using a neural network, which takes as input the data x_t and the embedding at time-step t, and is used to predict the mean $\mu_\theta(x_t, t)$ and covariance $\Sigma_\theta(x_t, t)$.

After the synthetic sample produced through DDPM's synthesis method, a mixed samples-driven manner is designed to integrate the synthesis and real-measurement, as it is illustrated in Fig. 3. Thus, the mixed sample y_{MIX}can be determined as follows:

$$\begin{aligned} y_{\text{MIX}} &= \mathcal{M}\left(\tilde{x}_0, x_0\right) \\ &= \mathcal{M}\left(\tilde{x}_T \cdot \prod_{t=1}^{T} \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)), x_0\right) \end{aligned} \tag{13}$$

where $\mathcal{M}$ denotes the mixed samples-driven manner, x_0 is the original small size of sample, and $\tilde{x}_0$ is the sample, which is produced by DDPM's synthesis technique.

As mentioned previously and illustrated in Fig. 3, DenseNet is used to enhance the accuracy of damage identification. This is achieved by streamlining the information, strengthening reusable features, and upgrading the features themselves. The output of the rth layer, denoted as z_r, can be determined as follows:

$$z_r = H_r\left(\left[z_0, z_1, \ldots, z_{r-1}\right]\right) \tag{14}$$

where the function H_r (where r is equal to $1, 2, 3, \ldots$) is composed primarily of batch normalization, the ReLU, and convolution. The features produced in layers $0, 1, \ldots, r-1$ are concatenated and represented as $[z_0, z_1, \ldots, z_{r-1}]$.

The output $\tilde{z_r}$ for a small batch data, denoted by B, can be computed as follows:

$$z_{\text{BN}} = \frac{\widehat{z_r} - \mu_{\text{B}}}{\sqrt{\sigma_{\text{B}}^2 + \varepsilon}} \tag{15}$$

$$\tilde{z_r} = \gamma \cdot z_{\mathrm{BN}} + \beta \tag{16}$$

where, the normalization operation of the input data, represented by z_{BN}, is utilized in this case. The rth layer input of BN, denoted by $\widehat{z_r} = [z_0, z_1, \ldots, z_{r-1}]$, is also taken into consideration. The mean and variance of the batch data, indicated by μ_{B} and σ_{B}^2, respectively, are both relevant to the calculation. Additionally, the learnable reconstruction parameters of scale and shift, represented by β and γ, respectively, allow the network to learn and retrieve the original distribution of features. These parameters enable the network to learn and retrieve the original distribution of features.

The damage diagnosis module utilizes a soft-max classifier to train on the feature vectors obtained from the previously mentioned DenseNet-based model. These feature vectors come from a mixed sample, denoted by $\{(y_{\mathrm{MIX}}^{(i)}, l^{(i)}); i \in 1, \ldots, N, l^{(i)} \in 0, \ldots, k-1\}$, where $y_{\mathrm{MIX}}^{(i)}$ represents the ith feature vector of the input sample and l^i is the corresponding label. For each sample in this set, the probability of belonging to a particular category is given by the following equation:

$$p\left(l^{(i)} = j \mid y_{\mathrm{MIX}}^{(i)}; \theta\right) = \frac{e^{\theta_j^T y_{\mathrm{MIX}}^{(i)}}}{\sum_{j=1}^{k} e^{\theta_j^T y_{\mathrm{MIX}}^{(i)}}} \tag{17}$$

where the probability that $y_{\mathrm{MIX}}^{(i)}$ belongs to the jth category, denoted by $p(l^{(i)} = j \mid y_{\mathrm{MIX}}^{(i)}; \theta)$, represents the probability that the sample is divided into different types of damage. This probability is equivalent to the probability that the sample belongs to the jth category.

The function for evaluating damage using soft-max is expressed as follows:

$$\begin{aligned} h_\theta\left(x_{\mathrm{MIX}}^{(i)}\right) &= \begin{bmatrix} p\left(l^{(i)} = 1 \mid y_{\mathrm{MIX}}^{(i)}; \theta\right) \\ p\left(l^{(i)} = 2 \mid y_{\mathrm{MIX}}^{(i)}; \theta\right) \\ \vdots \\ p\left(l^{(i)} = k \mid y_{\mathrm{MIX}}^{(i)}; \theta\right) \end{bmatrix} \\ &= \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T y_{\mathrm{MIX}}^{(i)}}} \begin{bmatrix} e^{\theta_1^T y_{\mathrm{MIX}}^{(i)}} \\ e^{\theta_2^T y_{\mathrm{MIX}}^{(i)}} \\ \vdots \\ e^{\theta_k^T y_{\mathrm{MIX}}^{(i)}} \end{bmatrix} \end{aligned} \tag{18}$$

where θ represents the model parameters and $\sum_{j=1}^{k} e^{\theta_j^T y_{\mathrm{MIX}}^{(i)}}$ is the operation that normalizes the probabilities.

IV. Real-World Field Testing Case Study for Experimental Verification

In this section, an experimental investigation is developed for verifying the efficiency of damage assessment with CFCS-cored ACCC wires, and its flowchart is shown in Fig. 4. First, as shown in Section IV-A, a sample comparison between DDPM's synthetic sample and data collected from a testing field in the real world. The data acquisition and data preprocessing are, respectively, specified in Sections IV-B and IV-C. Afterward, the damage assessment and results-based analysis is provided in Section IV-D. Section IV-E sketches the training process and comparative analysis. Furthermore, through using t-SNE technique, the captured feature manifold is visualized, and Section IV-F illustrates typical indicators for evaluating the proposed framework.

TABLE I
Quantitative Comparison of Real Measurement and Synthesis

Type of defect	Fracture	Sawing	Shifting	Splitting
Cosine similarity	0.981	0.995	0.985	0.997

A. Sample Comparison for Real-Measurement and Synthesis

A comparative case study is conducted to evaluate the quality of a synthetic sample produced using DDPM's synthesis method. The synthetic sample is compared to real-world data collected from a testing field to determine the accuracy and reliability of the synthetic sample. As shown in Fig. 5, any discrepancies or variations between the synthetic sample and the real-world data may be identified by comparing the two sets of data.

The real-measurement and synthesis of the four types of failures, including shifting, splitting, fracture, and sawed, are shown in Fig. 5. The real-measurement and synthesis images provide a detailed look at the severity of the failures and their pattern of occurrence in the experiments. They can identify the distinctive characteristics of each kind of failure, as shown in the zoomed region of Fig. 5. While there may be some small variations among the samples, the majority of them exhibit similar failure characteristics. This information is useful for understanding the causes behind various failure patterns and for developing strategies to prevent future failures. Real-measurement and synthesis images can be utilized to gain valuable insight into how these failures occurred and then utilize that information to improve the quality of the synthetic sample or refine the DDPM model. Furthermore, this makes it possible to analyze the failures in a more precise and consistent manner, making it easier to identify trends and potential explanations of the failures. In short, this comparison is crucial in general to make sure that the synthetic sample generated by the DDPM model accurately reflects the real-world data that it is intended to simulate.

The degree of similarity between real-measurement and synthesis is quantified using a common metric, such as cosine similarity [26], and the analysis's results are presented in Table I. Due to the cosine similarity values being all above 0.980, the quantitative analysis reveals that the synthesis sample from the four types of failure are very comparable to the real measurement.

B. Data Acquisition

To measure the X-ray images in ACCC wires with CFCSs core, a data acquisition system is developed, which can be seen in Fig. 6. The data acquisition system comprises three modules, such as the data collection modules, the data transmission module, and the data analysis module. In data collection module, see Section I in Fig. 6, an image collection component equipped with a portable and cold cathode X-ray source TXR-C1R150P-08 as well as a moving robot is designed to climb

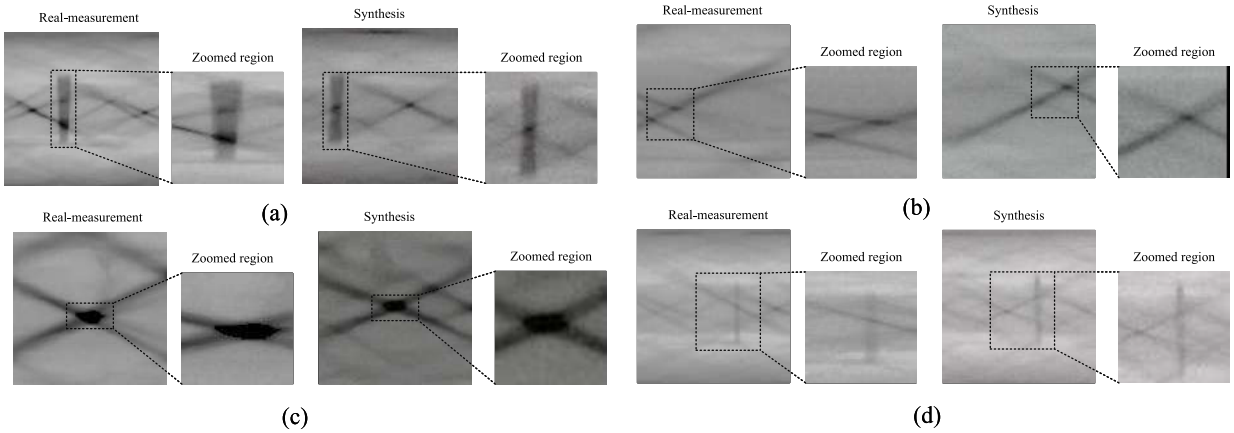

Fig. 5. Sample comparison of synthesis and real-measurement. (a) X-ray image of a shifting sample and it is zoomed region. (b) X-ray image of a splitting sample and it is zoomed region. (c) X-ray image of a fracture sample and it is zoomed region. (d) X-ray image of a sawed sample and it is zoomed region.

on the ACCC wires. The robot is composed of a pair of independent rubber wheels, and it is powered by a dc motor with reducers. The average movement speed is 6.5 m/min. The portable X-ray generator and a digital radiography dashboard are incorporated at the bottom of the data collecting robot, and the deployment location allows the robot to pass obstacles without interfering with the vibration damper. The data of X-ray images could then be uploaded to cloud storage via WIFI graphics transmission (see Section II in Fig. 6). Finally, the measured data can be delivered to a workstation for fault detection and analysis (see Section III in Fig. 6).

In [1], it outlines the major failure accidents of ACCC wires with CFCS core, which can be classified as tensile failure in strain clamps and bending fracture produced by nonstandard tools in manufacture. An ACCC wire cross section diagram can be found in Fig. 7. The typical damaged types can be illustrated in Fig. 8. Fig. 8(b) is an X-ray image of a shifting, as well as cracks with random lines that provide a little risk. Fig. 8(c) depicts an X-ray image of a splitting of a crack with irregular lines that pose a medium threat. Fig. 8(d) illustrates a fracture in which the carbon fiber core has ruptured and the strained section is the exterior aluminum wire. This is a huge weakness. sawed is illustrated in Fig. 8(e) as a vacancy formed by the carbon fiber core being pulled apart following fracture. This condition is exceedingly weakness.

C. Data Preprocessing

Several procedures, including as segmentation, straightening, and Gray-level normalization, should be performed to guarantee the consistency of the training data in position and gray distribution (see in Fig. 9). First, due to gravity's influence, the ACCC wires are not horizontally parallel to the ground in the measured image. Then, straightening and shareholding are necessary to extract straight ACCC lines, enabling ACCC to be located consistently in every image. The coordinates of the centered pixels on image for every vertical line serve as a reference for straightening, and translation is conducted for each line based on the vertical offset. Following that, variations in X-ray incident angle and detector plan response cause brightness inconsistency in the horizontal direction, leading defects to emerge differently in various positions. Thus, the brightness and contrast is utilized for remedying the inconsistency in single image or different images. Lastly, vertical integration is therefore replaced with vertical template convolution to alleviate discontinuities. To achieve similar color distribution, histogram specifications are assessed for contrast discrepancies in various images.

According to Fig. 7, it is discovered that the defect position is determined by the central portion of the raw X-ray image. Consequently, the raw image can be divided into small patches using a sliding window with a length of window size l_{win} and an overlap of l_s. In Fig. 9, it is shown that the patches are overlap to prevent defects along the patch boundary. Since the exterior aluminum wire gap will create interference dark stripes that could be misinterpreted as defects. As a result, the data preprocessing procedure should concentrate only the carbon core area. According to their morphological properties, there are four different types of defects: shifting, splitting, fracture, and sawed, as shown in Fig. 8. In this real-world field testing, the presence of noise, spurious edges, and other unexpected environmental interference typically reduces the quality of the measured X-ray image. The amount of X-ray images that have been collected is therefore quite limited. The measurement of 400 raw X-ray images involves five different types of samples, including intact, shifting, splitting, fractured, and sawed samples. Due to the insufficient defective samples, the DDPM is utilized to produce additional varieties of samples. Ultimately, a total of 6000 raw samples are obtained. The arrangement of these samples between the training and test datasets is shown in Table II, and the size of each sample is $128 \times 128 \times 1$.

D. Damage Assessment and Results-Based Analysis

In this section, the current state-of-the-art models, such as ShuffleNet [27], RepVGG [28], ResNet50 [29], ResNet101 [29], and DenseNet [30] are selected for comparison to evaluate the performance of the proposed architecture in damage identification of CFCS-cored ACCC wires.

An indicator like test accuracy is typically utilized to reveal the capacity of deep learning models in damage categorization.

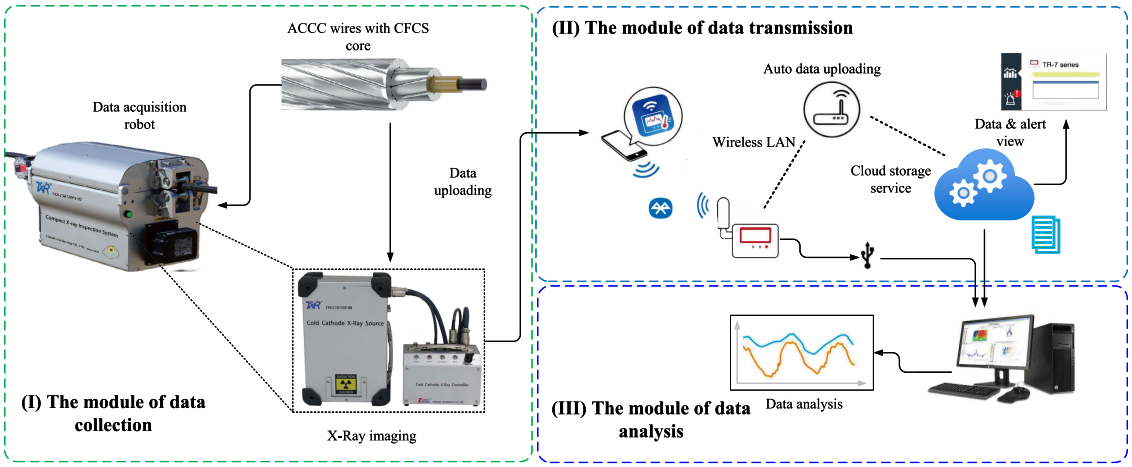

Fig. 6. Damage detection schematic in field deployment testing.

TABLE II
ARRANGEMENT OF TRAINING AND TEST DATASETS

Models	Categories	Training datasets	Test datasets
Existing models	Fracture	400 (real)	100 (real)
	Intact	1200 (real)	100 (real)
	Sawing	400 (real)	100 (real)
	Shifting	400 (real)	100 (real)
	Splitting	400 (real)	100 (real)
Proposed models	Fracture	400 (real) + 800 (synthetics)	100 (real)
	Intact	1200 (real)	100 (real)
	Sawing	400 (real) + 800 (synthetics)	100 (real)
	Shifting	400 (real) + 800 (synthetics)	100 (real)
	Splitting	400 (real) + 800 (synthetics)	100 (real)

Note: Existing models are denoted as ShuffleNet, RepVGG, ResNet50, ResNet101 and DenseNet

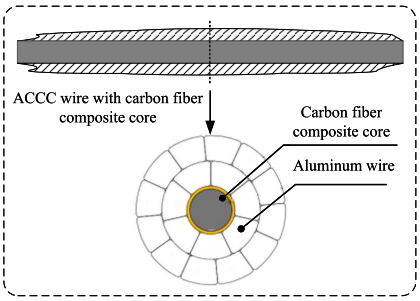

Fig. 7. Cross section diagram of ACCC wires.

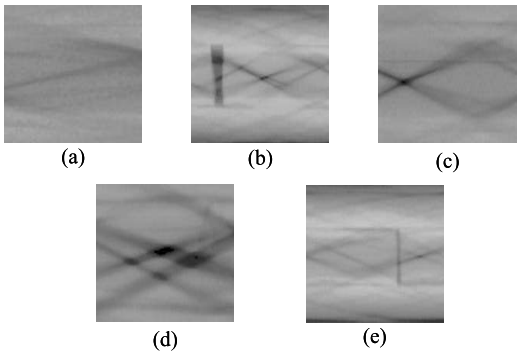

Fig. 8. Typical damaged ACCC wires with CFCS cores. (a) X-ray image of an intact sample. (b) X-ray image of a shifting sample. (c) X-ray image of a splitting. (d) X-ray image of a fracture sample. (e) X-ray image of a sawed sample.

The damage identification results of ShuffleNet, RepVGG, ResNet50, ResNet101, DenseNet are depicted in Fig. 10. The

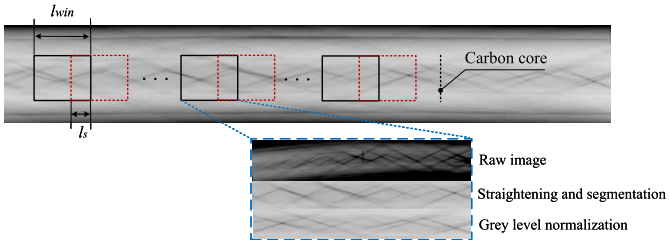

Fig. 9. Data preprocessing schematic illustration.

test accuracy approaches approximately 91.0% for model like ShuffleNet, yet it is unsteady with a greater fluctuation, and the average test accuracy falls to the lowest levels with 89.1% (see in Fig. 11). In contrast to ShuffleNet and RepVGG, the test accuracy for ResNet-based algorithms like ResNet50 and ResNet101 is nearly steady and achieved at around 90.5%. DenseNet is connected to each layer directly and in a feed-forward manner, enabling effective features flow and gradient propagation. This makes DenseNet is very beneficial for tasks like segmentation and image classification. As a whole, the maximum test accuracy of DenseNet is 91.1%. Compare to typical DenseNet, in the zoomed region in Fig. 10, the test accuracy of our proposed models achieves the greater performance, which is roughly to 93.1%, and this indicates that it is performing very well on the test dataset. By feature manifold learning among the mixed samples, which incorporates the features from synthesis sample and a limited number of real field testing sample, our proposed model may

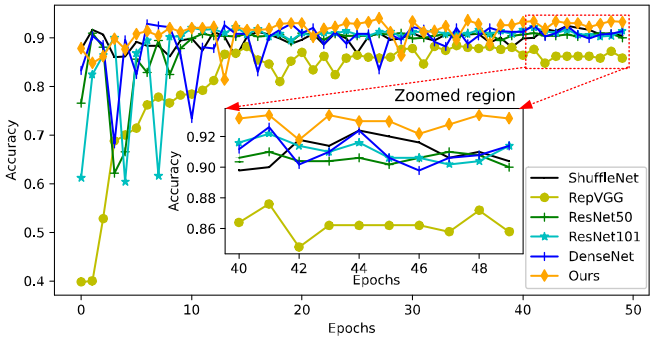

Fig. 10. Comparison of model accuracy across different types of models.

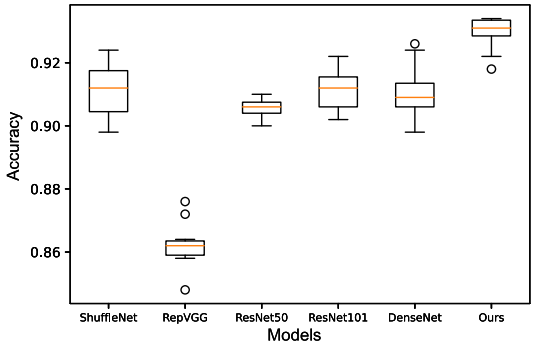

Fig. 11. Box plot for accuracy comparison.

gain more broad diagnostic information and develop into a more effective classifier for damage diagnoses. Furthermore, to visualize the test accuracy distribution and identifying potential outliers of the aforementioned models, a box plot is depicted in Fig. 11. As can be observed, the test accuracy ranges for ShuffleNet and RepVGG are broader, whereas those for ResNet50, DenseNet and our proposed model show a very minor variations.

E. Training Process and Comparative Analysis

A determination of the hyperparameters is essential for the successful training of the proposed architecture. In the earlier work, Chen et al. [14] and Ho et al. [23] offers a number of valuable learning curves that are beneficial for debugging and determining hyperparameters. After each set of pooling layers and fully-connected layers, the activation function used in this case is the ReLU. For the input layer and each activation layer, batch normalization is performed with a momentum value of 0.9. Each convolution layer has 12 input filters added to it. Sixteen basic blocks and 16 transition blocks are connected to each dense block, as it is shown in Fig. 3. The training uses an adaptive moment estimation (Adam) optimizer with a learning rate of 0.0001, and it runs for 50 epochs. The growth rate is set to 12.

In the testing process, there is a correlation between epochs and loss, as seen illustrated in Fig. 12. In the early stages, with steadily increasing epochs, loss value has a significant downward trend in training process. Nevertheless, as the epochs approached 40, the decreasing trend of loss value flattened out. This could be due to the performance of the models discussed above increases with the number of iterative trainings, whereas when a certain number of times is reached,

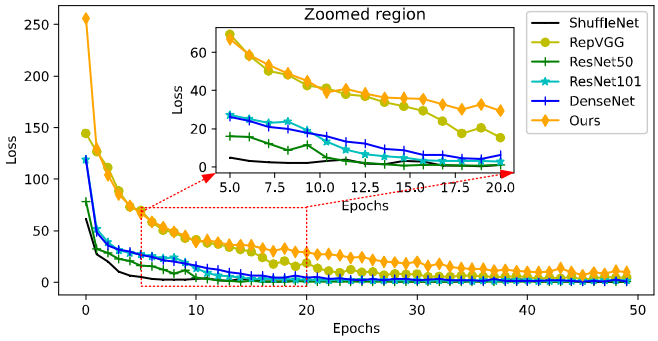

Fig. 12. Training loss in different types of models.

the models performance trends to stabilize. At this stage, extra training is a time and computing resources waster that does not boost better results. Therefore, appropriate epochs are advantageous for training models. The epochs in our proposed model are stable and coincide with those in common models like ShuffleNet. In contrast, models such as RepVGG, ResNet-based, and DenseNet achieves better, as seen in Fig. 12.

In this case, the algorithms are implemented with PyTorch and are running on a platform with an AMD Ryzen 9 5950X 16-core CPU, 64G of DDR4 memory, and an Nvidia RTX3090Ti GPU for the proposed architecture's training.

F. Quantitative Assessment Analysis

To visualize and interpret the captured latent features by the aforementioned models, a dimensionality reduction technique (t-SNE [14]) is utilized for comparison. The feature manifold representation created by t-SNE can be seen in Fig. 13, which offers a clear and understandable perspective of the underlying data. Using t-SNE, it is possible to quickly identify data clusters and patterns that may not have been immediately visible in the past. This enables us to comprehend the data better and aids in decision-making.

The corresponding captured feature manifolds for the typical models, such as ShuffleNet and RepVGG, are illustrated in Fig. 13(a) and (b), receptively. In both models, the scatter points for the intact condition are well-clustered together or separable. This indicates that the models are able to accurately identify and distinguish the intact condition from other states. However, the scatter points for the other states, such as the fracture and splitting modes, do not cluster together. Instead, they are intersected and overlapped with each other, which imply that the models may have difficulty distinguishing between these states. This is most likely a result of the data's intrinsic complexity and the difficulties in precisely identifying and interpreting the hidden properties in these states. To further evaluate the classification models, the confusion matrix is employed to categorize the captured feature into various labels, as shown in Fig. 14(a) and (b). As an illustration, the feature manifold of different types of failures in the RepVGG model are overlapped, with the exception of the intact state. This implies that this model may have difficulty distinguishing between these different types of failure modes.

Likewise, the visualization of the feature manifold representation for the ResNet-based models is depicted in

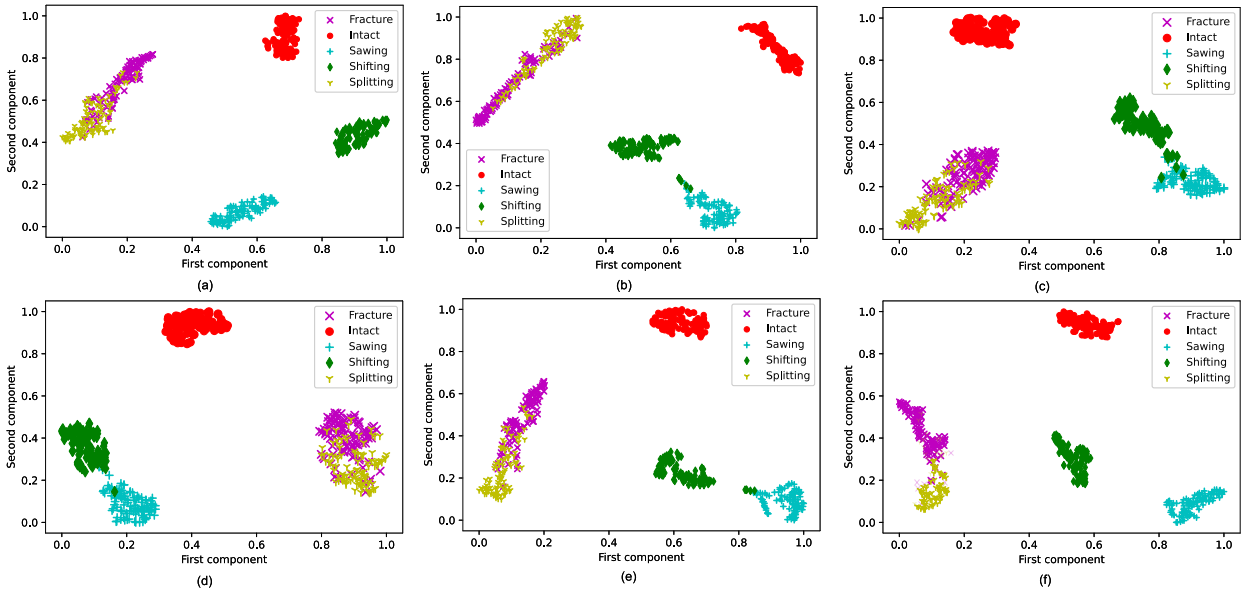

Fig. 13. Feature visualization through t-SNE. (a) ShuffleNet. (b) RepVGG. (c) ResNet50. (d) ResNet101. (e) DenseNet. (f) Ours.

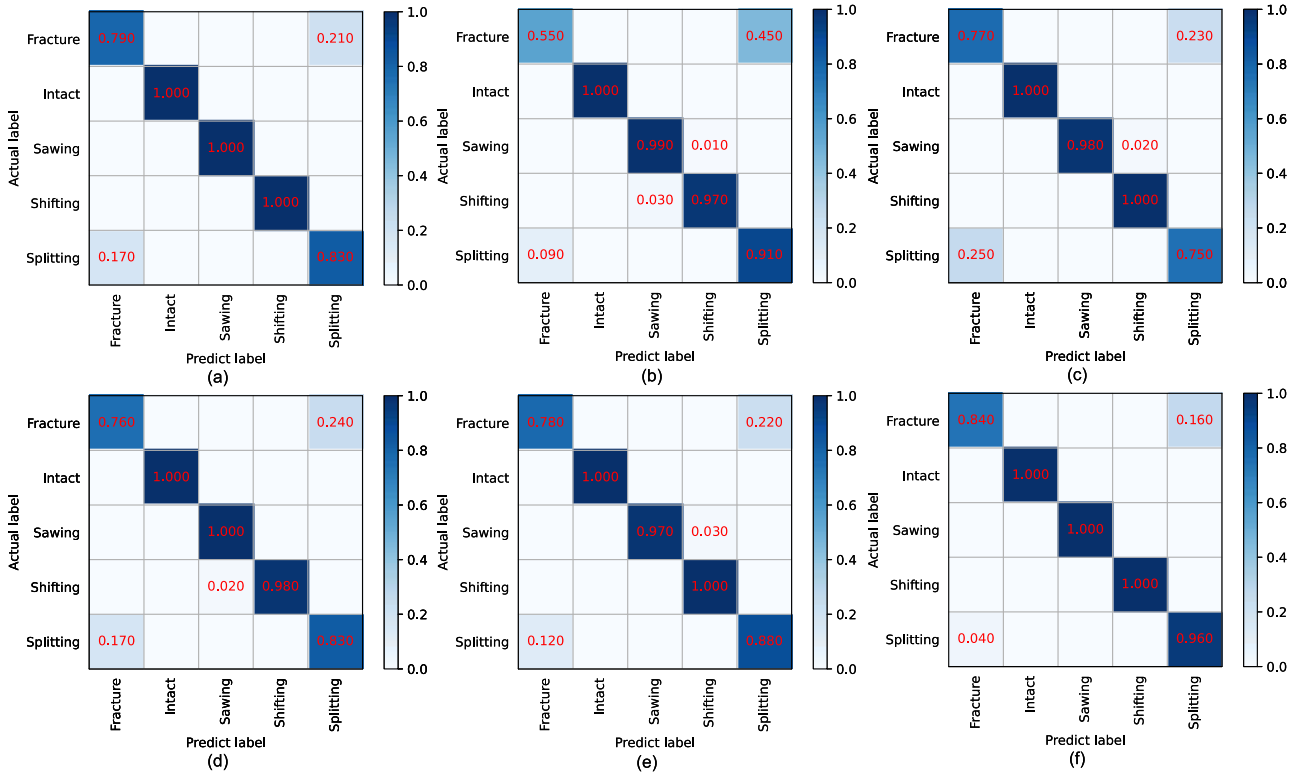

Fig. 14. Classifying performance through confusion matrix. (a) ShuffleNet. (b) RepVGG. (c) ResNet50. (d) ResNet101. (e) DenseNet. (f) Ours.

TABLE III
$F1$ SCORE OF DIFFERENT APPROACHES

Approaches	ShuffleNet	RepVGG	ResNet50	ResNet101	DenseNet	Ours
F1 score	0.924	0.873	0.857	0.876	0.919	0.939

Fig. 13(d) and (e). Numerous manifestations of the failures, such as splitting and fracture, extensively overlap. The set of feature manifold overlap indicates that the model is unable to effectively capture the underlying patterns and relationships in the data, which are necessary to discriminate between various forms of failure.

However, since relatively few features manifold distribution of splitting and fracture are involved, as shown in Fig. 13(f), the feature manifold distribution from almost all failures can be clearly distinguished by our proposed model. This reveals that our proposed model can accurately capture and discriminate the latent characteristics generated

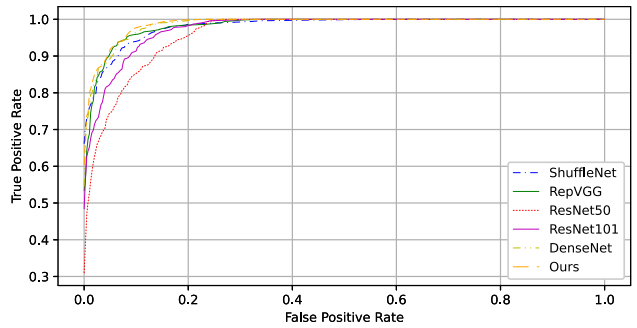

Fig. 15. ROC curve for evaluating different kinds of models.

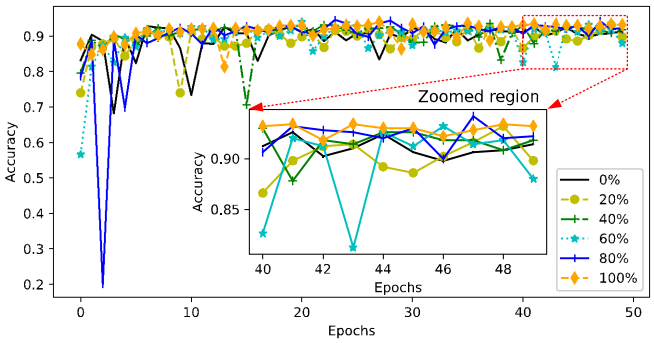

Fig. 16. Ablation experiment.

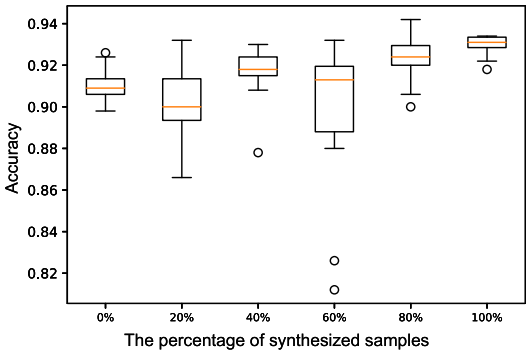

Fig. 17. Box plot for ablation experiment.

by different failure scenarios. According to the results from Figs. 13(a)–(f) and 14(a)–(f), the failures of splitting and fracture are particularly difficult to identify for all models. This is probably because modeling this kind of information effectively is difficult since the data are inherently complicated. Overall, the feature manifold visualization reveals valuable insight into the models' performance and highlights the effectiveness of our proposed model in accurately capturing and interpreting the latent features of the data.

A typical metric for assessing the effectiveness of classification models is the $F1$ score. It incorporates precision and recall, two crucial metrics for evaluating a model's performance. Recall is the percentage of genuine positives among all real positives, whereas precision is the percentage of real positives among all predicted positives. A model that has a high $F1$ score has successfully balanced precision and recall. In the context of the comparison in Table III, an $F1$ score of 0.939 for the proposed model indicates that it has achieved a high level of performance in terms of accurately recognizing positive cases and avoiding false positives. The proposed model has shown superior performance in terms of balancing precision and recall while compared to existing models like ShuffleNet, RepVGG, ResNet50, ResNet101, and DenseNet. This reveals that the proposed model may be more efficient at differentiating these various damages from CFCS.

The receiver operating characteristic (ROC) curve is utilized to evaluate the classifications by examining its shape and the area under curve (AUC), and the results are illustrated in Fig. 15. It is easy to see that the proposed model achieves higher AUC by examining the ROC curve's shape which is shown in Fig. 15. It reveals that the proposed model is more effective at categorizing the true and false category of the features captured from CFCSs.

G. Ablation Experiment

This section describes an experiment conducted to evaluate the proposed model's effectiveness through ablation analysis. The experiment involves gradually weighing the percentage of synthesized samples, and the results are depicted in Fig. 16. The results reveal that the percentage of synthesized samples plays a crucial role in improving the model's classification performance. Specifically, in Fig. 17, the mixed sample with 100% synthetic samples achieves a higher average accuracy of 93.1%. Conversely, in other scenarios, the accuracy is lower, with the synthesized samples containing 20% synthetic samples achieving the lowest accuracy of 89.8%.

V. Conclusion

In this study, a mixed samples-driven methodology based on DDPM are proposed for identifying damage in CFCSs. The proposed methodology allows for the improvement of small sample size robustness by integrating the DDPM module. The mixed samples-driven architecture is constructed by integrating and processing the synthetic sample together with a few real samples collected during real-world testing. According to experimental evidence, the proposed methodology for detecting and diagnosing damage in CFCS achieved an average accuracy of 93.1%, which is greater than 3.3% compared to current methods. In future work, the utilization of multiple sensors for failure forecasting and health condition monitoring may be explored.

References

[1] Y. Hu et al., "Automatic defect detection from X-ray scans for aluminum conductor composite core wire based on classification neutral network," *NDT & E Int.*, vol. 124, Dec. 2021, Art. no. 102549.

[2] M. K. Yücel, M. Legg, V. Kappatos, and T. H. Gan, "An ultrasonic guided wave approach for the inspection of overhead transmission line cables," *Appl. Acoust.*, vol. 122, pp. 23–34, Jul. 2017.

[3] R. J. Hamidi, S. H. Hosseinian, S. H. H. Sadeghi, and Z. Qu, "A novel approach to utilize PLC to detect corroded and eroded segments of power transmission lines," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 746–754, Apr. 2015.

[4] H. Zhang et al., "Optical and mechanical excitation thermography for impact response in basalt-carbon hybrid fiber-reinforced composite laminates," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 514–522, Feb. 2018.

[5] L. Wang, "Electromagnetic induction holography imaging for stroke detection," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 34, no. 2, pp. 294–298, 2017.

[6] L. Cheng and G. Y. Tian, "Surface crack detection for carbon fiber reinforced plastic (CFRP) materials using pulsed eddy current thermography," *IEEE Sensors J.*, vol. 11, no. 12, pp. 3261–3268, Dec. 2011.

[7] A. de Santos Sierra, C. S. Ávila, J. G. Casanova, and G. B. D. Pozo, "A stress-detection system based on physiological signals and fuzzy logic," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4857–4865, Oct. 2011.

[8] Y. Fan, P. Kong, H. Qi, H. Liu, and X. Ji, "A surface acoustic wave response detection method for passive wireless torque sensor," *AIP Adv.*, vol. 8, no. 1, Jan. 2018, Art. no. 015321.

[9] F. Lv, P. Zhang, Z. Tang, Y. Yue, and K. Yang, "A guided wave transducer with sprayed magnetostrictive powder coating for monitoring of aluminum conductor steel-reinforced cables," *Sensors*, vol. 19, no. 7, p. 1550, Mar. 2019.

[10] K. Balasubramaniam et al., "Global and local area inspection methods in damage detection of carbon fiber composite structures," *Measurement*, vol. 187, Jan. 2022, Art. no. 110336.

[11] Y. Gong, J. Luo, H. Shao, and Z. Li, "A transfer learning object detection model for defects detection in X-ray images of spacecraft composite structures," *Compos. Struct.*, vol. 284, Mar. 2022, Art. no. 115136.

[12] Y. Yan et al., "Non-destructive testing of composite fiber materials with hyperspectral imaging—Evaluative studies in the EU H2020 fibreuse project," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022.

[13] K. Liu, Q. Yu, Y. Liu, J. Yang, and Y. Yao, "Convolutional graph thermography for subsurface defect detection in polymer composites," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[14] P. Chen, Y. Li, K. Wang, M. J. Zuo, P. S. Heyns, and S. Baggeröhr, "A threshold self-setting condition monitoring scheme for wind turbine generator bearings based on deep convolutional generative adversarial networks," *Measurement*, vol. 167, Jan. 2021, Art. no. 108234.

[15] P. Chen, Y. Li, K. Wang, and M. J. Zuo, "An automatic speed adaption neural network model for planetary gearbox fault diagnosis," *Measurement*, vol. 171, Feb. 2021, Art. no. 108784.

[16] P. Chen, Y. Li, K. Wang, and M. J. Zuo, "A novel knowledge transfer network with fluctuating operational condition adaptation for bearing fault pattern recognition," *Measurement*, vol. 158, Jul. 2020, Art. no. 107739.

[17] Y. Gao, G. Song, S. Li, F. Zhen, D. Chen, and A. Song, "LineSpyX: A power line inspection robot based on digital radiography," *IEEE Robot. Autom. Lett.*, vol. 5, no. 3, pp. 4759–4765, Jul. 2020.

[18] R. Wei, H. Wei, D. Chen, L. Xie, Z. Wang, and Y. Hu, "Defect detection for aluminium conductor composite core X-ray image with deep convolution network," *J. Phys., Conf. Ser.*, vol. 1633, no. 1, Sep. 2020, Art. no. 012166.

[19] Y. Gao, L. Gao, X. Li, and X. V. Wang, "A multilevel information fusion-based deep learning method for vision-based defect recognition," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 3980–3991, Jul. 2019.

[20] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2256–2265.

[21] I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[22] S. Bond-Taylor, P. Hessey, H. Sasaki, T. P. Breckon, and C. G. Willcocks, "Unleashing transformers: Parallel token prediction with discrete absorbing diffusion for fast high-resolution image generation from vector-quantized codes," in *Computer Vision—ECCV 2022*, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds. Cham, Switzerland: Springer, 2022, pp. 170–188.

[23] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840–6851.

[24] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.

[25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

[26] P. Xia, L. Zhang, and F. Li, "Learning similarity with cosine similarity ensemble," *Inf. Sci.*, vol. 307, pp. 39–52, Jun. 2015.

[27] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6848–6856.

[28] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "RepVGG: Making VGG-style ConvNets great again," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13733–13742.

[29] J. Xu, Y. Pan, X. Pan, S. Hoi, Z. Yi, and Z. Xu, "RegNet: Self-regulated network for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 25, 2022, doi: 10.1109/TNNLS.2022.3158966.

[30] D. O'Neill, B. Xue, and M. Zhang, "Evolutionary neural architecture search for high-dimensional skip-connection structures on DenseNet style networks," *IEEE Trans. Evol. Comput.*, vol. 25, no. 6, pp. 1118–1132, Dec. 2021.

Peng Chen received the M.S. degree in mechatronic engineering and the Ph.D. degree in mechanical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2016 and 2020, respectively, under the supervision of Prof. Ming J. Zuo.

From 2019 to 2020, he was an International Scholar with the Department of Mechanical Engineering, Katholieke Universiteit Leuven (KU Leuven), Leuven, Belgium. He was a Visiting Research Scholar with the Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, South Africa, in 2018. He is currently an Assistant Professor with the College of Engineering, Shantou University, Shantou, Guangdong, China. His research interests comprise signal and acoustics processing, artificial intelligence includes deep learning, vision computing, image or audio super-resolution, intelligent interactions, and prognostics and health management (PHM). Their applications have applicability in bullet trains, wind turbines, industrial robotics, carbon fiber-reinforced composites, and other crucial components of smart manufacturing systems involving intelligent sensing and perception, fault diagnosis, and health condition monitoring.

Chaojun Xu was born in Chongzuo, Guangxi, China, in 1999. He received the B.S. degree in measurement and control technology and instrumentation from Beijing Jiaotong University, Beijing, China, in 2021. He is currently pursuing the M.S. degree in mechanical engineering with the College of Engineering, Shantou University, Shantou, Guangdong, China, under the supervision of Dr. Peng Chen.

His research interests include signal processing, computer vision, and their applications to the detection and prognosis of carbon-fiber composites for wind turbine structures.

Zhigang Ma was born in Huaian, Jiangsu, China, in 1998. He received the B.S. degree in vehicle engineering from Soochow University, Suzhou, China, in 2021. He is currently pursuing the M.S. degree in mechanical engineering with the College of Engineering, Shantou University, Shantou, Guangdong, China, under the supervision of Dr. Peng Chen.

His research interests include vibration signal processing, computer vision, and their applications to intelligent manufacturing systems, such as wind turbines and robotic arm.

Yaqiang Jin received the M.S. degree in mechanical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018, under the supervision of Prof. Ming J. Zuo, and the Ph.D. degree in mechanical engineering from the University of Lyon, Lyon, France, in 2022, under the supervision of Prof. Jerome Antoni.

He is currently a Post-Doctoral Fellow with the Qingdao University of Technology, Qingdao, China. He also works as a Senior Vibration Engineer with Mingserve Technology Company Ltd., Qingdao. His research interests include condition monitoring, signal processing, and fault diagnosis of rotating machinery, mainly on time–frequency analysis and stochastic modeling.