



A Guided-Wave Based Damage Diagnosing Method with Energy Spectrum and Multi-scale Network

Lyu Shuaishuai^(✉), Yang Yu, and Wang Binwen

Aircraft Strength Research Institute of China, Xi'an 710065, China
647817545@qq.com

Abstract. As a promising approach of structural health monitoring (SHM), Guided-wave is being widely used for damage diagnosis of carbon fiber reinforced composite (CFRP) in aeronautics. The traditional methods usually employ threshold for damage detection. However, as Guided-wave is easily influenced by structural and environmental factors, the man-made thresholds may vary significantly from case to case, which deteriorates its universal application. To overcome this weakness, a deep-learning-based method is proposed. Firstly, the sample is produced in the form of grey scale, so that its label is determined objectively rather than by human judgement. Secondly, it employs a multi-scale convolutional neural network with dedicated settings of kernel size and number of convolutional layers, which can effectively extract damage features from multi-dimensional signals while keeping the negative influences of increased network width low. Additionally, Gradient-class activation mapping (GRAD-CAM) technique is used to reveal the sample areas where the deep learning model concerns, which helps to interpret the physical significance of damage features. The experimental results demonstrate that the accuracy of damage identification and localization reaches 92.6%, which is superior to the traditional methods.

Keywords: Composite material · Damage diagnosis · Guided wave · Convolutional neural network · Grey scale

1 Introduction

Carbon fiber reinforced polymer composite (CFRP) have been widely used in the primary structures of aircrafts [1]. However, they are susceptible to the low-velocity impact of blunt objects, such as tools and runway debris. This kind of impact usually degrades as much as 40% of the laminate's compressive strength by inducing large areas of delamination inside with no observable dent from the surface [2]. Apparently, this is of severe safe hazard, and requires reliable approach of structural health monitoring (SHM) [3].

Guided wave is a very potential candidate of SHM, as it has the nature of long-range propagation and sensitivity to typical damages (such as crack, debonding and delamination) of aeronautic thin-walled structures [4, 5]. The mechanism of damage

recognition is by differentiating the signal variations of two guided waves before and after the damage event, and its general method is to quantify this variation as Damage Index (DI) and compare it with the DI threshold [6–8]. However, as the guided wave is so susceptible to geometrical configurations of structure and environmental factors such as temperature and loading conditions that the DI threshold is usually determined by individual experience, and the quality of damage recognition is heavily expertise-dependent and not fitted for general application.

As the intrinsic work of damage diagnosis is to identify and extract damage-related signals from a large amount of overwhelming noise signals, deep learning method is spontaneously adopted by many researchers in the hope of enabling the damage recognition more accurate and less individual-experience-dependent [9–14]. Tian Guo [15], et al., presented a model for crack identification of metallic beam. It used the vibrating waves at diverse modules as input, and employed multi-scale CNN \ ResNet \ Fully Connected Layer to alleviate the side effect induced by noise and data missing. Its accuracy of crack recognition reaches 90% with a good manifest of generalization. However, this good result is obtained partly due to the relative simplicity in this case, and a large number of numerical samples collected through finite element simulation, which, however, is a general obstacle for many cases. For instance, Sbarufatti [16] used ANN model to identify crack and quantify its size in aluminum plates. Merely 38 physical cracks and 410 numerical ones are generated. Therefore, some particular measure such as data augmentation has to be adopted to deal with the problem of inadequate samples. Liang Xu [17] investigated the influential factor of structural dispersion specimen after specimen in crack detection. 6 aluminum hinges with identical structural configuration were loaded to collect guided-wave signals during the period of crack initiation and propagation. The deep learning model consisted of 3 convolutional layer, 1 pooling layer, 2 fully connected layers and 1 output layer.

Although all these learning models presented above claimed that they have a good performance in accuracy, the accuracy is reported to achieve from internal testing only, which means that the models have not been verified by completely new samples, and the over-fitting situation exists somewhat.

Therefore, inadequate samples and poor sample-labelling strategy relying on individual experience, are the two major challenges for the application of deep learning method in SHM.

The study in this paper endeavors to address these two problems. Firstly, it improves the form of sample from a single piece of wave signals to a grey scale by aligning all pieces of signals over the monitoring area side by side. This enables the updated sample to be labelled by objective truth instead of individual judgement. Secondly, a multi-scale convolutional neural network was proposed to extract damage features from multi-dimensions to avoid over-fitting problems caused by insufficient data. All these approaches together serve to improve the accuracy and reliability of the deep learning model, especially when no individual expertise is involved and only limited samples are available.

2 Sample and Model Design

2.1 Sample of Grey Scale

The stiffened panel tested is made of CCF300/BA9916, and its geometric configuration is demonstrated in Fig. 1. A grid of 9 piezoelectricity (PZT) sensors forms a rectangle monitoring network, which is partitioned into 4 regions as A, B, C and D. Region A and C are on the skin, while Region B and D over stiffeners. PZT sensor could work as either guided-wave generator or receiver, and any two sensors apart would generate a wave-propagating channel (namely monitoring path). Therefore, every sensor in the network in turn activates wave signals for other sensors to receive, and, as to say, 6 monitoring paths for each region in Fig. 1 are established.

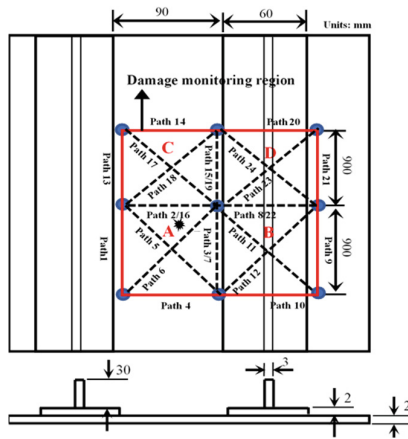


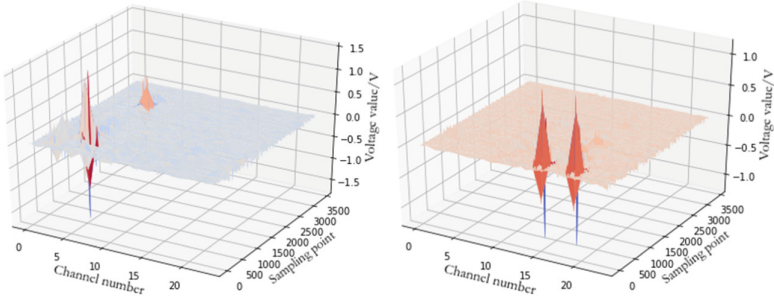
Fig. 1. Structure diagram of Lamb wave monitoring network

The fundamental mechanism of damage recognition is by differentiating the signal variations of two guided waves along the same path before and after the susceptible damage event, and its general method is to quantify this variation as Damage Index (DI) and to compare it with the DI threshold [18, 19]. Since guided wave is so susceptible to many factors, such as the distance from the damage to the path, this plausible method faces an uncompromising embarrassment. For example, if the damage occurs at point P in Region A (seen Fig. 1), the DI value of respective paths in Region A varies significantly. How to determine the DI threshold? It is quite difficult to establish a general rule, and, in most cases, it ends up with individual experiences [20].

By far, this practice is equally used for sample labelling in deep learning model. As there is not a fixed DI threshold, the sample state (damaged or undamaged) also has to be determined by individual experiences. However, for deep learning model, sample labelling is one of the most critical work. The operation like this will have significantly negative impact and largely mislead model training.

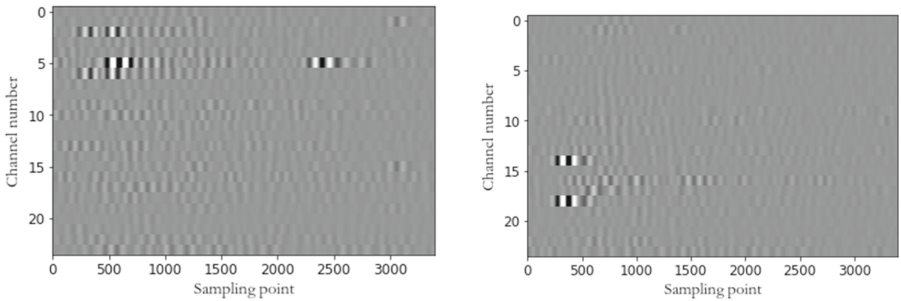
In contrast to a single monitoring path, the integration of all monitoring paths will reverse this situation completely [21]. When 24 pieces of signal variations obtained from

the monitoring network in Fig. 1 (the signals from paths on the region boundaries are repeated) are aligned side by side (seen Fig. 2 in terms of energy distribution of scattered wave), the sample in this new pattern represents the state of the whole area instead of a single path, and consequently the sample could be labelled according to objective truth (damage occurs in this area or not). This practice completely avoids the artificial intervention in sample labelling. Additionally, the new sample presents in a form of image (Fig. 3), which could take full advantage of the latest developments of the deep learning model which is particularly good at image recognition and disposition.



(a) The damage in Region A (b) The damage in Region C

Fig. 2. The energy distribution of the scattered wave in 24 monitored paths



(a) The damage is at position 1 (b) The damage is at position 2

Fig. 3. The corresponding grey scales of Fig. 2

For the convenience of image disposition, the 24 pieces of signal variations are transferred to integrated grey scale map (IGM). Each IGM has 5 types of labels: 0, 1, 2, 3, 4 and 5, representing undamaged, damaged in Region A, B, C and D, respectively.

2.2 Data Augment

The shortage of sample is the leading hazard for the application of deep learning model in SHM. The deep learning model is expected to manifest the capability of reliable damage

recognition while being immune to other disturbing factors such as structural dispersion. This capability relies upon a large amount of samples, which, however, is exactly the weak point in SHM. Expensive experiment cannot be afforded, and the solution is using virtual samples at one hand, while developing the adaptive deep learning model at the other hand. This paper focuses on the latter one and tries to improve it from several aspects.

Some practical problems are encountered when deep learning methods are applied:

- shortage of data samples. Limited by expensive and rare structures [24], the quantity of data sample usually cannot meet the requirement of model training;
- severe class skewness. In real application, the “damaged” signals generated are much less than the “undamaged” ones;
- Non-uniform data length. In practical application, the setting of data length varies from case to case. However, the uniform data length is essential as the input format for Convolutional Neural Network (CNN).

In order to tackle above problems, the module of data augment is introduced to increase the amount of data samples, as well as to alleviate class skewness.

- for the data, of which the length is shorter than the requirement, some meaning-less Fig. 4 are added to extend its length;
- for the data, of which the length is longer, a stride C_i of sliding window is defined in Eq. (1) to cut out segments from original data (schematic illustration in Fig. 4).

$$stride_{C_i} = \left\lceil MS * \frac{|samples\ labeled\ C_i|}{\max_{j=1}^m |samples\ labeled\ C_j|} \right\rceil \tag{1}$$

where, MS is the maximum stride length. It is a super parameter, and its value ranges between 50 and 300 empirically. C is the label set, $C = [C1, C2, \dots, Cm]$, where m is the pattern of label. In this paper, $C = [C1, C2, C3, C4, C5] = [undamaged, Region\ A, B, C, D]$. The symbol of absolute value is to obtain the magnitude of that labelled samples.

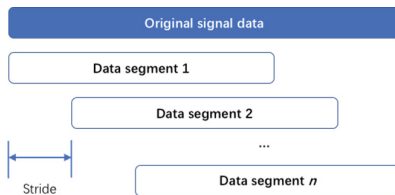


Fig. 4. Schematic diagram of using sliding window step to intercept the original data

2.3 Multi-scale Network Design

2.3.1 The Convolutional Kernel

Unlike general images, the IGM is extremely ill-proportional. Its dimensions are $24 * n$, in which n is the sample rate of the wave-signal, usually 3,000–5,000 points. Therefore, the receptive field is specifically designed, of which the parameters m , n , $f1$ and $f2$ are the length, width, and the stepping interval in length and width direction, respectively (Fig. 5). There are 3 patterns of convolution kernel, $3 * 16$, $6 * 16$, and $3 * 32$ at hand. As in each region there are 6 paths and usually no more than 3 paths of them are mostly affected by a damage, the width is set to be 3 or 6. The length also needs to be set properly. Over-long kernel will lead failure of characteristic extraction, while over-short kernel will undermine the generalization. In this study, 16 and 32 are preliminarily selected.

In order to acquire the optimal result, 3 patterns of kernel are tested separately or in a form of combination. The working mechanism of single and multi-scale kernels are demonstrated in Fig. 5a, b, respectively.

2.4 The Model Configuration

The model constitutes of 3 major modules, data augmentation, semantic feature extraction, and classification. The difference between Fig. 6a, b lies on the number of kernels, and the constitution of the parameter layer is presented in Fig. 6c.

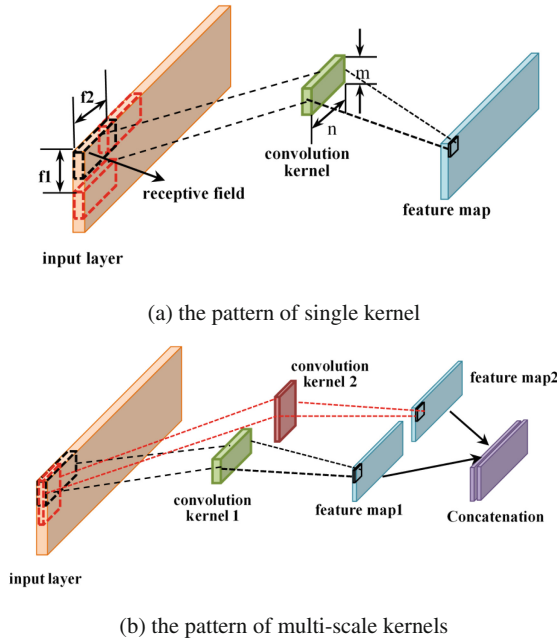


Fig. 5. The pattern of kernel

The semantic feature extraction is comprised of 3 convolutional layers, while for the pattern of multi-scale kernels an additional concatenation is adopted afterwards. This is because the feature map of the multi-scale kernels processed after 3 convolutional layers is significantly larger than that of the single one, and the additional convolutional layer helps to condense the size and reduce the parameter quantity at fully connected layer.

In the parameter layer, Relu is used as activation function in case of gradient explosion and gradient diffusion, while Normalization and “Dropout” for sparsely processing of the training parameters of each layer.

Two fully connected layers constitute of the classification. The first one extracts the top features and still uses Relu as activation function. The latter one finalizes the classification and uses SoftMax as activation function, which is presented as.

$$S_i = \frac{e^{V_i}}{\sum_j e^{V_j}} \quad (2)$$

where S is the output of SoftMax, V is the output vector of the last fully connected layer, V_i is the i th element in V , which has 5 elements in this paper. The position of the greatest element of the five is regarded as the classifying result.

The loss function is presented as

$$L(w) = - \sum_{i=1}^m \sum_{k=1}^K 1\{y^{(i)} = k\} \log S_k(w) + \frac{\lambda}{2} \|w\| \quad (3)$$

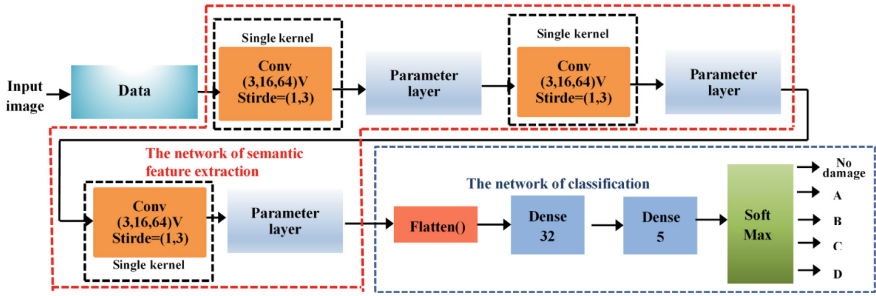
where m is the amount of samples in each batch, k is the sample type, $y^{(i)}$ is the real value of the i th sample, λ is the normalizing coefficient, w is the weighted matrix. $1\{y^{(i)} = k\}$ equals 1 when the sample is positive, otherwise equals 0.

Adam is used to update weight values. It is capable of taking the first and second orders of estimation into account, and consequently stabilizes the optimization process and avoids gradient diffusion.

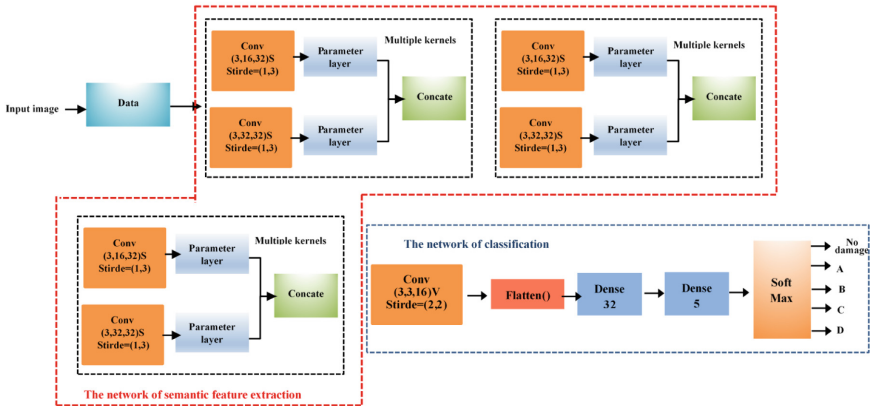
3 Experiment

A group of 6 CFRP stiffened panels (S1-S6) with the same geometrical configuration as shown in Fig. 1 are prepared, in which S1-S4 and S5-S6 are used for model training and testing, respectively.

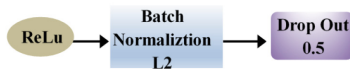
As confined by expensive specimens, false damages in the way of adhering clay is used. The size of clay is retained fluctuates from 8 mm * 8 mm to 15 mm * 15 mm. Each time when the clay is adhered, the specimen is scanned and corresponding guided-wave signal produced is specified as the damaging data. After every ten times of damaging data collection, one baseline data is collected, which is realized by scanning the specimen with no adhesion of clay. Following this way, 2614 damaging data and 270 baseline data are obtained from all 6 specimens. Random pairing of one damaging data and one baseline data collected from the same specimen is conducted, and consequently the signal variation is obtained, which is labelled as A, B, C or D according to its damaging position. Meanwhile, the signal variation produced by any pair of two baseline data from



(a) Model with single-scale kernel



(b) Model with multi-scale kernels



(c) parameter layer

Fig. 6. Model configuration

the same specimen is labelled as undamaged. Following this way, total 5911 samples are produced, in which 5592 are collected from S1–S4 and 319 from S5 to S6 (Table 1). The producing process of experimental data lasts 21 days, during which temperature fluctuated within 3°, and 4 PZT sensors failed and were replaced.

At the process of data augment, the dimensions of sliding window are defined as 24 * 3250, and the stride as 50. The reason of setting the window length equal to 3250 is because almost all the important wave information lies between 150 and 3250. Such operation can secure that each piece of augmented data will always cover this section. Favored by data augment, as many as 22,368 pieces of samples from S1-S4 are produced. Among them 20,000 pieces are randomly selected for training. After every use of 128

Table 1. The number of collected data and generated samples of S1 ~ S6

Test article	Damage signal	Baseline signal	Sample (totality)	Sample (label = 0)	Sample (label = 1)	Sample (label = 2)	Sample (label = 3)	Sample (label = 4)
S1	650	66	1574	400	390	158	496	130
S2	650	66	1574	400	390	158	496	130
S3	595	61	1222	147	410	190	350	125
S4	595	61	1222	147	410	190	350	125
S5	62	8	160	35	29	28	49	19
S6	62	8	159	35	28	28	49	19
Sum	2614	270	5911	1164	1657	752	1790	548

samples for training, the model parameters are updated, and once over 20,000 sample are used, the model is validated by the left 2,368 samples. This process is repeated until convergence.

Once the training is finished, the model is verified by the samples generated from S5 and S6. The result is presented in Fig. 7, and the accuracies for the 5 patterns reach 97%, 96%, 93%, 85% and 92%, respectively. The average accuracy rate is 92.6%. The relatively lower rate occurs at Label 3, although, the root cause is the inadequate samples collected in Label 4, which accounts for merely 9% of the total samples and undermines the model training effect. Nevertheless, the general accuracies are excellent, especially at the case when no individual experience is involved.

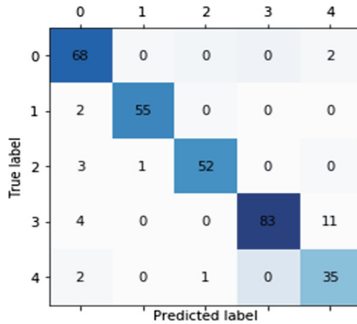


Fig. 7. Mixed matrix of verification

4 Discussion

4.1 The Selection of Kernel

3 patterns of kernel are tested separately or in a form of combination. For the pattern of single-scale kernel, all 3 models converge on the order of 10 iterations. The models

with the kernel of 3×16 and 3×32 have similar performance. Their loss functions stop at 0.10 and 0.093, respectively. With 6×16 kernel the loss function of model stops at 0.15, which does not perform as good as the other two. Therefore, in the pattern of double kernels, only 3×16 and 3×32 are adopted. The converging rate slows down, contrasted to that of the single-scale kernel. Nevertheless, the recognizing rate is observably improved. Its final loss function is 0.062 after around 40 iterations. For the model with 3 kernels, it doesn't converge at all. Therefore, only the pattern of double kernels of 3×16 and 3×32 are used in the following analysis.

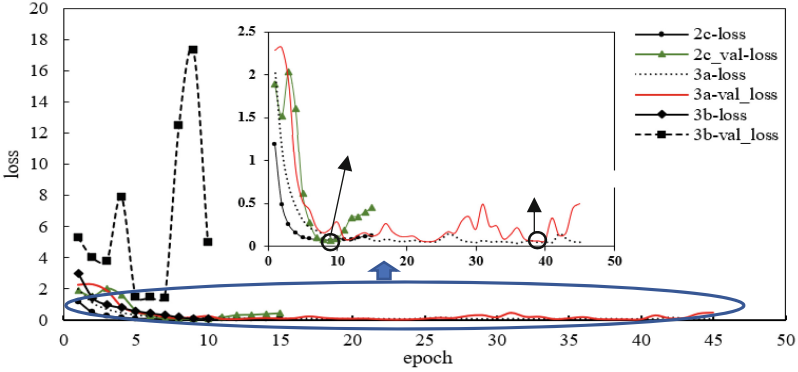


Fig. 8. The loss function curves of the training process of three models (2c represents single-scale kernel 3×16 ; 3a represents multi-scale kernels 3×16 and 3×32 ; 3b represents multi-scale kernels 3×16 , 6×16 and 3×32 ; loss represents the loss function of training set; val_loss represents the loss function of validation set)

4.2 The Generalizing Capability

The nature of generalization, referred in this paper, is confined to the capability of overcoming the structural uncertainty when dealing with a group of specimens with same geometric configurations.

Here the Euclid distance of among extracted features is adopted as the judgement of generalization, of which the lower the value is, the greater similarity the extracted features share, and the better the generalization is [17]. The reason why the way of comparing the final damage-recognizing results between the deep-learning and traditional method is not adopted is that the recognition quality of traditional method is highly expertise-dependent (such as the determination of threshold value of Damage Index). Contrastd to the result comparison, the Euclid distance method is more objective and intrinsic.

For traditional method, the extracted feature matrix consists of 24 monitoring paths, each of which presents 8 patterns of Damage Index, including Cross correlation, Spatial phase difference, Spectrum loss、Central spectrum loss, Differential curve energy, Normalized Correlation Moment, Differential signal energy, Root mean square deviation, respectively. Therefore, its matrix dimensions are 24×8 , and the specific matrix for

specimen S5 and S6 are designated as $[X_5]_{24 \times 8}$ and $[X_6]_{24 \times 8}$, respectively. The Euclid distance (namely, the similarity) is expressed as

$$D_{56}^X = \frac{1}{192} \sqrt{\sum_{i=1}^{24} \sum_{j=1}^8 (x'_{5ij} - x'_{6,ij})^2} \quad (3)$$

$$x'_{k,ij} = x_{k,ij} / \max([X_k]), k = 5, 6 \quad (4)$$

where, $x_{k,ij}$ is the element of $[X_k]_{24 \times 8}$ in the i th row and k th column.

For the deep learning method, the extracted feature is set to be the output of the first fully connected layer (Dense 32 in Fig. 6), which is one dimensional string of 32 elements. Following the naming rule, the specific matrix for specimen S5 and S6 are $[Y_5]_{1 \times 32}$ and $[Y_6]_{1 \times 32}$, respectively. Their Euclid distance is defined as

$$D_{56}^Y = \frac{1}{32} \sqrt{\sum_{i=1}^{32} (y'_{5i} - y'_{6,i})^2} \quad (5)$$

$$y'_{k,i} = y_{5,i} / \max([Y_k]), k = 5, 6 \quad (6)$$

where, $y_{k,i}$ is the i th element of $[Y_k]_{1 \times 32}$. 10 “damages” are produced in either specimen, and each “damage” in S5 has an identical counterpart in S6 in terms of damaging location and size. Consequently, 10 Euclid distances are acquired from either specimen, and their respective average values are $\overline{D_{56}^X} = 0.0108$, and $\overline{D_{56}^Y} = 0.0053$. This proves that the deep learning method may extract features with greater similarity from different individuals, and is more likely to be immune to the geometric uncertainty. Furthermore, this immunity, with more samples, is believed to grow up more.

4.3 The Advantage of Multi-scale Kernels

Figure 9a is an example of the grey scale representing undamaged, which is diagnosed as damaged in Region A and undamaged (the correct diagnosis) by single and multi-scale kernels, respectively. Their corresponding heatmaps are presented in Fig. 9b, c. It is seen that the single-scale kernel concentrates on the greyest zone alone, and consequently submits the false diagnosis. As a contrast, the “multiple” pattern takes all fluctuations over the whole image into account. It is found that the undamaged state generally presents in the form of grey scale with a characteristic of irregular and widely spread fluctuation. Therefore, the “multiple” pattern, which is good at global scanning, is more favored.

4.4 The Denoising Capability

Figure 10 is an example of the grey scale representing the damaged in Region B. Due to poor attachment of Sensor 9, considerable noise is produced on the last monitoring path (seen at bottom right of Fig. 10a). From 10(b) it is seen that the deep learning model successfully differentiate it from other normal damaged signals. This proves the model with the capability of noise suppression to a degree.

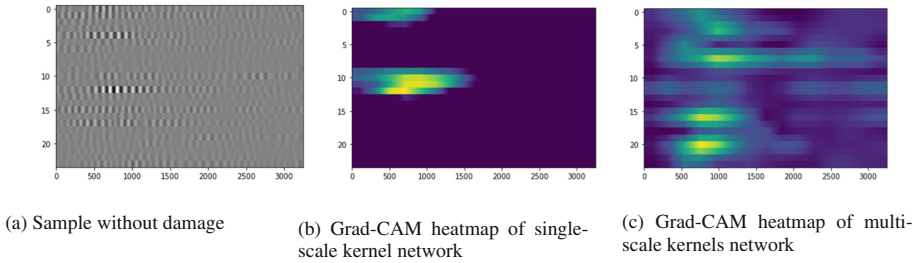


Fig. 9. Comparison of heatmaps between single-scale kernel and multi-scale kernels networks

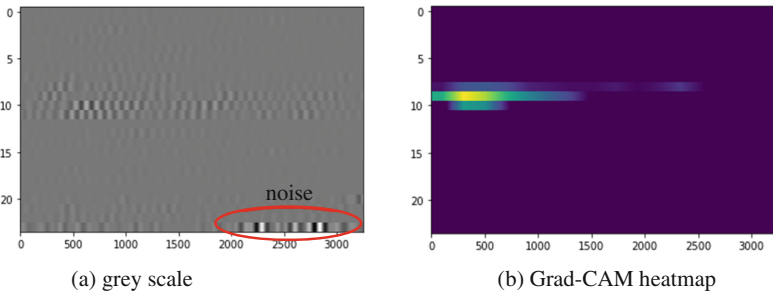


Fig. 10. Sample of damage in Region B with noise generated in Region D

5 Summary

Guided wave is a very potential candidate of structural health monitoring. In this paper, an innovative deep learning model is proposed and it proves to help the damage recognition of carbon fiber reinforced composite using guided wave more accurately and reliably. The major contributions include:

- It uses the integrated grey scale of all guided-wave-propagating paths instead of a single path before as the sample for deep learning model. This practice enables the sample to be labelled according to objective truth rather than individual experience, which may largely mislead the model training;
- The model is elaborately designed with multiple convolutional kernels and the module of data augment, which alleviates the natural shortage of inadequate sample collection in the field of structural health monitoring.

Both measures above help the deep learning model to have a good performance in terms of generalization. Individual experience is eliminated during the labelling process, and the recognizing accuracy reaches 92.6% in the verification.

References

1. Shanyi, D.: Advanced composite materials and aerospace. *Acta Mater Compos Sinica* **24**(1), 12–24 (2007)
2. Craven, R., Iannucci, L., Olsson, R.: Delamination buckling: a finite element study with realistic delamination shapes, multiple delaminations and fibre fracture cracks. *Compos. A Appl. Sci. Manuf.* **41**(5), 684–692 (2010)
3. Xinlin, Q., Yishou, W.: Structural health monitoring technology and its Application in aerospace field. *Exp. Mech.* **27**(5), 517–526 (2012)
4. Purekar, A.S., Pines, D.J.: Damage detection in thin composite laminates using piezoelectric phased sensor arrays and guided Lamb wave interrogation. *J. Intell. Mater. Syst. Struct.* **21**(10), 995–1010 (2010)
5. Lei, Q., Shenfang, Y., Xiaoyue, Z., et al.: A time reversal focusing based impact imaging method and its evaluation on complex composite structures. *Smart Mater. Struct.* **20**(10), 105014 (2011)
6. Ng, C.T., Veidt, M.: A Lamb-wave-based technique for damage detection in composite laminates. *Smart Mater. Struct.* **18**(7), 074006 (2009)
7. Ong, W.H., Chiu, W.K.: Redirection of Lamb waves for structural health monitoring. *Smart Mater. Res.* **2012** (2012)
8. Memmolo, V., Monaco, E., Boffa, N.D., et al.: Guided wave propagation and scattering for structural health monitoring of stiffened composites. *Compos. Struct.* **184**, 568–580 (2018)
9. Sbarufatti, C., Manson, G., Worden, K.: A numerically-enhanced machine learning approach to damage diagnosis using a Lamb wave sensing network. *J. Sound Vib.* **333**(19), 4499–4525
10. Worden, K., Manson, G.: The application of machine learning to structural health monitoring. *Philo-Sophical Trans.* **365**, 515–537 (1851)
11. Atashipour, S.A., Mirdamadi, H.R., Hemasian, M.H., et al.: An effective damage identification approach in thick steel beams based on guided ultrasonic waves for structural health monitoring applications. *J Intell. Mater. Syst. Struct. SACE* (2012)
12. Yan, G.: A Bayesian approach for damage localization in plate-like structure using Lamb waves. *Smart Mater. Struct.* **22**(3), 035012 (2013)
13. Su, Z., Lin, Y.: Lamb wave-based quantitative identification of delamination in CF/EP composite structures using artificial neural algorithm. *Compos Struct* **66**(1–4), 627–637
14. Agarwal, S., Mitra, M.: Lamb wave based auto-matic damage detection using matching pursuit and machine learning. *Smart Mater. Struct.* **23**(8), 085012
15. Guo, T., Wu, L., Wang, C., et al.: Damage detection in a novel deep-learning framework: a robust method for feature extraction. *Struct. Health Monitor. Int. J.* **19**(2), 424–442 (2020)
16. Sbarufatti, C., Manson, G., Worden, K., et al.: A numerically-enhanced machine learning approach to damage diagnosis using a Lamb wave sensing network. *J. Sound Vib.* **333**(19), 4499–4525 (2014)
17. Xu, L., Yuan, S., Chen, J., et al.: Guided wave-convolutional neural network based fatigue crack diagnosis of aircraft structures. *Sensors* **19**(16), 3567 (2019)
18. Yao, P., Zheng, B., Dawood, M., et al.: Real time monitoring of spot-welded joints under service load using lead zirconate titanate (PZT) transducers. *Smart Mater. Struct* **26**(3), 35–59 (2017)
19. Guoqiang, L., Xiasheng, S., Yingchun, X.: T-type reinforcement damage monitoring for composite materials based on Lamb wave and Hilbert transform. *Acta Mater Compos Sinica* **31**(3), 818–823 (2014)

20. Dworakowski, Z., Ambrozinski, L., Packo, P., et al.: Application of artificial neural networks for compounding multiple damage indices in Lamb-wave-based damage detection. *Struct. Control. Health Monit.* **22**(1), 50–61 (2015)
21. Guoqiang, L., Yingchun, X., Hua, Z., et al.: Probabilistic imaging method for damage identification of composite Reinforced panels. *Acta Mater Compos Sinica* **35**(002), 311–319 (2018)