applied sciences

MDPI

Article

Surface Damage Identification of Wind Turbine Blade Based on Improved Lightweight Asymmetric Convolutional Neural Network

Li Zou [1,2,3,*], Haowen Cheng [1,2,3] and Qianhui Sun [1,2,3]

[1] Software Technology Institute, Dalian Jiaotong University, Dalian 116028, China; 13543048304@163.com (H.C.); s14741493670@outlook.com (Q.S.)
[2] Liaoning Key Laboratory of Welding and Reliability of Rail Transportation Equipment, Dalian Jiaotong University, Dalian 116028, China
[3] Dalian Key Laboratory of Welded Structures and Its Intelligent Manufacturing Technology (IMT) of Rail Transportation Equipment, Dalian 116028, China
* Correspondence: lizou@djtu.edu.cn

Abstract: Wind turbine blades are readily damaged by the workplace environment and frequently experience flaws such as surface peeling and cracking. To address the problems of cumbersome operation, high cost, and harsh application conditions with traditional damage identification methods, and to cater to the wide application of mobile terminal devices such as unmanned aerial vehicles, a novel lightweight asymmetric convolution neural network is proposed. The network introduces a lightweight asymmetric convolution module based on the improved asymmetric convolution, which applies depthwise separable convolution and channel shuffle to ensure efficient feature extraction capability while achieving a lightweight design. An enhanced Convolutional Block Attention Module (CBAM) embedded with a spatial attention module with a selective kernel, enhances the acquisition of spatial locations of damage features by combining multi-scale feature information. Experiments are carried out to verify the efficacy and the generalizability of the network proposed for the recognition task. A comparison experiment of common lightweight networks based on transfer learning is also conducted. The experimental results show that the lightweight network proposed in this article has better experimental metrics, including 99.94% accuracy, 99.88% recall, and 99.92% precision.

Keywords: damage identification; convolutional neural network; asymmetric convolution; depthwise separable convolution; attention mechanism

check for updates

Citation: Zou, L.; Cheng, H.; Sun, Q. Surface Damage Identification of Wind Turbine Blade Based on Improved Lightweight Asymmetric Convolutional Neural Network. *Appl. Sci.* **2023**, *13*, 6330. https://doi.org/10.3390/app13106330

Received: 11 April 2023
Revised: 10 May 2023
Accepted: 18 May 2023
Published: 22 May 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Research Background

As a new type of clean energy, wind energy is renewable and pollution-free, which overcomes the shortcomings of traditional fossil energy's environmental pollution and non-renewable status, and conforms to the needs of sustainable development in today's world. According to the GWEC (Global Wind Energy Council, headquartered in Brussels, Belgium.) Global Wind Report 2023 [1], globally, 77.6 GW of new wind power capacity was connected to the grid in 2022, bringing total installed wind power capacity to 906 GW, up 9 percent year-on-year. The global onshore wind market added 68.8 GW last year, with China accounting for 52 percent. In addition, the generating cost of wind energy is reducing gradually, and the competitiveness with traditional fossil energy for power generation is also significantly enhanced. In addition, with the continuous development of science and technology and the tilt of energy policies around the world, wind power generation is developing rapidly, and wind electricity technology will become increasingly mature.

As one of the most significant parts of the wind turbine, the wind turbine blade (WTB) will directly affect the power generation efficiency. The blades of wind turbines

that are in continuous operation are susceptible to deterioration and damage from wind, sand, rain, snow, and seawater. Damage such as surface shedding and cracks may occur, which affect unit performance and bring safety risks. The purpose of WTB surface damage identification is to ensure that it can monitor WTBs and ensure the blades are operating in good condition before a catastrophic event or disaster for the wind turbine occurs. It is also possible to check the health of the blades during the operation of the wind turbine. Inefficient identification and improper maintenance can lead to prolonged downtime, loss of power generation, blade replacement, and high economic losses. Furthermore, analyzed from a cost perspective, the manufacturing cost of three blades is considered to be equivalent to 15–20% of the total manufacturing budget of a wind turbine [2]. The cost of replacing a new blade could be as expensive as $200,000 [3]. Exploring effective WTB surface damage identification techniques therefore has significant engineering worth and practical applications.

1.2. Related Works

Many scholars have been paying constant attention to the problem of WTB surface damage identification, and have proposed many related damage identification methods. At present, the common methods of WTB damage recognition are roughly divided into two categories: the convolutional-neural-network-based damage image recognition approach and the conventional damage recognition method.

Traditional damage identification methods mainly use sensors to detect signals of WTBs for identifying damage, including thermal imaging, vibration analysis, acoustic emission, optical fiber sensing, and other recognition technologies.

The infrared thermography technique is based on the thermal wave theory, which uses the optical radiation on the blade surface to analyze the heat transfer of various parts of the blade and convert it into a visual graph for analyzing the health of the blades. Hwang et al. proposed a continuous-line laser thermography technique [4] and a continuous-wave line laser thermography system [5] for non-destructive WTB monitoring. The system works under rotational conditions by generating thermal waves and using an infrared camera to record the propagation of corresponding waves. In their experiment, the method was able to correctly detect more than 95% of the rotating blade damage. Sanati et al. [6] studied two thermography techniques, including passive and active pulsed and step heating and cooling thermography. When the component is damaged, the vibration analysis detection technology analyzes and determines structural damage by displaying frequencies and modes of physical parameters. Dolin'ski et al. [7] used the Finite Element Method (FEM) and Laser Scanning Vibrometry (LSV) to determine the size and location of stratification in WTBs. Hoell et al. [8] proposed a data-driven vibration-based damage detection method, which uses multivariate Damage Sensitive Features (DSFs) extracted from acceleration responses to conduct damage detection. The experimental results showed that the best damage detection accuracy of 79.2% was obtained using sequential projection pursuit (SPP)-converted DSFs. When WTBs are iced, Zhang et al. [9] used a random forest classifier to detect this by analyzing data from wind farms with combined vibration signals in supervisory control and data acquisition systems. Acoustic emission detection technology observes the emissive sound signal when the blade is working normally. When cracks or other mechanical damage occur, the noise signal becomes unstable and time-varying. The highest accuracy of 97.4% was proved. Tang et al. [10] studied the feasibility of in-service monitoring of the structural health of blades by acoustic emission. Additionally, with the proposed lowering of the threshold from 45 dB to 40 dB, the signal-to-threshold ratio would increase from 5 to 10, which would allow an increase in the detection range from 0.35 m to 3 m, thus reducing the number of sensors required to 8 for basic monitoring and 32 for localization. Beale et al. [11] outlined an adaptive wavelet packet denoising algorithm that applies to numerous SHM technologies and includes acoustics, vibrations, and acoustic emission. Additionally, the success rate of damage detection was improved by 60% after applying the proposed adaptive wavelet packet denoising algorithm.

Xu et al. [12] performed health monitoring of a 59.5 m long composite WTB under fatigue loads using an Acoustic Emission (AE) technique. Fiber-sensing detection mainly uses fiber gratings, which have good photosensitivity. When the external environment changes, the projected wavelength of the grating changes, so the deformation of the blade is obtained by measuring the change in wavelength. Tian et al. [13] proposed a novel non-baseline defect detection methodology based on fiber Bragg grating (FBG) in WTB. Additionally, the authors provide the results of the mean damage detection error (MDDE) comparison between the proposed method and the conventional strain energy method. For single damage detection, the MDDE of the proposed method is 0.12786, and for dual damage detection, the MDDE of the proposed method is 0.024546. These results indicate that the proposed method has excellent accuracy and reliability in detecting wind turbine blade damage. Coscetta et al. [14] developed a distributed fiber-optic strain sensor for blade monitoring. Wen et al. [15] developed an online monitoring system based on FBG sensors along with fiber optic rotary joints (FORJ). The experimental results of the above traditional damage identification methods are good, but there are environmental and cost problems, so the applications are often unsatisfactory in actual engineering. For example, if the damage occurs close to the blade tip and the natural frequency and vibration response change very little, it is difficult for vibration analysis techniques to identify effectively. Other traditional damage identification methods also have disadvantages, such as difficult operation, harsh environmental conditions, a single type of damage identification, the inability to carry out damage identification during normal operation, and more.

In recent years, with the computing power of image processing equipment breaking new highs and the popularity of high-definition UAV, both of them have pushed WTB surface damage recognition technology to a new level, and the damage image recognition method based on a convolutional neural network has gradually become a hot research topic. It has more advantages compared with traditional damage identification methods. For example, it does not require expensive sensing equipment and relieves the cost pressure. It can work in the harsh environmental conditions of the seaside, high altitude, and optical fiber. It automatically distinguishes the difference between different types of damage to achieve a variety of damage categories. It enables damage detection while the wind turbine is running. These advantages have attracted much attention to the damaged image recognition method based on convolutional neural networks in practical deployment. Additionally, it has also attracted more and more scholars to engage in the research of improving this recognition method.

Guo et al. [16] proposed a hierarchical identification framework for WTB, which consists of a Harr-AdaBoost step for region suggestion and a convolutional neural network (CNN) classifier for damage detection and fault diagnosis. Yang et al. [17] proposed an image recognition model based on a deep learning network for the automatic extraction of image features and the accurate and efficient detection of WTB damage. Liu et al. [18] proposed a stiffness prediction method based on CNN and LETM networks considering time series stiffness data under the fatigue test. Yang et al. [19] applied the ResNet50 algorithm to identify the surface damage of WTBs based on UAV machine vision. X. Chen et al. [20] used CNN to predict the icing fault of WTBs. Yuan et al. [21] proposed a Wavelet transform and Full CNN (WaveletFCNN) to automatically obtain multi-scale wavelet characteristics from the time domain and frequency domain. X. Ran et al. [22] proposed an improved Attention and Feature Balanced YOLO (AFB-YOLO) algorithm based on YOLOv5s to enhance the real-time detection of small target defects in WTBs. Y. Yu et al. [23] proposed an image-based fault diagnosis method for WTBs and used CNN for depth feature extractor learning. H. Guo et al. [24] proposed a vision-based approach for blade tip detecting and positioning. More precisely, they detected each structure of the wind turbine by combining the Mask R-CNN detector and shape constraints. Q. Chen et al. [25] proposed an attention-mechanism-based CNN to classify the common damage to WTBs. Our team proposed ED Net [26] in the early stage. ED Net applied improved methods such as Enhanced Asymmetric Convolution Block (EAC Block) and Double Pooling Concatenated Input

SE Block (DPCI_SE Block), which made ED Net achieve good experimental indicators in recognition tasks, where the accuracy rate could reach 99.23% and the recall rate could reach 99.43%. The above method is a CNN image classification based approach for wind turbine blade surface damage detection. In their proposed CNN, the authors achieved good accuracy rates of 92.85% to 99.23% on their respective prepared datasets. However, their proposed CNNs were based on VGGNet, AelxNet, ResNet, DenseNet, and ShuffleNet and focus more on the improvement of classification accuracy with little consideration for further improvement of their CNNs in terms of lightness. Although their improved methods achieve significant accuracy improvements compared to the original models, they also inevitably increase the number of parameters, with some proposed CNNs reaching 20 M or more parameters. The large number of parameters in these CNNs limits their deployment on mobile devices such as UAVs.

Currently, inspection workers can use high-definition UAVs to obtain clear visual data on the surface of WTBs, and then visually identify WTB surface damage themselves, so it incorporates the inspector's subjective factors into the identification results. However, a trend has gradually formed of directly transmitting the collected visual data to mobile terminal devices, such as mobile phones, and using artificial intelligence to identify blade surface damage. The recognition accuracy rate of artificial intelligence is higher than that of human eyes, and can further reduce the burden of human work. In recent years, with the computing capacity of mobile terminals having improved significantly, the storage space can also be increasingly large. This provides base conditions for deploying CNNs in mobile terminals. However, at the same time, it still faces many challenges. For example, the large model size limits the embedding of the CNN in mobile phone software, and abundant parameters mean that mobile phones require too much time to carry out forward computation on the CNN, resulting in poor user experience. However, the lightweight nature of CNNs enables complex network models to migrate to run on devices with limited computing power. The lightweight model technology allows many complex models with high precision to be applied on more devices and expands the application field of deep learning methods. In terms of economic efficiency, the lightweight CNN model requires less computational and storage resources for deployment, thereby reducing the hardware cost and energy cost associated with deployment. Therefore, according to the actual needs, a new lightweight CNN is proposed in this article for the WTB surface damage identification.

1.3. Contributions

Since CNNs are widely used in related tasks in the computer vision field, their internal structure is constantly updated and replaced with the performance requirements of actual missions, so their feature extraction and feature selection capabilities are gradually improved. AC Block [27] originated from the ICCV conference in 2019. This module enhances the robustness of image information through asymmetric convolution and obtains better feature extraction ability. The CBAM [28] was proposed at the ECCV conference in 2018. It combines channel and spatial attention, strengthening feature selection ability from both dimensions at the same time. Therefore, AC Block and CBAM stand out in various image-recognition-related tasks [29–32]. However, both of them have relative defects. AC Block has the problem of many training parameters and a large amount of calculations during training. Additionally, CBAM lacks the multi-scale perception of information when selecting efficient information for the space. Therefore, in this article, considering these two existing problems, a novel Lightweight Asymmetric Convolutional Block combined with an Enhanced CBAM network (LACB_ECBAM Net) for WTB surface damage identification is constructed. The details are as follows.

Firstly, the Lightweight Asymmetric Convolutional Block (LAC Block) is proposed. In particular, we use depthwise separable convolution [33] and channel shuffle [34] to make lightweight improvements to AC Block. It reduces the amount of the parameters in the network training, decreases the computation, and boosts the training speed of the model. Channel shuffle can strengthen the communication of channel information

between groups and compensate for the loss of feature information caused by depthwise separable convolution.

Secondly, the Selective Kernel Spatial Attention Module (SK_SAM) is proposed. Specifically, we use the convolution of the selective kernel [35] to improve the Spatial Attention Module (SAM) [28]. The convolution with an adaptive receptive field will extract features of different sizes more effectively. It enhances the capability of the spatial attention model to capture the position for useful information.

Thirdly, the Enhanced CBAM (E_CBAM) is proposed. It is an improvement of CBAM. Its interior consists of DPCI_SE Block [26] and SK_SAM. By analyzing the combined channel and spatial dimension, it can better focus on the significant features of the image and suppress the unnecessary regional response.

2. Models and Principles

2.1. The Basic Structure of LACB_ECBAM Net

LACB_ECBAM Net is a novel CNN using lightweight asymmetric convolution and enhanced CBAM. Its overall structure is shown in Figure 1.

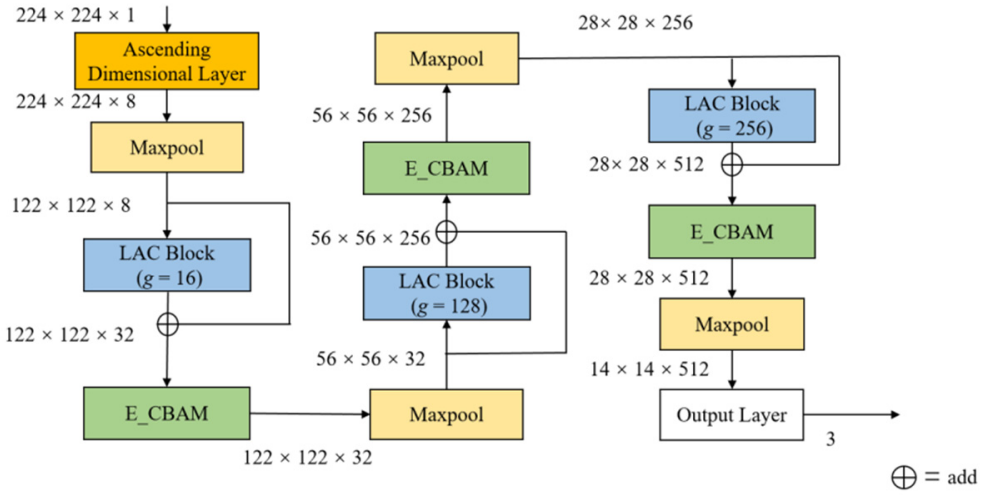


Figure 1. Schematic diagram of backbone-enhanced convolution kernel.

The overall network structure consists of five sub-structures, which are marked by five colors in Figure 1, including the ascending dimensional layer, the max pooling layer, the LAC Block, the E_CBAM, and the output layer. The ascending dimensional layer is composed of a normal convolutional layer with a kernel size of 3, a stride of 1, and a padding of 1, in addition to a batch normalization layer. The max pooling layer has a size of 3, a stride of 2, and padding of 1. The LAC Block based on the AC Block consists of g groups of depthwise separable convolution and channel shuffle, where g takes values of 16, 128, or 256. The E_CBAM is an improved version composed of DPCI_SE Block and SK_SAM instead of CAM as well as SAM. The output layer is composed of a normal convolutional layer with a kernel size of 1 and a stride of 1, in addition to the global average pooling layer, which is designed to replace the large parameter fully connected layers.

2.2. LAC Block

2.2.1. AC Block

In the use of CNNs, after the image is flipped or rotated, the image features extracted by standard square convolution will change, and the segmentation results of the model for the same object may have deviations, resulting in the reduction of the genericization of the model. In the CNN, the features of the image extracted by standard square convolution will change after the image is flipped or rotated, and the segmentation result of the model for the same object may be biased, resulting in a decrease in the generalization ability of the model. Asymmetric convolution is different from standard square convolution. Its

structure is asymmetric and has the characteristics of invariant scale and transformed features. Asymmetric convolution has better robustness when inputting rotational distortion. Asymmetric convolution uses the relationship between weight and weight space position in the CNN convolution kernel to strengthen the ordinary convolution layer and adds horizontal and vertical convolution operations to improve the robustness of the model to flip and rotate images. As for the adding position of asymmetric convolution, authors in the literature [27] demonstrated through pruning experiments that the importance of weights at different spatial locations in the convolution kernel is different, and removing weights from the backbone results in more accuracy reduction than the weights from the corners. A couple of asymmetric kernels can then be inserted at the locations of the matching square kernel at the skeleton to construct a skeleton-enhanced kernel, as shown in Figure 2.

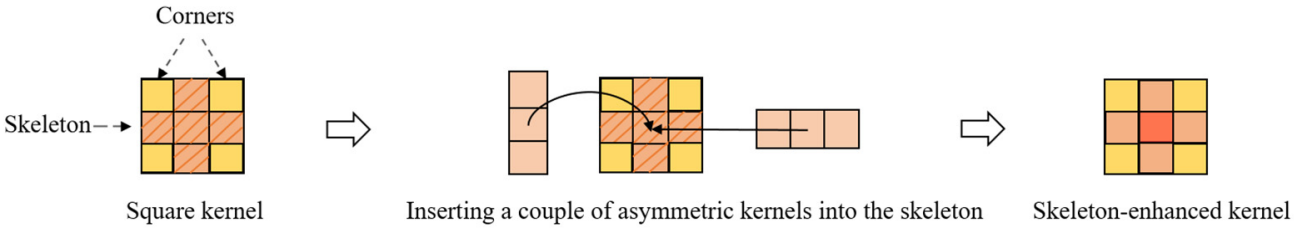


Figure 2. Schematic diagram of backbone-enhanced convolution kernel.

The construction of AC Block [27] is performed by adding a pair of asymmetric convolutions based on the convolution of the square kernel, forming three convolution branches in parallel, batch normalizing the feature information extracted by the three convolution branches, and then adding up and summarizing. This can highlight the capital feature information at the kernel skeleton during the convolution operation and improve the robustness of the model. The asymmetric convolution module adds additional asymmetric convolution layers to learn the input features during the training process and fuses the asymmetric convolution parameters with the standard convolution parameters after training. Therefore, it can enhance the representation ability of the post-training convolution kernel without introducing additional parameters and computation.

2.2.2. Depthwise Separable Convolution

Depthwise separable convolution is a lightweight enhancement to traditional convolutional neural networks. Using this structure can reduce the number of parameters during model training, reduce the amount of computation, and boost the model-trained speed. Feature acquisition and feature composition for the ordinary convolutional layer are performed in one step, while depthwise separable convolution decomposes this into two steps. First, the depthwise convolution [33] uses the convolution kernel with input channel 1 to extract features. Additionally, then, the pointwise convolution [33] uses the convolution kernel of size 1 to combine features. This process is shown in Figure 3.

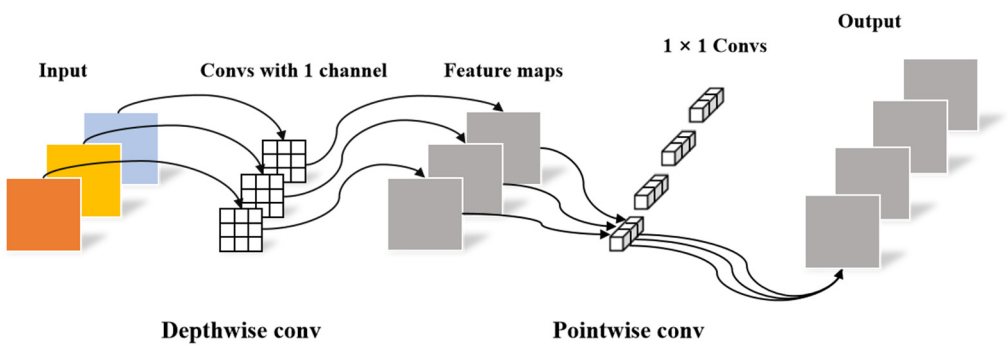


Figure 3. Diagram of the depthwise separable convolution process.

Supposing the feature map has a length and width equal to D_{in}, its size is $D_{in} \times D_{in} \times M$, where M refers to the number of channels in the input feature map. Similarly, it is assumed that the feature map has a height and width equal to D_{out}. However, the output feature map's size is $D_{out} \times D_{out} \times N$, where N refers to the number of channels of the output feature map. If the convolution kernel size is $D_K \times D_K$, after passing the standard convolution, the parameter quantity of the standard convolution layer is $P_{standard}$; see Formula (1).

$$P_{standard} = D_K \times D_K \times M \times N \tag{1}$$

The parameter quantity of the depthwise convolution is $P_{depthwise}$; see Formula (2).

$$P_{depthwise} = D_K \times D_K \times 1 \times M \tag{2}$$

The parameter quantity of the pointwise convolution is $P_{pointwise}$; see Formula (3).

$$P_{pointwise} = 1 \times 1 \times M \times N \tag{3}$$

The parameter quantity of the depthwise separable convolution and the standard convolution part is compared as in Formula (4).

$$\frac{P_{standard}}{P_{depthwise} + P_{pointwise}} = \frac{D_K \times D_K \times M \times N}{D_K \times D_K \times 1 \times M + 1 \times 1 \times M \times N} = \frac{1}{N} + \frac{1}{{D_K}^2} \tag{4}$$

Previous experiments [33] demonstrated that the amount of parameters is 8- to 9-fold fewer than the amount of standard convolution when using a 3×3 depthwise separable convolution, and the accuracy is only slightly less.

2.2.3. Channel Shuffle

Group convolution is applied in depthwise separable convolution. However, the group convolution refers to convolution within each group, so there is no information on interaction between groups. For information to sufficiently interact between groups, the author of ShuffleNet [34] proposed to group the results after each group convolution, then exchange the subgroups within their respective groups. This process allows the information between different information groups to blend and enrich the feature information between groups, as shown in Figure 4. The premise of the channel shuffle is to divide the input image into groups by channel. However, the depthwise convolution in the depthwise separable convolution can be regarded as a group convolution with the number of groups equal to that of input channels. Thus, it is possible to combine depthwise convolution with channel shuffle. While reducing the amount of computation and parameters of CNN, the problems that there is no interaction between each channel's information, leading to less learning content being solved. Therefore, channel shuffle enables ShuffleNet to achieve both lightweightness and performance.

2.2.4. The Structural Principle of the LAC Block

Many scholars have applied and improved the AC block since its discovery. High applicability is one of the highlights of the AC block, and it can be used to replace almost any convolution in a CNN. It makes the replaced CNN reach a new height in feature extraction capability.

However, the three convolutional branches in the AC block are all ordinary convolutions, so depthwise separable convolution can be used to construct an LAC Block with lightweight improvement. Figure 5 depicts the architecture of the AC Block and LAC Block. As the network's depth expands, we switch from the original three branches' conventional convolution to a group convolution, where the number of groups g and g takes 16, 128, and 256, respectively. This corresponds to the depthwise convolution in the depthwise separable convolution. Then, we add a 1×1 convolution before each group convolution, which corresponds to the pointwise convolution in the depthwise separable convolution.

It is mentioned in the literature [33] that different before-and-after order of depthwise convolution and pointwise convolution has an equivalent influence on the effect, so our design is reasonable. After the depthwise separable improvement of the three convolutional branches, we designed a channel shuffle for each convolutional branch. Then, each of them performs group convolution again. The channel shuffle reorders the current feature information so that the feature information of each group in the next group convolution comes from the different groups in the last group convolution. In this way, the information from different information groups can be blended and enrich the feature information within each group. Additionally, we included the GELU [36] activation function in place of the original ReLU activation function. The very effective activation function known as GELU has been applied in the Transformer model, which has lately gained popularity. It uses the idea of random regular, which is intuitively more in line with the understanding of nature. Additionally, it can assist in improving the model's degree of nonlinearity.

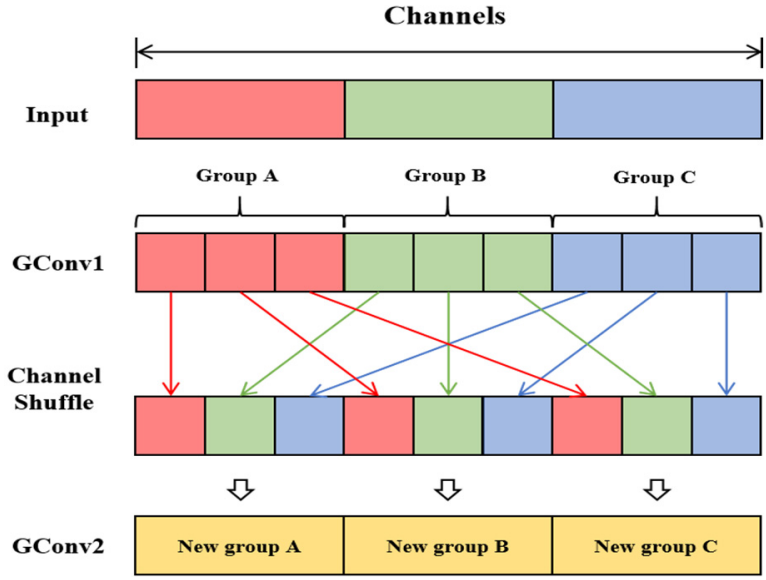
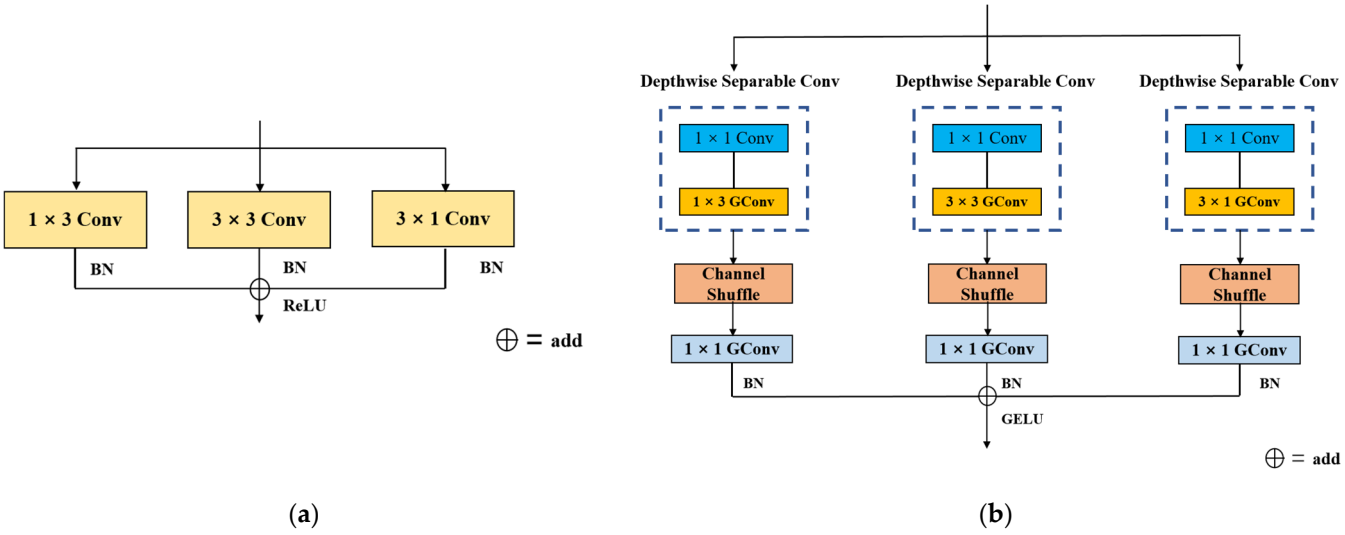


Figure 4. Schematic diagram of the channel shuffle process.



(**a**) (**b**)

Figure 5. Structure diagram of AC Block and LAC Block. (**a**) AC Block. (**b**) LAC Block.

2.3. SK_SAM

2.3.1. SAM

The attention mechanism can make CNN notices adaptively more crucial information, so the attention mechanism is a vital way to achieve adaptive attention in the network. In general, the attention mechanism can be divided into channel attention mechanism, spatial attention mechanism, and a combination of the two. channel attention may be relatively abstract for human vision, but spatial attention can be visualized in the distributed learning weights of attention. The main feature map obtained by convolution corresponds image, which should be the focus of the network. The CBAM is not only combined with the channel attention mechanism but also combined with the spatial attention mechanism. Compared with the previous simple network which only pays attention to channels, it has been significantly improved.

The Channel Attention Module (CAM) of CBAM is mainly concerned with which of the many feature maps contains more valid information. However, SAM focuses on which part of a feature map information is more significant, as a supplement to the CAM. First, the average pooling layer and the max pooling layer are applied along the channel dimension of each feature point of the input feature map and contact together the output feature maps of the two. Next, the feature map of SAM is obtained after passing through the convolution layer with the convolution kernel size of 7×7. Then, the Sigmoid function is used to activate, and a spatial attention diagram is obtained. The weight of each feature point in the input feature map is distributed at the same time. Finally, Weighting the learning weights to the original input feature map by multiplication. The derivation of the relevant formula is given in the literature [28].

2.3.2. The Convolution of Selective Kernel

SK Net [32] relies on the convolution of the selective kernel, which can use different receptive fields for convolution kernels used for different inputs. Finally, different parameter weights are obtained and the output is processed adaptively. The convolution of the selective convolution kernel consists of three parts: Split, Fuse, and Select. Their functions are to generate different size branches of convolution kernel, aggregate different branch information to obtain a global representation of selection weights, and combine feature maps from kernels of various sizes based on selection weights.

The overall process is shown in Figure 6. In the Split phase, two groups of dilated convolution with different convolution kernel sizes are used to extract features from the input which, respectively, represent the receptive fields of different sizes. In the Fuse phase, the information for the two branches, U_1 and U_2, is added to integrate by using the feature and then using the global max pooling layer to embed the global information S. After that, adding a fully connected layer to reduce the dimension, a fully connected layer calculation is performed first, followed by batch standardization, and the ReLU activation function is calculated to obtain the vector Z. In the Select phase, the SoftMax operation for vector Z is performed to obtain vectors a and b. Then, the vectors a and b are multiplied with the original features U_1 and U_2, and the resulting features added to obtain V. Because the sum of the corresponding values in vectors a and b is 1, the weight is set for each branch feature map to strengthen the necessary feature information and suppress the non-necessary feature information. The derivation of the relevant formula is given in the literature [35].

2.3.3. The Construction Principle of the SK_SAM

After many experiments, the authors of [28] concluded that the experimental results of convolution kernel size of 7×7 in SAM were the most ideal. However, the size of the receptive field has a great impact on feature extraction. If the current receptive field is too large, it will cause the detected object to be ignored, making the object become the background, and the object feature cannot be extracted. The accuracy of recognition will be impacted if the receptive field is too limited, since too much local information will be

gathered at the expense of global information. Therefore, better feature extraction ability will be obtained if the receptive field during network training can be adjusted adaptively. Generally speaking, the size of the receptive field is more directly reflected in the size of the convolution kernels. Therefore, if we want to extract features more accurately, the convolution of the selective convolution kernel may obtain better experimental results.

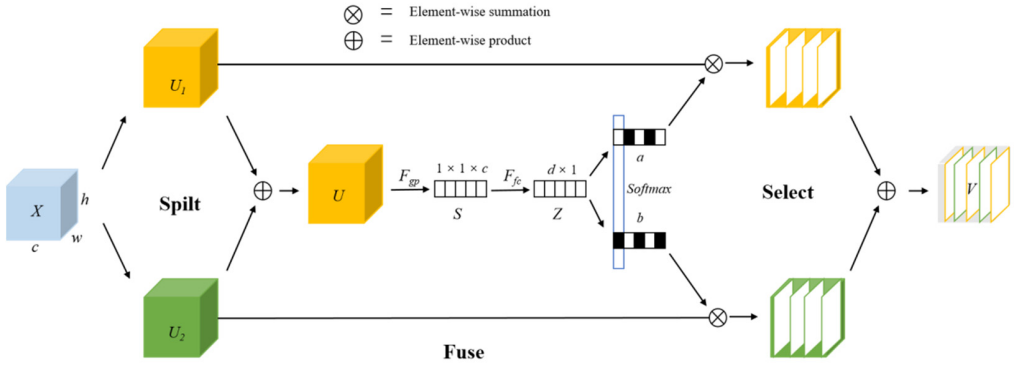
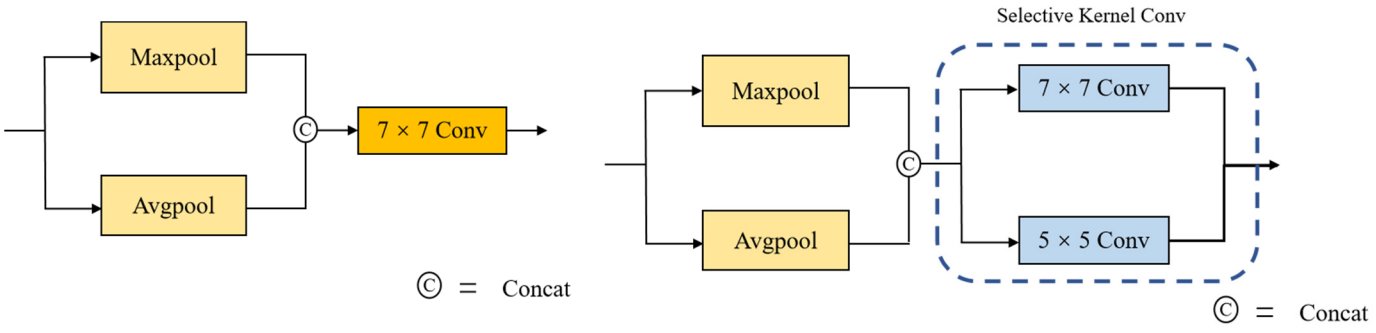


Figure 6. Diagram of the overall Selective Kernel process.

In this article, we replace the 7 × 7 convolution in the SAM with the convolution of the selective kernel to form SK_SAM. The structures of SAM and SK_SAM are shown in Figure 7: convolution of the selective kernel of two convolutional branches, one of which has three sequential convolutions of size 3, and the combined receptive field is equal to the convolution of size 7. In contrast, the other convolution branch has two sequential convolutions of size 3, and their combined receptive field is equal to the convolution of size 5. The receptive field scales of the two convolutional branches are different. By weighted processing of the feature information of channels with different scales, the network pays more attention to the feature information of WTB surface damage. This makes the network carry out adaptive selection and fusion for features with different scales, which is conducive to establishing a more efficient and accurate model.



(**a**) (**b**)

Figure 7. Structure diagram of SAM and SK_SAM. (**a**) SAM. (**b**) SK_SAM.

2.4. E_CBAM

The core of CBAM is that it applies the combination of CAM and SAM to process the input feature map. The SAM can make the neural network pay more attention to the pixel regions that play a decisive role in classification while ignoring the irrelevant regions. Additionally, the CAM takes into account how the feature map's channels relate to one another. Using the CBAM structural principle, we propose the Enhanced Convolutional Block Attention Module (E_CBAM) in this article. In E_CBAM, the DPCI_SE Block is selected as the new CAM, and SK_SAM proposed in this article is selected as the new SAM. The structure comparison of CBAM and E_CBAM is shown in Figure 8.

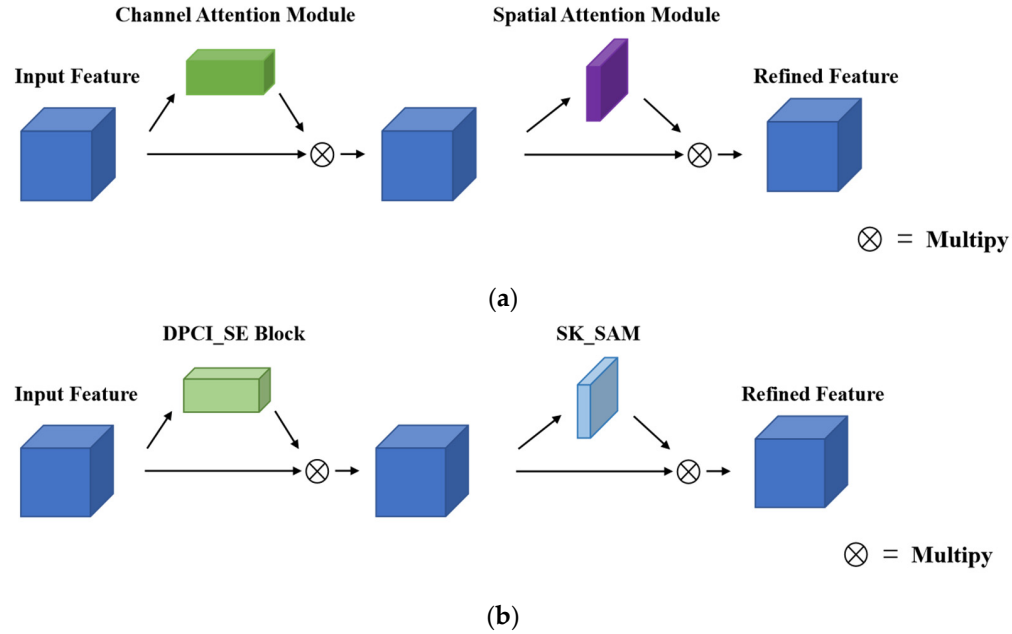


Figure 8. Structure diagram of CBAM and E_CBAM. (**a**) CBAM. (**b**) E_CBAM.

A dual pooling layer [26] was implemented before the SE Block, and DPCI_SE Block is an enhancement built on the SE Block. A parallel structure made up of an average pooling layer with a max pooling layer is known as a dual pooling layer. Likewise, SE Block should use the GELU activation function in place of ReLU activation. The average pooling layer can retain the global feature information and the max pooling layer can retain the local feature information. The two parts of the feature information are finally summarized through the concatenation operation to obtain richer feature information. Meanwhile, the GELU activation function brings better nonlinear expression ability to SE Block. Experimental results show that DPCI_SE Block is a simple and efficient CAM for WTB damage identification, so it is applied to the E_CBAM proposed in this article.

In E_CBAM, feature weights are considered in two separate stages. The first stage is the DPCI_SE block, which improves the network performance and interpretation by learning the importance of each channel in the features using global average pooling and fully connected layers after obtaining an adequate feature map through a channel-based attention mechanism, and enhancing the important features in the feature map and reducing the interference of non-important features. The second stage is followed by SK_SAM, which is based on the application of the convolution of selective kernel through a spatial-based attention mechanism that can adaptively learn the weights of the feature map through multi-scale convolution kernels and attention mechanisms to effectively capture features at different scales and focus on important feature regions to improve the differentiation of feature spatial locations and classification performance. The related derivation of the formulas are detailed in [35].

3. Experiments and Analysis

3.1. Dataset Construction

In this article, a JPG image dataset of WTB surface damage is constructed. The images of the dataset were obtained from the WTB surface damage image collected by the team using a high-definition UAV at a wind farm in Liaoning, China. After obtaining the original images collected by the high-definition UAV, the collected original images were cropped according to a uniform image size to meet the requirements of image classification. Then, image enhancement was performed on the unclear images. Finally, the training and test sets were split and built after data augmentation using a relatively modest number of damage types.

The raw image data consist of 1074 images, which contained both normal WTB surface images and WTB surface damage images. Figure 9 illustrates the deterioration in the form of cracks as well as surface shedding.



(**a**) (**b**) (c)

Figure 9. Three kinds of blade images. (**a**) Normal. (**b**) Cracks. (**c**) Surface shedding.

The size of the original images is 5472 × 3684 pixels. However, the input was too large for the training and testing of this network, which causes an excessive computational burden, and contains excessive and irrelevant background. However, if the image size is too small, the difference between the image is inconspicuous after subsequent data augmentation. This results in the decline of the training effect and affects the final experimental index of the network. Therefore, for the determination of image size, we referred to numerous sample set sizes commonly used for CNNs, as shown in Table 1.

Table 1. Common sample size.

CNN	Sample Size/(w,h)
AlexNet [37]	224 × 224
VGG-16 [38]	224 × 224
VGG-19 [38]	224 × 224
GoogLeNet [39]	229 × 229
MobileNet [33]	224 × 224
MobileNet_V2 [40]	224 × 224
ShuffleNet	224 × 224
ShuffleNet_V2 [41]	224 × 224
Xception [42]	299 × 299

Considering the random region clipping of the images during the subsequent data augmentation, the original image was scaled down to 336 × 336 pixels. Then, image enhancement was performed for images that were not sharp enough due to interference from objective factors such as shooting angle and lighting conditions. A total of two image enhancement methods were applied, including histogram equalization and median filtering. Specifically, histogram equalization enhances the contrast between the damage location and its surroundings, while median filtering eliminates the noise in the image. Next, data augmentation was performed, mainly using OpenCV as an image processing tool. First, the image was randomly cropped with a crop frame of 224 × 224 pixels in size, then clockwise, counterclockwise rotation, and mirror flip were carried out. Finally, random salt-and-pepper noise was added.

After data augmentation, the number of images reached 4729, among which the number of WTB for normal, cracks, and surface shedding images were 1738, 1449, and 1542, respectively. Finally, the above three kinds of images were divided into the training set and the testing set according to the quantity ratio of 7:3, and the construction of the WTB surface damage dataset was finally completed.

3.2. Experimental Scheme

In this article, three experiments were designed, all based on the LACB_ECBAM Net surface damage image dataset, but with different purposes. The first experiment was a training and testing of LACB_ECBAM Net, aiming to validate its effectiveness for

LACB_ECBAM Net surface damage identification. The second experiment performed a comparison test by applying LAC Block and E_CBAM to AlexNet to verify the generalizability of the improved method. The third experiment compared various common lightweight networks based on transfer learning to measure the performance difference between LACB_ECBAM Net and other lightweight networks.

For the experimental environment, the operating system was Windows 10 Professional Edition, Python 3.6.5, and the deep learning framework was Pytorch 1.7.1, developed by Facebook Ai Research (Silicon Valley, CA, USA, https://pytorch.org/, accessed on 10 September 2022). The CPU was an 11th Gen Intel Core i7-11800 H 2.30 GHz, 16 GB of RAM, and the GPU was an NVIDIA GeForce RTX 3060 Laptop GPU-6 GB. We used CUDA 11.2 and CUDNN 8.2.4, developed by NVIDIA (Santa Clara, CA, USA, https://developer.nvidia.com/cuda-zone, accessed on 10 April 2013), as training acceleration kits for the network model. In terms of experimental assessment, the effectiveness of the WTB damage identification approach was assessed using four evaluation metrics often employed in the field of classification, namely accuracy, recall, precision, and F1 score.

3.3. The Experiment of Training and Testing LACB_ECBAM Net

When the LACB_ECBAM Net model is trained, different batch sizes and initial learning rates can lead to changes in the output results of the model. To discover how the training batch size setting and the initial learning rate setting affect the training effect of LACB_ECBAM Net, this experiment compared the training accuracy of the network under different settings using control variables to select the most suitable training batch size and initial learning rate. In addition, the experiment used cross-entropy as the loss function and was optimized using the better-performing Adam gradient descent optimizer with the default decay rate.

The initial learning rates were set as 0.0005, 0.001, and 0.0015, respectively. After four epochs, the learning rate fell by 0.65 times the current rate as they all followed the learning rate reduction method with a fixed stepping length. A total of 20 epochs were trained, and each epoch was then fed to the model as a test set to test the training effect of the current epoch and recorded. Figure 10 shows the influence of different initial learning rates on the accuracy of network training under the condition that the batch sizes are set to 16, 32, and 64.

According to the analysis of the variation trend and convergence of LACB_ECBAM Net model accuracy in the figure, the initial learning rate suitable for LACB_ECBAM Net model training is 0.001 and the batch size is 32. On the same condition, an overhigh initial learning rate will accelerate the convergence rate of accuracy in the early stage, but it will easily fall into the plateau of repeated jumping in the later stage and affect the final accuracy. A too-low initial learning rate will lead to the overall convergence rate being slow and not meeting the experimental requirements. Similarly, setting the batch size too large or too small can also negatively affect the training effect of the model. At the same time, it can be found that if appropriate training parameters are set, the accuracy of LACB_ECBAM Net will eventually converge to 99.957%. This shows that LACB_ECBAM Net has an effective damage identification ability for WTB, which can meet the needs of practical engineering.

3.4. The Experiment of Improved Method Applied to AlexNet

This article needs to further verify the generalizability of LAC Block and E_CBAM. To achieve the effect of control variables, we chose AlexNet as the basic model and used LAC Block, AC Block, E_CBAM, and CBAM for the structural transformation of AlexNet. We added the WTB surface damage classification layer at the last layer; see Figure A1. The training mode with random model initialization parameters was adopted. Each model was trained 10 times, with 15 epochs per training, and tested after every epoch. The final and average accuracy, recall, precision, and F1 score were recorded. The results of six experiments are shown in Table 2.

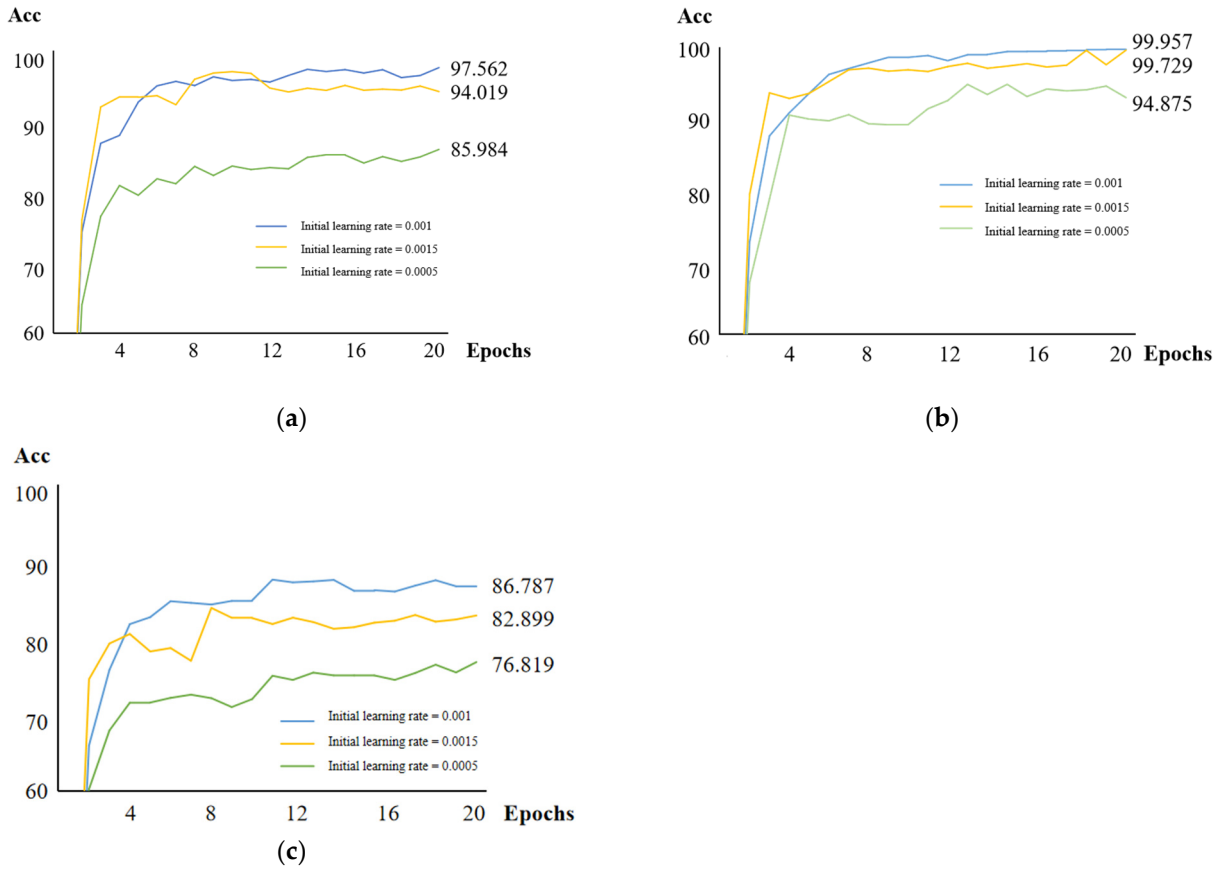


Figure 10. LACB_ECBAM Net's accuracy growth at different batch sizes: (**a**) 16, (**b**) 32, and (**c**) 64.

Table 2. Comparison of experimental results based on AlexNet.

Experimental Items	CNN	Accuracy/%	Recall/%	Precision/%	F1 Score	Params/M
a	AlexNet	85.36	84.49	86.41	85.44	58.2936
b	AlexNet + AC Block	86.29	84.55	86.69	85.61	63.8156
c	AlexNet + LAC Block	86.72	85.26	87.05	86.15	55.9607
d	AlexNet + CBAM	89.41	87.83	88.27	88.05	58.3018
e	AlexNet + E_CBAM	91.68	90.49	92.53	91.50	58.3020
f	AlexNet + LAC Block + E_CBAM	94.46	93.15	94.63	93.88	55.9690

As shown in Table 2, both LAC Block and E_CBAM have been effectively improved. According to the analysis of experiment a and experiment f, AlexNet improved by LAC Block and E_CBAM not only improves the accuracy as well as other evaluation metrics significantly but also reduces the number of parameters. According to the analysis of experiment b and experiment c, LAC Block's accuracy rate increased by 0.43%, recall rate increased by 0.71%, accuracy increased by 0.36%, and F1 score increased by 0.54% compared with AC Block. It is worth noting that the AC Block replacement brings more parameters to AlexNet, but the LAC Block replacement reduces the original number of parameters. LAC Block is lightweight. The quantity of network training parameters is decreased by the use of depthwise separable convolution, which inevitably has a negative impact on the accuracy of the model and other performance indicators. However, LAC Block's combination of channel shuffle and depthwise separable convolution enhances the exchange of information between different groups of channels, making up for the loss. This means that the accuracy and other evaluation metrics do not decrease, the number of parameters can be reduced, and it is a lightweight network. From the analysis of experiment

d and experiment e, E_CBAM is 2.27% more accurate than CBAM, with 2.66% more recall, 4.26% more precision, and 3.45% better F1 score. This indicates that E_CBAM is more targeted than CBAM in the task of fan blade surface damage identification. The difference between the cracks and surface shedding is small, but the depth, shape, and length of intraclass features vary greatly. In the E_CBAM, on the one hand, DPCI_SE Block enables the selection of channel information to obtain more abundant local and global information and ensure more sufficient feature dimensions, so that texture features such as cracks can be perceived more effectively. On the other hand, the application of the selective kernel in SK_SAM forms a multi-scale adaptive receptor field, allowing an appropriate-sized convolution kernel to extract features of that size, which enables complex features to be captured more efficiently in space.

In conclusion, the combination of DPCI_SE Block and SK_SAM strengthens the selection of effective features and the suppression of ineffective or weak features for the surface damage characteristics of WTB. Even though there is a modest rise in the total amount of parameters, performance improves considerably as a result.

3.5. The Experiment of Common Lightweight Networks Based on Transfer Learning

Transfer learning is the successful application of image classification knowledge learned in large data sets to new target classifications to be solved. The model relies on the current optimal network model after transfer training, and the model in this article is constructed by fine-tuning the network layer structure, which is faster and easier than building and training a new randomly initialized network.

To measure the overall performance of LACB_ECBAM Net, this article compares the common pre-trained lightweight models in the field of image recognition with LACB_ECBAM Net, including ResNeXt50 (32 × 4d) [43], Xception, MobileNet_V2, ShuffleNet_V2, and EfficientNet-B0 [44]; their structures are shown in Figure A2. The pre-trained network weight parameters from the ImageNet data set were added to the aforementioned models to ensure that the experiment was fair, and the last layer was added to the WTB surface damage classification layer. Each model was trained 15 times, with each training being 20 epochs, and was tested after every epoch. The final and average accuracy, recall, precision, F1 score, and training time were recorded each time, and their average values were taken as the basis for experimental analysis in Table 3.

Table 3. Contrast experiment of common lightweight networks.

CNN	Accuracy/%	Recall/%	Precision/%	F1 Score	Params/M	Training Time
ResNeXt50 (32 × 4d)	98.63	98.19	97.42	97.80	24.42	21 min 52 s
Xception	99.86	99.83	99.96	99.89	22.77	19 min 17 s
MobileNet_V2	98.33	98.51	98.72	98.61	3.40	16 min 9 s
ShuffleNet_V2	99.48	98.86	99.36	99.11	2.3	15 min 17 s
EfficientNet-B0	98.53	97.78	98.47	98.12	4.99	15 min 45 s
LACB_ECBAM Net (ours)	99.94	99.88	99.92	99.90	0.58	15 min 29 s

Table 3 shows that the accuracy of the LACB_ECBAM Net proposed in this article is 99.94%, placing it top and ahead of Xception, which is ranked second by 0.08%. This represents a significant advantage in both LACB_ECBAM Net and Xception accuracy over other networks. Xception is significantly optimized at all levels compared to models in the same family. Then, in terms of recall and precision, Xception ranked first, higher than LACB_ECBAM Net. LACB_ECBAM Net's recall is 99.92%, which ranks first, higher than Xception's. In the identification of surface damage of WTB, LACB_ECBAM Net should be more inclined to identify all damaged WTB, preferring to spend a small number of human resources for secondary inspection. Therefore, a higher recall rate is also one of the advantages of LACB_ECBAM Net. Secondly, from the perspective of parameter quantity, the parameter quantity of LACB_ECBAM Net is only 0.58 M, which is the least among the models involved in the experiment. LACB_ECBAM Net replaces the huge fully connected

layer with the global mean pooling layer, so it has an advantage in the number of parameters. In addition, most convolution in LACB_ECBAM Net is improved by depthwise separable convolution, which reduces the number of overall parameters. Finally, in terms of training time, LACB_ECBAM Net is at an average level, and the training time used is 15 min 29 s, which is slower than ShuffleNet_V2 because there are too many "multi-path" [41] structures designed in LACB_ECBAM Net, mainly in LAC Block. It is easy to cause network fragmentation, which will reduce the degree of parallelism of the model [41], and slow the corresponding speed, which is also the problem that LACB_ECBAM Net needs to solve in follow-up research. Additionally, the experimental indicators of LACB_ECBAM Net are stronger than ResNeXt50 (32 × 4d), MobileNet_V2, and EfficientNet-B0.

In summary, from the overall experimental results, compared with other networks in the task of WTB surface damage identification, LACB_ECBAM Net's accuracy, recall, and number of parameters are excellent, but there is still room for improvement in training time.

4. Conclusions

The WTB surface damage identification is an important part of wind turbine maintenance, and there are some relevant identification methods. However, the identification methods based on CNNs have the advantages of low cost, simple operation, and high recognition accuracy, and have therefore attracted the attention of many scholars. With continuous in-depth research, efficient and novel CNN-related structures continue to emerge. For example, the AC Block proposed in 2019 uses asymmetric convolution to obtain better feature extraction capabilities. The CBAM proposed in 2018 applies a novel channel and spatial combined attention mechanism, further improving the feature selection ability. Therefore, this article draws on the valuable ideas of the former two. Additionally, according to the actual task requirements, we make improvements to the problems of the number of parameters of the AC Block needing to be reduced and CBAM lacking multi-scale perception of the damaged space location. A lightweight CNN model of LACB_ECBAM Net with LAC Block and E_CBAM as the main highlights are proposed for the WTB surface damage identification.

The article conducted three related experiments on LACB_ECBAM Net, and the conclusions are as follows:

(1) The LACB_ECBAM network is effective for the WTB surface damage identification task. On the dataset proposed in this article, containing 4729 images and 3 categories, the accuracy reached 99.94−99.96%, which can meet the practical engineering needs.
(2) The LAC Block and E_CBAM used by LACB_ECBAM Net have outstanding generalization ability. On the present study task, AlexNet improved by LAC Block showed a 1.36% improvement in accuracy, 0.77% improvement in recall, 1.86% improvement in precision, and about 2.33 M reduction in the number of parameters. Comparatively, AlexNet improved by E_CBAM showed a 6.32% improvement in accuracy, 6.00% improvement in recall, and 6.12% improvement in precision.
(3) LACB_ECBAM Net has excellent comprehensive performance in the task of WTB surface damage identification. It was superior to the other CNNs participating in the experiment, such as Xception, in terms of accuracy, recall, and the number of parameters.

However, LACB_ECBAM Net also has shortcomings, such as precision needing to be improved and training time needing to be further reduced, which are problems to be solved in our follow-up work.

Author Contributions: Conceptualization, H.C.; methodology, H.C. and Q.S.; validation, L.Z.; formal analysis, L.Z.; investigation, H.C.; resources, H.C.; data curation, H.C. and Q.S.; writing—original draft preparation, L.Z. and H.C.; writing—review and editing, L.Z.; supervision, L.Z.; project administration, L.Z.; funding acquisition, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant 52005071 and the Applied Basic Research Program Project of Liaoning Province under Grant 2023JH2/101300236.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets analyzed during the current study are available from the corresponding author upon reasonable request. The data that support the findings of this study are openly available in "Wind-turbine-blade" at https://github.com/KaKoYu007/Wind-turbine-blade, accessed on 10 April 2023.

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

Appendix A

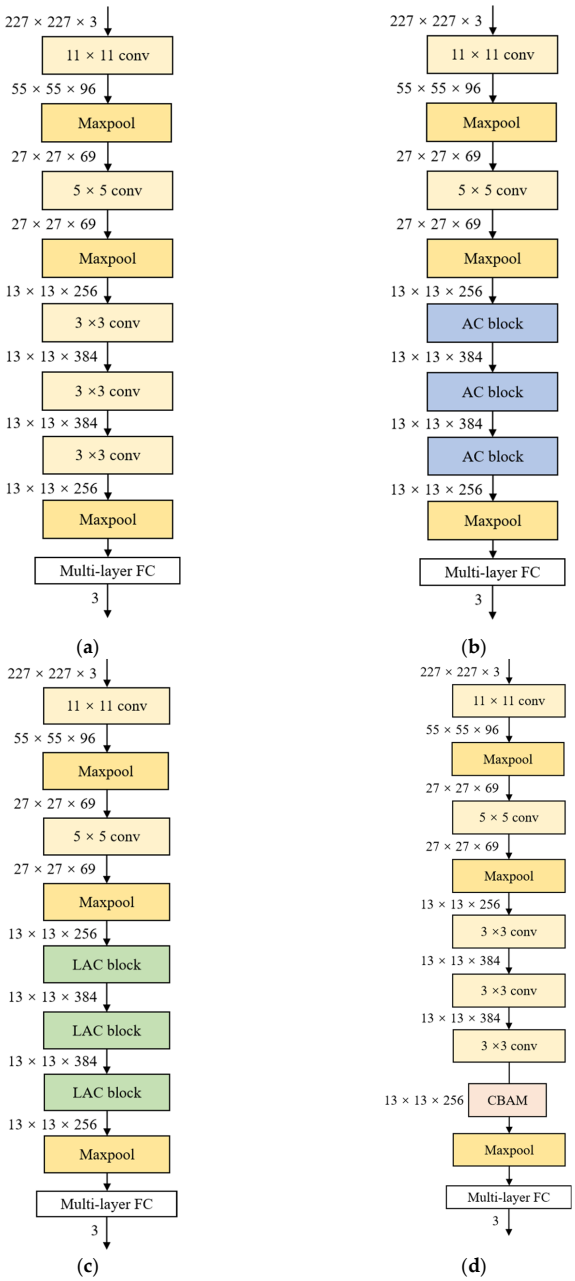


Figure A1. *Cont.*

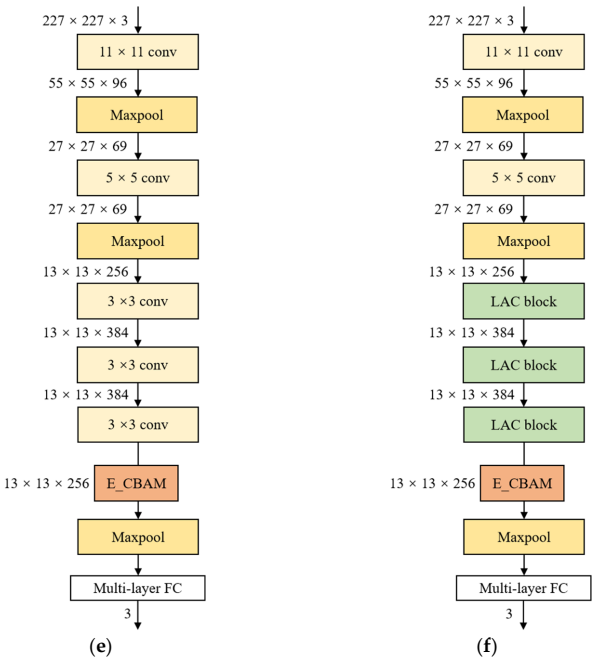
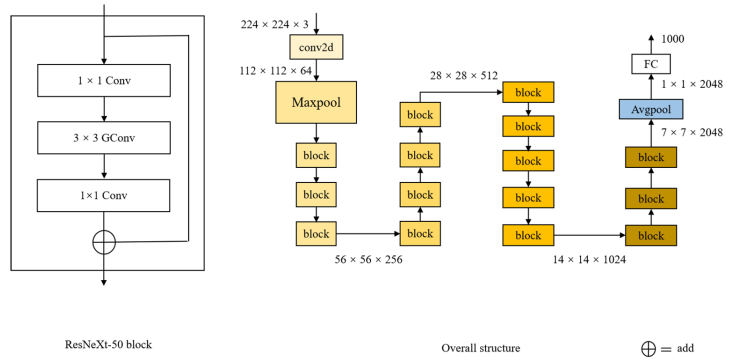
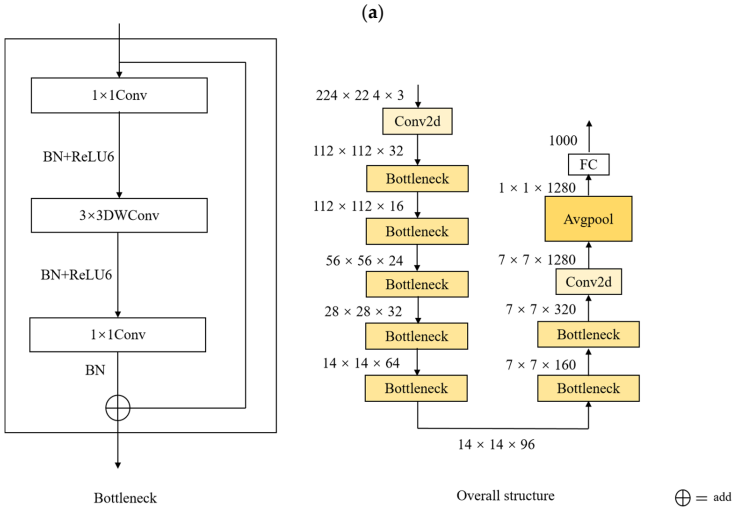


Figure A1. Structure diagram of AlexNet's modifications. (**a**) Original AlexNet. (**b**) AlexNet + AC Block. (**c**) AlexNet + LAC Block. (**d**) AlexNet + CBAM. (**e**) AlexNet + E_CBAM. (**f**) AlexNet + LAC Block + E_CBAM.

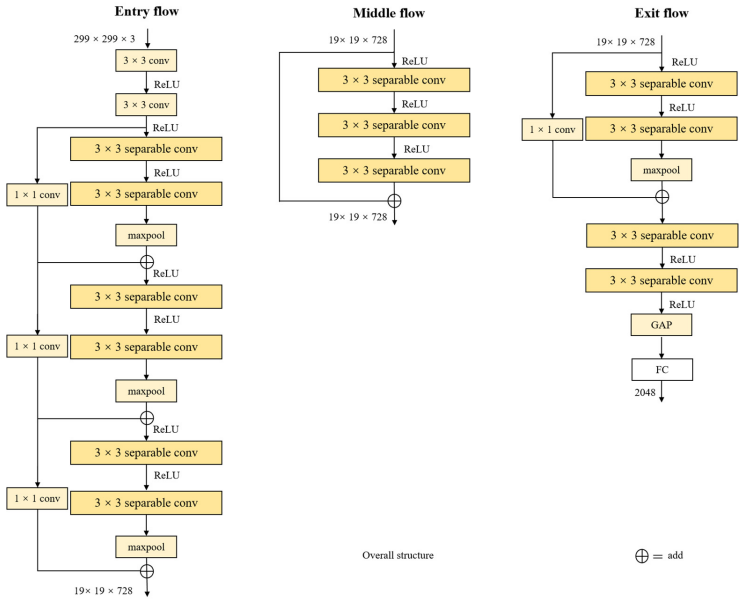


(**a**)

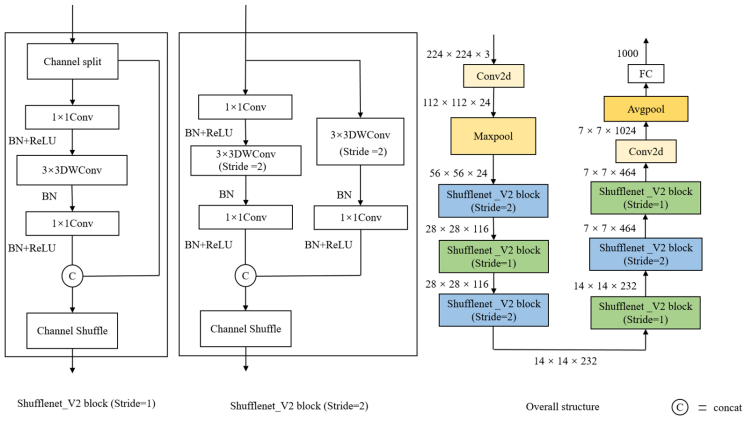


(**b**)

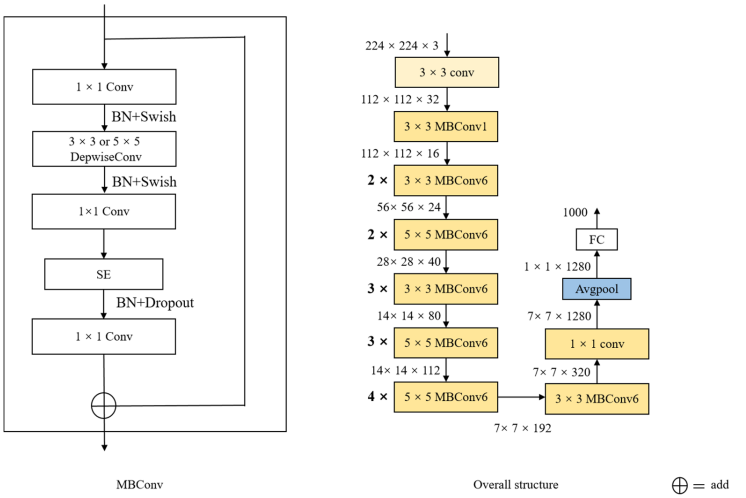
Figure A2. *Cont.*



(**c**)



(**d**)



(**e**)

Figure A2. Main structure diagram of the six lightweight networks. (**a**) ResNeXt50 (32 × 4d). (**b**) Xception. (**c**) MobileNet_V2. (**d**) ShuffleNet_V2. (**e**) EfficientNet-B0.

References

1. Global Wind Report 2023—Global Wind Energy Council. Available online: https://gwec.net/globalwindreport2023/ (accessed on 6 May 2023).
2. Shohag, M.A.S.; Hammel, E.C.; Olawale, D.O.; Okoli, O.I. Damage mitigation techniques in wind turbine blades: A review. *Wind Eng.* **2017**, *41*, 185–210. [CrossRef]
3. Mishnaevsky, L., Jr.; Thomsen, K. Costs of repair of wind turbine blades: Influence of technology aspects. *Wind Energy* **2020**, *23*, 2247–2255. [CrossRef]
4. Hwang, S.; An, Y.K.; Sohn, H. Continuous line laser thermography for damage imaging of rotating wind turbine blades. *Procedia Eng.* **2017**, *188*, 225–232. [CrossRef]
5. Hwang, S.; An, Y.K.; Sohn, H. Continuous-wave line laser thermography for monitoring of rotating wind turbine blades. *Struct. Health Monit.* **2019**, *18*, 1010–1021. [CrossRef]
6. Sanati, H.; Wood, D.; Sun, Q. Condition monitoring of wind turbine blades using active and passive thermography. *Appl. Sci.* **2018**, *8*, 2004. [CrossRef]
7. Doliński, Ł.; Krawczuk, M.; Żak, A. Detection of delamination in laminate wind turbine blades using one-dimensional wavelet analysis of modal responses. *Shock Vib.* **2018**, *2018*, 1–15. [CrossRef]
8. Hoell, S.; Piotr, O. Sequential projection pursuit for optimised vibration-based damage detection in an experimental wind turbine blade. *Smart Mater. Struct.* **2018**, *27*, 025007. [CrossRef]
9. Zhang, L.; Liu, K.; Wang, Y.; Omariba, Z.B. Ice detection model of wind turbine blades based on random forest classifier. *Energies* **2018**, *11*, 2548. [CrossRef]
10. Tang, J.; Soua, S.; Mares, C.; Gan, T.H. An experimental study of acoustic emission methodology for in service condition monitoring of wind turbine blades. *Renew. Energy* **2016**, *99*, 170–179. [CrossRef]
11. Beale, C.; Niezrecki, C.; Inalpolat, M. An adaptive wavelet packet denoising algorithm for enhanced active acoustic damage detection from wind turbine blades. *Mech. Syst. Signal Process.* **2020**, *1142*, 106754. [CrossRef]
12. Xu, D.; Liu, P.F.; Chen, Z.P. Damage mode identification and singular signal detection of composite wind turbine blade using acoustic emission. *Compos. Struct.* **2021**, *255*, 112954. [CrossRef]
13. Tian, S.; Yang, Z.; Chen, X.; Xie, Y. Damage detection based on static strain responses using FBG in a wind turbine blade. *Sensors* **2015**, *15*, 19992–20005. [CrossRef] [PubMed]
14. Coscetta, A.; Minardo, A.; Olivares, L.; Mirabile, M.; Longo, M.; Damiano, M.; Zeni, L. Wind turbine blade monitoring with Brillouin-based fiber-optic sensors. *J. Sens.* **2017**, *2017*, 9175342. [CrossRef]
15. Wen, B.; Tian, X.; Jiang, Z.; Li, Z.; Dong, X.; Peng, Z. Monitoring blade loads for a floating wind turbine in wave basin model tests using Fiber Bragg Grating sensors: A feasibility study. *Mar. Struct.* **2020**, *71*, 102729. [CrossRef]
16. Guo, J.; Liu, C.; Cao, J.; Jiang, D. Damage identification of wind turbine blades with deep convolutional neural networks. *Renew. Energy* **2021**, *174*, 122–133. [CrossRef]
17. Yang, X.; Zhang, Y.; Lv, W.; Wang, D. Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier. *Renew. Energy* **2021**, *163*, 386–397. [CrossRef]
18. Liu, H.; Zhang, Z.; Jia, H.; Li, Q.; Liu, Y.; Leng, J. A novel method to predict the stiffness evolution of in-service wind turbine blades based on deep learning models. *Compos. Struct.* **2020**, *252*, 112702. [CrossRef]
19. Yang, P.; Dong, C.; Zhao, X.; Chen, X. The surface damage identifications of wind turbine blades based on ResNet50 algorithm. In Proceedings of the 39th Chinese Control Conference (CCC), Nanning, China, 23–25 July 2021.
20. Chen, X.; Lei, D.; Xu, G. Prediction of icing fault of wind turbine blades based on deep learning. In Proceedings of the 2nd International Conference on Automation Electronics and Electrical Engineering (AUTEEE), Zhuhai, China, 9–11 August 2019.
21. Yuan, B.; Wang, C.; Luo, C.; Jiang, F.; Long, M.; Yu, P.S.; Liu, Y. WaveletAE: A wavelet-enhanced autoencoder for wind turbine blade icing detection. *arXiv* **2019**, arXiv:1902.05625.
22. Ran, X.; Zhang, S.; Wang, H.; Zhang, Z. An Improved Algorithm for Wind Turbine Blade Defect Detection. *IEEE Access* **2021**, *10*, 122171–122181. [CrossRef]
23. Yu, Y.; Cao, H.; Liu, S.; Yang, S.; Bai, R. Image-based damage recognition of wind turbine blades. In Proceedings of the 2nd International Conference on Advanced Robotics and Mechatronics (ICARM), Nanjing, China, 15–17 December 2017.
24. Guo, H.; Cui, Q.; Wang, J.; Fang, X.; Yang, W.; Li, Z. Detecting and positioning of wind turbine blade tips for uav-based automatic inspection. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019.
25. Chen, Q.; Liu, Z.H.; Lv, M.Y. Attention Mechanism-based CNN for Surface Damage Detection of Wind Turbine Blades. In Proceedings of the International Conference on Machine Learning Cloud Computing and Intelligent Mining (MLCCIM), Xiamen, China, 27–29 May 2022.
26. Zou, L.; Cheng, H. Research on Wind Turbine Blade Surface Damage Identification Based on Improved Convolution Neural Network. *Appl. Sci.* **2022**, *12*, 9338. [CrossRef]
27. Ding, X.; Guo, Y.; Ding, G.; Han, J. ACNet: Strengthening the Kernel Skeletons for Powerful CNN via Asymmetric Convolution Blocks. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 20–26 October 2019; pp. 1911–1920.

28. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
29. Ma, R.; Wang, J.; Zhao, W.; Guo, H.; Dai, D.; Yun, Y.; Ma, D.; Li, L.; Hao, F.; Bai, J. Identification of Maize Seed Varieties Using MobileNetV2 with Improved Attention Mechanism CBAM. *Agriculture* **2022**, *13*, 11. [CrossRef]
30. Chen, L.; Yao, H.; Fu, J.; Ng, C.T. The classification and localization of crack using lightweight convolutional neural network with CBAM. *Eng. Struct.* **2023**, *275*, 115291. [CrossRef]
31. Shi, Y.; Ma, D.; Lv, J.; Li, J. ACTL: Asymmetric convolutional transfer learning for tree species identification based on deep neural network. *IEEE Access* **2021**, *9*, 13643–13654. [CrossRef]
32. Liu, Y.; Zhou, J.; Qi, W.; Li, X.; Gross, L.; Shao, Q.; Zhao, Z.; Fan, X.; Li, Z. ARC-Net: An efficient network for building extraction from high-resolution aerial images. *IEEE Access* **2020**, *8*, 154997–155010. [CrossRef]
33. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
34. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
35. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 510–519.
36. Hendrycks, D.; Gimpel, K. Bridging nonlinearities and stochastic regularizers with Gaussian error linear units. *arXiv* **2016**, arXiv:1606.08415.
37. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
38. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.15564.
39. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Rabinovich, A.; Erhan, D. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Zurich, Switzerland, 6–12 September 2014.
40. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–21 June 2018.
41. Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
42. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807.
43. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
44. Mingxing, T.; Le Quoc, V.E. Rethinking Model Scaling for Convolutional Neural Networks. *arXiv* **2019**, arXiv:1905.11946.

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