

# Wood Defect Classification Based on Lightweight Convolutional Neural Networks

Yonghua Xie \* and Jiaxin Ling

Different types of wood defects correspond to different processing methods. Good classification means can transform defective boards into practical boards after appropriate processing. The detection accuracy of the wood surface defects is particularly important for improving the utilization rate and speed of processing the boards. The RegNet stands out in the field of computer vision. It automatically designs the network model based on the design space and applies it to wood defect detection, which can improve the classification accuracy. When the convolutional structure of the RegNet network is applied to industrial detection and classification, the problems of long real-time detection time and large algorithm parameters persist. This study focuses on collecting wood material images of common coniferous and broad-leaved trees in Northeast China with three types of defects: wormholes, slip knots, and dead knots. To improve the allocation of computing resources, based on the RegNet network model, an attention mechanism module was added, and the Ghostconv structure was introduced. The structure quickly and accurately highlighted the types of wood defects, improved the classification accuracy, reduced the parameters of the network, and exhibited generalization ability. To verify the performance of the improved network, MobileNet-v2, EfficientNet, and Vision-Transformer networks were introduced for comparative analysis. The improved RegNet network had smaller weight and higher accuracy, with a classification accuracy of 96.58%.

DOI: 10.15376/biores.18.4.7663-7680

Keywords: Board defect; Neural Architecture Search; RegNet; Attention mechanism

Contact information: College of Computer and Control Engineering, Northeast Forestry University, Harbin, 150040, People's Republic of China; \*Corresponding author: zdhxyh@163.com

## INTRODUCTION

Wood defects refer to the general names of various characteristics that reduce the commodity and use value of wood. According to the formation mode, wood defects can be divided into growth, biological hazard, and processing defects (Liu 2008; Diao *et al.* 2012; Luo and Sun 2019; Yan *et al.* 2022). Different types of board defects have different wood processing methods, which are convenient for the subsequent processing, production, and daily use of wood. The traditional method of manually identifying wood defects involves a large workload and is time-consuming, which is a huge demand on human and material resources. With improvements in computing technology, convolutional neural network provides convenient conditions for image recognition and classification. Many excellent algorithm models have significantly progressed in visual recognition. Zhang *et al.* (2021) designed an improved LeNet-5 model that can accurately classify common defects and defect-free samples of wood by increasing the network depth and batch normalization. Shi

*et al.* (2020) proposed a method combining adaptive component analysis and deep migration feed forward neural network to effectively migrate the spectrum and defect corresponding knowledge of other tree species to the target classifier and improve the performance of the classifier. Riana *et al.* (2021) used K-means to segment wood defect images, extracted six texture and shape features using the gray level co-occurrence matrix (GLCM) method, and calculated the shortest distance between the feature values of the test set and the training set using *the Euclidean distance method to achieve the purpose of identifying wood defects* (Riana *et al.* 2021). At the 2017 ICLR conference, Barret Zoph and his research team presented the NAS algorithm for the first time, using the controller RNN model to create a string describing the network structure, and then using the policy gradient algorithm to update the control variables to maximize the accuracy of the created network (Zoph and Le 2016). After experimental verification, in some fields of artificial intelligence, such as image classification (Zhang *et al.* 2021) target detection (Wang *et al.* 2020), semantic segmentation (Hu *et al.* 2022), and other fields, Guo *et al.* (2020) proposed a multichannel image segmentation method for wood knot defects based on the TVCV model. This method controls the range of the active contour *via* regularization parameters to capture the object. Furthermore, it uses the multi-channel weighted level set as the initial value, iterates with the TVCV model, and finally considers the calculated value as the segmentation result. The segmentation effect of the model is better. The aforementioned model is designed manually, and the design of the network model is mainly reflected in the application of network layer operators. Furthermore, the performance optimization process of the model is accidental and uncertain.

Network design has changed from manual exploration process to automatic network design, and several automatic search network structure models have been developed. For example, Gong *et al.* (2019) combined a neural architecture search (NAS) with a generative adversarial network (GAN), defined the search space of GAN network structure change, used an RNN controller to guide structure search, and combined parameter sharing and dynamic reset strategy to improve training speed. Sun *et al.* (2020) proposed a method to solve the image classification problem by using a genetic algorithm to evolve the structure and connection weight initialization value of a deep convolution neural network. In this algorithm, an effective variable length gene coding strategy is designed to represent the different building blocks and potential optimal depth of convolution neural network. Li *et al.* (2020) proposed a sequential greedy architecture search (SGAS) algorithm to automatically design the architecture of CNN and GCN. By considering the heuristic criteria of edge importance, selection certainty and selection stability, they solved the two-level optimization problem in NAS in a greedy manner. Although the aforementioned method uses an automatic design network algorithm, it still has the disadvantage of relying on a manually designed network model structure.

The RegNet network is an automatic search design model proposed by Kaiming He (Radosavovic *et al.* 2020). The network and its related network models exhibited good results in the field of computer vision, which can automatically design the network framework in a design space to simplify the task of deep learning. However, the application of rapid detection of wood defects still persists with the problem of long real-time detection time. In this study, the goal was to use the RegNet network with an increased attention mechanism to identify the types of wood defects, such as wormholes, slip knots and dead knots, and adjust the partial convolutional structure of the bottleneck network to further improve the classification accuracy of wood defects, reduce the detection time and reduce the algorithm parameters. Based on the requirements of computer hardware, in this study,

a reference is provided for the industrial application of the rapid classification of wood defects.

## EXPERIMENTAL

### Data Set Construction

Data set construction was based on the early wood defect research sample set of the project team (Xie 2014), which was supplemented with common coniferous and broad-leaved samples from Northeast China. The team collected wormhole, live knot, and dead knot samples to establish a small database that was divided into training, validation, and test sets in a 6:2:2 ratio. Owing to the small number of samples in the dataset, it was not suitable for deep learning. The samples in the training, validation and test sets were processed using rotation, translation, scale transformation, and gray transformation to expand the sample database (Ling and Xie 2022). After capacity expansion, the training set contained 3,310 sample data, the validation set contained 1,054 sample data, and the test set contained 1,054 sample data. Some samples of the dataset are shown in Fig. 1.

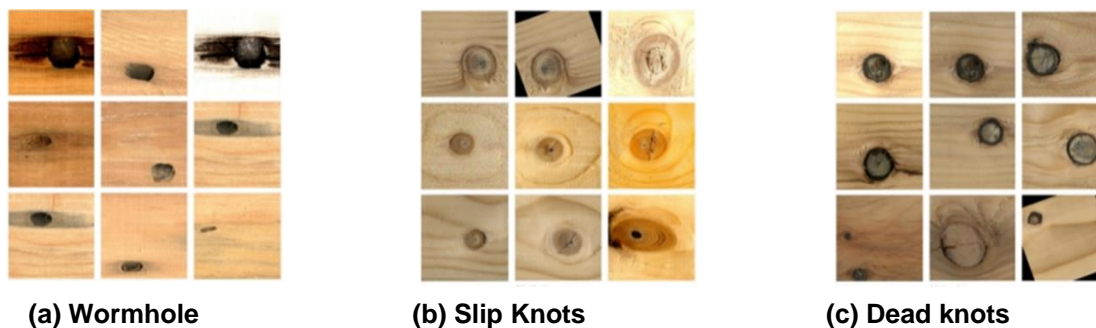
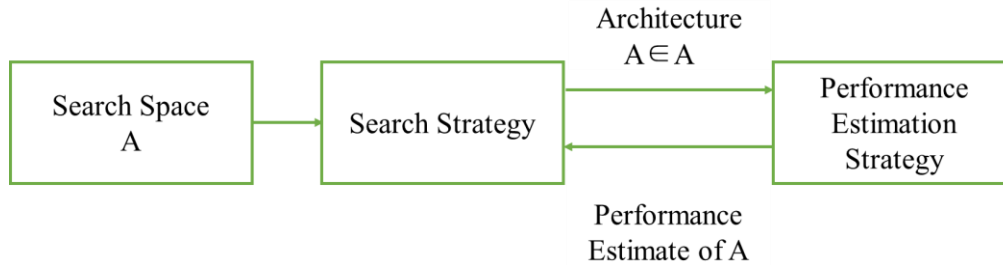


Fig. 1. Samples of defective part of a wood

### NAS Algorithm

Prior to its introduction, the RegNet model was used. First, the neural architecture search algorithm (NAS) was introduced. In deep learning, hyperparameters are divided into two categories: training parameters and parameters that define the network structure. The training parameters include the learning rate, batch size, attenuation weight, and other parameters. The automatic tuning of the training parameters belongs to hyperparameter optimization (HO). The parameters defining the network structure include the layer structure operator, convolution kernel size, dimension, and dispersion degree. Automatic optimization of network structure parameters is generally termed neural architecture search (NAS).

The classic NAS algorithm includes three aspects: search space, search strategy, evaluation, and prediction. The search strategy is selected based on search space. When the strategy is determined, the designed network structure is evaluated and predicted, and it continues to strengthen the search for a better network model, which corresponds to the NAS algorithm. Figure 2 shows the rules of the NAS algorithm.



**Fig. 2.** NAS algorithm

### *Search space*

In principle, using the NAS algorithm, a space composed of all potential network structure models is defined as the search space. The neural network architecture searches the network model structure parameters that can be optimized, including the number of layers  $n$ , type of operation performed by layers, and the super parameters in the layer structure.

### *Search strategy*

Given the search space, the best neural network structure can be determined using the corresponding search strategy. The common search strategies include reinforcement-based learning, evolutionary algorithms, Bayesian optimization algorithms, and gradient-based methods (Jin *et al.* 2019).

### *Evaluation and prediction*

The purpose of the search strategy involves designing a good network structure and ensuring that the network structure exhibits the highest accuracy in the test set. To strengthen the search process, the search strategy must evaluate the performance of a given network structure.

## **RegNet Model**

The RegNet model uses the neural architecture search algorithm (NAS) to obtain a low-dimensional design space composed of a simple rule network in a relatively unconstrained design space by using the man-machine rotation method. This design space is termed as RegNet space (Han *et al.* 2020). The central idea of the RegNet design space is that the width and depth of the model are determined by a quantitative linear function. The structure block diagram of RegNet network is shown in Fig. 3.

The network is mainly composed of stem, body, and head.

### *Stem*

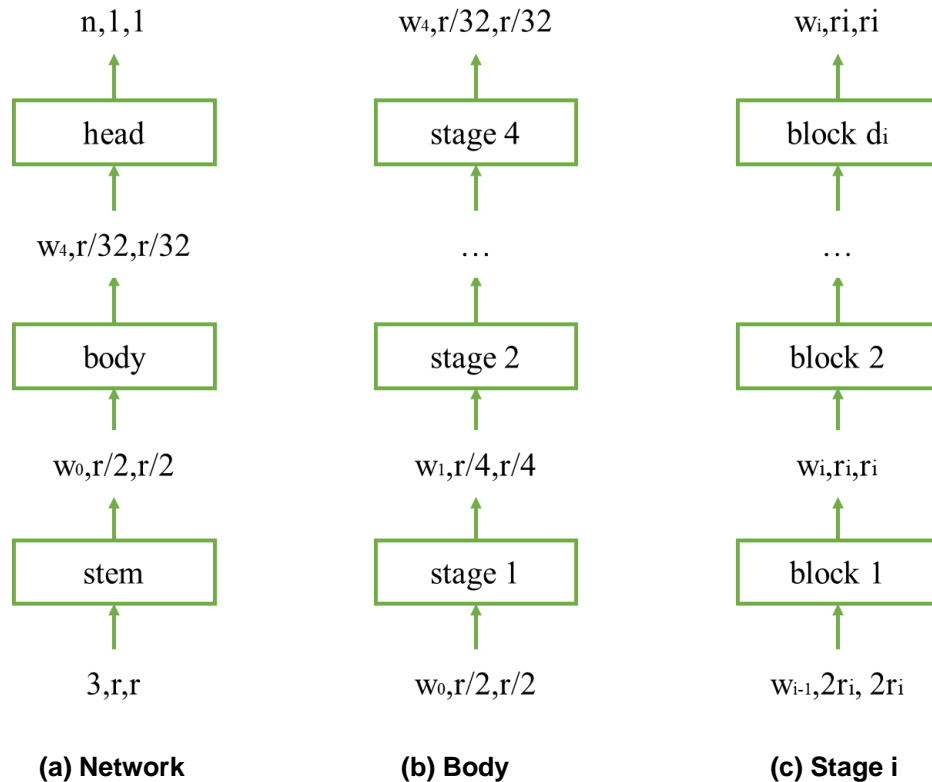
The stem is composed of a convolution layer containing a BN layer and ReLU activation function. The size of the convolution kernel is  $3 \times 3$ . The step length is two and the number of convolution kernels is 32.

### *Body*

The body is composed of four-stage stacks, and each stage is stacked with a large number of blocks. In this stage, in the main branch and shortcut branch of the first block part of the stage, there is a group convolution with two steps, and the number of convolution steps for the other block parts corresponds to one.

### Head

The head is composed of a global average pooling layer and full connection layer as part of the classifier in the classification network.



**Fig. 3.** Processing flow of the RegNet network model

The main branch of the block module of the RegNet model is a  $1 \times 1$  convolution kernel. The group convolution of 1 (including BN and ReLU) and one convolution kernel is  $3 \times 3$ . The group convolution of three (including BN and ReLU) and one convolution kernel is  $1 \times 1$  (including BN). On the shortcut branch, when stride = 1, the input data information is not processed. When stride = 2, the input data is processed *via* a convolution kernel  $1 \times 1$  (including BN) to down sample the information. The block module structure of the RegNet model is shown in Fig. 4.

In Fig. 4,  $r$  denotes the resolution, which can be understood as the height and width of the characteristic matrix. When stride = 1, the input and output  $r$  are equal, and when stride = 2, the output  $r$  is half that of the input. Specifically,  $w$  denotes the number of channels in the characteristic matrix,  $g$  denotes the group width in the group convolution,  $b$  denotes the bottleneck ratio, and the number of channels in the output characteristic matrix is reduced to the number of channels  $1/b$  of the output characteristic matrix. The image data of wood defects are input into the group convolution module of the stem part of the RegNet network, and feature extraction, recognition, and classification were performed according to the network structure.

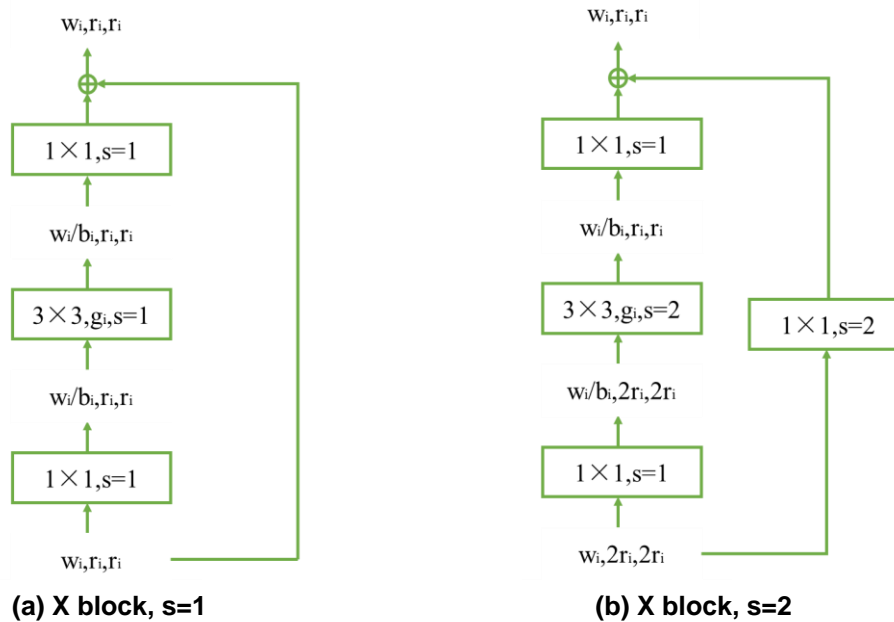


Fig. 4. Block module structure in RegNet mode

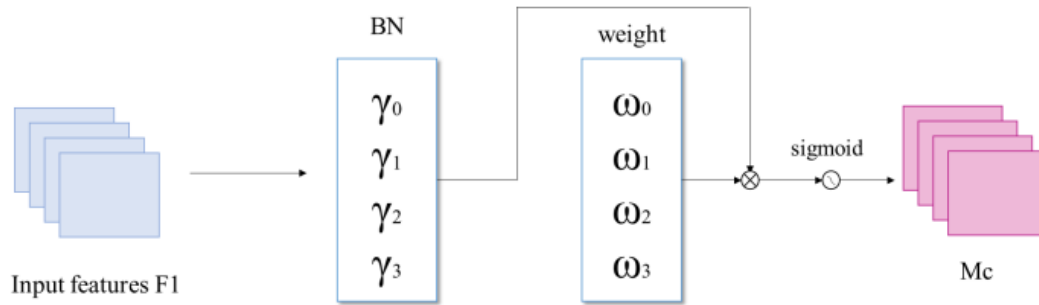
### Improved RegNet Model

To solve the problem of limited computing power, computing resources are allocated to more important tasks, and the problem of information overload is solved simultaneously. In this study, an attention mechanism module was added to the RegNet network. In computer vision, the attention mechanism module was mainly divided into a spatial attention module, channel attention module, position pixel attention module, and hybrid attention module. The spatial attention module adjusts the self-attention of each position of the feature map,  $(x, y)$  two-dimensional adjustment, to ensure that the model focuses on the areas worthy of higher attention. The channel attention module allocates resources to each convolution channel and adjusts the  $z$ -axis in one dimension. The position pixel attention module focuses more on the correlation between the pixels on the feature map and other pixels and realizes a global response to the output.

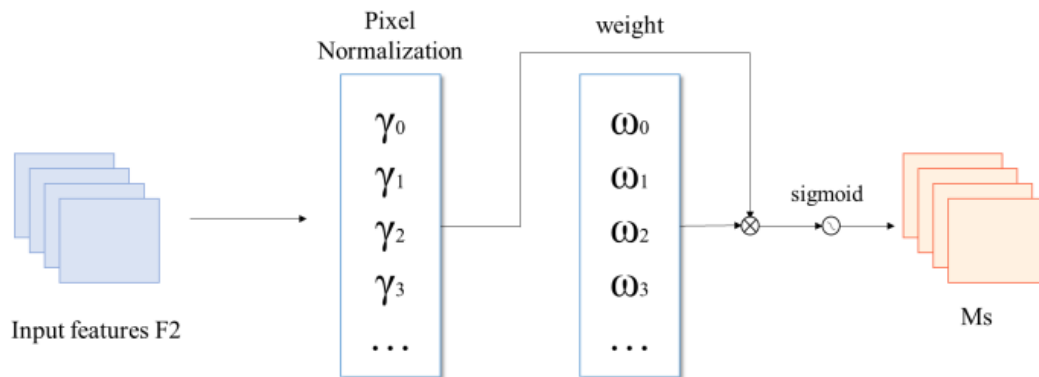
Squeeze-and-excitation networks belong to the channel attention module. The attention mechanism was divided into two steps: squeeze and exception. In squeeze step, the global compressed feature of the current feature map is obtained by performing global average pooling on the feature map layer. The excitation step is used to obtain the weight of each channel in the feature map *via* the two-layer fully connected bottleneck structure. Furthermore, weighted feature map is used as the input of the next layer network.

Efficient channel attention also belongs to the channel attention mechanism. Efficient channel attention avoids the dimensionality reduction due to channel compression, uses 1D convolution to efficiently realize local cross-channel interaction, and extracts the dependencies between channels.

After the group convolution structure in the block module of the RegNet model network, in this study, the attention mechanism module NAM (Liu *et al.* 2021) was added. Specifically, NAM adopts the module integration mode, including the channel attention submodule and spatial attention submodule. A schematic diagram of the channel attention of the NAM module is shown in Fig. 5(a), and a spatial attention schematic diagram is shown in Fig. 5(b).



(a) Channel attention sub module of NAM module



(b) Spatial attention sub module of NAM module

**Fig. 5.** Attention sub module of NAM module

In the channel attention submodule, the scaling factor is added in batch normalization. The scaling factor reflects the degree of change in each channel and its importance. The mathematical relationship of the scaling factor is as follows,

$$B_{out} = \text{BN}(B_{in}) = \gamma \frac{B_{in} - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} + \beta \quad (1)$$

where  $B_{in}$  and  $B_{out}$  denote batch input and output feature data;  $\text{BN}(\cdot)$  denotes the Batch Normalization function;  $\mu_B$  and  $\sigma_B$  denote the average value and standard deviation of batch  $B$ , respectively;  $\gamma$  denotes the scale transformation coefficient of affine transformation;  $\beta$  denotes the displacement transformation coefficient of affine transformation; and  $\varepsilon$  prevents the variance from being 0, resulting in invalid calculation. The calculation principle of the channel attention submodule is as follows:

$$M_c = \text{sigmoid} (W_\gamma(\text{BN}(F_1))) \quad (2)$$

where  $F_1$  denotes the input feature;  $M_c$  denotes the output feature ;  $\gamma$  denotes the scale factor of each channel;  $\text{BN}(\cdot)$  denote the Batch Normalization function;  $\text{sigmoid}(\cdot)$  denotes the sigmoid function; and the weight  $W_\gamma$  represents the scale factor of the BN and importance of the pixels. The calculation formula can be expressed as follows,

$$W_\gamma = \frac{\gamma_i}{\sum_{j=0} \gamma_j} \quad (3)$$

where  $\gamma_i, \gamma_j$  denotes the scale transformation coefficient of affine transformation. The calculation principle of the spatial attention submodule is shown in as follows,

$$M_s = \text{sigmoid}(W_\lambda (\text{BN}_s(F_2))) \quad (4)$$

where  $F_2$  denotes the input feature;  $M_s$  denotes the output feature;  $\text{BN}_s(\cdot)$  denotes the Batch Normalization function;  $\text{sigmoid}(\cdot)$  denotes the sigmoid function;  $\lambda$  denotes the scale factor of each channel, and the weight  $W_\lambda$  is as follows:

$$W_\lambda = \frac{\lambda_i}{\sum_{j=0} \lambda_j} \quad (5)$$

To suppress the small weight and prevent overfitting, a regularization term is added to the loss function, as shown in Eq. 6. In the equation,  $x$  denotes the input,  $y$  denotes the output,  $W$  denotes the network weight,  $l(\cdot)$  denotes the loss function,  $f(\cdot)$  denotes the activate function,  $g(\cdot)$  denotes the L1 norm penalty function, and  $p$  denotes the punishment for balancing  $g(\gamma)$  and  $g(\lambda)$  follows:

$$\text{Loss} = \sum_{(x,y)} l(f(x, W), y) + p \sum g(\gamma) + p \sum g(\lambda) \quad (6)$$

The output feature map of the convolution layer usually includes considerable redundancy. A large number of experiments demonstrated that the generation of these redundant feature maps individually with a large amount of traffic and parameters is a waste of computing and hardware storage resources (Han *et al.* 2020). To solve the problem of redundancy in the output characteristic diagram of the convolution layer, in this study, the Ghostmodule was introduced into the body part of the RegNet network. The Ghostmodule is a method of model compression, *i.e.*, while ensuring network accuracy, it can effectively reduce the network parameters and amount of calculation for improving the calculation speed and reducing the delay effect. In the improved model, the Ghostmodule was embedded in the bottleneck part of the RegNet network to replace some group convolution structures in the network.

The operating principle of the Ghostmodule convolution is shown in Fig. 6. The calculation of the module includes two steps. The first part is the ordinary convolution of the input features and the second part is the separation convolution of the results of the first part. Based on the calculation results of the two parts, the feature fusion between channels is conducted *via* concat. The block module structure in the improved RegNet model is shown in Fig. 7.



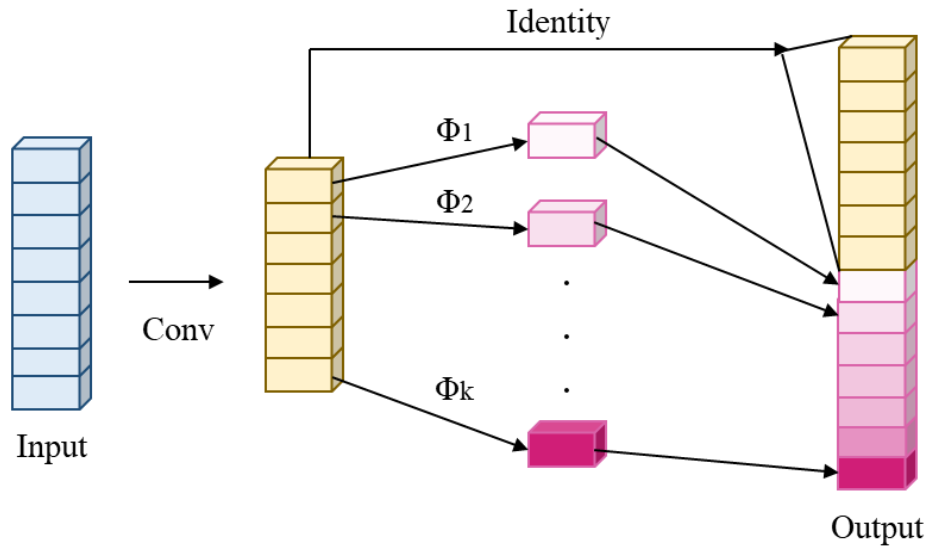


Fig. 6. Ghost module principle

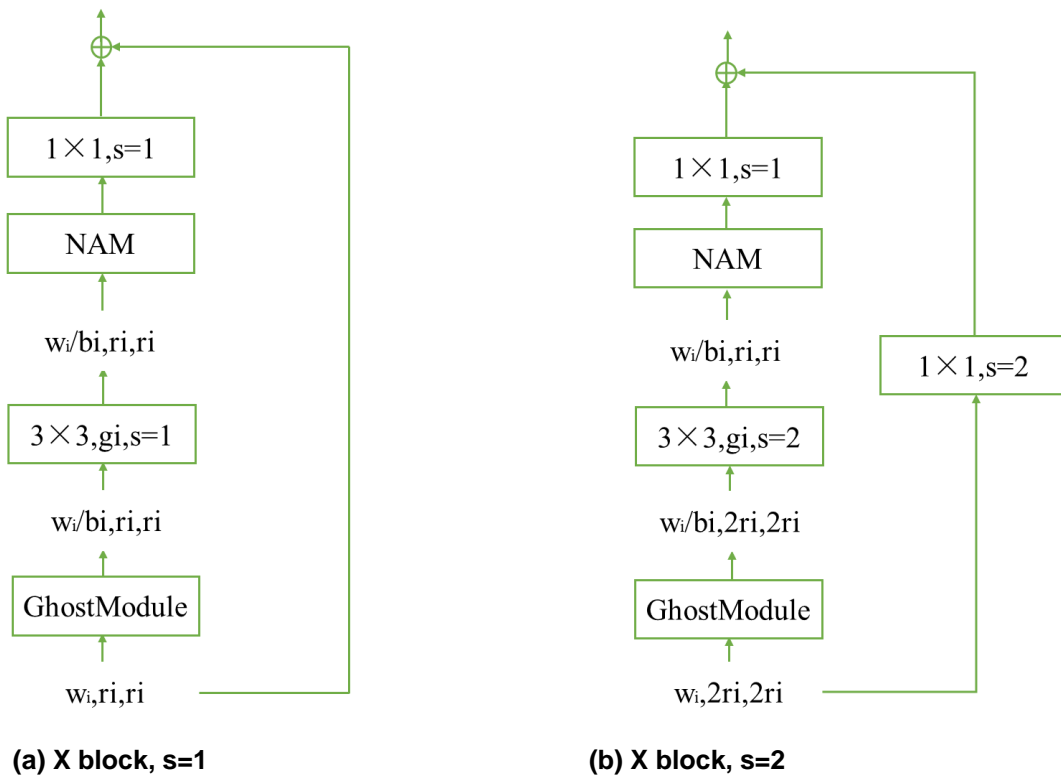


Fig. 7. Block module structure in the improved RegNet model

**Experimental Process**

The hardware and software environment of the experiment is shown in Table 1. The training set samples in the wood defect sample dataset were input into the above test network to obtain the training weight. The validation set samples were placed in the wood defect samples in each network to automatically adjust the appropriate super parameters.

Finally, the test set samples were placed into each network model with the best training, and the most suitable network model was obtained for this dataset. During the experiment, it was necessary to control the consistency of other super parameters, and finally, to compare the test set accuracy of each model, number of network parameters, and various evaluation performance indexes of the model.

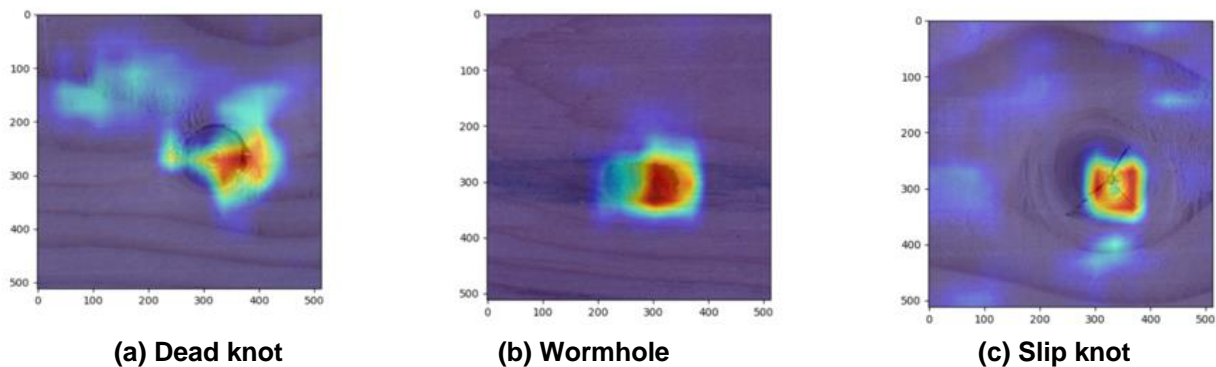
**Table 1.** The Hardware and Software Environment of the Experiment

Experimental environment	Version Details
GPU model	NVIDIA GTX980
CPU model	Intel(R)Core (TM) i5-6300HQ CPU@2.30GHz
Memory size	4G
System Version	Windows 10
programming language	Python 3.8.0
Python version	Pytorch 1.7.1

The experiment was completed in two steps. First, the classification performance of RegNet, RegNet+SE, RegNet+ECA, RegNet+NAM, and GhostRegNet+NAM networks were compared to select high-performance classification networks. Then, the selected network, MobileNet-v2, EfficientNet, and Vision-Transformer networks were selected twice to find the best network model suitable for the wood defect classification.

## RESULTS AND DISCUSSION

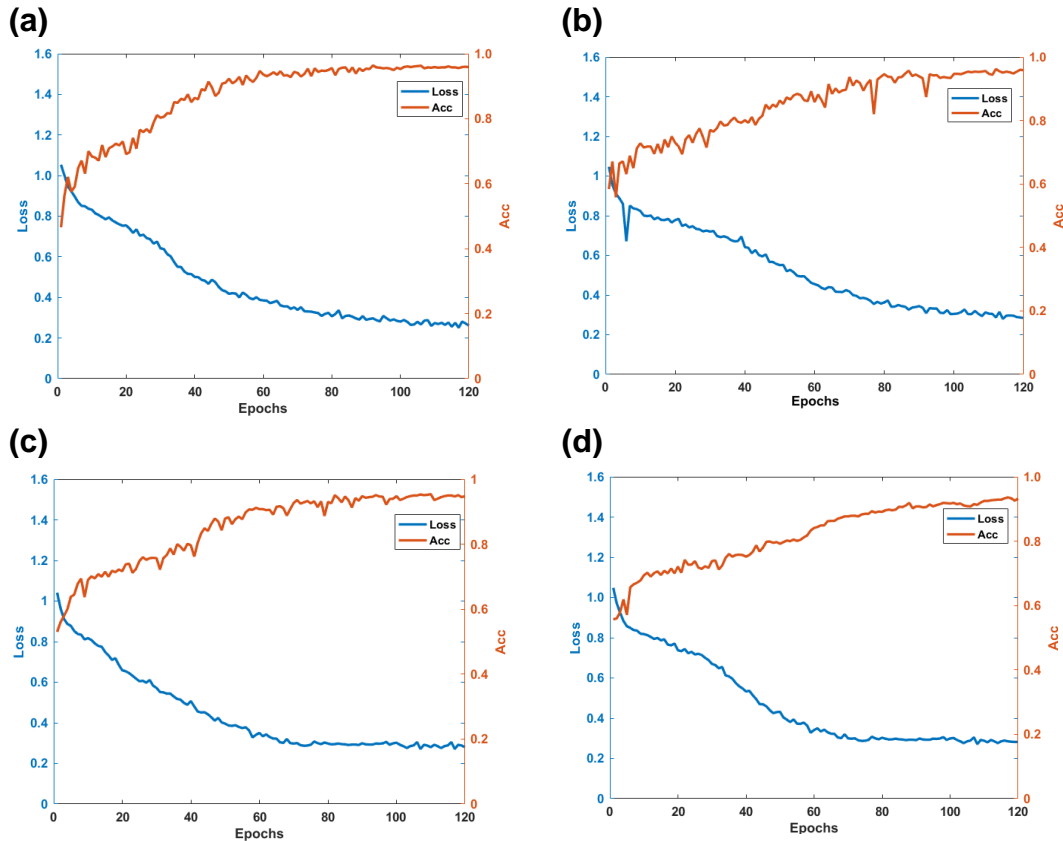
To demonstrate the role of the attention mechanism in network applications, the test set samples of various wood defects were input into the GhostRegNet+NAM network model, and the visual thermal diagram of the Nam attention mechanism was drawn by the Grad-CAM network, as shown in Fig. 8.



**Fig. 8.** Thermodynamic diagram of attention mechanism of various samples in the test set

## Experimental Data and Results

The training set samples were placed into the test network and the number of experimental epochs was set to 120, the number of learning rate was set to 0.001, and the number of batch size was set to 32 for comparative experiments.



**Fig. 9.** Loss function and accuracy of networks (a) Loss function and accuracy of GhostRegNet+NAM network (b) Loss function and accuracy of RegNet network (c) Loss function and accuracy of RegNet+SE network (d) Loss function and accuracy of RegNet+ECA network

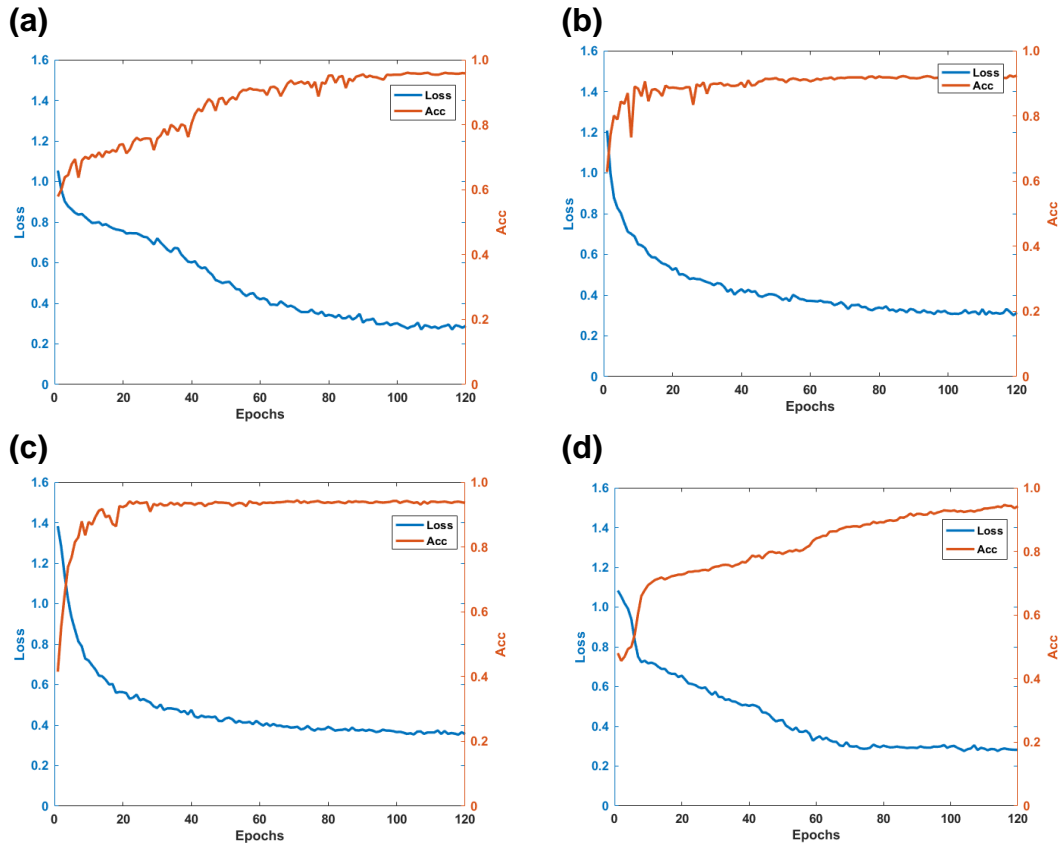
(1) RegNet network realized wood defect classification. In the 108<sup>th</sup> epoch, the accuracy of the validation set was highest at 95.2%. The training set loss function and validation set accuracy of the RegNet network are shown in Fig. 9(b).

(2) RegNet+SE network classified the classification of wood defects. In the 105<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 95.9%. The training set loss function and validation set accuracy of the RegNet+SE network are shown in Fig. 9(c).

(3) RegNet+ECA network realized the classification of wood defects. In the 116<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 95.2%. The training set loss function and validation set accuracy of the RegNet+ECA network are shown in Fig. 9(d).

(4) RegNet+NAM network realized wood defect classification. In the 113<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 96.2%. The training set loss function and validation set accuracy of the RegNet+NAM network are shown in Figure 10(a).

(5) GhostRegNet+NAM network realized wood defect classification. In the 107<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 96.3%. The training set loss function and validation set accuracy of the GhostRegNet+NAM network are shown in Fig. 9(a).



**Fig. 10.** Loss function and accuracy of networks (a) Loss function and accuracy of RegNet+NAM network (b) Loss function and accuracy of MobileNet-v2 network (c) Loss function and accuracy of EfficientNet network (d) Loss function and accuracy of Vision-Transformer network

The training set samples were placed into the Mobilenet-v2, EfficientNet, and Vision-Transformer networks. The number of experimental epochs was set to 120, the number of learning rate was set to 0.001, and the number of batch size was set to 32 for comparative experiments.

(1) MobileNet-v2 network realized wood defect classification. In the 118<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 92.4%. The training set loss function and validation set accuracy of MobileNet-v2 network are shown in Fig. 10(b).

(2) EfficientNet network realized wood defect classification. In the 103<sup>rd</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 94.2%. The training set loss function and validation set accuracy of EfficientNet network are shown in Fig. 10(c).

(3) Vision-Transformer networks realized wood defect classification. In the 115<sup>th</sup> epoch, the accuracy of the validation set was the highest, with a maximum accuracy of 94.6%. The training set loss function and validation set accuracy of Vision-Transformer network are shown in Fig. 10(d).

### Model Evaluation Index

The weight of the highest accuracy on each network model training set was saved, which was convenient for the test set to identify and predict the category of wood defect

samples. TP represents the positive sample, predicted as positive by the model, TN represents the negative sample, predicted as negative by the model, FP denotes the negative sample, predicted as positive by the model, and FN denotes the positive sample predicted as negative by the model. A confusion matrix diagram can be drawn based on the prediction and classification of the various samples. Using TP, TN, FP, and FN, performance indices, such as accuracy, recall, and specificity were obtained, where the accuracy rate denotes the ratio of the number of positive samples correctly predicted to the number of positive samples predicted as follows:

$$\text{precision} = \frac{TP}{TP + FP} \tag{7}$$

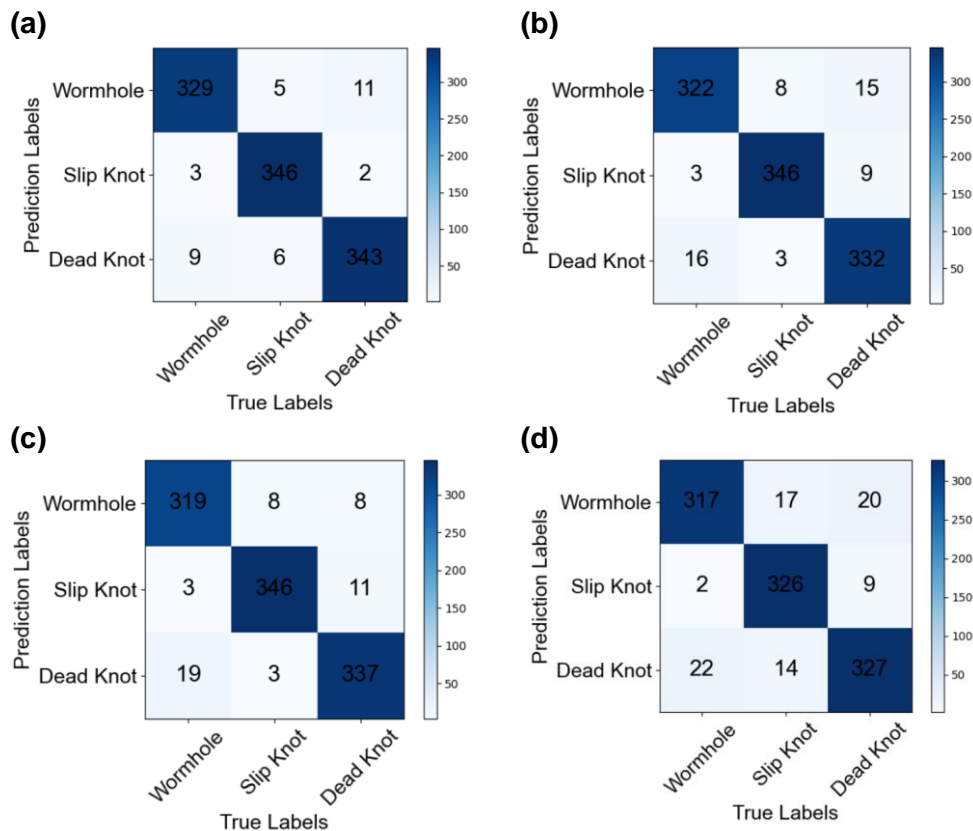
The recall rate refers to the ratio of the number of correctly predicted positive samples to the total number of real samples.

$$\text{recall} = \frac{TP}{TP + FN} \tag{8}$$

*Specificity*

As the specificity increases, the probability of false detections decreases.

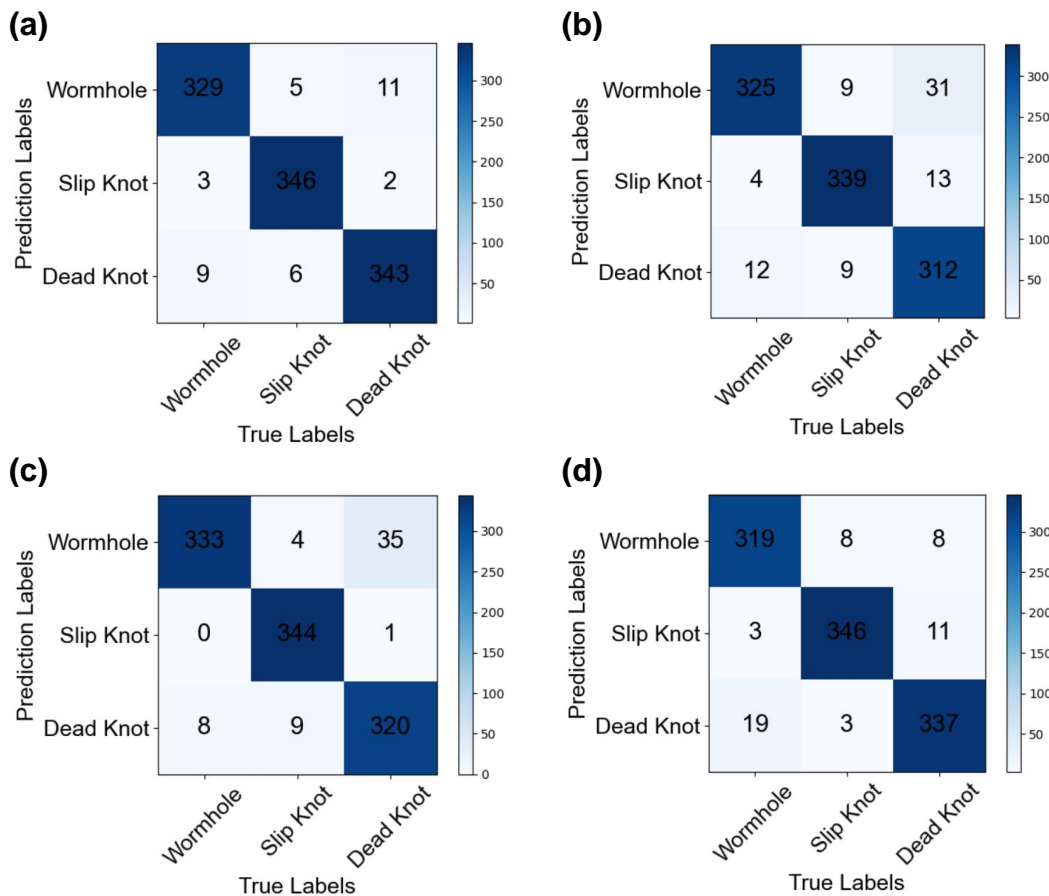
$$\text{specificity} = \frac{TN}{TN + FP} \tag{9}$$



**Fig. 11.** Networks confusion matrix (a) GhostRegNet + NAM network confusion matrix (b) RegNet network confusion matrix (c) RegNet+SE network confusion matrix (d) RegNet + ECA network confusion matrix

### Model Performance Analysis

The predictions of RegNet, RegNet+SE, RegNet+ECA, RegNet+NAM, GhostRegNet+NAM, MobileNet-v2, EfficientNet, and Vision-Transformer networks on the test set of wood defects are shown in the confusion matrix Figs. 11 and 12, respectively. The abscissa of the confusion matrix figure represents the real label of the sample and the ordinate represents the predicted label of the sample. By using the confusion matrix predicted by each model on the test set and color depth of the color block of the confusion matrix, it can be intuitively observed that the prediction effect of the GhostRegNet+NAM network is better, and there are a large number of positive samples predicted as positive by the network model.



**Fig. 12.** Networks confusion matrix (a) RegNet + NAM network confusion matrix (b) MobileNet-v2 network confusion matrix (c) EfficientNet network confusion matrix (d) Vision-Transformer network confusion matrix

The performances of four RegNet network improved schemes are compared in Table 2. The GhostRegNet+NAM network exhibited the highest precision, recall, and specificity values for the three types of wood defects and lower false detection probability. The samples of the wood defect test were set into the sample library of the wood defect and the optimal weight pre-training model is trained on the set to obtain the recognition accuracy, network parameter quantity, and average image recognition time of four RegNet network improved schemes. A comparison of the data is presented in Table 3. Compared with other improved schemes, the parameters of GhostRegNet+NAM network are

significantly reduced, the average image recognition time was significantly reduced, and the recognition accuracy was improved.

**Table 2.** Performance Comparison of RegNet network, RegNet+SE network, RegNet+ECA network, RegNet+NAM network, and GhostRegNet+NAM network

Model	Types of Defects	Precision	Recall	Specificity
RegNet	Wormhole	0.933	0.944	0.968
	Slip Knot	0.966	0.969	0.983
	Dead Knot	0.946	0.933	0.973
RegNet+SE	Wormhole	0.952	0.935	0.978
	Slip Knot	0.961	0.969	0.98
	Dead Knot	0.939	0.947	0.968
RegNet+ECA	Wormhole	0.895	0.930	0.948
	Slip Knot	0.967	0.913	0.984
	Dead Knot	0.901	0.919	0.948
RegNet+NAM	Wormhole	0.942	0.959	0.972
	Slip Knot	0.963	0.952	0.981
	Dead Knot	0.935	0.93	0.967
GhostRegNet+NAM	Wormhole	0.954	0.965	0.978
	Slip Knot	0.986	0.969	0.993
	Dead Knot	0.958	0.963	0.979

**Table 3.** Comparison of Accuracy, Network Parameters, and Average Image Recognition Time of Regnet Network, Regnet+SE Network, Regnet+ECA Network, Regnet+NAM Network, and Ghostregnet+NAM Network on the Test Set

Model	Accuracy	Parameter	Infer
RegNet	94.88%	2.32M	0.046s
RegNet+SE	95.07%	2.8M	0.047s
RegNet+ECA	92.03%	2.32M	0.043s
RegNet+NAM	94.69%	2.32M	0.036s
GhostRegNet+NAM	96.58%	1.86M	0.034s

**Table 4.** Performance Comparison of Mobilenet-V2 Network, Efficientnet Network, Vision-Transformer Network And Ghostregnet+NAM Network

Model	Types of Defects	Precision	Recall	Specificity
MobileNet-v2	Wormhole	0.89	0.953	0.944
	Slip Knot	0.952	0.95	0.976
	Dead Knot	0.937	0.876	0.97
EfficientNet	Wormhole	0.895	0.977	0.945
	Slip Knot	0.997	0.964	0.999
	Dead Knot	0.95	0.899	0.976
Vision-Transformer	Wormhole	0.944	0.935	0.973
	Slip Knot	0.977	0.969	0.989
	Dead Knot	0.934	0.949	0.966
GhostRegNet+NAM	Wormhole	0.954	0.965	0.978
	Slip Knot	0.986	0.969	0.993
	Dead Knot	0.958	0.963	0.979

To verify the performance of GhostRegNet+NAM network model, this study introduced the comparative analysis of MobileNet-v2, EfficientNet, and Vision-Transformer networks, as shown in Table 4.

According to Table 4, the recall rate, specificity, and false detection probability of GhostRegNet+NAM network were still higher than those of MobileNet-v2, EfficientNet, and Vision-Transformer networks. The recognition accuracy, network parameter quantity, and average image recognition time of the MobileNet-v2, EfficientNet, and Vision-Transformer networks are presented in Table 5.

**Table 5.** Comparison of Accuracy, Network Parameters, and Average Image Recognition Time of MobileNet-v2 Network, EfficientNet Network, Vision-Transformer Network and GhostRegNet+NAM Network on the Test Set

Model	Accuracy	Parameter	Infer
MobileNet-v2	92.60%	4.01M	0.070s
EfficientNet	94.59%	2.23M	0.081s
Vision-Transformer	95.16%	85.80M	0.081s
GhostRegNet+NAM	96.58%	1.86M	0.034s

Compared with the other two networks, the accuracy of GhostRegNet+NAM network was as high as 96.58%, the amount of network parameters was only 1.86 M, and the average image recognition time was 0.034 s. While reducing the model weight, the detection speed and accuracy are improved.

From Table 5, it can be seen that GhostRegNet+NAM network achieved higher accuracy and could accurately classify plate defect models. At the same time, the classification time was relatively short. Wood defect detection is often used for industrial optimizing saw for online sorting of woods. Modern industrial production lines have track feed speeds of up to 120m/min, and it is very important to improve the speed even at millisecond level. Meanwhile, the params of GhostRegNet+NAM network are minimal, which can alleviate computational pressure for computational power.

## CONCLUSIONS

1. The GhostRegNet+NAM network model can distinguish the types of wood defects, and the classification accuracy is high.
2. To verify the performance of the model, MobileNet-v2, EfficientNet, and Vision-Transformer networks are introduced for comparison. GhostRegNet+NAM model is still better than the comparison networks with less parameters.
3. The GhostRegNet+NAM network model in distinguishing the types of sample defects with high accuracy and fewer network parameters is determined. This can effectively reduce the detection time and reduce the hardware requirements of the algorithm, which is robust and practical.
4. A training overfitting phenomenon was encountered. By adding a regularization term in the process of calculating the loss function of the attention mechanism NAM, this problem can be effectively solved. However, during the learning process of the training set, the loss function fluctuated slightly.



## ACKNOWLEDGEMENTS

This study is funded by a program of Heilongjiang Province Postdoctoral Exit Research Initiation Fund Funded Project (2022/415858) and a program of the Natural Science Foundation of Heilongjiang Province (LH2020C034).

## Conflicts of Interest

The author declares that there is no competing interest.

## REFERENCES CITED

- Diao, Z. D., Zhang, D. Y., and Cao, J. (2012). "Research on wood sorting image segmentation based on genetic algorithms and mathematical morphology," *Automation Technology and Applications* 31(01), 59-62. DOI: 10.3969/j.issn.1003-7241.2012.01.016.
- Gong, X. Y., Chang, S. Y., Jiang, Y.F., and Wang, Z. Y. (2019). "AutoGAN: Neural architecture search for generative adversarial networks," DOI: 10.48550/arXiv.1908.03835
- Guo, K., Huang, Y., Yang, N., Tan, G., Dong, Q. Q., and Cheng, Y. Z. (2020). "Multichannel wood defect image segmentation algorithm based on TVCV model," *Forestry Machinery and Woodworking Equipment* 48, 22-26. DOI: 10.13279/j.cnki.fmwe.2020.0100
- Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., and Xu, C. (2020). "GhostNet: More features from cheap operations," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. DOI: 10.1109/CVPR42600.2020.00165
- Hu, Y. F., Belkhir, N., Angulo, J., Yao, A., and Franchi, G. (2022). "Learning deep morphological networks with neural architecture search," *Pattern Recognition* 131. DOI: 10.48550/arXiv.2106.07714
- Jin, H. F., Song, Q. Q., and Hu, X. (2019). "Auto-Keras: An efficient neural architecture search system," *Knowledge Discovery & Data Mining*. DOI: 10.48550/arXiv.1806.10282
- Li, G., Qian, G., Delgadillo, I. C., Müller, M., Thabet, A., and Ghanem, B. (2020). "SGAS: Sequential greedy architecture search," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. DOI: 10.1109/CVPR42600.2020.00169
- Ling, J. X., and Xie, Y. H. (2022). "Research on wood defects classification based on deep learning," *Wood Research* 67. DOI: 10.37763/WR.1336-4561/67.1.147156
- Liu, C. R. (2008). "Development status and prospects of China's wood industry," *Shaanxi Forestry* 141(03), 15-16. DOI: CNKI: SUN: SXNI.0.2008-03-012.
- Liu, Y., Shao, Z., and Teng, Y. (2021). "NAM: Normalization-based attention module," DOI: 10.48550/arXiv.2111.12419
- Luo, W., and Sun, L. P. (2019). "Support vector machine learning classification of wood defects using local binary patterns and directional gradient histogram fusion features," *Journal of Northeast Forestry University* 47(06), 70-73. DOI: 10.13759/j.cnki.dlxb.2019.06.014

- Radosavovic, I., Kosaraju, R. P., Girshick, R., He, K. M., and Dollár, P. (2020). "Designing network design spaces," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. DOI: 10.1109/CVPR42600.2020.01044
- Riana, D., Rahayu, S., Hasan, M., and Anton (2021). "Comparison of segmentation and identification of *Swietenia mahagoni* wood defects with augmentation images," *Heliyou* 7(6). DOI: 10.1016/j.heliyon.2021.e07417
- Shi, G.Y., Cao, J., and Zhang, Y. Z. (2020). "Research on near infrared recognition method of wood defects based on Transfer Learning," *Journal of Electrical Machinery and Control* 24. DOI: 10.15938/j.emc.2020.10.018
- Sun, Y.N., Xue, B., Zhang, M. J., and Yen Gary, G. (2020). "Evolving deep convolutional neural networks for image classification," *IEEE Transactions on Evolutionary Computation* 24. DOI: 10.1109/TEVC.2019.2916183
- Wang, N., Gao, Y., Chen, H., Wang, P., Tian, Z., Shen, C., and Zhang, Y. (2020). "NAS-FCOS: Fast neural architecture search for object detection," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI:10.1109/CVPR42600.2020.01196
- Xie, Y. H. (2014). "Research on the application of digital image processing technology in wood surface defect detection," *Northeast Forestry University*
- Yan, F., Zhang, J. H., Yao, Y.C., and Liu, J. (2022). "Improved UNET wood defect image segmentation method," *Forestry Machinery and Woodworking Equipment* 50, 41-45. DOI: 10.13279/j.cnki.fmwe.2022.0005.
- Zhang, S., Wang, Y. B., Yang, T., and Li, M. (2021). "Research on identification method of typical defects on wood surface based on improved LeNet-5 model," *Wood Science and Technology* 35, 31-37. DOI: 10.12326/j.2096-9694.2021024
- Zhang, K., Feng, X. H., and Guo, Y.R. (2021). "Overview of deep convolutional neural network models for image classification," *Chinese Journal of Image Graphics* 26 (10): 2305-2325. DOI: 10.11834/jig.200302.
- Zoph, B., and Le, Q. V. (2016). "Neural architecture search with reinforcement learning," *CoRR*. DOI: 10.48550/arXiv:1611.01578

Article submitted: March 30, 2023; Peer review completed: September 2, 2023; Revised version received: September 10, 2023; Accepted: September 22, 2023; Published: September 27, 2023.

DOI: 10.15376/biores.18.4.7663-7680