

RESEARCH

Open Access



# Accurate identification of cashmere and wool fibers based on enhanced ShuffleNetV2 and transfer learning

Yaolin Zhu<sup>1</sup>, Ran Liu<sup>1\*</sup>, Gang Hu<sup>2</sup>, Xin Chen<sup>1</sup> and Wenya Li<sup>3</sup>

\*Correspondence:  
xpu\_lr@163.com

<sup>1</sup> School of Electronics and Information, Xi'an Polytechnic University, Xi'an, China

<sup>2</sup> School of Science, Xi'an University of Technology, Xi'an, China

<sup>3</sup> School of Textile Science and Engineering, Xi'an Polytechnic University, Xi'an, China

## Abstract

Recognizing cashmere and wool fibers has been a challenging problem in the textile industry due to their similar morphological structure, chemical composition, and physicochemical properties. Traditional manual methods for identifying these fibers are inefficient, labor-intensive, and inaccurate. To address these issues, we present a novel method for recognizing cashmere and wool fibers using an improved version of ShuffleNetV2 and Transfer Learning, which we implement as a new cashmere and wool classification network (CWCNet). The approach leverages depthwise separable dilated convolution to extract more feature information for fiber classification. We also introduce a new activation function that enhances the nonlinear representation of the model and allows it to more fully extract negative feature information. Experimental results demonstrate that CWCNet achieves an accuracy rate of up to 98.438% on our self-built dataset, which is a 2.084% improvement over the original ShuffleNetV2 model. Furthermore, our proposed method outperforms classical models such as EfficientNetB0, MobileNetV2, Wide-ResNet50, and ShuffleNetV2 in terms of recognition accuracy while remaining lightweight. The method is capable of extracting more information on fiber characteristics and has the potential to replace manual labor as technological advancements continue to be made. This will greatly benefit engineering applications in the textile industry by providing more efficient and accurate fiber classification.

**Keywords:** Cashmere and wool, Transfer learning, Activation function, Data augmentation, Classification

## Introduction

Cashmere is finer than wool, with a smoother surface. The scales of cashmere are neatly arranged, resembling bamboo nodes. In contrast, wool has a rough surface and irregular arrangement of scales that form a spiral pattern. Cashmere and wool exhibit remarkable similarities in terms of their morphological structure, chemical composition, and physical and chemical properties. This similarity poses a challenge when it comes to distinguishing between cashmere and wool. Furthermore, the emergence of hybrid breeds of goats and sheep has further complicated matters, as it has led to variations in both cashmere and wool fibers. As a result, the difference between

these two fibers becomes smaller and more difficult to identify accurately. Moreover, compared to cashmere, wool's price is relatively low. However, cashmere is a luxurious and rare textile material. As the demand for cashmere increases with the rise in living standards, many dishonest traders use wool to counterfeit cashmere, disrupting the market for financial gain. Therefore, there is an urgent need for a fast and stable method to identify cashmere and wool fibers. Currently, cashmere and wool fiber identification is performed manually with the aid of a microscope, which is not only costly but also subjective. To address this issue, we propose a method to identify cashmere and wool fibers based on enhanced ShuffleNetV2 and Transfer Learning. By utilizing deep learning, we can perform efficient, reliable, and rapid identification of cashmere and wool fibers.

In recent decades, numerous domestic and international scholars have put forward various identification methods grounded on the distinctions between cashmere and wool fibers. These methods can be categorized into three main categories, namely, visual characteristics of fiber surfaces such as image processing methods, manual microscope observation methods [1], and deep learning methods, fiber composition-based methods such as near-infrared spectroscopy [2] and proteomics detection [3, 4], and genetic characteristics-based methods that include DNA analysis [5]. Currently, the research is mainly focused on two directions, namely, image processing methods and deep learning methods. These two research areas are considered the current hotspots in the field.

In 2015, Yuan et al. [6] presented a texture analysis-based discrimination method that utilized improved Tamura texture features to extract six texture parameters, including roughness, contrast, orientation, linearity, regularity, and roughness, from the final texture image. A BP neural network was used for classification achieving a recognition rate of 81.17%. In 2017, Lu et al. [7] proposed a fast identification method for similar fibers of wool and cashmere based on a visual bag-of-words model. The SIFT algorithm was utilized to extract local features from fiber morphology and generate visual words, while the SVM algorithm was used for fiber image classification and recognition based on visual words, with an average accuracy of 95.4%. In 2019, Xing et al. [8] calculated the counting box dimension and information dimension of fiber binary images using a fractal algorithm to derive fiber fineness. They also used a morphological and texture feature fusion strategy to classify the obtained morphological features using the K-means clustering algorithm, achieving an accuracy of 97.47%. In the same year, two preprocessing methods for fiber images [9] were adopted by proposing the morphological features (diameter) obtained using an interactive measurement algorithm and texture features using the Gray-level co-occurrence matrix (GLCM) [10] and other texture features, which were used for classification using the K-means clustering algorithm. An accuracy of 94.29% was finally achieved. In 2020, Zhu et al. [11] proposed an optimal parameter selection method based on the fusion of morphological and texture features. The fiber diameter and texture features were first extracted and fused, and the five dimensional feature vector that best characterized fiber information was selected. The fiber recognition accuracy was up to 96.7% using the Fisher classifier for classification. In the same year, a texture feature analysis method for cashmere and wool fibers was presented based on the Gray-level co-occurrence matrix and Gabor [12, 13] wavelet transform [14] to extract texture features in the frequency domain and transform domain.

The feature fusion strategy used the weight method and the Fisher classifier was utilized for classification, achieving an accuracy of 93.33%.

In 2017, Fei Wang et al. and colleagues [15] presented a method for discriminating between different fibers using convolutional networks and deep learning. They employed Alex-Net to integrate local and global features and used a sigmoid classifier for preliminary classification. The optimal weights of the network were recorded based on the validation results, which were then used to construct a new classification network. The new network was trained for 50 iterations and achieved an accuracy of 92.1%. In 2018, Wang et al. and colleagues [16] proposed a method for cashmere and wool fiber recognition based on the Fast RCNN model. They first used a sigmoid classifier to obtain an initial rough classification and model weights. They then extracted features using the Fast RCNN method and augmented the overall features using partial features. The network from the preliminary classification round was used for cashmere and wool image classification and achieved an accuracy of 95.2%.

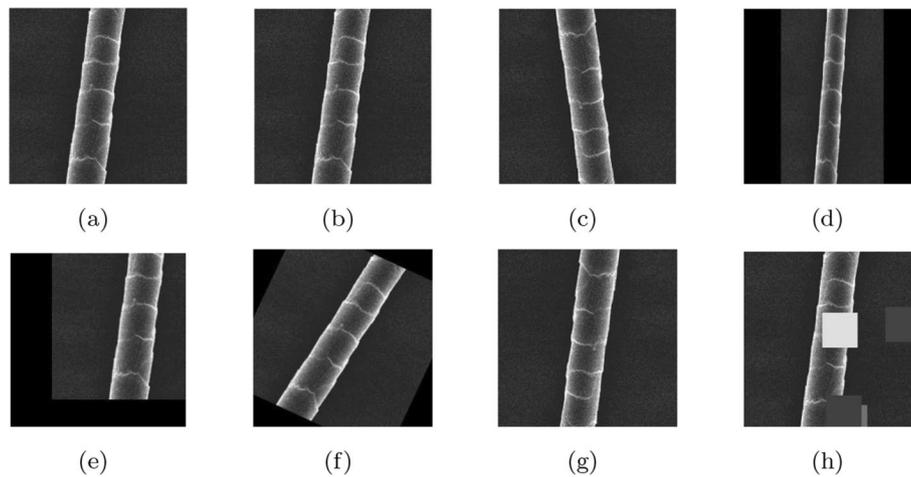
In 2021, Lu et al. [17] proposed a residual network-based method to identify cashmere and wool fibers. They utilized various parameter initialization methods, including Kaiming initialization [18], pre-training weights fine-tuning, and pre-training weights with frozen tuning for each layer except the fully connected layer, to compare their performances. Ultimately, their approach achieved an accuracy of 97.1%. In the same year, D. Agrawal et al. [19] presented a method for recognizing cashmere and wool fibers using an integrated model and transfer learning. Specifically, they built an integrated model based on ResNet50 and VGG16 and experimentally achieved an accuracy of 97.32% with a standard deviation of 0.89% and a training loss of 0.107.

Traditional image processing techniques typically use single features such as morphology, texture, and spectral lines for fiber identification. However, these methods cannot obtain complete fiber image information, resulting in low accuracy rates. The use of multi-feature fusion algorithms may lead to an increase in computation due to the increasing number of feature parameters and data dimensionality, leading to slower training speeds, larger errors, and lower recognition rates. The field of natural image classification and object recognition has seen significant advancements with the development of convolutional neural network algorithms, providing a new research method and ideas for the vision-based detection and classification of cashmere wool fibers. Therefore, in this paper, we propose a method for identifying cashmere and wool fibers based on enhanced ShuffleNetV2 and Transfer Learning, using deep learning methods to perform efficient and reliable fast cashmere and wool fiber identification. To solve the problems of feature information loss and inadequate feature extraction caused by the “Dead ReLU Problem” in cashmere and wool fiber recognition, ShuffleNetV2 is improved using an improved activation function, Depthwise Separable Dilated Convolution, and Transfer Learning to achieve high accuracy and a low parameters of cashmere and wool fiber recognition.

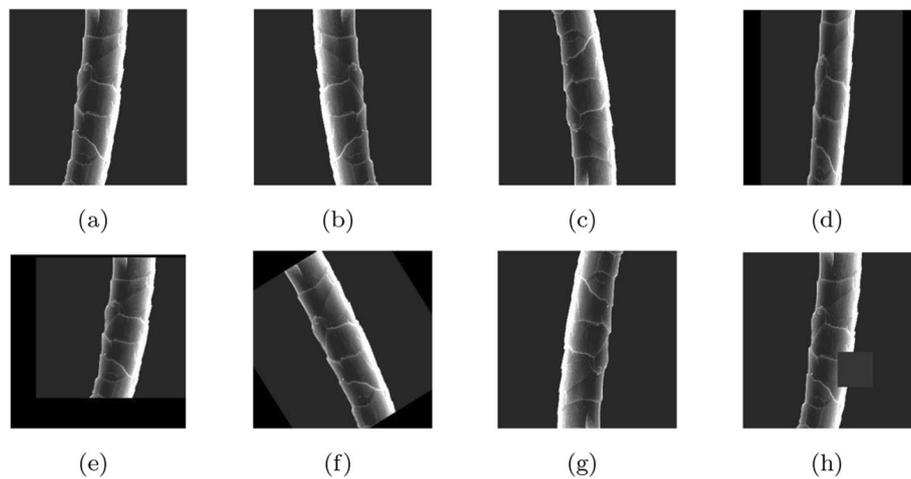
## **Materials and methods**

### **Dataset and experimental environment**

The dataset utilized in this study was obtained through scanning electron microscopy and consisted of 550 images of cashmere and 550 images of wool. To effectively



**Fig. 1** Original cashmere image and data enhancement image



**Fig. 2** Original wool images and data enhanced images

extract fiber features utilizing convolutional neural network methods, a substantial dataset is required. However, the initial dataset fell short of the necessary amount. To address this limitation, this paper implements data augmentation techniques to increase the dataset size, aiding in reducing overfitting and improving the model's generalization performance [20]. The effectiveness of the data augmentation is shown in Fig. 1 and Fig. 2: figure (a) is the original image; (b) is obtained by (a) horizontal flip with 50% probability of random horizontal flip; (c) is obtained by (a) vertical flip with 50% probability of random horizontal flip; (d) is unequal scaling of (a); (e) is (a) random erasure of (a) with 25% probability of erasure range is 2%–1/3; (f) is (a) 0–30° rotation of the original image; (g) is a horizontal flip followed by a vertical flip of (a); and (h) is random masking. By utilizing these methods, the final dataset contained 4,400 images of both cashmere and wool fiber.

**Table 1** GPU server parameters

Hardware and software	
Memory	24G
Processor	Intel(R) Xeon(R) Gold 6330 CPU @ 2.00GHz
Graphics	RTX 3090 (24GB)
Operating	Linux Ubuntu 18.04
Cuda	11.0
Pytorch	1.7.0

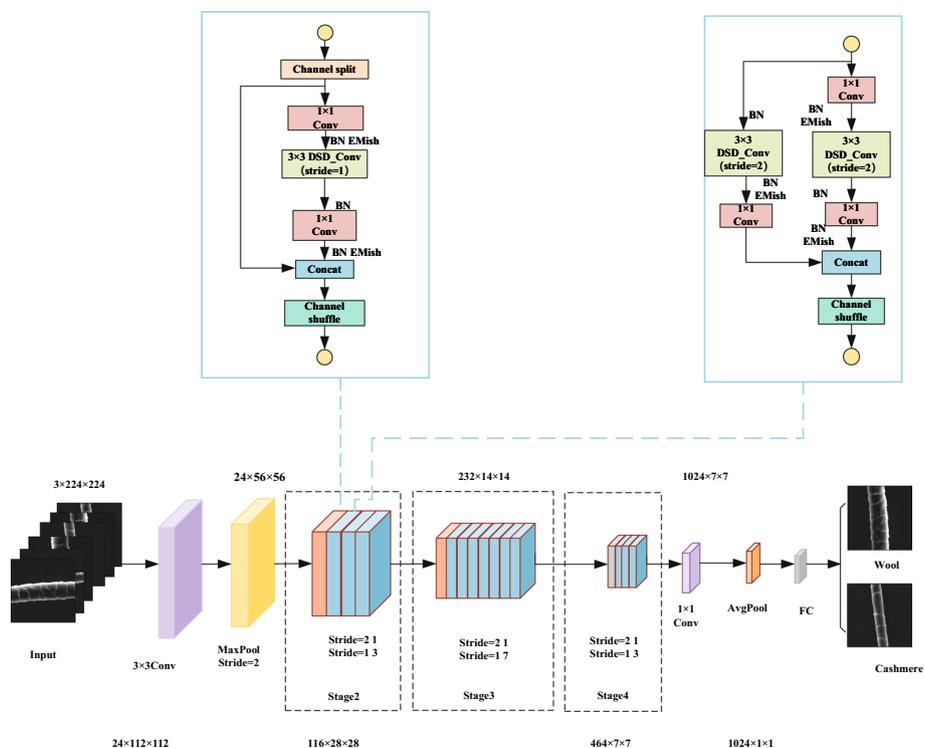
**Table 2** Parameters of ShuffleNetV2 model

Name	Parameters
Input	224x224
Learning Rate Update Strategy	poly
Slover	SGD
Learning rate	0.001
Loss	Cross Entropy Loss
Batch size	64
Epoch	200

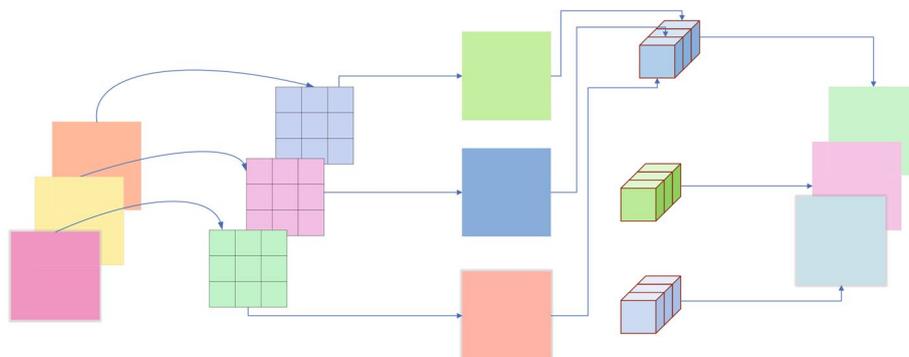
The PyTorch deep learning framework was utilized in the experiments, and all programs were executed on a GPU server (Ubuntu 18.04 system, 24 G RAM and RTX3090 graphics card). Jupyter notebook was employed as the development platform, and the specifications of the GPU server are presented in Table 1. The training algorithm hyperparameters used in this study include, but are not limited to, the learning rate, training time, optimizer, and batch size. The specific settings are illustrated in Table 2. With  $224 \times 224$  input image size, the initial learning rate and training epoch were configured to 0.001 and 200, respectively. The optimizer enables random gradient descent learning rate updates using SGD. To enhance model performance and accelerate convergence, the batch size was set to 64. The learning rate was dynamically adjusted via the learning rate update strategy. The Cross Entropy Loss function was adopted to optimize the model parameters in this study. Pre-training parameters were also incorporated into the model training procedure to expedite model fitting. Random seeds were fixed to ensure experiment reproducibility.

### ShuffleNet V2

ShuffleNetV2 [21], which was proposed by Ma et al. in 2018, is a convolutional neural network that strikes a good balance between speed and accuracy. It analyzes the factors influencing the inference speed of the model and proposes a more efficient basic block by taking two important metrics into account: memory access cost (MAC) and degree of parallelism (DP). By reducing the number of model parameters and computation while improving inference speed and detection accuracy, ShuffleNetV2 is able to achieve effective results. However, although it is a lightweight network with relatively low recognition accuracy, the approach proposed in this paper is an improvement on the original ShuffleNetV2. By extracting richer microscopic feature information while maintaining



**Fig. 3** Overall block diagram of the enhanced ShuffleNetV2



**Fig. 4** Depthwise separable convolution

its lightweight nature, the method achieves high-accuracy classification of cashmere and wool fibers. Figure 3 shows the overall structure of the enhanced ShuffleNetV2.

**Improvement methods**

**Depthwise separable dilated convolution (DSD\_Conv)**

Convolutional neural networks utilize convolutional operations for extracting regional features of an image layer by layer. This method continuously deepens the feature depth while narrowing the feature range, which enables the neural network to learn the image features more efficiently. The depthwise separable convolution (shown in Fig. 4) is composed of two parts - depthwise convolution (DW) and pointwise convolution (PW).

Depthwise convolution is different from normal convolution in that it convolves a channel of the feature map using only one convolution kernel and the number of convolution kernels equals to the number of channels. Therefore, the number of parameters and operation cost is greatly reduced. In depthwise convolution, each channel in the feature map is convolved using one convolution kernel while in pointwise convolution, a 1\*1 sized convolution kernel is used to mix feature information from different channels at the same spatial location.

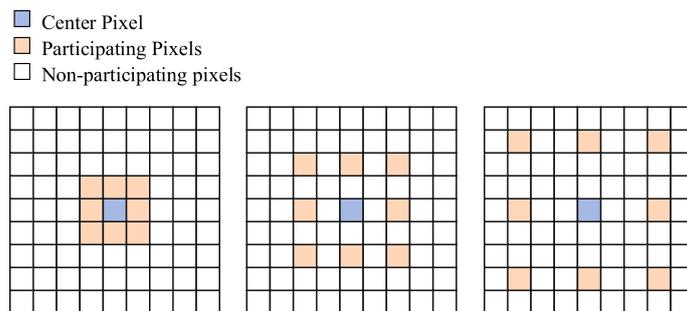
Depthwise separable convolution not only reduces computational complexity in a model but it also greatly reduces its size. Equation (1) outlines the comparison of the number of training parameters between ordinary convolution and depthwise separable convolution. This comparison assumes that the input feature map size is  $D_f \times D_f$  pixels, the number of input channels is  $M$ , the convolution kernel size is  $D_K \times D_K$  pixels, and the number of output channels is  $N$ . By analyzing the number of parameters, it becomes clear that as the number of output channels increases, depthwise separable convolution reduces the number of parameters more than ordinary convolution.

$$\frac{D_K \times D_K \times M \times D_f \times D_f + M \times N \times D_f \times D_f}{D_K \times D_K \times M \times N \times D_f \times D_f} = \frac{1}{N} + \frac{1}{D_K^2} \tag{1}$$

Dilated convolution [22] can help to expand the receptive field without relying on a pooling layer. This allows each convolutional output to contain a wider range of fiber feature information, resulting in better network performance. Basically, the dilation rate determines the number of intervals processed by the convolutional kernel in the convolutional layer.

The standard convolution, as shown on the left of Fig. 5, uses a dilation rate of 1. On the other hand, the 3x3 convolution, as shown on the right of Fig. 5, uses a dilation rate of 2 and has the same size as the standard convolution with a 5x5 Receptive Field. By embedding the dilation convolution in the depth-separable convolution, the receptive field of the convolution kernel is expanded without adding parameters. The depthwise separable dilated convolution effectively extracts deeper fiber features from the image, thereby improving the network’s fiber discrimination capabilities.

The use of depthwise separable convolution significantly reduces the number of parameters while maintaining accuracy. However, convolutional neural networks



**Fig. 5** Schematic diagram of different cavity rate receptive fields

sometimes lose spatial hierarchical information during training due to the small receptive field. Dilated convolution expands the receptive field, enabling each convolution output to contain a broader range of information. This paper employs depthwise separable dilated convolution (as shown in Fig. 6) to expand the receptive field and reduce computational effort. A larger receptive field contains richer image information and more complete fiber feature information, making it more favorable for feature extraction.

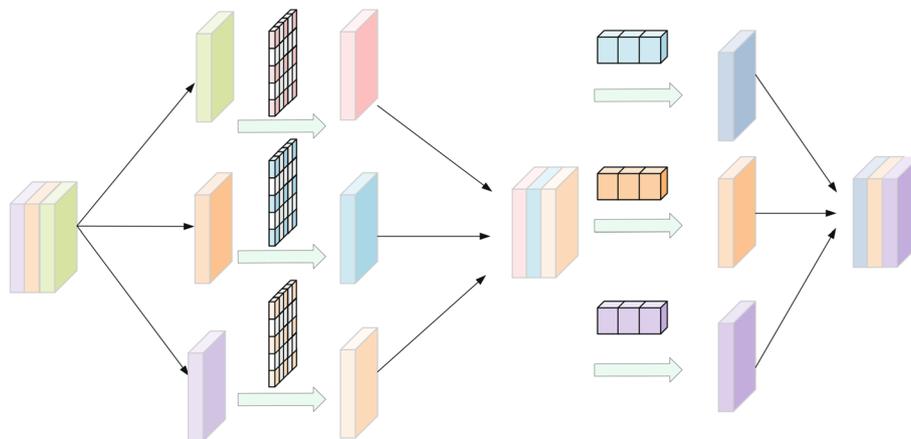
Equation (1) reveals that under identical circumstances, using a convolution kernel of 3×3, depth-separable cavity convolution is 8 to 9 times less computationally intensive than standard convolution. ShuffleNetV2 substantially reduces training time and the computational cost of parameter updates while maintaining image classification accuracy stability. However, ShuffleNetV2 suffers from negative feature information loss in the network layer’s activation function, and this paper improves upon the original activation function.

**EMish activation function**

The activation function [22] imbues the convolutional neural network with nonlinear factors, effectively enhancing the model’s expressive power and ultimately leading to improved classification results. ShuffleNetV2 uses the ReLU activation function [23], which offers rapid convergence and effectively mitigates gradient dispersion. However, this function has limitations such that, as the number of training rounds increases, the weights of some neurons remain unchanged and the neural network is unable to learn from negative value inputs, resulting in the discarding of negative feature information during feature extraction. This phenomenon is known as the ‘Dead ReLU Problem’ and weakens the non-linear representative capability of the convolutional neural network. The Mish activation function, on the other hand, is a self-regularizing non-monotonic function that overcomes these limitations and effectively extracts negative feature information. The function is expressed as:

$$Mish(x) = x \cdot \tan h(\ln(1 + e^x)) \tag{2}$$

Mish [24] has no upper bound, making it less likely for the model to suffer from gradient disappearance during the training process. This also accelerates the training process.



**Fig. 6** Depthwise separable dilated convolution

Furthermore, the lower bound characteristic of Mish assists in achieving a strong regularization effect. The property of having a lower bound enables the achievement of strong regularization effects. The Mish activation function possesses several advantages:

- (1) It is not completely truncated at negative values, thus ensuring better information inflow.
- (2) Its positive value is unbounded, with gradient tending to 1 in the left and right limits, thus avoiding gradient saturation [25].
- (3) The gradient descent of Mish is superior, ensuring the smoothing of each point to the maximum extent possible [26].

Compared to ReLU, Mish has stronger nonlinear characterization ability and is not subject to the ‘Dead ReLU Problem.’ Furthermore, Mish outperforms ReLU in terms of classification accuracy [24], which is why it is used as the activation function in this model. When  $x$  is greater than 0, Mish maps  $x$  to a new value space without discarding the original data. When  $x$  is less than 0, the exponential function  $e^x$  in the activation function becomes smaller and smaller as  $x$  tends to negative infinity, ultimately resulting in the activation function selectively losing some negative data [27]. However, although the Mish activation function improves neural network accuracy and stability, its ability to fit different network models and data distributions is limited due to the loss of some negative data. To better fit the cashmere and wool fiber identification problem, this paper proposes an improved version of Mish, EMish, as shown in Equation (3). EMish dynamically adjusts the activation function’s saturation region within a certain range by introducing parameters, which alleviates the loss of negative data, thus optimizing Convolutional Neural Networks and enhancing performance.

$$EMish(x) = (x + \alpha) \cdot \tan h(\ln(1 + e^{x+\alpha}) - \beta) \quad (3)$$

In determining the values of the parameters, an experimental approach is employed. Firstly,  $\beta$  is determined by incrementing from 0 to 1 in increments of 0.1. The value of  $\beta$  is selected based on the accuracy rate derived from multiple experimental results, ensuring that the activation function intersects the point (0, 0). Next,  $\alpha$  is determined by using recognition accuracy as a measure; Table 3 reveals that the EMish activation function

**Table 3** Experimental results of Mish activation function under different parameters

	$\alpha$	$\beta$	Accuracy(%)	Train time(s)
①	0.154	0.1	96.364	556.5
②	0.289	0.2	95.312	575.2
③	0.413	0.3	95.833	595.3
④	0.527	0.4	96.875	585.5
⑤	0.636	0.5	95.833	580.2
⑥	0.739	0.6	95.312	617.2
⑦	0.840	0.7	94.792	560.8
⑧	0.937	0.8	96.354	562.7
⑨	1.033	0.9	96.354	563.4
⑩	1.127	1.0	95.312	562.8

has the highest recognition accuracy at  $\beta$  and  $\alpha$  of 0.4 and 0.527, respectively. The final expression for EMish is shown in Equation (4).

$$EMish(x) = (x + 0.527) \cdot \tan h(\ln(1 + e^{x+0.527})) - 0.4 \quad (4)$$

Fig. 7 illustrates a comparison between the activation function before and after improvement. Specifically, Fig. 7a displays the original ReLU activation function, Fig. 7b showcases the Mish activation function, and Fig. 7c demonstrates the proposed EMish activation function. Upon comparing these three activation functions, it is discernible that the EMish function enables negative values that would typically be discarded to continue to be passed down. This, in turn, ameliorates the negative impact on the neural network resulted from discarding negative values.

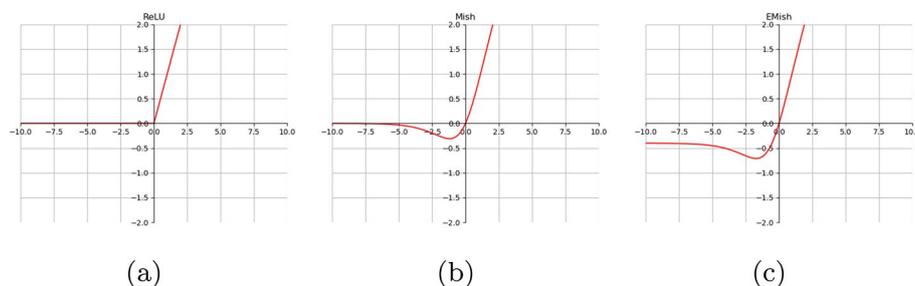
### Transfer learning

Transfer learning applies the knowledge acquired from image classification on large datasets (such as ImageNet) to a new classification task [28], This method offers the following advantages over direct model training:

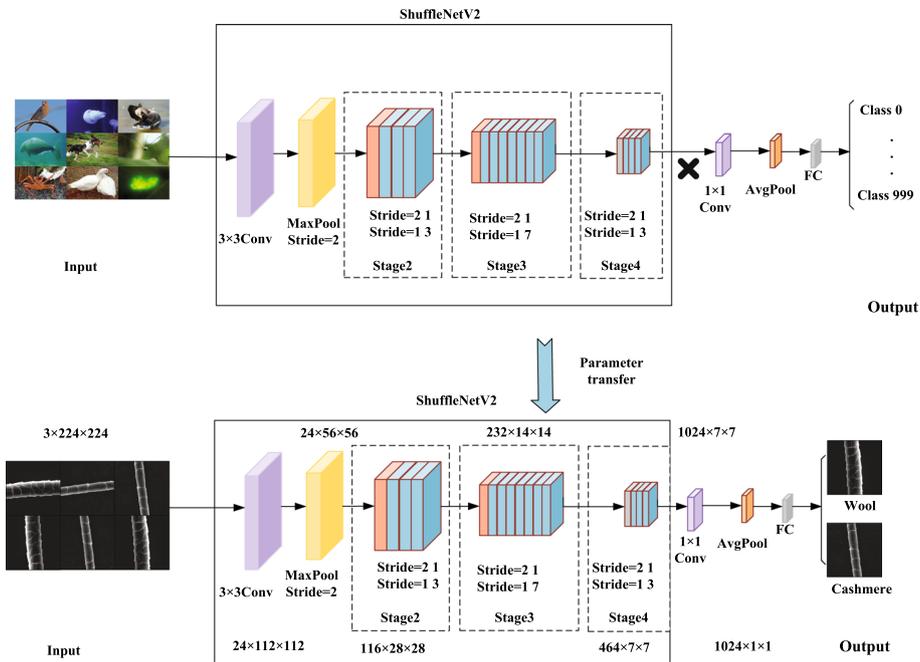
- (1) Transfer Learning can address the issue of insufficient Deep Learning samples by fine-tuning the pre-training weights to retrain on new datasets. This is far simpler than training the model from scratch.
- (2) Pre-trained network models can significantly reduce the training time as they have already learned rich features and do not require extensive data to be retrained.

Nowadays, there are two commonly used approaches for Transfer Learning: one directly uses the pre-trained weights as a solution for a new classification task [28], while the other fine-tunes the network weights by only training the weights closer to the output and freezing the remaining layers [29].

This paper uses ShuffleNetV2 as a pre-trained model for Transfer Learning, employing model structure optimization and parameter fine-tuning. Transfer Learning is performed using pre-training parameters from the original ShuffleNetV2, which was trained on the ImageNet dataset [30] and can extract fiber image features (as depicted in Fig. 8).



**Fig. 7** Comparison of activation functions



**Fig. 8** Schematic diagram of Transfer Learning

## Experiments and analysis of results

### Evaluation criteria

In order to assess the model's performance, various metrics were utilized, including Accuracy, Precision, Recall, and F1-score, which were computed utilizing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), as specified in equations (5) through (8). Furthermore, multiple dimensions were used to gauge the model's efficacy, including the number of model parameters (Params), training time, ROC curves, and PR curves.

Accuracy(Acc) indicates the proportion of correctly predicted samples to the total samples, as shown in Equation (5).

$$Accuracy = \left( \frac{TP + TN}{TP + TN + FP + FN} \right) \times 100\% \quad (5)$$

Precision(Pre) indicates the proportion of samples predicted as positive samples species predicted correctly to the total samples, as shown in Equation (6).

$$Precision = \left( \frac{TP}{TP + FP} \right) \times 100\% \quad (6)$$

Recall(Rec) represents the proportion of all positive samples that are predicted to be positive, as shown in Equation (7).

$$Recall = \left( \frac{TP}{TP + FN} \right) \times 100\% \quad (7)$$

The F1\_Score is the summed average of Pre and Rec, as shown Equation (8).

**Table 4** Experimental results of different cavitation rates

Dilated rate	1	2	3
Accyrcy(%)	95.312	95.833	94.271
Precision(%)	95.153	95.893	94.821
Recall(%)	95.288	95.588	93.603
F1-sorce(%)	95.218	95.729	94.074
Increase(%)	0	+0.521	-1.042
Train time(s)	539.2	552.8	564

**Table 5** Experimental results using different activation functions in ShuffleNetV2

Activation Function	Accyrcy(%)	Increase(%)	Train time(s)
ReLU	95.312	0	539.2
Mish	96.354	+1.042	574.9
EMish	96.875	+1.563	595.5

$$F1\_Score = 2 \left( \frac{Pre \cdot Rec}{Pre + Rec} \right) \times 100\% \quad (8)$$

#### Performance evaluation of ordinary convolution and depthwise separable dilated convolution

Dilated convolution increases the receptive field without using pooling, allowing each convolution output to contain a larger range of information without sacrificing feature spatial resolution, obtaining richer image features and enhancing the ability of the network to identify fibers.

In order to examine the impact of different dilated rates on the accuracy of cashmere wool fiber identification, this experiment compared the performance of dilated convolutions with varying rates in the ShuffleNetV2 model. Table 4 illustrates the results of the experiment, indicating that the highest recognition accuracy was attained using a dilated convolution with a rate of 2, increasing accuracy by 0.521%. Alternatively, a dilated convolution with a rate of 3 yielded diminished recognition accuracy due to an excessive dilated rate, resulting in an inadequate number of sampling points per unit range and an associated loss of correlation characteristics. Ultimately, the ShuffleNetV2 model incorporating a dilated convolution of rate 2 delivered superior recognition accuracy.

#### Enhanced activation function performance evaluation

The activation function is crucial in providing nonlinearity to the convolutional neural network, thereby increasing the model's expressiveness. This study examines the impact of activation functions on the accuracy of cashmere wool fiber recognition. Specifically, we compare the performance of ReLU, Mish, and EMish activation functions on the ShuffleNetV2 model. From Table 5 presents the comparison results of the different activation functions. Notably, Mish activation function outperforms ReLU activation function by 1.042 in terms of accuracy, while EMish improves accuracy by 1.563

**Table 6** Experimental results of the improved activation function in other models

Model	Activation Function		
	ReLU	Mish	EMish
MobileNetV3	94.792%	95.312%	95.833%
Increase	0	+0.52	+1.041
ResNet50	80.208%	85.417%	87.500%
Increase	0	+5.209	+7.292
DenseNet121	83.333%	85.938%	89.062%
Increase	0	+2.605	+5.729
VGG19	75.000%	80.208%	81.771%
Increase	0	+5.208	+6.771

**Table 7** Comparison of time cost of activation functions

Model	Activation Function		
	ReLU	Mish	EMish
MobileNetV3	554.4s	563.0s	592s
ResNet50	1021.8s	1150.2s	1239.7s
DenseNet121	1059.6s	1318.7s	1425.1s
VGG19	1735.4s	2019.1s	2087.6s

percentage points. These findings highlight the superiority of EMish activation function, which enables better transmission of negative data and effectively addresses the ‘Dead ReLU Problem’, ultimately leading to faster model convergence and improved classification accuracy.

To validate the efficacy of the suggested activation function, we employed it in several other deep learning models. As shown in Table 6, we compared the performance of distinct activation functions in networks such as MobileNetV3, ResNet50, DenseNet121, and VGG19. Our findings suggest that the ReLU activation function delivers superior results in the MobileNetV3 network model, with an accuracy rate of 94.792%.

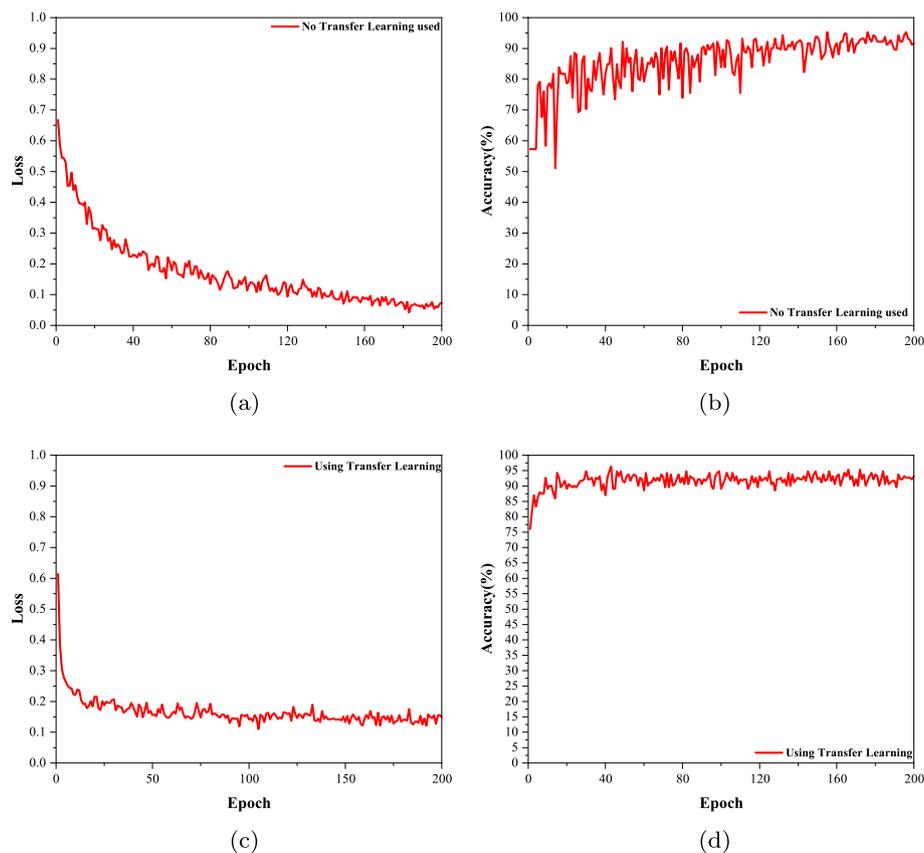
Among the four networks, the EMish activation function was found to outperform the other activation functions across the board. In fact, with ResNet50, it achieved the highest recognition accuracy on the validation set. The EMish function showed a remarkable improvement of 7.292% over the ReLU activation function. Table 6 indicates that the accuracy of the ReLU function on the MobileNetV3 network is slightly higher and in close proximity to that of using the Mish and EMish functions. However, the accuracy of both the Mish and EMish functions on the other network models was observed to be stronger than that of the ReLU function. Furthermore, the EMish function’s accuracy surpassed that of the other baseline activation functions across the board. It is evident that the proposed improvement method in this paper is practical and highly effective.

Table 7 compares the time costs of EMish function with other activation functions in four network models using the data-enhanced dataset. For the purpose of comparison, the experimental results are shown to one decimal place. From the table, it can be observed that ReLU function has the lowest time cost since it only retains positive data. On the other hand, the time cost for EMish function is relatively higher compared to

other activation functions. This increase in time cost is attributed to the corresponding increase in complexity in the data distribution and the number of neural network layers for the EMish function. When dealing with complex data distributions and neural networks with many layers, relatively more time will be required. However, the increase in time cost remains within an acceptable range when considering factors such as the magnitude of the increase in time and the accuracy rate performance indexes.

### Model performance evaluation with or without Transfer Learning

In order to examine the impact of transfer learning on the recognition of cashmere and wool fibers, we compared the original ShuffleNetV2 with ShuffleNetV2 that utilized pre-trained weights (TL-ShuffleNetV2). Pre-trained weights allowed for better training parameters at the onset of training, thus enhancing the model's performance. Figure 9 displays the loss and accuracy curves of the ShuffleNetV2 model with transfer learning and the original ShuffleNetV2 model. We observed that transfer learning significantly accelerates the network's convergence. During initial training, the TL-ShuffleNetV2 model achieved lower loss values and higher accuracy, and rapidly reached a smooth accuracy around the 40th epoch. By contrast, the original ShuffleNetV2 model had a slower convergence of loss and accuracy curves, with the model's accuracy peaking only at around the 150th epoch. Our experimental outcomes



**Fig. 9** Comparison of validation accuracy and training loss with and without Transfer Learning

demonstrate that TL-ShuffleNetV2 quickens network convergence and reduces training time. As the model has acquired relevant generic features from its original task, using transfer learning for cashmere and wool fiber identification involves only incremental learning without overfitting to new data, thus improving the model's accuracy and generalization ability.

#### Network improvement ablation experiments

To evaluate the impact of each improvement point on network performance, this study integrates them into ShuffleNetV2 and conducts ablation experiments. The experiments include the implementation of Depthwise Separable Dilated convolution(DSD\_Conv), Transfer learning, and EMish activation function. Table 8 displays the training results comparison. It shows that despite being a lightweight network, ShuffleNetV2 can still achieve 95.312% recognition accuracy. After implementing DSD Conv, the evaluation indexes remain relatively stable, while the Receptive Field of the convolutional kernel is expanded, facilitating the extraction of detailed features that aid in cashmere and wool fiber classification, and the accuracy improves by 0.521%. The use of EMish mitigates the adverse effect of the ReLU activation function dropping negative values on the neural network, thus optimizing the deep neural network and improving its performance, with a model accuracy improvement of 1.563%. Furthermore, the implementation of Transfer learning reduces training time and enhances the model's generalization ability, preventing it from overfitting to new data, and improving the accuracy by 1.042%. Ultimately, with DSD\_Conv, Transfer learning, and EMish activation functions integrated, CWCNet achieves 97.971% accuracy on cashmere and wool fiber datasets, with a 2.084% accuracy improvement over the benchmark network.

#### Scaled performance evaluation for different training and test sets

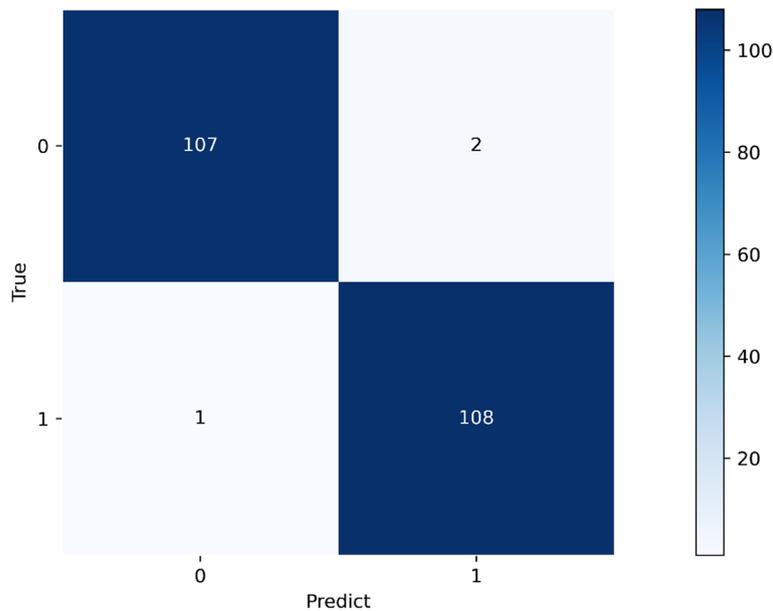
Table 9 presents a comparison of the experimental results for the data-enhanced dataset using different training and test sets. Among them, the accuracy rate reaches 98.438% with a relatively short training time when the ratio of training set to validation set is set at 9:1. These results indicate that the best recognition performance is obtained when the ratio of training set to test set is set at 9:1.

**Table 8** Results of ablation experiments

Model	Evaluation Indicators			Accuracy(%)	Increase(%)	Train time(s)
	Factor					
	DSD_Conv	EMish	Transfer Learning			
①	×	×	×	95.312	0	539.2
②	✓	×	×	95.833	+0.521	580.0
③	×	✓	×	96.875	+1.563	585.5
④	×	×	✓	96.354	+1.042	546.0
⑤	✓	✓	✓	97.971	+2.084	587.1

**Table 9** Experimental results of different training and test sets

train:test	Accyrcy(%)	Precision(%)	Recall(%)	F1-sorce(%)	Train time(s)
5:5	95.703	95.821	95.563	95.668	622.7
6:4	97.656	97.641	97.566	97.603	607.5
7:3	96.875	96.867	96.911	96.874	582.2
8:2	97.917	97.871	97.871	97.871	587.1
9:1	98.438	95.000	99.090	96.909	624.6



**Fig. 10** Heat map of the confusion matrix

**Improving model performance evaluation**

To analyze the misidentification among categories, this study employed the confusion matrix to illustrate the susceptibility to confusion among different categories. A total of 218 images from the test set were classified and identified by our trained model, with the resulting confusion matrix heat map presented in Fig. 10. The label for cashmere is designated as 0, while wool is labelled as 1. Figure 10 illustrates that the fibre recognition model performs optimally for cashmere; however, a few images are misclassified as wool. Conversely, the model functions slightly less effectively for wool, with two images being wrongly classified as cashmere. By combining the confusion matrix with the original dataset images, Fig. 11 shows that coarsening of cashmere fibers and the existence of superfine wool are the primary sources of misclassification of images as belonging to alternative categories. Overall, the majority of test data exists along the diagonal of the confusion matrix, indicating that CWCNet is proficient at categorizing most images accurately. Therefore, although not all images are classified precisely, misclassified images remain in the minority.

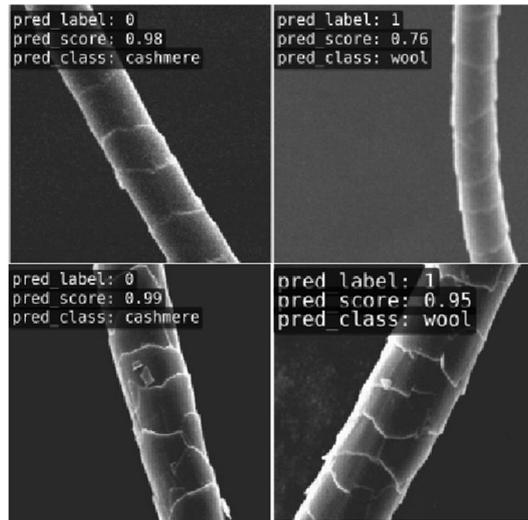


Fig. 11 Single sample test results

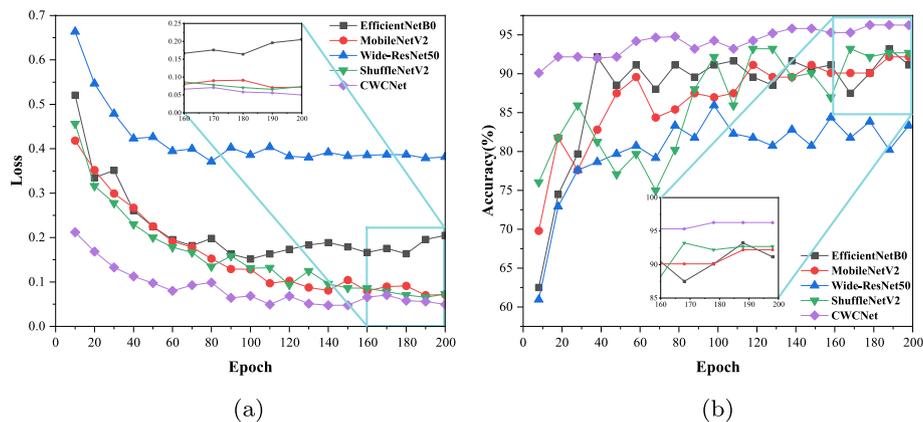


Fig. 12 Comparison of loss and accuracy of different models

**Performance comparison of the improved model with other models**

To verify the effectiveness of the improved ShuffleNetV2 model, we compared it to several classical convolutional neural networks, namely EfficientNetB0, MobileNetV2, Wide-ResNet50, and ShuffleNetV2. We recorded the loss on the training set and the accuracy on the test set during the training process to observe the training of the models and ensure that each model completes training with convergence. The numerical comparison of the final training results is shown in Table 10. From Fig. 12, we observed that both EfficientNetB0 and Wide-ResNet50 models tended to be unstable in the first 40 epochs. The MobileNetV2 model fluctuated for the first 68 epochs, with its accuracy fluctuating around 80%, and gradually stabilized at 90% after the 69th epoch. When the training reached the 11th epoch, the CWCNet model stabilized, and the accuracy reached 95%. Through experiments and analysis, we found that the accuracy of the improved ShuffleNetV2 network model is at least 3% better than other commonly used models. As shown in Fig. 13, CWCNet outperforms the

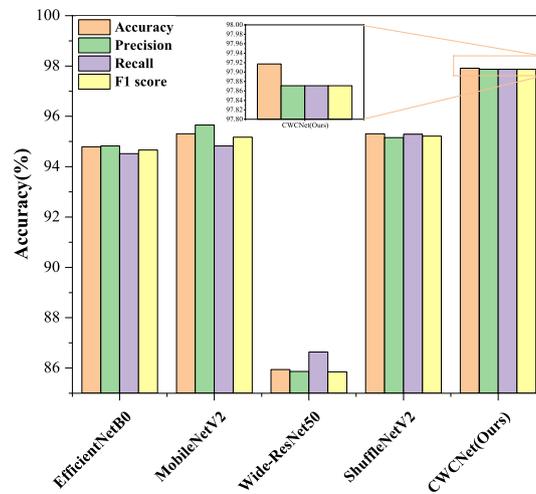


Fig. 13 Comparison of each evaluation index of different models

Table 10 Experimental results of different models

Models	EfficientNetB0	MobileNetV2	Wide-ResNet50	ShuffleNetV2	CWCNet
Accyrcy(%)	94.792	95.312	85.938	95.312	98.438
Precision(%)	94.821	95.658	85.865	95.153	95.000
Recall(%)	94.523	94.823	86.640	95.288	99.090
Params	4.01M	2.23M	66.84M	2.51M	2.51M
FLOPs	0.02G	0.32G	11.44G	0.15G	0.15G
Train time(s)	914.6	614.9	1280.3	539.2	624.6

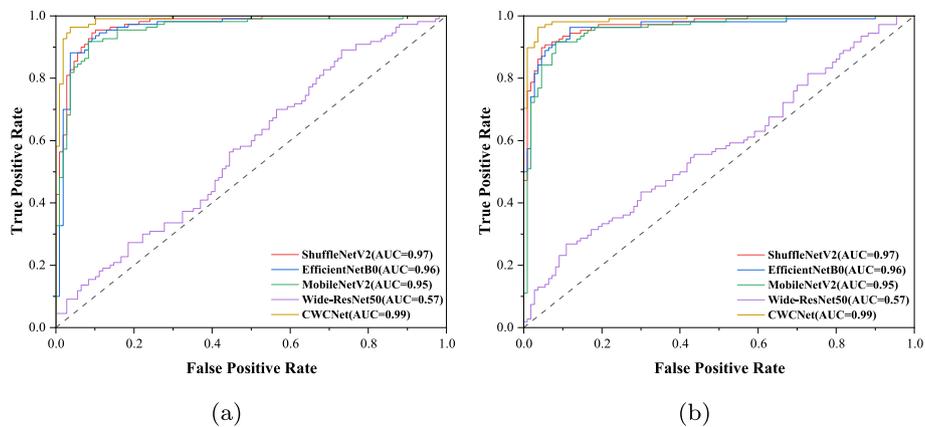
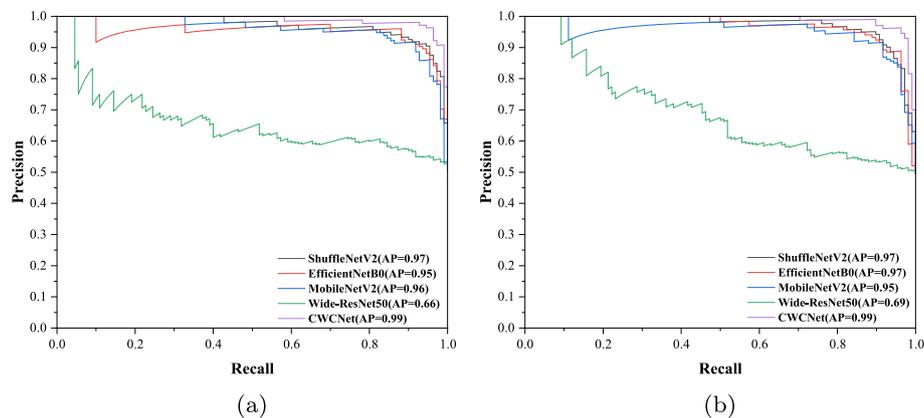
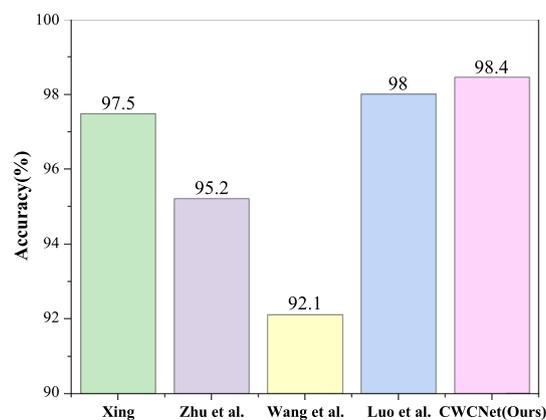


Fig. 14 Comparison of model's ROC curves

classical model across the board in evaluation metrics such as Accuracy, Precision, Recall, and F1 score. From the training curves and evaluation metric comparison plots, it is evident that CWCNet has a higher accuracy in the cashmere and wool classification tasks than all comparison models, demonstrating the superior performance of CWCNet over the other models.



**Fig. 15** Comparison of model's PR curves



**Fig. 16** Comparison of other researchers' fiber identification methods

To further verify the classification performance of the CWCNet model, we compared the ROC and PR curves of the test set classification model. The experimental results are demonstrated in Fig. 14 and Fig. 15. Figure 14 depicts the ROC curves of the CWCNet fiber classification model. By leveraging the enhanced dataset, the area under the ROC curve (AUC) is a metric used to judge the classification effectiveness. The current AUC results for classification using different network models are 0.97, 0.96, 0.95, 0.57, and 0.99. On the other hand, Fig. 15 demonstrates the PR curve of the CWCNet fiber classification model, which indicates the relationship between classifier accuracy and recall. The larger the area under the PR curve line (AP), the better the model classification. These results indicate that the CWCNet fiber classification model proposed in this study outperforms other models.

After comparing the method proposed in this paper with existing studies, as shown in Fig. 16, it was found that Xing et al. achieved a recognition accuracy of up to 97.47% for cashmere and wool fibers using a combination of feature extraction and machine learning. Zhu et al. investigated the classification of these fibers using the maximum inter-class difference and achieved a recognition accuracy of 95.20%. Wang et al. improved the Alexnet model to achieve a recognition accuracy of 92.10%. Luo

et al. used a residual network model to achieve a recognition accuracy of 98% for cashmere and wool fiber recognition. However, compared to these methods, the proposed method in this paper achieved a maximum accuracy of 98.438%, which clearly indicates its effectiveness. Therefore, it was concluded that our proposed method outperforms other researchers' algorithms.

### Conclusion and future work

The current methods for classifying cashmere and wool fibers rely heavily on manual visual identification, which is susceptible to subjective factors. Applying artificial intelligence techniques to this task can improve efficiency and reduce labor costs. In this paper, we propose a model called CWCNet for cashmere and wool fiber classification, which avoids the need for manual feature extraction using traditional methods. CWCNet is an improved model based on ShuffleNetV2, optimized using Depthwise separable dilated convolution and a new activation function (EMish), and trained with Transfer Learning. Experimental results show that the Depthwise separable dilated convolution expands the Receptive Field, enhances the feature information extraction ability of ShuffleNetV2, and improves classification accuracy. The EMish activation function effectively improves the performance of deep neural networks, showing strong robustness and high stability, improving the classification accuracy of the model, simply and effectively alleviating the phenomenon of 'Dead ReLU Problem', and accelerating the convergence speed of the model. The Transfer Learning method can better initialize the training, accelerate the convergence of the model network, and improve the model's performance in a shorter period of time. Under the same experimental conditions, the model in the paper outperforms the EfficientNetB0, MobileNetV2, Wide-ResNet50, and ShuffleNetV2 models for classification. However, this paper has some limitations. Most of the cashmere and wool fiber datasets used so far consist of images of single fibers with simple backgrounds. In real environments, images with complex backgrounds and possibly containing multiple fibers are often obtained. Thus, future research plans to collect more images of cashmere and wool fibers in real environments, expand the fiber image dataset, further optimize the model, improve its performance and robustness, and establish an end-to-end cashmere and wool fiber classification model to improve its practical value.

### Abbreviations

CWCNet	Cashmere and wool classification network
DSD_Conv	Depthwise separable dilated convolution
EMish	Enhanced Mish
GLCM	Gray-level co-occurrence matrix
DW	Depthwise convolution
PW	Pointwise convolution
MAC	Memory access cost
DP	Degree of parallelism
TL-ShuffleNetV2	Transfer Learning ShuffleNetV2

### Acknowledgements

Not applicable.

### Author Contributions

Conceptualization, YZ, RL and GH; investigation, YZ and WL; methodology, YZ; acquisition of data, WL; software, RL; supervision, YZ and GH; validation, RL, XC, and WL; visualization; writing-original draft, YZ, RL, and GH; writing-review and editing, YZ, RL, XC, WL, and GH. All authors have read and agreed to the published version of the manuscript. All authors read and approved the final manuscript.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the natural science basic research key program funded by Shaanxi Provincial Science and Technology Department (No. 2022JZ-35 and No. 2023-JC-ZD-33), the key research program industrial textiles Collaborative Innovation Center Project of Shaanxi Provincial Department of Education (No. 20JY026), Science and Technology plan project of Yulin City (No. CXY-2020-052) and Science and Technology plan project of Xi'an City (No. 23DCYJSGG0008-2023).

**Availability of data and materials**

The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 13 May 2023 Accepted: 21 September 2023

Published online: 06 October 2023

**References**

1. Tonetti C, Varesano A, Vineis C, Mazzuchetti G. Differential scanning calorimetry for the identification of animal hair fibres. *J Ther Anal Calori*. 2015;119:1445–51.
2. Zhou J, Yu L, Ding Q, Wang R. Textile fiber identification using near-infrared spectroscopy and pattern recognition. *Autex Res J*. 2019;19(2):201–9.
3. Zhou J, Wang R, Wu X, Xu B. Fiber-content measurement of wool-cashmere blends using near-infrared spectroscopy. *Appl Spectrosc*. 2017;71(10):2367–76.
4. Geng R-Q. Species-specific pcr for the identification of goat cashmere and sheep wool. *Mol Cell Probes*. 2015;29(1):39–42.
5. Hamlyn P, Nelson G, McCarthy B. Wool-fibre identification by means of novel species-specific dna probes. *J Text Inst*. 1992;83(1):97–103.
6. Yuan SL, Lu K, Zhong YQ. Identification of wool and cashmere based on texture analysis. *Key Eng Mat*. 2016;671:385–90.
7. Lu K, Zhong Y, Li D, Chai X, Xie H, Yu Z, Naveed T. Cashmere/wool identification based on bag-of-words and spatial pyramid match. *Text Res J*. 2018;88(21):2435–44.
8. Xing W, Deng N, Xin B, Wang Y, Chen Y, Zhang Z. An image-based method for the automatic recognition of cashmere and wool fibers. *Measurement*. 2019;141:102–12.
9. Xing W, Xin B, Deng N, Chen Y, Zhang Z. A novel digital analysis method for measuring and identifying of wool and cashmere fibers. *Measurement*. 2019;132:11–21.
10. Li F, Yuan L, Zhang K, Li W. A defect detection method for unpatterned fabric based on multidirectional binary patterns and the gray-level co-occurrence matrix. *Text Res J*. 2020;90(7–8):776–96.
11. Zhu Y, Huang J, Wu T, Ren X. An identification method of cashmere and wool by the two features fusion. *Int J Cloth Sci Technol*. 2022;34(1):13–20.
12. Zhang Q, Li H, Li M, Ding L. Feature extraction of face image based on lbp and 2-d gabor wavelet transform. *Math Biosci Eng*. 2020;17(2):1578–92.
13. Lohithashva B, Aradhya VM, Guru D. Violent video event detection based on integrated lbp and glcm texture features. *Rev d'Intelligence Artif*. 2020;34(2):179–87.
14. Zhu Y, Huang J, Wu T, Ren X. Identification method of cashmere and wool based on texture features of glcm and gabor. *J Eng Fibers Fabr*. 2021;16:1558925021989179.
15. Xing W, Liu Y, Xin B, Zang L, Deng N. The application of deep and transfer learning for identifying cashmere and wool fibers. *J Nat Fibers*. 2022;19(1):88–104.
16. Wang F, Jin X. The application of mixed-level model in convolutional neural networks for cashmere and wool identification. *Int J Clothing Sci Technol*. 2018;30(5):710–25.
17. Luo J, Lu K, Chen Y, Zhang B. Automatic identification of cashmere and wool fibers based on microscopic visual features and residual network model. *Micron*. 2021;143: 103023.
18. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2015;1026–1034.
19. Agrawal D, Minocha S, Namasudra S, Kumar S. Ensemble algorithm using transfer learning for sheep breed classification. In: *2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, 2021;199–204. IEEE.
20. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):1–48.
21. 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)*, 2018;116–131.

22. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv*. 2015. <https://doi.org/10.48550/arXiv.1511.07122>.
23. Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. *Towards Data Sci*. 2017;6(12):310–6.
24. Misra D. Mish: a self regularized non-monotonic activation function. *AarXiv*. 2019. <https://doi.org/10.48550/arXiv.1511.07122>.
25. Chen Z, Liu Y, Chen C, Lu M, Zhang X. Malicious url detection based on improved multilayer recurrent convolutional neural network model. *Secur Commun Netw*. 2021;2021:1–13.
26. Kalayeh MM, Shah M. Training faster by separating modes of variation in batch-normalized models. *IEEE Trans Pattern Anal Mach Intell*. 2019;42(6):1483–500.
27. Tanaka M. Weighted sigmoid gate unit for an activation function of deep neural network. *Pattern Recognit Lett*. 2020;135:354–9.
28. Lee SH, Goëau H, Bonnet P, Joly A. New perspectives on plant disease characterization based on deep learning. *Comput Electr Agric*. 2020;170: 105220.
29. Barbedo JGA. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput Electr Agric*. 2018;153:46–53.
30. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84–90.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Yaolin Zhu** Yaolin Zhu, Professor, postdoctorate of Northwestern Polytechnical University and master tutor of School of Electronics and Information, Xi'an Polytechnic university. He was mainly engaged in the research of embedded system design, digital media and 3D visualization technology.

**Ran Liu** Ran Liu is a graduate student of School of Electronics and Information, Xi'an Polytechnic university. The research direction is image processing.

**Gang Hu** Gang Hu, Professor, master tutor of School of Science, Xi'an University of Technology. He is mainly engaged in research on computational geometry, graphic image processing, CAGD/CAD, artificial intelligence and computational intelligence.

**Xin Chen** Xin Chen, Associate Professor, master tutor of School of Electronics and Information, Xi'an Polytechnic university. He focuses on communication signal processing, wireless channel modelling, 5G Telematics, and the Internet of Things.

**Wenya Li** Wenya Li, Lecturer of School of Textile science and engineering, Xi'an Polytechnic university. she is mainly engaged in research on new spinning technology, fancy yarn, new textile equipment and new technology.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---