

Animal Fur Recognition Algorithm Based on Feature Fusion Network

Peng Liu^{1,2*}, Tao Lei^{1,2}, Qian Xiang¹, Zexuan Wang¹, Jiwei Wang³

Abstract

China is a big country in animal fur industry. The total production and consumption of fur are increasing year by year. However, the recognition of fur in the fur production process still mainly relies on the visual identification of skilled workers, and the stability and consistency of products cannot be guaranteed. In response to this problem, this paper proposes a feature fusion-based animal fur recognition network on the basis of typical convolutional neural network structure, relying on rapidly developing deep learning techniques. This network superimposes texture feature - the most prominent feature of fur image - into the channel dimension of input image. The output feature map of the first layer convolution is inverted to obtain the inverted feature map and concat it into the original output feature map, then Leaky ReLU is used for activation, which makes full use of the texture information of fur image and the inverted feature information. Experimental results show that the algorithm improves the recognition accuracy by 9.08% on Fur_Recognition dataset and 6.41% on CIFAR-10 dataset. The algorithm in this paper can change the current situation that fur recognition relies on manual visual method to classify, and can lay foundation for improving the efficiency of fur production technology.

Key Words: Convolutional Neural Network, Fur Image Recognition, Texture Features, Activation Function, Feature Fusion.

I. INTRODUCTION

China is a major producer and consumer of fur luxury goods. Fur products have also gone out of the exclusive rights of the rich and powerful, and the public's demand for fur has increased year by year [1]. Fur products are increasingly close to people, and consumers' preference for fur products is also on the rise year by year [2]. However, the identification of animal fur in the fur production process still relies on the visual identification of skilled workers, which is easily affected by external factors, and it is difficult to ensure the consistency and stability of the product [3-4]. The method of manual visual identification can no longer be matched with the automated fur production process [5]. Zhong Y. et al. studied fur recognition using the compound microscopy [6]. Zoccola M. et al. investigated near-infrared spectroscopy method and used this method for fur recognition [7]. Baichoo N. et al. achieved recognition by analyzing the DNA of fur [8]. Chen H. et al. investigated near-infrared spectroscopy and chemometric models for animal fur recognition [9]. Lu K. et al. verifies that the bag-of-words models and spatial pyramid match is an effective ap-

proach to identify the animal fur [10]. Vineis C. et al. studied fur recognition using peptide analysis by ultra-performance liquid chromatography/electrospray-mass spectrometry [11].

With the widespread use of image acquisition equipment, images have gradually become the main information exchange method for human beings, and a large number of image databases have been generated [12]. Relying on the value of big data, machine learning shows a wide range of applications in related fields [13]. A key step in the early stages of machine learning is to manually design a variety of effective features. The degree of discrimination of feature expression will have a decisive influence on the effect of the algorithm. Traditional methods usually rely on artificially designed features, which have many disadvantages [14]. For example, in the face of different scenes and different types of data, it is necessary to design different artificial features, resulting in high labor cost and time cost, and high requirements on experience of feature design [15-16]. Deep learning [17] is an emerging branch of machine learning. Instead of manually selecting features, deep learning uses several layers of networks to perform hierarchical feature

Manuscript received January 23, 2022; Revised March 08, 2022; Accepted March 10, 2022. (ID No. JMIS-22M-01-001)

Corresponding Author (*): Peng Liu, lpeng@cuit.edu.cn

¹Shaanxi University of Science and Technology, Xi'an, Shaanxi, China, lpeng@cuit.edu.cn

²Shaanxi Joint Key Laboratory of Artificial Intelligence, Xi'an, Shaanxi, China, lpeng@cuit.edu.cn

³Graduate School of Science and Technology, Hirosaki University, Hirosaki, Japan, h18ds257@hirosaki-u.ac.jp

extraction on input data and learns data features in this process.

With the rapid improvement of the computing capacity of computer hardware and the rapid popularization of large-capacity storage devices and graphics processors, neural network-based recognition algorithms have gradually become an emerging main direction [18-19]. Based on the research on animal neurons, Yann LeCun designed convolutional neural network in 1998. Drawing on the characteristics of multi-layer perceptron, the network used local connection and weight sharing to carry out calculation, which effectively reduced the parameter scale and greatly reduced the amount of calculation. Alex et al. proposed AlexNet model [20] in 2012, which raised much concern over convolutional neural network in the field of image processing. AlexNet network uses block convolution for parallel training. The Oxford Visual Geometry Group, together with Google Research, proposed the VGGNet network [21], among which VGG-16 network performed better in a number of classification tasks. Microsoft proposed ResNet network [22], which can train deeper networks through jumping connections between layers, so that the problem of gradient disappearance can be effectively alleviated.

In the image recognition task, the arrival of deep learning method has greatly improved the recognition accuracy. Deep learning does not need feature extraction and recognition training as two links, and can automatically extract effective features during training, with very strong feature extraction capability. Moreover, the model obtained after training has strong generalization, with rotation invariance and translation invariance. The weight sharing structure is the biggest characteristic of CNN, which greatly reduces the number of network parameters. CNN shows very good robustness and expansibility, but when CNN is needed to complete image recognition tasks, there will be problems such as gradient disappearance or gradient explosion and over-fitting. In the background of increasing production and consumption of fur, this paper studies the automatic recognition of fur in the process of animal fur production, and proposes an algorithm of animal fur recognition based on

feature fusion network.

II. PROPOSED METHOD

2.1. Convolutional Neural Network Framework

Convolutional neural network is a machine learning model with strong adaptability, which is a kind of supervised learning. Based on the idea of biological neurons, the network adopts the structure of biological neural network and the way of weight sharing. By constantly extracting local features of data, the network gradually maps from shallow to deep to high-level feature space step by step, and then completes the classification task. A typical network structure is shown in Fig. 1. Convolutional neural network obtains the probability vector $Y \in R^{C \times 1}$ through a series of hidden layer mappings. The length of the probability vector is C , where C represents the number of categories. Each value in vector Y corresponds to the probability of a class, and one-hot encoding is adopted. Finally, the maximum value method is used to determine the category corresponding to the sample.

The hidden layer is composed of the following four parts:

(1) Convolutional layer: The main function of the convolutional layer is feature extraction. The feature map output by the convolutional layer is the result of the weighted summation of the feature map of the previous layer and the convolution kernel, and the addition of the bias term. The calculation formula is shown in formula (1), and the convolution operation can extract the local features of the image.

$$Z^{l+1}(i,j) = g\{\sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [Z_k^l(s_0 i + x, s_0 j + y) w_k^{l+1}(x,y)] + b\}, \quad (1)$$

where b is the deviation, Z^l and Z^{l+1} are respectively l layer feature maps, $Z(i,j)$ is the feature map pixels, K is the number of corresponding channels, f is the size of the convolution kernel, s_0 is the convolution stride (Stride), g is the non-linear activation function (Fig. 2).

(2) Activation function: The activation function makes

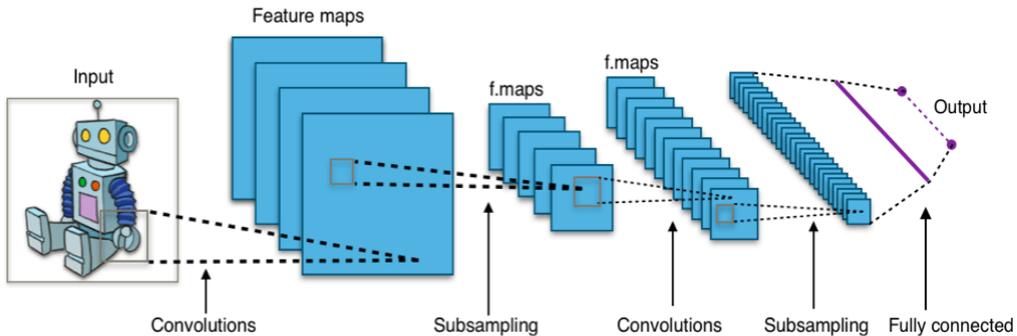


Fig. 1. Schematic diagram of convolutional neural network structure.

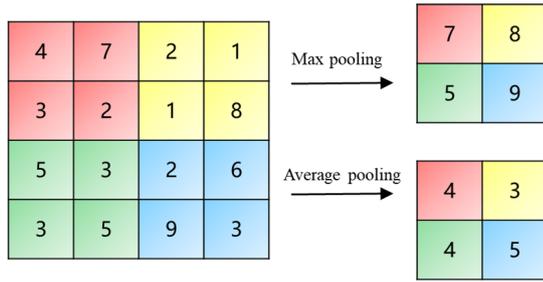


Fig. 2. Schematic diagram of maximum pooling operation and average pooling operation.

nonlinear correction for each convolution. Commonly used activation functions include ReLU, Tanh and Sigmoid.

(3) Pooling layer: The pooling layer down-sampling the feature map. Down-sampling can reduce the computational complexity while maintaining the features. Pooling operations mainly include average pooling, maximum pooling, and random pooling, the last of which is less used.

Its general expression is formula (2).

$$A_k^l(i, j) = \left[\sum_{x=1}^f \sum_{y=1}^f A_k^l(s_0i + x, s_0j + y)^p \right]^{\frac{1}{p}} \quad (2)$$

In the formula, s_0 is the step size, pixel (i, j) , and p is the pre-specified parameter.

(4) Local response normalization layer (LRN): This layer performs the contrast normalization operation on the feature maps that have completed the pooling operation.

The ReLU function, also called the Rectifier function, is an activation function often used in neural networks. Its definition is shown in equation (3), and its derivative is defined in equation (4). The image is shown in Fig. 3:

$$f(x) = ReLU(x) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad (3)$$

$$f'(x) = ReLU'(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (4)$$

The unilateral suppression feature of ReLU can block the propagation of negative gradients and lead the network to a certain sparse expression ability. However, when the learning rate is large, it may cause a certain number of neurons in the network to irreversibly die, which seriously affects the performance of the network, as shown in formula (5). In order to improve the ReLU activation function, scholars

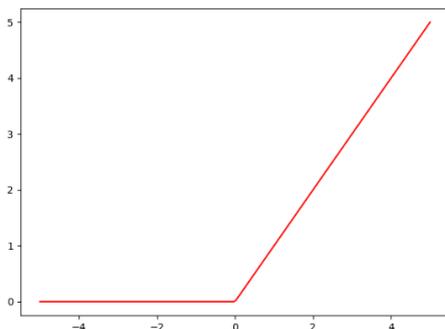


Fig. 3. ReLU function.

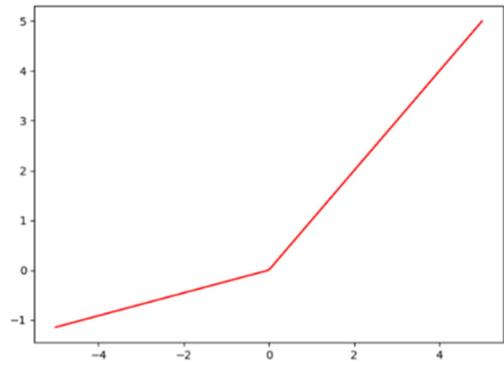


Fig. 4. Leaky ReLU function.

proposed the Leaky ReLU function.

$$\omega_{i+1} = \omega_i + \eta \Delta \omega x_i \quad (5)$$

If $x_i < 0$, then $\eta \Delta \omega x_i \ll 0$, the output is 0 after passing the activation function, and the neuron output is always 0 in subsequent iterations.

Leaky ReLU, also known as the Leaky ReLU function, this activation function improves the shortcomings of the ReLU function that is relatively fragile during training. When the input $x < 0$, the function assigns a non-zero slope to all negative values, thus when the neuron's corresponding value is negative, there can also be a non-zero gradient to update the parameters to avoid permanent inactivation. The Leaky ReLU function is shown in equation (6), and the function image is shown in Fig. 4.

$$LeakyReLU = \begin{cases} x, & x > 0, \\ ax, & x \leq 0, \end{cases} \quad (6)$$

where a is a small constant.

Leaky ReLU uses some negative features, but its utilization is still not sufficient. Aiming at the problem of insufficient utilization of negative features by the activation function, the network model in this paper integrates negative features, as shown in Fig. 6 for details. Network, which is conducive to better use of negative features, can improve the accuracy of fur recognition.

2.2. Texture Feature Descriptor

Researchers have conducted a lot of studies on image texture information. Texture is the performance of the regular arrangement of image pixels, which is very stable, robust to environmental noise and lighting, and can not change due to image rotation and movement [23]. The LBP (local binary pattern) operator is a mainstream descriptor for quantitative analysis of texture information, and it is widely used. Its calculation method is based on a 3×3 area definition. The center value of the area is used as the threshold of the local area, and the surrounding 8 values are compared with the threshold to calculate the LBP value of the local area [24], as shown in Fig. 5. The calculation method

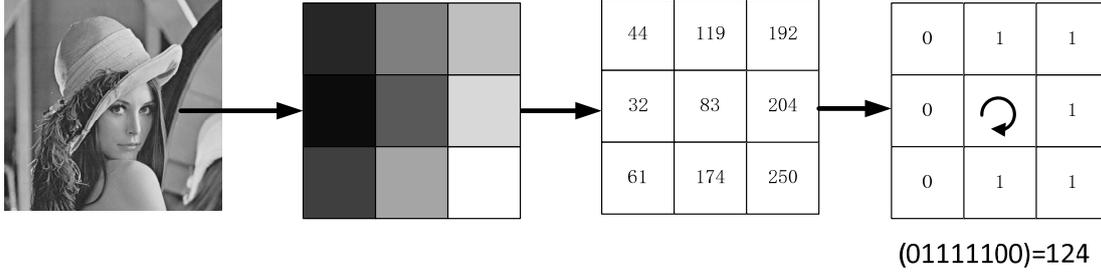


Fig. 5. LBP feature descriptor principle.

is shown in formula (7) and (8). Combining the LBP information of multiple local areas will get the LBP feature map of the entire image. Therefore, changes in the texture features of small local areas will have an impact on the final overall feature situation.

$$LBP(x_a, x_b) = \sum_{p=0}^{p-1} 2^p (i_p - i_b), \quad (7)$$

$$f(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{else,} \end{cases} \quad (8)$$

where (x_a, y_b) is the center point coordinates, p represents the p -th pixel in the neighborhood, i_p is the gray value of the neighborhood pixel, i_b is the gray value of the center point, and $f(x)$ represents the symbolic function.

2.3. Feature Fusion Network Framework and Algorithm Process

This article proposes an improved model design. We have three improvements to the first convolutional layer:

- (1) The input image is fused with LBP feature map.
- (2) Invert the feature map of first layer and fuse it with the feature map before the inversion to form a new feature map.
- (3) The first layer of convolution uses the Leaky ReLU function to activate the new feature map.

The improvement in this paper can effectively integrate

the texture features of the image, and effectively retain and increase the positive propagation of negative features, so as to have a positive impact on subsequent recognition and improve the accuracy of recognition. The subsequent convolutional layer, pooling layer, and fully connected layer are operated in a standard manner. The output of the fully connected layer is sent to the Softmax classifier, and finally the category of the fur image is obtained. The network structure of this article finally uses three convolutional layers and three pooling layers alternately, plus a fully connected layer, and finally uses the Softmax classifier to calculate the probability of the corresponding category of the fur. The network structure is shown in Fig. 6.

The algorithm flow of the feature fusion network is mainly divided into two stages of training and testing. In the network training stages. First, the network parameters are initialized, and then the final output result is calculated forward according to the initialized parameters, and the gradient is calculated backward according to the deviation between the output result and the true value. Update all parameters along the negative gradient direction, and then judge whether the end condition of training is reached. After the end condition is reached, the training phase is completed, and the network model is trained.

In the testing phase of the network, it is necessary to use

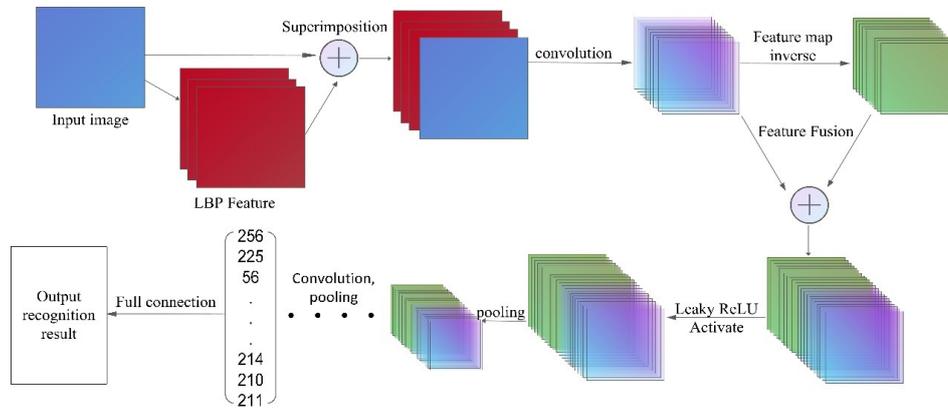


Fig. 6. Feature fusion network structure diagram.

Algorithm 1. Animal fur recognition algorithm based on feature fusion network.

Algorithm steps (network test phase)

Input: test image I

Output: category results $i \in (1, N)$

- (1): Calculate the corresponding LBP feature maps of three single channels of the test image I .
- (2): LBP feature maps and original image are superimposed and fused by channel dimension.
- (3): Perform a convolution operation on the fused image to obtain the feature map of the first layer of convolution.
- (4): Reverse the feature map of the first layer of convolution to obtain the corresponding negative feature map.
- (5): The feature map of the first layer of convolution and the corresponding negative feature map are superimposed and fused by the channel dimension.
- (6): Use the Leaky ReLU function to activate the fused feature map.
- (7): Perform a pooling operation on the result after activation.
- (8): Complete subsequent convolution and pooling operations.
- (9): Pull the characteristic graph to the growth vector.
- (10): Use the softmax classifier to output the classification results.

the trained model to output the classification results according to the input test images. The process is shown in Algorithm 1.

III. EXPERIMENTAL RESULTS AND ANALYSIS

The experiment in this article is performed on the Intel(R) Core i7-8700 CPU 3.20 GHz, the graphics card is RTX 2080Ti, and the operating system is 64bit Win10. The algorithm is completed on the PyTorch framework and implemented with Python 3.7.

3.1. Dataset

The experiment in this paper was carried out with the animal fur dataset and the CIFAR-10 dataset respectively. The animal fur dataset uses the Fur_Recognition dataset, which contains more than 31,000 labeled images, the size of each image being 225×225 , divided into training set and testing set. This dataset collected and established by our laboratory. The number of fur images in training set and testing set is 5:1 Divide. The dataset includes American sheep skin (AMS), Australian sheep skin (AUS), French rabbit skin (FRS), domestic rabbit skin (RAB) and domestic rex rabbit skin (REXRAB), 5 animal fur images, as shown in Fig. 7, classified according to the name, and generating a name library.

CIFAR-10 is a dataset for studying object recognition compiled by the Hinton artificial intelligence team. A single image is 32×32 in size, and each category includes 6,000

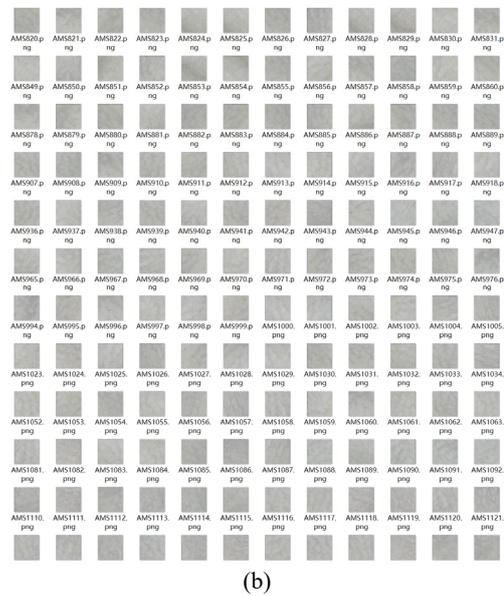
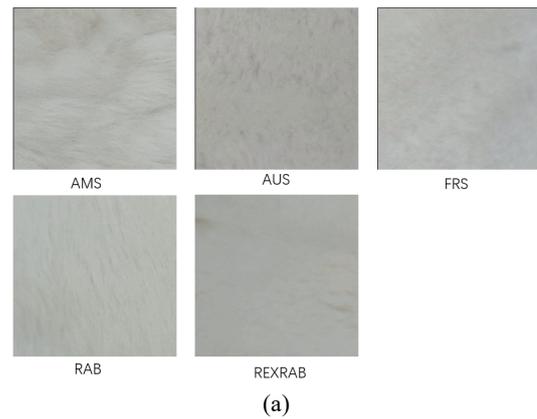


Fig. 7. Partial image of fur dataset.

images. The data set is divided into 5:1 Divide. 50,000 images are used for training and 10,000 images are used for testing. The dataset contains RGB images of 10 types of objects such as truck, frog, deer, airplane, bird, ship, horse, dog, cat and automobile, as shown in Fig. 8. The details of the datasets are shown in Table 1.

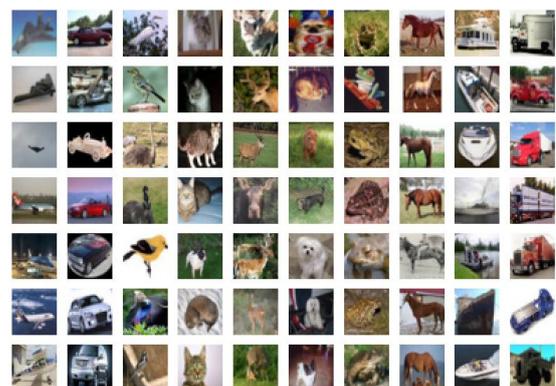


Fig. 8. Partial image of CIFAR-10.

Table 1. The details of datasets.

Dataset	Category name	Category number	Training set	Testing set
Fur Recognition dataset	American sheep skin (AMS)	6240	5200	1040
	Australian sheep skin (AUS)	6240	5200	1040
	French rabbit skin (FRS)	6240	5200	1040
	Domestic rabbit skin (RAB)	6240	5200	1040
	Domestic rex rabbit skin (REXRAB)	6240	5200	1040
CIFAR-10 dataset	Airplane	6000	5000	1000
	Automobile	6000	5000	1000
	Bird	6000	5000	1000
	Cat	6000	5000	1000
	Deer	6000	5000	1000
	Dog	6000	5000	1000
	Frog	6000	5000	1000
	Horse	6000	5000	1000
	Ship	6000	5000	1000
	Truck	6000	5000	1000

3.2. Comparison and Selection of LBP Features

LBP features can describe textures well and can effectively distinguish different types of fur images. We compare four different LBP operators through experiments. Table 2 compares the number of patterns of the four operators under three sampling situations LBP_8^1 , LBP_8^2 and LBP_{16}^3 .

When we select LBP equivalent mode, the recognition accuracy is higher than other modes.

3.3. Comparison of Experimental Results

The initial learning rate of the network model in this paper is 1×10^{-4} , and every 2 cycles, the learning rate decays by 0.2, the momentum factor is 0.99, and the weights are updated by backpropagation through stochastic gradient descent. Table 3 and Table 4 are the parameter settings of each layer applied to the fur data set and the CIFAR-10 dataset, respectively.

In order to determine the optimal size of the first-layer convolution kernel, this paper compares the effects of different size of the first-layer convolution kernel on the recognition accuracy through experiments. For the fur dataset and the CIFAR-10 dataset, four first-layer convolution

Table 2. Number of different LBP operator modes.

	Original LBP	Round LBP	Rotation invariant mode	Equivalent model
LBP_8^1	256	256	9	59
LBP_{16}^2	-	65536	17	243
LBP_{16}^3	-	16777216	25	555

Table 3. Parameter settings of each layer of the network (corresponding to the fur dataset).

Layer type	Convolution kernel size/pooling domain size	Number of kernels	Step size	Activation function / Pooling method
Input	-	-	-	-
Conv1	5×5	32	1	Leaky ReLU
Pool1	2×2	-	2	Max pooling
Conv1	3×3	32	1	ReLU
Pool1	2×2	-	2	Max pooling
Conv1	3×3	64	1	ReLU
Pool1	2×2	-	2	Max pooling
FC	-	-	-	-
Output	-	-	-	-

Table 4. Network layer parameter settings (corresponding to CIFAR-10 dataset).

Layer type	Convolution kernel size / Pooling domain size	Number of kernels	Step size	Activation function / Pooling method
Input	-	-	-	-
Conv1	5×5	32	1	Leaky ReLU
Pool1	2×2	-	2	Max pooling
Conv2	5×5	32	1	ReLU
Pool2	2×2	-	2	Max pooling
Conv3	5×5	64	1	ReLU
Pool3	2×2	-	2	Max pooling
FC	-	-	-	-
Output	-	-	-	-

kernels are set respectively.

Calculate the corresponding recognition accuracy rate, as shown in Table 5. The experimental results show that for the fur dataset, when the first layer convolution kernel is set to 3×3, the recognition accuracy is the highest; for the CIFAR-10 dataset, when the first layer convolution kernel is 3×3 and 5×5, the recognition accuracy rate is relatively

Table 5. Convolution kernel size comparison experiment of the first layer.

Dataset	The size of the first convolution kernel	Accuracy
Fur_Recognition dataset	9×9	76.5
	7×7	78.9
	5×5	80.6
	3×3	82.2
CIFAR-10 dataset	9×9	75.7
	7×7	78.8
	3×3	79.9
	5×5	80.5

close. When the first layer convolution kernel is 5×5, the recognition accuracy rate is slightly higher.

Through the comparative experiments on the Fur_Recognition dataset and the CIFAR-10 dataset, it can be seen that when the original convolutional neural network is used and the ReLU function is used for the activation operation, the recognition accuracy is low. The experimental results are shown in Fig. 9 and Fig. 10. When the activation

function is replaced with Leaky ReLU, the recognition accuracy is improved to a certain extent because some negative features are used. When the activation function is unchanged, the ReLU function is still used to activate, and the LBP feature is integrated before the first layer of the network convolution, and after the convolution is integrated into the inverted feature map, the accuracy is improved compared with the original network. On the fur dataset an increase of 4.18%, an increase of 3.23% on the CIFAR-10 dataset. When the feature fusion network is used, the LBP features and the features that are inverted after the first layer convolution are respectively integrated into the original network, and the ReLU function is replaced with Leaky ReLU to perform the activation function, the recognition accuracy improves greatly. Experimental results show that the algorithm improves the recognition accuracy by 9.08% on Fur_Recognition dataset and 6.41% on CIFAR-10 dataset. Table 6 shows the recognition accuracy of every category in the form of confusion matrix, which shows that the recognition accuracy of the algorithm in this paper is higher than that of the manual visual method in every category.

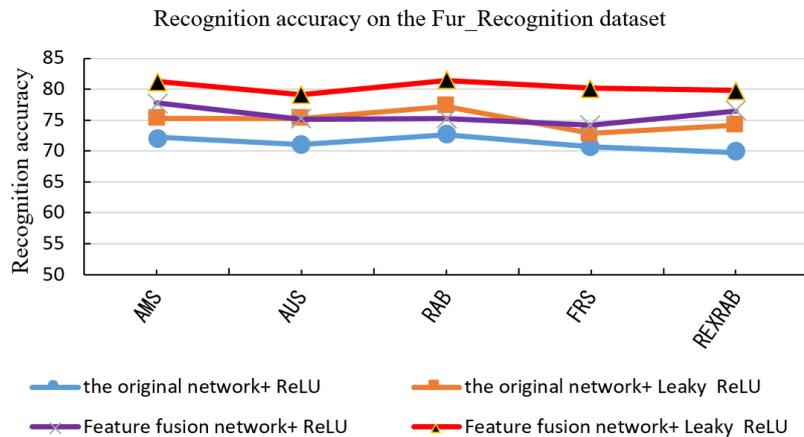


Fig. 9. Recognition accuracy on the Fur_Recognition dataset.

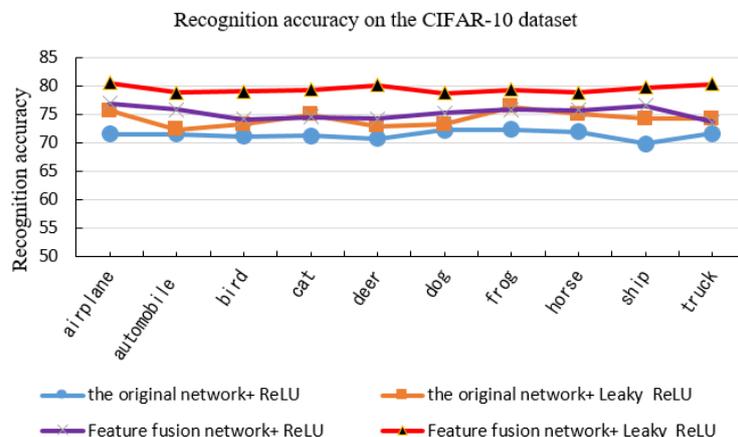


Fig. 10. Recognition accuracy on the CIFAR-10 dataset.

Table 6. Confusion matrix of fur image recognition.

Category	AMS	AUS	RAB	FRS	REXRAB	Line_sum	Pacc%
AMS	854	85	31	37	33	1040	82.1
AUS	80	851	38	31	40	1040	81.8
RAB	29	27	854	68	62	1040	82.1
FRS	33	26	72	847	62	1040	81.4
REXRAB	23	26	56	61	874	1040	84.0
Col_sum	118	116	125	107	97	5200	
Uacc%	83.8	83.8	81.3	81.1	81.6		

IV. CONCLUSION

This paper proposes an animal fur recognition network based on feature fusion. Compared with a typical convolutional neural network, this network extracts LBP features from the original image before the first layer of convolution, and integrates the original image channel dimension to participate in the convolution operation. After the first layer of convolution, by inverting the feature map and using the Leaky ReLU function to activate, the negative features are retained, which makes the feature information of forward propagation more abundant, and finally improves the accuracy of image recognition. This paper verifies the recognition effect of the network on two datasets, which are significantly higher than the typical original network. Applying this network to the fur production process can realize the automatic identification of fur, avoiding the problems of high labor cost, consistency and stability of the fur enterprise's manual visual identification, and promoting the improvement of fur processing efficiency. The disadvantage is that the feature fusion idea of this article has not been tested on large-scale networks. In the next research, we will apply the two feature fusion methods plus the Leaky ReLU function to large-scale networks to address the problem of fur recognition on large-scale networks.

REFERENCES

- [1] C. P. Martin, "Low salt preservation of Australian sheepskins," *Journal of the Society of Leather Technologists and Chemists*, vol. 105, no. 1, pp. 9-16, 2021.
- [2] Y. Wang, Q. Xia, Q. Liu, H. Dai, and Z. Zhang, "Study on the dry-cleaning process of mink fur based on sub-critical solvent," *Journal of the American Leather Chemists Association*, vol. 116, no. 9, pp. 312-316, Sep. 2021.
- [3] S. Çolak and M. Kaygusuz, "Dry heat resistance of leathers of different tannages," *Journal of the Society of Leather Technologists and Chemists*, vol. 105, no. 3, pp. 124-131, 2021.
- [4] V. Sivakumar, "Approaches towards tannery modernization and up-gradation: Leather industry 4.0," *Journal of the American Leather Chemists Association*, vol. 116, no. 2, pp. 4-6, Feb. 2021.
- [5] M. Mehta, Y. Liu, R. Naffa, M. Waterland, and G. Holmes, "Changes to the collagen structure using vibrational spectroscopy and chemometrics: A comparison between chemical and sulfide-free leather process," *Journal of the American Leather Chemists Association*, vol. 116, no. 11, pp. 379-389, Nov. 2021.
- [6] Y. Zhong, K. Lu, J. Tian, and H. Zhu, "Wool/cashmere identification based on projection curves," *Textile Research Journal*, vol. 87, no. 14, pp. 1730-1741, Sep. 2017.
- [7] M. Zoccola, N. Lu, R. Mossotti, R. Innocenti, and A. Montarsolo, "Identification of wool, cashmere, yak, and angora rabbit fibers and quantitative determination of wool and cashmere in blend: A near infrared spectroscopy study," *Fibers and Polymers*, vol. 14, no. 8, pp. 1283-1289, Sep. 2013.
- [8] N. Baichoo and J. D. Helmann, "Recognition of DNA by fur: A reinterpretation of the fur box consensus sequence," *Journal of Bacteriology*, vol. 184, no. 21, pp. 5826-5832, Nov. 2002.
- [9] H. Chen, Z. Lin, and C. Tan, "Classification of different animal fibers by near infrared spectroscopy and chemometric models," *Microchemical Journal*, vol. 144, pp. 489-494, Jan. 2019.
- [10] K. Lu, Y. Zhong, D. Li, X. Chai, H. Xie, Z. Yu, and T. Naveed, "Cashmere/wool identification based on bag-of-words and spatial pyramid match," *Textile Research Journal*, vol. 88, no. 21, pp. 2435-2444, Aug. 2018.
- [11] C. Vineis, C. Tonetti, S. Paoletta, P. D. Pozzo, and S. Sforza, "A UPLC/ESI-MS method for identifying wool, cashmere and yak fibres," *Textile Research Journal*, vol. 84, no. 9, pp. 953-958, Jun. 2014.
- [12] H. Lianhua, X. Chengyi, and Z. Feng, "Research on Sheepskin contour extraction method based on computer vision measurement technology," *Journal of the American Leather Chemists Association*, vol. 116, no. 8, pp. 267-276, Aug. 2021.
- [13] M. Aslam, T. M. Khan, S. S. Naqvi, G. Holmes, and R. Naffa, "Learning to recognize irregular features on leather surfaces," *Journal of the American Leather Chemists Association*, vol. 116, no. 5, pp. 169-180, May 2021.
- [14] C. T. Nguyen and M. Nakagawa, "An improved segmentation of online English handwritten text using recurrent neural networks", in *Proceedings of the 2015 3rd IAPR Asian Conference on Pattern Recognition*

(ACPR), Kuala Lumpur, Malaysia, Nov. 2015, pp. 176-180.

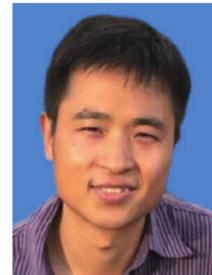
- [15] C. T. Nguyen, B. Indurkha, and M. Nakagawa, "A unified method for augmented incremental recognition of online handwritten Japanese and English text," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 23, no. 1, pp. 53-72, Jan. 2020.
- [16] N. Eligüzcel, C. Çetinkaya, and T. Dereli, "A state-of-art optimization method for analyzing the tweets of earthquake-prone region," *Neural Computing and Applications*, vol. 33, no. 21, pp. 14687-14705, May 2021.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-44, May 2015.
- [18] L. Zhang, P. Liu, and J. A. Gulla, "Dynamic attention-integrated neural network for session-based news recommendation," *Machine Learning*, vol. 108, no. 10, pp. 1851-1875, Jan. 2019.
- [19] W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li, "Development of convolutional neural network and its application in image classification: A survey," *Optical Engineering*, vol. 58, no. 4, 40901, Apr. 2019.
- [20] A. Krizhevshy, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 26th Annual Conference on Neural Information Processing Systems 2012*, Lake Tahoe, NV, Dec. 2012.
- [21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, CA, May 2015.
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, Jun. 2016, pp. 770-778.
- [23] P. Liu, J. M. Guo, K. Chamnongthai, and H. Prasetyo, "Fusion of color histogram and LBP-based features for texture image retrieval and classification," *Information Sciences*, vol. 390, pp. 95-111, Jun. 2017.
- [24] L. Liu, L. Zhao, C. Gue, L. Wang, and J. Tang, "Tu xiang wen li fen lei fang fa yan jiu jin zhan he zhan wang [Texture classification: State-of-the-art methods and prospects]," *Acta Automatica Sinica*, no. 4, pp. 584-607, 2018.

AUTHORS



Peng Liu received the bachelor's degree from Nanjing University of Posts and Telecommunications, Nanjing 210023, China, and the master's degree from University of Electronic Science and Technology of China, Chengdu 610054, China. He is currently pursuing the Ph.D. degree at Shaanxi University of Science and Technology, Xi'an 710021, China.

His current research interests include computer vision and pattern recognition.



Tao Lei received the Ph.D. degree in Information and Communication Engineering from Northwestern Polytechnical University, Xi'an, China. From 2012 to 2014, he was a Postdoctoral Research Fellow with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. He is currently a Professor with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Tech-

nology. His current research interests include image processing, pattern recognition, and machine learning.



Qian Xiang received the master's degree from the Jiangxi Normal University of Science & Technology. From 2015 to 2020, he worked at Fuyang Normal University, China. He is currently pursuing the Ph.D. degree at Shaanxi University of Science and Technology, Xi'an 710021, China. His research interests include signal processing and speech enhancement.



Zexuan Wang is pursuing his master's degree at Shaanxi University of Science and Technology, Xi'an 710021, China. His research interests lie in computer vision and machine learning. His research directions are crowd counting and neural network optimization.



Ji-wei Wang is currently a Ph.D. candidate at Graduate School of Science and Technology, Hirosaki University, Japan. His research interests include machine learning and parameter identification.