

Adaptive low-resolution palmprint image recognition based on channel attention mechanism and modified deep residual network

Xuebin Xu^{1,2}, Kan Meng^{1,2*}, Xiaomin Xing^{1,2}, and Chen Chen^{1,2}

1 School of Computer Science and Technology, Xi'an University of Posts & Telecommunications,
Xi'an Shaanxi 710121, China

2 Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University
of Posts & Telecommunications, Xi'an Shaanxi 710121, China

[e-mail: 1076539078@qq.com]

*Corresponding author: Kan Meng

*Received December 5, 2021; revised January 12, 2022; revised February 14, 2022; accepted March 10, 2022;
published March 31, 2022*

Abstract

Palmprint recognition has drawn increasingly attentions in the past decade due to its uniqueness and reliability. Traditional palmprint recognition methods usually use high-resolution images as the identification basis so that they can achieve relatively high precision. However, high-resolution images mean more computation cost in the recognition process, which usually cannot be guaranteed in mobile computing. Therefore, this paper proposes an improved low-resolution palmprint image recognition method based on residual networks. The main contributions include: 1) We introduce a channel attention mechanism to refactor the extracted feature maps, which can pay more attention to the informative feature maps and suppress the useless ones. 2) The ResStage group structure proposed by us divides the original residual block into three stages, and we stabilize the signal characteristics before each stage by means of BN normalization operation to enhance the feature channel. Comparison experiments are conducted on a public dataset provided by the Hong Kong Polytechnic University. Experimental results show that the proposed method achieve a rank-1 accuracy of 98.17% when tested on low-resolution images with the size of 12dpi, which outperforms all the compared methods obviously.

Keywords: Deep learning, Low-resolution palmprint recognition, attention mechanism.

1. Introduction

Biological characteristics, which are based on one or more biological behaviors or physical characteristics of the identification method, currently have a wide range of social research value [1]. With the development of online shopping in recent years, biometrics has been widely used in real life. Common biometrics include fingerprints, veins, face, ears, iris, gait, and voice. The palmprints in these biometrics have rich ridges, fingerprints, main lines, and wrinkles, and these features are unique to individuals and are difficult to modify. Thus, palmprint recognition has extensive research in personal identity verification.

Palmprint recognition generally focuses on the research of ordinary resolution (about 100dpi) and high resolution (above 500dpi), and few people pay attention to the research of lower resolution, such as about 50dpi [2, 3, 4]. [5] The study set the resolution between 10dpi and 20dpi. In the case of 12dpi, the recognition rate of this paper's method is only 30.00%, while our method has 98.17%. [6, 7] Proposed the best neural network (MobileNetv3, EfficientNet-b3, DenseNet121) for palmprint recognition at present, which can achieve 100% recognition rate under the original resolution. With the development of Internet technology, people enjoy the convenience brought by technological advancement and at the same time continue to enable relevant technology companies to develop products with higher customer experience. When the user's collection device does not meet the standard resolution, the previous palmprint recognition rate will be reduced, and when users use mobile devices to collect palmprint information and then upload the collected palmprint information to the server, the palmprint recognition rate will decrease due to network instability. Therefore, low-resolution palmprint research is a new research direction, which has broad research prospects and value [8].

In order to solve above problems, we propose a general framework with channel features, which can obtain a very deep trainable low-resolution palmprint recognition network which named Low-resolution palmprint recognition network based on channel attention mechanism (LPCANet). The framework flow chart of this network is shown in Fig. 1, which includes three stages: construction of deep residual network (LPCANet) combined with channel attention mechanism, deep discriminative feature extraction and classification. When the pixels of the palmprint image are very low, the conventional network will filter out useful information during feature extraction, resulting in a low palmprint recognition rate. At the same time, with the deepening of the network layer, the learning ability of the network will be enhanced, and the extraction features of the network will be better. However, with the deepening of the network layer, the problem of network degradation will also be brought, which has a great impact on the low-resolution palmprint recognition. We introduce a new residual learning structure to overcome the degradation problem caused by the deeper the network. At the beginning of feature extraction, the attention mechanism is adopted to adaptively readjust the features of each channel by modeling the interdependence between feature channels. This CA (channel-wise attention) mechanism enables our proposed network to focus on more useful channels and enhance the ability of discriminative learning.

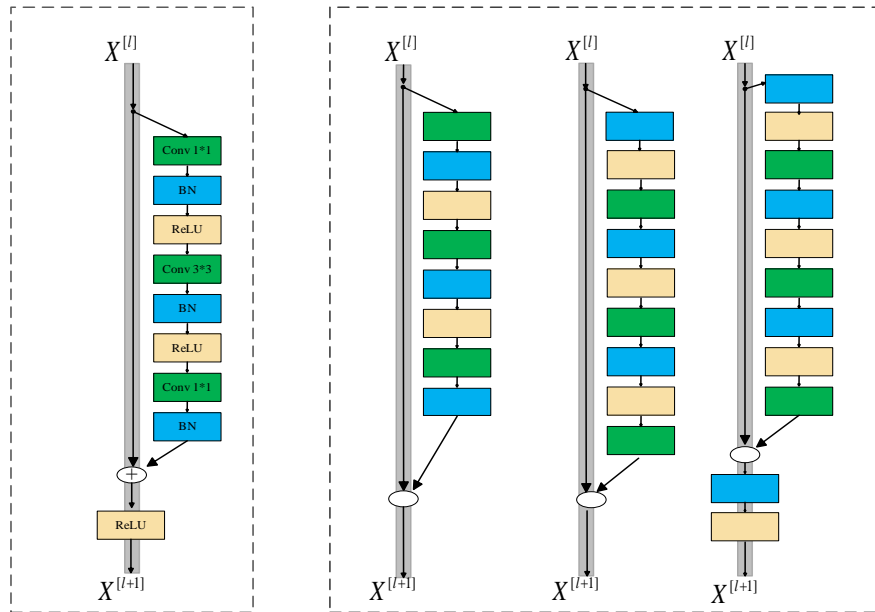


Fig. 1. Residual Building block architectures: (a) original [9]; (b) proposed ResStage.
(* the first BN in the first Middle Resblock is eliminated in each stage).

In short, the contribution of this article is divided into the following three aspects: (1) we propose a deep residual network (IRCANet) with unique residual blocks for low-resolution palmprint recognition. (2) this paper introduces the CA mechanism, the feature is adaptively readjusted by considering the interdependence between feature channels. (3) In the last layer of the network, before the features are input to the linear layer, we changed the structure of the previous network, using the maximum pooling method to retain the main features, extract the texture information of the features, and maximize the retention of the low-resolution image feature information.

Next, in the second section, we will introduce the knowledge of the related residual network and the channel attention mechanism. In the third section of our work, we will introduce the low-resolution palmprint image recognition model. Finally, in the fourth section, we will show our experimental results on the palmprint database of Hong Kong Polytechnic University. Finally, in the fifth section, we will summarize our work.

2. Related Work

2.1 Resnet Network

The emergence of convolutional neural networks [10, 11] has continuously increased the accuracy of image classification. Convolutional neural networks combine different feature extractors and classifications in an end-to-end manner. The higher the feature extractor, the more Network depth to support. More and more evidences show that the deeper the network layer, the more accurate the recognition effect of the network. However, as the network depth increases, there will be a problem of gradient disappearance [12, 13, 14]. In the first dozen or so layers, the problem of gradient disappearance can be solved by normalization [12, 15, 16,

17], but as the number of network layers reaches very deep, the problem of gradient disappearance will become very serious. The specific performance of gradient disappearance is that it climbs to the highest point first and then drops rapidly, which is similar to over-fitting, but not caused by over-fitting [18, 19].

The resnet network introduces the residual module to explicitly adapt the stacked network layers to the residual mapping. The residual network is implemented by a feedforward neural network with fast connection. This connection does not increase additional parameters or increase calculations. Complexity, which solves the problem of network gradient degradation caused by higher depth to a certain extent. However, resnet did not really solve the problem of network gradient degradation [20]. The ResStage group structure proposed by us divides the original residual block into three stages, and we stabilize the signal characteristics before each stage by means of BN normalization operation to enhance the feature channel.

2.2 Channel Attention Mechanism

For a picture with three channels, the weight of the information provided by each channel is different. The channel attention is to weight the convolutional feature channels to further improve the expressive ability of the features, because we are studying Low-resolution palmprint recognition, so when introducing the channel attention mechanism, we only use the maximum pooling method to retain useful sample information, which is different from the original channel attention mechanism.

3. Low-resolution palmprint recognition network based on channel attention mechanism

3.1 Network Architecture

As shown in Fig. 2, our LPCANet mainly consists of four parts: primary feature extraction, LPCANet structure deep feature extraction, max pooling, fully connected layer classification.

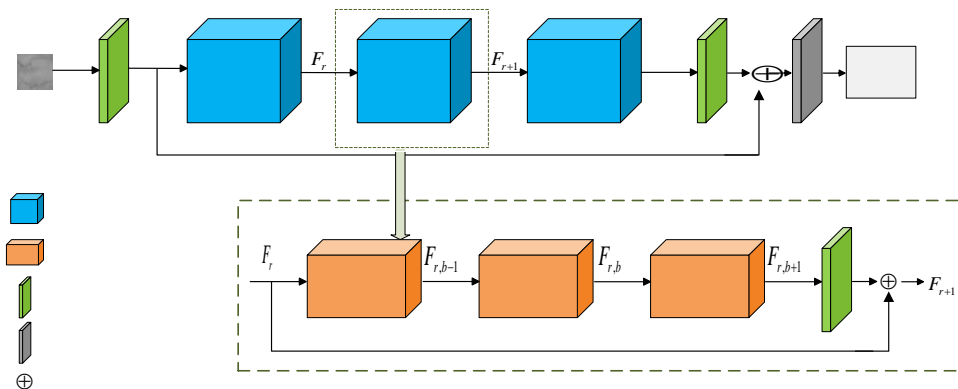


Fig. 2. Network architecture of our Low-resolution palmprint recognition network based on channel attention mechanism (LPCANet).

Let's denote U_o and U_c as the input and output of LPCANet. Convolutional layer can extract image features, so we use a convolutional layer (Conv) to extract relatively shallow data features from palmprint images:

$$F_0 = C_{PF}(U_o) \quad (1)$$

Where $C_{PF}(\cdot)$ represents the convolution operation of feature extraction. F_0 is then on the basis of shallow extraction, we use our LPCANet module to further and more in-depth extraction of features, so we can further have:

$$F_1 = C_{RCAG}(F_0) \quad (2)$$

Where $C_{RCAG}(\cdot)$ denotes our proposed new ResStage Group structure (RCAB module), which extracts the deep features in the channel through broadening the residual and prominent the important feature. By this process, the unimportant feature of the low-resolution palmprint image could be bypassed by the network, so that the deep feature could be more distinguished, we treat it as max pooling layers input:

$$F_2 = C_{UP}(F_1) \quad (3)$$

Where $C_{UP}(\cdot)$ and F_2 denote max pooling and further deep feature respectively. The role of the pooling layer is to filter out useful features from many features to prevent over-fitting problems in classification tasks. The pooling layer ensures that the main features are retained, reduces the dimensionality, reduces the amount of calculation, and enhances the expressive ability of the network to reduce the noise points in the information. The final features are classified by the fully connected layer:

$$U_C = C_{Fc}(F_2) = C_{LPCANet}(U_o) \quad (4)$$

Where $C_{Fc}(\cdot)$ and $C_{LPCANet}(\cdot)$ denote the fully connected layer and the main function of our LPCANet module. The fully connected layer integrates the extracted useful features together.

Then LPCANet is optimized with a loss function. The cross-entropy loss function is a common classification model. When combined with the sigmoid function, it can avoid the decrease of the learning rate of the mean square error loss function during gradient descent, and is used to predict the distribution of different categories during the LPCANet training process. The loss function is as follows:

$$L(x, class) = -\log \frac{\exp(x[class])}{\sum_j \exp(x[j])} = -x[class] + \log(\sum_j \exp(x[j])) \quad (5)$$

Where x and $class$ denotes the parameter set of our network. We use a stochastic gradient descent algorithm with momentum to optimize the loss function, and finally converge to a better local extreme point or even a global extreme point.

3.2 ResStage group structure

Next, we discuss the detailed information of the ResStage group structure we designed in the resnet network (Fig. 1), which overflows many residual blocks into three stages, stabilizing the signal characteristics before each stage, because we have the signal will use BN normalization operation, which enhances the characteristic channel, improves the information flow through the network, and provides a better path for information to propagate through the network layer.

We divide the network into stages. Each stage mainly contains three parts. The BN and ReLU of each stage are used to stabilize the signal and prepare for the next stage. In the last layer of our Start ResBlock, we use the BN layer. The benefits of doing so are also obvious. It normalizes the complete signal and improves the learning ability of the network. At the same time, we can eliminate the BN layer of Middle ResBlock, because we used the BN operation in the Start ResBlock part to make the signal more standardized. Our purpose in doing this is to realize the efficient circulation of information into the next network, while also keeping the information in a stable and controllable state.

3.3 Channel Attention mechanism

The low-resolution palmprint recognition method based on CNN does not pay much attention to the relationship between channels, resulting in poor experimental results under low-resolution. We used channel attention (Fig. 3) to master the relationship between channels and extract effective information of each channel, so as to solve the problem of too little feature information of low-resolution palmprint collection.

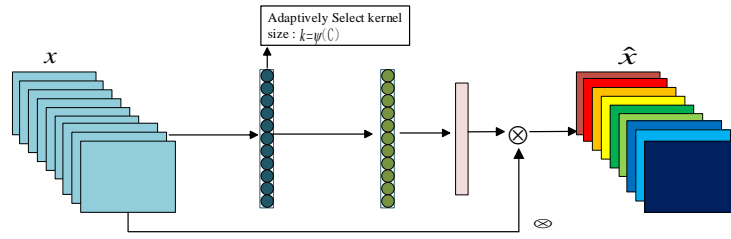


Fig. 3. Channel attention mechanism.

Collecting the characteristics of each channel is the first step of the channel attention mechanism and it is also a key step. Two points are critical in channel attention collection. First, the low-resolution palmprint image information contains rich low-frequency information and high-frequency information. The low-frequency information is relatively flat, while the high-frequency is full of palmprint image edges, textures and other details. Secondly, each filter in the Conv layer uses a local receptive field to operate, so the output after convolution cannot extract context information outside the local area.

Based on above analysis, we replaced the original network's combination of adaptive average pooling and adaptive maximum pooling with only adaptive maximum pooling to bring channel global information into the channel. As shown in Fig. 3, let $x_c \in R^{H \times W \times C}$ is the input parameter, which has C feature maps with the size of $H \times W$. Useful channel information statistics $z_c \in R^C$ can be obtained by shrinking through channel dimensions $H \times W$. The following formula can show this process:

$$z_c = H_{MP}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (6)$$

Where $x_c(i, j)$ is the value at the position (i, j) of the feature z_c . $H_{MP}(\cdot)$ denotes the max pooling function, Statistically useful channels can be regarded as the best description of local information. This effect is very obvious, and it is conducive to the expression of the entire image feature [21]. In addition to max pooling, there are some complex fusion techniques, but

for low-resolution palmprint recognition, only max pooling has the best effect.

In order to fully obtain the relationship between channels from the maximum pooling information, we introduced the idea of gate control. As expressed in [21], the gate control should be able to learn the nonlinear mutual exclusion relationship between each channel, and emphasize the characteristics of multiple channels. Here we use a simple gating mechanism with a sigmoid function.

$$s = f(W_U \delta(W_D z_c)) \quad (7)$$

Where $f(\cdot)$ and $\delta(\cdot)$ respectively represent the sigmoid function and the ReLU [22] function. W_D is the weight set of a Conv layer, and its function is to connect to some nodes of the previous layer, learn local features and share weights. After the shallow features are activated by the ReLU function, and then amplified by the channel, the resulting weight set can be expressed as W_U . After this operation, We get the final channel statistical information, which can be readjusted and input operations are performed x_c :

$$\hat{x}_c = s_c \cdot x_c \quad (8)$$

Where s_c and x_c are the scaling factor and the feature map that the c-th channel represents.

3.4 Residual Channel Attention Block (RCAB)

To sum up, the design of ResStage expands the feature channel, making the main part of the network focus on more information components of low-resolution palmprint image features, which has better results in deeper networks, and the channel attention extracts the inter-channel The interaction relationship further enhances the feature extraction ability at low-resolution. We integrate CA into ResStage and propose ResStage channel attention Group (RCAG). For the b-th ResStage in r-th RCAG, we have:

$$F_{r,b} = F_{r,b-1} + H_{r,b}(X_{r,b}) \cdot X_{r,b} \quad (9)$$

where $H_{r,b}$ denotes the function of channel attention $F_{r,b-1}$, $F_{r,b-1}$ are the input and output of RCAB, which learns the residual $X_{r,b}$ from the input.

4. Experimental results and discussion

In this section, the low-resolution palmprint database of Hong Kong PolyU is used to testify the effectiveness of the algorithm. We use the LPCANet model described in Section 3.1 to extract deep features and classification. Then we evaluate the model performance of different resolutions and compare it with other deep networks (eg: ResNet [23], GoogleNet [24]...) at the same resolution. The LPCANet model is evaluated on NVIDIA Tesla V100 SXM2 16G. The software environment is Pytorch1.5.1.

4.1 Residual Channel Attention Block (RCAB)

The low-resolution palmprint database of Hong Kong PolyU is used to testify the effectiveness of the algorithm. The PolyU palmprint dataset is considered to be one of the most authoritative public palmprint datasets. Our data set has a total of 6000 images in each band, the resolution

of the image is 352×288 (<100 dpi), including 500 different palm print information. When the data was collected, there were 195 males with an age distribution of 20 to 60 years old. The collection time interval was 9 days each time to ensure that the conditions of the data set were the same. Here, all the images are pretreated, and the extracted Region of Interesting (ROIs) are finally normalized to size 128×128 .

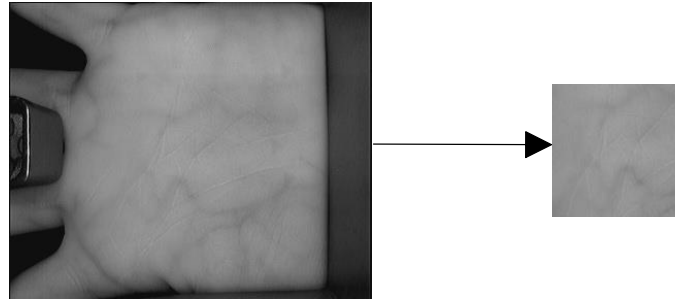


Fig. 4. Original palmprint image and ROI.

4.2 Effect of pooling method on LPCANet

According to the analysis in Section 3.3 of this paper, pooling methods are the key factor. In this section, its effects on our LPCANet are evaluated. We train our model by setting three pooling methods. The results of the experiment can be obtained from **Table 1**. First of all, we fix all remaining parameters of networks, and our model achieves the best results when adopting the max pooling. Additionally, when we adopt the two kinds of mixed pooling, the results are unsatisfactory. The above analyses show the effectiveness of the max pooling in obtaining stable and better results.

Table 1. Accuracy in different pooling methods

Pooling Methods	Different resolution ratio								
	128	96	84	72	60	48	36	24	12
Maxpoolig	99.77	99.57	99.53	99.83	99.73	99.73	99.73	99.7	98.17
Avgpooling	99.6	99.67	99.67	99.7	99.6	99.67	99.7	99.47	74.67
Max+Avg	99.73	99.77	99.83	99.83	99.7	99.77	99.97	99.77	90.53

4.3 Network performance in different depth

Our experimental results show that the LPCANet network model has a good recognition effect when the depth is 50, 101, and 152, and the recognition effect is the best when the depth is 50. When the depth increases from 101 to 152, the network degenerates, and the result is not as good as the 50-layer effect.

Although the recognition accuracy starts to decrease at the depth of 152, the results show that it is not caused by optimization problems. As the depth increases, the results are better. All these indicate that LPCANet has no degradation problems. The only limitation comes from data and computing resources. It can be seen that the low-resolution palmprint recognition performance of the 50-layer LPCANet is very good than other depths.

Table 2. Performance results at different depths

LPCANet layer	Acc(%)	Loss(%)	Params(M)	GFLOPs(G)
50	97.97	0.0808	0.54	25.06
101	94.97	0.2011	1.14	44.05
152	40.30	0.6545	1.69	59.69

4.4 Comparison with state of the art moth as the resolution decrease

We first evaluate the model performance on different networks' convergence effect with decreasing resolution. Using palmprint database for training on IResNet [19], ResNet [23] and GoogleNet [24] networks, the experimental results are also obvious. As shown in Fig. 5, in the range of 120-60dpi, each network has a good effect on palmprint image recognition. At 60dpi, the distinction of each network began to be significant. We can see that when the resolution is 30dpi the GoogleNet accuracy drops the fastest. At 30dpi, the IResNet and ResNet accuracy has also begun to decline. However, due to its unique residual group structure and attention mechanism, LPCANet can still maintain an accuracy of up to 98% at 10dpi. Our model is compared with IResNet, GoogleNet and ResNet models under the same conditions, and the results are shown in Table 3. From the results, it can be concluded that the channel attention mechanism we added has a significant impact on the model. In the case of image resolution is low, it can assess and calculate the correlation of the low-resolution images between each channel, and then obtain the important degree of each channel, finally endow the different weights for each channel. In this way, the useful features of low-resolution images can be better preserved and the irrelevant important features can be suppressed.

Table 4 compares our network and MobileNetv3 [6, 7], EfficientNet-b3 [2, 3], DenseNet121 [6, 7] recognition rate, when the image resolution is set to 72 dpi, 48 dpi, 32 dpi, 24 dpi, 12 dpi, due to the network (EfficientNet-b3, DenseNet121) input limitation, we can only do the resolution experiment of 32dpi in these two network experiments, so the data of these two networks are empty at 24dpi and 12dpi. Experimental results show that with the decrease of resolution, the accuracy of each network also decreases, but our network generally tends to be stable and the accuracy remains above 98%.

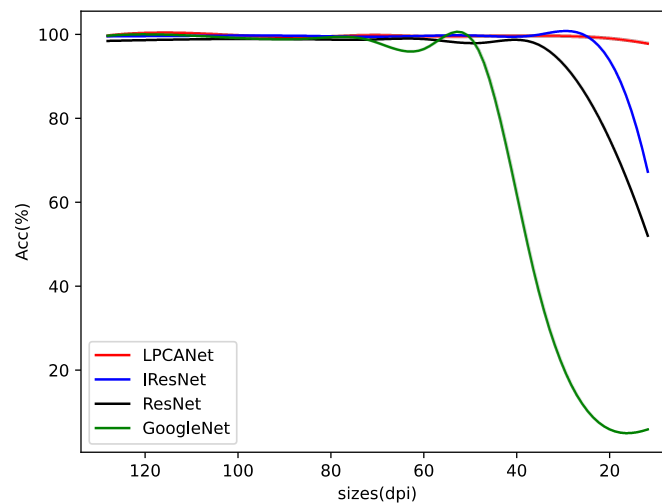
**Fig. 5.** The performance of different networks with decreasing resolution.

Table 3. Comparison (%) with state-of-the-art CNNs on PolyU validation set using 10×10 dpi

Network	Top-1 err	Top-5 err
LPCANet	1.83	0.23
IResNet	24.73	7.53
ResNet	37.3	18.47
GoogLeNet	91.4	77.03

Table 4. Comparison (%) of recognition results of four kinds of networks at different resolutions

Network	Different resolution ratio				
	72	48	32	24	12
MobileNetv3	99.73	99.07	28.17	4.53	0.80
EfficientNet-b3	99.83	99.80	0.97	-	-
DenseNet121	99.63	98.80	95.37	-	-
LPCANet	99.83	99.73	99.70	99.76	98.17

4.5 Network convergence under different resolutions

As the number of iterations increases, the training results of each network on the data set gradually converge. The computational results of the second experiment, LPCANet give us the best average recognition accuracy on each resolution, compared to other exacted network used for palmprint recognition tasks, the use of channel attention mechanism effectively extract channel features, so that the network to converge in a less number of iterations. **Figs. 6** shows the performance of the LPCANet (50 layers) as compared with other networks in different resolutions. Since GoogLeNet does not have layer 101, **Figs. 7** shows the convergence effect comparison between LPCANet (101 layer), IResNet and ResNet. From **Figs 6 and 7**, we can see that our method's convergence rate is the fastest, and convergence accuracy is the highest compared with other methods in different resolutions. Especially, in case of low-resolution, our method is obviously better than other methods on convergence and accuracy. When the experimental data's resolution is set 12dpi, our method can converge rapidly and the recognition accuracy is much higher than other networks. Then, the network structure with different layers, our experimental results also have the advantages, when the network layer is set 100 and the resolution is 12 dpi, the advantages is most obvious, which indicates that our method can get the advantages of consistence in different network structures. The experimental results are better than former in **Figs 6 and 7** due to two improvements. The first is that we proposed the ResStage group: we use BN normalization operation to stabilize the signal characteristics before each stage, enhance the feature channels and improve the information flow on the Internet. The second is the channel attention mechanism: it evaluates and calculates the correlation between channels in low-resolution images and then assigns different weights to each channel. These two improvements will enable our network to better recognize low-resolution images.

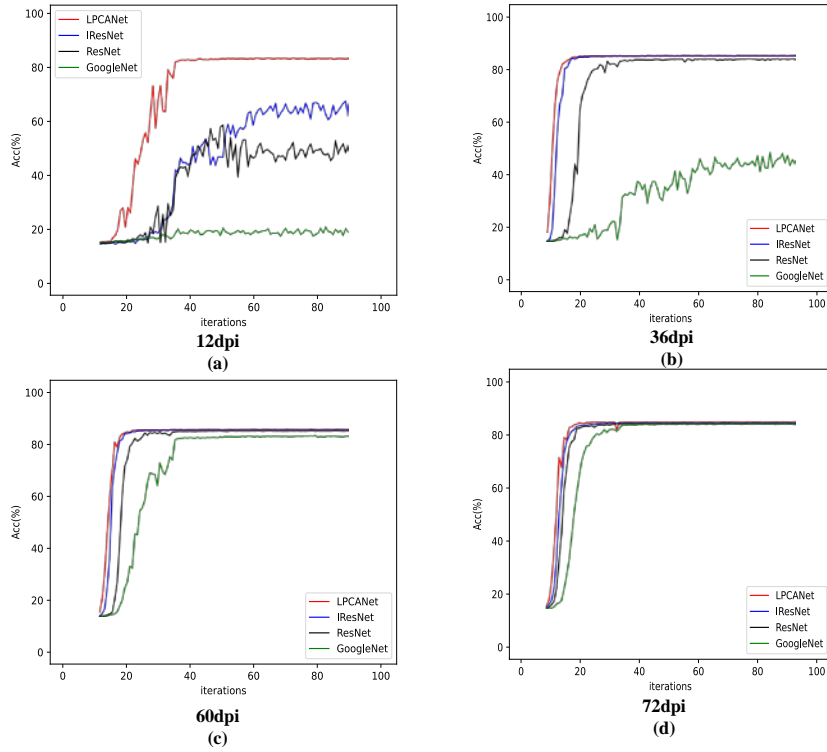


Fig. 6. Comparison of LPCANet (50 layers) and other networks at different resolutions.

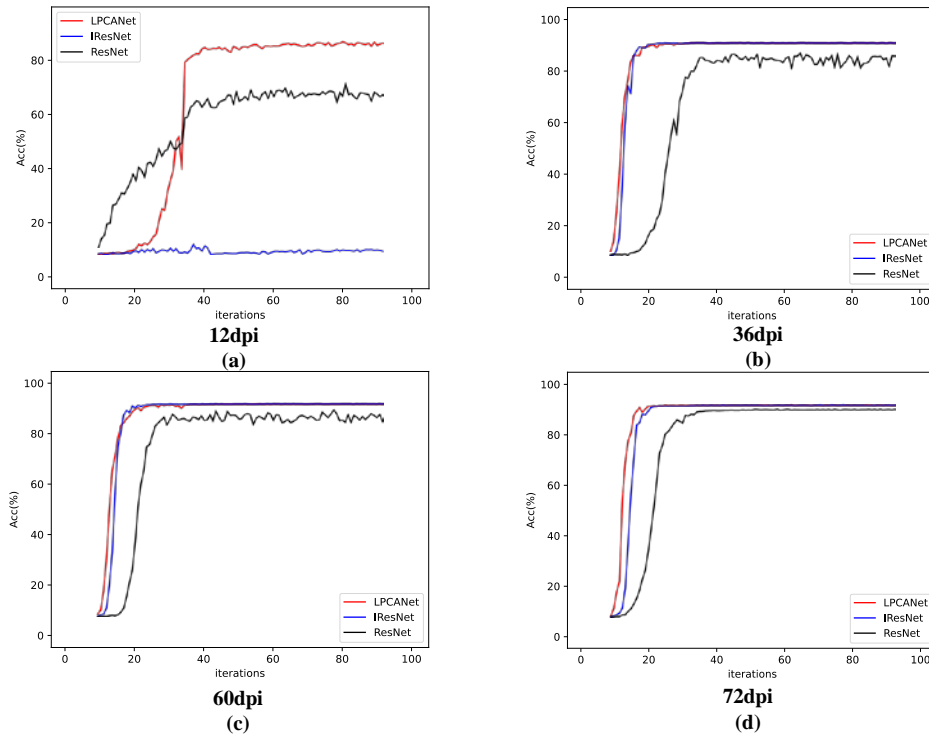


Fig. 7. Comparison of LPCANet (101 layers) and other networks at different resolutions.

4. Conclusions

The network LPCANet we proposed in this paper is mainly to solve the problem of low palmprint recognition at low-resolution. Through our network, we can retain useful features to the greatest extent during feature extraction, through the channel attention mechanism and the maximum pooling operation, and we will still have gradient descent problems when the number of layers in the resnet is deeper, combined with the iresnet network design out of our own network. The experimental results on the Hong Kong Polytechnic University database show that our proposed method can achieve satisfactory accuracy in the palmprint recognition rate with a sufficiently small template size, which is very suitable for practical production applications.

Appendix

Disclosures

The authors declare no conflict of interest.

Acknowledgments

Thanks to my teachers and classmates for giving me guidance on my studies.

References

- [1] Pankanti, Sharath, *Biometrics: Personal Identification in Networked Society*, Springer Publishing Company, Incorporated 1999.
- [2] K, J. A., Jianjiang, F., "Latent palmprint matching," *IEEE transactions on pattern analysis and machine intelligence*, 31(6), 1032-1047, 2009. [Article \(CrossRef Link\)](#)
- [3] Zhang, D., Kong, W., You, J., Wong, M., "Online palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), 2003. [Article \(CrossRef Link\)](#)
- [4] Wei, S., Zhang, D., "Automated personal identification by palmprint," *Optical Engineering*, 37(8), 2359-2362, 1998. [Article \(CrossRef Link\)](#)
- [5] Li C, Benezheth Y, Nakamura K, et al., "A robust multispectral palmprint authentication algorithm and its evaluation for embedded applications," *Journal of Systems Architecture*, 88, 43-53, 2018. [Article \(CrossRef Link\)](#)
- [6] Michele A, Colin V, Santika D D., "Mobilenet convolutional neural networks and support vector machines for palmprint recognition," *Procedia Computer Science*, 157, 110-117, 2019. [Article \(CrossRef Link\)](#)
- [7] Jia W, Gao J, Xia W, et al., "A performance evaluation of classic convolutional neural networks for 2D and 3D palmprint and palm vein recognition," *International Journal of Automation and Computing*, 18(1), 18-44, 2021. [Article \(CrossRef Link\)](#)
- [8] Wu, X., Zhang, D., Wang, K., "Fisherpalms based palmprint recognition," *Pattern Recognition Letters*, 24(15), 2829-2838, 2003. [Article \(CrossRef Link\)](#)
- [9] He, Kaiming, et al., "Deep residual learning for image recognition," in *Proc. of the IEEE conference on computer vision and pattern recognition*, 2016.
- [10] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. of NIPS*, 2012. [Article \(CrossRef Link\)](#)
- [11] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, 1(4), 541-551, 1989. [Article \(CrossRef Link\)](#)

- [12] S. Hochreiter, *Untersuchungen zu dynamischen neuronalen netzen*, Diploma thesis, TU Munich, 1991.
- [13] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, 5(2), 157–166, 1994. [Article \(CrossRef Link\)](#)
- [14] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. of AISTATS*, 2010.
- [15] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Muller, "Efficient backprop," in *Neural Networks: Tricks of the Trade*, Springer, 1998, pp. 9–50. [Article \(CrossRef Link\)](#)
- [16] A. M. Saxe, J. L. McClelland, and S. Ganguli, "Exact solutions to the nonlinear dynamics of learning in deep linear neural networks," *arXiv:1312.6120*, 2013. [Article \(CrossRef Link\)](#)
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. of ICCV*, 2015. [Article \(CrossRef Link\)](#)
- [18] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. of CVPR*, 2015. [Article \(CrossRef Link\)](#)
- [19] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," *arXiv:1505.00387*, 2015.
- [20] Duta, I. C., Liu, L., Zhu, F., Shao, L., "Improved Residual Networks for Image and Video Recognition," in *Proc. of 2020 25th International Conference on Pattern Recognition (ICPR)*, 2021. [Article \(CrossRef Link\)](#)
- [21] Minaee, S., Wang, Y., "Palmpoint recognition using deep scattering network," in *Proc. of IEEE International Symposium on Circuits & Systems*, pp 1-4, 2017. [Article \(CrossRef Link\)](#)
- [22] Yue, F., Li, B., Yu, M., Wang, J., "Hashing Based Fast Palmpoint Identification for Large-Scale Databases," *IEEE Transactions on Information Forensics & Security*, 8(5), 769-778, 2013. [Article \(CrossRef Link\)](#)
- [23] He, Kaiming, et al., "Deep residual learning for image recognition," in *Proc. of the IEEE conference on computer vision and pattern recognition*, 2016. [Article \(CrossRef Link\)](#)
- [24] Szegedy, C., Wei, L., Jia, Y., Sermanet, P., Rabinovich, A., "Going deeper with convolutions," in *Proc. of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. [Article \(CrossRef Link\)](#)



Xuebin Xu received Ph.D. degree in the department of computer science from Xi'an Jiaotong University, PR China in 2010. He has worked as a research associate professor of University of Florida, USA. He is currently as a research scientist of Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing. His research interests include machine learning, bioinformatics, biomedical image processing and information fusion. He has published over 100 papers in top international journals and conferences.



Kan Meng received the B.E degree in in the School of Computer Science, Xi'an University of Posts and Telecommunications, in 2015. He is currently working toward the M.S. degree in the School of Computer Science, Xi'an University of Posts and Telecommunications. His research interests include machine learning, bioinformatics, biomedical image processing and information fusion.



XiaoMin Xing received the B.E degree in the School of Information Engineering, Xi'an University of Arts and Sciences, in 2020. She is currently working toward the M.S. degree in the School of Computer Science, Xi'an University of Posts and Telecommunications. Her research interests are machine learning, bioinformatics, biomedical image processing and information fusion.



Chen Chen received the B.E degree in Huanggang Normal University, in 2021. She is currently working toward the M.S. degree in the School of Computer Science, Xi'an University of Posts and Telecommunications. Her research interests are machine learning, bioinformatics, biomedical image processing and information fusion.