Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-5, Issue-11, November 2023
DOI: 10.53469/jrse.2023.05(11).17

Racial Identification and Classification Based on Improved ResNet50

Shuxing Lan, Xingguo Jiang, Yang Fan, Guojun Lin

Sichuan University of Science and Engineering, Zigong, 643000, China
tonny_jiang@suse.edu.cn

Abstract: ***Face recognition research focuses on acquiring facial soft biometrics such as identity, gender, age, ethnicity, and expression. One specific area of interest is face race recognition. This study addresses the facial race recognition challenges for Asian, African, Caucasus, and Indian individuals. It proposes an improved deep learning model called TCGF-ResNet, which builds upon the ResNet50 architecture. To address data category imbalance, the model utilizes the focus loss function instead of the cross-entropy loss function. Additionally, it incorporates the GELU activation function to enhance the model's effectiveness. Furthermore, the feature extraction module introduces the convolutional attention module (CBAM) to assign weights to different regions in the face feature map, thereby enhancing attention to relevant features. To evaluate the model, it was tested on the BUPT balanced polygon dataset using transfer learning, achieving a test accuracy of 98.9% in face race recognition.***

Keywords: Face recognition, Race recognition, Deep learning, Residual network, ResNet50, Transfer learning**.**

1. Introduction

Soft biometric features, such as age, gender, expression, race, etc., have been used in combination with hard biometrics to improve face recognition performance [1]. In the field of human-computer interface (HCI), computers can provide speech recognition or present different options to users based on soft biometrics [2]. While racial categories are directly identifiable by the human eye, the same is not true for machines. Race is a major biometric trait that has major implications for racial identification in fields as diverse as law enforcement, surveillance video, visas, advertising and social media analysis. Most security agencies, such as law enforcement and intelligence, are currently leaning towards automated predictive systems where identifying information about an individual's soft biometrics can enhance the security of the system. For example, when looking for fugitives, if race can be screened first, it will reduce the scope of the search. In computer vision, racial classification using face images has received a lot of attention, mainly involving two topics: the first topic is mainly divided into black, white, Asian, European, etc. [3, 4]. For example Ahmed et al. [5] discussed the problem of detecting race in four different ethnic groups, namely Indians, Caucasians, Africans and Asians. The second theme focuses on small ethnic groups, which can be people of the same nationality. For example, Wang et al.[6] classified eight ethnic groups in China.

In order to further improve the ability to identify different races, this paper uses ResNet50 as the backbone network to propose a residual attention network model for different race recognition. The model builds a deep learning model by using large-scale data (BUPT dataset), which has large differences in pose, expression, age, and back-ground.

The improved network model is tested on this data set using the method of trans-fer learning, and compared with the original network and the classic networks. The experimental results show that the network model proposed can be competent for the problem of race recognition, and is better than the compared networks.

2. Related Work

In the past ten years, there have been a lot of racial identification research, and researchers have proposed various methods to solve racial detection and classification. For example, in 2011, Tin et al. [7] studied the linear transformation technology of PCA and LDA to extract ethnic features. In 2011, Lagree et al. [8] studied the influence of iris characteristics on racial characteristics, and proved that soft indicators can be used in mobile environments. In 2012, Xie et al. [9] used skin color features to classify Caucasians, Africans, Indians and Asians. In 2013, Ding et al. [10] used texture and shape features to identify races. The technology is divided into three stages, including FS, classification and face detection.

The facial features extracted by a typical manual design system are difficult to deal with the pictures generated under unlimited imaging conditions, which has great limitations. Therefore, in recent years, many researchers have begun to study and use deep learning methods. In 2016, Wang et al. [11] collected face datasets from MORPH-II, CMU-MultiPIE, CASIA-WebFace, CASIA-PEAL and other sources, and used convolutional neural network (CNN) to analyze Chinese and African. The Chinese are categorized and ethnically categorized as Han, Uyghur, and non-Han. In 2016, Narang et al. [12] studied race estimation using CNNs in the visible and near-infrared bands at night. In 2017, Srinivas et al. [13] used a self-designed dataset to explore the sub-ethnic classifications of China, Japan, Korea, South Asia, and the Philippines. In 2020, Greco et al. [14] used the VMER dataset, which contains more than 3 million facial images, and used the ResNet-50, VGG-Face, VGG-16 and MobileNet v2 network frameworks to assess in depth African American, East Asian Humans, Caucasian Lat-ins, and Asian Indians. In 2021, Khan et al. [15] used the DCNN method to develop a face segmentation technology, which divided facial photos into seven different categories of skin, nose, eyes, hair, back, mouth and eyebrows for labeling, and the method can also

identify races. Although these methods have the ability to identify races, there is still a lot of room for improvement in their accuracy. Therefore, in order to further improve the recognition accuracy, this paper proposes a new deep learning residual network (TCGF-ResNet) model.

3. Ethnic Category Recognition Model

3.1 Improved ResNet50

The paper regards race detection as a classification problem, and uses ResNet50 as the backbone network, and introduces the CBAM attention mechanism to improve the feature extraction ability of the model for faces, and uses the improved focus loss function to improve the problem of unbalanced dataset categories, and finally uses the Softmax classifier classifies different races, as shown in Figure 1. As shown in Figure 1, the network structure uses an initial convolutional layer, maximum pooling, four residual blocks and a CBAM attention mechanism to form a model feature extraction module. Among them, the initial convolutional layer consists of a convolutional kernel of 7×7, with a step size of 2, batch normalization, and a Gelu activation function. The size of the input picture data is 224×224×3. CBAM is added into the feature extraction module, and the channel attention module and spatial attention module are used to distribute the weight of each region in the face feature map, so as to increase the attention to effective features and compress unimportant features. 12 CBAMs are added between the 12 Identity Blocks, which can refine the intermediate feature mapping in-to a target feature that is more representative of the input. By the global maximum pooling layer and the global average pooling layer, therefore, the model prediction speed can be improved. At the same time, a focal loss function is introduced to estimate the loss of TCGF-ResNet, so that samples of different ethnic groups use different weights to enhance the generalization ability of the model and achieve the effect of optimizing the model. Finally, the feature is converted into a probability value through the SoftMax layer and output for ethnic recognition. To ensure compatibility, the input face photos need to be preprocessed at the beginning.

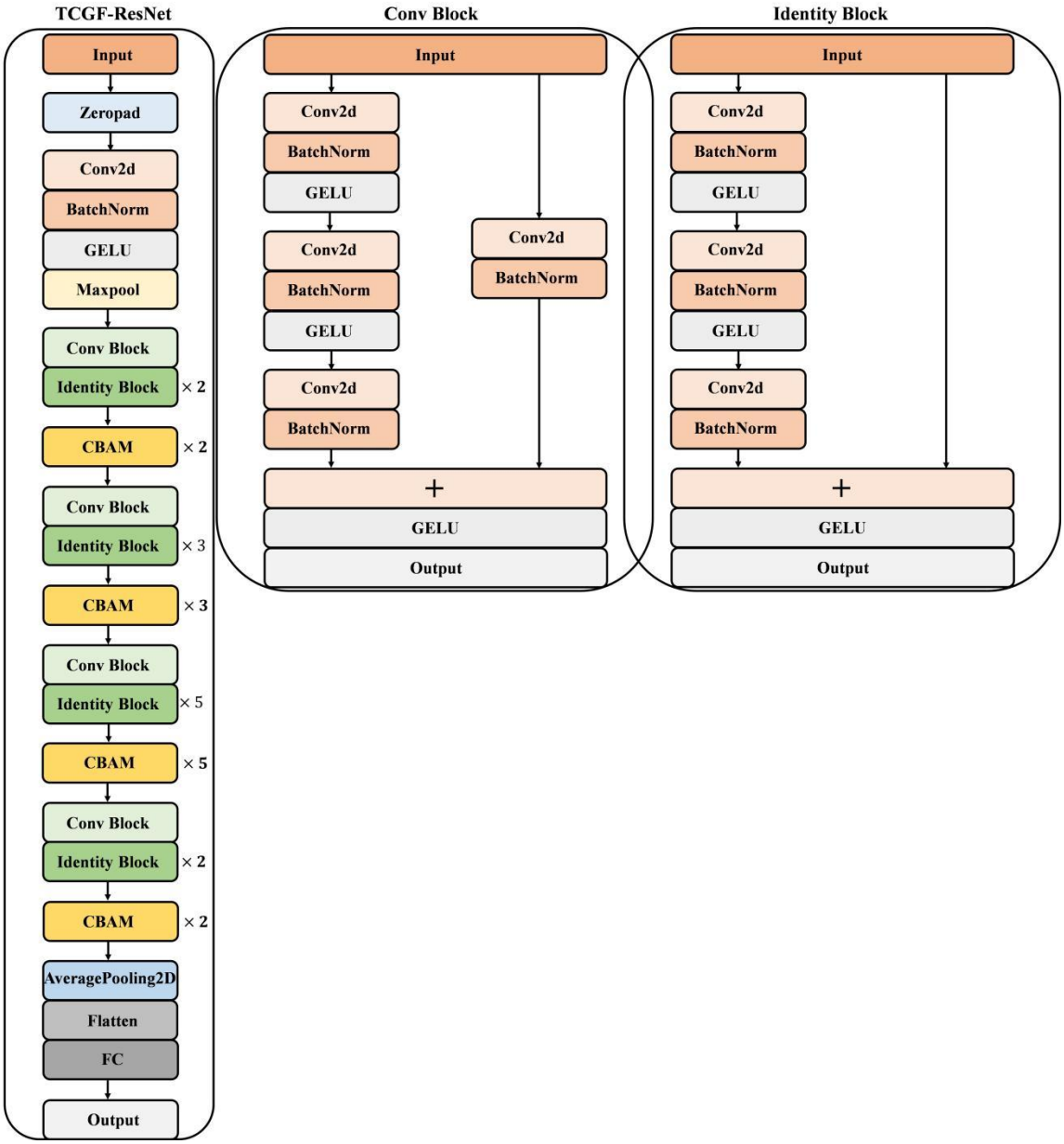


Figure 1: TCGF-ResNet network structure

GELU introduced the idea of random regularization into activation, which is a probabilistic description of neuron input, intuitively more in line with natural understanding [16]. Therefore, the experimental effect is better than the original RELU activation function. So, the paper uses GELU activation function to replace the original RELU activation function. Its calculation formula is as follows:

$$GELU(x) = xP(X \le x) = x\Phi(x) \quad (1)$$

Among them, $\Phi(x)$ is the probability function of the normal distribution, and the normal distribution N (0,1) can be simply used.

The identity residual block is concatenated at multiple inputs and outputs of the network, and there are Conv2d and Gelu operations in each layer. Therefore, in the training part of the convolutional neural network, the adjustment of the parameters of the previous layer will change the distribution of the input data of the latter layer, so it will not only increase the complexity of training to affect the training speed of the network, but also may increase the fitting risk. To solve this problem, a batch normalization layer can be added to keep the input of each layer in the same channel, so as to solve the problem of difficult network training, thereby effectively improving the convergence speed and stability of the network. Therefore, each feature extraction is performed with batch normalization before nonlinear activation.

CBAM is a lightweight framework that combines the channel attention mechanism module and the spatial attention mechanism module. The channel attention module is able to highlight the representative information provided by the image, and the spatial attention module focuses on the representative regions that contribute to the image. The operation process of CBAM is generally divided into two parts. First, the global maximum pooling and average pooling are performed on the input according to the channel, and the two one-dimensional vectors after pooling are sent to the fully connected layer operation and added to generate a one-dimensional channel attention. At the same time, channel attention is multiplied with input elements to obtain the adjusted feature map[17]. Then, the feature map is globally maximum pooled and average pooled according to space, and the two two-dimensional vectors generated by pooling are splicing and convolving to finally generate two-dimensional spatial attention, and then the spatial attention and feature map are multiplied by elements to get the final feature. The specific structure of the CBAM attention module is shown in Figure 2.

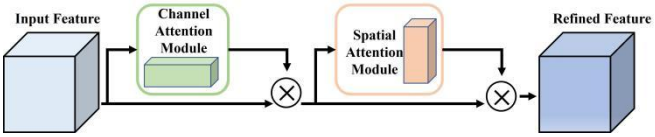


Figure 2: CBAM attention module

The channel attention module is shown in Figure 3. Each channel of the feature map is considered as a feature detector [18]. Pass the input features through two parallel MaxPool layers and AvgPool layers, and change the feature map from CHW to C11 size, and then superimpose through the multilayer perceptron module and ReLU activation function to generate two channel attention feature maps. And add the two results element by element to get finally the output result of the channel attention module through a sigmoid activation function. Its calculation formula is as follows:

$$M_C(F) = \sigma\left(MLP\left(Avgpool(F)\right)MLP\left(MaxPool(F)\right)\right) \quad (2)$$

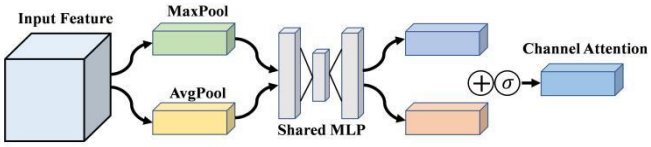


Figure 3: Channel attention module

The spatial attention module is shown in Figure 4. Applying a pooling operation along the channel axis was proved to be very effective in highlighting informative regions [19]. The channel attention module outputs two 1HW feature maps through maximum pooling and average pooling, and performs channel splicing operations on the two feature maps. Then, the channel dimension is reduced to 1 by a convolutional layer with a kernel size of 7, followed by a sigmoid activation to obtain the feature map of the spatial attention module. Its calculation formula is as follows:

$$M_s(F) = \sigma\left(f^{7\times7}\left(\left[AvgPool(F);MaxPool(F)\right]\right)\right) \quad (3)$$

In the BUPT data set, the recognition of each type of race will only have a small number of pictures with feature recognition confusion, and the number is much smaller than the correct pictures. Therefore, in order to improve the recognition accuracy of a small number of pictures, the Focal loss function is improved to avoid the problem that the Cross Entropy loss function ignores other categories while optimizing one category, and avoid the overfitting phenomenon caused by the model during training [20]. The Focus loss function is improved, and an adjustment factor is added on the basis of the original Focal loss function to adjust the weight of difficult and easy samples during model training, thereby alleviating the problem of difficult-to-classify samples. Its calculation formula is as follows:

$$FL(p,\hat{p}) = -\left(\alpha(1-\hat{p})^{\gamma}\, plog(\hat{p}) + (1-\alpha)\hat{p}^{\gamma}(1-p)\log(1-\hat{p})\right) \quad (4)$$

Among them, α and $1-\alpha$ are used to control the proportion of positive and negative samples respectively, and their value range is [0,1]. The parameter γ is a focusing parameter, which can reduce the weight of samples that are easy to classify and increase the attention of samples that are not easy to classify, thereby reducing the loss. Its γ [0,+∞].

3.2Transfer Learning

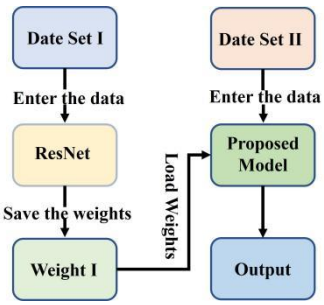


Figure 5: Migration learning process

Transfer learning is a method for solving similar unknown problems in different fields, which solves the shortcomings of deep learning that require a large number of samples to train the model. The use of transfer learning can effectively reduce the dependence of the model on data, shorten the calculation time, and improve the efficiency of the model. Therefore, in order to further reduce the impact of data on model performance and speed up the convergence speed of the model, transfer learning is used for training. The specific training process is shown in Figure 5. In the BUPT dataset, dataset I consists of 48,000 photos that are distinct from the final training, while dataset II consists of 48,000 images that do not intersect with dataset I. Train the ResNet50 model on dataset I to obtain a pre-trained weight, and then load the pre-trained weight on the TCGF-ResNet model for training to achieve the purpose of transfer learning.

4. Experiment and Results

4.1 Data Preprocessing

In 2019, Wang et al. [21] released the BUPT Equalized Face dataset. The dataset, developed by Beijing University of Posts and Telecommunications, contains a total of 1.3 million images of different races. Each race has about 320K+ images. The photos were taken from 28,000 people, 7,000 of each race. All of these images were captured in an unconstrained environment with variations in expression, age, pose and background. Figure 6 shows the BUPT samples for each category separately.

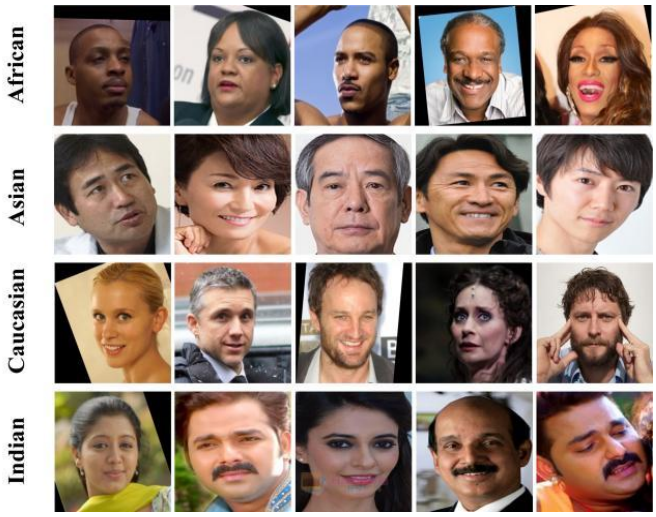


Figure 6: Example of BUPT dataset

Computer resources limited the use of the entire dataset in the experiments. Therefore, the dataset is downsampled to fit the capacity of the computer. A total of 48,000 images are selected for training, and the training set consists of 40,000 images, 10,000 for each class. The test set consists of 8000 images, 2000 for each class. The selected pictures are mutually disjoint.

Since images appear in different poses, they need to be preprocessed to ensure uniformity, thereby improving the recognition ability of the model. Image preprocessing consists of five different sub-processes, namely

- Face recognition and alignment
- resize image
- RGB to grayscale conversion
- average centered
- Normalize to the range [0, 1]

The images undergo a 2D alignment process in which the steps involved are as follows:

1) Using dlib to recognize facial landmarks [22], estimate the angle between the line passing through the eyes and the horizon.

2) The image is rotated around the center of the eye by the angle determined in step 1.

3) Crop the target image around the face area so that the eyes appear at 32% of the total width from the edges and 38% of the height from the top. Finally, the resulting image is scaled to 224*224 pixels and converted to grayscale.

4) Clipping faces that yaw higher on the bottom area.

4.2 CBAM Impact on the ResNet50 model

Through the channel attention module and the spatial attention module, respectively, CBAM highlights the representative data the image provides and the representative regions that contribute to the image. In order to give each area in the face feature map weights and maximize attention to useful characteristics while compressing unnecessary features, CBAM is introduced to the feature extraction module. The paper analyzes the effect of CBAM on the performance of the ResNet50 model, and its loss curve is shown in Figure 7.

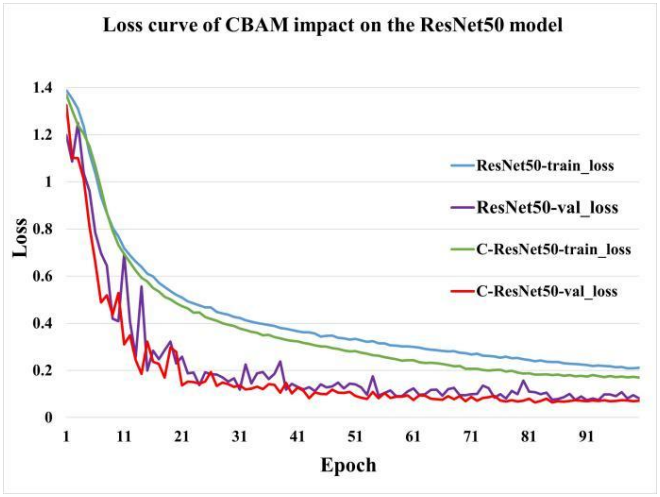


Figure 7: Loss curve of CBAM influence on the performance of ResNet50 model

It can be seen from FIGURE 7 that adding CBAM to the ResNet50 model can improve the convergence speed of the model and make the model convergence more stable. Both the loss of the training set and the loss of the test set are lower than those of the original ResNet model, making the model more effective on face race recognition.

4.3 Choice of Activation Function and Optimizer

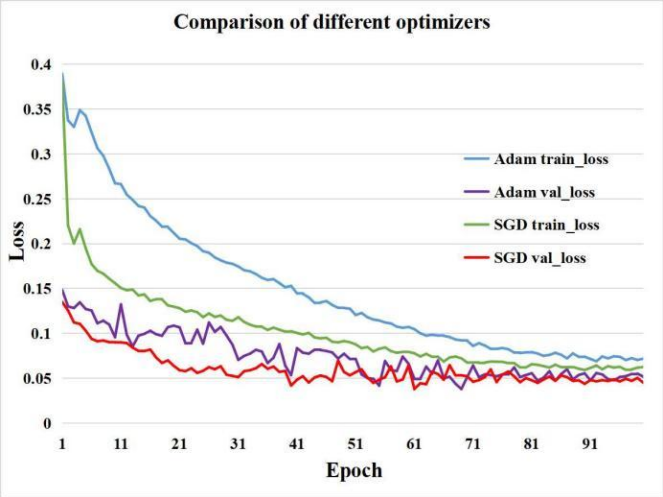


Figure 8. Adam and SGD training model loss curve

Activation functions and optimizers are important components of convolutional neural networks and affect the performance of the model. The paper analyzes and compares the performance of the two optimizers, SGD and Adam. The loss curves are shown in Figure 8. The initial learning rate of the two optimizers is set to 0.01.

It can be seen from Figure 8 that the SGD optimizer reduces the loss faster and the effect is better. Therefore, SGD is chosen as the optimizer of the model.

Table 1 shows the performance of GELU and RELU activation functions with SGD as the optimizer. It can be seen from Table 1 that after 100 rounds, in the case of GELU activation, the loss and recognition rate of the test set are significantly better than RELU, so GELU is chosen as the activation function.

Table 1: Performance of different activation functions on the model

Activation function name		Training set		Test set	
		Loss	Accuracy (%)	Loss	Accuracy (%)
RELU	100 Epochs	0.070	97.2	0.041	98.7
GELU	100 Epochs	0.078	96.8	0.038	98.9

4.4 Comparative Analysis of Transfer Learning

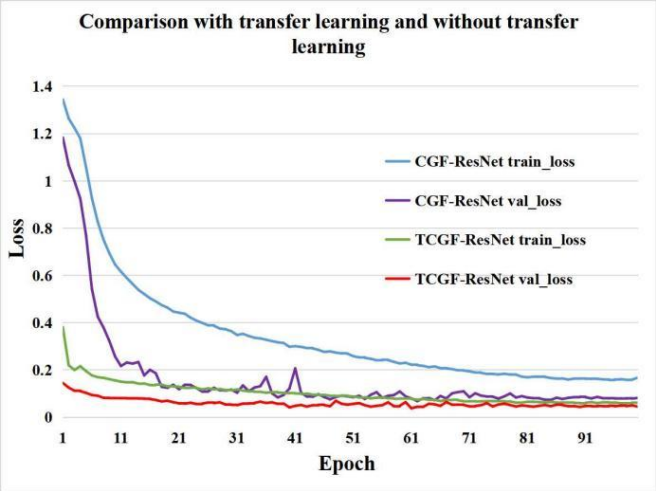


Figure 9: Comparison curve of the impact of transfer learning on model performance

In order to verify the impact of the proposed transfer learning method on model performance, the CGF-ResNet model is used as the basis to compare transfer learning with non-transfer learning, and its loss curve is shown in Figure 9.

From Figure 9, it can be seen that the training method of transfer learning makes the convergence speed of the model faster, the fitting effect is better, the accuracy rate is higher, and the training loss is reduced, so that the classification effect is better.

4.5 Comparative Analysis of Loss Functions

To improve the problem of unbalanced dataset categories, the Cross Entropy loss function is replaced with an improved Focal loss function. The loss curves of the two loss functions are shown in Figure 10.

It can be seen, from Figure 10, that the improved focal loss function makes the test set loss of the model slightly better than the cross-entropy loss function, so the Focal loss function is chosen as the loss function of the model.

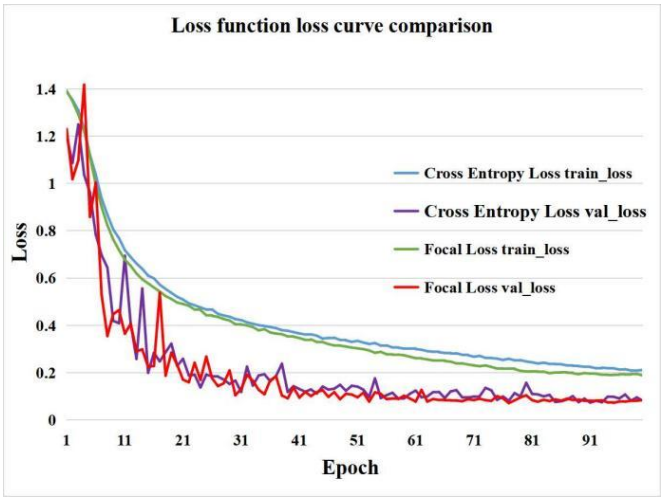


Figure 10: The influence curve of Focal loss function and Cross Entropy loss function on model performance

4.6 TCGF-ResNet Model Test

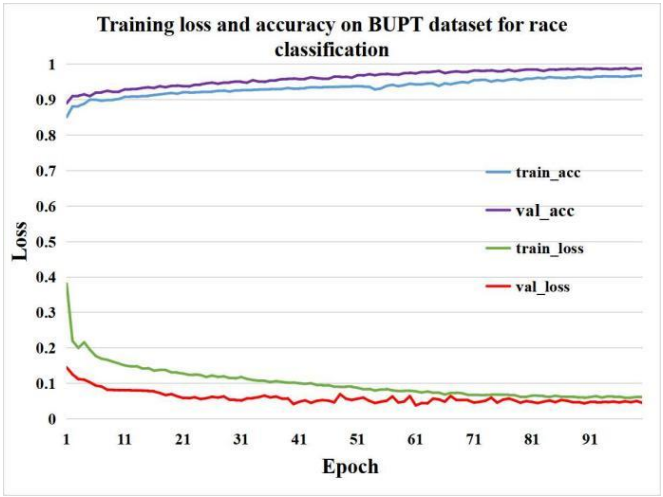
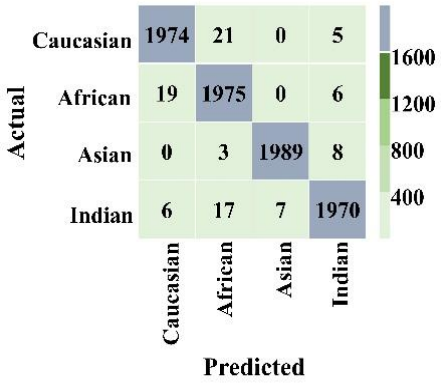
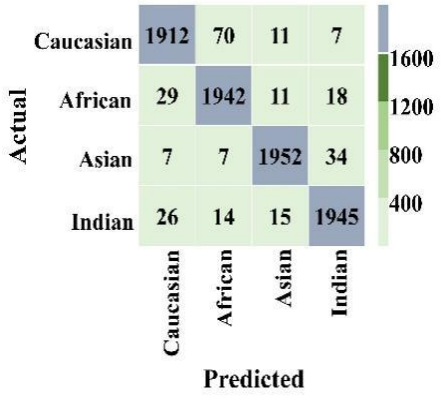


Figure 11: Model training loss and accuracy curve

Using the GELU activation function and the optimizer SGD, the Focal loss function is introduced to estimate the loss of TCGF-ResNet, where is set to 0.25 and is set to 2, and the CGF-ResNet model is trained for 100 rounds by transfer

learning. Figure 11 shows the variation of training set and test set loss and recognition rate throughout the training. The model achieved a recognition rate of 96.8% on the training set and 98.9% on the test set after 100 rounds.

Table 3: TCGF-ResNet compared to other methods tested on the BUPT dataset

Method	Parameter Values	Acc	Recall	Confusion Matrix
Proposed Model	TrainSet:40000 × 224 × 224 × 3 Val/Test Set:8000 × 224 × 224 × 3 Optimizer: SGD (lr=0.01) Epochs:100	98.85%	98.85%	
(Ahmed et al.,2020)	TrainSet:400000 × 40 × 40 × 1 Val/Test Set:32000 × 40 × 40 × 1 Optimizer: Adam Epochs:40	97%	97%	

4.7 Network Comparison

To further verify the performance, the popular network models ResNet50, Mobilenet and VGG16 were tested and compared with the improved network models. TCGF-ResNet has a network model verification accuracy of 98.9% and a verification loss of 0.038. The ResNet50 network has a verification accuracy rate of 97.2% and a verification loss of 0.080; The Mobilenet network has a 95.4% verification accuracy rate and a 0.129 verification loss; The VGG16 network has a 96.6% verification accuracy rate and a 0.102 verification loss. The experimental data are shown in Table 2. The results show that the proposed model has the highest verification accuracy and the lowest verification loss.

When testing the proposed model on the BUPT dataset, the average accuracy rate can reach 98.85%, and the precision and recall rate of all classes are higher than 98.5%, which shows that the model has unbiased learning for all classes. Comparing the TCGF-ResNet model with other methods tested on the BUPT dataset. The results are shown in Table 3.

The confusion matrix of different ethnic categories obtained through the TCGF-ResNet model is shown in Table 3. The model identified 1974 instances as Caucasian, 1975 instances as African, 1989 instances as Asian, and 1970 instances as Indian. The model recognition accuracy rate is 98.9%, and the recall rate is 98.85%. By comparing with other methods, it can be seen that the proposed network model outperforms other methods on the BUPT dataset, achieving the best accuracy.

Table 2: Network Model Performance Comparison

Network model	Epoch	Training set		Test set	
		Loss	Accuracy (%)	Loss	Accuracy (%)
TCGF-ResNet	100	0.078	96.8	0.038	98.9
ResNet50	100	0.226	90.9	0.080	97.2
Mobilenet	100	0.248	90.2	0.129	95.4
VGG16	100	0.206	91.9	0.102	96.6

5. Conclusion

Face race recognition is considered as a classification problem. Based on the original ResNet50 model, a new deep learning convolutional neural network (TCGF-ResNet) is proposed. The TCGF-ResNet network model is compared with the popular model on the BUPT dataset. The network model ResNet50, Mobilenet and VGG16 are compared experimentally. Also, compared to other methods tested on the BUPT dataset. The final results show that the proposed model achieves the best performance with a verification accuracy of 98.9%, and the superiority of the TCGF-ResNet model in ethnic recognition is demonstrated through rigorous experimental studies. Therefore, the race identification and classification method based on the improved ResNet50 proposed in the paper can be used for race recognition and provide a reference for subsequent race recognition and classification methods. Today, the growing mixed-race population is further narrowing the borders, and the identification of less diverse races is expected to be further improved.

Acknowledgement

This paper was supported by the Scientific Research Foundation of Sichuan University of Science and Engineering under Grant 2019RC12.

References

[1] Jain, A.K.; Dass, S.C.; Nandakumar, K., In Soft biometric traits for personal recognition systems, International conference on biometric authentication, 2004; Springer: 2004; pp. 731-738.
[2] Hu, Y.; Fu, Y.; Tariq, U.; Huang, T.S., In Subjective experiments on gender and ethnicity recognition from different face representations, International Conference on Multimedia Modeling, 2010; Springer: 2010; pp. 66-75.
[3] Masood, S.; Gupta, S.; Wajid, A.; Gupta, S.; Ahmed, M., Prediction of human ethnicity from facial images using neural networks. In Data Engineering and Intelligent Computing; Springer: 2018; pp. 217-226.
[4] Anwar, I.; Islam, N.U., Learned features are better for ethnicity classification. arXiv preprint arXiv:1709.07429 2017.
[5] Ahmed, M.A.; Choudhury, R.D.; Kashyap, K., Race estimation with deep networks. J King Saud Univ-Com 2020.
[6] Wang, C.; Zhang, Q.; Duan, X.; Gan, J., Multi-ethnical Chinese facial characterization and analysis. Multimed Tools Appl 2018, 77, (23), 30311-30329.
[7] Tin, H.H.K.; Sein, M.M., Race identification for face images. ACEEE Int. J. Inform. Tech 2011, 1, (02), 35-37.
[8] Lagree, S.; Bowyer, K.W., Ethnicity Prediction Based on Iris Texture Features. MAICS 2011, 14, (1), 225-230.
[9] Xie, Y.; Luu, K.; Savvides, M., In A robust approach to facial ethnicity classification on large scale face databases, 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2012; IEEE: 2012; pp. 143-149.
[10] Ding, H.; Huang, D.; Wang, Y.; Chen, L., In Facial ethnicity classification based on boosted local texture and shape descriptions, 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2013; IEEE: 2013; pp. 1-6.
[11] Wang, W.; He, F.; Zhao, Q., In Facial ethnicity classification with deep convolutional neural networks, Chinese Conference on Biometric Recognition, 2016; Springer: 2016; pp. 176-185.
[12] Narang, N.; Bourlai, T., In Gender and ethnicity classification using deep learning in heterogeneous face recognition, 2016 International Conference on biometrics (ICB), 2016; IEEE: 2016; pp. 1-8.
[13] Srinivas, N.; Atwal, H.; Rose, D.C.; Mahalingam, G.; Ricanek, K.; Bolme, D.S., In Age, gender, and fine-grained ethnicity prediction using convolutional neural networks for the East Asian face dataset, 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 2017; IEEE: 2017; pp. 953-960.
[14] [Greco, A.; Percannella, G.; Vento, M.; Vigilante, V., Benchmarking deep network architectures for ethnicity recognition using a new large face dataset. Mach Vision Appl 2020, 31, (7), 1-13.
[15] Khan, K.; Ali, J.; Uddin, I.; Khan, S.; Roh, B., A Facial Feature Discovery Framework for Race Classification Using Deep Learning. arXiv preprint arXiv:2104.02471 2021.
[16] Hendrycks, D.; Gimpel, K., Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415 2016.
[17] Tian, Q.; Song, Q.; Wang, H.; Hu, Z.; Zhu, S., In Verification Code Recognition Based on Convolutional Neural Network, 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2021; IEEE: 2021; pp. 1947-1950.
[18] Zeiler, M.D.; Fergus, R., In Visualizing and understanding convolutional networks, European conference on computer vision, 2014; Springer: 2014; pp. 818-833.
[19] Zagoruyko, S.; Komodakis, N., Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer. arXiv preprint arXiv:1612.03928 2016.
[20] Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P., In Focal loss for dense object detection, Proceedings of the IEEE international conference on computer vision, 2017; 2017; pp. 2980-2988.
[21] Wang, M.; Deng, W.; Hu, J.; Tao, X.; Huang, Y., In Racial faces in the wild: Reducing racial bias by information maximization adaptation network, Proceedings of the ieee/cvf international conference on computer vision, 2019; 2019; pp. 692-702.

Author Profile

Shuxing Lan received his B.S. degree from Sichuan University of Science and Engineering in 2021. He is currently a graduate student in the School of Automation and Electric Information, Sichuan University of Science and Engineering, Sichuan, China. His current research interests include Image Processing and Deep Learning.

Xingguo Jiang is currently working as an Associate professor in the School of Automation and Electric Information, Sichuan University of Science and Engineering, Sichuan, China. Prior to that, he was an Associate professor in the School of Information and Communication, Guilin University of Electronic Technology, Guangxi, China. He received the Ph.D. degree in 2007 from the Institute of Optics and Electronics, Chinese Academy of Sciences,

Chengdu, China and the MS degree in 2003 from Chongqing University, Chongqing, China. His current research interests include Image Processing, Intelligent Information Processing, and Deep Learning.

Yang Fan received his B.S. degree from Sichuan University of Science and Engineering in 2021. He is currently a graduate student in the School of Automation and Electric Information, Sichuan University of Science and Engineering, Sichuan, China. His current research interests include Image Processing and Deep Learning.

Guojun Lin graduated from Zhejiang University of Technology in July 2001. He graduated from Southwest Jiaotong University with a master's degree in March 2008. September 2008, Shenzhen Tianpai Electronics Co., LTD. Software engineer; In December 2014, he graduated from University of Electronic Science and Technology of China. Jan. 20015-present, Lecturer, Sichuan University of Science and Engineering.