IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 February 2023, accepted 13 March 2023, date of publication 16 March 2023, date of current version 23 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3257986

RESEARCH ARTICLE

A Real-Time Face Detection Method Based on Blink Detection

HUI QI[1], CHENXU WU[1], YING SHI[1,2], XIAOBO QI[1], KAIGE DUAN[1], AND XIAOBIN WANG[1]
[1]School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China
[2]School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China

Corresponding author: Hui Qi (qihui@tynu.edu.cn)

This work was supported in part by the Basic Research Program (Free Exploration) of Shanxi Province under Grant 20210302123334, and in part by the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi under Grant 2021L443.

ABSTRACT Face anti-spoofing refers to the computer determining whether the face detected is a real face or a forged face. In user authentication scenarios, photo fraud attacks are easy to occur, where an illegal user logs into the system using a legitimate user's picture. Aiming at this problem and the influence of illumination in real-time video face recognition, this paper proposes a real-time face detection method based on blink detection. The method first extracts the image texture features through the LBP algorithm, which eliminates the problem of illumination changes to a certain extent. Then the extracted features are input into the ResNet network, and the facial feature extraction is enhanced by adding an attention mechanism is added to enhance the face feature extraction. Meanwhile, the BiLSTM method is used to extract the temporal characteristics of images from different angles or at different times to obtain more facial details. In addition, the fusion of local and global features is realized by SPP pooling, which enriches the expression ability of feature maps and improves detection accuracy. Finally, the eye EAR value is calculated by the face key point detection technology to achieve face anti-spoofing, and then the real-time face recognition against fraud is realized. The experimental results show that the algorithm proposed in this paper has good accuracy on NUAA, CASIA-SURF and CASIA-FASD datasets, which can reach 99.48%, 98.65% and 99.17%, respectively.

INDEX TERMS Attention mechanism, BiLSTM, face recognition, SPP.

I. INTRODUCTION

With the continuous development of computer image processing technology, face recognition, as a means of identity verification, is widely used in all walks of life for its flexibility and convenience. Live real-time face comparison and monitoring are required in various life scenes such as online learning, driving test, and fatigue driving.

Reference [1] uses a combination of LBP and CNN to extract features through a scale-invariant LBP histogram. Then the ULBP is used for feature dimension reduction and is combined with a convolutional neural network for real-time face recognition, which has good real-time performance and effectiveness. However, when there is a large-scale deflection of the face angle, the model cannot perform the detection and recognition work well. In reference [2], the AdaBoost algorithm is used to automatically detect face images, and the circular LBP operator is used to extract the texture features of the preprocessed face image, so as to avoid the influence of light changes on the gray level. However, the correct face recognition rate of this method is not too high. Reference [3] designs a real-time face detection and tracking algorithm based on MTCNN and applies it to the ARM platform. The method has a faster running speed on embedded systems, but the recognition rate on datasets such as CFP-FP and Mega Face is still not high.

Due to the widespread application of biometrics, biometric spoofing attacks have also become a threat to human security. A spoofing attack means that an attacker forges the biometrics of a legitimate user in an attempt to enter the system as an illegal identity. In the field of face recognition, attackers can easily obtain the face information of legitimate users on the Internet. Therefore, the technology of face anti-spoofing comes into being. In the process of face recognition, it can

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

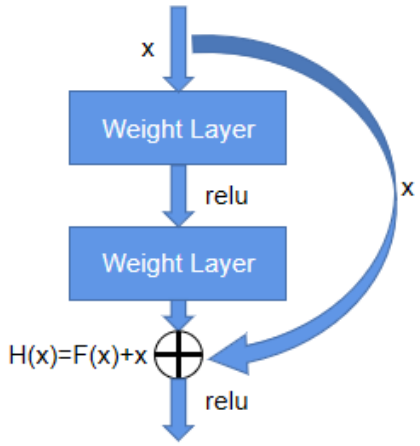
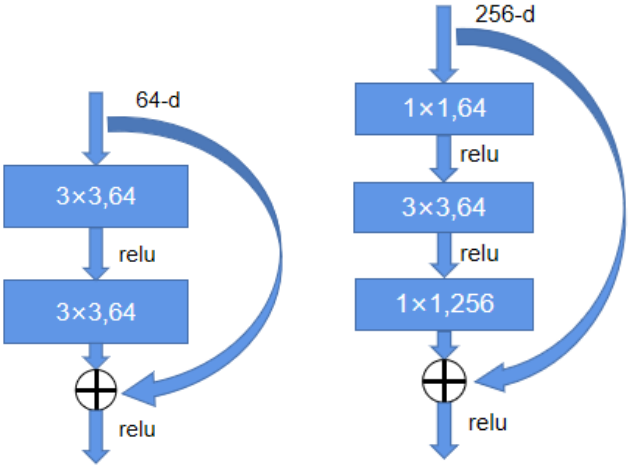


FIGURE 1. Residual network unit.



(a) Building Block (b) Bottleneck Block

FIGURE 2. ResNet residual network structure.

not only detect the face of the target but also judge the living. When the system detects a living body, the face recognition results are effective.

In reference [4], RGB image, depth image, and infrared image of faces are respectively input into three identical residual network structures to extract features, and then the fusion is carried out to distinguish whether it is alive or not. However, with the change of attack mode and external environment, the extracted features may change, resulting in misjudgment. In reference [5], LBP and multi-layer DCT are applied to face images to obtain LBP-MDCT features, and then part of the face images are input into CNN to obtain CNN features. Finally, the two features are input into the support vector machine to obtain the classification results. It can effectively combine the local and global information of the image to achieve a better recognition result. In the literature [6], face anti-spoofing has changed from the traditional two-class problem to a multi-class problem. A dual stream feature fusion network model has been constructed to fuse the feature vectors extracted from two different color spaces, RGB and YCrCb to improve the feature representation ability of the network model. However, due to the differences in brightness and illumination information of face images, the recognition effect is slight poor.

In this paper, a real-time face detection method based on blink detection called LBAS_Resnet50 (L stands for LBP, B for BiLSTM, A for Attention Mechanism, and S for SPP) is proposed considering the influence of lighting, posture, background environment, and other factors on real-time face recognition, as well as the fraud attack problem in the process of face recognition. Firstly, LBP is used to extract texture features to reduce the influence of illumination changes. Secondly, based on ResNet, the important facial features are extracted through the Attention mechanism, and BiLSTM is used to capture the sequence features to further improve the accuracy of feature selection. Thirdly, the SPP pooling technology fully considers the local and global information to enhance the network generalization ability. Finally, the face key point detection technology is used to determine whether the target is alive by calculating the EAR value of the eyes, so as to achieve the anti-fraud function. The accuracy of this can reach 99.48%, 98.65%, and 99.39% on NUAA, CASIA-SURF, and CASIA-MFSD, respectively.

The proposed model has the following contributions:

1.The method LBAS_Resnet50 first extracts the texture features of the image by LBP, which eliminates the problem of illumination variation to some extent.

2.The proposed model LBAS_Resnet50 uses ResNet50 as the base network to enhance the face feature extraction by adding an attention mechanism.

3.The LBAS_Resnet50 model uses the BiLSTM method to extract temporal features of images from different angles or different times. In addition, the fusion of local features and global features is achieved by SPP pooling, which enriches the expression capability of the feature map.

4.The eye EAR value is calculated by face key point detection technique to achieve face anti-spoofing.

II. REVIEW OF RELATED ALGORITHMS

A. RESIDUAL NEURAL NETWORK

With the continuous development of CNN [7], people continue to increase the number of convolutional layers in order to obtain deep-level features. However, it is not always feasible to enhance the learning ability of the network by increasing the number of network layers. Because when the number of network layers reaches a certain depth, the problem of vanishing random gradient will occur if the number of network layers is increased. At the same time, it will also lead to a decrease in the accuracy of the network. In order to solve this problem, the residual neural network (ResNet) came into being.

The key to the ResNet network is the residual unit in its structure, as shown in Figure 1 below. The residual network unit includes cross-layer connections, and the curves in the figure can directly pass the input across layers and perform the equivalent mapping. Then, add the result of the

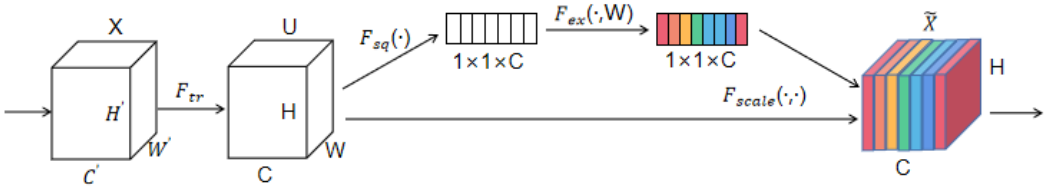


FIGURE 3. SENet module.

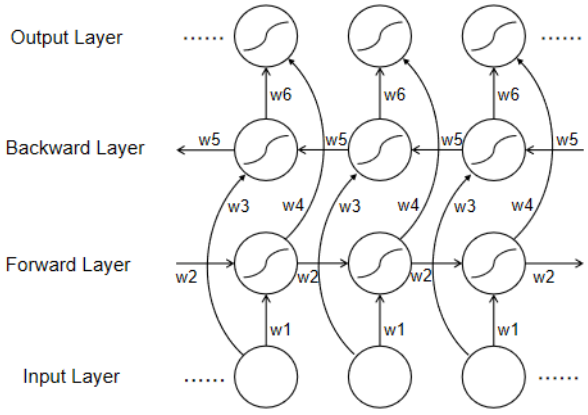


FIGURE 4. BiLSTM network structure diagram.

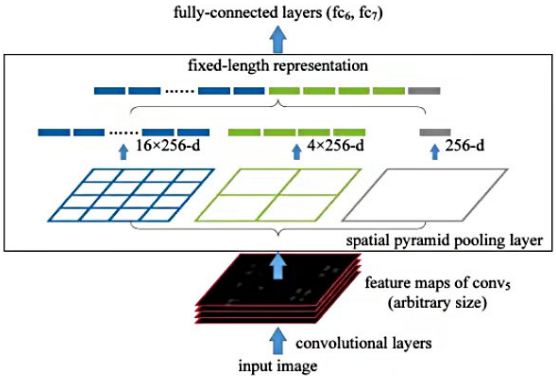


FIGURE 5. SPP model.

convolution operation. Assuming that the input image is x, the output is H(x), and the output after convolution in the middle layer is a nonlinear function of F(x), then the final output is H(x) = F(x) + x. Such an output can still undergo nonlinear transformations.

There are two main types of residual structures in the ResNet network. One is the Building Block, which is suitable for shallower ResNet networks. The other is Bottleneck, which is suitable for ResNet networks with deeper layers. Its structure is shown in Figure 2. The method proposed in this paper uses the ResNet50 network structure, in which the residual network structure is Bottleneck.

B. ATTENTION MECHANISM

The attention mechanism [8] stems from the study of human vision. In cognitive science, due to bottlenecks in information processing, humans selectively focus on a portion of all information while ignoring other visible information. Therefore, the attention mechanism mainly determines which part of the input the network needs to focus on and allocates limited information processing resources to the important part. In deep learning, the attention mechanism can achieve with the help of weight vectors.

In this paper, we use Squeeze-and-Excitation Network (SENet), which is a channel attention mechanism that can capture more important information about face features. Its structure is shown in Figure 3. The transformation operation is actually a standard convolution operation. An important operation of SENet is the Squeeze operation (F_{sq}), which performs a global average pooling operation on the input features, and compresses the H × W × C features into a 1 × 1 × C size. Another important operation is the Excitation operation (F_{ex}), which performs two full connection operations on the result of the Squeeze operation, then uses Sigmoid activation to obtain the weight matrix, and finally multiplies (F_{scale}).

C. BI-DIRECTIONAL LONG SHORT-TERM MEMORY

The full name of LSTM is Long Short-Term Memory [9], which is a type of RNN. Due to its design characteristics, LSTM is very suitable for modeling time series data. BiLSTM is the abbreviation of Bi-directional Long Short-Term Memory, which is composed of forwarding LSTM and backward LSTM. Both are often used to model contextual information. The BiLSTM network structure diagram is shown in Figure 4, w_1-w_6 represents 6 shared weights.

In this paper, by adding BiLSTM to extract bidirectional sequence features based on images, the amount of information available to the network model is increased, the context information of the algorithm is improved, and the accuracy of video face recognition is improved.

D. SPATIAL PYRAMID POOLING

As we know, basically all CNNs [10] require a fixed input data size. But when dealing with pictures, most of them are of different sizes and have different pixel values. For the same batch of data, if it has to be processed into images of the same size after a certain amount of cropping, there may be some problems. For example, when some areas are cropped, there will be repetitions, which invisibly increases the weight of the area. This paper uses Spatial Pyramid Pooling (SPP) to solve the problem that the input image size of CNN must be fixed so that the input image aspect ratio and size can be arbitrary.

FIGURE 6. 68-point feature map of human face.

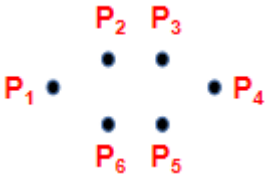


FIGURE 7. Eye feature annotation points.

FIGURE 8. Example of NUAA dataset.

SPP [11] pooling performs three convolution operations on the input feature maps, as shown in Figure 5. The following figures respectively perform 4×4, 2×2, and 1×1 convolution operations. These output feature maps are then flattened into a one-dimensional array. Then, these one-dimensional arrays are concatenated. Finally, the feature map obtained by splicing is sent to the fully connected layer.

E. LOCAL BINARY PATTERN

Local Binary Pattern (LBP) [12] is an operation used to describe the local texture features of images. It has significant advantages such as rotation invariance and grayscale invariance. The original LBP operator is defined in a 3×3 window, which takes the center pixel of the window as the threshold and compares it with the gray value of the adjacent 8 pixels. If the surrounding pixel value is greater than the center pixel value, the position is marked as 1, otherwise, it is marked as 0. This results in an 8-bit binary number. Usually, it is also converted into decimal, that is, LBP code, a total of 256 kinds. We use this value as the LBP [5] value of the center pixel of the window to reflect the texture information of this 3×3 area. Its calculation formula is as follows.

$$LBP(x_c, y_c) = \sum_{p=1}^{8} s(I(p) - I(c)) * 2^p \tag{1}$$

Among them, p represents the p-th pixel in the window except for the center pixel. I(c) represents the gray value of the center pixel. I(p) represents the gray value of the p-th pixel in the field. The formula for s(x) is as follows:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & otherwise \end{cases} \tag{2}$$

F. FACE KEY POINT TECHNOLOGY

This paper selects the 68 feature point model of Dlib face detection, and the face feature map is shown in Figure 6. In the process of detecting faces with 68 feature points [13], each eye is represented by six feature points, as shown in Figure 7. We numbered the six feature points clockwise from the position of the left corner of the eye and calculated the Eye Aspect Ratio (EAR). When the human eye is opened, the EAR fluctuates up and down a certain value, and when the human eye is closed, the EAR decreases rapidly and theoretically will be close to zero. EAR is calculated as follows.

$$EAR = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2||P_1 - P_4||} \tag{3}$$

III. THE ALGORITHM OF THIS PAPER

A. EXPERIMENTAL ENVIRONMENT AND PREPROCESSING

The experiments are carried out on the platform of i7-6000CPU, with 3.40 GHz and 32G memory. Using Python, matrix computing NumPy. The backend uses TensorFlow to develop the improved ResNet network.

Before data training, images need to be preprocessed. The frontal_face_detector built-in dilb is used to detect faces and adjust the image to the size of 224×224. The original data is divided into the training set and test set according to the probability of 7:3, and labels are divided into real and fraudulent faces.

B. DATASETS

This paper adopts the NUAA dataset, CASIA-SURF dataset [14], and CASIA-MFSD dataset for experiments. The NUAA dataset is a photo-printing fraud dataset that contains images of 15 individuals with 500 images of each individual. Its resolution is 640×480, as shown in Figure 8, the upper part is the real face, and the lower part is the fraudulent face.

CASIA-SURF contains a large number of datasets of multiple forms (RGB, Depth, IR) of people of different ages. It contains a total of 1,000 different targets and has 21,000 videos. Among them, the RGB image resolution is 1280×720, and the Depth and IR are 640×480. Each sample will record a real video and six attack videos. Types of attacks include covering areas such as the eye, nose, mouth, etc. As shown in Figure 9, the upper part is a real face, and the lower part is a fraudulent face.

CASIA-FASD is a face anti-fraud database containing video and photo prints. It was collected using cameras of three different qualities, and 50 volunteers took part in the data collection. Each volunteer recorded 12 videos at different resolutions and lighting conditions. The dataset has a total

FIGURE 9. Example of the CASIA-SURF dataset.

FIGURE 10. Example of the CASIA-FASD dataset.

of 600 video recordings. The sample picture is shown in Figure 10, the upper part is a real face, and the lower part is a fraudulent face.

C. EVALUATION INDICATORS

This paper selects the accuracy rate for evaluation. Accuracy [15] is our most common evaluation metric and easy to understand. It is the number of pairs of samples divided by the total number of samples. Generally speaking, the higher the accuracy, the better the classifier. Its calculation method is as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP is the number of predicted positive classes as positive classes, FN is the number of predicted positive classes as negative classes, FP is the number of predicted negative classes as positive classes, and TN is the number of predicted negative classes as negative classes.

For better evaluation, this paper also selects the half-error rate HTER as the evaluation index for face anti-spoofing. The lower the HTER index value, the better the performance of the algorithm model. The calculation formula is as follows.

$$HTER = \frac{FAR + FRR}{2} \tag{5}$$

Among them,

$$FAR = \frac{FP}{FP + TN} \tag{6}$$

$$FRR = \frac{FN}{FN + TP} \tag{7}$$

FAR is the false acceptance rate, and FRR is the false rejection rate.

D. ALGORITHM MODEL

The proposed algorithm network model LBAS_Resnet50 is shown in Figure 11. First, preprocess the image, using the frontal_face_detector in the dlib library to detect the face, and adjust the image size to 224 × 224. Since the LBP features of real face images and fraudulent face, images are different, and the LBP algorithm has the advantages of rotation invariance and gray invariance, which effectively avoids the influence of illumination changes, we extract the texture features of the input images through the LBP algorithm. Then the extracted texture features are input into the network model based on the ResNet50 network. Compared with other network architectures, ResNet solves the problem of vanishing gradients caused by deepening network layers. In the improved ResNet50 network called Resnet50-BAS, BiLSTM is firstly added to extract the bidirectional feature information of the image, to better consider the local and global features. Then, the classical channel attention mechanism SeNet is added, which gives higher weight to important features, so that we can recognize the important features of pictures more clearly. Finally, SPP pooling is used in the network to improve the robustness of the model.

In the LBAS_Resnet50 model, after several experiments, we found that the experimental results are not good when the threshold is less than 0.9. The experimental results fluctuate a lot when the threshold is larger than 0.9, and the results are not good. Therefore, we chose 0.9 as the threshold value. In addition, before performing LBP texture feature extraction, we first performed pre-processing by cropping the original image to 224 × 224 size, so that the texture image after LBP algorithm is also 224 × 244 size. Therefore, the texture image will look slightly larger than the original image. The structural model of ResNet50-BAS is shown in Figure 12. Part of the LBP texture map (selected from the NUAA dataset example) is shown in Figure 13, the left is the texture map of the real face, and the right is the texture map of the photo.

According to the above network model, the selected dataset is trained, and then the trained model is used for real-time anti-fraud face detection. First, real-time face images are extracted from the camera. Then the extracted image is fed into the trained model to get the classified face image. Finally, in order to improve the effect of living detection, the classified images are further detected by eye blink to further judge whether they are living, and the final conclusion is drawn.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL RESULTS

LBAS_Resnet50 model uses the Adam optimizer, which has fast convergence and simple computation. The learning rate is 0.001. At the same time, the cross entropy loss function is used to improve the accuracy of the model.

The model proposed in this paper is trained for 85 epochs on the NUAA dataset, 80 epochs on the CASIA-SURF dataset, and 80 epochs on the CASIA-FASD dataset, with batches of 128. After training, the accuracy rates can reach 99.48%, 98.65%, and 99.39%, respectively, and the loss function gradually tends to 0. The average accuracy on the three data sets was 99.17%. Its accuracy and loss function change curves are shown in Figures 14-16. It can be seen from the figure that after a certain period of training, a higher accuracy

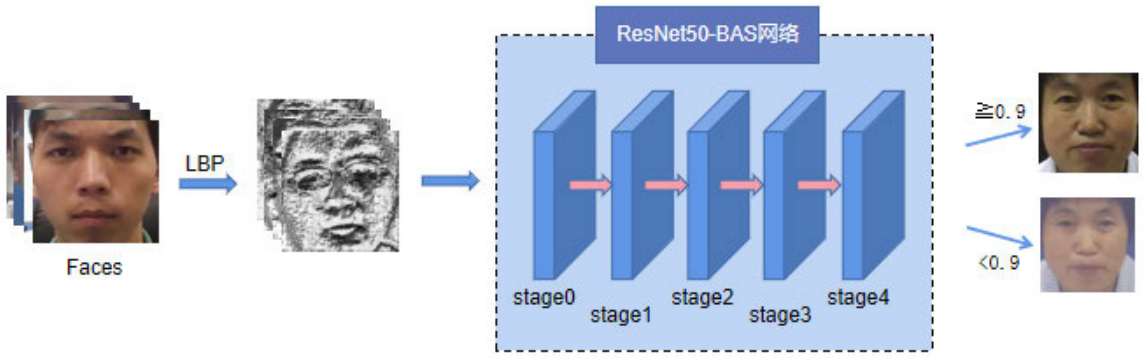


FIGURE 11. LBAS_Resnet50 model diagram.

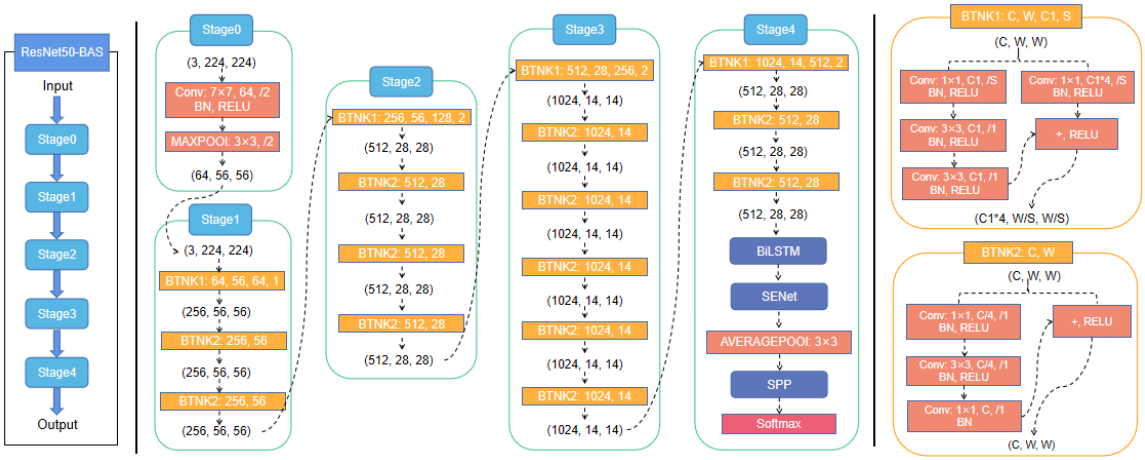


FIGURE 12. ResNet50-BAS structure.

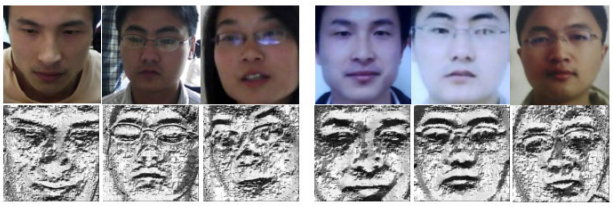


FIGURE 13. Texture map.

rate can be achieved, which gradually tends to be stable with less fluctuation.

From the above experimental results, it can be seen that the LBAS_Resnet50 model has achieved good results on NUAA, CASIA-SURF, and CASIA-FASD datasets, indicating the effectiveness and universality of the model.

B. COMPARISON WITH OTHER ALGORITHMS

The comparison between the proposed LBAS_Resnet50 model and other real-time face recognition models is shown in Table 1. It can be seen that the model proposed in this paper has a higher recognition rate than other real-time face recognition models. Because the model proposed in this paper uses the LBP algorithm, the influence of illumination is avoided to a certain extent. At the same time, BiLSTM is also added, which fully considers the timing characteristics between pictures for the recognition of real-time video. In addition, the SENet mechanism is also added to improve the extraction of important features and improve the efficiency of real-time face recognition.

TABLE 1. Comparison of different real-time face recognition models.

Model	ACC/%
Gabor_LBPH [16]	92
CNN [17]	97.78
AdaBoost_LBP [18]	92.16
FNet [19]	99.12
LBAS_Resnet50	**99.17**

The HTER comparison between the LBAS_Resnet50 model proposed in this paper and other in vivo detection models are shown in Table 2. It can be seen that the model proposed in this paper has a good effect on face anti-spoofing.

C. ABLATION EXPERIMENT

The algorithm proposed in this paper contains four ablation factors, BiLSTM, SENet, SPP, and ResNet50. In this paper, the following ablation analysis is performed to demonstrate the effectiveness of the algorithm.

a) ResNet50 is used as the base network structure.

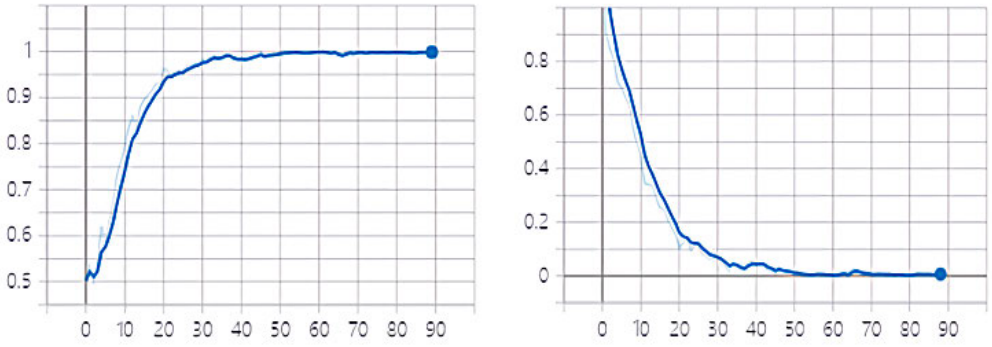


FIGURE 14. Accuracy and loss curves of NUAA dataset.

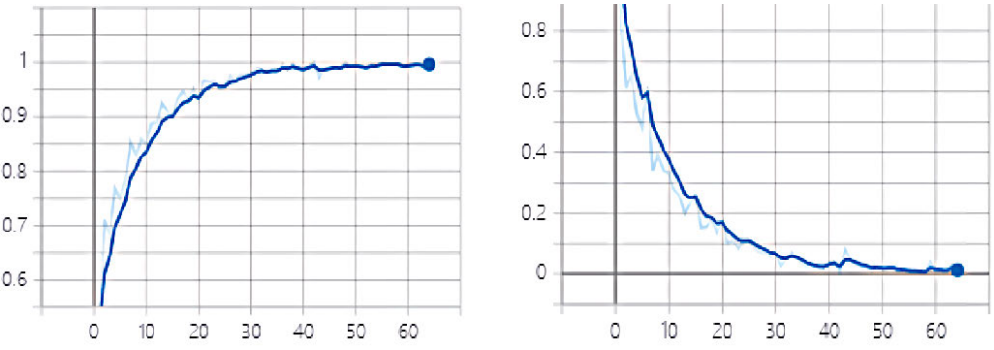


FIGURE 15. Accuracy and loss curves of CASIA-SURF dataset.

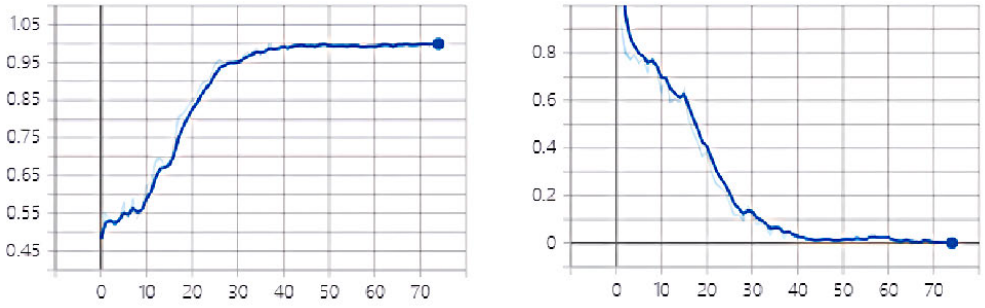


FIGURE 16. Accuracy and loss curves of CASIA-FASD dataset.

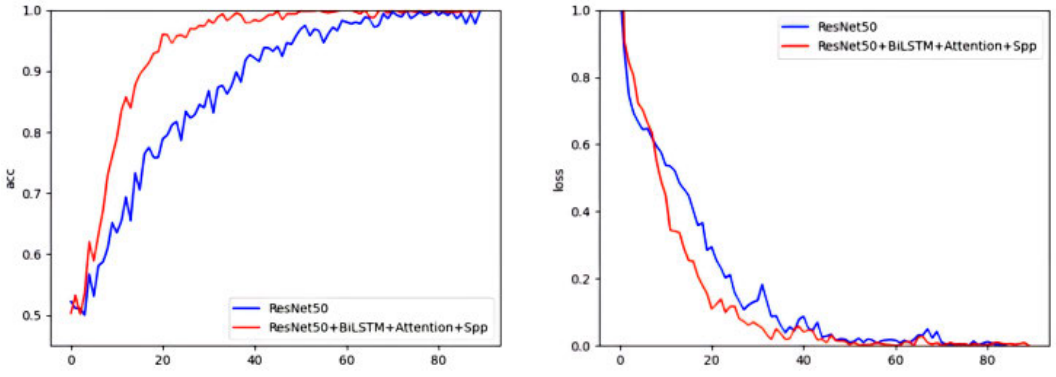


FIGURE 17. Comparison of NUAA dataset.

b) Add the BiLSTM module to the ResNet50 structure.
c) Add the SENet module to the ResNet50 structure.
d) Add the SPP module to the ResNet50 structure.
e) Add the BiLSTM, SENet, and SPP modules to the ResNet50 network, which is the final algorithm designed in this paper.

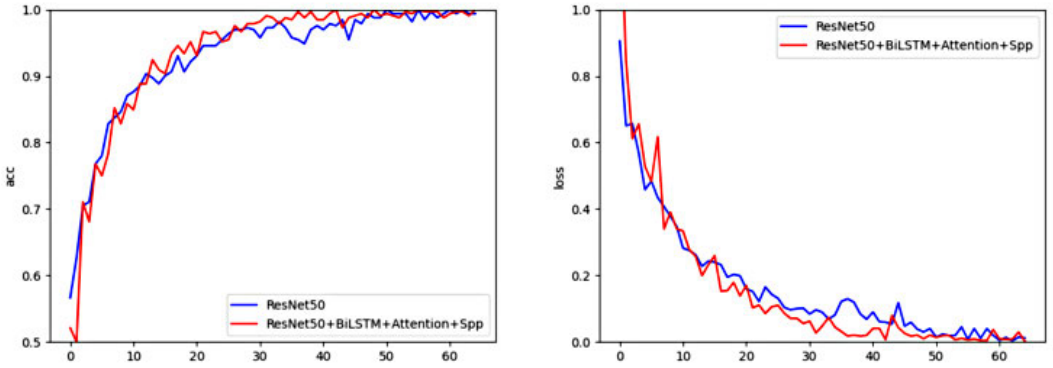


FIGURE 18. Comparison of CASIA-SURF dataset.

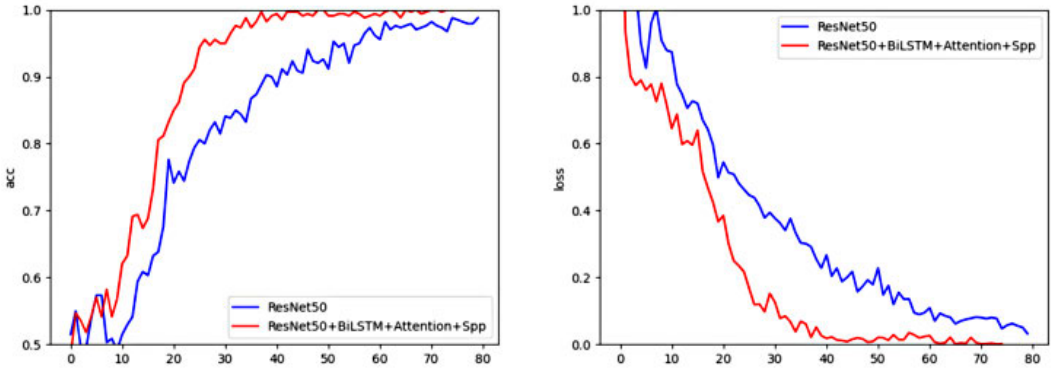


FIGURE 19. Comparison of CASIA-FASD dataset.

TABLE 2. Comparison of different in face anti-spoofing models.

Model	NUAA/%	CASIA-FASD/%	CASIA-SURF/%
HWT_GAN [20]	15.01	24.05	-
CNN_LBP-MDCT [21]	-	2.98	-
Base Net-Fusion [6]	-	2.961	-
LBP_Attention [22]	-	2.081	-
CADG [23]	8.3	16.2	-
ELM-LRF [24]	6.64	5.17	-
CNN [25]	3.24	-	-
IDA+SVM [26]	7.41	-	-
FeatherNetB [27]	-	-	2.54
ResNet18+Concat [28]	-	-	4.7
ResNet18+SE [28]	-	-	2.4
LBAS_Resnet50	**2.034**	**1.974**	**2.31**

We have validated them on the NUAA, CASIA-SURF, and CASIA-FASD datasets, respectively, and the experimental accuracies are shown in Table 3. The comparison result plots with the algorithm in this paper for item a) correspond to Figures 17-19, respectively. From the above, it is clear that the proposed algorithm in this paper has better results and stronger generalization ability.

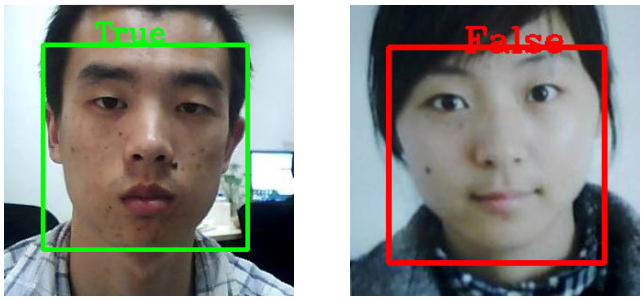
D. SINGLE-SHEET MODEL RECOGNITION RESULTS

The recognition effect of a picture randomly selected by LBAS_Resnet50 on the NUAA dataset is shown in figure 20, where (a) represents the recognition effect of a real face and (b) represents the recognition effect of a false face. LBAS_Resnet50 can achieve correct classification by randomly extracting images from the NUAA dataset.

We use OpenCV to call the camera and use dilb for face detection. Then use the model of this paper and blink detection for real-time recognition. One is a real face, and the other is a face holding a mobile phone photo. The results are shown in Figure 21, (a) is the recognition result of the real face, and (b) is the result of the fake face recognition. It can be seen

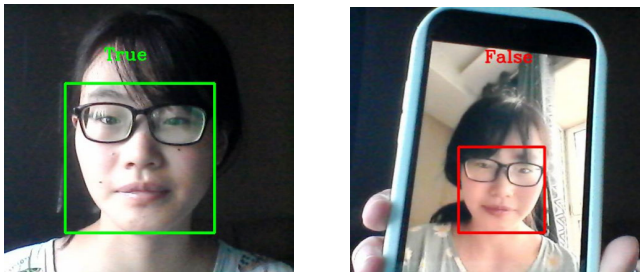
TABLE 3. Comparison of different in face anti-spoofing models.

Model	NUAA/%	CASIA-SURF/%	CASIA-FASD/%
ResNet50	99.04	98.34	97.45
ResNet50+BiLSTM	72.62	69.76	70.46
ResNet50+SENet	95.32	96.24	95.67
ResNet+SPP	98.86	96.73	97.89
LBAS_Resnet50	**99.48**	**98.65**	**99.39**



(a) real face (b) fake face

FIGURE 20. NUAA identification result.



(a) real face (b) fake face

FIGURE 21. Real-time identification results.

that the LBAS_Resnet50 model has a good effect on real-time video recognition of real faces and non-real faces.

V. CONCLUSION

This paper proposes a real-time face detection method based on blink detection called LBAS_Resnet50 to solve the problems of illumination and expression changes in the process of real-time face recognition. The model takes ResNet50 as the basic network structure and sends the texture features extracted by the LBP algorithm into the basic network to improve the tolerance to illumination in the recognition process. Then by adding BiLSTM to obtain context information, it is convenient to extract time series features, so as to improve the accuracy of real-time recognition. At the same time, the channel attention mechanism is added to extract key feature information and assign important weights, and SPP pooling is used to improve the robustness of the model. Finally, the real face is judged by eye blink detection. The experimental results indicate that the method proposed in this paper has a good effect on the accuracy of anti-spoofing real-time face recognition. Due to the different structures of paper, electronic device screens and real faces, the facial images acquired by cameras differ in brightness and illumination information. In the next research, we will consider efficiently separating brightness and reflected light features from RGB images to further improve model performance. In addition, we will consider applying sparse representation to deep learning based on face recognition.

REFERENCES

[1] X. Jie, S. Chao, and G. Chunhe, "A face recognition method combining LBP and CNN under a real-time system," *J. Heilongjiang Univ. Sci. Technol.*, vol. 28, no. 6, pp. 692–696, 2018.
[2] C. Weiyue, "Research on face recognition technology based on real-time video," *Electron. Technol. Softw. Eng.*, vol. 15, pp. 128–130, Jan. 2021.
[3] F. Guokang, L. Jun, and W. Yaoru, "Real-time face recognition on ARM platform based on deep learning," *Comput. Appl.*, vol. 39, no. 8, pp. 2217–2222, 2019.
[4] Z. Yang, X. Jun, J. Yongqiang, W. Yigang, and L. Yingchun, "Face liveness detection algorithm based on multimodal feature fusion," *Radio Eng.*, vol. 52, no. 5, pp. 738–744, 2022.
[5] L. Wei, Z. Wanling, and X. Shijun, "Face liveness detection algorithm based on LBP-MDCT and CNN," *J. Appl. Sci.*, vol. 37, no. 5, pp. 609–617, 2019.
[6] H. Xinyu, Y. Fan, Z. Pei, Z. Zhao, Z. Boli, L. Jianhua, and X. Lizhen, "Silent living detection algorithm based on multi-classification and feature fusion," *J. Zhejiang Univ.*, vol. 56, no. 2, pp. 263–270, 2022.
[7] K. Yinghui, W. Zhihan, and C. Chuan, "Real-time video face recognition based on convolutional neural network (CNN) and CUDA acceleration," *Sci. Technol. Eng.*, vol. 16, no. 35, pp. 96–100, 2016.
[8] L. Panpan, W. Wangli, and S. Zhanquan, "Multi-feature fusion face detection based on attention mechanism," *Inf. Control*, vol. 50, no. 5, pp. 631–640, 2021.
[9] C. Bojie, "A prediction method for optimizing LSTM neural network model," *J. Light Chem. Univ.*, vol. 35, no. 5, pp. 78–86, 2022.
[10] X. Jianliang, Z. Mingan, M. Jianhui, and F. Kunli, "Based on a convolutional neural networks real-time face recognition method research," *J. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 11–20, 2020.
[11] K. B. Pranav and J. Manikandan, "Design and evaluation of a real-time face recognition system using convolutional neural networks," *Proc. Comput. Sci.*, vol. 171, pp. 1651–1659, Mar. 2020.
[12] F. Deeba, H. Memon, F. Ali, A. Ahmed, and A. Ghaffar, "LBPH-based enhanced real-time face recognition," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 1–7, 2019.
[13] Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu, "Face anti-spoofing using patch and depth-based CNNs," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 319–328.
[14] A. Liu, Z. Tan, J. Wan, S. Escalera, G. Guo, and S. Z. Li, "CASIA-SURF CeFA: A benchmark for multi-modal cross-ethnicity face anti-spoofing," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 1178–1186.
[15] A. Liu, X. Li, J. Wan, Y. Liang, S. Escalera, H. J. Escalante, M. Madadi, Y. Jin, Z. Wu, X. Yu, Z. Tan, Q. Yuan, R. Yang, B. Zhou, G. Guo, and S. Z. Li, "Cross-ethnicity face anti-spoofing recognition challenge: A review," *IET Biometrics*, vol. 10, no. 1, pp. 24–43, Jan. 2021.
[16] Z. Wei, C. Gang, H. Gang, and Y. Shi, "Real-time face recognition system based on Gabor wavelet and LBPH," *Comput. Technol. Develop.*, vol. 29, no. 3, pp. 47–50, 2019.

[17] Z. Dian, W. Haitao, J. Ying, and C. Xing, "Research on real-time face recognition algorithm based on the lightweight network," *Comput. Sci. Explor.*, vol. 14, no. 2, pp. 317–324, 2020.
[18] L. Ce, L. Lan, X. Shuxing, Y. Jing, and D. Shaoyi, "A live face detection algorithm using hypercomplex wavelet generative adversarial network," *J. Xi'an Jiaotong Univ.*, vol. 55, no. 5, pp. 113–122, 2021.
[19] T. Cai, F. Chen, W. Liu, X. Xie, and Z. Liu, "Face anti-spoofing via conditional adversarial domain generalization," *J. Ambient Intell. Humanized Comput.*, pp. 7–13, May 2022, doi: 10.1007/s12652-022-03884-z.
[20] H.-H. Chang and C.-H. Yeh, "Face anti-spoofing detection based on multi-scale image quality assessment," *Image Vis. Comput.*, vol. 121, May 2022, Art. no. 104428.
[21] I. Chingovska, "On the effectiveness of local binary patterns in face anti-spoofing," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, 2012, pp. 1–7.
[22] D. Xiong and W. Hongchun, "Face liveness detection algorithm based on deep learning and feature fusion," *Comput. Appl.*, vol. 40, no. 4, pp. 1009–1015, 2020.
[23] R. Cai, "DRL-FAS: A novel framework based on deep reinforcement learning for face anti-spoofing," *IEEE Trans. Inf. Forensics Security* vol. 16, pp. 937–951, 2021.
[24] C. Xiangyun and W. Xiaopeng, "A live face detection method using extreme learning machine," *Sensors Microsyst.*, vol. 40, no. 12, pp. 122–124&132, 2021, doi: 10.13873/J.1000-9787(2021)12-0122-03.
[25] A. Alotaibi and A. Mahmood, "Deep face liveness detection based on nonlinear diffusion using convolution neural network," *Signal, Image Video Process.*, vol. 11, no. 4, pp. 713–720, May 2017.
[26] D. Wen, H. Han, and A. K. Jain, "Face spoof detection with image distortion analysis," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 746–761, Apr. 2015.
[27] Z. Yizhou and W. Hao, "Lightweight network-based and near-infrared face live detection algorithm," *Modern Comput.*, vol. 16, pp. 122–127, Jan. 2021.
[28] S. Zhang, X. Wang, A. Liu, C. Zhao, J. Wan, S. Escalera, H. Shi, Z. Wang, and S. Z. Li, "A dataset and benchmark for large-scale multi-modal face anti-spoofing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 919–928.
[29] Y. Shi, H. Qi, X. Mu, and M. Hou, "Adjustable fuzzy rough reduction: A nested strategy," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–15, Jul. 2021.
[30] Y. Shi, H. Qi, X. Qi, and X. Mu, "An efficient hyper-parameter optimization method for supervised learning," *Appl. Soft Comput.*, vol. 126, Sep. 2022, Art. no. 109266.
[31] H. Qi, Y. Shi, X. Mu, and M. Hou, "Knowledge granularity for continuous parameters," *IEEE Access*, vol. 9, pp. 89432–89438, 2021.

HUI QI received the master's degree in computer science from Shanxi University, in 2009. She is currently a Professor with Taiyuan Normal University. Her research interests include machine learning, data mining, and image processing.

CHENXU WU received the bachelor's degree in the Internet of Things from Taiyuan Normal University, in 2020, where she is currently pursuing the master's degree. Her research interests include image processing and deep learning.

YING SHI received the master's degree in computer science from Shanxi University, in 2015. She is currently a Lecturer with Taiyuan Normal University. Her research interests include machine learning and image processing.

XIAOBO QI received the Ph.D. degree in computer science from Shanxi University, in 2021. She is currently a Lecturer with Taiyuan Normal University. Her research interests include image processing, distributed computing, and data analysis.

KAIGE DUAN received the bachelor's degree in computer science and technology from Taiyuan Normal University, in 2019, where he is currently pursuing the master's degree. His research interests include machine learning and data mining.

XIAOBIN WANG received the bachelor's degree in software engineering from the Jinzhong College of Information, in 2019. He is currently pursuing the master's degree with Taiyuan Normal University. His research interests include intelligent data analysis and processing.