IoT Device Identity Authentication Method Based on rPPG and CNN Facial Recognition

Liwan Wu, Chong Yang*
Information Technology Center, Guangzhou Health Science College, Guangzhou, 510450, China
Office, Guangzhou Health Science College, Guangzhou, 510450, China

***Abstract*—This study aims to address the insufficient model recognition accuracy and limitations of authentication techniques in current IoT authentication methods. The research presents a more accurate face video image authentication technique by using a new authentication method that combines convolutional neural networks (CNN) and remote Photoplethysmography (rPPG) volumetric tracing. This method comprehensively analyzes facial video images to achieve effective authentication of user identity. The results showed that the new method had higher recognition accuracy when the light was weak. The new method performed better in ablation experiments. The error rate was 1.12% lower than the separate CNN model and 1.73% lower than the rPPG model. The half-error rate was lower than the traditional face authentication recognition model, and the method had better performance effect. Meanwhile, the images with high similarity showed better recognition stability. It can be seen that the new method is able to solve problems such as the recognition accuracy in identity authentication, but the recognition effect under extreme conditions requires further research. The research provides a new technical solution for the authentication of Internet of Things devices, which helps to improve the security and accuracy of the authentication system. By combining the CNN model and rPPG, the research not only improves the recognition accuracy in complex environments, but also enhances the system's adaptability to environmental changes. The new method provides a new solution for the advancement of Internet of Things authentication technology.**

Keywords*—Internet of Things; identity authentication; facial recognition; remote photoplethysmography; error rate

I. Introduction

With the rapid development of the Internet of Things (IoT) technology, the number of IoT devices has increased dramatically. These IoT devices have become important application devices in various fields, such as daily life, industrial production, and urban infrastructure [1]. However, with the popularity of IoT devices, the communication and data exchange between the devices also provide challenges for security and authentication [2]. Ensuring legitimate authentication of IoT devices is essential to protect data and system security [3].

Authentication, as a core issue in the field of information security, has been proposed and implemented in various approaches. Face recognition technology, as a biometric method, is potentially important in the authentication of IoT devices [4]. This technique not only provides highly accurate authentication, but also reduces the reliance on traditional passwords and keys, thus improving system security.

However, traditional face recognition methods usually require specialized hardware devices, which are extremely sensitive to lighting conditions and environmental factors. This limits their application scope in IoT devices. In addition, some traditional face recognition methods perform poorly in terms of recognition accuracy and error rate, which poses difficulties for the application of face recognition technology.

Therefore, this research investigates the face recognition authentication method using remote Photoplethysmography (rPPG) and Convolutional Neural Network (CNN). Firstly, utilizing the powerful fitting ability of neural networks, a facial pose recognition method using neural networks is designed to address the facial occlusion and recognition in facial recognition. Secondly, rPPG technology is used to solve the insufficient recognition accuracy and poor facial information in the authentication process of IoT devices. This research is divided into six sections. Section II reviews the domestic and international research. Section III constructs the research method. Results, discussion and conclusion is given in Section IV, IV and VI respectively.

II. Literature Review

Identity authentication is a means of authenticating users through facial recognition and fingerprint authentication. It is widely used in some IoT devices. Therefore, many experts and scholars have conducted extensive research on identity authentication of facial recognition devices. Mengjuan Zhai et al. developed a new scheme based on chameleon hash value and self-updating secret sharing to address the user privacy protection. The new solution was characterized by editable blockchain, providing users with fine-grained and fair editing functions. It could be applied with only a small additional cost. Compared with traditional centralized authentication schemes, the new scheme could better protect user privacy while providing more refined and fair services. However, there was still relatively little research on user physical characteristics in this study. Therefore, this study seeks new identity authentication methods for further research [5]. Yu Pingping et al. proposed a novel gesture recognition and identity verification algorithm based on continuous hidden Markov models and optical flow methods to address information security issues in the power IoT. The optical flow method was applied to extract features from preprocessed dynamic gesture information. A user dynamic gesture model using CHMM was established, which could improve the dynamic gesture recognition accuracy. The research results indicated that the new method had advantages in the identity verification accuracy, with higher recognition accuracy compared with traditional methods. However, the study only achieved this by recognizing user gestures, which was unable to achieve faster and more accurate authentication recognition through facial recognition [6]. Xin Xu et al. proposed a novel biometric identity verification strategy based on music induced autobiographical memory electroencephalogram to address the identity verification. The research results indicated that it had high uniqueness, which was suitable for identity verification applications. However, the method used in the study was only address the biometric authentication, which did not fully address the entire process from identification to authentication [7]. Zhiguo Qu et al. proposed a novel quantum identity authentication protocol ground on three photon error correction codes to address the anti-interference problem of quantum identity authentication under quantum channel noise. The research results indicated that the protocol could effectively resist the interference of noise on information transmission in quantum channels, which had good anti-interference performance. Meanwhile, the new protocol maintained better security against various eavesdropping attacks. However, the study was not effective in improving the accuracy of authentication [8].

Jaiswal, Kokila Bharti et al. proposed a fusion-based new method to address the impact of non-uniform lighting on rPPG measurement results. The new method combined RGB and multi-scale Retinex color spaces to generate prominent spatiotemporal maps. The experimental results showed that the proposed method achieved excellent results in both inter database and internal database testing in public databases. This method could improve the data analysis of rPPG, but the method used in the study had insufficient security [9]. Tomasz Szabala et al. developed a new method to obtain remote optical heart movement data from a standard camera on a personal computer. The research results indicated that the image intensity changes generated by tracking blood volume changes in microvascular tissues using visible light cameras could effectively estimate the heart pulse. The new method was effective in detecting human pulse changes, but there were still shortcomings in the research of face information data [10]. Feng Qi et al. proposed a distributed and efficient key distribution protocol that did not require secure channel assumptions to address the inherent issues of identity cryptography in the IoT and ad-hoc networks. The research results indicated that the new protocol was maliciously secure under weaker assumptions. The new method could effectively solve the IoT data authentication security [11]. Gao, Zhigang et al. proposed a user authentication method based on button time interval groups to address the high cost, and low accuracy in mobile device user authentication. The research results indicated that the new method had high accuracy, low cost, and sustainable authentication. It could effectively solve the shortcomings of existing identity verification methods based on button dynamics. The research could effectively improve the low recognition accuracy of user authentication, but there was still a lack of security [12].

In summary, there are still many issues with current identity authentication methods for devices, such as security, recognition accuracy, etc. Therefore, to build a relatively complete facial recognition and identity authentication system, CNN and rPPG models are used to design the facial recognition and identity authentication method.

III. IoT Device Authentication Model Building

This chapter mainly analyzes the application of CNN in facial recognition. At the same time, a facial recognition identity authentication system integrating CNN and rPPG methods is built on the basis of the rPPG method. Through systematic analysis, this research can improve the facial recognition identity authentication system.

A. Facial Recognition Analysis Based on CNN

Facial recognition is a detection and analysis technique for recognizing and authenticating facial features of individuals. With the rapid development of the IoT, it has become the main means of identity authentication. As a feed-forward neural network, CNN is mainly used in image recognition analysis due to its ability to recognize image features during training. The CNN structure includes input, convolution, pooling, fully connected, and output layers. The input layer mainly processes the input data to ensure that the current data can be analyzed and processed by the neural network structure. The pooling layer is mainly used to reduce over-fitting by reducing the data dimensionality. The fully connected layer is mainly applied to connect and analyze the data of the upper and lower layers, facilitating the training of subsequent classifiers. This is also a processing layer for improving the ability of the entire model. The output layer mainly outputs the data processing results of the current network model [13]. The CNN structure diagram is shown in Fig. 1.

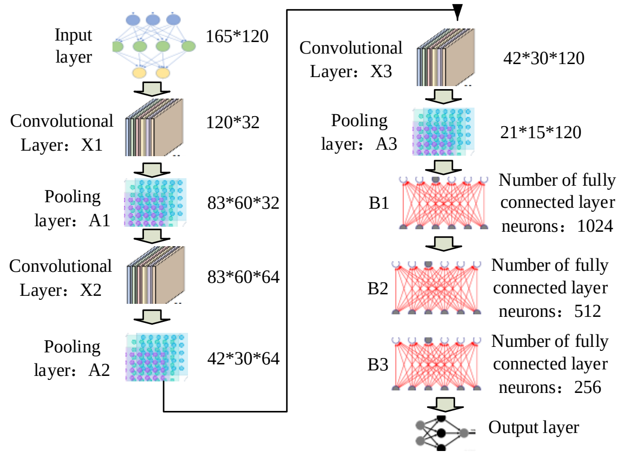


Fig. 1. CNN structure diagram.

In Fig. 1, X1, X2, and X3 are convolutional layers of the CNN, mainly used for extracting and analyzing different image features. A1, A2, and A3 are pooling layers of the current results, mainly used to reduce the dimensionality of the feature network, lower the computational complexity, and overcome over-fitting. B1, B2, and B3 are fully connected layers of the network model, which mainly extract features from the model results to accelerate the classification effect. Therefore, the data output structure of CNN is shown in Eq. (1) [14].

$$y_k = f\left(\sum_{j=1}^{m} w_{kj} x_j + b_k\right) \tag{1}$$

In Eq. (1), y_k represents the output data. $x_1, x_2, \cdots, x_m$ represent the input data. b_i represents the size of the bias. $w_{k1}, w_{k2}, \ldots w_j$ represent the activation function. CNN enhances image recognition capability through convolution operations. Therefore, the one-dimensional convolution of CNN is shown in Eq. (2).

$$c_{cn} = f\left(x * w_{cn} + b_{cn}\right) \tag{2}$$

In Eq. (2), f represents the activation function of the convolution. x refers to the input data size. w_{cn} refers to the convolution value. b_{cn} represents the bias size of the convolution kernel. c_{cn} represents the output value of the convolutional layer in one-dimensional space. The normalized probability distribution of CNN is displayed in Eq. (3).

$$p\left(x\right)_i = \frac{e^{z_i}}{\sum_k e^{z_i}}, i = 1, 2, ..., k \tag{3}$$

In Eq. (3), k refers to the number of classified data. z_i refers to the number of neurons in the output layer that have not been activated. $p(x)_i$ represents the normalized probability of the model. At this point, the cross entropy loss function of the CNN is shown in Eq. (4) [15].

$$H\left(p, q\right) = -\sum_x p(x) \log q(x) \tag{4}$$

In Eq. (4), $p(x)$ represents the distribution definition. $q(x)$ represents the distribution definition that has not been predicted. $H(p, q)$ represents the loss value of uncrossed entropy. In model analysis, the collected image data is subjected to feature processing. Eq. (5) shows the CNN fusion method for image data feature processing.

$$F_k = \sum_{i=1}^{n} f_i \tag{5}$$

In Eq. (5), F_k represents the image data fused with feature data using a separate convolutional layer. k represents the number of pooling layers A2 and A3. f_i represents the feature image data on this channel. n represents the number of channels. When k is A2, the number of channels is 64. When k is A3, the channels are 120. When the number of feature fusion layers increases, the coordinate transformation is shown in Eq. (6).

$$V(x_n, y_n) = V_{(x?\frac{F_w}{T_w}, y?\frac{F_h}{T_h})} \tag{6}$$

In Eq. (6), $V(x, y)$ represents the size of pixels when the image coordinate is (x, y). F_w, F_h refer to the width and height of the feature image. T_w, T_h refer to the width and height of the target image. The feature fusion obtained by adjusting the number of layers is shown in Eq. (7) [16].

$$F_N = \alpha F_{P2} + (1 - \alpha) F_{P3} \tag{7}$$

In Eq. (7), F_N represents the classification feature data after multi-layer fusion. F_{P2}, F_{P3} represent the feature set of classification data after increasing the number of layers. α represents the weight coefficient. The weight values of the algorithm are mainly used for matrix analysis of facial features and other data from different facial images. Eq. (8) represents the weight vector matrix.

$$f(x) = x \square y^{''} \tag{8}$$

In Eq. (8), $f(x)$ is the weight vector matrix. x represents the matrix definition of the sample. $y^{''}$ represents the defined vector size. Analyzing the matrix vector representation of two images can achieve weight size analysis, as presented in Eq. (9) [17].

$$f^{'}(x) = x^{'} \square y^{''}, f^{'}(x) = f(x) \tag{9}$$

In Eq. (9), $f^{'}(x) = x^{'} \square y^{''}$ represents the weight vector matrix of another image. When the weights of two facial images are equal, the algorithm can learn the true weight size. To improve the feature vector extraction ability of the algorithm for facial data, the compensation vector and weight vector are multiplied to obtain the final vector extraction result, as displayed in Eq. (10).

$$H(x) = (x_i + a_i) y_i^{''} \tag{10}$$

In Eq. (10), x_i represents the size of the original vector. a_i represents the compensation vector. $y_i^{''}$ represents the vector definition of weights. $H(x)$ represents the final feature vector extraction.

B. *Analysis of Device Identity Authentication System Based on rPPG and CNN*

In device identity authentication, there are some background shadows and unevenness in the facial area of the image during the recognition process, which leads to recognition [20] errors in facial information data. There will also be authentication results that are not real people. Therefore, using only CNN models for identity authentication can lead to recognition errors and insufficient clarity of the entire system. Therefore, to improve the detection ability of live facial data, the rPPG feature analysis algorithm is added to the research. This method can use image background information to enhance the color and color difference information data in facial image feature extraction, thus converting colors and other details in facial images and improving the performance of identity recognition. The current algorithm framework is shown in Fig. 2.

In Fig. 2, the structure of the algorithm mainly includes a neural network module and an rPPG module. In the neural network module of the algorithm framework, a detection model is first used to detect video images and other face data. The regional image of the face is analyzed through key positions and localization. Afterwards, the analyzed image data is transmitted to the color feature extraction area and appearance extraction area. The data is trained and analyzed through a model classifier. In the rPPG module, the matrix data is mainly fused by using remote optical volume description technology to extract power features of facial region signals and analyze spectral features. By training the classifier model, the probability weight size is calculated, which is the best weight value for the facial image. Finally, image recognition authentication is completed by weighting the two classifiers [18].

rPGG is a technology that can measure human blood heart rate and other factors. When light shines on the human body, some capillaries and hemoglobin can absorb some of the light. Cardiac fluctuations can alter the hemoglobin levels in different regions of the human body. This technology can capture this change and feedback it into the model system. The working principle of rPPG is shown in Fig. 3.

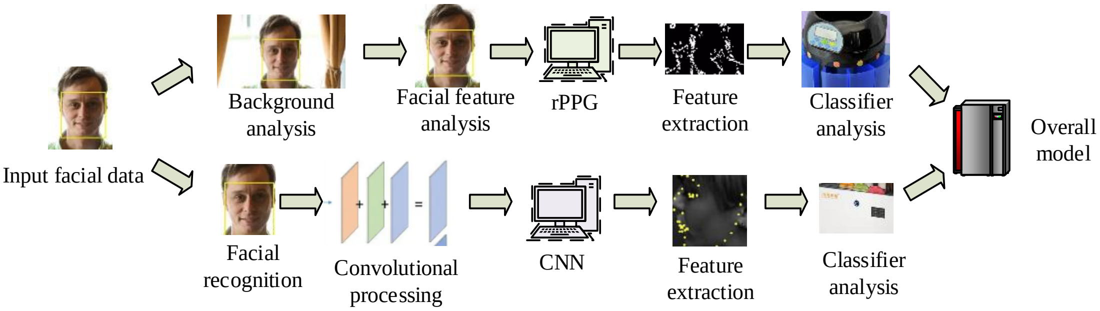


Fig. 2. Schematic diagram of model framework.

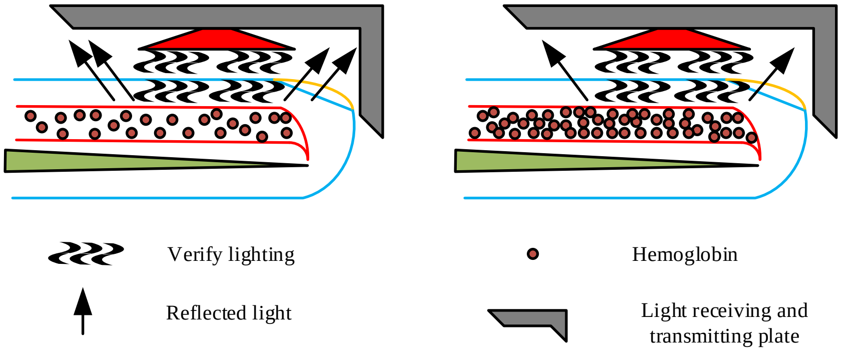


Fig. 3. Schematic diagram of rPPG working principle.

In Fig. 3, when light is emitted from the instrument, it is absorbed by hemoglobin in the human body through the skin. A portion of the light that is not absorbed is directly fed back to the emitting surface. In Fig. 3, the amount of unabsorbed light decreases as the amount of hemoglobin in the skin area increases. Therefore, the feedback light obtained is reduced. Because the number of proteins in different positions of the face varies, this method can describe the data of from different positions of the face. At the same time, background information and lighting information can affect facial signal recognition during rPPG face authentication. Therefore, during facial authentication, the facial background signal of the face is analyzed and recognized to improve the recognition performance of the model.

When extracting features from facial data, the image data information is preprocessed firstly. The processing method mainly involves eliminating the influence of ambient light on signal changes, that is, removing some redundant data signal information. The other is to eliminate random signal noise on adjacent images. This random noise can cause inaccurate model recognition. Finally, the identified heart rate standard sometimes exceeds the normal heart rate range of the human body. Therefore, it is necessary to remove heart rate signals that exceed the normal heart rate range during processing. After completing data extraction, the algorithm needs to use Fourier transform to convert the signal into frequency-band and time-domain. Some real facial image data can be recognized through spectral feature transformation, thereby improving the recognition performance between real and virtual faces.

In simulated face authentication, there are similarities in the faces, which are caused by subtle differences in the faces of different people. Therefore, to meet the needs of most facial recognition, it is necessary to improve feature recognition capabilities and the stability and consistency of image data authentication. Thus, similarity analysis is added to the model, as shown in Eq. (11).

$$x = \bigcup_{\substack{i,j=1,\cdots,N \\ i\le j}} \rho(S_i, S_j) \tag{11}$$

In Eq. (11), $\rho(S_i, S_j)$ represents the similarity of the measured signal. The signal information is represented by S_i, S_j. $\cup$ represents the continuous calculation. To improve the similarity of the entire signal, the correlation spectrum of similarity is taken to the maximum value, as shown in Eq. (12) [19].

$$\rho(S_i, S_j) = \max | f\{s_i \bullet s_j\} | \tag{12}$$

In Eq. (12), f represents the Fourier transform. $\bullet$ represents the correlation operator. The remaining parameters are the same as above. Finally, the regional signal of the face can be obtained through correlation calculation, while reducing the influence of random noise. The data processing and classification results of the entire rPPG module are shown in Fig. 4.

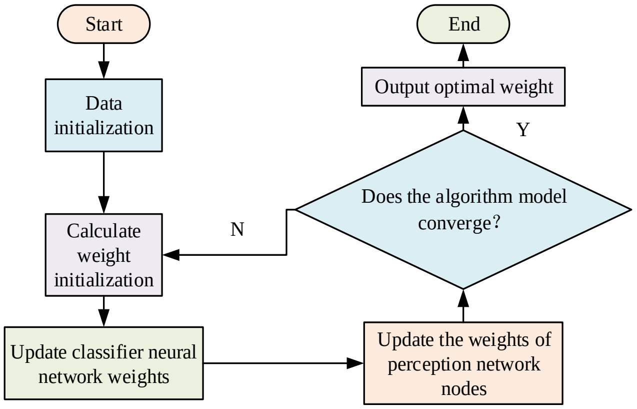

Start
Data initialization
Calculate weight initialization
Update classifier neural network weights
Update the weights of perception network nodes
Does the algorithm model converge?
N
Y
Output optimal weight
End

Fig. 4. Data processing flow of rPPG module.

In Fig. 4, the rPPG module first initializes the data information randomly. The weights of the classifier and the perception network are initialized. Afterwards, the classifier and neural network weights are updated through the model. The weights of the perception network nodes are fixed to determine whether the current model is converging. If it converges, the process ends. If it does not, the weights are calculated and the optimal weight size is output, thereby obtaining the face authentication process of the rPPG module. In the data collection and analysis of the entire system, two modules use different classifiers to collect image data. Therefore, when collecting and analyzing data, different classifiers are used to analyze the data information. The process is displayed in Fig. 5.

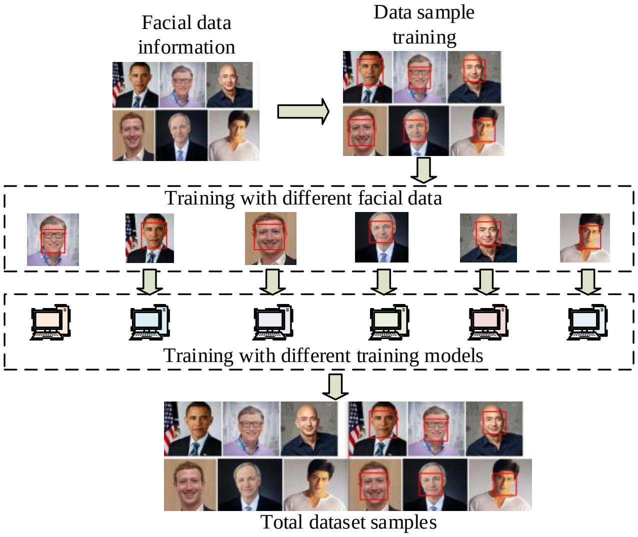

Facial data information
Data sample training
Training with different facial data
Training with different training models
Total dataset samples

Fig. 5. Schematic diagram of data collection process.

In Fig. 5, when face authentication involves multiple data face samples, the model first trains the sample data. Afterwards, the data is randomly sampled and distributed into 1 to n sampling datasets. The datasets are then trained using relatively weaker classifiers. The trained datasets are combined before being trained on strong datasets. The datasets trained in this way can achieve relatively good data collection and classification. The complete authentication system process is shown in Fig. 6.

In Fig. 6, the system module consists of three main parts: network training module, data acquisition module, and image processing and analysis module. The network training module mainly processes and analyzes facial videos and image data that require authentication. CNN and rPPG are used to process and analyze image video data. The data acquisition module mainly analyzes and processes the facial video data that needs to be collected to ensure that it can be processed by the model currently. The final image processing module is to detect, recognize, and authenticate the current image data, then process and analyze it. The processed data is fed back into the system to complete the facial recognition and authentication process.

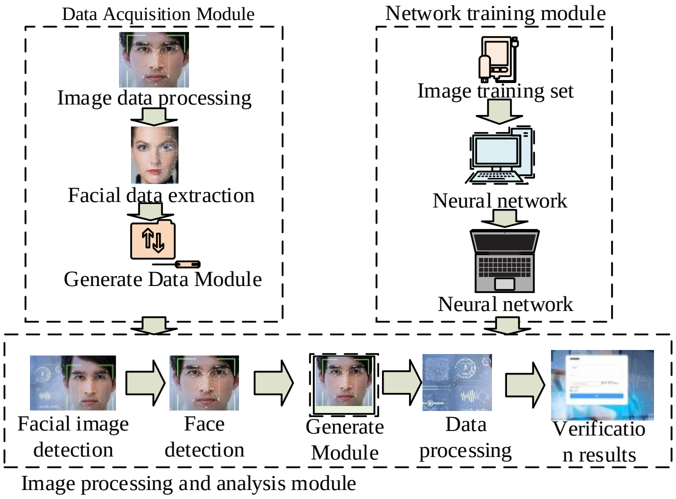


Fig. 6. Diagram of authentication system process module.

IV. Results

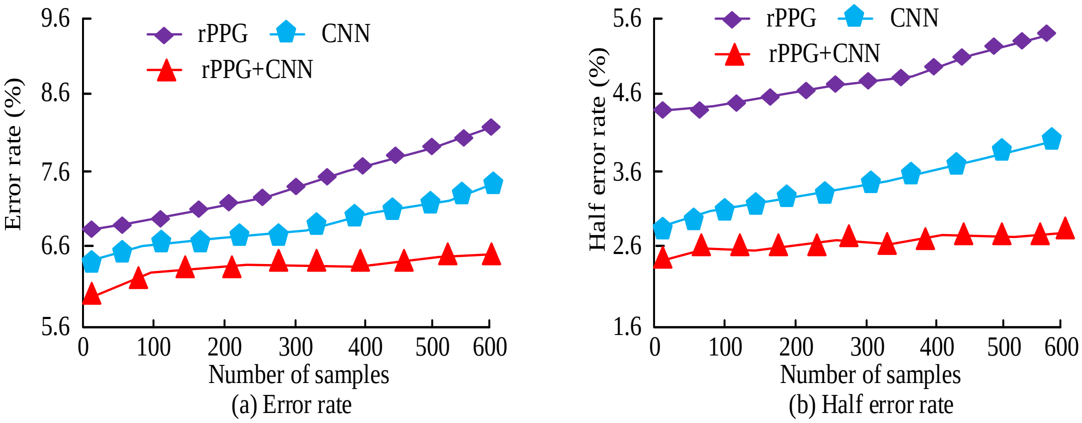
To test the authentication performance and the algorithm performance, the publicly available facial video dataset is selected. The same pixel size is 360*240, and the number of faces selected is 50, totaling 1200 video images. A total of 1200 video images are selected, including 50 user image videos that require authentication. The dataset is divided into two parts, with 600 video image data for training and testing. Each trained neural network has the same parameters. To test the recognition accuracy of the current research method on facial video images in different backgrounds, three models with different iteration times are selected for comparison. Table I displays the results.

In Table I, when the iteration was the same, the recognition accuracy of weak light and warm color backgrounds was relatively higher. When the iterations were 10, the recognition accuracy of strong light backgrounds was 0.79% lower than that of the highest warm light backgrounds. The decrease in accuracy was relatively small. This may be due to the influence of lighting on the image data. The recognition accuracy of the model performed better at different iterations, with the highest recognition accuracy of 92.31% at 15 iterations for warm light backgrounds. This may be due to the better training effect of the model with higher iterations. To test the ablation performance of the current usage method, the error rate analysis is performed on the rPPG model and CNN model used separately, as shown in Fig. 7. The half error rate represents the acceptable probability of an error and the average value of the error probability. The small value indicates that the model is better.

Fig. 7 (a) displays the error rates of different models. As the validation recognition samples increased, the sample error rate has increased. Among the three models, the proposed method had the lowest error rate. The average overall error rate was around 5.72%, while the average error rates of the other two models were 7.45% and 6.84%. The proposed method was 1.12% lower than the CNN model and 1.73% lower than the rPPG model. In the comparison of the half error rates in Fig. 7 (b), the half error rate of the proposed method was lower. The change was also the same as the error rate. The overall performance of the research method was improved after incorporating some advanced models, which also indicated that the two models optimized each other. To compare the recognition performance of different methods, Local Binary Pattern - Three Orthogonal Planes (LBP-TOP), Long Short Term Memory-CNN (LSTM-CNN), and Visual Geometry Group (VGG) are compared. Fig. 8 displays the results.

TABLE I. RECOGNITION ACCURACY OF DIFFERENT SCENES UNDER THE SAME STEP SIZE AND DIFFERENT ITERATION TIMES

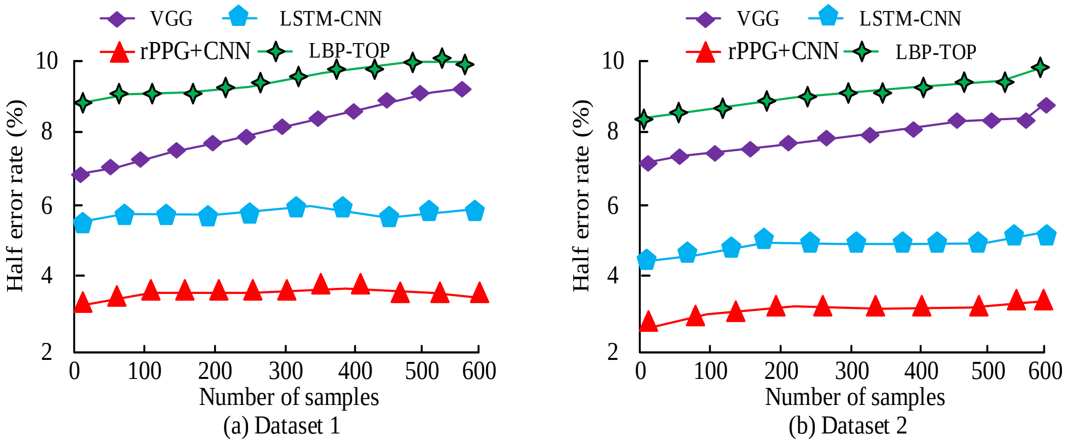
Scene	Strong light			Weak light			Warm light		
Iterations	5	10	15	5	10	15	5	10	15
Model step size	10	10	10	10	10	10	10	10	10
Recognition accuracy (%)	80.25	84.51	90.23	81.24	84.66	92.03	81.25	85.3	92.31



(a) Error rate

(b) Half error rate

Fig. 7. Ablation experiments for error rate and half error rate.



(a) Dataset 1

(b) Dataset 2

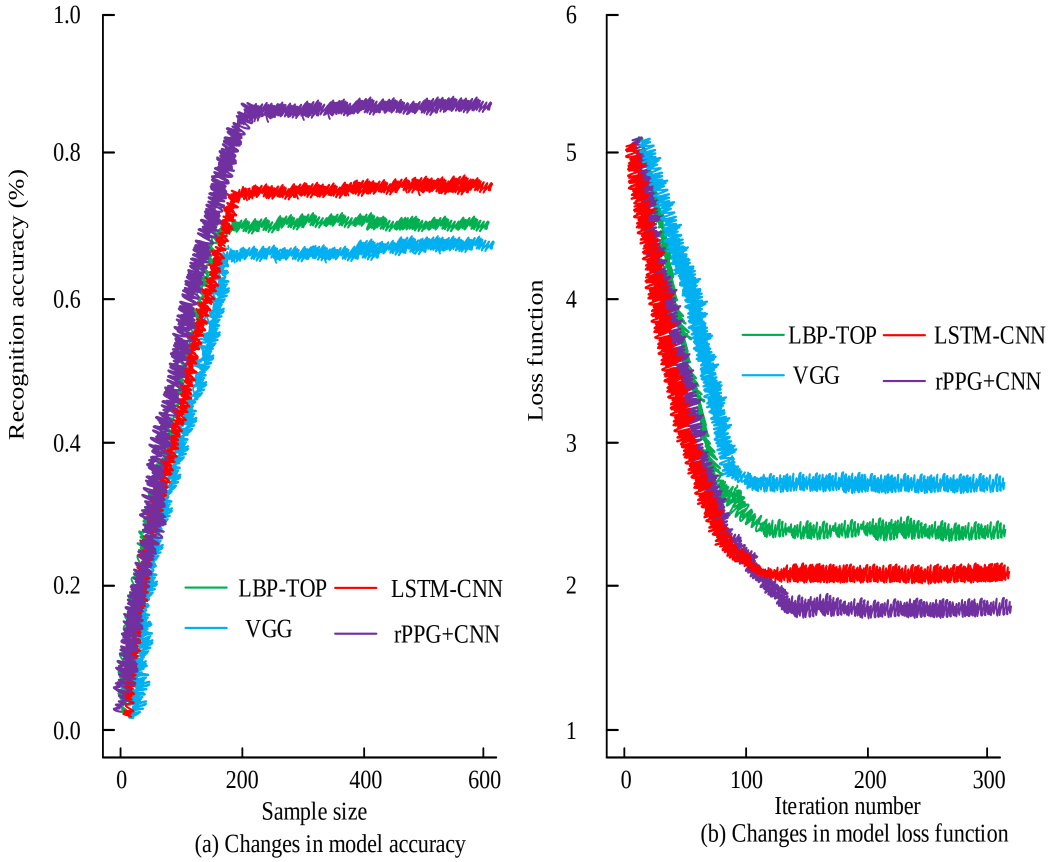
Fig. 8. Comparison of half error rates in different model recognition.

In Fig. 8(a), in dataset 1, the half error rate value of the three models increased with the increase of dataset size. However, the increase of CNN was relatively small. The average half error rates of LBP-TOP, LSTM-CNN, VGG, and rPPG+CNN models were 9.24%, 8.15%, 5.84%, and 3.21%, respectively. The half error rate of the proposed method was lower, with LBP-TOP, LSTM-CNN, and VGG models being 6.03%, 4.94%, and 2.27% lower, respectively. In Fig. 8(b), the variation trend of several models in dataset 2 was the same as that in Fig. 8(a). The average half error rate was basically the same. This indicates that the half error rate of these models does not change much in different datasets, which may be due to the relatively stable models. To verify the generalization ability of the current research method, the data performance of different models is analyzed. Table II displays the results.

In Table II, the half error rate obtained from different datasets for the testing and training sets of different models was not the same. When dataset 2 was used as the testing set, the model had a lower half error rate. This may be due to differences in algorithm stability during training. Among the four models, the rPPG+CNN had the lowest half error rate and better performance. To test the recognition accuracy and model loss function changes of different models, the obtained results are shown in Fig. 9.

TABLE II. COMPARISON RESULTS BETWEEN DIFFERENT METHOD DATASETS AND TEST SETS

Model	Testing set	Training set	Half error rate (%)
LBP-TOP	Dataset 1	Dataset 2	50.1
	Dataset 2	Dataset 1	49.3
LSTM-CNN	Dataset 1	Dataset 2	45.6
	Dataset 2	Dataset 1	46.3
VGG	Dataset 1	Dataset 2	61.5
	Dataset 2	Dataset 1	49.8
rPPG+CNN	Dataset 1	Dataset 2	42.5
	Dataset 2	Dataset 1	37.1



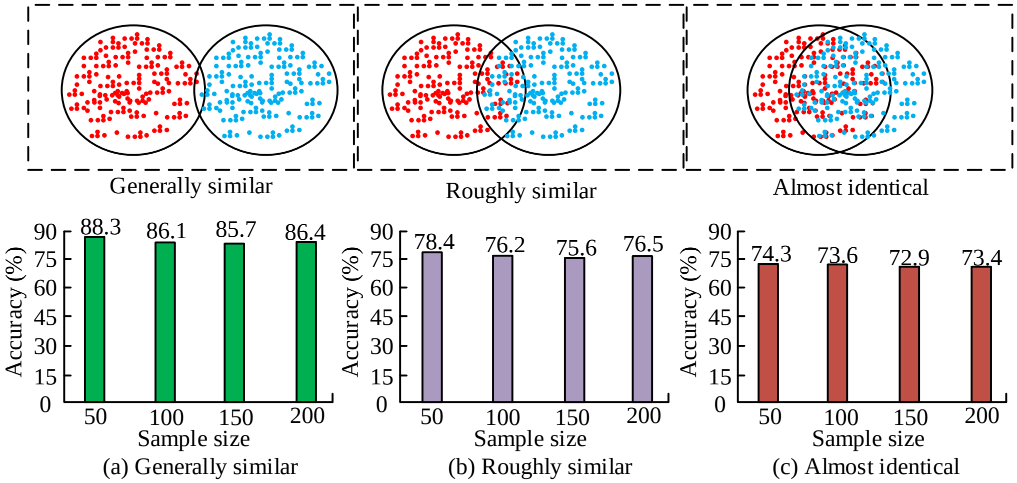
(a) Changes in model accuracy

(b) Changes in model loss function

Fig. 9. Accuracy and loss function changes of four models.

In Fig. 9(a), among the accuracy trends of the four models, the accuracy increased with the increase of sample size and then tended to a stable state. At this time, the accuracy of the LBP-TOP was 67.2%, the LSTM-CNN was 74.6%, the VGG was 63.4%, and the rPPG+CNN was 89.5%. The rPPG+CNN had the highest accuracy among the four models. In Fig. 9(b), the loss function decreased with increasing iterations and then tended to stabilize. The minimum loss function value of the rPPG+CNN was only 1.8, indicating that its model was more stable. To verify the authentication performance of the proposed method, similar facial video images in the dataset are used as the validation dataset to analyze the facial image authentication, as shown in Fig. 10.

In Fig. 10, when the facial similarity was low, the recognition accuracy was higher, with the highest value being 88.3%. After increasing the similarity, the recognition accuracy slightly decreased. When the similarity was almost identical, the recognition accuracy decreased significantly. However, from the analysis in Fig. 10, the recognition accuracy was still at a high level after increasing similarity, indicating that the overall authentication performance of the model was good.



(a) Generally similar (b) Roughly similar (c) Almost identical

Fig. 10. Changes in similarity accuracy of samples.

V. Discussion

Face recognition technology has been widely used in the field of biometrics. Traditional authentication methods, such as passwords and tokens, although commonly used, carry the risk of being stolen or forgotten. Their high operational complexity makes them unsuitable for large-scale deployment in IoT environments. In recent years, biometrics have been recognized as a powerful tool for solving the authentication problem of IoT devices due to its convenience and security. For this purpose, the study uses CNN and rPPG techniques and applies them to face identification on IoT devices.

In comparison of different background colors, face recognition accuracy is higher in warm backgrounds, which may be due to inaccurate testing of face data caused by lighting effects. Secondly, in 15 iterations, the recognition accuracy of warm colored backgrounds is relatively high, reaching 92.31%, which may be due to the increase in the number of iterations resulting in more accurate face data. In the comparison of the error rate for different models, the change in the error rate of the model used in the study increases with the increase of sample size, which may be due to the increase in the number of samples resulting in a decrease in the overall recognition effect. The error rate of the research model is lower. This may be due to the added rPG technology [21] improving model performance. In the comparison of half error rate of the three models, the half error rate of the research method is lower, which may be due to the high accuracy of facial authentication recognition in the research model. In the comparison of error rate values of different models, the model used in the study has a lower half error rate, which may be due to its ability to better handle data. In the comparison of half-error rate of different models, the performance effect of the research model is better than the individual model, which may be due to different technologies improving the model performance. In the comparison of the accuracy rate change, the used model has the highest accuracy rate, which may be because the model used can better process facial data. In the variation of loss function values in several models, the designed model has lower loss function value, which may be due to the more stable performance of the research model. In the similarity comparison of different face data, the recognition performance of the authentication similarity of the research model is better, as the algorithm currently used in research can recognize blood vessels in the face.

In summary, the model used in the current study has better performance and recognition effect in the face recognition authentication of IoT devices. The model has better face recognition authentication effect and recognition accuracy, which has a better guiding role for face recognition authentication afterward.

VI. Conclusion

This study mainly focused on the facial identity authentication of IoT devices. The CNN and rPPG face detection technology were used to build a new device facial identity authentication system. Firstly, a facial recognition and identity authentication system based on CNN and rPPG was analyzed and constructed. Then, the performance and feasibility of the current system were verified through comparative analysis among different models. The research results indicated that the recognition accuracy of the proposed model varied under different color backgrounds. The algorithm had higher recognition accuracy under weaker lighting conditions. In the comparison of error rates, the rPPG+CNN model had the lowest error rate, which was 1.12% lower than the CNN and 1.73% lower than the rPPG. The half error rate of rPPG+CNN in different comparison methods was 6.03%, 4.94%, and 2.27% lower than those of LBP-TOP, LSTM-CNN, and VGG, respectively. When testing and training on different datasets, the model performed better when dataset 2 was the testing set. Among the four different comparison methods, the rPPG+CNN had the best accuracy and overall performance. At the same time, when comparing similar faces, the method used in the study had relatively stable recognition accuracy when the facial similarity was high. The accuracy was at a relatively high level. Although this study has achieved many results in facial recognition identity authentication, further improvement should be improved. For example, the background and datasets used in the experiment are relatively small. More and larger data will be analyzed in the future. At the same time, future research will also analyze data from different devices. In addition, the study is less analyzed for different scenarios of real faces, so different face authentication scenarios will be analyzed and detected in the subsequent study. Finally, in the study only focuses on IoT devices. Therefore, different devices will be analyzed for face authentication in the subsequent study.

FUNDING

The research is supported by: Guangzhou Education Bureau 2022 Guangzhou Higher Education Teaching Quality and Teaching Reform Project Famous Teacher Studio Project，Guangzhou Health Vocational Education Research Famous Teacher Studio, No.: 2022MSGZS020；Guangzhou Teaching Achievement Cultivation Project，Exploration and Practice of the "Lesson Plan Post" Teaching Model Based on Modular Case Teaching of Work Process Professional Courses，No.:2020123311.

REFERENCES

[1] Gupta D S, Islam S H, Obaidat M S, Hsiao Kuei-Fang. A Novel Identity-based Deniable Authentication Protocol Using Bilinear Pairings for Mobile Ad Hoc Networks. Ad Hoc & Sensor Wireless Networks, 2020, 47(1-4):227-247.

[2] Zheng L, Song C, Zhang R, Baoqing Lv, Yujin Liu, Meng Cui. Design and analysis of telemedicine authentication protocol. International Journal of Sensor Networks, 2021, 37(3):198-208.

[3] Chen Y, Chang T, Liu W. Improved SRP algorithm and bidirectional heterogeneous LTE-R authentication key. IET Communications, 2023, 17(11):1300-1309.

[4] Ante L, Fischer C, Strehle E. A bibliometric review of research on digital identity: Research streams, influential works and future research paths. Journal of Manufacturing Systems, 2022,62(6):523-538.

[5] Zhai M, Ren Y, Feng G, Xinpeng Zhang.Fine-Grained and Fair Identity Authentication Scheme for Mobile Networks Based on Blockchain. China Communications, 2022, 19(6):35-49.

[6] Sun Y, Du Z, Cao N, Du Zheng. An identity authentication method for ubiquitous electric power internet of things based on dynamic gesture recognition. International Journal of Sensor Networks, 2021, 35(1):57-67.

[7] Xu X, Jiang L, Xu T. Identity Authentication Based on Music-Induced Autobiographical Memory EEG. Journal of circuits, systems and computers, 2022, 31(11):1-16.

[8] Qu Z, Liu X, Wu S. Quantum identity authentication protocol based on three-photon quantum error avoidance code in edge computing. Transactions on Emerging Telecommunications Technologies, 2020, 33(6):3945-3964.

[9] Jaiswal K B, Meenpal T. rPPG-FuseNet: Non-contact heart rate estimation from facial video via RGB/MSR signal fusion. Biomedical signal processing and control, 2022, 78(Sep.):1-9.

[10] Szabaa T. Exploratory Study on Remote Photoplethysmography using Visible Light Cameras. PRZEGLĄ D ELEKTROTECHNICZNY, 2023, 99(1):282-285.

[11] Feng Q, He D, Wang H, Wang Ding. Multi-party key generation protocol for the identity-based signature scheme in the IEEE P1363 standard for public key cryptography. IET Information Security, 2020, 14(4):724-732.

[12] Gao Z, Diao W, Huang Y, Xu Ruichao, Lu Huijuan, Zhang Jianhui. Identity authentication based on keystroke dynamics for mobile device users. Pattern Recognition Letters, 2021, 148(7):61-67.

[13] Wei Z, Liu F, Masouros C, H. Vincent Poor. Fundamentals of Physical Layer Anonymous Communications: Sender Detection and Anonymous Precoding. IEEE Transactions on Wireless Communications, 2021, 21(1):64-79.

[14] Madarkar J, Sharma P, Singh R P. Sparse representation for face recognition: A review paper. IET Image Processing, 2021, 15(2):1825-1844.

[15] Ergin S, Isik S, Gulmezoglu M B. Face Recognition by Using 2D Orthogonal Subspace Projections. Traitement du Signal, 2021, 38(1):51-60.

[16] Mohanty V, Thames D, Mehta S. Photo Sleuth: Identifying Historical Portraits with Face Recognition and Crowdsourced Human Expertise. The ACM Transactions on Interactive Intelligent Systems, 2020, 10(4):1-36.

[17] Xu X, Li Y, Jin Y. Hierarchical discriminant feature learning for cross-modal face recognition. Multimedia tools and applications, 2020, 79(45/46):33483-33502.

[18] Zhao F, Li J, Zhang L, Li Zhe, Na Sang-Gyun. Multi-view face recognition using deep neural networks. Future Generation Computer Systems, 2020, 111(2):375-380.

[19] Mokayed, H., Quan, T. Z., Alkhaled, L., & Sivakumar, V. Real-time human detection and counting system using deep learning computer vision techniques. Artificial Intelligence and Applications. 2023, 1(4): 221-229.

[20] Balfaqih M. A Hybrid Movies Recommendation System Based on Demographics and Facial Expression Analysis using Machine Learning. extraction. 2023;14(11).

[21] Balfaqih M, Altwaim A, Almohammedi AA, Yusof MH. An Intelligent Movies Recommendation System Based Facial Attributes Using Machine Learning. In2023 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA) 2023, 10(10):1-6.