

# Score Level Fusion for Iris and Periocular Biometrics Recognition Based on Deep Learning

Yufei Wang

School of Computer Science and Engineering  
Xi'an University of Technology  
Xi'an, Shaanxi 710032, China  
E-mail: 649728545@qq.com

Bo Liu

School of Computer Science and Engineering  
Xi'an University of Technology  
Xi'an, Shaanxi 710032, China  
E-mail: 2502341201@qq.com

Songze Lei

School of Computer Science and Engineering  
Xi'an University of Technology  
Xi'an, Shaanxi 710032, China  
E-mail: lei-sz@xatu.edu.cn

Huan Zuo

School of Computer Science and Engineering  
Xi'an University of Technology  
Xi'an, Shaanxi 710032, China  
E-mail: 1094282550@qq.com

Yonggang Li

School of Computer Science and Engineering  
Xi'an University of Technology  
Xi'an, Shaanxi 710032, China  
E-mail: 1984130837@qq.com

**Abstract**—Traditional iris recognition has high recognition accuracy and low misrecognition rate. However, in the case of mobile terminal or distance, the image resolution and image quality decrease, and the recognition rate also decreases. To solve the above problems, this article is based on deep learning technology, on the basis of single mode state recognition, from different levels of multimodal integration, the iris and the eyes in the score level fusion recognition research, put forward the adaptive dynamic weighted score fusion method, to determine the weighing values can adaptive algorithm of the modal, without artificial specified, dynamic weighting algorithm more flexible, stronger applicability. Experimental results of

casIA-Iris-LAMP and CasIA-Iris-Distance Iris database in Chinese Academy of Sciences show that the proposed fusion algorithm has higher recognition accuracy and better recognition performance than the single mode recognition algorithm and the traditional fractional fusion method, which proves the effectiveness of the algorithm.

*Keywords*-Multimodal Fusion; Iris Recognition; Periocular Recognition; Score Level Fusion

## I. INTRODUCTION

With the development of the Internet, human identification technology has also become more

important. Traditional identification methods such as passwords and ID cards have many defects and cannot meet people's needs for security. Biometric-based identification technology uses the inherent physiological characteristics of the human body, such as: iris, face, fingerprint, voice print, gait, etc., to automatically identify a person through a computer, which is more secure and convenient, with the advantages of not forgetting, not losing and irreplaceable. As a biometric feature, the iris has the advantages of high uniqueness, stability and non-intrusiveness.

However, in some specific situations, such as mobile, long distances and weak constraints (where the person is moving slowly), it is difficult to capture a complete and clear iris texture, which will lead to reduced recognition rates. To solve this problem, the periocular area can be fused with the iris for multimodal fusion recognition. The periocular area, as a local feature of the face, is perfectly integrated with the iris position. When the iris image is acquired using a mobile device, it may not be possible to acquire the whole face, whereas the periocular region image can be acquired simultaneously with the iris, making the periocular region a better choice compared to the face. Therefore it is important to study the fusion of iris and periocular recognition methods to improve recognition accuracy and stability and to overcome the anti-forgery nature of unimodal feature recognition techniques.

Liu et al [1] were the first to use convolutional neural networks for iris recognition and obtained more superior recognition results. mohtashim Baqar et al used an improved deep belief network for iris recognition [2]. Zhao et al used a capsule network in the network structure and improved the network structure for iris feature extraction and recognition to improve the problem of small iris samples [3] Zhao and Kumar et al [4]

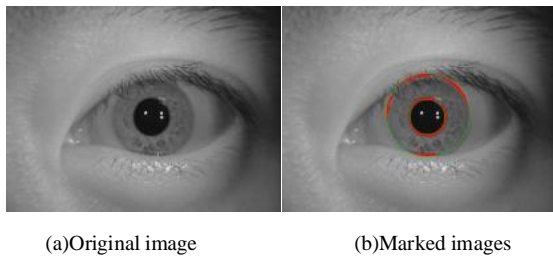
proposed a new network model that recovered periocular features by combining semantic information, where higher matching accuracy was achieved using a small number of samples and better recognition results were obtained. proenca and Neves et al [5] improved the performance of periocular recognition based on discarding iris and sclera features in visible light and recognising other regions around the eye. Performance of periocular recognition. Chen et al [6] used additive rules to fuse two biometric features, face and iris, in the score level for fusion recognition, and used a wavelet neural network classifier for classification recognition. The experimental results demonstrated that multimodal biometric verification is more reliable and accurate than single biometric verification. Son et al [7] fused two biometric features, face and iris, in the feature layer Direct linear discriminant analysis was used in the fusion to reduce the dimensionality of the feature information, and finally nearest neighbours were used for classification and recognition. Teddy Ko et al [8] discussed the impact of multiple fusion methods and image quality on the overall recognition accuracy on fingerprints, faces, and irises. Zhang et al [9] designed a deep feature fusion network using a Maxout structure for feature extraction that The weighted fusion on the feature layer for iris and face achieved good recognition results, and the model occupies less storage space and is suitable for mobile.

Therefore, this paper proposes an adaptive weighted fusion method based on iris and periocular score layers based on deep learning techniques. It has been experimentally demonstrated that the fusion method can significantly improve the recognition accuracy compared with unimodal recognition.

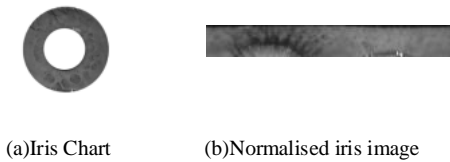
## II. FUSION USION METHODS

### A. Iris and periocular image pre-processing

Iris image preprocessing mainly includes two steps: segmentation and normalisation. In this paper, the iris is segmented using the Osiris V4.1 system [10]. Osiris uses the Viterbi algorithm [11], which is able to perform at low resolution, and which is able to quickly locate the iris contour. The green circle indicates the outline of the iris and the red marker is the non-iris portion of the annular region. This is shown in Fig.1.



Figur 1. Iris pre-processing diagram



Figur 2. Iris normalisation

The iris normalisation process is carried out with the aid of a rubber sheet model, which is then used to obtain a rectangular iris image for classification recognition. As shown in Fig.2, where (a) shows the original image of the iris after segmentation, as a circular region, and (b) shows the normalised iris image, after which the iris is converted to a rectangular image.

Pre-processing of the periocular area includes periocular angle correction and periocular normalisation. The periocular image is

downsampled from  $640 \times 480$  pixels to  $224 \times 224$  pixels by normalisation. In order to prevent the periocular image from not being wide or tall enough during normalisation, this paper uses bi-cubic interpolation, using the surrounding skin texture to fill in. A normalised periocular image with higher feature differentiation is obtained.

### B. Traditional score level fusion methods

Traditional score layer fusion methods include those based on combined strategies [12][13] and manual weighting methods. The commonly used combinatorial decision strategy fusion methods are: maximum value, minimum value, multiplication rule, and addition rule.

#### 1) Maximum value rule

$$f_i = \max n_i^m \quad (1)$$

#### 2) Minimum value rule

$$f_i = \min n_i^m \quad (2)$$

#### 3) Multiplication rules

$$f_i = \prod_{m=1}^M n_i^m \quad (3)$$

#### 4) Addition rule

$$f_i = \sum_{m=1}^M n_i^m \quad (4)$$

In the above equation  $n_i^m$  denotes the matching score of the  $m$ th modality after normalization.  $f_i$  denotes the score after fusion.

#### 5) Manual setting of weights

In this paper, before fusing the two modalities of iris and periocular, the weights of both are divided and the iris and periocular weights are set to weighted fusion with weighting ratios from 1:9 to 9:1 to obtain the fusion accuracy under each weighting ratio. The weighting formula is shown in Equation (5).

$$F_{score} = \alpha I_{score} + \beta E_{score} \quad (5)$$

$F_{score}$  is the fused score,  $I_{score}$  denotes the iris matching score and  $E_{score}$  denotes the periocular matching score.  $\alpha$  and  $\beta$  are the weighting coefficients for the iris and periocular modal scores respectively, taking values in the range [0,1], and  $\alpha+\beta=1$ , set by hand. The method was used to find the best matching weighting coefficients through multiple experiments.

### C. Our iris and periocular score level fusion

Score layer fusion is a method to fuse the scores of different biometric images or data after feature extraction and feature matching. The fused score is then used for identification. The main idea of score layer fusion is to use combinatorial decision strategy, which includes addition, multiplication, weighted sum, Max and min rules, etc.. Compared with the feature layer fusion, the score layer fusion does not damage the integrity of each biometric recognition, and is easier to carry out the fusion. The fusion method is simple and the fusion speed is fast. However, the score layer fusion will lose a lot of feature information, and it is difficult to achieve higher recognition effect.

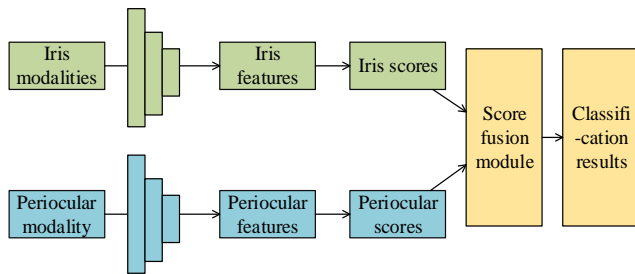
Biological characteristics themselves have very rich characteristic information, and there are obvious differences in the characteristics of different organisms. In single mode recognition,

due to the limitation of acquisition equipment, the surrounding environment, acquisition distance and other influences, the collected biometric image data received serious interference, resulting in a lot of adverse impression for the final recognition. Therefore, multiple modal biometrics can be combined to enhance the diversity of feature information and improve the recognition performance. According to the level of fusion, pixel layer fusion contains the most abundant feature information. Due to the complexity and diversity of data types, it is difficult to operate during the fusion, and data cannot be effectively fused. Feature layer fusion is to fuse the extracted feature data together on the basis of the extracted feature data. Therefore, some features of the fused data are lost. Score layer fusion uses the score obtained after feature matching, and feature information loss is large. However, the score data is easy to obtain and the fusion operation is simple. It has become the most widely used multi-mode fusion technology in deep learning. One difficulty of score layer fusion is that too many distinguishing features are lost in the fusion process and too little feature information can be used, which limits the fusion performance. In order to make up for the loss of characteristic information on the score layer as much as possible. In this chapter, when selecting score layer fusion, more feature information is extracted through deep learning, and the fusion strategy is improved by adding weights to improve the performance of score layer fusion recognition.

Score layer fusion is the most commonly used method in multimodal fusion. The fusion makes full use of the score information of different modalities and does not result in dimensional disasters. In this paper, we propose an algorithm for fusion of iris and periocular score layers based on a combined decision strategy with set

weighting factors. The fusion problem of the iris and periocular modalities is regarded as a combined score determination problem, and the fusion rules are analysed in depth. By conducting theoretical and experimental studies on the information of each modality during fusion and the fused scores, a fusion formula that can automatically obtain the weight coefficients of each modality is finally obtained.

Fig.3 shows a schematic diagram of iris and periocular fusion in the score layer. First, the iris is extracted by performing iris segmentation on the acquired periocular image. Image pre-processing is performed on the iris and periocular images to obtain the two modalities of iris and periocular that can be used for recognition. Next, the deep learning algorithm is used to extract and match the features of the iris and periocular modalities to obtain the iris and periocular scores respectively. Finally, the scores of the two modalities are fused and recognised by the score fusion module to obtain the final classification result.



Figur 3. Schematic diagram of scorer layer fusion

Aiming at the problems exposed by the traditional combinatorial decision fusion strategy, a weighted score layer fusion algorithm is proposed in this section. The algorithm fully considers the importance of the matching score of each mode to the recognition result. By dynamically adding a weight coefficient to the fraction of each mode, the fusion recognition is

carried out at last. The algorithm studies the importance of each modal fraction to the recognition result, which improves the recognition accuracy significantly, alleviates the over-fitting problem, and improves the overall recognition performance.

The most widely used existing method is to use manually set filters to obtain features and finally a distance metric to obtain a matching score. Commonly used distance similarity methods are: cosine distance, Hamming distance and Euclidean distance. In this paper, we start from deep learning and feature extraction of iris and periocular images by building deep neural networks. In this paper, ResNet18 is used as the feature extraction network with Softmax function and Categorical\_crossentropy loss function to achieve the matching of iris and periocular modal scores.

In the last layer in the neural network, this paper uses the Softmax function to implement the function of neural network classification. In a multi-classification network, Softmax implements the function of normalisation and the output is a vector of predicted probabilities, indicating the probability of the input data belonging to each category, its beneficial to the data fusion study in this paper. The Softmax formula is expressed as shown in equation (6).

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \tag{6}$$

In the above equation, z denotes the input data of the last layer and z\_i denotes the ith data in z. After Softmax calculation on the input data, the obtained data is normalised to between [0,1], which is the probability value of each class.

The recognition study in this paper belongs to

the domain of multi classification, so Categorical\_crossentropy is used as the loss function of the fusion network to calculate the loss of real and predicted data, and its expression is shown in Equation (7).

$$L = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_{i=0}^{N-1} \sum_{c=0}^{C-1} y_{ic} \log p_{ic} \quad (7)$$

In the above equation, N denotes the number of overall samples and K denotes the number of final classifications. The  $y_{ic}$  is a symbolic function taking the value 0 or 1. The meaning of  $p_{ic}$  is the predicted probability c of the first category.

In this paper, score layer fusion is performed in a weighted manner. The iris and periocular are fused on the score layer by an automatic weight setting method. By setting the weights at the time of fusion, it is hoped to obtain better recognition results than under unimodal recognition and simple rule fusion.

On the acquisition of adaptive weights. In this paper, adaptive weighting is applied to the iris and periocular scores by analysing the correlation between each modality and ultimately the degree of importance to the recognition result. The formula for obtaining the weights for the iris is shown in Equation (8).

$$I_{weight} = \frac{I_{score}}{I_{score} + E_{score}} \quad (8)$$

The formula for obtaining the periocular weights is shown in Equation (9).

$$E_{weight} = \frac{E_{score}}{I_{score} + E_{score}} \quad (9)$$

The overall fusion formula is shown in Equation (10).

$$F_{score} = I_{weight} \times I_{score} + E_{weight} \times E_{score} \quad (10)$$

$F_{score}$  is the fused score,  $I_{score}$  denotes the iris matching score and  $E_{score}$  denotes the periocular matching score. The weights are set by adaptive weighting so that each modality plays a maximum role in the recognition result when fused. Thus, the recognition effect is improved.

### III. EXPERIMENTAL RESULTS AND ANALYSIS

In this paper, the CAS database is selected for the exploration of score layer fusion algorithm. The experimental results of periocular and iris recognition under unimodality are shown first, followed by the fusion experimental results of additive rule, maximum value rule and minimum value rule respectively, and finally the experimental results after score layer weighting are shown.

#### A. Experimental protocol

The experimental hardware system version of this paper is Linux Ubuntu16.04. A Nvidia GeForce RTX2080Ti GPU is used for accelerated computing to speed up the convergence of the network model and increase the computing speed. In addition, version 3.6.8 of python environment was chosen. The Tensorflow framework was used for network construction. The specific configuration is shown in Table 1. This table mainly describes the configuration of the processor, graphics card, operating system,

framework and compiler.

TABLE I. ENVIRONMENT CONFIGURATION

Hardware Environment	Processor	Inter(R) XeonGold b6254
	Graphics card	Nvidia GeForce RTX2080Ti
Software Environment	Operating System	Linux Ubuntu 16.04
	Framework	Tensorflow-GPU
	Programming language	Python3.6.8
	Compiler	PyCharm

In this study, two Iris datasets from Chinese Academy of Sciences were selected: CasIA-Iris-LAMP and CasIA-Iris-Distance. The CasIA-Iris-LAMP dataset includes 791 categories of Iris images, with a total of 15,820 images. In these Iris images, there are also an equal number of corresponding periocular images. Casia-iris-distance data set includes 284 categories of Iris images, with a total of 5073 images. Among these Iris images, an equal

number of corresponding periocular images are also included. The above data set is divided in a 3:1:1 ratio. The Casia-Iris-LAMP dataset contains 791 categories, including 9492 training sets, 3164 validation sets and 3164 test sets respectively. Casia-iris-distance data set contains 284 categories, including 3160 training sets, 1024 verification sets and 889 test sets. The image size is 224 pixels  $\times$  224 pixels. Detailed parameters are shown in Table 2.

TABLE II. EXPERIMENTAL DATA SETS

Dataset	Category	Iris images	Periocular images	Training set	Validation set	Test set
CASIA-Iris-Lamp	791	15820	15820	9492	3164	3164
CASIA-Iris-Distance	284	5073	5073	3160	1024	889

In this paper, score layer fusion is used. After obtaining pre-processed iris image and pre-processed eye image, model parameters need to be set in advance. Before fusion, iris and periocular recognition models in single mode were obtained, and the model training parameters were shown in Table 3. The input dimension is 224 $\times$ 224 $\times$ 3, the number of images input by the model is 32, the number of steps trained by the model is 2000, and the gradient descent function used by the model is RMSprop function.

TABLE III. EXPERIMENTAL CONFIGURATION

Relevant parameters	Parameter settings
input_shape	224 $\times$ 224 $\times$ 3
batch_size	32
epoch	2000
optimizers	rmsprop
learning_rate	10 <sup>-3</sup>
rho	0.9
decay	10-6

Before the experiment, we selected the iris data set of Chinese Academy of Sciences,

preprocessed the iris image and periocular image in advance, and configured the experiment in detail. In the performance test of score layer fusion, this paper mainly uses the method of calculation accuracy to evaluate the accuracy of single mode algorithm and score layer fusion algorithm.

### B. Results of fusion experiments

The experiments in this section include unimodal periocular and unimodal iris recognition, as well as multimodal score layer fusion. The fusion of score layers compares the traditional methods based on the maximum value, minimum value, multiplication rules, addition rules in the simple combination strategy and the adaptive weighted rules proposed in this paper, respectively, and also compares the results of iris and periocular recognition in unimodality. The experimental results are shown in Table 4.

TABLE IV. COMPARISON OF EXPERIMENTAL RESULTS

Dataset	Fusion rule	Recognition accuracy
CASIA-Iris-Lamp	Maximum value	0.992055
	Minimum value	0.955830
	Multiplication	0.992017
	Additive	0.992373
	Unimodal periocular	0.982617
	Unimodal iris	0.981252
	Adaptive weighting	0.993831
CASIA-Iris-Distance	Maximum value	0.961754
	Minimum value	0.866141
	Multiplication	0.972512
	Additive	0.961754
	Unimodal periocular	0.933633
	Unimodal iris	0.894263
	Adaptive weighting	0.985376

Several conclusions can be drawn from the experiments in Table 4, as follows.

1) In terms of the dataset, the recognition accuracy on the CASIA-Iris-Lamp dataset is higher than that on the CASIA-Iris-Distance dataset, regardless of whether the recognition is performed under multimodal recognition or unimodal recognition. Moreover, the recognition effect of unimodal periocular is better than that of unimodal iris recognition.

2) In terms of modality, the recognition accuracy of score layer fusion was above 98.9% on both CASIA-Iris-Lamp dataset and 98.2% on both unimodal recognition accuracies. On the CASIA-Iris-Distance dataset, the recognition accuracy of score layer fusion was above 97% in all cases, and the unimodal recognition accuracy was lower. Therefore, it can be concluded that the recognition accuracy of periocular and iris score layer fusion is higher and better than that of unimodal periocular and unimodal iris recognition, except for the minimum value fusion rule.

3) In terms of fusion rules, it was shown by experimental results on both datasets. The traditional score layer fusion has the best recognition by the multiplicative rule and the worst recognition by the minimum value rule. The adaptive weighted score layer fusion proposed in this paper achieves the best recognition results.

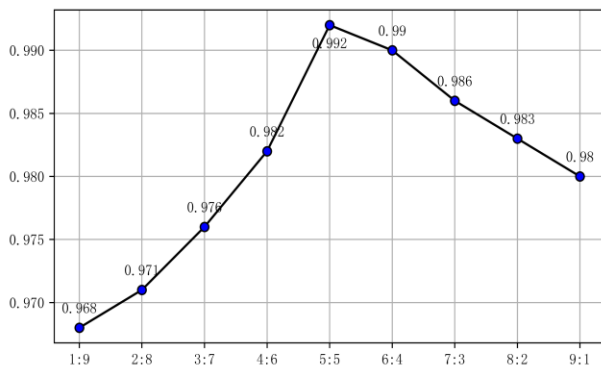
In this paper, we set the weight range to [0.1, 0.9] in the manual weighted fusion experiments for iris and periocular in the score layer. The weights are set and calculated in order from 1:9 to 9:1 for the iris and periocular. The weights were first set to 0.1 for the iris and 0.9 (1-0.1) for the periocular area, and were updated at a rate of 0.1 each time. The results of the experiments are shown in Table 5.



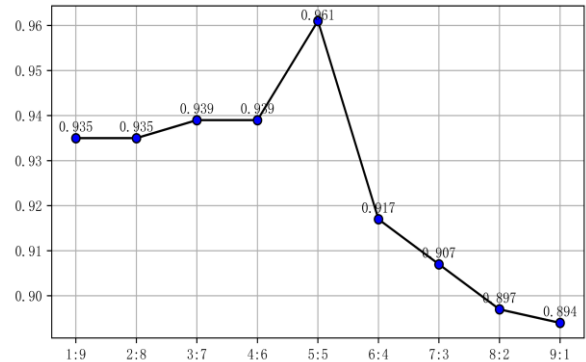
TABLE V. COMPARISON RESULTS OF DIFFERENT WEIGHTS

DataSet	CASIA-Iris-Lamp	CASIA-Iris-Distance
1:9	0.968541	0.935883
2:8	0.971083	0.935883
3:7	0.976167	0.939257
4:6	0.982840	0.939257
5:5	0.992373	0.9617547
6:4	0.990149	0.917885
7:3	0.986653	0.907761
8:2	0.983158	0.897637
9:1	0.980934	0.894263

Table 5 combined with Fig.4 and Fig.5 shows that the fused recognition accuracy steadily increased as the weight of the iris increased until it reached an optimum of 99.2% at a weight of 0.5 for the iris. This experiment shows that on both the CASIA-Iris-Lamp and CASIA-Iris-Distance datasets, the best fusion scores were obtained for both iris and periocular modalities at a weighting ratio of 1:1. As the weight of the iris continued to increase, the fusion recognition rate gradually decreased to 98%. The experiments show that the recognition effect after fusion is strongly influenced by the iris. Good performance was also obtained with manually weighted fusion.



Figur 4. Fusion plot of different weights on CASIA-Iris-Lamp



Figur 5. Schematic representation of the fusion of different weights on CASIA-Iris-Distance

#### IV. CONCLUSIONS

In this paper, score layer fusion of iris and periocular is studied from the perspective of multimodal fusion. In response to the degradation of recognition rate due to image quality degradation in mobile or remote situations, a score layer fusion method based on adaptive weighting was proposed, where the weights of the modalities can be determined adaptively. Experimental comparisons of the methods were carried out in the CAS iris library. The experimental results show that the adaptive weighted score layer fusion achieves better recognition results, demonstrating the effectiveness of the algorithm in this paper, and also proving that the accuracy of the score layer-based multimodal fusion algorithm is significantly better than that of the single-modal recognition algorithm.

#### ACKNOWLEDGEMENT

This work was supported by National Joint Engineering Laboratory of New Network and Detection Foundation (grant no. GSYSJ2016008). The experiments in this paper use the CASIA-Iris collected by the Chinese Academy of Sciences' Institute of Automation (CASIA).

## REFERENCES

- [1] Liu N , Zhang M , Li H , et al. DeepIris: Learning Pairwise Filter Bank for Heterogeneous Iris Verification [J]. Pattern Recognition Letters, 2015, 82(2):154-161.
- [2] Baqar M, Ghani A , Aftab A , et al. Deep belief networks for iris recognition based on contour detection[C]// International Conference on Open Source Systems & Technologies. IEEE, 2017.
- [3] T. Zhao, Y. Liu, G. Huo and X. Zhu, "A Deep Learning Iris Recognition Method Based on Capsule Network Architecture," in IEEE Access, vol. 7, pp. 49691-49701, 2019, doi: 10.1109/ACCESS.2019.2911056.
- [4] Zijng Zhao,Ajay Kumar 0001. Accurate Periocular Recognition Under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network. [J]. IEEE Trans. Information Forensics and Security, 2017, 12(5).
- [5] Hugo Proença, João C. Neves. Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning Frameworks. [J]. IEEE Trans. Information Forensics and Security, 2018, 13(4).
- [6] Chen C H, Chu C T. Fusion of Face and Iris Features for Multimodal Biometrics[C]// International Conference on Advances in Biometrics. Springer-Verlag, 2006.
- [7] Son B, Lee Y . Biometric Authentication System Using Reduced Joint Feature Vector of Iris and Face[C]//International Conference on Audio-&Video-based Biometric Person Authentication. Springer-Verlag, 2005.
- [8] Ko T .Multimodal biometric identification for large user population using fingerprint, face and iris recognition[C]// Applied Imagery & Pattern Recognition Workshop. IEEE, 2005.
- [9] Qi, Zhang, Haiqing, et al. Deep Feature Fusion for Iris and Periocular Biometrics on Mobile Devices[J]. IEEE Transactions on Information Forensics & Security, 2018.
- [10] Othman N, Dorizzi B, Garcia-Salicetti S.OSIRIS: An open source iris recognition software [J]. Pattern Recognition Letters, 2016, 82(oct.15):124-1
- [11] Sutra G, Garcia-Salicetti S, Dorizzi B. The Viterbi algorithm at different resolutions for enhanced iris segmentation[C]// Iapri International Conference on Biometrics. IEEE, 2012.
- [12] Snelick Robert,Uludag Umut,Mink Alan,Indovina Michael,Jain Anil. Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. [J]. IEEE transactions on pattern analysis and machine intelligence, 2005, 27(3).
- [13] Anil Jain,Karthik Nandakumar,Arun Ross. Score normalization in multimodal biometric systems [J]. Pattern Recognition, 2005, 38(12).