J. Meas. Sci. Instrum., 2024, 15(1): 54-63.
DOI: 10.62756/jmsi.1674-8042.2024006
http://jmsi.publish.founderss.cn/
jmsi@nuc.edu.cn

Using dynamic data augmentation and contrastive learning for iris verification

HE Landi, JI Dezan, DONG Xingchen, SU Mingxin, ZHOU Weidong*

School of Microelectronics, Shandong University, Jinan 250100, China

***Corresponding author:** ZHOU Weidong (wdzhou@sdu.edu.cn)

Received: September 29, 2023 **Revised:** November 18, 2023 **Accepted:** December 26, 2023

Abstract: Iris verification has gained extensive attention because of its uniqueness, stability, and non-invasiveness. Deep learning techniques have made significant progress in the field of iris verification. By using convolutional neural networks (CNNs), features of iris images can be automatically extracted and learned, enabling high-precision identity verification. However, challenges such as intra-class variability and limited dataset size can compromise verification accuracy. To address these issues, we proposed an iris verification method based on dynamic data augmentation and contrastive learning. Four data augmentation strategies were carefully designed for online iris enhancement and dataset expansion, and the accuracy of iris verification was further improved by using data augmentation probability scheduler (DAPS). The MobileNetV3 was employed as the backbone network, which was optimized with contrastive learning for 3-channel iris pairs. The proposed method was evaluated on two benchmark iris databases, CASIA-V4-Interval and CASIA-V4-Thousand, achieving high accuracies of 99.85% and 98.82%, respectively. Experimental results demonstrate that the proposed method can achieve competitive performance with a small number of training samples.

Key words: iris verification; contrastive learning; convolutional neural network (CNN); data augmentation

0 Introduction

Biometric recognition is a method used to verify identities and can be used as a replacement for traditional identity recognition methodologies such as passwords, keys, and ID cards, which are easily lost or forged. Biometric recognition can quantify and encode a person's fingerprint, voiceprint, brainwave, palm print, face, walking posture, body or behavioral features, and has excellent anti-forgetting and anti-loss characteristics. However, many bodily features are at risk of coercion or forgery, and some features will gradually lose effectiveness over time. Compared with other biometrics, iris possesses many advantages including uniqueness, randomness, stability, and non-invasive acquisition. Once the structure of the iris is formed, it stays stable during infancy and does not change with the increase of age, making it unique for each individual[1].

Iris has been widely used in personal identity verification due to its specificity, stability, and non-invasiveness[2]. A complete iris recognition system usually includes iris image acquisition, database management, iris localization, segmentation, normalization, feature extraction, and feature matching. Feature extraction and feature matching are the core parts of the recognition system and directly affect the system's performance[3].

In recent years, deep learning models such as convolutional neural networks (CNNs) have been utilized for iris recognition. CNNs are capable of extracting features from highly complex datasets, and various techniques have been proposed to improve CNNs, such as batch normalization, data augmentation, regularization, dropout, and different activation functions. To reduce computational complexity of CNNs, improved pre-training network structures have been designed[4]. Previous studies on iris recognition have shown that neural network-based methods can effectively learn the inherent features of iris images[5,6], and outperform classical iris matching methods such as IrisCode in terms of accuracy and robustness[7].

In this study, we proposed an accurate and fast iris verification method based on dynamic data augmentation and contrastive learning. Contrastive learning[8] was adopted by transforming the original multi-class task into a binary classification task. MobileNetV3[9] was employed as the backbone network for iris verification to

achieve a smaller network size and shorter inference time. Multiple online data augmentation strategies were designed to provide more training data without changing the size of the training set. Moreover, the use of dynamic data augmentation effectively mitigated the drawbacks of traditional data augmentation. Experimental results demonstrate that our approach can achieve competitive performance with a small number of training samples.

Our contributions in this study can be summarized as follows.

1) Proposing a contrastive learning-based iris verification method, which is optimized for 3-channel iris pairs.

2.) Applying four effective data augmentation techniques to perform online augmentation of iris images and increase the data space in the training stage by adding more negative examples and refreshing their combinations.

3) Proposing the data augmentation probability scheduler (DAPS) to further enhance the accuracy of the model by dynamically adjusting data augmentation probability during the training phase.

The remainder of this paper is organized as follows: In Section 1, we review related works on iris recognition and deep learning-based verification methods. Section 2 introduces our proposed iris verification method in detail. In Section 3, we introduce the iris dataset used in our experiments and analyze the experimental results. Finally, Section 4 concludes the paper with a summary of our contributions.

1 Related work

Effectively extracting iris features from iris images is crucial for improving the accuracy of iris recognition. Daugman et al. [10-13] proposed the use of integro-differential operator to determine the inner and outer boundaries of the iris to obtain an iris template, which was then normalized using Daugman's rubber sheet model. 2D Gabor filter was subsequently used to encode the iris template into IrisCode. The Hamming distance between different IrisCodes was utilized to generate recognition and verification results. Subsequently, more handcrafted feature engineering methods were proposed, including log-Gabor filter[14], discrete cosine transform (DCT) [15], principal component analysis (PCA) [16,17], gray level co-occurrence matrix (GLCM) [18,19], local binary patterns (LBP) [20,21], and scale-invariant feature transform (SIFT) [21,22]. These well-designed feature extraction methods can achieve significant performance under ideal conditions, but they are difficult to achieve desirable recognition results when dealing with non-ideal data such as blur, eyelash and eyelid occlusion[23,24].

The powerful ability for feature extraction of deep neural networks has been demonstrated by their tremendous success in the field of computer vision. By utilizing fine-tuning techniques to incorporate off-the-shelf weights pre-trained on other large datasets, overfitting caused by insufficient data can be effectively avoided. Nguyen et al.[25] adopted five pre-trained CNNs for feature extraction and trained a support vector machine (SVM) as a classifier, and then classified using an SVM. The results demonstrate that weights pre-trained on other large-scale datasets can be well-suited for iris recognition tasks. Deep neural networks can also be directly trained on iris datasets to generate an end-to-end feature classifier. Gangwar et al.[26] used a carefully designed convolutional neural network for feature extraction and used softmax as the output layer of the network, and stochastic gradient descent (SGD) as the optimizer. Compared with traditional classifier methods, this approach enables the classifier to directly learn from iris datasets, making it more robust in cross-sensor and cross-dataset experiments.

However, classifiers that utilize softmax as the output layer can only recognize identity classes that have been previously trained, which means that the networks must be re-trained when a new class is added[27]. Iris similarity networks was developed by taking pairs of samples as input and comparing the similarity between the two samples[28]. During training, the dataset generated by all possible combinations of training samples is fed into the neural network. Therefore, this approach can achieve high accuracy with few training samples, and there is no need to retrain when new classes are added. Nianfeng et al.[2] proposed a pairwise network that accepts two input images from two heterogeneous sources. They used six-fold data augmentation to correct rotation variance and achieved competitive performance in both cross-resolution and cross-sensor iris verification. Liu et al.[28] proposed a hybrid framework approach to reduce the computational burden of large-scale iris identification scenarios.

2 Proposed method

The overview of the proposed iris verification framework is shown in Fig. 1. Iris verification can be considered as a binary classification task. First, iris

images are preprocessed, then three preprocessed images are selected according to specific rules and combined into a 3-channal iris pair, which is stored in the iris database. During training, the three images in each iris pair undergo individual online data augmentation, and the probability of data augmentation is dynamically controlled by the proposed DAPS. Subsequently, the augmented 3-channal iris pairs are fed into CNN to determine whether the images in the pair come from the same eye. The four main components of the framework, namely preprocessing, dynamic data combination, online data augmentation, and CNN classification, will be discussed in detail in the following sections.

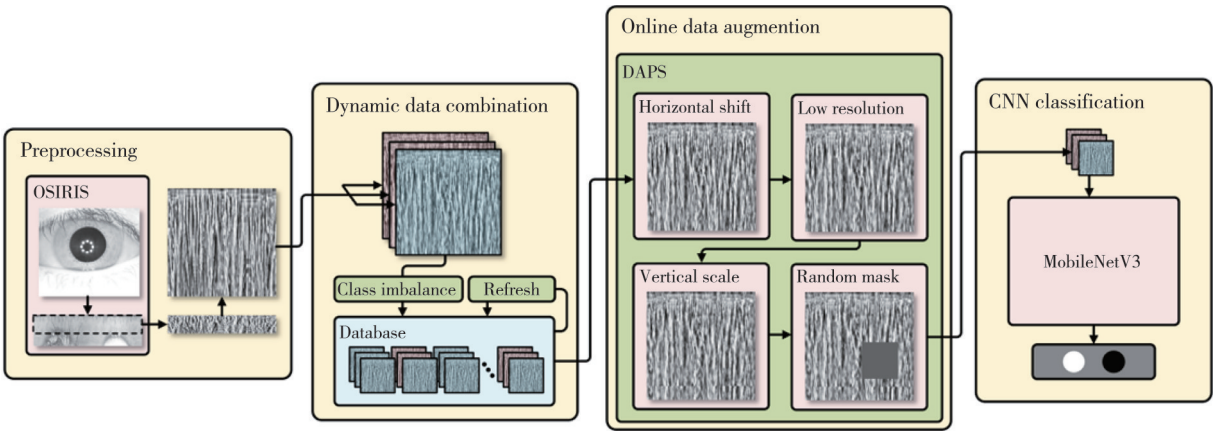

Fig. 1 Overview of proposed iris verification framework

2.1 Preprocessing

As shown in Fig. 2, the preprocessing of iris images involves four steps: Location and segmentation, Normalization, enhancement, and transformation. We utilized the open-source tool OSIRIS[29] to perform the segmentation and normalization. OSIRIS recognizes the iris boundary and segments the iris region, then maps the segmented image into polar coordinates to generate a size-invariant iris band. In this study, all iris images are normalized to a size of 103×360 iris band, and to reduce the interference from eyelids and eyelashes on recognition, only the first 60 rows of the iris band are retained.

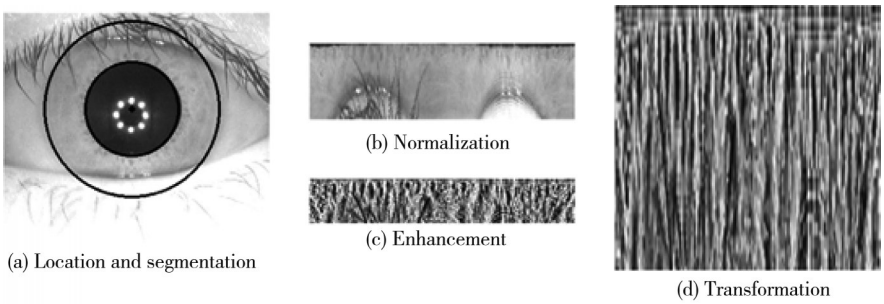

Fig. 2 Examples of the outputs of different preprocessing steps

Histogram equalization (HE) adjusts the gray distribution of normalized images to enhance the contrast. This technique can reduce the impact of different illumination environments on iris images, but it may result in decreased details in bright and dark regions for images with excessively concentrated gray values[30]. Therefore, we used contrast limited adaptive histogram equalization (CLAHE) [30] with carefully tuned parameters to enhance the contrast of iris images and preserve more details. Fig. 3 shows normalized images with different enhancement methods, as well as their corresponding gray-level histograms and cumulative distribution function (CDF) curves. Experimental results demonstrate that the use of CLAHE can achieve a smoother gray-level distribution and significantly improve accuracy.

In the transformation step, we first scaled the pixel values of images to the range of (0, 1) and resized the image to a resolution of 224×224 pixels. To speed up the convergence of the model, we transformed the image data into a standard Gaussian distribution. Before training, we randomly sampled the resized images from the training set, and computed their mean and standard deviation. Then, we normalized all resized images using the pre-calculated mean and standard deviation.

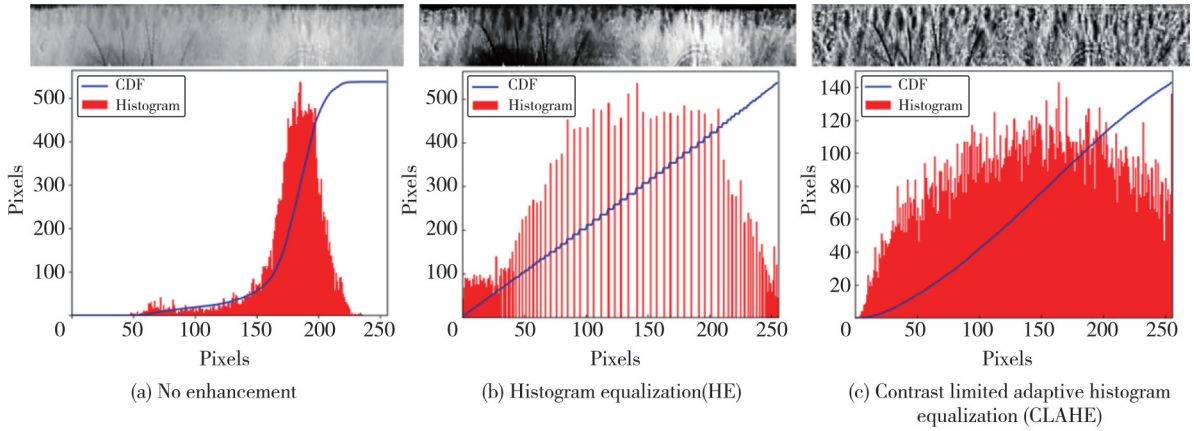

Fig. 3 Illustration of gray histogram and CDF curve of different iris enhancement methods

2.2 Dynamic data combination

Contrary to the normal 2-channel iris pairs as input images for contrastive learning, we proposed 3-channel contrastive learning for iris verification. We used $(X_1, (X_{21}, X_{22}))$ to represent the 3-channel iris pair, where X_1 is considered as the base image and is fixed in channel 0; X_{21} and X_{22} are contrastive images from the same class of different images, and are fixed in channel 1 and channel 2, respectively. In positive pairs, the base images and the contrastive images are from the same class, while in negative pairs, the base images and the contrastive images are from different classes.

In existing public iris datasets, iris images from both eyes of the same person are collected. But in actual iris recognition scenarios, only one iris image is usually used for recognition. According to our experimental findings, iris images from both eyes of the same person are more similar than iris images from both eyes of different people. Hence, we consider iris images from both eyes of the same person as different classes. But in negative pairs, the base images and the contrastive images are not from the same person.

Assuming there are C classes (from different people), each class has P iris images, then at most $C \times P \times P \times (P-1)$ positive pairs and $C \times P \times (C-1) \times P \times (P-1)$ negative pairs can be generated. Typically, the number of negative pairs is far greater than the number of positive pairs. To ensure the balance between positive and negative sample numbers, we randomly selected $C \times P \times P \times (P-1)$ negative pairs (under the condition of ensuring the same probability of occurrence for the base image) and all positive pairs to put into the iris database.

To expand the diversity of negative samples during training phase, we proposed a dynamic data combination approach, which refreshes the iris database with new negative pairs after a fixed number of training epochs. This approach allows the neural network to encounter a wider range of negative pairs without increasing the size of the iris database. In our experiments, we refer to the parameter controlling this refresh as "refresh epochs".

Typically, different classes have equal importance, and class balance is crucial for ordinary image classification tasks. However, reducing the risk of accidental unauthorized access is essential for iris recognition tasks. Hence, we deliberately introduce class imbalance by increasing the number of negative pairs several times to improve the false rejection rate (FRR) with a low false acceptance rate (FAR).

2.3 Online data augmentation

Data augmentation is a prevalent strategy deployed in the deep learning, aiming to enlarge the size of the training dataset. By using this approach, the model is capable of recognizing the data and its variations that are not present in the original dataset, thus improving its robustness and generalizability. The traditional data augmentation approach involves generating new samples from the existing data, which may require additional storage space.

Online data augmentation addresses these limitations by conducting the augmentation process in real-time, during the training stage. This approach is helpful to dynamically increase each batch of data fed into the model, thus generating ever-changing and diverse datasets. Furthermore, online data augmentation can affect the model's performance in real-time.

The learning rate scheduler is a commonly utilized optimization technique in deep learning. It enables the automatic adjustment of the learning rate of a model based on its performance during the training stage, thus improving training efficiency and preventing overfitting.

Inspired by the learning rate scheduler, we proposed

the DAPS to dynamically adjust the augmentation probability during the training phase in a novel manner. The DAPS significantly improves accuracy and stability while reducing overfitting. In our work, we set the probability of each augmentation method to be equal to p, and adjust it during the training phase using the step decay method, where the learning rate is decreased by a fixed factor. The formula of p in training phase can be expressed as:

$$p = \frac{\left\lfloor 10\left(p_0 + \frac{p_1 - p_0}{Epoch} Epoch_{\text{cur}} \right) \right\rfloor}{10}, \tag{1}$$

where p_0 denotes the initial augmentation probability, p_1 denotes the final augmentation probability, $Epoch_{\text{cur}}$ denotes the current training epoch, and $Epoch$ denotes the total number of training epochs.

In this work, four augmentation methods were designed to simulate changes in iris images under different conditions, including horizontal shift, low resolution, random mask, and vertical scale. The details of each augmentation method are described as follows:

1) Horizontal shift

Due to head movement, two iris images from the same eye may present a difference that rotates around the pupil as the center. Since the iris image has been mapped to polar coordinates, we simulated this difference by horizontally shifting the iris image. Fig.4 depicts the iris images after horizontal shift. The horizontal shift amount will be randomly selected within $[-45, 45]$, meaning that it can simulate an iris rotating clockwise or counterclockwise by up to 45°. An example of the horizontally shifted iris image is shown in Fig.4(b).

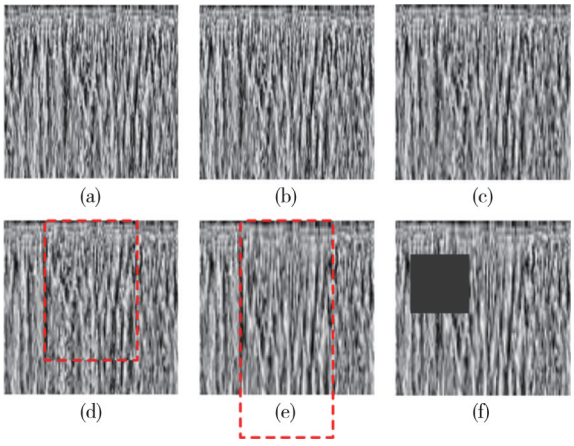

Fig. 4 Output examples of four augmentation methods. (a) an example of the preprocessed original iris image. (b), (c) and (f) are examples of iris images after applying horizontal shift, low resolution, and random mask, respectively. (d) and (e) are two examples of vertical scale, where heights of images are h=168 and h=280, respectively

2) Low resolution

When the subject is too far away from the iris camera during acquisition, the iris may occupy a smaller proportion in the image, resulting in a lower clarity of the normalized iris image. To better overcome the impact of varying acquisition distances, we simulated low-quality iris images by reducing the clarity of iris images. Firstly, we reduce the size of the iris image to a 112×112 resolution, and then increased it to a 224×224 resolution without using any interpolation methods. The iris image after reducing the resolution is shown in Fig.4(c).

3) Vertical scale

To mitigate the impact of distortions in iris images caused by changes in shooting angles, iris segmentation algorithm imprecision, pupil constriction, and so on, we performed vertical scaling on a portion of the iris image, either stretching or compressing it. We randomly selected an image segment of width $w \in [75, 224]$ and scale the height h of the image from the original 224 to $[168, 280]$. All scaling operations were performed using bilinear interpolation. If the height of the iris image was reduced, the missing parts were filled with the original image. Fig.4(d) and (e) show examples with $h = 168$ and $h = 280$, respectively.

4) Random mask

To mitigate the impact of obstructions such as glare and debris on iris images and improve model robustness, we implemented random masking of the images with a 75×75 mask. The mask pixel values $v \in [v_{\min}, v_{\max}]$, where $v_{\min}$ is the minimum pixel value in the iris image and $v_{\max}$ is the maximum pixel value in the iris image. The resulting iris image after random masking is shown in Fig.4(f). It is noteworthy that when the mask partially extends beyond the boundaries of the iris image, the portion of the mask beyond the boundary is automatically wrapped to the other side of the iris image.

2.4 Network architecture

Due to the requirement for real-time recognition and the limited computational capabilities of edge devices used for iris recognition in practical applications, we chose MobileNetV3 for iris recognition, which is designed specifically for mobile devices. MobileNetV3 could achieve high classification accuracy and fast recognition speed while requiring low computational resources. The network architectures used in this experiment are two versions of MobileNetV3, namely MobileNetV3 small and MobileNetV3 large. The difference between the two versions mainly lies in the

scale and complexity of the network, and its trade-off between computing resources, speed and accuracy. MobileNetV3 Small is a lightweight version, which is suitable for mobile devices or edge devices with limited computing resources. It usually has fewer parameters and less computation, so as to achieve faster reasoning speed in the resource-limited environment, but it may sacrifice some classification accuracy. MobileNetV3 Large is a larger and more complicated version, usually with more parameters and calculation. It may show higher classification accuracy in some cases, but it needs more computing resources and inference time accordingly. We used SGD as the model optimizer, with momentum and weight decay set to be 0.5 and 0, respectively. The learning rate starts from 10^{-3} and decreases to 10^{-9} over 700 epochs, controlled by a cosine annealing schedule.

3 Experiments and analysis

3.1 Datasets

To validate the effectiveness of our proposed method, we conducted a series of experiments on two publicly available iris databases: CASIA-V4[31]-Interval and CASIA-V4-Thousand[31]. We partitioned each dataset into training and testing sets, ensuring that the classes in the training set and testing set are non-overlapping.

CASIA-V4-Interval is one of the most widely used publicly available iris datasets with high quality images, which contains 2 639 images from 395 eyes. Since the number of images per class (eye) is different, we selected the first 8 images from classes with 8 or more images for the training set, and the first 6 images from classes with 6 or 7 images for the training set. Therefore, the training set consists of 93 classes with a total of 774 images, while the testing set consists of 193 classes with a total of 1 158 images.

CASIA-V4-Thousand is the first publicly available iris dataset with one thousand subjects, containing 20 000 images from 2 000 eyes, with 10 images per class. We selected the right eyes of the first 600 subjects for our experiments. Compared to CASIA-V4-Interval, CASIA-V4-Thousand introduces more challenges to iris recognition algorithms due to the presence of glasses, specular reflections, and other interferences. After removing some classes that OSIRIS failed to work properly on, we kept 476 classes and selected the first 8 images from each class. We uniformly sampled 150 classes from the 476 classes as the training set and the remaining 326 classes as the testing set.

3.2 Performance evaluation

To comprehensively evaluate the performance of the proposed iris verification algorithm, we adopted accuracy, equal error rate (EER), FAR, and FRR as the evaluation metrics. In the testing phase, we set the ratio of positive and negative samples to 1∶3 to properly reflect the performance of the algorithm. Unless otherwise specified, the hyper parameters in the training phase were set as follows: Dataset: CASIA-V4-Interval, Training classes=40, Training epochs=700, Positive samples: Negative samples=1:1, DAPS: p_0=0.3, p_1=0, and Refresh epoch=1.

We evaluated the feature extraction ability of our model by adjusting the number of classes in the training set. Table 1 shows the results of our method on different datasets and models. It can be observed that our method achieved excellent performance even with a small number of training samples, and achieved the best EER of 0.16% and 0.84% on CASIA-V4-Interval and CASIA-V4-Thousand, respectively. Although MobileNetV3 Large achieved the best accuracy, MobileNetV3 Small performed better in more cases. The performance gap between the two models is not obvious, and the model choice also needs to consider the performance limitations of the actual deployment terminal. Considering the weak computing power of the edge side, MobileNetV3 small may be a better choice in most cases.

Table 1 Experimental results on different datasets, models, and number of training classes

Model	Dataset	Training classes	Testing classes	$P:N^*$	Accuracy/%	EER/%	FRR@FAR=0.1%/%	FRR@FAR=0.01%/%
MobileNetV3_large	CASIA-V4-Interval	10	193	1:3	99.19	0.81	1.79	37.55
MobileNetV3_large	CASIA-V4-Interval	20	193	1:3	99.54	0.47	0.85	4.44
MobileNetV3_large	CASIA-V4-Interval	40	193	1:3	**99.85**	0.16	0.19	0.55
MobileNetV3_large	CASIA-V4-Interval	93	193	1:3	99.78	0.19	0.23	0.42
MobileNetV3_small	CASIA-V4-Interval	93	193	1:3	99.83	**0.16**	**0.18**	**0.30**
MobileNetV3_large	CASIA-V4-Thousand	93	326	1:3	98.67	1.01	2.00	3.62
MobileNetV3_small	CASIA-V4-Thousand	93	326	1:3	98.20	1.25	2.82	5.57
MobileNetV3_large	CASIA-V4-Thousand	150	326	1:3	98.53	0.99	2.06	3.85
MobileNetV4_small	CASIA-V4-Thousand	150	326	1:3	98.82	0.84	1.56	2.84

As mentioned earlier, the iris enhancement method has a significant impact on the performance of the model. Table 2 shows the results under different enhancement methods. Compared to HE and no HE, CLAHE can bring huge performance improvements, especially in the two metrics of FRR@FAR=0.1% and FRR@FAR=0.01%.

Table 2 Performance comparison of different data enhancement methods

Enhancement method	Accuracy/%	EER/%	FRR@FAR=0.1%/%	FRR@FAR=0.01%/%
CLAHE	99.75	0.24	0.37	0.99
HE	98.85	1.10	4.33	79.43
no HE	96.56	3.12	73.52	98.50

Fig. 5 illustrates the impact of positive-to-negative sample ratio and refresh epoch on model performance. When the number of training classes is low, imbalanced datasets can have a positive effect on model performance. However, this effect may not be as effective when training with a larger number of classes. Moreover, it could be observed that a faster refresh rate for data combination leads to better performance in Table 3.

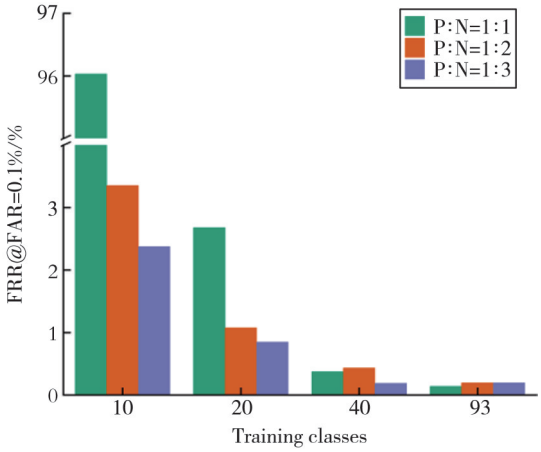

Fig. 5 Ablation experiments conducted on different positive-to-negative sample ratios

Table 3 Ablation experiments conducted on different refresh epochs

Training classes	Refresh epoch	Accuracy/%	EER/%	FRR@FAR=0.1%/%	FRR@FAR=0.01%/%
40	1	99.75	0.24	0.37	0.99
40	5	99.70	0.34	0.54	1.14
40	25	99.70	0.32	0.53	1.07
40	140	99.61	0.45	0.72	1.94
40	0	99.38	0.46	0.90	42.09

We further investigated the impact of different parameters of DAPS on the model performance. Fig. 6 illustrates the data augment probability curve of DAPS with different parameters during the training phase, and the results are presented in Table 4. In particular, when $p_0=p_1$, it means that the data augment probability remains unchanged during the training phase. We found that the most important thing for achieving good results is to lower the p_1 to 0 before the end of training, which means setting $p_1=0$. The second most important thing is to set an appropriate p_0.

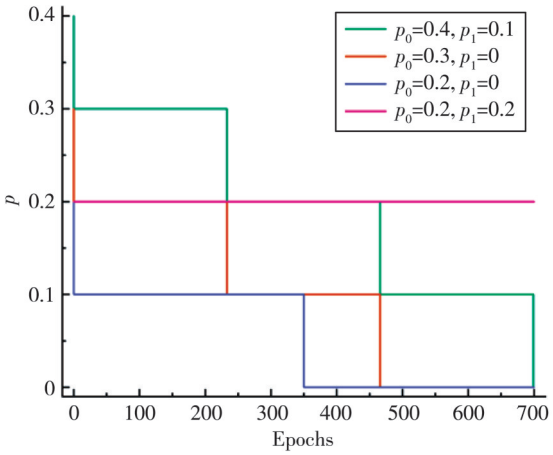

Fig. 6 Data augment probability curves for different parameters during training phase

Table 4 Ablation experiments of different DAPS parameters

No.	DAPS	Accuracy/%	EER/%	FRR@FAR=0.1%/%	FRR@FAR=0.01%/%
0	$p_0=0.5$, $p_1=0$	99.67	0.35	0.66	2.77
1	$p_0=0.4$, $p_1=0$	99.64	0.40	0.67	1.64
2	$p_0=0.4$, $p_1=0.1$	99.35	0.56	1.20	44.09
3	$p_0=0.3$, $p_1=0$	99.75	0.24	0.37	0.99
4	$p_0=0.3$, $p_1=0.1$	99.21	0.73	1.78	78.94
5	$p_0=0.2$, $p_1=0$	99.70	0.35	0.52	1.11
6	$p_0=0.3$, $p_1=0.3$	99.36	0.39	0.93	56.83
7	$p_0=0.2$, $p_1=0.2$	99.32	0.41	0.93	76.63
8	$p_0=0.1$, $p_1=0.1$	99.39	0.33	0.65	32.88
9	$p_0=0$, $p_1=0$	99.03	0.84	1.65	3.30

Although the experimental results demonstrate that our data augmentation method is effective in improving accuracy, high data augment probability can lead to an increase in FRR@FAR=0.01%. DAPS solves this problem by dynamically adjusting the data augment probability during the training phase.

As shown in Table 5, we conducted a comparative analysis against other algorithms for a comprehensive evaluation of the performance of our proposed method. Our proposed approach demonstrated superior performance on two distinct datasets in comparison to other algorithms. Notably, our approach achieved enhanced performance while utilizing datasets with smaller sizes as compared to the algorithms being compared.

Table 5 Comparison of the performances of proposed method with other algorithms

Studies	Year	Method	Evaluation protocol	EER/%	FRR@FAR=0.1%/%
Othman et al.[29]	2016	Iriscode	CASIA-V4-Thousand: 602 classes (Test: 602)	3.5	
Proença et al.[32]	2019	VGG-19 based CNN	CASIA-V4-Interval: 2 000 classes (Train: 1 000, Test: 1 000)	3	
Kien et al.[33]	2020	NAS	CASIA-V4-Thousand: 1 000 classes (Train: 700, Test: 300)	1.54	
Wei et al.[34]	2021	CUL	CASIA-V4-Thousand: 1 000 classes (Train: 900, Test: 100)	0.85	2.29
Liu et al.[28]	2021	Condensed 2-ch CNN	CASIA-V4-Interval: 233 classes (Train: 33, Test: 200)	0.76	1.24
Proposed Method	2023	DAPS	CASIA-V4-Interval: 233 classes (Train: 40, Test: 193)	0.16	0.19
			CASIA-V4-Thousand: 476 classes (Train: 150, Test: 326)	0.84	1.56

4 Conclusions

In this study, we proposed an iris verification method based on dynamic data augmentation and contrastive learning. The MobileNetV3 was employed as the backbone network, which was optimized with contrastive learning under 3-channel iris pair scenarios. To enhance the robustness of the model, four effective data augmentation strategies were employed to perform online enhancement of iris images, and contrast limited adaptive histogram equalization was utilized to extract more detailed iris features. Data diversity was enhanced in the training phase by increasing the number of negative examples and refreshing the combination of negative examples. Moreover, we further improved the accuracy of the model by using the proposed DAPS. DAPS can not only ensure the improvement of model robustness, but also reduce the interference of noise. The proposed method maintains high performance under low training sample size conditions, achieving competitive accuracies of 99.85% and 98.82% on CASIA-V4-Interval and CASIA-V4-Thousand, respectively.

Acknowledgement

This work was supported by National Natural Science Foundation of China (No. 62271291); and Key Program of the Natural Science Foundation of Shandong Province (No. ZR2020LZH009).

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] BOWYER K W, HOLLINGSWORTH K, FLYNN P J. Image understanding for iris biometrics: A survey. Computer Vision and Image Understanding, 2008, 110(2): 281-307.

[2] LIU N, ZHANG M, LI H, et al. DeepIris: Learning pairwise filter bank for heterogeneous iris verification. Pattern Recognition Letters, 2016, 82: 154-61.

[3] De MARSICO M, PETROSINO A, RICCIARDI S. Iris recognition through machine learning techniques: A survey. Pattern Recognition Letters, 2016, 82: 106-15.

[4] ALINIA LAT R, DANISHVAR S, HERAVI H, et al. Boosting iris recognition by margin-based loss functions. Algorithms, 2022, 15(4): 118. https://doi.org/10.3390/a15040118

[5] MARAM G A, LAMIAA A E. Convolutional neural network based feature extraction for iris recognition. International Journal of Computer Science & Information Technology, 2018, 10: 65–78.

[6] ZHAO T, LIU Y, HUO G, et al. A deep learning iris recognition method based on capsule network architecture. IEEE Access, 2019, 7: 49691-49701.

[7] KONG A W K, ZHANG D, KAMEL M S. An analysis of iriscode. IEEE Transactions on Image Processing, 2010, 19(2): 522-532.

[8] KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition//The 32nd International Conference on Machine Learning, July 6-11, 2015, Lille, France. New York: Curran Associates, Inc., 2015, 2: 1-8.

[9] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3//IEEE/CVF Conference on Computer Vision, October 27-November 1, 2019, Seoul, South Korea. Piscataway, N. J.: IEEE, 2019: 1. arXiv: 1905.02244

[10] DAUGMAN J. New methods in iris recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics. 2007, 37(5): 1167-1175.

[11] DAUGMAN J. How iris recognition works. IEEE Transactions on Circuits & Systems for Video Technology, 2004, 14(1): 21-30.

[12] DAUGMAN J. Information theory and the iriscode. IEEE Transactions on Information Forensics and Security, 2015,

11(2): 400-409.

[13] DAUGMAN J G. High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(11): 1148-1161.

[14] YAO P, LI J, YE X, et al. Iris recognition algorithm using modified log-gabor filters//The 18th International Conference on Pattern Recognition, August 20-24, 2006 Hong Kong, China. Piscataway, N. J.: IEEE, 2006, 4: 461-464.

[15] SARHAN A M. Iris recognition using discrete cosine transform and artificial neural networks. Journal of Computer Science, 2009, 5(5): 369.

[16] SHI J X, GU X F. The comparison of iris recognition using principal component analysis, independent component analysis and Gabor wavelets//The 3rd International Conference on Computer Science and Information Technology, April 23-25, 2010. Piscataway, N. J.: IEEE. 2010: 5563947.

[17] JASIM Y A, AL-ANI A A, AL-ANI L A. Iris recognition using principal component analysis//The 1st Annual International Conference on Information and Sciences, November 20-21, 2018, University of Fallujah, Iraq. Piscataway, N. J.: IEEE, 2018: 00028.

[18] RASOOL R A. Iris feature extraction and recognition based on gray level co-occurrence matrix (GLCM) technique. International Journal of Computer Applications, 2018, 181(25): 15-17.

[19] HAJARI K, GAWANDE U, GOLHAR Y. Neural network approach to iris recognition in noisy environment. Procedia Computer Science, 2016, 78: 675-682.

[20] RASTI P, DANESHMAND M, ANBARJAFARI G. Statistical approach based iris recognition using local binary pattern. DYNA-Ingeniería e Industria, 2017, 92(1): 1.

[21] LOWE D G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110.

[22] ALONSO-FERNANDEZ F, TOME-GONZALEZ P, RUIZ-ALBACETE V, et al. Iris recognition based on sift features//The 1st IEEE International Conference on Biometrics, Identity and Security, September 22-23, Tampa, USA. 2009. Piscataway, N. J.: IEEE, 2009: 1-8.

[23] SAHMOUD S A, ABUHAIBA I S. Efficient iris segmentation method in unconstrained environments. Pattern Recognition, 2013, 46(12): 3174-3185.

[24] REN X, PENG Z, ZENG Q, et al. An improved method for Daugman's iris localization algorithm. Computers in Biology and Medicine, 2008, 38(1): 111-115.

[25] NGUYEN K, FOOKES C, ROSS A, et al. Iris recognition with off-the-shelf CNN features: A deep learning perspective. IEEE Access, 2018, 6: 18848-18855.

[26] GANGWAR A, JOSHI A. DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition//IEEE International Conference on Image Processing, September 25-28, 2016, Phoenix, Arizona, USA. 2016: 2301-2305.

[27] NGUYEN K, PROENCA H, ALONSO-FERNANDEZ F. Deep learning for iris recognition: A Survey. 2022. arXiv: 221005866.

[28] LIU G, ZHOU W, TIAN L, et al. An efficient and accurate iris recognition algorithm based on a novel condensed 2-ch deep convolutional neural network. Sensors, 2021, 21(11): 3721.

[29] OTHMAN N, DORIZZI B, GARCIA-SALICETTI S. OSIRIS: An open source iris recognition software. Pattern Recognition Letters, 2016, 82: 124-131.

[30] SANPACHAI H, MALISUWAN S. A study of image enhancement for iris recognition. Journal of Industrial and Intelligent Information, 2015, 3(1): 61-64.

[31] CASIA-V4. [2023-08-29]. Available online: http://biometrics. idealtest. org/

[32] PROENçA H, NEVES J C. Segmentation-less and non-holistic deep-learning frameworks for iris recognition//IEEE/CVF Conference on Computer Vision and Pattern Recognition, October 27-November 1, 2019, Seoul, South Korea. Piscataway, N. J.: IEEE, 2019: 1-10.

[33] NGUYEN K, FOOKES C, SRIDHARAN S. Constrained design of deep iris networks. IEEE Transactions on Image Processing, 2020, 29: 7166-7175.

[34] WEI J, HE R, SUN Z. Contrastive uncertainty learning for iris recognition with insufficient labeled samples//IEEE International Joint Conference on Biometrics, August 4-7, 2021, Shenzhen, China. Piscataway, N. J.: IEEE, 2021: 1-8.

使用动态数据增强和对比学习进行虹膜验证

贺兰迪，纪德赞，董兴辰，苏明鑫，周卫东*

山东大学 微电子学院，山东 济南 250101

摘　要：虹膜验证因其独特性、稳定性和非侵入性而受到广泛关注。深度学习技术在虹膜验证领域取得了重要的进展，通过使用卷积神经网络（Convolutional neural network，CNN），可以自动提取和学习虹膜图像的特征，实现高精度的身份验证。然而，类内变异性和有限的数据规模等挑战可能会影响验证准确性。为了解决这些问题，我们提出了一种基于动态数据增强和对比学习的虹膜验证方法。设计了四种数据增强策略，用于在线虹膜增强和数据集扩展，通过使用数据增强概率调度器（Data augmentation probability scheduler，DAPS），进一步提高了虹膜验证的准确性。采用MobileNetV3作为骨干网络，并通过对比学习对其进行优化，用于处理3通道的虹膜对。提出的方法在两个基准虹膜数据库，CASIA-V4-Interval和CASIA-V4-Thousand上进行了评估，准确性分别达到了99.85%和98.82%。实验结果表明，在训练样本数量较少的情况下，该方法可获得具有竞争性的性能。

关键词：虹膜验证；对比学习；卷积神经网络；数据增强

引用格式：HE Landi，JI Dezan，DONG Xingchen，et al. Using dynamic data augmentation and contrastive learning for iris verification. Journal of Measurement Science and Instrumentation，2024，15(1)：54-63.