The Application of Improved Scale Invariant Feature Transformation Algorithm in Facial Recognition

Yingzi Cong
School of Continuing Education, Criminal Investigation Police University of China, Shenyang, 110854, China

***Abstract*—Currently, face recognition models suffer from insufficient accuracy, stability, and computational efficiency. To address this issue, an improved feature extraction algorithm on the ground of Haar wavelet features and scale invariant feature transformation algorithm is proposed. In addition, the study also combines this algorithm with deep belief networks to construct an improved facial recognition model. The effectiveness of the proposed improved feature extraction algorithm was verified, and it was found that the recognition accuracy of the algorithm was 94.2%, which is better than other comparative algorithms. In addition, the study also conducted empirical analysis on the improved facial recognition model and found that the recognition accuracy of the model was 0.92, and the feature matching time was 2.6 seconds, which was better than other comparative models in terms of performance. On the ground of the above results, the proposed facial recognition model has significantly improved recognition accuracy and efficiency compared to traditional models. It can provide theoretical reference for improving the universality of facial recognition applications in different fields.**

Keywords*—*Haar wavelet features; scale invariant feature transformation algorithm; deep belief network; facial recognition; performance improvement

I. INTRODUCTION

With the rapid development of social economy and the improvement of overall social living standards, more and more people are beginning to realize the importance of information security. At the same time, with the rapid iteration of computer hardware and software, more and more target face recognition technologies have emerged [1-2]. Facial recognition technology has begun to be widely used in people's daily lives. This includes but is not limited to security monitoring, facial payment, and social media login. Using facial biometrics as an important information component to identify the target object is beginning to become a very popular verification method [3]. Using facial features as a medium to identify target groups has also become a new type of identification method. This technology combines advanced scientific and technological means such as information science with related basic disciplines to achieve the identification of target groups. However, in real-life environments, the recognition accuracy of face recognition technology may be affected by undesirable factors such as obstructions, light changes, and various scale transformations. Due to the influence of hardware and other related factors, traditional face recognition technology is currently unable to meet the huge and changing needs of the urrent society [4]. The development of society has also prompted users to no longer be satisfied with the efficiency and accuracy of current face recognition. How to improve the accuracy of face recognition under non-ideal conditions has become an important topic studied by many scholars in the field of biometrics. Scale Invariant Feature Transform (SIFT) can generate feature descriptors that are insensitive to illumination and occlusion, and has properties such as affine, rotation, and scale invariance. These characteristics make SIFT useful in image recognition and classification and also Excellent performance in tasks [5]. Deep Belief Network (DBN), as a deep learning model, can automatically learn more stable feature representations from a large amount of data through a multi-level neural network structure, thereby improving the accuracy of face recognition, performance and robustness [6]. In view of this, the experiment innovatively proposes a feature extraction algorithm that integrates SIFT and DBN, and applies it to facial feature recognition to extract facial features more effectively and accurately, thereby improving the accuracy and matching of face recognition, speed, and look forward to providing more technical support for face recognition related work.

The article can be divided into five sections. Section II is the literature review, elaborates on the current development status and application fields of SIFT technology, DBN technology and face recognition; Section III is a research method, which combines SIFT technology and DBN technology to jointly deal with the problems of face recognition. Section IV is the result analysis, mainly to verify the performance of the built model. Section V is the conclusion and discussion, which is mainly a summary statement of the entire manuscript.

II. RELATED WORKS

With the advancement of social technology, deep learning algorithms have been widely applied in various fields. To address the issue of registration performance being easily affected by synthetic aperture radar during remote sensing image registration. Paul et al. proposed a remote sensing image registration algorithm that suffers from the SIFT structural tensor, and validated its effectiveness. It was found that this algorithm can strengthen the accuracy and precision of remote sensing image matching through feature classification and recognition [7]. To improve the analysis sensitivity and speed of volatile organic compounds in water, Perkins and Langford proposed a selective ion flow tube mass spectrometry method on the ground of SIFT. The effectiveness of this method was verified, and it was found that the non-fermentation traditional method has better linearity and relative standard error, and can be used as a rapid screening tool [8]. Lee MKI et al. proposed a tissue pathology morphology evaluation model using SIFI algorithm and convolutional neural network to address the issue of strong dependence on expert experience and subjective consciousness in tissue immunochemical pathology detection. The effectiveness of the model was verified, and it was found that its evaluation results were highly consistent with expert evaluation results, significantly improving the manpower and material resources of pathological diagnosis [9]. To improve the accuracy and effectiveness of the prediction of the remaining service life of bearings, this study used DBN to construct a bearing remaining service life prediction model, and verified the effectiveness of the model. It was found that the model could accurately forecast the remaining service life of bearings and has practical application value [10]. To improve the anti-interference ability of phase sensitive time-domain reflection against single disturbance events, Liu et al. proposed a DBN based interference event classification and recognition model, and verified its effectiveness. It was found that the model can effectively identify five types of single disturbance events with a recognition accuracy of 90.94%, which has practical application value [11].

In recent years, FR has not only been extensively utilized in many aspects, but research on FR has also become a hot research direction. To improve the robustness of FR methods, Zhang et al. proposed an adaptive margin FR model on the ground of convolutional neural networks. The effectiveness of the model was verified, and it was found that it can be widely applied in FR in various scenarios, with certain universality and robustness [12]. In response to the problem that FR models are susceptible to different conditions such as lighting, facial expressions, posture, and occlusion, a study proposes to use local binary patterns to verify the effectiveness of the FR model. The effectiveness of the improved model is verified, and it is found that the model has more robust performance under complex lighting conditions compared to traditional models [13]. Due to the insufficient recognition performance of traditional facial image recognition models in nighttime dark scenes, Sun et al. proposed a FR model on the ground of near-infrared technology and validated its effectiveness. It was found that this model has certain competitive performance compared to the latest methods [14]. For strengthening the anti-interference performance of FR models against occlusions, Ma et al. presented a FR method on the ground of second-order degree constraints and verified its effectiveness. It was found that compared with non-deep learning methods, this method possesses the best recognition rate [15].

To sum up, SIFT algorithm and DBN technology have been widely used in many fields. Both technologies have huge application prospects. At the same time, there are more and more related technologies related to face recognition. However, there are few studies on using SIFT technology and DBN technology to identify facial features at the same time. Although face recognition technology faces many challenges in complex real-life environments and unconstrained environments, it also has unprecedented development opportunities in today's rapid development and progress of information. In recent years, with the continuous progress of society, the public's demand for facial feature extraction and recognition technology has become more and more urgent. In view of this, the study introduces DBN technology and combines it with the SIFT algorithm to jointly deal with the problem of face recognition and improve the accuracy and recognition rate of face recognition.

III. CONSTRUCTION OF A FR MODEL WITH IMPROVED SCALE INVARIANT FEATURE CHANGE ALGORITHM

To improve the accuracy and reliability of FR, this study constructs the Haar SIFT algorithm on the ground of Haar, and combines this improved algorithm with DBN to construct an improved FR model.

A. *Improved SIFT Algorithm on The Ground of Haar Features*

The SIFT algorithm is an optimization algorithm on the ground of stochastic iterative function inversion. It can transform optimization problems into function inversion problems and use random sampling methods to approximate the inverse function of the objective function, thereby finding the optimal solution [16]. The core idea of the SIFT algorithm is for searching the solution space through random sampling and function inversion for finding the optimal solution [17]. Therefore, the SIFT algorithm has good application performance in continuous optimization problems and complex nonlinear problems. It has extensive applications in engineering design, combinatorial optimization, machine learning, and neural network training [18]. The calculation process of SIFT is shown in Fig. 1.

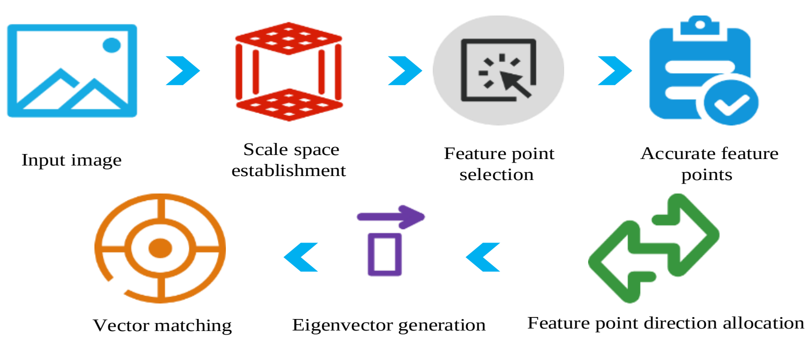


Fig. 1. Calculation flow of SIFT.

As shown in Fig. 1, the SIFT algorithm for extracting facial features mainly includes four parts: scale space extremum detection, key point localization, direction determination, and description of key points. In the step of scale space extremum detection, the SIFT algorithm constructs a Gaussian difference pyramid to detect the extreme points of the image at various scales, which may correspond to different features of the face. The calculation is showcased in Eq. (1).

$$L(x,y,\sigma)=G(X,Y,\sigma)*I(x,y) \tag{1}$$

In Eq. (1), $L(x,y,\sigma)$ serves as the scale space. $I(x,y)$ serves as the original image. σ serves as the scale of change. $G(X,Y,\sigma)$ represents the iterative Gaussian function with varying scales, and its calculation formula is showcased in Eq. (2).

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{\frac{(x-m)^2+(x-2/n)^2}{2\sigma^2}} \tag{2}$$

In Eq. (2), m and n represent the magnitude and number of changes in the scale of change. At this stage, the study conducted downsampling of the input image on the ground of scale changes, resulting in different images. Then the study constructs a pyramid like structure of these images from bottom to top and from large to small, to obtain a pyramid model of the images. Subsequently, the study utilized Gaussian difference functions for extreme value detection in the scale space, and the calculation formula is showcased in Eq. (3).

$$\begin{aligned}D(x,y,\sigma)&=\left(G(x,y,k\sigma)-G(x,y,\sigma)\right)*I(x,y,\sigma)\\&=L(x,y,k\sigma)-L(x,y,\sigma)\end{aligned} \tag{3}$$

In Eq. (3), $D(x,y,\sigma)$ represents the Gaussian difference function. k represents the multiple of the scale size between two adjacent scale spaces that generate a Gaussian difference space. Finally, on the ground of the detection of extreme values in the scale space, the SIFT algorithm accurately locates the detected extreme points to determine their accurate positions, which are considered as key points of the face. The formula for calculating the extreme points of the SIFT algorithm is shown in Eq. (4).

$$\begin{cases}D(\hat{X})=D+\frac{1}{2}\frac{\partial D^T}{\partial X}\hat{X}\\ \hat{X}=(x,y,\sigma)^T\end{cases} \tag{4}$$

In Eq. (4), $\hat{X}$ represents the offset relative to the interpolation center. ∂ represents the distance between adjacent points. In addition, on the ground of keypoint localization, the SIFT algorithm will determine a dominant direction for each keypoint, which can correspond to the directionality of facial features, which will help with subsequent feature description and matching. The calculation formula for its dominant amplitude and direction is shown in Eq. (5).

$$\begin{cases}m(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2}\\ \theta(x,y)=\tan^{-1}(((L(x,y+1)-L(x,y-1))/(L(x+1,y)-L(x-1,y))\end{cases} \tag{5}$$

In Eq. (5), θ represents the direction of change. After determining the direction, the SIFT algorithm will extract relevant feature descriptors on the ground of the image information around the key points. These descriptors can accurately describe the texture, shape, and other features around the key points, thus forming a representation of facial features. Through the above four parts, the SIFT algorithm can effectively extract facial features, which can play an important role in tasks such as FR and detection, improving the accuracy and stability of facial related tasks. Although the SIFT algorithm can handle complex nonlinear optimization problems and has global search capabilities, it can find the global optimal solution. And its robustness and computational difficulty are relatively low, which can be applied to facial image feature extraction. However, the SIFT algorithm requires high continuity and differentiability of the objective function and is not suitable for non-continuous or non-differentiable problems. The search process of the SIFT algorithm relies on random sampling and may fall into local optima. Therefore, this study aims to improve the algorithm on the ground of Haar wavelet features to enhance the search ability of the SIFT algorithm. Haar wavelet feature is a feature extraction method on the ground of wavelet transform, which was proposed by Alfred Haar in 1909. Haar wavelet features can be used for image processing and pattern recognition tasks, especially suitable for FR. The Haar wavelet features are showcased in Fig. 2.

Fig. 2 showcases that Haar wavelet features can divide an image into small blocks, calculate the differences between pixels in different rectangular regions, and combine the features of all small blocks into a feature vector. The calculation formula for Haar features is shown in Eq. (6).

$$feature=\sum_{i\in(i\ldots N)}\omega_i \operatorname{Re}ctSum(r_i) \tag{6}$$

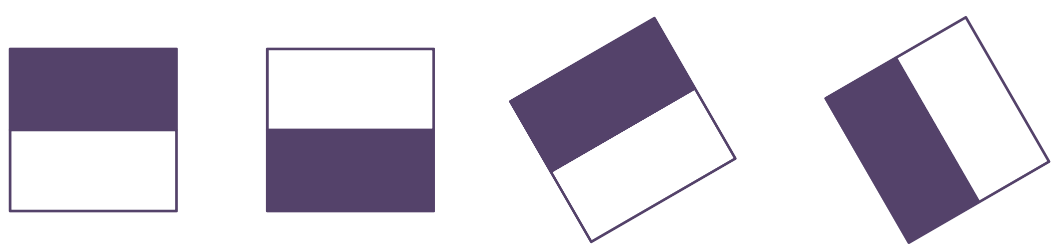


Fig. 2. Haar wavelet characteristics.

In Eq. (6), r_i represents a number in the Haar feature array. $\mathrm{Re}\,ctSum(r_i)$ represents the grayscale integration of the region image of r_i. ω_i represents the weight of r_i. N represents the number of rectangles in the region. This study imposes conditional limitations on Haar features to determine their effectiveness. Firstly, the areas with different weights have inverse proportional restrictions, and their calculation formula is shown in Eq. (7).

$$\begin{cases} \omega_1 Area(r_1) = -\omega_2 Area(r_2) \\ \omega_1 = -1; \quad \omega_2 = Area(r_1) / Area(r_2) \end{cases} \tag{7}$$

In Eq. (7), the weight must be a different sign within two regions and inversely proportional to the area of the region. In addition, for the convenience of calculating the integral image, the study limits the number of Haar features, and the calculation formula is shown in Eq. (8).

$$\begin{cases} r_1 \subset r_2 \\ r_2 \subset r_1 \end{cases} \tag{8}$$

Finally, to ensure the simplicity and computational speed of Haar features in image search operations, the study sets the number of rectangles that make up the feature area to 2. In addition, the study also utilized Sobel filters to accelerate the running speed of the SIFT algorithm. The calculation formula for Sobel filters is shown in Eq. (9).

$$G_\sigma = \sqrt{(G_{x,\sigma})^2 + (G_{y,\sigma})^2} \tag{9}$$

In Eq. (9), $G_{x,\sigma}$ serves as the horizontal derivative of the image after operation. $G_{y,\sigma}$ serves as the vertical derivative of the image after operation. The calculation formula for gradient direction is shown in Eq. (10).

$$R_{\sigma 1} = \arctan \frac{G_{y,\sigma}}{G_{x,\sigma}} \tag{10}$$

The calculation formula for image gradient amplitude is shown in Eq. (11).

$$G\sigma 1 = \sqrt{(G_{x,\sigma_1})^2 + (G_{y,\sigma_1})^2} \tag{11}$$

In Eq. (11), G_{x,σ_1} represents the horizontal derivative of the gradient amplitude image in the scale space. G_{y,σ_1} represents the vertical derivative of the gradient amplitude image in the scale space. The process of the improved SIFT algorithm on the ground of region segmentation is shown in Fig. 3.

As shown in Fig. 3, the study first preprocesses the input image, including image grayscale, denoising, and enhancement, to improve the accuracy of subsequent processing. Subsequently, the study searched for possible feature point positions on preprocessed images by constructing a set of image pyramids of different scales. The third step is the detection of extreme points. On the ground of the proposed results in the scale space, the image is segmented into different regions, and within each region, Haar wavelet features are used to locate feature points. For each feature point, Sobel is used to calculate the gradient direction of its surrounding area to determine its direction. Subsequently, the study compared the Haar wavelet feature vectors of feature points and used Euclidean distance to match similar feature points. This study screens and filters feature points on the ground of matching results to remove unreliable or redundant feature points. Finally, the research will output the matching results.

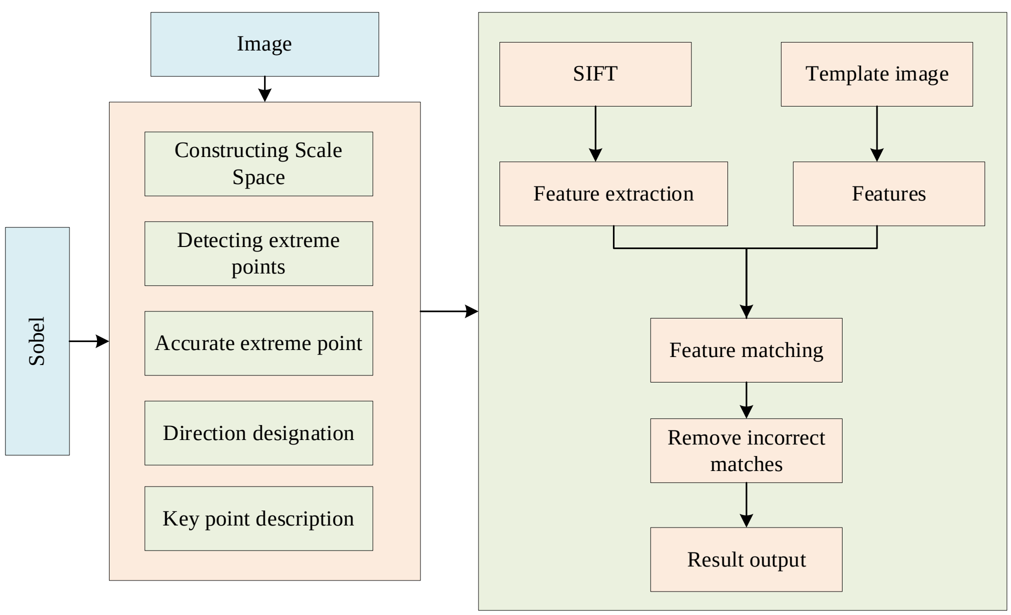


Fig. 3. Improved SIFT algorithm process on the ground of region segmentation.

B. FR Model on The Ground of Improved SIFI Algorithm

FR model is a technology that recognizes and verifies personal identity through facial images or videos. It is on the ground of the uniqueness and unforgeability of facial features, and is implemented using computer vision and pattern recognition techniques [19]. The application fields of FR models are very extensive, which can be applied in security fields such as access control systems, identity verification, and crime investigation. In addition, FR can also be applied to human-computer interaction and has a wide range of applications [20]. The traditional FR model is shown in Fig. 4.

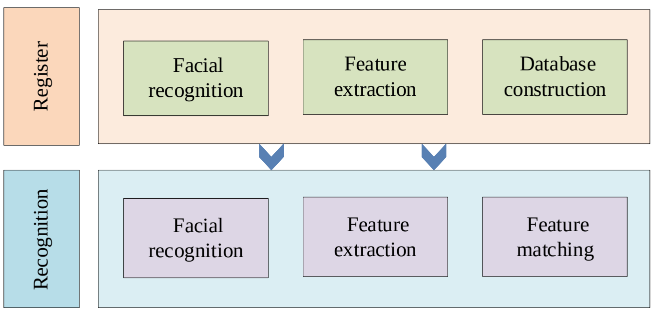


Fig. 4. The traditional FR model.

As shown in Fig. 4, a typical FR model consists of four main components: face acquisition recognition sensor, feature extraction, matcher, and database. The FR sensor is responsible for obtaining facial image samples from the real world and transmitting the captured images to the system for subsequent processing. In the step of feature extraction, the system preprocesses the obtained facial image to remove noise and interference, and extracts key features from the image. These features can include facial contours, eyes, nose, mouth, etc. The next step is the matcher step, where the system will compare and match the extracted facial features with existing facial features. Finally, for the database, the system will store the already entered facial features in the database for comparison and matching with subsequent facial features. However, traditional FR models still have certain shortcomings, including susceptibility to external environments and non-rigid deformations, making it difficult to achieve accurate matching. In addition, there are differences in facial appearance among different ages, genders, and races. Traditional models may not be able to adapt well to these diversity and variability, resulting in a decrease in accuracy and a significant increase in time costs. Therefore, to improve the accuracy, stability, and operational efficiency of FR performance, this study uses the Haar SIFT algorithm to extract features, and then uses DBN to classify and match features in images, achieving fast and effective FR. DBN is a deep learning model consisting of multiple stacked Restricted Boltzmann Machines (RBMs), which can be used for tasks such as feature learning, data dimensionality reduction, and model generation. To achieve effective classification of feature images by DBN, the study first trains DBN using the constructed feature library. The relevant training is showcased in Fig. 5.

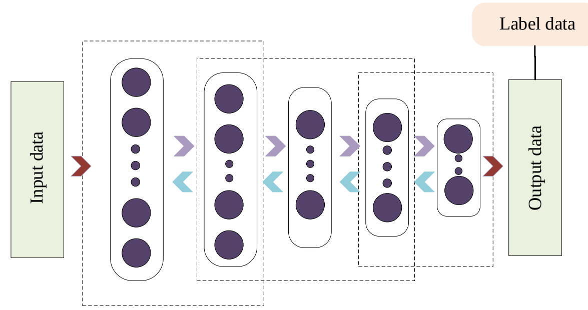


Fig. 5. Training process of DBN.

As showcased in Fig. 5, the training of DBN contains two stages: pre training and fine-tuning. In the pre training stage, the study first uses the first RBM as the underlying layer of the DBN. By performing unsupervised learning on the training data, this RBM learns low-level features of the data. Then, the hidden layer of the RBM is used as the visible layer of the next RBM to continue unsupervised learning. This is stacked layer by layer until all RBMs have been trained. During the training process of each RBM, the study also used contrast divergence to maximize the likelihood function. In the fine-tuning stage, research is conducted to fine tune the entire network using backpropagation algorithms to maximize the likelihood function of labeled data. This stage is supervised, allowing research to train using labeled data with the goal of adjusting network parameters to better adapt to specific tasks. The pre training process of DBN allows the network to learn abstract feature representations of data layer by layer, which helps to solve the gradient vanishing problem in deep neural networks. The fine-tuning process further optimizes the performance of the entire network to adapt to specific tasks. After completing the training of DBN, the study combines this algorithm with Haar SIFT algorithm to construct an improved FR network. The basic architecture of the improved FR network constructed through research is shown in Fig. 6.

As shown in Fig. 6, the improved FR model framework constructed in the study mainly includes two main parts: the construction of facial feature vector library and feature matching. In the construction module of the feature vector library, this study first performs facial detection on the images in the database to remove non facial regions as much as possible. This can reduce the useless feature vectors detected in other aspects such as background, and strengthen the computational efficiency. Subsequently, the study utilized an improved SIFT algorithm to extract and represent detected faces, and generated SIFT feature descriptors to form a feature vector library. Finally, the study groups the feature vectors in the feature vector library for training, and uses DBN to divide the feature vectors into different face categories, thereby establishing a training model for each face category. In the facial feature matching module, the research first performs facial detection on the face to be recognized, and uses an improved SIFT algorithm to extract feature descriptors. Subsequently, the study used DBN to calculate the measurement distance between the feature descriptors in the unknown face and each feature vector obtained in the database. DBN was used to compare the face to be recognized with the face in the feature vector library to find the most similar feature vector. Finally, research is conducted to determine which person the face to be recognized belongs to on the ground of the facial category to which the matched feature vectors belong. Through the above framework and workflow, the FR model can achieve classification and recognition of facial images, which can be extensively utilized in many aspects like secure access control, facial payment, and facial authentication.

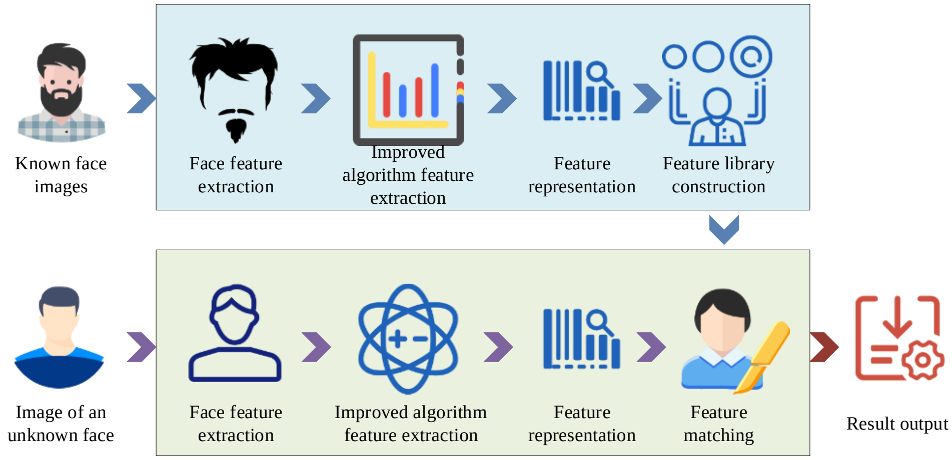


Fig. 6. Improved basic architecture of FR network.

IV. Result and Discussion

For testing the effectiveness of the proposed improved SIFT algorithm and the FR models on the ground of Haar SIFT and DBN, performance comparison experiments and empirical analysis were conducted.

A. Validation of The Harr SIFT Algorithm's Effectiveness

For testing the effectiveness of the proposed Harr SIFT algorithm for feature extraction, this study conducted feature extraction performance comparison tests using SIFT, fused SIFT and K-Means Scale Invariant Feature Transform (Means SIFT), as well as SIFT and Random Forester Scale Invariant Feature Transform (RF SIFT). In order for the experiment to proceed smoothly, ensure that the parameters of all equipment are consistent. The selection of equipment and parameters used in the experiment is as follows: the implementation platform is Fastone, the operating system is Windows 10, the operating environment is UNIX, the running computer memory is 64G, the central processor is i7-8700, the central processor frequency is 4.2Hz, and the graphics processor It is NPU, the processor graphics card is RTX-2070, the data storage platform is MySQL, and the data statistics software is SPSS 26.0. The comparison indicators are the accuracy, precision, extraction time, F1 value, and loss function value of feature extraction. The experimental environment used in the study is showcased in Table I.

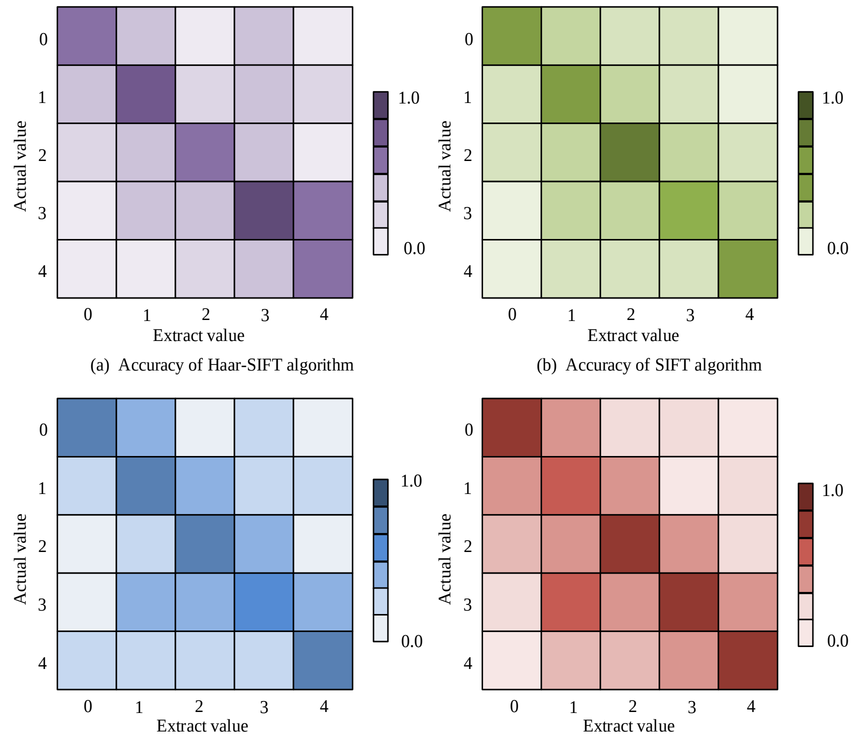
The study randomly selected facial images of five volunteers from the ORL facial database as experimental subjects and conducted feature extraction performance tests on them. The accuracy comparison results of various feature extraction algorithms are shown in Fig. 7.

Fig. 7(a), (b), (c), and (d) showcase the accuracy comparison results of Haar SIFT, SIFT, Means SIFT, and RF SIFT feature extraction algorithms, respectively. As shown in Fig. 7, the Haar SIFT feature extraction algorithm has the highest accuracy, at 94.2%, which is 3.6% higher than the SIFT algorithm, 8.3% higher than the Means SIFT algorithm, and 6.5% exceeding the RF SIFT algorithm. In terms of the above results, it can be concluded that the proposed Harr SIFT feature extraction algorithm has better accuracy in feature extraction and practical application value. The comparison results of the accuracy and feature extraction time of each feature extraction algorithm are shown in Fig. 8.

TABLE I. Table of the Experimental Environments

Device name	Specification parameters
Language form	C++ language
Java runtime environment Java	AMD Radeon56400G@4.8GHz
Internal storagememory	32GB
Operating system	Ubuntu 18.0
Experimental platform	ClassBench
Data set	The ORL human face database
Learning rate	0.001
Iterations	600-2000

As shown in Fig. 8, the feature extraction accuracy of each exploration extracted feature extraction algorithm is 92.68%, which is better than other comparative algorithms and has better stability. In addition, the Haar SIFT algorithm possesses more excellent feature extraction performance than other algorithms, with a feature extraction time of 2.6 minutes, including the establishment and extraction of feature libraries. On the ground of the above results, it demonstrates that the Haar SIFT algorithm has the best feature extraction accuracy and efficiency performance, and has practical application value. The F1 values and loss functions of each algorithm are shown in Fig. 9.



(a) Accuracy of Haar-SIFT algorithm (b) Accuracy of SIFT algorithm

(c) Accuracy of the Means-SIFT algorithm (d) Accuracy of RF-SIFT algorithm

Fig. 7. Comparison of the accuracy results of the feature extraction algorithm.

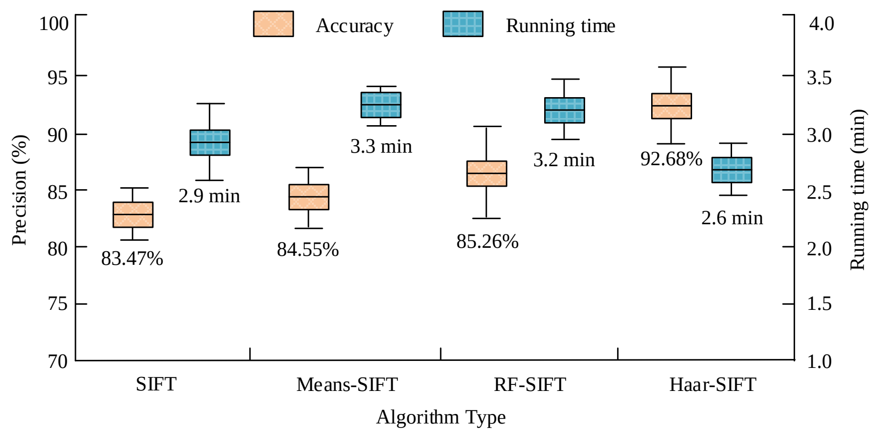
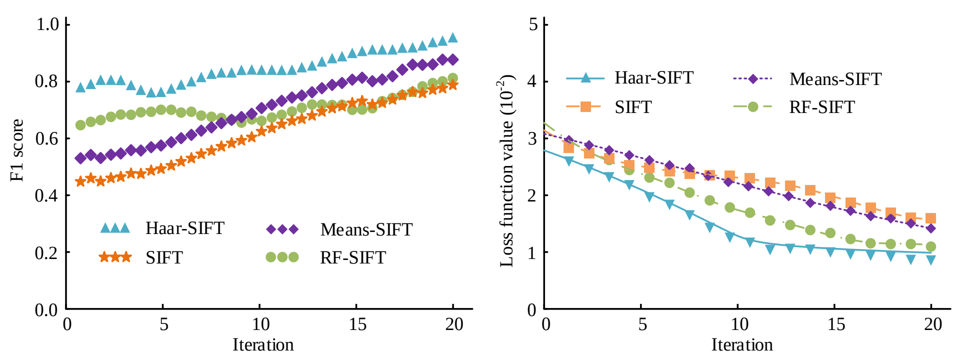


Fig. 8. Accuracy of each feature extraction algorithm and the comparison results of the feature extraction time.



(a) F1 score of each algorithm (b) Loss function value of each algorithm

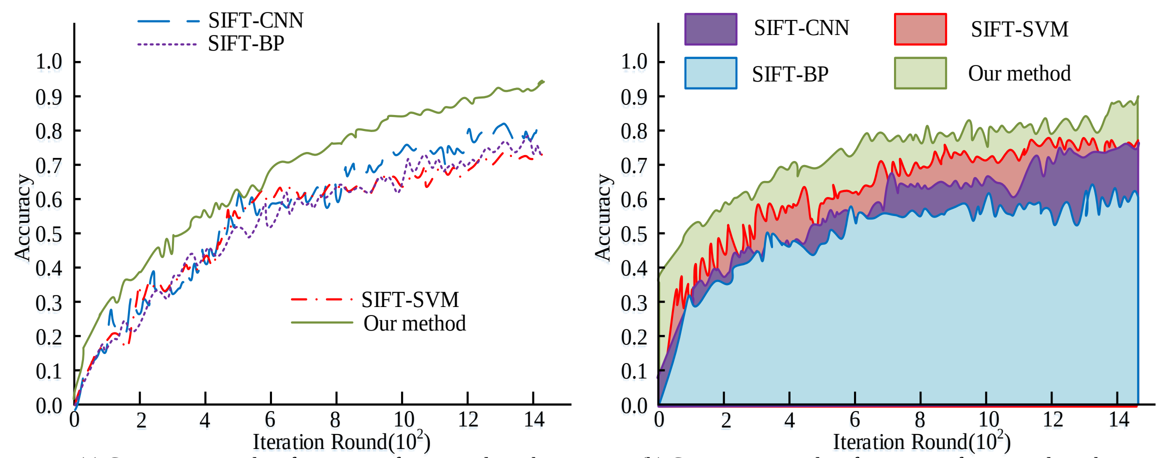
Fig. 9. Comparison results of F1 score and loss function value.

Fig. 9(a) shows the F1 value comparison results of the feature extraction algorithm, as shown in Fig. 9(a). As the number of iterations increases, the F1 values also increase. The F1 value of Haar SIFT proposed in the study is the highest, at 0.89. Fig. 9(b) shows the comparison results of the loss function values of the feature extraction algorithm, as shown in Fig. 9(b). The proposed Haar SIFT algorithm has a lower loss function value curve than other algorithms, and its convergence speed is also faster. On the ground of the above results, it can be concluded that the Haar SIFT algorithm has better convergence performance and practical application value.

B. Empirical Analysis of A FR Model Integrating Haar SIFT and DBN Algorithms

For testing the effectiveness of the FR model on the ground of Haar SIFT algorithm and DBN proposed in the study, the ORL face database and FERET face database were used to verify its effectiveness. The comparative models are FR models on the ground of SIFT and Scale Invariant Feature Transform Convolutional Neural Network (SIFT-CNN), SIFT and Scale Invariant Feature Transform Back Propagation (SIFT-BP), and SIFT and Support Vector Machines (SIFT-SVM). The comparison indicators are accuracy, recognition time, and recall rate. The relevant outcomes of recognition accuracy of many facial comparison models in ORL and FERET facial databases are shown in Fig. 10.

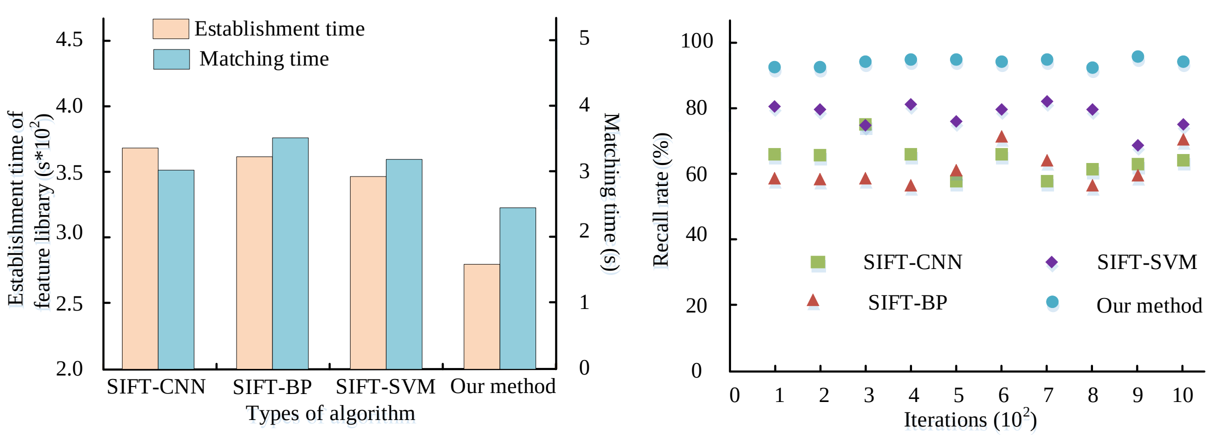
Fig. 10(a) shows the accuracy comparison results of various face comparison models on the ORL face database. As shown in Fig. 10(a), the recognition accuracy of each FR model grows with the number of iterations. Among them, the accuracy of the proposed FR model can reach 0.92, which is higher than the accuracy of other models. Fig. 10(b) shows the accuracy comparison results of various face comparison models on the FERET face database. As shown in Fig. 10(b), the recognition accuracy of each FR model increases with the number of iterations. Among them, the accuracy of the proposed FR model can reach 0.9, which is higher than the accuracy of other models. In terms of the above results, it can be concluded that the FR model proposed in the study has better recognition performance than other models in different datasets and is more stable, which can be applied to practical FR. The comparison results of running time and recall of each FR model on ORL and FERET face databases are shown in Fig. 11.



(a) Comparison results of accuracy of various algorithms in ORL face database

(b) Comparison results of accuracy of various algorithms in FERET face database

Fig. 10. Recognition accuracy of each face comparison model in the ORL, namely the FERET face database.



(a) Running time of each comparison model

(b) Precision of each comparison model

Fig. 11. Comparison results of the running time and recall rate of each FR model.

Fig. 11(a) shows the comparison results of the running time of various FR methods. As showcased in Fig. 11(a), the running time of the FR model includes the time for establishing the feature library and the time for feature matching. The proposed FR model has a shorter feature library establishment time compared to other models, which is 254.3s. Meanwhile, the feature matching time of the FR model proposed in the study is also shorter than other models, which is 2.6 seconds, and its operating efficiency is better. Fig. 11(b) shows the comparison results of recall rates for various FR methods. As shown in Fig. 11(b), the proposed FR model has a better recall rate of 98.2% compared to other models, and its stability is also better. On the ground of the above results, the proposed FR model has better operational efficiency and recall performance compared to other comparative models. In addition, for further validating the effectiveness of the proposed FR model, the study evaluated the satisfaction of each FR model through expert ratings. The expert satisfaction rating results of each model is shown in Fig. 12.

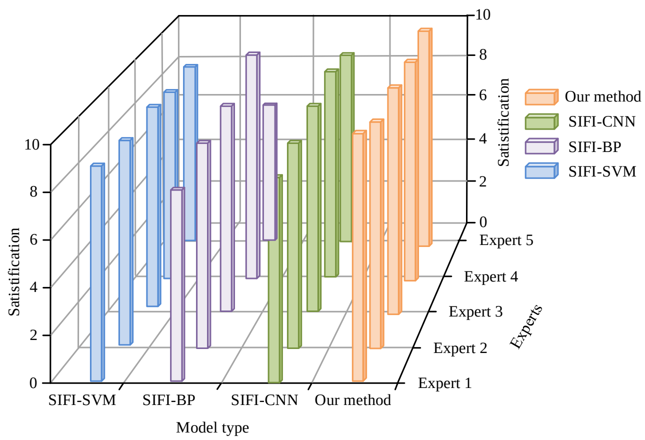


Fig. 12. Results of expert satisfaction scores.

As shown in Fig. 12, the average score of the proposed facial model in the study is 8.7 points, the average satisfaction of the SIFT-SVM based facial model is 7.6 points, the average satisfaction of the SIFT-BP based facial model is 7.5 points, and the average satisfaction of the SIFT-CNN based facial model is 7.7 points. In summary, the expert rating for the motion interaction control model on the ground of Haar SIFT and DBN proposed in the study is the highest, indicating that this model has better practical application value compared to other models.

C. Discussion

Due to the limitations of face image sampling in practical applications and the differences in the surrounding environment, image recognition will be affected by non-rigid changes in illumination and expression, as well as errors caused by non-primary data redundancy. Many algorithms still have problems with these problems. There are some shortcomings. The experiment proposes to recognize face images based on the SIFT sparse deep belief network model, and sparsely represent the features extracted by the SIFT algorithm, making the extracted features more effective and improving the accuracy of face recognition. At the same time, the deep belief network is combined with the training classification recognition to avoid redundant information and reduce the time of training the network. The above experimental results show that the SIFT sparse deep belief network model has a better recognition effect on facial expression changes. This recognition model effectively improves the face recognition effect and matching rate.

V. Conclusion

To further improve the accuracy, stability, and operational efficiency of FR models, a Haar SIFT feature extraction algorithm on the ground of Haar features and SIFT algorithm is proposed. This algorithm is combined with DBN to construct a FR model on the ground of Haar SIFT algorithm and DNB. The effectiveness of the Haar SIFT algorithm proposed in the study was verified, and it was found that the recognition accuracy of the algorithm was 94.2%, the feature extraction accuracy was 92.68%, and the feature extraction time was 2.6 minutes, which was better than other comparative algorithms. In addition, the study also validated the effectiveness of the FR model that integrates Haar SIFT and DBN, and found that the accuracy of the model on the ORL face database can reach 0.92, the accuracy on the FERET face database can reach 0.9, and the recall rate is 98.2%, which is better than other comparative models. In addition, the study also found that the feature library establishment time of the FR model integrating Haar SIFT and DBN is 254.3s, and the feature matching time is 2.6s, which is shorter than other models. In terms of the above results, the FR model proposed in the study has higher recognition accuracy and better recognition efficiency compared to traditional models, and its accuracy performance is consistent when facing different databases. However, research also has certain limitations. The SIFT algorithm typically requires a significant amount of computing resources and time to extract and match features. Therefore, it has certain limitations on devices with limited resources. Therefore, future research directions will improve on the basis of Haar SIFT to reduce computational complexity, improve real-time performance and availability of algorithms.

References

[1] Arzykulov S, Nauryzbayev G, Tsiftsis T A. Performance Analysis of Underlay Cognitive Radio Non-Orthogonal Multiple Access Networks. IEEE Transactions on Vehicular Technology, 2019, 68(9):9318-9322. DOI: 10.1109/TVT.2019.2930553.

[2] Whiteley R, Napier C, Dyk N V. Clinicians use courses and conversations to change practice, not journal articles: is it time for journals to peer-review courses to stay relevant? British Journal of Sports Medicine, 2020, 55(12):651-652. DOI: 10.1136/bjsports-2020-102736.

[3] Paul S, Udaysankar D, Naidu Y. An efficient SIFT-based matching algorithm for optical remote sensing images.Remote sensing letters, 2022, 13(12):1069-1079. DOI: 10.1080/2150704X.2022.2121186.

[4] Singh K, Neelima A, Tuithung T. Robust perceptual image hashing using SIFT and SVD. Current Science, 2019, 117(8):1340-1344. DOI: www.jstor.org/stable/27138450.

[5] Langford V S. Real-Time Monitoring of Volatile Organic Compounds in Ambient Air Using Direct-Injection Mass Spectrometry. LC GC North America, 2022, 40(4):174-179. DOI: 10.56530/lcgc.na.nf7066t5.

[6] Yan X, Shi Z, Li P. IDCF: information distribution composite feature for multi-modal image registration. International journal of remote sensing, 2023, 44(5):1939-1975. DOI: 10.1080/01431161.2023.2193300.

[7] Paul D S, Divya S V, Pati U C. Structure Tensor Based SIFT Algorithm for SAR Image Registration. IET Image Processing, 2019, 14(11):929-938. DOI: 10.1049/iet-ipr.201 9.0568 www.ietdl.org.

[8] Perkins M J, Langford V S. Standard Validation Protocol for Selected Ion Flow Tube Mass Spectrometry Methods Applied to Direct Headspace Analysis of Aqueous Volatile Organic Compounds. Analytical Chemistry, 2021, 93(24):8386-8392. DOI: 10.1021/acs.analchem.1c01310.

[9] Lee M K I, Rabindranath M, Faust K. Compound computer vision workflow for efficient and automated immunohistochemical analysis of whole slide images. Journal of clinical pathology, 2022, 76(7):480-485. DOI: 10.1136/jclinpath-2021-208020.

[10] Hu C H, Pei H, Si X S. A Prognostic Model Based on DBN and Diffusion Process for Degrading Bearing. IEEE Transactions on Industrial Electronics, 2019, 67(10):8767-8777. DOI: 10.1109/TIE.2019.2947839.

[11] Liu M, Wang X, Liang S. Single and composite disturbance event recognition based on the DBN-GRU network in 9-OTDR. Applied optics, 2023, 62(1):133-141. DOI: 10.1364/AO.477642.

[12] Zhang Z, Gong X, Chen J. Face recognition based on adaptive margin and diversity regularization constraints. IET Image Processing, 2021, 15(5):1105-1114. DOI: 10.1049/ipr2.12089.

[13] Chen T, Gao T, Li S. A novel face recognition method based on fusion of LBP and HOG. IET Image Processing, 2021, 15(14):3559-3572. DOI: 10.1049/ipr2.12192.

[14] Sun R, Shan X, Zhang H. Data gap decomposed by auxiliary modality for NIR-VIS heterogeneous face recognition. IET image processing, 2022, 16(1):261-272. DOI: 10.1049/ipr2.12350.

[15] Ma X, Ma Q, Ma Q. Robust face recognition for occluded real-world images using constrained probabilistic sparse network. IET image processing, 2022, 16(5):1359-1375. DOI: 10.1049/ipr2.12414.

[16] Chen Z, Chen J, Ding G. A lightweight CNN-based algorithm and implementation on embedded system for real-time face recognition. Multimedia systems, 2023, 29(1):129-138. DOI: 10.1007/s00530-022-00973-z.

[17] Wei Y, Weng Z. Research on TE process fault diagnosis method based on DBN and dropout. The Canadian Journal of Chemical Engineering, 2020, 98(6):1293-1306. DOI: 10.1002/cjce.23750.

[18] Su Z, Yang J, Li P, Jing J. A precise method of color space conversion in the digital printing process based on PSO-DBN. Textile Research Journal, 2022, 92(10):1673-1681. DOI: 10.1177/004051752110672.

[19] Ren B, Zhang M, Xu S. DBN-Catalyzed Regioselective Acylation of Carbohydrates and Diols in Ethyl Acetate. European Journal of Organic Chemistry, 2019, 29:4757-4762. DOI: 10.1002/ejoc.201900776.

[20] Nsugbe E. Toward a Self-Supervised Architecture for Semen Quality Prediction Using Environmental and Lifestyle Factor. Artificial Intelligence and Applications. 2023, 1(1): 35-42. DOI: 10.47852/bonviewAIA2202303.