IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 26, 2022, accepted April 9, 2022, date of publication April 13, 2022, date of current version April 20, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3167143

Face Recognition Method Based on Siamese Networks Under Non-Restricted Conditions

CUNLI SONG AND SHOUYONG JI
College of Software, Dalian Jiaotong University, Dalian, Liaoning 116052, China
Corresponding author: Cunli Song (scunli@163.com)

This work was supported in part by the Scientific Research Funding Projects of Liaoning Provincial Department of Education, China, under Grant LJKZ0489 and Grant LJKZ0486; and in part by the Liaoning Provincial Doctoral Research Start-Up Foundation, China, under Grant 2019-BS-042.

ABSTRACT Face recognition is an important part of computer vision, and has a vital role to play in public safety. Because of the problems of low accuracy and low efficiency in the case of uneven image quality and occlusion due to sudden changes in light in face recognition, we design a Siamese Neural Network based on Local Binary Pattern (also called LBP) and Frequency Feature Perception. The network is based on the Siamese neural network and adopts the Uniform LBP algorithm and Frequency Feature perception to achieve face recognition under non-restricted conditions. the LBP algorithm can eliminate the effect of lighting on the image and provide vector-level input to the network model; the frequency feature perception divides the image features into low-frequency features and high-frequency features, and compresses the low-frequency features in the Siamese neural network to increase the network's recognition efficiency while exchanging information with high-frequency features to retain their feature data while eliminating target noise data. This maintains the recognition rate of the network and improves the computational speed of the network. Simulation experiments are conducted on the standard face datasets CASIA-WebFace, Yale-B, and LFW, and compared with other network models. The experimental results show that the proposed SN-LF network structure in this paper can improve the recognition accuracy of the algorithm and obtain a better recognition accuracy.

INDEX TERMS Artificial intelligence, face recognition, Siamese neural network, LBP algorithm, frequency feature perception.

I. INDRODUCTION

With the development of artificial intelligence, target recognition technology has also been developed rapidly, especially in pedestrian recognition and vehicle recognition, as security awareness has increased [1]–[2], prompting the rising demand for public and personal information security. And in public safety through pedestrian recognition, some accidents can be better tracked and ranked. Non-restrictive face recognition means that face data are mostly collected in real environmental conditions, integrating various influencing factors such as facial pose, illumination, expression, and background, and recognizing these face data. In practical applications face recognition is often limited by the influence of some series of noisy data such as illumination and facial occlusion, which leads to the occurrence of poor recognition accuracy, wrong detection, and missed detection by face recognition algorithms. Non-restrictive face recognition face detection and recognition are mainly used in public security real-time monitoring areas and are of great significance [3] for applications such as capturing suspects. The main problems of non-restrictive face recognition at the present stage are: (1). In an environment without restrictive conditions, it is easy to generate more noise interference, which leads to a decrease in the accuracy of face recognition; (2). Noisy data can cause the algorithm to consume a lot of time in the process of noise removal, which affects the efficiency of face recognition; (3). In non-restrictive face recognition, due to the collection in the real environment, the face may be partially or completely occluded, resulting in missed detection and false detection. In this paper, we design a Siamese neural network model based on Local Binary Pattern

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

and Frequency Feature Perception model to solve the above problem

Summarizing the classical face recognition algorithms, we can find that the principal component analysis (PCA) [4] method reduces the computational complexity by transforming the matrix, but there are external interference factors such as occlusion, which will destroy the real spatial structure when there is interference, resulting in the inability to obtain the real subspace structure of the data. In the case of more noise interference, Tian *et al.* [5] innovatively proposed to add sparse representation to the image deblurring algorithm in the preprocessing stage, which is used to solve the current situation that the traditional algorithm has a poor denoising effect on blurred photos under Non-restrictive face recognition conditions, resulting in the decrease of recognition accuracy. Moreover, the traditional denoising algorithm requires a large number of relevant reference images to provide templates, which do not apply in the case of face recognition under non-restrictive face recognition conditions. Tian Lei *et al*. use the a priori knowledge of non-local self-similarity of images to first use the sparse representation to perform the deblurring coding operation on the images, and then use the local image blocks obtained from the sparse representation as to the input of the data dictionary by using the unsupervised clustering algorithm to The image data dictionary is finally constructed by using principal component analysis algorithm to generalize the similarity of image blocks in the image. The traditional classical algorithm will generate a large amount of redundant data while processing noisy data, too much redundant data will interfere with the recognition effect of the algorithm, and too much redundant data will increase the computational cost. With the proposal of compressed sensing [6] (CS, Compressed Sensing), the theory proposes that the sampling rate of the signal depends on the sparsity and irrelevance of the signal rather than the width of the signal. Compressed sensing uses compression while sampling and uses signal reconstruction at the output to decompress the data, which avoids generating a large amount of redundant data at the sampling end and reduces the storage cost and computational [7] cost.

In real situations, the face of the target person may be included in the unrestricted case, in which case it may cause the face recognition algorithm to miss the detection process or even lead to the wrong detection. Zhang [8] proposed an intermediate mode for mutual conversion between visible light and near-infrared based on the joint constraints of the image layer and the feature layer to improve the recognition rate, using a generative adversarial network to constrain the features of visible and infrared light images separately. Liu *et al.* [9] proposed to use adversarial networks to increase the number of samples and design a joint function to optimize the generative network, and then add the feature information of pedestrian key points to reduce the influence brought by the background as a way to improve the model efficiency. Zhao *et al.* [10] propose the key features of multi-pose faces based on vector machines for the low recognition efficiency. Use filters to extract the key points of facial features, obtain the 3D information of facial features, normalize the grayscale of facial features, and realize face recognition under non-constrained conditions by calculating the distance between their facial features.

Deep learning is developed and innovated based on neural networks. With the development of deep networks, a series of deep learning network models such as convolutional neural networks, Siamese neural networks, and deep confidence networks have been proposed. The principle of a deep confidence network [11] is to use a greedy algorithm to train multi-layer Restricted Boltzmann machines (RBM) layer by layer, and the output of the previous layer is used as the input of the next layer to train the image for unsupervised learning, which improves the speed and accuracy of model training, and solves the problem of local optimal solution very well. Since then, deep neural network development has entered a more rapid development stage. Face recognition also ushered in rapid development. Liang [12] and other scholars proposed to improve the deep confidence network to process the image, using the texture features obtained by preprocessing the image with the LBP algorithm and then using them as the input of the deep confidence network to assist the deep confidence network for unsupervised training to improve the efficiency of image recognition, but the method is not efficient in recognizing those targets with excessive noise data interference and is prone to the problem of missed detection. And it is easy to produce the situation of missing detection.

Siamese neural network is a simple and efficient network model framework, Siamese network was first proposed to be used to compare the signature authentication of U.S. checks, but the development of the Siamese neural network was stalled due to hardware limitations until the 2010 year, Hinton [13] used Siamese neural network on the field of face recognition by combining convolutional neural network with the Siamese network to form Siamese neural network. In the 2015 year, Zagoruyko and Komodakis [14] Using Siamese neural networks to calculate image similarity, Siamese neural networks have been rapidly developing in target recognition and tracking since then with the advantages of simple model structure, a small number of parameter settings, and coexistence of multiple network models. Xu *et al.* [15] proposed to introduce the Inception model into the Siamese network and add the cyclic learning rate optimization strategy to accelerate the training speed. Wu [16] combined Siamese Neural Network and Convolutional Neural Network, using the method of local response value normalization to make the mutation of the eigenvalue with strong feedback greater than the eigenvalue with small feedback so that the characteristics of the target are more obvious and the target is accurately identified. Shen *et al*. [17] used Siamese networks as classifiers for image recognition and proposed adding spatial transformation networks to Siamese neural networks to make adaptive changes to images to improve accuracy to solve the disturbances such as deformation and rotation that occur in

the target in images. However, to make the Siamese network better for classification, the required benchmark samples need to be extracted manually, which consumes more human resources.

There are also some studies to improve the performance of deep neural network algorithms and models by using regularization to enhance their generalization ability. Zheng [32] *et al.* use the idea of regularization in deep learning from the perspective of penalty, in the pre-training phase, using invisible regularization to optimize feature boundaries by penalizing samples, in the training phase, using the density difference between samples to detect outliers, and improving the small batch SGD by data intervention The algorithm performance is improved by data intervention to improve the regularization effect of small batch SGDs and thus improve the generalization ability. On the other hand, they improve the generalization ability of the network structure by data augmentation in the training and testing phases [33]. Deep learning generalization has also been applied to the secondary data augmentation method based on spectral interference to automatic modulation classification [34], where the radio signal is enhanced by data augmentation with the original signal, and then another data augmentation is performed in the training and testing phases to improve the generalization ability of the model and improve the performance of the algorithm.

To solve the problem of poor detection accuracy that occurs in pedestrian recognition under Non-restrictive face recognition conditions, this paper proposes SN-LF for Non-restrictive face recognition of pedestrian images. In this paper, the texture features of the target are obtained by image pre-processing using the LBP algorithm, and then the features are classified into high-frequency features and low-frequency features using frequency feature perception. Training is performed using Octiva Convolution. The LBP algorithm is also used to extract the texture features of the target to ensure the texture and structure of the target, which can accurately recognize the image, and the frequency features are used to enhance the receptive domain of the pixel space, thus improving the recognition performance.

The main contributions of this article are summarized below.

- To solve the problem of low detection accuracy in pedestrian recognition under non-limiting conditions, an output structure of vector-level feature data is proposed to provide input for subsequent vector-level feature data while removing noisy data in the pre-processing stage, eliminating most of the noisy data caused by illumination.
- The feature constraint period divides the feature data into high-frequency and low-frequency data, increases the perceptual field of high-frequency data, decreases the acceptance domain of low-frequency features, and reduces the acceptance of noisy data.
- During feature extraction, high-frequency features and low-frequency features are allowed to update their information while exchanging information with each other to ensure that the information of the image itself is not missing after data fusion.

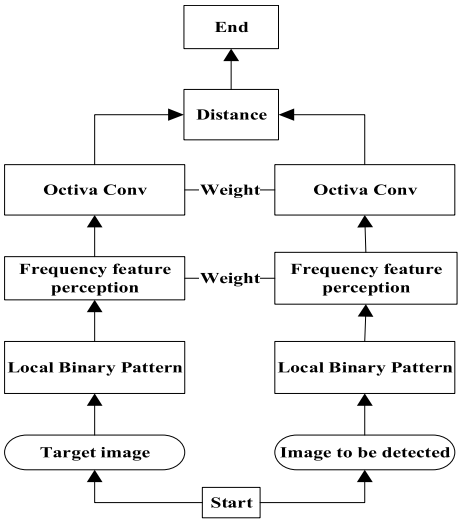

FIGURE 1. SN-LF algorithm flow.

II. SIAMESE NETWORK-BASED ALGORITHM FOR NON-RESTRICTED PEDESTRIAN FACE RECOGNITION

The algorithm flow in this paper is shown in Fig. 1. The two branches in the Siamese neural network are divided into the search branch and the model branch, and the LBP algorithm is added to the search branch of the Siamese network to preprocess the image to be detected, then the frequency features are used to convolve the image operation to make the spatial acceptance domain of the image target features larger, and finally, the pixel-level face features are learned by the Octiva convolution layer, while the model branch of the Siamese network extracts the target image from the target library, extract the target image, do the same operation on the target image, and finally use the Euclidean distance to determine whether the target in the two images is the same. The algorithm flow is as follows.

1) Pre-processing of the image with noise reduction by reducing the image dimension to 32*32 and histogram equalization using the LBP algorithm.
2) Training of the feature data layer by layer by using the training samples as input to the Siamese neural network.
3) Using high and low-frequency convolution for the input feature images in the Siamese neural network to obtain a frequency convolution image, the width and height of the frequency convolution image are 1/2 of the width and height of the original feature image, making the perceptual field of high frequency features in the feature image larger and the spatial acceptance rate of low frequency features lower, to achieve the recognition rate without reducing the recognition speed while improving the recognition rate.

4) The convolutional layers are divided into4 groups, corresponding to the width and height of the feature image, so that the image feature data are exchanged for data and frequency, which are width in high frequency - width in low frequency, width in high frequency - high in low frequency, high in high frequency - width in low frequency, and high in high frequency - high in low frequency. Making it possible to exchange information between high and low frequencies so that the data of the feature images are not lost.
5) When the Siamese neural network is trained, the two outputs in the test sample are compared to determine the Euclidean distance of the two image feature vectors, and if the Euclidean distance of the two image feature vectors is within the threshold value, they are the same target, and if it is greater than the threshold value, they are not the same target.

A. IMAGE PRE-PROCESSING BASED ON LBP ALGORITHM

The LBP algorithm uses the central pixel point as the threshold value, and the binary coding of its surrounding pixels is unchanged and insensitive to light changes, so it has gray invariance, and the LBP algorithm rotates the domain point to get multiple LBP values and selects the smallest value as the feature value, so LBP has the feature of rotation invariance, and LBP extracts features is a method of extracting features without parameters, so the LBP algorithm does not need to be calculated by adjusting parameters, which reduces the difficulty and complexity of the algorithm in the operation process, and at the same time can extract the main features of the image completely.

In this paper, we use the characteristics of gray invariance and rotation invariance in the LBP algorithm for image preprocessing, so that the domain pixel points are compared with the central pixel points, and then the encoding of the central pixel points is obtained, finally, according to the encoding of the pixel points, it is then converted to decimal encoding, and the algorithm formula is as follows.

$$T = t\left(s\left(g_0 - g_1\right), s\left(g_1 - g_c\right), \ldots, s\left(g_8 - g_c\right)\right) \quad (1)$$

$$s\left(x\right) = \begin{cases} 1\ x \geq 0 \\ 0\ x \leq 0 \end{cases} \quad (2)$$

$$LBP = \sum_{i=0}^{i-1} s\left(g_i - g_c\right) 2^i \quad (3)$$

where equation (1) is the binary code for computing the LBP algorithm, T is the obtained LBP binary code, and $t(x_1, x_2, \ldots, x_n)$ is the encoding function, and g_c is the value of the intermediate pixel point, and $g_0 \ldots g_8$ is the pixel value of 8 in the field around the center point. In equation (2), $s\left(x\right)$is the judgment function. Combine equation (1) and equation (2), let each pixel value is compared with the center pixel value, when the field pixel value is less than the value of the center pixel point, $s\left(x\right)$ is equal to 0, when the field pixel value is greater than or equal to the center pixel value, $s\left(x\right)$ is equal to 1. Eq. (3) is the decimal number to obtain the LBP algorithm. i is the number of domain pixel points around the center pixel point. Combining Eq. (2) and Eq. (3), the LBP value encoding of the center point is converted from binary to decimal, and Fig. 2 shows the LBP algorithm conversion diagram.

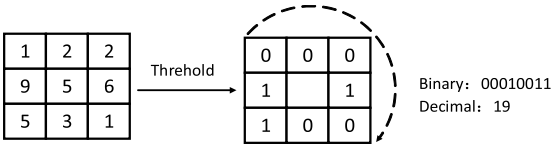

FIGURE 2. Transformation form of LBP algorithm.

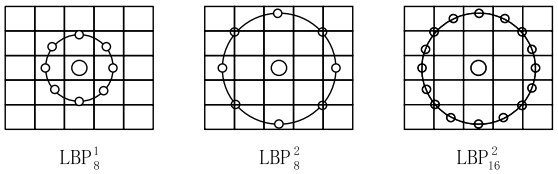

FIGURE 3. Equivalence model of LBP.

The traditional LBP texture pattern types are2^i To reduce the image dimensionality, this paper refers to the literature [18] and uses LBP Uniform Patterns instead of traditional LBP, the use of the advantages of the circular field, which is no longer limited to 8 one domain pixel points but can obtain multiple domain pixel points to meet the needs of various sizes and different frequency texture features. As shown in Fig.3

The LBP equivalence model reduces the LBP texture feature types to $i * (i - 1) + 3$ kinds and its algorithm is formulated as

$$\begin{cases} x_p = x_c + R * cos(\dfrac{2\pi p}{P}) \\ y_p = y_c - R * sin(\dfrac{2\pi p}{P}) \end{cases} \quad (4)$$

$$f\left(x_c, y_c\right)_{P,R} = \sum_{p=0}^{P} s(g_p - g_c)2^p \quad (5)$$

where x_c, y_c are the circle center coordinates, R is the radius of the sampled circular field, P is the number of points on this circular field that need to be sampled, p is the pth sampled point among the total P sampled points, x_p, y_p are the point$p$ coordinates, g_p is the grayscale value of point p sampled on the circular field, g_c is the grayscale value of its circle center, $f\left(x_c, y_c\right)_{P,R}$ is the texture feature of the image in the circular domain with the center coordinates of x_c, y_c and the radius of R. The original image and the result calculated by the LBP algorithm are shown in Fig. 4.

To be able to express the spatial structure characteristics of the human face, the face map is divided into M blocks, and each block of the face histogram is noted as R_m. obtain the spatial histogram features of each face sub-image, and connect them to get the spatial histogram of the complete face

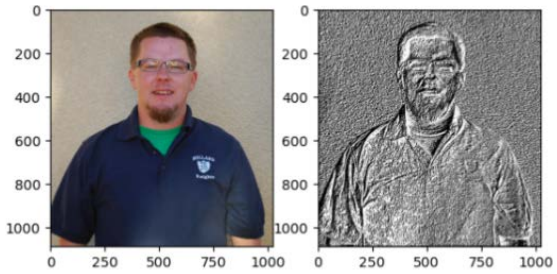

FIGURE 4. The left image is the original image, and the right image is the feature map obtained by the LBP algorithm transformation.

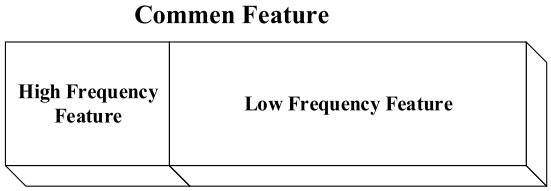

FIGURE 5. Frequency characteristics of the current image.

image.

$$R_m = f_m(x_c, y_c)_{P,R} \tag{6}$$

$$H = [R_m] = [f_m(x_c, y_c)_{P,R}], m = 1, 2 \ldots M \tag{7}$$

where R_m is the LBP histogram of image block m, and $f_m(x_c, y_c)_{P,R}$ is the texture feature of image block m in Eq (5). and H denotes the LBP histogram of the m block of face images connected to get the complete face image space histogram. The texture features of the image obtained by using the LBP operator can not only reduce the effect of illumination significantly but also express the regional features of the face image and form the global features of the face image when the regional features are connected.

B. FEATURE CONSTRAINTS

Although most of the effects from lighting can be handled in the preprocessing stage, there will still be false detection and missed detection due to noisy data interference such as occlusions. We divide the image features into high-frequency features and low-frequency features according to different feature mappings based on frequency feature perception [19], as shown in Fig 5, The current frequency features of the image are high-frequency features and low-frequency features have the same height and width, but the perceptual field of low-frequency features is significantly higher than that of high-frequency features.

Use information exchanges between high-frequency features and low-frequency features. Set the size of the low-frequency part to (0.5h,0.5w), and the length and width to be exactly half of the high-frequency part (h, w). We set the hyperparametersα to control the ratio of high and low-frequency feature segmentation. The algorithm formula is.

$$X \in \mathbb{R}^{c*h*w} \tag{8}$$

$$X^H \in \mathbb{R}^{(1-\alpha)c*h*w} \tag{9}$$

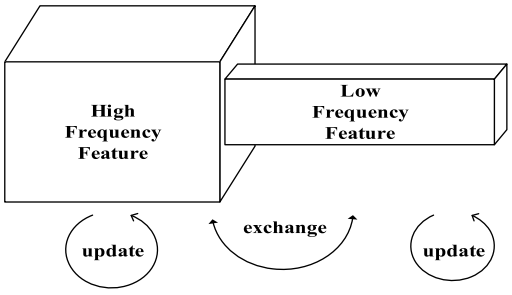

FIGURE 6. Frequency-aware feature extraction.

$$X^L \in \mathbb{R}^{ac*\frac{h}{2}*\frac{w}{2}} \tag{10}$$

where R is the set of all features, X is the set of high-frequency features and low-frequency features, h, w is the height and width of the features, c is the channel number, X^H is the high-frequency feature of the image by threshold α captures the precise detail features in the image, as shown in Eq. (9), X^L is the low-frequency features of the image, slowing down the low-frequency features from changing the spatial dimension by reducing the width and height. as shown in Eq. (10). After frequency feature extraction, reducing the width and height of the low-frequency features proportionally can reduce the perceptual field of the low-frequency features, while exchanging information between the low-frequency features and the high-frequency features, so that the frequency features of the reduced image can preserve most of the information of the original image.

In the feature update operation, the high-frequency and low-frequency features are updated within the corresponding frequencies. The feature exchange operation will update the information of high-frequency and low-frequency features between different frequencies. Therefore, the high frequency features include not only its information processing but also the mapping from low-frequency to high-frequency and vice versa. Frequency-aware features can obtain a larger receptive field of low-frequency features while exchanging high and low-frequency information, and the obtained image features can interpret the original image comprehensively. Compared with general feature extraction, it effectively doubles the receptive field and obtains more contextual feature information, which further improves the learning ability of Siamese neural networks and improves recognition performance [20]. As shown in Fig. 6.

C. FEATURE DISCRIMINATOR

There are two identical network structures in Siamese neural networks, and the VGG16 network structure is mainly used for the classification of images, so this paper uses the VGG16 network structure for feature extraction of images. Training with the same weights and parameters in both network structures, one as a model for training and one as a search for training. Input an image in each of the two networks, then train the two images, and finally, compare the distance between the two images to perform image recognition [21].

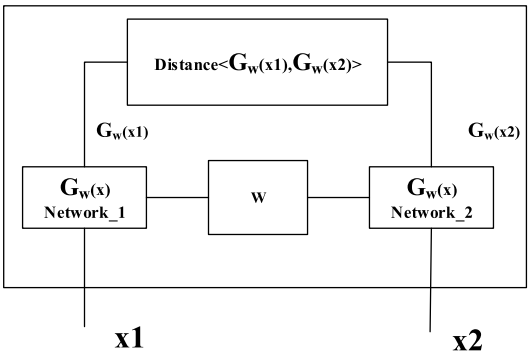

FIGURE 7. Structural model of Siamese neural network.

It applies to problems such as target detection under large data sets and multi-target recognition under small data sets, and its model structure is shown in Fig. 7

In the Siamese network, metric learning is used to discriminate [22] the feature similarity of two images. First, as shown in Fig. 7, Input an image in Network_1, enter the convolutional layer, pooling layer, and fully connected layer through the Siamese neural network in turn, and finally output the feature vector $f\ (x_1)$ of this image, input another image in Network_2, perform the same operation on the image in Network_2, and finally get its feature vector $f\ (x_2)$. To be able to determine whether two pictures are the same person, we define a threshold (hyperparameter), if the distance between the encoding results of two pictures is less than the threshold, then these two pictures are one person if the distance is greater than the threshold, then the two pictures are not the same person, the algorithm formula is

$$d\ (x_1, x_2) = \| f\ (x_1) - f\ (x_2)\ \|_2^2 \tag{11}$$

$$\begin{cases} d\ (x_1, x_2) \geq w \\ d\ (x_1, x_2) < w \end{cases} \tag{12}$$

where $d\ (x_1, x_2)$ are the distance function, $f\ (x_1)$, $f\ (x_2)$ are the feature vectors of the images in Network_1 and Network_2, respectively. And Ew is the threshold value. By combining Eqs. (11) and (12), the distance difference between the feature vectors of two images is calculated, and then the difference is compared with the threshold value, if the difference is greater than the threshold value, the two images are not the same target, and if the difference is less than the threshold value, they are the same target.

To be able to have a good encoding of the input images and learn the corresponding parameters autonomously, the triple loss function will be used here. The triple loss function refers to the calculation of a loss function using three images: a fixed image A, a positive image P (which is the same person as the fixed image), and a negative image N (which is not the same person as the fixed image). So, we let the encoding distance $d\ (A, P)$ between the fixed image A and the positive image P be smaller than the encoding distance $d\ (A, N)$ between the fixed image A and the negative image N. That is, we want the images of the same person to be closer to each other, and the pictures of different people to be farther away from each other. The triple loss minimizes the distance between the fixed image A and the positive image P, which both have the same identity while maximizing [23] the distance between the fixed image A and the negative image N. The model may make the same encoding for different images during the learning process, which is equivalent to the distance between images as 0, so on top of that we add a hyperparameter α to avoid this situation, making a gap between $d(A, N)$ and $d\ (A, P)$ with the

$$d\ (A, P) + \alpha \leq d(A, N) \tag{13}$$

Combining the distance equation in the above equation is

$$\| f(A) - f(P)\ \|^2 + a \leq \| f(A) - f(N)\ \|^2 \tag{14}$$

Further, get that

$$\| f(A) - f(P)\ \|^2 - \| f\ (A) - f\ (N)\ \|^2 + a \leq 0 \tag{15}$$

The loss function is obtained as

$$L\ (A, P, N) = max(\| f\ (A) - f\ (P)\ \|^2 - \| f\ (A)\, x - f\ (N)\ \|^2 + a, 0) \tag{16}$$

Its cost function is the sum of the loss functions in all individuals.

$$J = \sum_{i=1}^{n} L(A) \tag{17}$$

Its overall network structure diagram is shown in Fig.8

Considering that the computation time is mainly affected by the image convolution time complexity $O(\sum_{l=1}^{D} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l)$, where D is the number of convolution layers the algorithm has, l is the lth convolution layer, K is the complexity of image convolution affected by the edge length of the convolution kernel, M is the edge length of the feature map output from each convolution kernel, C_l is the output from the *lth* convolution kernel Compared with the convolution operation, the algorithm complexity of LBP and frequency feature convolution operation is negligible, the number of iterations is iter, and the total computation time of its algorithm complexity is $O(iter \cdot \sum_{l=1}^{D} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l)$

III. EXPERIMENT

A. DATA SET AND PARAMETER SETTING

In this section, we experimentally verify whether the previously mentioned improving method is effective for target recognition under Non-restrictive face recognition conditions, with the classical dataset CASIA-WebFace [35]. This dataset is a semi-automatic way to collect face images on the IMDb website, which contains 10,575 people, a total of 494,414 images, and each target photo has a label. In the CASIA-WebFace dataset, each category has a main photo containing only the target. According to the label and the main photo, pictures containing the target are classified by the clustering method, and then the target image is cropped. Finally, wrongly classified photos are manually screened.

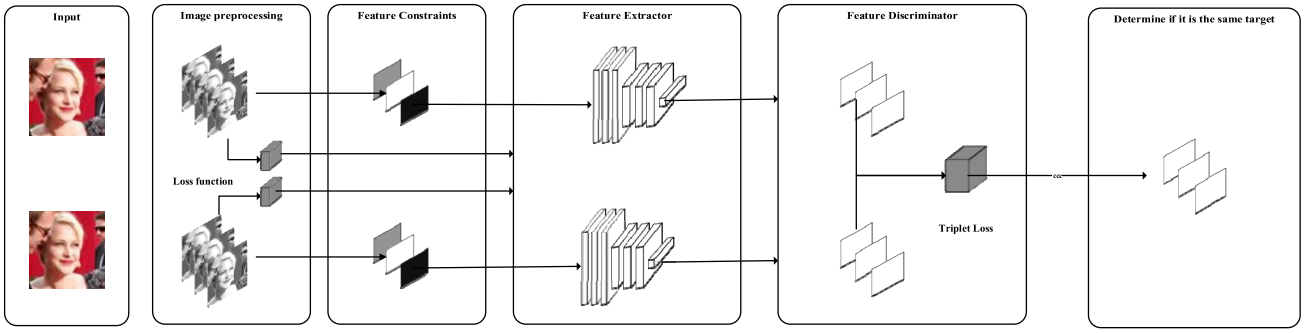

FIGURE 8. SN-LF network structure diagram.

FIGURE 9. Partial images of CASIA-WebFace dataset.

In this paper, 300 different characters are randomly selected in CASIA-WebFace, and 15 photos of each character are selected for experimental research. Some pictures of the data set are shown in Figure 9:

B. EXPERIMENTAL ENVIRONMENT AND PARAMETER SETTINGS

The experimental hardware environment is Intel core i3 CPU, Geforce 710m GPU, 2G memory, Win 10 operating system, Tensorflow deep learning framework, programming language, and version is python 3.6. The highest recognition rate of the network structure was determined by comparison when the number of layers of the network convolutional layers was 5. When the number of convolutional layers is constant, the frequency feature perception is performed on the convolutional layers, and the network structure with the highest recognition rate and lowest loss convergence is obtained by learning the image features, so the overall network structure is determined as 5 convolutional layers, 5 MaxPool layers, and 1 fully connected layer, and the network structures of Siamese neural network and Siamese Network based on LBP and Frequency Feature are noted

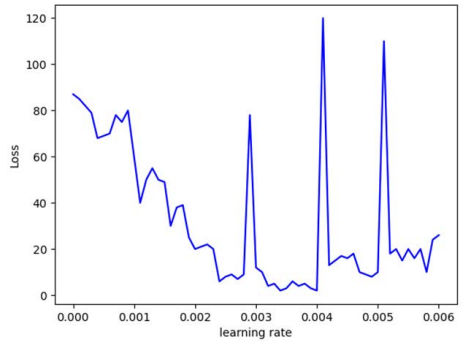

FIGURE 10. The variation of loss value with learning rate under Siamese neural network.

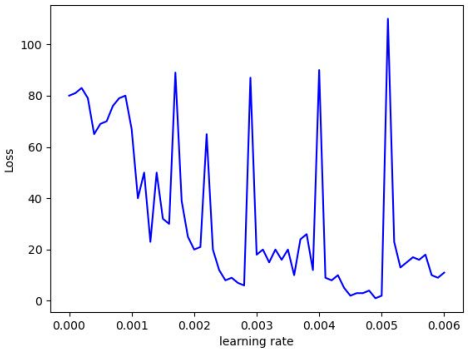

FIGURE 11. SN-LF loss value with learning rate.

as SN and SN-LF, respectively. Determine the learning rate range. The initial value of the learning rate is set to 1e-3 according to experience. After 5000 iterations of training, the learning rate is linearly increased at a growth rate of 1e-4 in each batch of the training process. Record the learning rate and loss values of the Siamese neural network and SN-LF under the CASIA-WebFace dataset respectively, and draw their relationship diagrams, as shown in Fig 10 and Fig 11:

Analysis of Fig 10 shows that the loss value decreases when the learning rate increases linearly, and once the learning rate increases to a certain degree the loss value increases instead, so the range of the periodic change

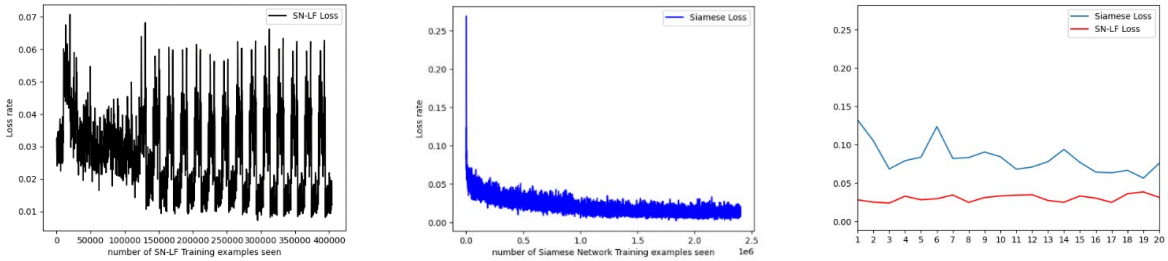

FIGURE 12. Comparison of the loss value of the algorithm in this paper and the Siamese neural network.

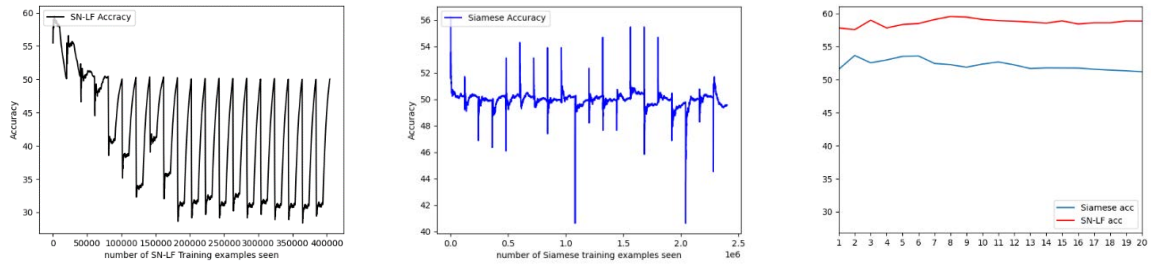

FIGURE 13. Accuracy comparison between the algorithm in this paper and Siamese neural networks.

of the learning rate is chosen as the interval where the corresponding loss value decreases rapidly, and the parameter settings in this paper refer to the literature[24], for the CASIA-WEBFACE dataset, the SN and the SN-LF structure based on the learning rate lower bound MinLR is chosen is 0.001, and the upper bound of learning rate MaxLR is 0.005.

As Fig 12 shown in the figure shows the change in the loss rate of this paper algorithm, although the loss rate of this paper algorithm fluctuates more, overall the loss rate of the algorithm is 0.04 around, and the fluctuation frequency is the same, which shows that the convergence of this algorithm is fast, and at the same time in the comparison with the classical Siamese network algorithm, it is found that the loss rate of our algorithm is significantly lower than the classical Siamese neural network, the loss rate of our algorithm is more stable, and the fluctuation of up and down is 0.03 around. As shown in Fig13, compared with the average accuracy of the Siamese neural network algorithm, the proposed algorithm is improved by about 6%, and the maximum difference is 10%.

In addition, it is necessary to set the parameters to change the learning rate to avoid the fitting condition. The literature [25] suggests that the learning rate should be changed after 2-10 iterations to prevent overfitting from occurring in too many iterations and to improve the detection efficiency. The experimental results show that there is little difference between the results when the step size is set to 2 times or 10 multiples of iterations. In this paper, we know that the efficiency of image recognition is the best when iterations are performed 5, so we use 5 iterations time.

C. COMPARATIVE EXPERIMENTS OF THE ALGORITHMS IN THIS PAPER ON THE CASIA-WEBFACE DATASET

In this paper, the Mean Average Precision (mAP) and the recognition rate are used to evaluate the performance of the algorithm under the CASIA-WEBFACE dataset. The comparison between this algorithm and the classical Siamese neural network algorithm shows that the similarity of this algorithm is improved by about 15% compared with the classical Siamese neural network, which indicates that the recognition effect of this algorithm on the target features in the recognition process is higher than that of the classical Siamese neural network. The experimental results are shown in Fig.14.

In this paper, the Mean Average Precision (mAP) and the recognition rate are used to evaluate the performance of the algorithm under the CASIA-WEBFACE dataset. The comparison between this algorithm and the classical Siamese neural network algorithm shows that the similarity of this algorithm is improved by about 15% compared with the classical Siamese neural network, which indicates that the recognition effect of this algorithm on the target features in the recognition process is higher than that of the classical Siamese neural network. The experimental results are shown in Fig.14.

Table 1 is the recognition rate of the traditional Siamese neural network and the algorithm SN-LF in different iterations, as can be seen from the table, in the same iteration case, SN-LF compared with the traditional Siamese neural network recognition rate increased by about 5%, in the same number of iterations, the recognition efficiency of the traditional Siamese network is not as good as the recognition efficiency of SN-LF, as the number of iterations increases,

FIGURE 14. Comparison of the experimental effect of classical Siamese neural network and this paper's algorithm (the left figure is the classical algorithm, the right figure is this paper's algorithm).

TABLE 1. Comparison of recognition rates between traditional Siamese neural network and SN-LF under different iterations.

Method	1000	2000	3000	5000
Siamese Network	65.89%	78.34%	82.16%	82.45%
SN-LF	70.23%	83.25%	87.12%	90.23%

the recognition rate of the traditional Siamese neural network is not as robust as the SN-LF algorithm compared to the recognition rate of SN-LF. This shows that in the image recognition of traditional Siamese Neural Networks when the target is in the case of occlusion, the recognition accuracy and efficiency will decrease, and the LBP algorithm in SN-LF can reduce the interference caused by high-frequency noise, while the frequency characteristic perception of the compression and information exchange of high-frequency and low-frequency features does not reduce the recognition accuracy while reducing the noise interference.

From the comparison of the mAP curves of classical Siamese neural network algorithm and SN-LF in Fig 15, it is found that the difference between the mAP curves of SN-LF and classical Siamese neural network algorithm is not much before the 1000 iterations, and with the increase of the number of iterations, the difference between them becomes larger and larger, reaching the maximum difference of 15% at 3000 iterations, which indicates that with the increase of the number of iterations, the convergence speed of SN-LF algorithm is faster than that of This indicates that the SN-LF algorithm converges faster than the classical Siamese neural network as the number of iterations increases. When the number of iterations increases to 5000 seconds, the mAP curve reaches about 90%, which indicates that the algorithm has strong robustness.

To explore the comparison of recognition effect between SN-LF and traditional face recognition algorithm in non-restricted face recognition, this paper refers to the literature [26], at the number of iterations of SN-LF is 3000 times, from the table 3 we know that the algorithm of this paper and different algorithms in the comparison of recognition rate, compared to other algorithms has improved 7.1%, 1.97%, 3.71%, 25.22%, 4.86%. The average image processing speed is improved by 0.186s, 0.141s, 0.017s, 0.195s, and 0.147s compared to other algorithms. In the literature 22. Its

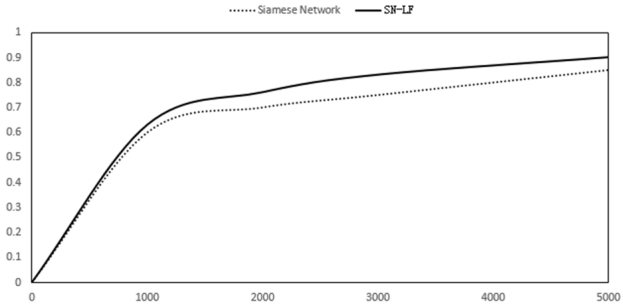

FIGURE 15. mAP curve of classical Siamese neural network algorithm compared with SN-LF.

TABLE 2. Comparison of recognition rates between different algorithms.

Method	Recognition rate	Average Time(s)
Resnet [36]	80.12%	2.043
TripleNet [37]	85.25%	1.978
FaceNet [38]	83.47%	1.854
FACE [25]	61.00%	2.032
Siamese Network	82.36%	1.984
SN-LF	87.22%	1.837

face algorithm has a recognition rate of 61%, but in the literature, its algorithm is complex and has more parameters and complex manual intervention. The comparison of mAP between different algorithms is shown in Fig 16. Combined with the table2, the algorithm in this paper is ahead of the traditional algorithm in both accuracy and mAP. As can be seen from Tables 1 and 2, the algorithm in this paper can maintain a stable recognition rate in complex samples, which shows that SN-LF can accurately recognize the target under the interference situation, thus further proving the robustness of the algorithm of this paper for pedestrian recognition under the unrestricted situation with high resistance to interference.

Through the experimental results of this paper's algorithm under the CIASIA-WebFace dataset, it is known that SN-LF using texture features as input helps to reduce redundant information compared to pixel-level input, while adding frequency feature perception makes the space utilization

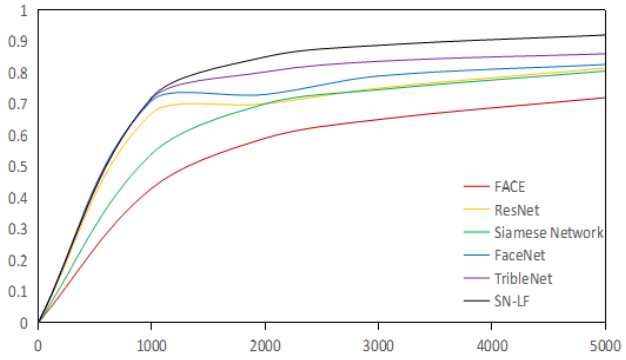

FIGURE 16. mAP curve comparison among algorithms.

TABLE 3. Comparison of accuracy rates for different iterations under the Yale-Bdataset.

Method	1000Next iteration	2000Next iteration	3000Next iteration	5000Next iteration
Siamese Network	75.89%	85.44%	92.36%	92.15%
SN-LF	80.23%	88.25%	94.12%	94.23%

TABLE 4. Comparison of the accuracy of different algorithms of Yale B.

Method	Recognition rate	Average Time(s)
Resnet [36]	80.85%	1.534
TripleNet [37]	89.53%	1.256
FaceNet [38]	91.48%	1.325
FACE [25]	91.10%	1.246
Siamese Network	92.20%	1.159
SN-LF	94.50%	1.147

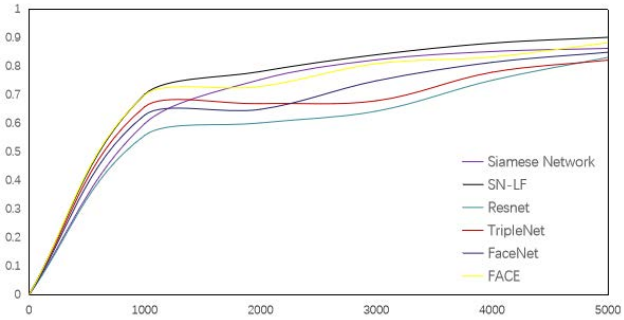

FIGURE 17. mAP curves of each algorithm under the Yale-B dataset.

improved, thus enabling the experimental accurate recognition of face images under non-unrestrictive conditions.

D. COMPARISON EXPERIMENTS OF OUR ALGORITHM ON YALE-B STANDARD FACE DATASET

To test the recognition accuracy of the face in the lighting situation of the algorithm in this paper, this paper uses the Yale-B standard face dataset to compare experiments with the traditional algorithm, the Yale-B dataset has 15 people, each person has 11 different expressions of the picture, and there are 3 different lighting situations, each person selects 6 pictures as the training sample, Table 3 is the image recognition effect of the Yale-B face library under different iterations. From table 3, we know that SN-LF is better for face recognition under illumination, and the recognition accuracy improves with increasing iterations until the image recognition accuracy no longer changes significantly at about 3000 iterations. From4 the table, we know that this algorithm improves 17.88%, 9.27%, 7.18%, 4.40%, and 2.20% in the Yale-B dataset compared with other algorithms, which shows that this algorithm has better recognition efficiency even under the illumination condition, and also shows that this algorithm can handle the noisy data well in the case of interference. And the average speed of each image processing has also improved compared to other algorithms, with the fastest improvement in 0.375 seconds. Also combined with Table 3, it is known that the algorithm in this paper has a faster convergence rate during the iterative process and better robustness compared to the traditional Siamese neural network. It shows that the algorithm of this paper also shows high performance under restrictive conditions.

To evaluate the performance difference between each algorithm and the algorithm of this paper, the mAP comparison between the algorithms in Fig. 17 shows that the algorithm in this paper also has better performance under the Yale B dataset. At about 1000 iterations, the performance between each algorithm is close to each other, with the increase in the number of iterations, there is a difference between each algorithm, some algorithms have a large fluctuation, which indicates that the anti-interference performance is reduced under the condition of the sudden change of light mutation. The algorithm in this paper uses the characteristics of the LBP algorithm that is insensitive to changes in light and can still have a better recognition effect in the case of sudden changes in light. As the number of iterations increases, the mAP value of SN-LF increases by nearly 20%, indicating that the algorithm in this paper has good robustness in the learning process. When the number of iterations reaches about 5000, the mAP value of SN-LF improves about 3%-10% compared with other algorithms which shows that the algorithm in this paper can extract more target features for learning during the iterations.

To verify the effectiveness of the method in this paper, this paper is compared with the mainstream and latest methods on the Yale-B dataset, which are LBP+DBN, IMISCNN [15], GWF [27], LTV [28], RLS (LOG-DCT & NPL-QI) [29], TT [30] and CLC [31], and the final experimental results are shown in Table.5

As can be seen from the table, when comparing mainstream and latest methods, the recognition performance of LBP+DBN slows down when the texture features are not obvious; IMISCNN and CLC have good performance in recognition, but overfitting will occur when there is target occlusion; LTV, RLS, TT, GWF are not sensitive to the processing of illumination changes. The method in this paper still has good processing ability when the

TABLE 5. Comparison between the algorithm in this paper and the mainstream algorithm.

Method	Recognition rate
LBP+DBN	90.7%
IMISCNN	92.8%
GWF	86.5%
LTV	82%
RLS(LOG-DCT	83.6%
RLS(NPL-QL)	86.9%
TT	82.6%
CLC	89.4%
SN-LF	95.2%

FIGURE 18. Loss rate variation curve with iterations.

illumination changes are too large and occlusion occurs, and the target features are retained. These fully verify the robustness of this paper to noisy data such as illumination and occlusion.

E. COMPARISON EXPERIMENTS OF THE ALGORITHM IN THIS PAPER ON THE LFW DATASET

To better determine the portability of the algorithm of this paper, the algorithm of this paper conducts comparison experiments on LFW, LFW dataset collects more than 13000 photos on the Internet, consisting of photos in natural scenes with different orientations, expressions, and lighting environments, there are more than 5000 people, among which 1680 people have more than two face pictures, each face photo has a uniquely identified name ID and serial number The LWF dataset is mainly tested for the recognition rate of human faces., we randomly select 300 people with a total of 720 photos to verify the accuracy of the algorithm in this paper. From Fig. 18, it can be seen that the loss rate of the algorithm tends to be stable and remains the same when the iterations are run 1000-1500 times, so this iteration is chosen at 1000 times.

The roc curve between SN-LF and Siamese in Figure19 and the accuracy of SN-LF is shown in Figure 20. shows that SN-LF rises faster compared to Siamese, and True Positive reaches 1 at around 0.2 in False Positive and stays there, indicating that the algorithm in this paper has better performance in face classification recognition compared to Siamese.

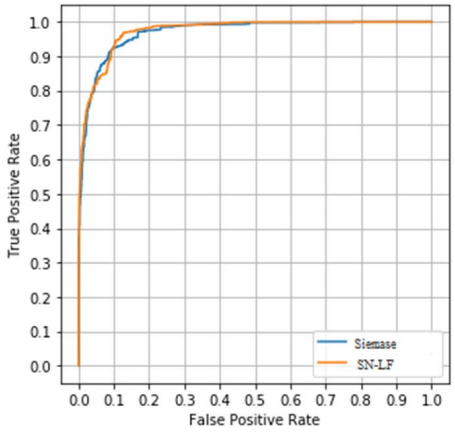

FIGURE 19. Comparison of roc curves of Siamese and SN-LF.

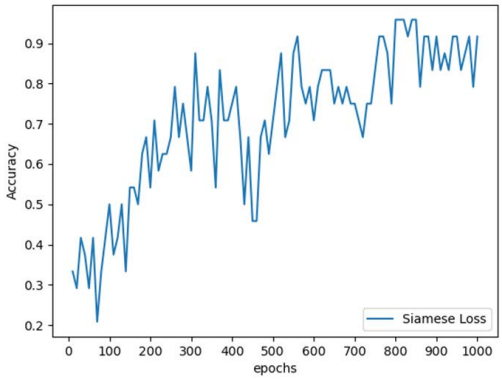

FIGURE 20. SN-LF accuracy curve with epoch.

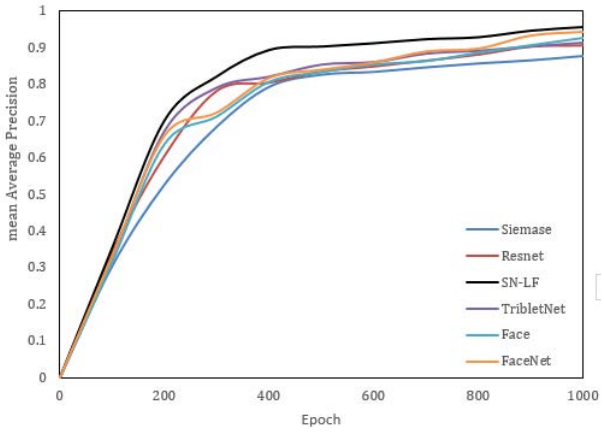

FIGURE 21. The variation of the mAP curve with epoch for each algorithm in the LFW dataset.

To evaluate the performance difference between each model algorithm and the algorithm in this paper, the mAP comparison between each algorithm in Fig 21 shows that the algorithm in this paper maintains a better performance during the iterative process, and at the end of the iteration, the algorithm in this paper is about 8% higher than the Siamese algorithm, which indicates that this algorithm also has a better recognition performance on the LFW dataset.

TABLE 6. Comparison between the algorithm in this paper and the mainstream algorithm.

Method	Recognition rate	Average Time(s)
Face[25]	90.67%	1.472
Resnet [36]	93.34%	1.386
TripleNet [37]	89.67%	1.452
FaceNet [38]	92.23%	1.521
Siamese	85.21%	1.386
SN-LF	94.54%	1.372

As can be seen from the table, when comparing this algorithm with the mainstream face recognition algorithms and frameworks in the case of 1000 iterations, it was found that the algorithm improved by 3.87%, 2.31%, 2.20%, 4.87%, 9.33% compared with other algorithms, and the average speed of each image processing was up to 0.149 seconds faster and the lowest was 0.014 seconds faster. It proves that the algorithm in this paper has a great improvement in the accuracy of face recognition, especially compared with Siamese Neural Networks by 9.33%, and the algorithm computation speed has improved compared with other algorithms, which proves that the improvement of this algorithm has a great improvement compared with the original algorithm.

IV. SUMMARY

Existing non-finite face recognition-based methods mainly start from the direction of face key points combined with denoising algorithms, which are prone to too fast convergence or poor robustness, and the algorithms are applied to a single scenario and cannot adapt to multiple scenarios. In this paper, we propose an architecture that combines frequency feature sensing, LBP, and Siamese neural networks, which makes the network converge faster and reduces the number of iterations required for the same recognition rate by changing the parameters. The algorithm first uses the LBP algorithm to pre-process the image to obtain texture features, which effectively reduces the interference caused by noisy data, and then inputs the texture features into the Siamese neural network, which improves the running speed and also enhances the recognition accuracy and robustness of the algorithm. The recognition accuracy is improved. Moreover, the proposed network has a simple structure and is suitable for face recognition of small-scale data sets under Non-restrictive face recognition conditions. The next problem to be solved is to reduce the algorithm running time and improve the efficiency of the algorithm, as well as to solve the case of recognizing multiple targets appearing in the image under Non-restrictive face recognition conditions.

REFERENCES

[1] K. Sundararajan and D. L. Woodard, "Deep learning for biometrics: A survey," *ACM Comput. Surv.*, vol. 51, no. 3, pp. 1–34, May 2018, doi: 10.1145/3190618.

[2] K. Nguyen, C. Fookes, S. Sridharan, M. Tistarelli, and M. Nixon, "Super-resolution for biometrics: A comprehensive survey," *Pattern Recognit.*, vol. 78, pp. 23–42, Jun. 2018, doi: 10.1016/j.patcog.2018.01.002.

[3] L. Nanni, G. Minchio, S. Brahnam, G. Maguolo, and A. Lumini, "Experiments of image classification using dissimilarity spaces built with Siamese networks," *Sensors*, vol. 21, no. 5, p. 1573, Feb. 2021, doi: 10.3390/s21051573.

[4] Q. Wang, "Study of robust principal component analysis and its applications," Ph.D. dissertation, Dept. Inf. Commun. Syst. Eng., Xidian Univ., Xian, China, 2019.

[5] L. Tian, "Research on feature learning based face recognition under unconstrained scenario," Ph.D. dissertation, Dept. Electron. Eng., Beijing Univ. Posts Telecommun., Beijing, China, 2018.

[6] V. J. Barranca, "Neural network learning of improved compressive sensing sampling and receptive field structure," *Neurocomputing*, vol. 455, pp. 368–378, Sep. 2021, doi: 10.1016/j.neucom.2021.05.061.

[7] J. Liu, "Research on the key technology of face recognition," Ph.D. dissertation, Dept. Inf. Commun., Beijing Univ. Posts Telecommun., Beijing, China, 2015.

[8] Y. Zhang, L. Tan, and J. Chen, "Cross-modality person re-identification based on joint constraints of image and feature," *Acta Autom. Sinica*, vol. 47, no. 8, pp. 1943–1950, Jan. 2021, doi: 10.16383/j.aas.c200184.

[9] Y. Liu, J. Jiang, M. Qi, H. Liu, and H. Zhou, "Video-based person re-identification method based on GAN and pose estimation," *Acta Autom. Sinica*, vol. 46, no. 3, pp. 576–584, Dec. 2018, doi: 10.16383/j.aas.c180054.

[10] Z. Zhao and G. Ni, "Automatic recognition algorithm for unconstrained multi-pose face key features under unqualified conditions," *Comput. Sci.*, vol. 46, no. 9, pp. 250–253, Sep. 2019.

[11] M. Roder, L. A. Passos, G. H. de Rosa, V. H. C. de Albuquerque, and J. P. Papa, "Reinforcing learning in deep belief networks through nature-inspired optimization," *Appl. Soft Comput.*, vol. 108, Oct. 2021, Art. no. 107466, doi: 10.1016/J.ASOC.2021.107466.

[12] S. Liang, Y. Liu, and L. Li, "Face recognition under unconstrained based on LBP and deep learning," *J. Commun.*, vol. 35, no. 6, pp. 154–160, Jun. 2014.

[13] N. Vinod and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. ICML*, Madison, WI, USA, Jun. 2010, pp. 807–814.

[14] S. Zagoruyko and N. Komodakis, "Learning to compare image patches via convolutional neural networks," in *Proc. IEEE-CVPR*, Boston, MA, USA, Jun. 2015, pp. 4353–4361.

[15] X. Xu, L. Zhang, B. Lang, and Z. Xia, "Research on inception module incorporated Siamese convolutional neural networks to realize face recognition," *Acta Electron. Sinica*, vol. 48, no. 4, pp. 643–647, Apr. 2020.

[16] H. Wu, "Face tracking using Siamese convolutional neural networks," *Comput. Eng. Appl.*, vol. 54, no. 14, pp. 175–179, Jul. 2018.

[17] Y. Shen, H. Wang, and Y. Dai, "Deep Siamese network-based classifier and its application," *Comput. Eng. Appl.*, vol. 54, no. 10, pp. 19–25, May 2018.

[18] X. Wu, X. Feng, A. Huang, X. Zhang, J. Dong, and L. Liu, "Survey on facial kinship verification," *Acta Autom. Sinica*, vol. 47, pp. 1–27, Apr. 2021, doi: 10.16383/j.aas.c201023.

[19] Y. Yang, W. Xing, S. Zhang, Q. Yu, X. Guo, and M. Guo, "A learning frequency-aware feature Siamese network for real-time visual tracking," *Electronics*, vol. 9, no. 5, p. 854, May 2020, doi: 10.3390/electronics9050854.

[20] Y. Chen, H. Fan, B. Xu, Z. Yan, Y. Kalantidis, M. Rohrbach, Y. Shuicheng, and J. Feng, "Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution," in *Proc. IEEE/CVF ICCV*, Seoul, South Korea, Oct. 2019, pp. 3434–3443.

[21] L. Fu, Y. Zhao, X. Sun, Z. Lu, D. Wang, and H. Yang, "Fast video object segmentation based on Siamese networks," *Acta Electron. Sinica*, vol. 48, no. 4, pp. 625–630, Apr. 2020.

[22] G. Shi and X. Zhao, "Object tracking algorithm based on jointly-optimized strong-coupled Siamese region proposal network," *J. Comput. Appl.*, vol. 40, no. 10, pp. 2822–2830, May 2020.

[23] H. Luo, W. Jiang, X. Fan, and S. Zhang, "A survey on deep learning based person re-identification," *Acta Autom. Sinica*, vol. 45, no. 11, pp. 2032–2049, Jan. 2019.

[24] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," Nov. 2014, *arXiv:1411.7923*.

[25] K.-C. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005, doi: 10.1109/TPAMI.2005.92.

[26] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition for uncontrolled pose and illumination changes," in *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 1, pp. 149–163, Jan. 2013, doi: 10.1109/TSMCA.2012.2192427.
[27] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010, doi: 10.1109/TIP.2010.2042645.
[28] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1519–1524, Sep. 2006, doi: 10.1109/TPAMI.2006.195.
[29] X. Xie, W. Zheng, J. Lai, P. C. Yuen, and C. Y. Suen, "Normalization of face illumination based on large-and small-scale features," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1807–1821, Jul. 2011, doi: 10.1109/TIP.2010.2097270.
[30] Y. Wu, Y. Jiang, W. Li, Z. Lu, Q. Liao, and Y. Zhou, "Generalized Weber-face for illumination-robust face recognition," *Neurocomputing*, vol. 136, pp. 262–267, Jul. 2014.
[31] X. Tu, F. Yang, M. Xie, and Z. Ma, "Illumination normalization for face recognition using energy minimization framework," *IEICE Trans. Inf. Syst.*, vol. E100.D, no. 6, pp. 1376–1379, Jun. 2017.
[32] Q. Zheng, M. Yang, J. Yang, Q. Zhang, and X. Zhang, "Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process," *IEEE Access*, vol. 6, pp. 15844–15869, 2018, doi: 10.1109/ACCESS.2018.2810849.
[33] Q. Zheng, M. Yang, X. Tian, N. Jiang, and D. Wang, "A full stage data augmentation method in deep convolutional neural network for natural image classification," *Discrete Dyn. Nature Soc.*, vol. 2020, pp. 1–11, Jan. 2020, doi: 10.1155/2020/4706576.
[34] Q. Zheng, P. Zhao, Y. Li, H. Wang, and Y. Yang, "Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification," *Neural Comput. Appl.*, vol. 33, no. 13, pp. 7723–7745, Jul. 2021.
[35] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," 2014, *arXiv:1411.7923*.
[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[37] W. Ma, Y. Cui, N. Shao, S. He, W.-N. Zhang, T. Liu, S. Wang, and G. Hu, "TripleNet: Triple attention network for multi-turn response selection in retrieval-based chatbots," Sep. 2019, *arXiv:1909.10666*.
[38] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823, doi: 10.1109/CVPR.2015.7298682.

CUNLI SONG was born in Shaanxi, China, in 1975. She received the M.S. degree from the School of Software, Dalian Jiaotong University, Dalian, China, in 2003, and the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, Dalian, in 2011. Since 2011, she has been an Associate Professor with the School of Software, Dalian Jiaotong University. She has authored more than 20 refereed papers. Her current research interests include evolutionary computation, multi-objective optimization, and flow shop scheduling.

SHOUYONG JI was born in Shandong, China, in 1995. He is currently pursuing the master's degree with the School of Software, Dalian Jiaotong University. His research interests include computer vision, machine learning, and deep learning.