IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY SECTION

Received 7 September 2022, accepted 16 October 2022, date of publication 28 October 2022, date of current version 9 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3218160

RESEARCH ARTICLE

Multiple Genetic Syndromes Recognition Based on a Deep Learning Framework and Cross-Loss Training

JIANFENG WANG[1], BO LIANG[2], LIJUN ZHAO[1], YUANFANG CHEN[3], (Member, IEEE), WEN FU[1], PEIJI YU[1], HONGBING CHEN[1], HONGYING WANG[4,5], GUOJIE XIE[6], TING WU[7], MUHAMMAD ALAM[8,9], (Senior Member, IEEE), HAITAO LV[5], AND LIN HE[10]

[1]Suzhou Life Intelligence Industry Research Institute, Suzhou 215125, China
[2]School of Biology and Food Engineering, Changshu Institute of Technology, Changshu 215506, China
[3]School of Cyberspace, Hangzhou Dianzi University, Hangzhou 310018, China
[4]Department of Clinical Laboratory, Children's Hospital of Wujiang, Suzhou, Jiangsu 215234, China
[5]Department of Clinical Laboratory, Children's Hospital of Soochow University, Suzhou, Jiangsu 215000, China
[6]Key Laboratory of Open Data of Zhejiang, Hangzhou 310013, China
[7]Hangzhou Innovation Institute, Beihang University, Hangzhou 311228, China
[8]Blockchain Laboratory of Agricultural Vegetables, Weifang University of Science and Technology, Weifang, Shouguang 310051, China
[9]School of Engineering, London South Bank University, London SE1 0AA, U.K.
[10]Bio-X Center, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200030, China

Corresponding authors: Lin He (helinhelin3@vip.163.com), Haitao Lv (haitaosz@163.com), and Muhammad Alam (alamm52@lsbu.ac.uk)

This work was supported in part by the Jiangsu province, China, under Grant WWK202020 and Grant F202119.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Medical Ethics Committee of Soochow University Children's Hospital under Approval No. 2021KS022.

ABSTRACT Many patients with genetic syndromes have special facial features, which boast significant potential value for clinical diagnosis. Deep learning and computer vision technology can be employed to diagnose genetic diseases by analyzing facial features of patients. As a matter of fact, the application of deep learning technology in the area of genetic diseases is confined owing to the difficulties of patient data acquisition. This study develops BioFace, a deep learning framework that can recognize multiple genetic diseases facial attributes based on limited datasets. BioFace is a deep neural network architecture designed premised on Resnet. To increase the weight of effective features and weaken the weight of invalid or unobvious features during extraction of facial features, we add Squeeze-and-Excitation (SE) blocks in the network. In combination with this network architecture, we designed a cross-loss training method based on transfer learning. This method can transfer the ability learned from the task of face identification to the task of recognition of genetic diseases facial attribute, and improve the inter-class distance of different genetic diseases and the intra-class distance of similar genetic diseases simultaneously. These render it possible for deep learning to be applied to recognition of multiple genetic diseases facial attribute with very small amount of data. In this research, we tested 10 syndromes with our framework and the Top-1 accuracy was 93.5%, which is the state-of-the-art in multiple genetic syndromes recognition research. In practical clinical applications, our framework and methods can be extended to the disease identification of more small datasets, potentially offering valuable assistance for the auxiliary clinical application of genetic diagnosis and other related genetic research.

INDEX TERMS Genetic syndromes, convolutional neural network, deep learning, transfer learning, cross-loss training.

The associate editor coordinating the review of this manuscript and approving it for publication was Kin Fong Lei.

I. INTRODUCTION

Genetic diseases affect almost 8% of the world's population [1] and have a significant impact on the entire life

course of those who are affected. Currently, there are over 7,000 kinds of known genetic diseases, but most patients with genetic diseases are unaware of their conditions, and neglect to go to professional medical institutions for diagnosis [2]. Prompt and early detection, prevention and treatment can help patients avoid some potential health-threatening problems, so the research on diagnosis and diagnostic methods of genetic diseases is of crucial significance and value [3].

Many genetic diseases affect the facial features of patients. Different genetic diseases carry their own unique facial features [4], [5], [6], [7], [8]. Premised on these facial features, clinical experts on genetic diseases can give some diagnosis or diagnostic suggestions accordingly. Nevertheless, diagnosis of genetic diseases for atypical clinical manifestations of some rare syndromes may be confined by the personal experience of clinical experts of genetic diseases. In some experimental studies, the expert's diagnostic accuracy for typical facial features of Cornelia de Lange Syndrome is only 87%, 54% for mild features [9], and the average accuracy is 77% [10]. In the study of Angelman Syndrome, the expert's diagnostic accuracy for its features is 71%, sensitivity is 60%, and specificity is 78% [11]. In general, on account of the rarity of some diseases and the numerous possibilities of diagnosis results, it takes a rather long time to realize a correct diagnosis [12]. Thus, many researches diagnose genetic diseases [13], [14], [15], [16], [17] by using different methods combined with computer vision.

Early research generally extracted artificial features by means of traditional image processing methods or traditional machine learning methods. Saraydemir et al. [18] proposed a representation extraction technology based on Gabor wavelet, then conducted Principal Component Analysis and Linear Discriminant Analysis (LDA), and then employed Support Vector Machine (SVM) and K-Nearest Neighbors, with 97.3% and 96% accuracy rates, respectively. Burcin and Vasif [19] used Local Binary Pattern (LBP) and template matching to detect Down's Syndrome from artificially cropped images, and applied local binary pattern to significant face markers to capture significant aspects of face, and manually marked all images with three-dimensional face representation, then classified by LDA and SVM respectively. Zhao et al. [4], [20], combined the representation methods based on local geometry and local texture to recognize Down's Syndrome, and proposed a local model of facial recognition based on hierarchical constraints. This method was used to test 130 subjects, including 50 Down's Syndrome patients and 80 healthy subjects.With the aid of the SVM, its accuracy reaches 97%. S Hadj-Rabia et al. studied the automatic recognition of XLHED phenotypes using face photographs [21]. In the research of Basel-Vanagaite [10], the Facial Dysmorphology Novel Analysis (FDNA) method was used to identify the face image of Cornelia de Lange, and the detection rate was 94%. Despite the promising results achieved in these efforts, they can identify only a single disease and can not meet the actual needs of clinical diagnosis. Kuru et al. [22] proposed a facial genotype phenotype diagnosis decision support system capable of detecting genetic diseases of patients. The system then tested 92 patients suffering from 15 different diseases, and the accuracy rate reached 53%. There is some other research on the identification and classification of multiple genetic diseases [13], [14], [23], but the accuracy of their results is too low to be applied in clinical applications.

In recent years, the Convolutional Neural Network (CNN) has been successfully applied to many image recognition tasks. Some researchers have also applied CNN to face recognition tasks for genetic diseases. Shukla et al. [24] proposed a framework to extract features using CNN, and SVM was used as a classifier to identify six genetic diseases. The average accuracy rate of the framework was 48% [25]. Yaron came up with a deep learning framework DeepGestalt [26] in face2gene application, which can recognize multiple genetic diseases. This work has collected over 26,000 cases of patients to train their model and achieved 61.3% top-1 accuracy in clinical test dataset. These works only used shallow neural networks, which cannot make full use of the feature extraction ability of deep learning and their results can not meet the needs of clinical diagnosis.

In consideration of the deficiencies in the above research in related areas, we put forward a new deep learning framework and cross-loss training strategy. Our method and framework can extract genetic disease face attribute features using only a small batch of data sets and achieve good recognition accuracy. Deep learning and improved Resnet64 are the foundations of our framework. In CNN, models with deeper layers can express more strongly, but the deeper network does not converge well since the gradient disappears during training. Thus, we cannot simply add many layers. The residual network can help solve this problem so that more layers can be built in the CNN. In comparison with face recognition methods, the task of identification of genetic diseases needs not merely local features of specific individuals, but more global information about a type of syndrome. We added SE block to Resnet to extract more global information in the deep layer [27], which helps to boost the effectiveness of extracting genetic disease features in tasks. Apart from that, the added block can help to enhance the weights of valid regions and weaken the weights of invalid regions. We also designed cross-loss training in the framework. Applying deep learning to the task of small data sets is extremely challenging. Thus, facial recognition model is taken as pre-training and fine-tuning is done with syndrome data [28]. However, there is a big gap between facial recognition and the identification of genetic diseases, and rich genetic disease features cannot be well extracted using a facial recognition model alone. To decrease the individual redundant information to be identified, we adopted the loss alternation method to train the model to maximize the interval between syndrome categories and minimize the individual interval in the same syndrome category. Finally,

we designed the improved network structure and carried out transfer learning in combination with cross-loss training, so that deep learning can achieve excellent recognition effects in recognition of genetic syndromes facial attribute in small data sets.

In this study, we tested our framework with 10 most common genetic diseases. Our framework is not confined to the recognition of these 10 genetic diseases. Rather, it is a universal framework model for recognition of various genetic diseases facial attribute. Combining more kinds of data, our framework can be extended to the recognition tasks of hundreds of genetic diseases with facial features. The contributions of this research are summarized as follows:

- We design a deep learning framework, which is applicable to recognition of multiple genetic syndromes. It is a Resnet64 network architecture designed premised on the deep residual network Resnet and added with SE block, which can increase the weight of effective features and weaken the weight of invalid or unobvious features during extraction of facial features, thus keeping redundant information to a minimum during recognition to gain a more ideal recognition effect.
- In combination with this network architecture, we designed a cross-loss training method for transfer learning. This method can transfer the ability learned from the task of face identification to the task of recognition of genetic diseases, and improve the inter-class distance of different genetic diseases and the intra-class distance of same genetic diseases simultaneously. This renders it possible for deep learning to be applied to the task of recognition of genetic diseases facial attritubes with a small data size.
- Combined with the above methods, our study achieved state-of-the-art performance in the task of identifying multiple genetic diseases. This makes it possible for recognition of genetic syndromes facial attribute to be applied in clinical practice.

II. METHODS

Our framework consists of face detection, image preprocessing, feature extraction and disease discrimination. As an end-to-end discrimination framework, our model is equivalent to a mapping function f(x), which inputs the patient's facial image and maps to the probability list of genetic disease discrimination. The overall flow is exhibited in the following figure:

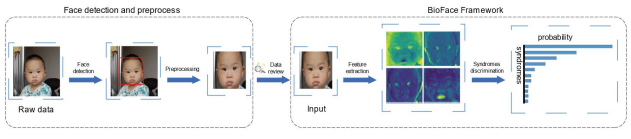

FIGURE 1. Overall flow of the genetic disease identification.

A. FACE DETECTION AND PREPROCESS

Face detection means that the input is an image containing human faces and the output is the bounding boxes of all faces. Face detection is supposed to be able to detect all faces in the image without missing or wrong detection. In order to get this result, firstly we use face detection algorithm to process all raw images. Then we manually review each detected result. For the wrong cases, we will conduct secondary detection or manually crop the face. After obtaining the face, the face is aligned. As the face in the original image may have differences in pose and position, the face should be aligned for the sake of unified processing afterwards. Our model is suitable for 2D images of patients without limiting scenes, in which the face detection algorithm is responsible for detecting the patient's facial landmarks and facial regions and correcting the face. The part of face detection and correction uses the algorithm of Multi-Task Convolutional Neural Network (MTCNN) [29] to put detection of face region and detection of face landmarks together for face detection and correction.

For the input original image with borders, this step is to detect the landmarks in the face, and then align the face based on these landmarks. The so-called landmarks are usually the positions of canthus, nose, contour of face and so on. With these landmarks, we can ''calibrate'' or ''align'' the face. It means that the original face may be crooked, and affine transformation is used here to ''straighten'' the face, so as to eliminate the errors caused by different postures [30].

Preprocessing of the detected face images principally includes face clipping, size change and other operations. The processed facial image is used to input the model to extract features and achieve disease discrimination.

When processing the raw images with MTCNN algorithm, only 93% of the faces were correctly detected, and the wrong results (7%) would be corrected in subsequent manual review. After these processes, we can create a dataset with valid faces to ensure that the data input to BioFace are 100% correct.

B. NETWORK ARCHITECTURE FOR IDENTIFICATION OF MULTIPLE GENETIC DISEASES

In the application of deep learning, the depth of network layers is crucial. The deeper the network layers, the better the feature extraction ability. However, for the reason of the disappearance of gradients, etc, the representation ability of the network will be weakened when the depth reaches a certain level. The emergence of batch standardization and residual network solves the problem to a certain extent through the jump connection [31].

For some tasks of computer vision, CNN extracts features through receiving field, which is much small in shallow layers, so it cannot extract global information unless through downsampling in the deeper layers. But in the shallow layer, the global information is helpful for improving the effectiveness, such as inception network. Squeeze and Exception [27] is a network block with light magnitude

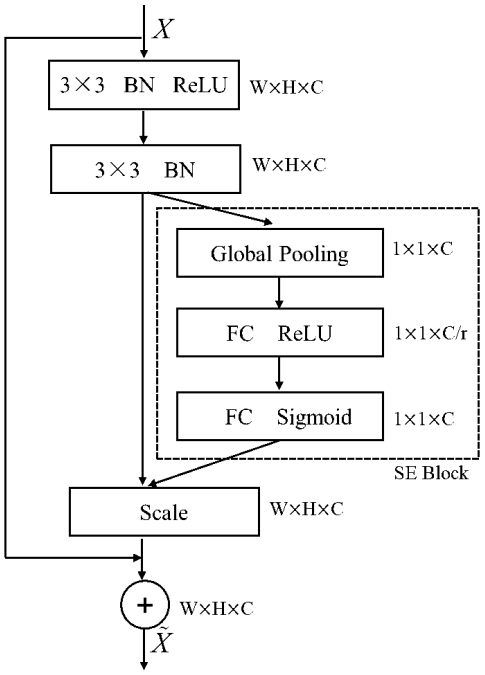

FIGURE 2. Structure of SE-Resnet Block. 3×3 means the size of convolution kernel. BN means the batch normalation layer and FC means fully connection layer. W×H×C and 1×1×C means the size of output feature map.

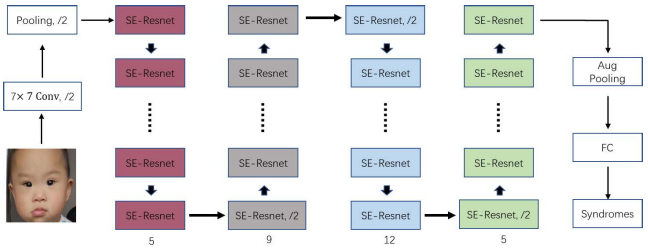

FIGURE 3. Overall network architecture. A SE-Resnet rectangular box represents a SE-Resnet block. The number under the box is the number of the blocks in each column. /2 means that the stride is 2.

and small resource consumption. We can utilize the global information extraction capability of SE Block to make the network fully use the global information of the shallow layer. This can boost the network's ability to extract global facial features of genetic diseases. The structure of SE-Resnet block is shown in Figure2. In comparison with task of facial recognition, the target of genetic disease identification requires more global information about the same diseases besides the local features of specific individuals. Moreover, the added SE Block can enhance the weights of valid regions and weaken the weights of invalid regions.

In the principle of residual network, we have designed the residual network architecture SE-Resnet64 that integrates SE-Resnet Block. The structure of the overall network is shown in Figure3.

The main structure of the network consists of four sets of SE-Resnet blocks, and the number of each set of SE-Resnet blocks is 5,9,12,5 respectively. After a convolution, the input image will enter the SE-Resnet sets to gradually extract the genetic disease attribute characteristics, and the extracted characteristics will enter the full connection layer after an average pooling, finally realizing the identification of multiple types of genetic diseases.

C. CROSS TRAINING FOR FACIAL RECOGNITION OF GENETIC DISEASES IN SMALL DataSets

Deep learning applications premised on small data sets are easily prone to over-fitting,which makes the network not have good generalization ability. The method of pre-training model combined with fine-tuning training can solve this problem [32]. Fine-tuning is based on a pre-training model that was trained on large data sets. On this basis, the target data set is used in fine-turning process to make the pre-training model have the ability to express the target data. For transfer learning with similar tasks, the fine-turned network can have better representation ability. We use large-scale facial recognition data in pre-training, then use genetic disease data to fine-tune. In face identification and recognition of genetic diseases, the two tasks are of some similarity, but there are also great differences. The target of face identification is to identify the different images of the same person, while recognition of genetic diseases is to identify face clusters with the same genetic disease. To reduce the individual redundant information in the task of face identification, we compared the loss functions softmax and A-softmax [33], expressed by the following formula (1) and formula (2), respectively. Softmax can be used to improve the inter-class spacing, but the optimization of intra-class spacing is relatively poor. A-softmax has a good effect on the optimization of intra-class spacing. In this work, we need a large intra-class spacing in pre-training to avoid the network learning the individual redundant information and a small intra-class spacing to keep the faces of the same diseases together as possible in the fine-turning stage.

$$L = \frac{1}{N}\sum_{i} -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right) \tag{1}$$

$$L_A = \frac{1}{N}\sum_{i} -\log\left(\frac{e^{\|\boldsymbol{x}_i\|\varphi(\theta_{y_i,i})}}{e^{\|\boldsymbol{x}_i\|\varphi(\theta_{y_i,i})} + \sum_{j\neq y_i} e^{\|\boldsymbol{x}_i\| cos(\theta_{j,i})}}\right) \tag{2}$$

$$\varphi(\theta_{y_i,i}) = (-1)^k cos(m\theta_{y_i,i}) - 2k, \theta_{y_i,i} \in [\frac{k\pi}{m}, \frac{(k+1)\pi}{m}] \tag{3}$$

In formula (1), f_j represents the j-th element in vector **f** (j $\in$ [1, K], K is the number of classes), N is the number of training samples. In formula(2), formula(3), k$\in$[0, m - 1], "m" is an integer greater than 1 and it controls the angular spacing. We use the loss function of softmax to train the facial recognition network with a large amount of facial recognition data, and then use the loss function of A-softmax to fine-tune the network with genetic diseases small datasets.

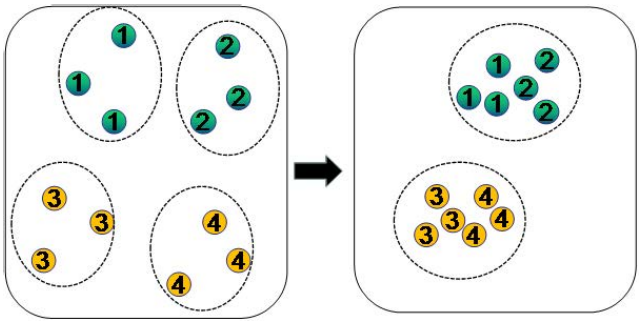

FIGURE 4. Process of optimizing the spacing among the subjects with cross-training. The circles with same number are the faces with same identity and the same color are the same syndromes.

In this way, we can optimally distinguish different types of genetic diseases by maximizing the inter-class spacing and minimizing the intra-class spacing. This process is shown as Figure4. The circles with the same number are classified as the same identity and the circles with the same color are classified as the same type of genetic disease. The network can optimize the spacing among classes but can not optimize the inter-class spacing well in the first stage. With the loss-alternating training in the second step, the network can minimize the distance among the subjects in the same classes and ignore the redundant information in the face identification.

III. EXPERIMENTS

A. DATASETS

We used Casia Webb-Face dataset [28] for facial recognition pre-training. Casia Webb-Face is the most commonly used public facial recognition dataset, containing 490,000 face photos of over 10,000 different people.

Our application BioFace is deployed by many Chinese healthcare professionals, and BioFace collects data through clinical users. For the test set, we only use cases in the clinical data that have clear clinical or molecular diagnostic results. The amount of clinical data is limited. Therefore, in addition to clinical data, our training set also collects some Internet data as a supplement,so that the number of test sets accounts for about 10 % of the total data. We totally collected face photos of 10 kinds genetic disease syndromes patients and data with very low resolution and no full front face are excluded. The collected data and numbers are exhibited in Table 1. This study also collected the data of healthy people from public dataset CAS-PEAL [34] as a category and added them to model learning.

B. EXPERIMENTAL SETUP

We use Pytorch for deep learning architechture. For initial method we use Xavier. For the optimizer, we use Stochastic Gradient Descent (SGD), momentum 0.5, weight decay of 5e-4. The batch size is 512 and the learning rate is 0.01. We train for 100 epochs and keep a constant learning rate in the training process. After 100 epochs, the pre-training with softmax loss is done and we test the model on Labeled Faces in the Wild (LFW) [35] dataset, which is very popular for person verification test, and the accuracy of the result is 98.4%. After the pre-training stage, we change the softmax loss function with A-softmax and train for 200 epochs. In this step, the ''m'' for A-softmax is set as 4. For all training, we use online data augmentation method, randomly rotation within 5 degrees and vertical and horizontal shifts randomly within 20 pixels, randomly scaled by 0.05 times of the size and shear transformation.

TABLE 1. Data situation.

No.	Name of genetic disease	Total quantity	Training set	Test set
1	Angelman syndrome	471	424	47
2	Cornelia de Lange syndrome	194	175	19
3	Downs syndrome	672	605	67
4	Sotos syndrome	141	127	14
5	CHARGE syndrome	162	146	16
6	Noonan syndrome	263	237	26
7	DiGeorge syndrome	280	252	28
8	Williams-Beuren syndrome	209	188	21
9	Rubinstein-Taybi syndrome	225	203	22
10	Fragile x syndrome	158	142	16
0	Health	400	360	40

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. THE EVALUATION OF CROSS TRAINING

We conducted comparative experiments to verify the effect of cross-training on the spatial distribution of data features. The first experiment does not use cross-training, that is, softmax is used in the pre-training and fine-tuning training stages. The second experiment uses the cross-training method described above. In order to make a fair comparison, all the experimental settings are the same except for the comparative components in the two experiments. When the training of the two experiments is completed, we take the output of the previous layer of classifier as the characteristic representation of the data. For visualization, we reduced the characteristic representation to two dimensions. The representation distribution for both is shown in Figure5. As can be seen from the figure, using the cross-training method, the data feature distribution has been significantly optimized. The intra-class data are more gathered and the inter-classe are more separate by cross-training compared with all using softmax loss function.

When the spatial distribution of data features is optimized, the performance of model recognition is improved accordingly by using cross-training compared with using softmax loss function in both pre-training and fine-turn stages. The test accuracy is exhibited in Table 2. Softmax Loss means that softmax is used in both face identification pre-training and the fine-tune of genetic disorders recognition. Through the experimental results, we can conclude that transfer learning based on our cross-training can make the model more suitable

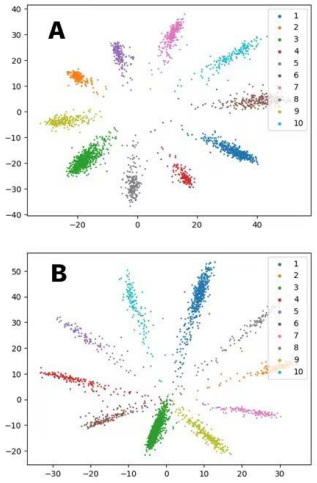

FIGURE 5. Visualization results of representation spatial distribution. The upper part is the distribution of softmax loss and the lower part is the distribution of cross-loss training.

TABLE 2. Cross-loss training and softmax loss comparison.

Training	Top-1 accuracy
Softmax Loss	90.6%
Cross-loss Training	93.5%

for the classification task of genetic disease attributes of small data sets.

B. IDENTIFICATION OF MULTIPLE GENETIC SYNDROMES

In order to evaluate the results, we use precision, recall as metrics, which are defined as below:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

TP or True Positive refers to the number of predicted results being the same with the annotation for a certain genetic syndrome. FP or False Positive is the number of samples of other types that are wrongly classified as samples of this type for a certain genetic syndrome. FN or False Negative is the number of samples of this type that are wrongly classified as samples of other types for a certain genetic syndrome. In addition, we calculate the Top-1 accuracy, which is defined as all correctly classified numbers divided by the total number of test sets.

In the research related to recognition of genetic syndromes, most of the work is only about identification of a single kind of disease or binary classifications, syndromic or normal, and it is rather difficult to acquire good accuracy in identification of multiple genetic disorders. Our framework is designed for accurate identification of multiple genetic disorders. In this work, we tested ten of the most common genetic disorders. Our framework application is not confined to the identification of these ten genetic diseases, but theoretically supports the accurate identification of hundreds of genetic diseases.

TABLE 3. Confusion matrix of test results (BioFace).

Label \ Prediction	0	1	2	3	4	5	6	7	8	9	10
0	40	0	0	0	0	0	0	0	0	0	0
1	1	44	0	1	0	0	0	0	1	0	0
2	0	0	19	0	0	0	0	0	0	0	0
3	0	1	0	66	0	0	0	0	0	0	0
4	0	0	0	0	13	0	0	0	0	0	1
5	0	0	0	1	0	14	0	0	0	1	0
6	0	1	0	0	1	0	23	1	0	0	0
7	0	0	0	1	0	0	1	25	0	1	0
8	0	1	0	1	0	0	0	0	19	0	0
9	0	0	1	0	0	1	0	0	0	20	0
10	0	0	0	0	0	0	0	0	0	1	15

TABLE 4. Confusion matrix of test results (DeePGestalt).

Label \ Prediction	0	1	2	3	4	5	6	7	8	9	10
0	23	0	0	0	1	0	0	1	1	0	0
1	9	32	1	3	0	1	1	3	1	2	0
2	0	0	17	0	0	0	0	0	0	0	0
3	1	3	0	55	0	3	0	0	1	0	0
4	3	2	0	2	9	0	0	0	3	0	1
5	0	0	0	1	0	8	1	2	0	0	0
6	0	0	0	0	1	1	19	2	1	2	1
7	1	3	1	3	0	3	0	18	1	1	1
8	2	2	0	0	3	0	0	1	6	2	1
9	1	0	0	1	0	0	3	1	2	14	0
10	0	5	0	2	0	0	2	0	0	0	18

TABLE 5. Results of test set.

No.	Syndromes	Precision (ours)	Recall (ours)	Precision (DeeP-Gestalt)	Recall (DeeP-Gestalt)
1	Angelman syndrome	93.6%	93.6%	68.1%	60.4%
2	Cornelia de Lange syndrome	95%	100.0%	89.5%	100.0%
3	Downs syndrome	94.3%	98.5%	82.1%	87.3%
4	Sotos syndrome	92.9%	92.9%	64.3%	45.0%
5	CHARGE syndrome	93.3%	87.5%	50.0%	66.7%
6	Noonan syndrome	95.8%	88.5%	73.1%	70.4%
7	DiGeorge syndrome	96.2%	89.3%	64.3%	56.3%
8	Williams-Beuren syndrome	95.0%	90.5%	37.5%	35.3%
9	Rubinstein-Taybi syndrome	87.0%	90.9%	66.7%	63.6%
10	Fragile x syndrome	93.8%	93.8%	81.8%	66.7%
0	Healthy	97.6%	100 %	57.5%	88.5%

To test our method, 10 % of the dataset is randomly reserved as the test set. The confusion matrix of the test results is displayed in Table 3. The test distribution of the model in the whole test set can be seen from this table. The precision and recall can be calculated from the confusion matrix and are shown in Table 5. For precision, the result of Rubinstein Taybi syndrome is 87.0%, and the results of other syndromes are higher than 92.9%. For recall, the result of CHARGE syndromes is 87.5%, Noonan is 88.5%, DiGeorge is 89.3%, and other syndromes are above 90%. In addition, our Top-1 accuracy rate of 10 diseases is 93.5%. The result is 94.3% if the healthy samples are included in the Top-1 accuracy calculation.

Among the research related to identification of various genetic diseases, DeePGestalt [26] is the most similar to our work. This research dose not disclose the image datasets, so we reproduced DeePGestalt to predict genetic diseases in the same training and test sets as BioFace. The confusion matrix of DeePGestalt is exhibited in Table 4, and the precision and recall is shown in Table 5. Besides, the Top-1 accuracy rate of DeePGestalt is only 71.0%, while the accuracy of BioFace is 93.5%. Our model can accurately identify the types of genetic diseases in clinical application.

V. CONCLUSION

In this study, we came up with an application framework for deep learning for recognition of genetic diseases based on small data sets. By adding SE block to the residual network Resnet64 and through cross-loss training, we can extract rich facial features of genetic diseases. This can improve inter-class and intra-class spacing during transfer learning of the network model, which can realize accurate identification based on small data sets. Our Top-1 accuracy for 10 genetic diseases is 93.5%. In the identification of multiple genetic diseases, our Top-1 accuracy is higher than that of related research work [26]. In future research, our framework will add more kinds of genetic diseases, and is expected to be applied to clinical auxiliary diagnosis of genetic diseases.

Data availability

To protect patient privacy, the facial images in this research are not publicly available.

Ethic statement

The authors affirm that all human research participants in clinical data provided informed consent, including the publication of the images in Figure1 and Figure3.

REFERENCES

[1] P. A. Baird, T. W. Anderson, H. B. Newcombe, and R. Lowry, "Genetic disorders in children and young adults: A population study," *Amer. J. Hum. Genet.*, vol. 42, no. 5, p. 677, 1988.

[2] T. Hart and P. Hart, "Genetic studies of craniofacial anomalies: Clinical implications and applications," *Orthodontics Craniofacial Res.*, vol. 12, no. 3, pp. 212–220, Aug. 2009.

[3] J. T. Pantel et al., "Advances in computer-assisted syndrome recognition by the example of inborn errors of metabolism," *J. Inherited Metabolic Disease, Off. J. Soc. Study Inborn Errors Metabolism*, vol. 41, no. 3, pp. 533–539, 2018.

[4] Q. Zhao, K. Okada, K. Rosenbaum, L. Kehoe, D. J. Zand, R. Sze, M. Summar, and M. G. Linguraru, "Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ICA," *Med. Image Anal.*, vol. 18, no. 5, pp. 699–710, Jul. 2014.

[5] J. E. Allanson, A. Bohring, H.-G. Dörr, A. Dufke, G. Gillessen-Kaesbach, D. Horn, R. König, C. P. Kratz, K. Kutsche, S. Pauli, S. Raskin, A. Rauch, A. Turner, D. Wieczorek, and M. Zenker, "The face of Noonan syndrome: Does phenotype predict genotype," *Amer. J. Med. Genet. A*, vol. 152, no. 8, pp. 1960–1966, Jul. 2010.

[6] E. Y. Gulec, Z. Ocak, S. Candan, E. Ataman, and C. Yarar, "Novel mutations in PTPN11 gene in two girls with Noonan syndrome phenotype," *Int. J. Cardiol.*, vol. 186, pp. 13–15, May 2015.

[7] M. Zenker, D. Horn, D. Wieczorek, J. Allanson, S. Pauli, I. Van der Burgt, H. G. Doerr, H. Gaspar, M. Hofbeck, G. Gillessen-Kaesbach, and A. Koch, "SOS1 is the second most common Noonan gene but plays no major role in cardio-facio-cutaneous syndrome," *J. Med. Genet.*, vol. 44, no. 10, pp. 651–656, 2007.

[8] H. Cave, A. Caye, N. Ghedira, Y. Capri, N. Pouvreau, N. Fillot, A. Trimouille, C. Vignal, O. Fenneteau, Y. Alembik, and J. L. Alessandri, "Mutations in RIT1 cause Noonan syndrome with possible juvenile myelomonocytic leukemia but are not involved in acute lymphoblastic leukemia," *Eur. J. Hum. Genet.*, vol. 24, no. 8, pp. 1124–1131, 2016.

[9] S. Rohatgi, D. Clark, A. D. Kline, L. G. Jackson, J. Pie, V. Siu, F. J. Ramos, I. D. Krantz, and M. A. Deardorff, "Facial diagnosis of mild and variant CdLS: Insights from a dysmorphologist survey," *Amer. J. Med. Genet. A*, vol. 152, no. 7, pp. 1641–1653, Jun. 2010.

[10] L. Basel-Vanagaite, L. Wolf, M. Orin, L. Larizza, C. Gervasini, I. D. Krantz, and M. A. Deardoff, "Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis," *Clin. Genet.*, vol. 89, no. 5, pp. 557–563, May 2016.

[11] L. M. Bird, W. Tan, and L. Wolf, "The role of computer-aided facial recognition technology in accelerating the identification of Angelman syndrome," in *Proc. 35th Annu. David W Smith Workshop*, 2014, pp. 1–16.

[12] A. K. Eurordis and F. Faurisson, *The Voice of 12,000 Patients. Experiences and Expectations of Rare Disease Patients on Diagnosis and Care in Europe*. Paris, France: Eurordis-Rare Diseases, 2009.

[13] S. Boehringer, T. Vollmar, C. Tasse, R. P. Wurtz, G. Gillessen-Kaesbach, B. Horsthemke, and D. Wieczorek, "Syndrome identification based on 2D analysis software," *Eur. J. Hum. Genet.*, vol. 14, no. 10, pp. 1082–1089, 2006.

[14] S. Boehringer, M. Guenther, S. Sinigerova, R. P. Wurtz, B. Horsthemke, and D. Wieczorek, "Automated syndrome detection in a set of clinical facial photographs," *Amer. J. Med. Genet. A*, vol. 155, no. 9, pp. 2161–2169, 2011.

[15] A. B. Dalal and S. R. Phadke, "Morphometric analysis of face in dysmorphology," *Comput. Methods Programs Biomed.*, vol. 85, no. 2, pp. 165–172, Feb. 2007.

[16] S. V. Vollmar and W. A. Kalender, "Reduction of dose to the female breast in thoracic CT: A comparison of standard-protocol, bismuth-shielded, partial and tube-current-modulated CT examinations," *Eur. Radiol.*, vol. 18, no. 8, pp. 1674–1682, Aug. 2008.

[17] G. Baynam, M. Walters, P. Claes, S. Kung, P. LeSouef, H. Dawkins, D. Gillett, and J. Goldblatt, "The facial evolution: Looking backward and moving forward," *Hum. Mutation*, vol. 34, no. 1, pp. 14–22, Jan. 2013.

[18] S. Saraydemir, N. Taspinar, O. Erogul, H. Kayserili, and N. Dinçkan, "Down syndrome diagnosis based on Gabor wavelet transform," *J. Med. Syst.*, vol. 36, no. 5, pp. 3205–3213, Oct. 2012.

[19] K. Burçin and N. V. Vasif, "Down syndrome recognition using local binary patterns and statistical evaluation of the system," *Exp. Syst. Appl.*, vol. 38, no. 7, pp. 8690–8695, Jul. 2011.

[20] Q. Zhao, N. Werghi, K. Okada, K. Rosenbaum, M. Summar, and M. G. Linguraru, "Ensemble learning for the detection of facial dysmorphology," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 754–757.

[21] S. Hadj-Rabia, H. Schneider, E. Navarro, O. Klein, N. Kirby, K. Huttner, L. Wolf, M. Orin, S. Wohlfart, C. Bodemer, and D. K. Grange, "Automatic recognition of the XLHED phenotype from facial images," *Amer. J. Med. Genet. A*, vol. 173, no. 9, pp. 2408–2414, Sep. 2017.

[22] K. Kuru, M. Niranjan, Y. Tunca, E. Osvank, and T. Azim, "Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics," *Artif. Intell. Med.*, vol. 62, no. 2, pp. 105–118, Oct. 2014.

[23] H. S. Loos, D. Wieczorek, R. P. Wurtz, C. V. D. Malsburg, and B. Horsthemke, "Computer-based recognition of dysmorphic faces," *Eur. J. Hum. Genet.*, vol. 11, no. 8, pp. 555–560, Aug. 2003.

[24] P. Shukla, T. Gupta, A. Saini, P. Singh, and R. Balasubramanian, "A deep learning frame-work for recognizing developmental disorders," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 705–714.

[25] Y. Gurovich, Y. Hanani, O. Bar, N. Fleischer, D. Gelbman, L. Basel-Salmon, P. Krawitz, S. B. Kamphausen, M. Zenker, L. M. Bird, and K. W. Gripp, "DeepGestalt–identifying rare genetic syndromes using deep learning," 2018, *arXiv:1801.07637*.

[26] Y. Gurovich, Y. Hanani, O. Bar, G. Nadav, N. Fleischer, D. Gelbman, L. Basel-Salmon, P. M. Krawitz, S. B. Kamphausen, M. Zenker, L. M. Bird, and K. W. Gripp, "Identifying facial phenotypes of genetic disorders using deep learning," *Nature Med.*, vol. 25, pp. 60–64, Jan. 2019.

[27] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

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

[29] X. Yin and X. Liu, "Multi-task convolutional neural network for pose-invariant face recognition," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 964–975, Feb. 2018.
[30] S. Garner, "Learning to program from scratch," in *Proc. 9th IEEE Int. Conf. Adv. Learn. Technol.*, vol. 25, Jul. 2009, pp. 451–452.
[31] 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.
[32] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–40, Dec. 2016.
[33] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 212–220.
[34] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 38, no. 1, pp. 149–161, Jan. 2008.
[35] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in *Proc. Workshop Faces Real-Life Images, Detection, Alignment, Recognit.*, 2008, pp. 1276–1284.

JIANFENG WANG received the M.Rs. degree from the University of Liverpool, U.K. He is currently the Director at the AI Department, Suzhou Life Intelligence Industry Research Institute. He has led several major AI medical projects in Suzhou, China. His research interests include machine learning, deep learning, and algorithm design.

BO LIANG received the Ph.D. degree in biology from Shanghai Jiao Tong University, China. He is currently an Associate Professor with the School of Biology and Food Engineering, Changshu Institute of Technology. He is also an Adjunct Associate Professor at Suzhou University and a Researcher at the National Assisted Reproduction and Eugenics Engineering Technology Research Center, China. Among several major projects he led, one of his products has completed the testing services of more than 30000 samples (including clinical samples), with a comprehensive clinical pregnancy rate of more than 70%, the clinical pregnancy rate of some hospitals even as high as 80%, and the abortion rate has been reduced from 33.5% to 6.9%, becoming an important promoter and a leader in the application of PGT clinical norms in China.

LIJUN ZHAO received the master's degree in biomedical engineering from Southern Medical University, China. She is currently an Image Processing Engineer at the Suzhou Life Intelligence Industry Research Institute. She has rich experience in biological image processing and has participated in many medical bioengineering projects.

YUANFANG CHEN (Member, IEEE) received the M.S. and first Ph.D. degrees from the Dalian University of Technology, China, and the second Ph.D. degree from University Pierre and Marie Curie, France. From 2009 to 2010, she was an Assistant Researcher with the Illinois Institute of Technology, USA, with Prof. Xiangyang Li. She is currently a Professor with Hangzhou Dianzi University, China. She is also a Researcher with Shanghai Jiaotong University, China. Her research interests include artificial intelligence of things, artificial intelligence security, and algorithm design.

WEN FU received the bachelor's degree in computer science from Hunan Normal University, China. He is currently a Machine Learning Engineer at Suzhou Life Intelligence Industry Research Institute. He has participated in many AI projects. His research interests include machine learning and image processing.

PEIJI YU received the M.B.A. degree from Hong Kong Metropolitan University. He is currently the Program Development and Design Director at the Suzhou Life Intelligence Industry Research Institute. His research interests include network engineering and algorithm design.

HONGBING CHEN received the bachelor's degree in biology from Chuzhou University and the master's degree in genetics from Suzhou University, China. He is currently the Genetic Department Manager at the Suzhou Life Intelligence Industry Research Institute. His research interests include child genetics and genetic variation.

HONGYING WANG is the Director and the Deputy Chief Technician at the Laboratory Department of Wujiang District Children's Hospital (Wujiang District, Children's Hospital of Soochow University). Her research interest includes medical genetics. She is a senior genetic consultant. She is a member of Precision Medicine Professional Committee of China Maternal and Child Health Association, the Rare Diseases Professional Committee of Jiangsu Research Hospital Association, the Gene Diagnosis and Standardization Branch of Asia–Pacific Medical Biological immunology Association, and the Reproductive Health Professional Committee of Suzhou Preventive Medicine Association.

She has published more than ten SCI and Chinese level papers as the first author, presided over two people's livelihood science and technology projects in Suzhou, won two new technology introduction awards in Suzhou, won one second prize of Medical Science and Technology Award of Suzhou Medical Association, won one national patent, and found nine cases of karyotype first reported in the world.

GUOJIE XIE is the Head of External Affairs at the Provincial Key Laboratory. He has long been engaged in digital transformation, digital government, and industrial internet field research, accumulating rich practice in project delivery, solution customization, and intellectual property protection. Moreover, his extensive experience in undertaking major projects in government and central state-owned enterprises allows him to lead the construction of several brand-new field products.

TING WU received the Ph.D. degree from Shandong University, China. He is currently serving as a Full Professor at the Hangzhou Innovation Institute, Beihang University, China. He has actively participated in several national funded projects as a principal investigator and a senior researcher. His research interest includes information security.

MUHAMMAD ALAM (Senior Member, IEEE) received the Ph.D. degree in computer science from the University of Aveiro, Portugal, in 2014. In 2009, he joined the Instituto de Telecomunicações-Aveiro (Portugal) as a Senior Researcher. In 2017, he joined Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, as an Assistant Professor of computer science. He is currently working as a Senior Lecturer in smart cities and secure systems (the IoT) at London South Bank University, U.K. His research interests include vehicle communication, the Internet of Things Security, and algorithm design.

HAITAO LV is a second-class Chief Physician, a Professor, an Academic Doctoral Tutor, the Deputy Dean at the Children's Hospital of Soochow University, and a leader of the diagnosis and treatment of congenital cardiovascular diseases, a key discipline in Jiangsu. He was rated as the key medical talent of ''rejuvenating health through science and education'' in Jiangsu Province, the 333 project in Jiangsu Province, the six talent peaks in Jiangsu Province, and the key health talent in Suzhou. In 2020, he was rated as the leading health talent in Suzhou. His main research direction: Kawasaki disease and vascular dysfunction; Clinical: interventional treatment of congenital heart disease in children. He presided over four National Nature Fund projects and three provincial and ministerial projects.

LIN HE is a Genetic Biologist in China. He is an Academician at the Chinese Academy of Sciences and the Third World Academy of Sciences. He is currently a Professor and the Doctoral Supervisor at the School of Life Science and Technology, Shanghai Jiao Tong University. His research interest includes biogenetics.