Research on Image Algorithm for Face Recognition Based on Deep Learning

Qiang Wu
School of Information Technology, Jiangsu Open University, Nanjing, 210036, China

***Abstract*—As people's requirements for applications are getting higher and higher, the recognition of facial features has been paid more and more attention. The current facial feature recognition algorithm not only takes a long time, but also has problems such as large system resource consumption and long running time in practical applications. Based on this, the research proposes a multi-task face recognition algorithm by combining multi-task deep learning on the basis of convolutional neural network, and analyzes its performance in four dimensions of face identity, age, gender, and fatigue state. The experimental results show that the multi-task face recognition algorithm model obtained through layer-by-layer progression takes less time than other models and can complete more tasks in the same training time. At the same time, comparing the best model M44 with other algorithms in four dimensions, it is found that the Mean Absolute Error lowest is 3.53, and the highest Accuracy value is 98.3%. On the whole, the multi-task face recognition algorithm proposed in the study can recognize facial features efficiently and quickly. At the same time, its training time is short, the calculation speed is fast, and the recognition accuracy is much higher than other algorithms. It is applied to intelligent driving behavior. Analysis, intelligent clothing navigation and other aspects have strong practical significance.**

Keywords—Multi task deep learning; face recognition; convolution neural network; multi task; dimension

I. Introduction

The rapid development of artificial intelligence has made it the core driving force of industrial transformation. It is also widely used in the fields of speech recognition and image recognition, and has gradually replaced human resources [1-3]. As an important pillar in the field of image recognition, face recognition has been widely used in business, identity authentication and other fields, and has gradually received attention. In addition, face recognition is no longer limited to a single task, and the needs of multi-task face recognition. It is also constantly improving [4-5]. In this context, many domestic and foreign scholars have conducted in-depth research on it. Khan AA et al. built a face recognition image model on the basis of neural network and integrated genetic algorithm and principal component analysis, which provided help for face matching in forensic investigation [6]. Srivastava S et al. proposed a new method of biometric authentication face recognition based on artificial neural network, which effectively reduces the error rate of face recognition [7]. Karanwal overcomes the problem of low image recognition rate in local binary patterns by proposing descriptors [8]. However, although the current feature matching algorithms using similarity measures in single task face recognition are simple and fast, they are difficult to robustly determine the threshold size. Although feature matching algorithms using feature subspaces can map intra class differences to subspaces for compression, their noise will also be mapped to subspaces and easily amplified. Although feature matching algorithms using statistical models have good robustness and identification, they are prone to data overfitting, and many deep learning methods also have shortcomings such as low accuracy and high hardware requirements. The methods in multitasking facial recognition have drawbacks such as multiple parameters, slow running speed, and high resource consumption. Based on this, the research proposes a multi-task deep learning algorithm by integrating deep learning on the basis of Convolutional Neural Network (CNN). The purpose is to effectively improve the accuracy of face recognition through a new face recognition algorithm. At the same time, it reduces its recognition time and provides effective suggestions for artificial intelligence face recognition.

The research is divided into six sections. Section I is the introduction, Section II is about related work. Research on Face Recognition image algorithm is mentioned in Section III, results and discussion in Section IV and Section VI concludes the paper.

II. Related Work

With the development of artificial intelligence, face recognition technology is gradually applied to all walks of life, and it is also a very important link in biometrics. Compared with fingerprint recognition, iris recognition and other recognition technologies, it can better meet the needs of users, and the recognition accuracy and recognition speed are very fast [9]. The current face recognition technology has many defects in practical application. How to improve face recognition has become the focus of current research. Based on this, scholars at home and abroad of Everbright have conducted in-depth research on it. Chen et al. proposed a new Collaborative Representation Based Fuzzy Discriminant Analysis (CRFDA) algorithm based on collaborative representation, and effectively extracted the relevant features of the image through dimension reduction, which effectively improved the feature extraction standard and improved the face recognition accuracy [10]. Alami et al. proposed a new model of quaternion discrete orthogonal matrix neural network, thereby reducing the time consumption of the model in face recognition training, and further effectively improving the accuracy of color face recognition [11]. Tripathi et al. effectively improved the accuracy of face age recognition by using descriptors in local gradient relationship patterns [12]. Singh et al. put forward a new and robust description of color texture, thus emphasizing the advantages of related color models in identifying color components, so as to help improve the accuracy of color face recognition [13]. Soni et al. proposed a hybrid optimization algorithm based on bird search to remove image noise and improve face recognition accuracy [14]. Sharma et al. designed a face recognition system based on field programmable gate array to improve the security of face recognition [15], aiming at the problem of imperfect user data protection at present. Long et al. have effectively improved the recognition performance of face recognition by using singular value decomposition (SVD) to solve the problem of poor face recognition technology [16].

In addition, Sharma et al. used the depth learning method to achieve multi-modal determination of personal recognition and improve the robustness of face identification [17]. On the basis of deep learning, Han et al. proposed a new personalized convolution method, which significantly improved the efficiency of face recognition [18]. Srivastava et al. used deep learning to build a hybrid model to improve the accuracy of personal face recognition involving violence [19]. Zhao et al. aimed at the problem that face recognition takes too long; they reduced the time consumption in face recognition by building an unsupervised deep learning network [20]. Vedantham has constructed a classifier by using depth learning to improve the accuracy of facial expression recognition [21]. Deshmukh et al., aiming at the low efficiency of biometric feature extraction in face recognition, built a multi-mode biological learning system using deep learning networks to improve the efficiency of biometric feature extraction [22]. Silva et al. have effectively improved the accuracy of individual identification of wild Asian elephants by using deep learning methods, thus providing help for the protection of wild Asian elephants [23].

From the research of scholars at home and abroad, the current face recognition technology has problems such as long model training time, high resource consumption in practical applications, and deep learning has a good effect on collaborative operation. Therefore, it is of great significance to study the algorithm proposed by combining the depth learning algorithm with human face recognition in practical application. In addition, the face recognition algorithm proposed in the study creatively uses the relevance of multi task reflection to achieve the multi task recognition requirements of the algorithm, which effectively changes the simplification of traditional face recognition.

III. Research on Face Recognition Image Algorithm based on Multi-Task Deep Learning

A. *Analysis of Convolutional Neural Networks in Multi-Task Deep Learning*

Aiming at the problems of slow training time and large system resource consumption for the current face attribute recognition task, the research conducted a related analysis on the face recognition image algorithm on the basis of multi-task deep learning. Deep learning includes two parts: deep and learning. Specifically, deep learning is a deep network model and has a suitable training learning method. The current fast-developing deep learning method is CNN, and for multi-task depth in terms of learning, convolutional neural networks are the basis for face recognition [24]. CNN uses a large amount of data for learning, and performs multiple multi-level convolution and non-linear mapping on it, so as to achieve the purpose of feature extraction or classification. It has the advantages of high accuracy and strong robustness. In other words, the working principle of CNN is to simulate the recognition process of the human brain [25]. Strictly speaking, CNN imitates the human brain to process information in more detail, and its principle structure is shown in Fig. 1.

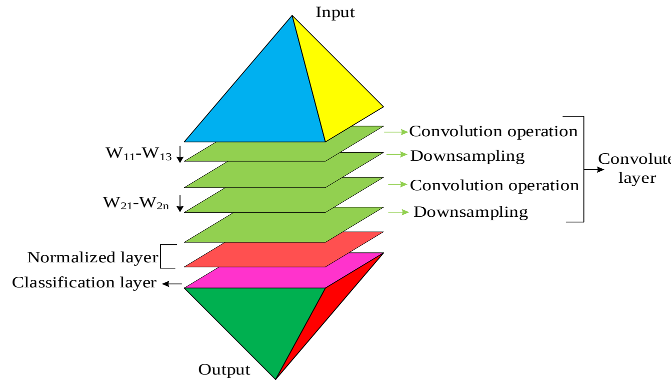


Fig. 1. Principle and structure of CNN.

It can be clearly found from Fig. 1 that the principle structure of CNN includes five levels of input, convolution, normalization, classification and output, which are progressive until the desired information is output. Specifically, after inputting the initial training data, CNN uses the reverse transfer algorithm to perform repeated learning, revise the parameters of different neurons, and finally obtain a convergent neural network model. On this basis, by inputting the image into the trained CNN model, a large number of abstract expressions ranging from low-level intuition to high-level can be obtained, and then through linear or nonlinear combination, accurate representation can be obtained to complete classification and feature extraction. Therefore, the basic structure of CNN can be divided into two parts, namely the feature extraction part and the fully connected part, and the more important ones are the convolution layer, the activation function layer and the fully connected layer (Softmax). At the signal level, the image of the convolution layer can be regarded as a visual signal, and its convolution operation can be understood as a process of filtering the signal, and its related calculation expression is shown in Eq. (1).

$$I(x, y'')*W(x, y'') = \sum_{s=-w}^{w}\sum_{t=-h}^{h}(s,t)\,I\left(x-s, y''-t\right) \tag{1}$$

In Eq. (1), I represents the pixel value; x represents the row; y'' represents the column; represents the convolution kernel; W the t parameter w represents the width of the convolution kernel; h represents s the height of the convolution kernel; At the matrix level, the convolution operation of the image is to use the convolution core to perform continuous convolution on the perceptual region of the image, and finally use it as the feature value of the region to obtain the feature map. The specific operation process is shown in Fig. 2.

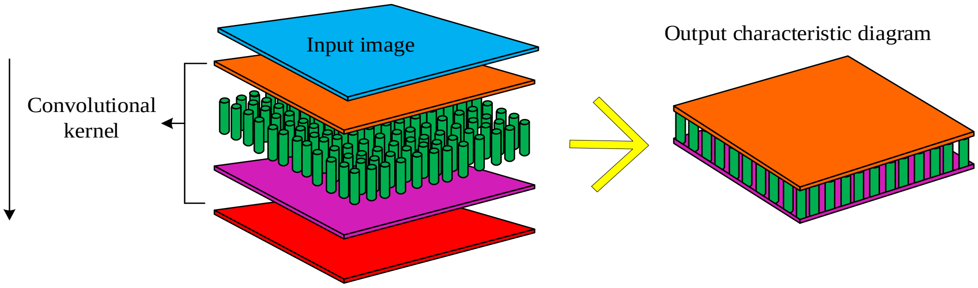


Fig. 2. Operation process of convolution.

It can be seen from Fig. 2 that the operation process of convolution is relatively simple. First, input the corresponding image into the convolution kernel, and then output the feature map after the operation. For the convolution layer, the convolution layer is usually composed of multiple convolution cores to better extract the image features. In CNN, the convolution layer uses a variety of convolution to check the local receiving area of the image, and obtains the feature maps of different convolution cores through the sliding convolution operation. It can be seen that CNN can effectively detect the features of the face in face recognition by recognizing the local feature map, thus improving the accuracy of face recognition. In this process, the corresponding calculation formula is shown in Eq. (2).

$$h_{ij}{}^{k} = f\left(\left(W_k * x'\right) + b_k\right) \tag{2}$$

In expression (2), x represents the convolution result, where i and j represents the position in the matrix, specifically represented as rows and columns, k represents the number of convolution kernels; f represents the activation function; x' represents the input image; b represents the convolution The corresponding bias of the kernel. The activation function layer is also known as the normalization layer. Its purpose is to use the relevant activation function to add corresponding nonlinear correlation factors, so as to ensure that the features are unique and invariant, thereby improving the overall expression ability of the model. The linear activation function is the Linear Rectification Function (RELU) [26]. Its related expression formula is shown in Eq. (3).

$$f(z) = \max(z, 0) \tag{3}$$

In Eq. (3), it $\max$ represents a comparison function that takes a larger value; z it represents the size of the input. However, in the actual training process of CNN, if the backpropagation presents a large gradient, if a RELU neuron flows through it, the neuron will always be in a "dead" state, which will as a result, the parameters of CNN are no longer continuously updated, thus consuming too many resources and related operations. The activation function selected for the study can be understood as a layer of network in the neural network, namely the Maxout activation function, which is relative to the sparse feature of the RELU activation function. It is more compact, and feature selection and dimensionality reduction can also be performed at the same time, and the calculation formula of its related neurons is shown in formula (4).

$$h_i(x) = \max_{j \in [1,k']} z'_{ij} \tag{4}$$

In Eq. (4), it $h_i(x)$ represents i the output of the neuron of the Maxout layer; it represents k' the number of input layers connected to the Maxout layer; j it represents the j input layer; z' it represents the input layer connected to the neurons of the Maxout layer. value, and its evaluation formula is shown in Eq. (5).

$$z'_{ij} = W_{\cdots ij} * x^T + b_{ij} \tag{5}$$

In Eq. (5), it $W_{\cdots ij}$ represents a certain convolution kernel among multiple convolution kernels; x^T it represents the input. For the Maxout layer, the dimensions of each CNN layer input must be consistent. On this basis, the maximum value is selected as the value of the same coordinate of the output layer after each input of the same coordinate is compared. This can be seen. The Maxout function is a piecewise function, and its input gradient formula is shown in Eq. (6).

$$\frac{\partial hi(x)}{\partial z'ij} = \begin{cases} 1 \quad if\ zij > ziq \\ When\ q \neq j\ and\ q \in [1,k] \\ 0 \quad otherwise \end{cases} \tag{6}$$

In Eq. (6), q represents the input value. It can be clearly seen from the Eq. (6) that the slope exhibited by each segment of the Maxout function is determined by the output value, so it presents a dynamic change and does not completely fix a certain value. In addition, the Softmax layer is essentially a nonlinear classifier. In deep learning, it is often used as the output layer of CNN to output a probability vector. The output probability correlation calculation formula of the Softmax layer is shown in Eq. (7).

$$p\left(y' = i \middle| x', \theta\right) = \frac{e^{\theta_j^T x}}{\sum_{j=1}^{k} e^{\theta_j^T x}} \tag{7}$$

In the output probability calculation expression (7), θ represents the bias term; y' represents the output category, T represents transposition. On the whole, the Softmax layer is to map the multiple outputs after the original convolution to the values 0 to 1 with the Sofmax function, and the accumulation of these values is 1, which satisfies the characteristics of probability. On this basis, it selects the final output. As a result, the largest prediction result can be selected to achieve multi-classification tasks, and in actual face recognition, multi-task face recognition can be effectively achieved. In face recognition technology, there are three very important technologies, which are the recognition of face age, gender, and fatigue status, which together lay the foundation for face recognition technology. Specifically, face recognition technology includes face-related feature extraction and comparison technology.

B. Research on Multi-Task Face Recognition Algorithm based on Convolutional Neural Network

CNN is a relatively common deep learning method, which is also the basis of multi-task face recognition algorithm. To understand multi-task face recognition algorithm, you need to understand the principle of multi-task learning. The current deep learning methods for multitasking mainly have the following two characteristics: one is structurally speaking, shallow parameter sharing between tasks; the other is that, in terms of impact, the common data characteristics hidden between different tasks are excavated. This requires that a model can handle multiple tasks at the same time, and also requires multiple tasks to work together to enhance the generalization ability. The overall structure is shown in Fig. 3.

It can be seen from the overall structure of multi task learning in Fig. 3 that its specific content includes model input, shared parameters and relevant parameters of specific tasks. By inputting relevant values in model input, shared parameters are obtained, and relevant parameters of specific tasks are output in subsequent output. An important premise of multi task learning is that there is a certain correlation between multiple tasks, so learning efficiency can be improved through the correlation between tasks. Although there is no clear definition, when looking for related tasks, a basic assumption is that the characteristics concerned by each task are interrelated, or the learning process is beneficial in different tasks. In the research, there is some correlation between facial features and features. Through the association training of multiple tasks, the accuracy of multiple tasks can be effectively improved. In the actual face recognition, the task is no single, and the improvement of the accuracy of multi task recognition can also effectively improve the overall recognition accuracy.

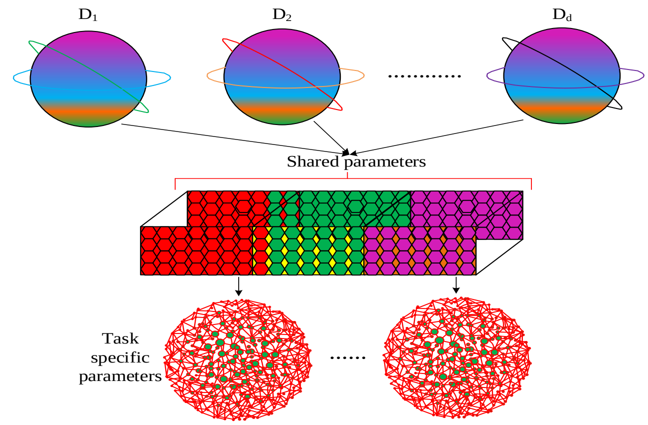


Fig. 3. The overall structure of multi task deep learning.

The traditional face recognition algorithm has the problem of low recognition task accuracy. Aiming at this problem, a multi-task deep learning face recognition algorithm is proposed. This algorithm uses the relatively strong correlation of face identity, age, and gender and fatigue recognition. To carry out research, in order to build a multi-task face recognition model in a short time, and to complete the relevant recognition tasks under the premise of maintaining high precision and high speed. High accuracy depends on the relevance of multiple learning tasks, and high-speed operation requires optimization of model parameters. Therefore, the research integrates CNN to construct a CNN network for multi-task face recognition. Its structure is shown in Fig. 4.

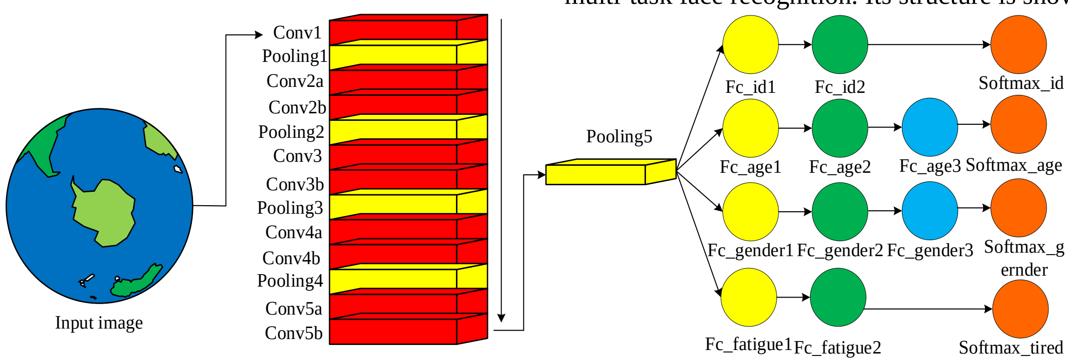


Fig. 4. Multi task face recognition network structure.

As can be seen from Fig. 4, the overall network structure of CNN has a total of nine convolutional layers, the pooling layer uses the maximum pooling function, the activation function selects the Maxout function, and all tasks in the training select the Softmax function to perform related classification training tasks for face recognition. It is worth noting that before constructing the actual network training, it is necessary to set the objective function and the training algorithm. This is because the work related to face recognition is closely related to the data, including identity, gender, age, etc. It is very difficult to generate a multi-label data set with information such as fatigue, fatigue, etc. Therefore, after setting the two in detail, the multi-label data set can be trained synchronously, and the multi-task training on the basis can be obtained. Same result. In the proposed multi-task deep learning face recognition algorithm, a brief description of the relevant forward propagation expression is shown in Eq. (8).

$$output_{k'} = f\left(W_{k'}X + b_{k'}\right) \tag{8}$$

In the expression Eq. (8), it k' represents the first k' face recognition task; $W_{k'}$ it represents the convolution related parameters. In contrast, the formula expression of backpropagation is shown in Eq. (9).

$$W_{k'} = W_{k'} - \eta_{k'}\frac{\partial L_{k'}}{\partial W_{k'}} \tag{9}$$

In the Eq. (9) of backpropagation, η it represents the learning rate; L it represents the loss of identifying characters. It is worth noting that the overall back-propagation formula of the recognition task can be divided into two parts: the independent parameter part of a single task, whose back-propagation formula is consistent with Equation (9), and the back-propagation formula of the common parameter segment of multiple tasks is as in Eq. (10).

$$W_{share} = W_{share} - \sum_{k'=1}^{M}\eta_{k'}\frac{T_k \partial L_{k'}}{\partial W_{share}} \tag{10}$$

In Eq. (10), it W_{share} represents the shared parameters between multiple tasks; M it represents the total number of face recognition tasks; and $T_{k'}$ it represents the weight of the task. It can be seen from Eq. (10) that the training strategy given by the study is consistent with the training of multiple tasks at the same time, except that there is a sequence in the training time of each task. Therefore, for different recognition tasks, the specific training strategy not only grasps the three important foundations of face recognition, but also introduces identity recognition. In face recognition, the forward propagation calculation formula of the deep learning model CNN structure is shown in Eq. (11).

$$o = g\left(f_n\left(W_n \cdots f_2\left(f_1\left(W_1 X_0 + b_1\right) + b_2\right) \cdots + b_n\right)\right) \tag{11}$$

In Eq. (11), o represents the forward propagation training value; W_n represents the n weight of b_n the first n layer; represents the displacement bias of the first layer; and $g()$ represents the final classification function. The ultimate goal of network training is to obtain the minimized loss function value loss, so the Softmax-loss function is selected to calculate the loss rate, and its specific formula is shown in Eq. (12).

$$Loss\left(y, o_y\right) = -\log\left(o_y\right) = -\log\left(\frac{e^{z_y}}{\sum_{j=1}^{m} e^{z_j}}\right) \tag{12}$$

In Eq. (12), it y represents the real label of the face image; o_y it represents the probability value of the position in the output probability vector when Softmax outputs y. In addition, in face age recognition, a unique loss function is selected to speed up the training process, and its specific expression is shown in Eq. (13).

$$G\_Loss\left(y, o_y\right) = \mu G(y, o_y) + \left(1-\mu\right) Loss\left(y, o_y\right) \tag{13}$$

In Eq. (13), it μ represents the weight of loss; $G(y, o_y)$ it represents the special loss function selected by the research, and its calculation formula is shown in Eq. (14).

$$G(y, o_y) = \left(1 - \exp\left(-\frac{o_y - y}{2\delta^2}\right)\right) \tag{14}$$

In Eq. (14), it δ represents the standard deviation of the age distribution in the training data set, and its calculation formula is shown in Eq. (15).

$$\delta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mu'\right)^2} \tag{15}$$

In Eq. (15), it N represents the total number of pictures of human faces; x_i represents i the age of the th face photo; μ' represents the mean value.

IV. Application and Performance Analysis of Face Recognition Image Algorithm based on Multi-Task Deep Learning

In order to verify the performance of the multi task deep learning face recognition algorithm proposed by the research, the research selects the face recognition training data set (CASIA Webface), human age recognition data set (IMDB-WIKI 500k+), face gender recognition data set (Celeba) and face fatigue recognition data set (bus driver face fatigue data set) before the experiment, and the test sets are respectively IJB-A, FG-NET, Celeba and human eye closure data set (CEW). It is worth noting that the data sets selected for the study are all found on the Internet, part of which are internet data (such as Wikipedia) and part of which are intranet data. In addition, in the experiment, the deep learner Aki chose Caffe, and the processor chose Intel (R) Core (TM) 7-7700 CPU @ 3.6GHz, with 8-core 8GB of memory; Select 64 bit Ubuntu 16 04 for the operating system; Choose NVIDIA GeForce GTX1070 graphics card with 8G memory; Select the Python 27 environment and Anaconda related scientific computing library for the development environment and tools. At the same time, before the experiment, the dataset was preprocessed for possible missing faces or imbalanced data, mainly through three steps: affine transformation, random cropping, and data balancing. The specific content of the dataset is shown in Table I.

It can be clearly seen from Table I that in the training data set selected for the experiment, the CASIA-Webface data set contains 494,414 identity photos of 10,575 people, and the training data set IJB-A data set contains the face pictures of 500 people, including 5,396 still images, 20,412 frames of video images; IMDB-WIKI 500k+ dataset contains 524,230 face images of 82,612 people, of which 460,723 are from the Internet Movie Database (IMDB) and 62,328 are from WiKi , its test set contains 1,002 photos of 1,297,028 people; the Celeba data set divides 202,599 face images of 10,177 people into training and test sets in a ratio of 9:1; the bus driver's face fatigue data set contains There are 47,500 photos of 10,100 individuals, including 10,000 photos of fatigued state, 30,000 photos of normal state, and 7,500 photos of standing wearing sunglasses. The test set CEW contains 2,423 photos of 40 individuals, including 1,192 photos of closed eyes, and 1,231 photos of closed eyes. Open-eye photos, closed-eye photos are obtained from a web crawler, and open-eye photos are obtained from the face database (Labeled Faces in the Wild, LFW). According to the four parts of the dataset, the research first analyzes the face recognition task, and its experimental results on the training set LFW are shown in Fig. 5.

TABLE I. SPECIFIC CONTENTS OF TRAINING DATA SET AND TEST DATA SET SELECTED BY THE EXPERIMENT

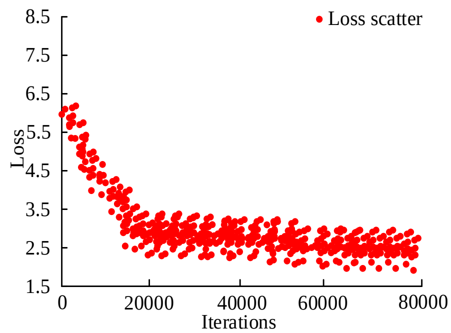
Training Set	**Number of people**	**Number of photos**			**Test Set**	**Number of people**	**Number of photos**	
CASIA-Webface	10575	494414			IJB-A	500	5396 (static state)	20412 (Video image)
IMDB-WIKI 500k+	82612	460723 (IMDB)	62328 (Wiki)		FG-NET	1297028	1002	
Celeba	10177	182339			Celeba	10177	20260	
Bus drivers face fatigue	10100	10000 (Fatigue state)	30000 (Normal state)	7500 (Wear sunglasses)	CEW	40	1192 (Closed eye photograph)	1231 (Eye opening photo)

In Fig. 5, Fig. 5(a) is the loss function curve obtained by using the loss function. Using this loss function curve, the model to be verified for face identification is divided into five models, which are respectively used M_{11} to M_{15} represent, and through the analysis of the five models The model's rate of change (Rate of Change, ROC) and the area under it (Area Under Curve, AUC) are compared, and Fig. 5(b) is obtained. It can be seen from Fig. 4 that the AUC value is the highest M_{13}, and the AUC value is 0.982 at this time. Therefore, the study uses it as the subsequent face age recognition pre-training model, and the relevant formula is used to obtain the threshold value of 0.35, and the accuracy rate is as high as 92 %. The experimental results show that the preprocessing model can effectively extract the identity features of the face. In the face gender experiment, the research uses the face identification with the highest AUC value M_{13}, trained for four hours in the same equipment environment, and obtained the corresponding loss curve, and also obtained five models according to the obtained loss function curve, respectively. Used M_{21} to M_{25} represent, and the corresponding face gender experiment is on the model with the best performance in the face age recognition task, after 10 minutes of training, the obtained 5 models are respectively used M_{31} to M_{35} represent, the two are respectively in LG- The experimental results on the NET and Celeba training sets are shown in Fig. 6.

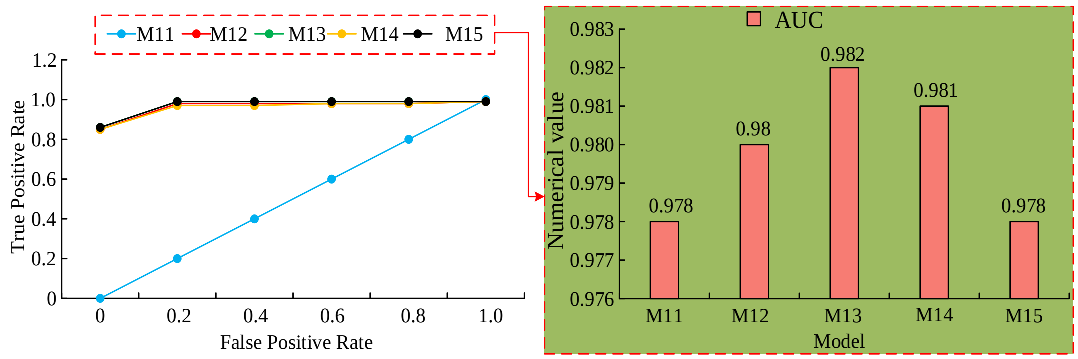
Fig. 6 (a) is the loss convergence scatter and Mean Absolute Error (MAE) obtained in the face age recognition training process, and Fig. 6(b) is the loss obtained in the face gender recognition training process Convergence curve and face recognition accuracy value (Accuracy, Acc). It can be seen from Fig. 6(a) that the MAE minimum value appears at the model M_{22}, and the algorithm has the highest accuracy at this time, so it can be used as the pre-training model for face gender recognition in Fig. 6(b). It can be seen from Fig. 6(b) that the Acc value M_{34} is the highest at the model, which is 0.983. The experimental results show that with the continuous optimization of the model accuracy, the training time for the algorithm to obtain loss convergence is decreasing, indicating that the error of the algorithm is decreasing, and it is continuously improving the accuracy of the model until it is optimal. Finally, the loss convergence curve was obtained after only two minutes of training in human fatigue state recognition. The experimental results on the CEW dataset and the total convergence time of the four identification experiments are shown in Fig. 7.

It can be clearly seen from Fig. 7 that the recognition accuracy of face fatigue M_{44} appears on the model, which is 0.983. In addition, in the introduction of three deep learning models (Facenet), Face Recognition Based on Visual Geometry Group Network (Vggface) and Levi G's four best models in the research itself From the comparison, it can be found that on the basis of the algorithm proposed by the research, the four tasks can complete the continuous training of the optimal mode within 10 hours, which is much shorter than the training time of deep CNN networks such as Facenet and Vggface; Compared with the face feature recognition method proposed by Levi [27] et al., this method can only recognize the age and gender of the face, but the algorithm proposed in the study can complete the face recognition identity within the same training time. On the basis of gender, it can also realize the recognition of face age and fatigue state. The research results show that the multi-task training method can effectively reduce the completion time of training tasks.

In order to further verify the effectiveness of the proposed algorithm, the research M_{44} evaluates the performance of the final multi-task face recognition model. Chen algorithm, Webface algorithm and block cipher algorithm (Gosudarstvennyi Standard, GOST) algorithm are introduced again in the experiment. The experimental results are shown in Table II.

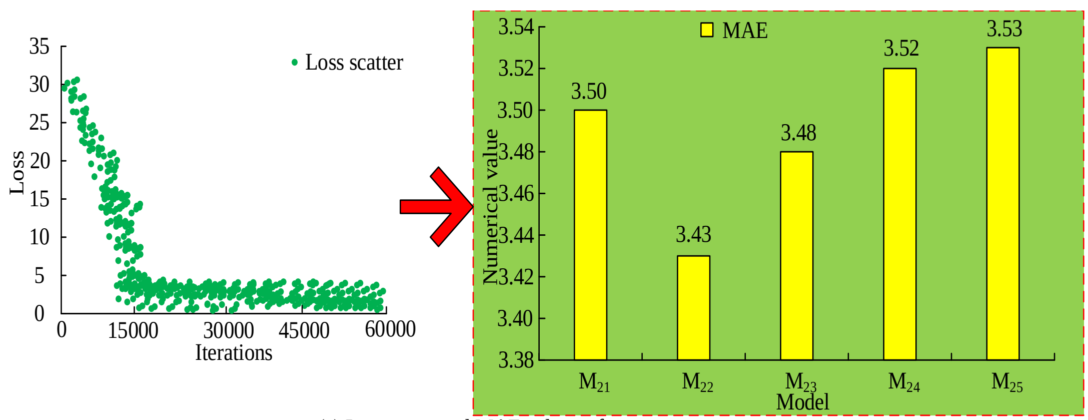


(a) Loss convergence graph in face recognition

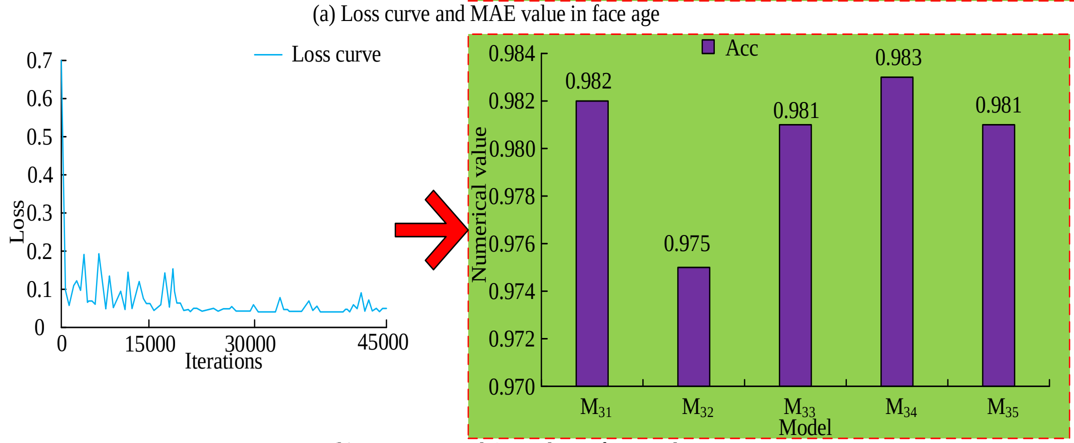


(b) ROC value change and AUC value of the model in 5

Fig. 5. Training process of face recognition.

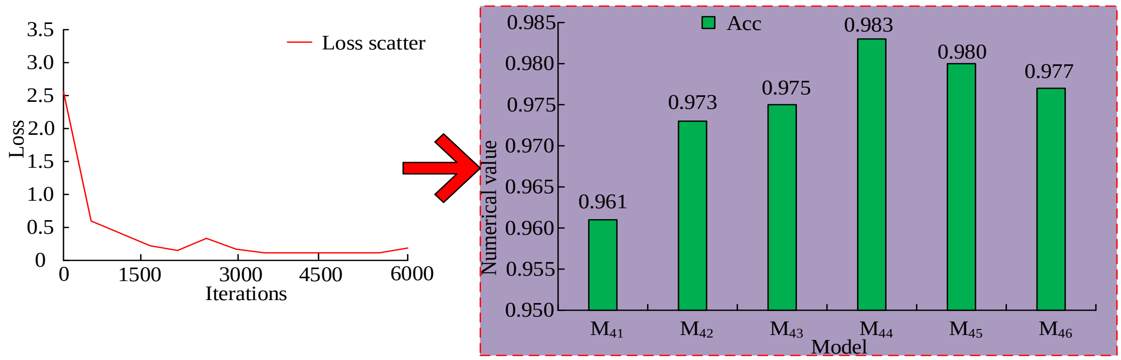


(a) Loss curve and MAE value in face age

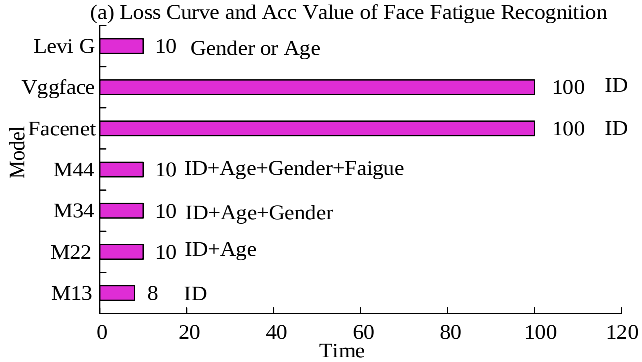


(b) Loss curve and Acc value in face gender

Fig. 6. Face age and gender training process.



(a) Loss Curve and Acc Value of Face Fatigue Recognition



(b) Convergence time of four kinds of identification to reach the best model and time of partial depth algorithm

Fig. 7. Experimental results on the CEW dataset and the time spent on four face recognition experiments and three depth algorithms.

TABLE II. PERFORMANCE COMPARISON OF MULTIPLE ALGORITHMS

Model	Feature extraction time	Number of parameters	Enter image dimensions	Accuracy (%)	FAR=0.1	FAR=0.01	FAR=0.001
$M44$	41ms	5555K	64×64×1	97.88 0	0.951	0.844	0.790
Chen	67ms	5004K	100×100×1	97.71 0	0.966	0.837	-
Webface	-	5011K	31×36×3	97.49 0	-	-	-
Facent	>558ms	140690K	224×224×3	99.62 0	0.950	0.834	0.750
Vggface	>528ms	134249K	224×224×3	97.27 0	0.936	0.804	0.603
GOST	-	-	-	-	0.620	0.400	0.190

As can be seen from Table II, in the LJB-A dataset with more complex actual scenes, M_{44} the model performance of the model is better than other models, and the accuracy reaches 97.880%. In addition, when the False Accept Rate (FAR) value is 0.01, the True Accept Rate (TAR) value at this time is 0.844, indicating that the M_{44} model has higher robustness in actual complex situations, which can distinguish face identities well. At the same time, the research compares the performance and accuracy of different algorithms (methods) in the recognition of face age, gender, and fatigue state. Among them, multi region convolution neural network (MR-CNN) and two ranking based methods are introduced into face age recognition, namely OH ranker and OR-CNN; panda recognition (PANDA-1), hyperspectral face recognition (Hyperspectral Face Recognition, Hyperface), LNet+ANet, based on deep learning are introduced in face gender recognition Face detection and alignment based on depth learning (MT) and walk recognition (Walk); in face fatigue recognition, Support Vector Machines (SVM) are added, and linear inverse Projection algorithm (Local Binary Patterns, LBP), Gabor and multimedia algorithm (Multi), the specific experimental results are shown in Fig. 8.

It can be clearly seen from Fig. 8 that M_{44} the value of the model MAE is 3.53, which is much lower than other algorithms; in addition, its ACC value is 98.3% in face gender and fatigue recognition, which is also higher than other algorithms. The experimental results are obvious. Compared with other face recognition algorithms and depth algorithms, the performance of the proposed algorithm has higher recognition performance, and it also speeds up the training time and improves the recognition accuracy.

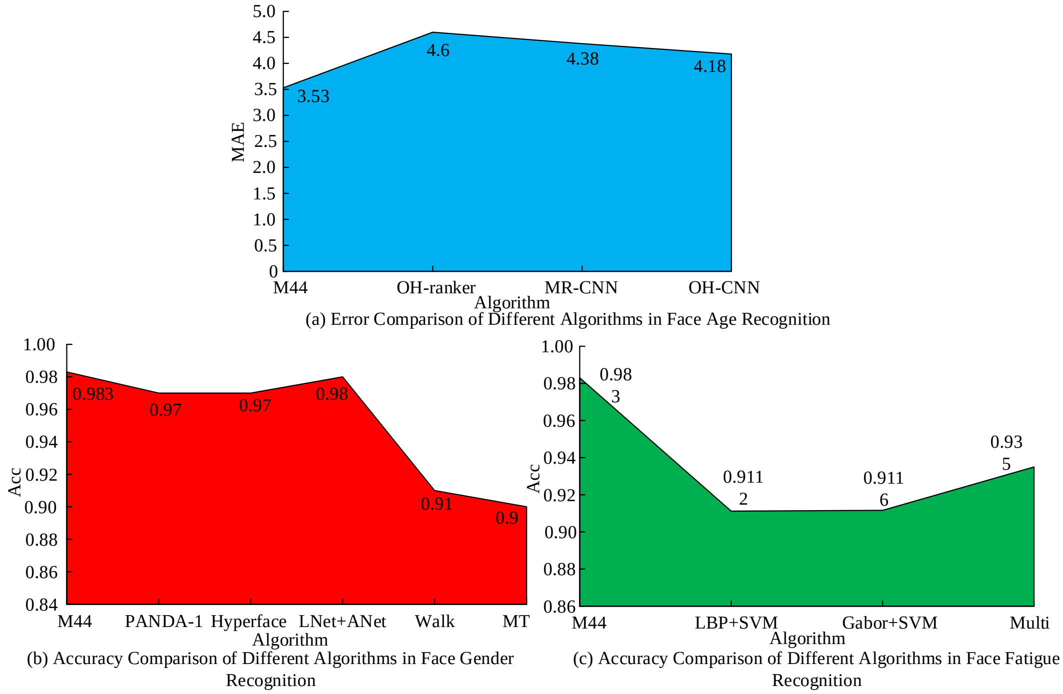


Fig. 8. Performance comparison of different algorithms in face recognition.

V. Results and Discussion

The demand for multitasking facial recognition has significantly improved in many current fields. In order to achieve multitasking facial recognition, existing technologies generally use multiple models in parallel to complete multiple different recognition tasks, which may have serious problems in some cases. Firstly, multiple models have slow training time and long application cycles. Secondly, in practical use, there is a problem of high system resource consumption, which also means high hardware requirements and increased usage costs. Based on this, a multi task face recognition algorithm was proposed by combining CNN with multi task deep learning.

The experimental results show that the Acc value is the highest at model M34, at which point it is 0.983. As the model accuracy is continuously optimized, the training time for the algorithm to achieve loss convergence is continuously decreasing. When the FAR value is 0.01, the correctly accepted proportional TAR value is 0.844, indicating that the M44 model has higher robustness in practical complex situations and can distinguish facial identities well, which is better than the results of Basil et al. [28]. In addition, the MAE value of the M44 model is 3.53, which is much lower than other algorithms. Its ACC value in facial gender and fatigue recognition is 98.3%, which is also higher than other algorithms. This result is also superior to the results of Vu et al. and Mohammed Ali et al. [29-30].

In the dataset LFW, the highest AUC value appears in M13, at which point the AUC value is 0.982; In the dataset LG-NET and Celeb, the minimum MAE value appears at model M22, the Acc value is the highest at model M34, and the accuracy of facial fatigue recognition appears at model M44, which is 0.983. This result shows that the research algorithm has high performance on different datasets, and the final optimal model has the best performance, indicating the generalization of the algorithm.

Overall, the algorithm proposed in the study not only consumes less training time, but also has higher accuracy and lower error values, indicating good performance.

VI. Conclusion

In order to solve the problems of long training time and high consumption of face recognition, the research proposes a multi-task face recognition algorithm based on CNN and multi-task deep learning, and conducts experimental analysis on its performance and application. Four dimensions of face identity, age, gender and fatigue state are used to verify the performance of the proposed algorithm. The experimental results show that the AUC value of the best model for face identity recognition is 0.982, indicating that the extracted face identity features are very effective; in face age recognition, the absolute error value of the best model is 3.43, while the in the gender experiment, the highest ACC value was 0.983, the average accuracy rate was 98.04%, and the average accuracy rate of fatigue state recognition also reached 97.5%. In addition, in the experiment of the same four dimensions, when comparing the performance and accuracy of it with other algorithms, the time-consuming feature of the proposed algorithm for face identification is only 41ms, which is much lower than other algorithms; face age recognition The MAE value of Acc is 3.53, the Acc value of gender recognition is 0.983, and the Acc value of fatigue state recognition is also 0.983. The three actually show far lower errors than other algorithms and far higher accuracy than other algorithms. On the whole, the proposed algorithm not only consumes less training time, but also has higher accuracy and lower error value, and has better performance. However, although the study utilized the Maxout function to extract features, there is still room for expanding the inter class spacing between different individuals in facial identity recognition tasks. Later work can incorporate the triplet loss function for further research.

FUNDINGS

The research is supported by: 2017 Jiangsu Education Informatization Project: Research on Building a New Lifelong Learning Platform in the Context of Internet+(20172075); The 13th Five Year Plan for Educational Science in Jiangsu Province: Research on Building Jiangsu Intelligent Lifelong Learning Service Platform in the "Internet+" Era (B-b/2018/01/33).

REFERENCES

[1] Liu Q, Wang Y, Ye B, Ding M. Recognition Confidence of Welding Seam Defects in TOFD Images Based on Artificial Intelligence. Automatic Control and Computer Sciences, 2022, 56(2): 180-188.

[2] Nasir N, Kansal A, Barneih F, Al-Shaltone O, Bonny T, Al-Shabi M, Al Shammaa A. Multi-modal image classification of COVID-19 cases using computed tomography and X-rays scans. Intelligent Systems with Applications, 2023, 17: 200160.

[3] Tahoun N, Awad A, Bonny T. Smart assistant for blind and visually impaired people//Proceedings of the 3rd International Conference on Advances in Artificial Intelligence. 2019: 227-231.

[4] Bonny T, Haq A. Emulation of high-performance correlation-based quantum clustering algorithm for two-dimensional data on FPGA. Quantum Information Processing, 2020, 19: 1-21.

[5] Tamilselvi M, Karthikeyan S. An ingenious face recognition system based on HRPSM_CNN under unrestrained environmental condition. Alexandria Engineering Journal, 2022, 61(6): 4307-4321.

[6] Khan AA, Shaikh AA, Shaikh ZA, Laghari AA, Karim S. IPM-Model: AI and metaheuristic-enabled face recognition using image partial matching for multimedia forensics investigation with genetic algorithm. Multimedia Tools and Applications, 2022, 81(17): 23533-23549.

[7] Srivastava S, Kumar A, Singh A, Prakash S, Kumar A. An improved approach towards biometric face recognition using artificial neural network. Multimedia Tools and Applications, 2022, 81(6): 8471-8497.

[8] Karanwal S. Robust local binary pattern for face recognition in different challenges. Multimedia Tools and Applications, 2022, 81(20): 29405-29421.

[9] Marmolejo-Ramos F, Murata A, Sasaki K, Yamada Y, Ospina R. Your Face and Moves Seem Happier When I Smile: Facial Action Influences the Perception of Emotional Faces and Biological Motion Stimuli. Experimental Psychology, 2020, 67(1): 14-22.

[10] Chen C, Zhou X. Collaborative representation-based fuzzy discriminant analysis for Face recognition. The Visual Computer, 2022, 38(4): 1383-1393.

[11] Alami AE, Berrahou N, Lakhili Z, Mesbah A, Berrahou A, Qjidaa H. Efficient color face recognition based on quaternion discrete orthogonal moments neural networks. Multimedia Tools and Applications, 2022, 81(6): 7685-7710.

[12] Tripathi RK, Jalal A SA robust approach based on local feature extraction for age invariant face recognition. Multimedia Tools and Applications, 2022, 81(15): 21223-21240.

[13] Singh C, Majeed S. Novel and robust color texture descriptors for color face recognition. Multimedia Tools and Applications, 2022, 81(15): 21313-21347.

[14] Soni N, Sharma EK, Kapoor A. Deep neural network and 3D model for face recognition with multiple disturbing environments. Multimedia Tools and Applications, 2022, 81(18): 25319-25343.

[15] Sharma V, Joshi A M. VLSI Implementation of Reliable and Secure Face Recognition System. Wireless Personal Communications, 2022, 122(4): 3485-3497.

[16] Long X, Zhang Z, Li YA singular value decomposition representation based approach for robust face recognition. Multimedia Tools and Applications, 2022, 81(6): 8283-8308.

[17] Sharma A, Jindal N, Thakur A, Rana PS, Garg B, Mehta R. Multimodal Biometric for Person Identification Using Deep Learning Approach. Wireless Personal Communications, 2022, 125(1): 399-419.

[18] Han C, Shan S, Kan M, Wu S, Chen X. Personalized Convolution for Face Recognition. International Journal of Computer Vision, 2022, 130(2): 344-362.

[19] Srivastava A, Badal T, Saxena P, Vidyarthi A, Singh R. UAV surveillance for violence detection and individual identification. Automated Software Engineering, 2022, 29(1): 1-28.

[20] Zhao R, Shi F. I2DKPCN: an unsupervised deep learning network. Applied Intelligence, 2022, 52(9): 9938-9951.

[21] Vedantham R. Adaptive increasing-margin adversarial neural iterative system based on facial expression recognition feature models. Multimedia Tools and Applications, 2022, 81(3): 3793-3830.

[22] Deshmukh S, Abhyankar A, Kelkar S. DCCA and DMCCA framework for multimodal biometric system. Multimedia Tools and Applications, 2022, 81(17): 24477-24491.

[23] Silva ED, Kumarasinghe P, Indrajith K, Pushpakumara TV, Vimukthi RDY, Zoysa K De, Gunawardana K, Silva S De. Feasibility of using convolutional neural networks for individual-identification of wild Asian elephants. Mammalian Biology, 2022, 102 (3): 909-919.

[24] Mageshkumar N, Lakshmanan L. An improved secure file deduplication avoidance using CKHO based deep learning model in a cloud environment. The Journal of Supercomputing, 2022, 78(13): 14892-14918.

[25] Kuo YL, Tang S C. Deep regression of convolutional neural network applied to resolved acceleration control for a robot manipulator: Transactions of the Institute of Measurement and Control, 2022, 44(4): 784-798.

[26] Liang C, Zhang Z, Wu Q, Li X. Barrier Lyapunov function-based robot control with an augmented neural network approximator. Industrial Robot: the international journal of robotics research and application, 2022, 49(2): 359- 367.

[27] Levi G, Hassner T. Age and gender classification using convolutional neural networks. Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2015: 34-42.

[28] Basil N, Raad M, Wazzan A N, Marhoon H M. Face recognition with real-time framing based on multi task convolutional neural network: a case study. Int. J. Mech. Eng, 2022, 7(2): 3170-3178.

[29] Vu H N, Nguyen M H, Pham C. Masked face recognition with convolutional neural networks and local binary patterns. Applied Intelligence, 2022, 52(5): 5497-5512.

[30] Mohammed Ali F A, Al-Tamimi M S H. Face mask detection methods and techniques: A review. International Journal of Nonlinear Analysis and Applications, 2022, 13(1): 3811-3823.