IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 April 2023, accepted 12 May 2023, date of publication 18 May 2023, date of current version 16 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3277495

RESEARCH ARTICLE

Face Recognition Model Optimization Research Based on Embedded Platform

WEI LI[1], JINBAO SUN[2], JING ZHANG[1], AND BOCHENG ZHANG[1]

[1]School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China
[2]Neusoft Corporation, Shenyang 110179, China

Corresponding author: Wei Li (liweizhisjb_1@163.com)

ABSTRACT The development of information technology has promoted the expansion of the application field of facial recognition technology. Its mainstream recognition methods rely on deep learning algorithms for calculation, but the problem of large data computation brought by its system makes it difficult to apply to embedded platform devices. As a result, this study focuses on improving recognition systems built on lightweight backend networks and builds an embedded platform system environment using multi-scale feature fusion, anchor box size optimization, the addition of channel attention mechanism weighted features, affine face alignment, and file compilation. The experimental results showed that when the number of iterations was 300, the loss value (0.46) of the improved embedded algorithm was much smaller than that of other comparison algorithms (1.42, 1.73, 2.01), and its ACC value (0.924) was significantly better than other comparison algorithms (0.915, 0.909, 0.894). The minimum system testing time consumedwas7 ms. This deep learning embedded facial recognition algorithm has high recognition accuracy, a fast running speed, and is less limited by environmental conditions and data types. It is ideally suited for use in embedded hardware devices, broadening the scope of equipment matching and facial recognition algorithms' applications. As a result, it is better suited to satisfy the demands of embedded devices and massive data processing jobs.

INDEX TERMS Embedded, face, deep learning, lightweight back-end network, channel attention mechanism.

I. INTRODUCTION

The rise of biometric recognition technology effectively compensates for the shortcomings of conventional identification methods by using external means and tools that produce subpar recognition results. This is due to the development of computer technology and the increased importance that people place on the privacy of information and data. Due to its simplicity in data collection, immediacy of recognition results, scalability, and interactivity, face recognition technology is employed in a variety of fields, including smart payments, access management, and criminal investigation [1], [2]. However, the effectiveness and accuracy of face recognition can be limited by objective conditions such as lighting conditions and device resolution, as well as subjective conditions such as expression posture and data collection angle [3], [4]. The emergence of deep learning technology has provided new tools and evaluation ideas for traditional face recognition. The model is robust to different illumination conditions and reduces the interference of shadow areas on face content recognition. The experimental results showed that the model-driven face image illumination processing algorithm can effectively recognise face images in different scenes and at different scales [5]. The deep learning model can technically handle facial recognition. The study suggests a deep learning, lightweight back-end network face recognition system and quantifies and improves various aspects of face recognition and detection while introducing an attention mechanism to weight the feature channels to better achieve feature extraction of key information about faces. This will help reduce the interference of lighting conditions, acquisition equipment, and shooting angles on recognition accuracy. In particular, face alignment and redundant information removal are accomplished using non-maximum

The associate editor coordinating the review of this manuscript and approving it for publication was Hosam El-Ocla.

and affine transformation, the lightweight back-end network MobileNet0.25 for feature fusion of various dimension levels, and image feature extraction. Subsequently, the size of the anchor frame was optimized to ensure better extraction of detailed features, and attention mechanisms were introduced on the basis of lightweight networks to weight the feature channel weights to improve recognition accuracy. The implementation of deep learning facial recognition technology based on embedded platforms through cross compilation environment design, embedded platform compilation, and other aspects, as well as improving the accuracy of feature extraction of important facial information.

II. LITERATURE REVIEW

The rise in low-cost mobile cameras and surveillance systems is fueling the continual growth of facial recognition's application spectrum while simultaneously presenting new tools and difficulties for use in the real world. Facial recognition has currently established itself as a key authentication technique. Most scholars have conducted research on image detection, facial information capture, and feature generation for recognition. Juneja K et al. improved their methods for face normalization, feature generation, and recognition and defined the performance and accuracy of different methods. The comparison between traditional methods and the latest methods has been achieved from a mapping perspective, providing a new exploration method for improving facial recognition systems [6]. Based on the differences in recognition accuracy of previous generations of facial recognition algorithms for different ethnic images, Cavazos J G Scholar analyzed the differences in the impact of data-driven and scene modeling on images, such as image quality, algorithm architecture, threshold decision-making, and demographic constraints. The scholar proposed using deep convolutional neural networks to classify datasets and improve algorithm recognition accuracy [7]. To enhance the effectiveness of the algorithm and lessen the interference of background information while increasing the efficiency of resource availability, Wang et al. applied a synaptic neural regeneration mechanism to the visual system. Experimental results showed that this improved model can effectively improve recognition accuracy based on covering multiple learning dimensions [8]. Goswami et al. propose to improve the deep learning algorithm with filter response behaviour to address the low robustness of the algorithm. Experimental results showed that this improved method can effectively reduce the impact of adversarial attacks on the performance of the algorithm, and the recognition results are better [9]. Keinert's team uses a non-convex induced penalty function and This simplified strategy can effectively represent the feature space in a low-dimensional way and has a high recognition performance [10]. He et al. use convolutional networks and coefficient classification to achieve dynamic matching of face features, avoiding the dependence of the original algorithm on a priori location information. The experimental results showed that the algorithm performs well in multiple databases and retains high integrity of details [11]. Wang proposes to redefine the search space with a profit-normalised loss function and a reward-oriented approach after analysing the core concept of feature differentiation. Recognition simulation experiments show that this improved method has better recognition accuracy [12].

Images are progressively becoming a form of communication thanks to advances in media and information technology. The recognition power of convolutional neural networks has expanded the scope of their use, but it is challenging to combine data from several dimensions. Therefore, Lou proposed a multi-level information fusion algorithm under convolutional computation to ensure the integrity of feature information data. The improved model can group feature extraction for network information data, and its recognition effect is better [13]. Rejeesh proposed a face inference recognition algorithm based on interest points and optimised and classified the parameters of key points with an adaptive genetic algorithm and fuzzy neural network. Simulation results showed that the algorithm can effectively recognize the detected images, and the recognition efficiency is significantly higher than other existing Algorithm [14]. Geet al. proposed to extract facial information from high-resolution images in order to improve the accuracy of low-resolution field face images and reduce the computational cost, and the results were used to fine-tune the low-resolution images. Simulation results showed that the face recognition efficiency accuracy of low-resolution images was higher and the memory consumption was less than 1 MB [15]. To address the current research on the face recognition mechanism of convolutional neural networks. To address the problem that the current research on the face recognition mechanism of convolutional neural networks is limited, Grm et al. labeled face target images with the help of a wild dataset and analyzed the mechanism of covariate factors affecting image recognition quality and feature extraction. The results showed that noisy data and missing pixels negatively affected the recognition performance of the model and that there was a limit to the quality of the image under compression changes. The results can provide more research inspiration for computer recognition of face images [16]. Considering the high-dimensional nature of face images, Annamalai collects data with the help of semantic analysis and implements feature extraction of facial information with LTP patterns and key point scaling, while the Firefly optimiser and DBN network perform dimensionality reduction and classification operations on the images. Experimental results show that the recognition algorithm improves accuracy by more than 5% [17]. As a way to ensure the clarity and accuracy of image information data recognition, Dorofeev proposes to implement 3D face recognition using convolutional networks and incorporate the idea of depth filtering into the algorithm model. Experimental results also showed that this improved method was successful in recognizing key features of 3D faces [18]. Wang et al. used the local binary histogram method for facial recognition system drone technology to better ensure the safety of citizens. When identifying the target object, this method designs a label and sends its image and position coordinate information to relevant personnel, with a recognition accuracy rate far exceeding 85% [19]. Yang scholar designed a face recognition attendance system based on real-time video processing, and analyzed and designed it from four aspects: accuracy, stability, truancy situation, and interface settings. The experimental results showed that compared to traditional check-in methods, the face recognition attendance system has a recognition accuracy of over 80%, effectively reducing the situation of students leaving early and skipping classes, and improving classroom efficiency [20]. According to Baloch scholars, facial recognition and motion detection are complex systems for connecting information that can operate simultaneously while also storing data, have a wide range of application spaces, and have the ability to link data [21]. In response to the changes in facial appearance features and recognition difficulties caused by age, scholar Zhao proposed a deep Argue invariant model for field face recognition. By designing a unified deep architecture pattern, facial data authenticity matching, and effective training strategies, recognition accuracy was improved. At the same time, a new large-scale cross Argue face recognition (CAFR) benchmark dataset was constructed. The results indicate that the recognition method exhibits good recognition generalization ability in unconstrained face recognition datasets [22]. Zhang et al. proposed an attention aware facial recognition method based on deep convolutional neural networks and reinforcement learning to address the interference of outdoor environments on facial recognition. Facial recognition is achieved through facial feature recognition and reinforcement training, as well as network feature embedding. The results showed that this method had good recognition performance in public facial verification databases [23]. Wang proposed that the information processing pathway of the human visual cortex can be simulated to achieve facial type image recognition on different backgrounds and sizes. At the same time, synaptic maintenance mechanisms and neuron regeneration mechanisms are introduced to reduce environmental background interference and improve network operation efficiency. The experimental results showed that the proposed WWN model can effectively recognize the type, position, and size of facial images [8].

In conclusion, the majority of researchers use convolutional neural algorithms to research and develop face recognition algorithms and improve recognition accuracy in terms of feature dimension extension, neural mechanism introduction, interest point recommendation, etc. Additionally, to increase recognition accuracy, the majority of researchers rely on the extraction of face feature information and hardware device design. Some of the research content involving recognition system design only uses the hardware device as an intermediary tool for data collection and input, which makes it difficult to meet the device's performance. Based on this, the study suggests building face recognition models for embedded platforms in deep learning mode and improving data processing, feature information extraction, target image recognition, and detection to increase the efficiency and accuracy of face recognition algorithms. This is done in order to offer fresh perspectives and methods for face recognition.

III. EMBEDDEDFACERECOGNITIONMODEIBASEDDESIGN

The study focuses on the characteristics and influencing factors of facial recognition. Firstly, it utilizes the lightweight backend network MobileFaceNet for facial detection and key point detection, and utilizes the introduction of feature pyramid and channel attention mechanisms to achieve facial detection and key recognition optimization, improving its feature extraction ability. And achieve facial feature matching using cosine similarity, and design recognition systems using embedded development environments to provide technical tools for verifying the performance of facial recognition network models.

A. FACE RECOGNITION MODEL CONSTRUCTION AND NETWORK OPTIMIZATION

Strengthening face feature recognition is the key to resolving this issue since face features are frequently employed due to their ease of capture and contactless nature, but their identification accuracy is easily influenced by the surroundings and the technology of the device. Neural networks are used for face recognition by pre-processing and extracting features from the input image and using network training to discriminate the information. However, the number of parameters involved in the neural network itself is excessive, and the resource consumption and time cost of most neural network models is high, making some of the algorithms improve their performance by increasing the depth of the network and reducing the complexity of data processing [24], [25].The convolution is divided into two parts by the MobileFaceNet network architecture: a non-cross-channel convolution and a feature-mergeable convolution, each of which has the ability to extract features. Each channel of the non-cross-channel convolution has its own independent convolution kernel. The ascending and descending of feature data dimensions is possible with feature processable convolution [26]. Figure 1 shows a diagram of the separable convolution kernel.

The equation for the standard convolution is shown in equation (1).

$$P1 = Dk^{*}Dk^{*}M^{*}N^{*}Dw^{*}Dh \tag{1}$$

In equation (1), $P1 = Dk^{*}Dk^{*}M^{*}N^{*}Dw^{*}Dh$ represents the width and height of the standard convolution, the number of convolution channels and the number of convolutions, respectively. The mathematical expression for calculating the depth-separable convolution divided into depth convolution and point convolution is given in equation (2).

$$\begin{aligned} P' &= \frac{Dk^{*}Dk^{*}M^{*}N^{*}Dw^{*}Dh + M^{*}N * Dw^{*}Dh}{Dk^{*}Dk^{*}M^{*}N^{*}Dw} \\ &= \frac{1}{N} + \frac{1}{D_k^2} \end{aligned} \tag{2}$$

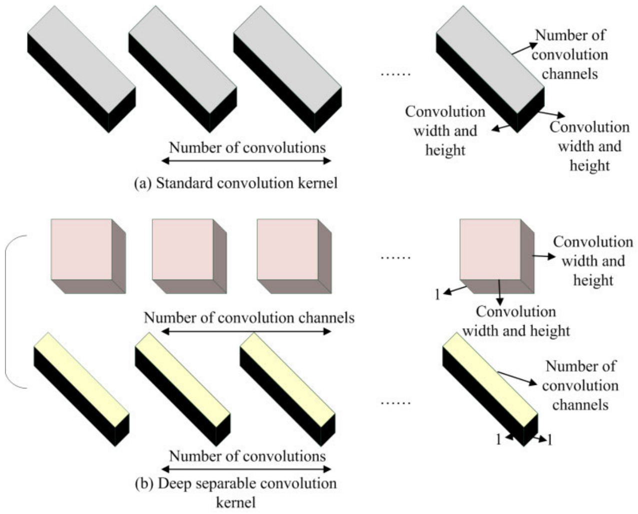

FIGURE 1. **Schematic diagram of traditional convolution structure and depth separable convolution core.**

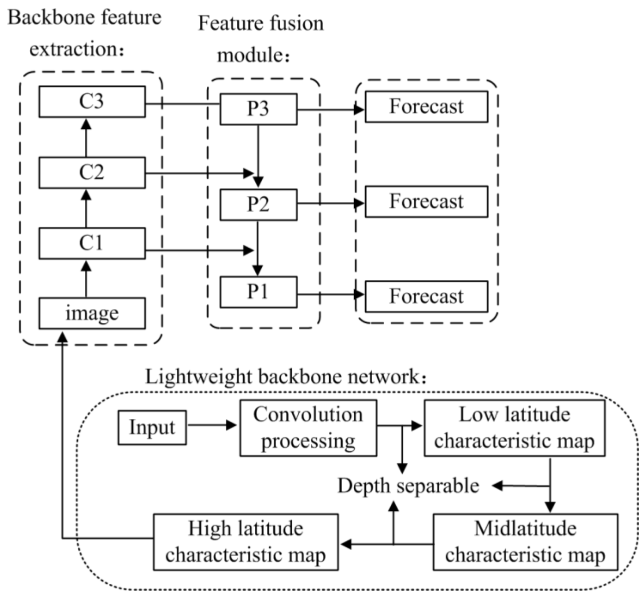

FIGURE 2. **Schematic diagram of face detection network framework.**

Compared with equation (1), the number of parameters and computational complexity of equation (2) have been reduced. The separable convolution in MobileFaceNet can effectively reduce the computational volume of the model and reduce the accuracy degradation caused by lightweight processing, which mainly plays the role of 1*1 convolution in the expansion and recovery of dimensionality to achieve the extraction of feature information and the preservation of information integrity. How to recognize a face in distinct photographs is crucial to achieving recognition accuracy since the visual impact of a face in different images changes significantly due to the variation in image shooting distance. Therefore, the study uses feature pyramids to fuse feature information maps of different scales to enhance their characterisation capabilities, and in the process, it achieves the prediction of high and low-dimensional features in order to significantly improve the accuracy of target detection. Figure 2 shows a schematic diagram of the face detection network framework.

The C1, C2, and C3 in Figure 2 represent the feature maps obtained by the face detection network framework through convolution, which represent low latitude, medium dimension, and high latitude feature information, respectively. P1, P2, and P3 represent the results of feature fusion from different dimensions [27], [28]. After upsampling, the high latitude feature map outputs a rich information map by adding it to the low latitude feature map. The feature information represented by the three dimensional levels of high, low, and medium has also huge variations, and the face detection network architecture is to convolve the distinct feature maps obtained according to their dimensional differences in multiples from bottom to top. When the channel dimensions are changed, the modest variations in the images are ensured by setting the number of convolutions for image merging to 1*1. The network framework also enables the fusion of multi-dimensional detection features. Face detection requires the determination of the specific location of the human figure in the image, and the detection network is mostly based on determining the reference frame with the help of an anchor window and the distance offset between the target frame and the reference frame as the basis for face location detection. Because of the embedded platform's restrictions on image size, it can be challenging to identify and recognize some target images with wide anchor windows. The study therefore reduces the size of the anchor frame in the face detection network structure so that it can better adapt to the platform properties and enhance the detection and recognition of small image details. A channel attention mechanism is introduced to obtain the importance and weighted features of different feature channels to effectively identify the features of important channels and reduce the proportion of feature information of non-important channels in due course. Figure 3 illustrates the structure of the channel attention mechanism.

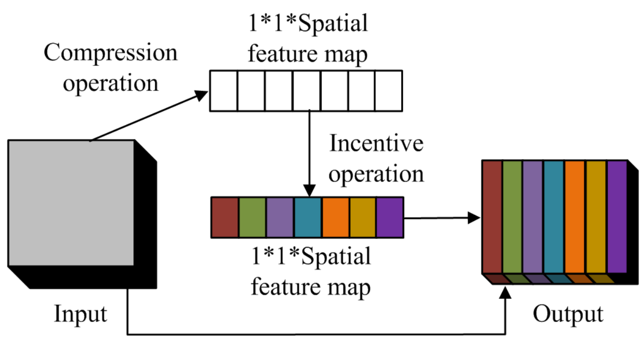

FIGURE 3. **Structure diagram of channel attention mechanism.**

The Sigmoid activation function is used to map the input values to obtain the weight coefficients for each channel [29], [30]. Figure 4 shows the structure of the battleneck convolution with the introduction of the channel attention mechanism.

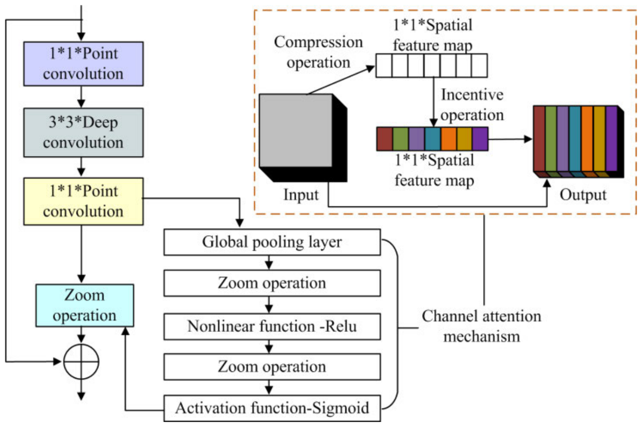

FIGURE 4. Schematic diagram of battleneck convolution structure introducing channel attention mechanism.

In Figure 4, a battleneck-structured graph is added on top of the MobileNet lightweight back-end network in order to guarantee high feature extraction performance for the model. This is accomplished primarily by increasing the number of channels through convolution, ensuring that the feature information of the input values can be presented more comprehensively, and recovering the dimensional information with convolution in order to improve the feature extraction performance of the network. The channel attention mechanism is a global pooling layer to extract the global feature information from the input values, followed by a fully connected layer to reduce and then increase the dimensionality of the feature information of different channels to obtain the correlation between the channels, and finally the activation function to calculate the sum of the feature weights of the channels [31], [32]. The battleneck convolutional structure with the introduction of the channel attention mechanism is effective in reducing the computational effort for extracting key feature information and improving its data recognition capability [33].

As there are differences in the size and overall scale of the target face, it is necessary to set multiple sizes of pre-selected frames to ensure that the detection effect is better suited to the diversity of human images. Three step sizes of 8, 16 and 32 are set for small, medium and large scale faces respectively. By setting the aspect ratio of the preselected box at each scale, each feature point can be represented in the preselected box and the mathematical expression of the coordinates in the preselected target box is shown in equation (3).

$$\begin{cases} R_x = P_w q_x(P) + P_x \\ R_y = P_h q_y(P) + P_y \\ R_w = P_w \exp(q_w(P)) \\ R_h = P_h \exp(q_h(P)) \end{cases} \tag{3}$$

In equation (3), P_x, P_y, P_w, P_h is the centre coordinates, width and height of the preselected frame, $q_x(P)$, $q_y(P)$, $q_w(P)$, $q_h(P)$ is the variable for network regression, and Rx, Ry, Rw, R_h is the centre coordinates, width and height of the predicted frame. It is necessary to filter and filter the target information of the preselected frames in order to reduce the overlap of the redundant frames. This is done using the non-maximum suppression algorithm since the overlap of the picture information of the preselected frames is an issue. The algorithm detects the degree of overlap between the candidate box with the highest confidence level and other candidate boxes, and compares the ratio of the intersection set of the two figures with a preset threshold size. If the degree of overlap is greater than the preset value, then the corresponding duplicate image candidate frame part is removed. Face alignment is one of the most important indicators that affects the recognition effect. The angle of the shot and the offset angle will make the face position not "centred" in the traditional sense [34]. The affine transformation does not adjust the position of the image before and after the transformation, but only presents a change in shape, as shown in equation (4) [35].

$$\begin{cases} u = a_1 x + b_1 y + c_1 \\ v = a_2 x + b_2 y + c_2 \end{cases} \tag{4}$$

In equation (4), u, v represents the coordinates of the original x, y transformed by the transformation factor a, b, c. equation can be expressed as a matrix expression $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ and is equivalent to equation (5) if the affine transformation does not require a translation operation.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos A & -\sin A \\ \sin A & \cos A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}$$

The matrix transformation allows for the alignment of face images by varying the key information extracted from the face with the standard values.The face has similar feature dimensions and is angularly differentiable, so the study uses cosine similarity, whose mathematical expression is shown in equation (6), to calculate the similarity calculation in order to recognize differences in face images. The face recognition network is used to calculate the distance between different features in order to recognize differences in face images.

$$\cos A = \frac{(b, c)}{\|b\| \, \|c\|} = \frac{\sum_{i=1}^{N} bi^{*}ci}{\sqrt{\sum_{i=1}^{N} bi^2}^{\,*} \sqrt{\sum_{i=1}^{N} ci^2}} \tag{6}$$

In equation (6), $b, c, \|b\|$, $\|c\|$ is the feature vector and its corresponding mode, A is the vector angle, (b, c) is the inner product of the feature vector, (b, c) is the dimensional eigenvalue of the vector, and cos is the cosine similarity, with higher values indicating higher similarity between the two features.

B. EMBEDDEDDEVELOPMENTENVIRONMENTCONSTRUCTIONANDRECOGNITIONSYSTEMDEPLOYMENTDESIGN

The face recognition system is mainly designed with the help of an embedded platform, and face recognition is achieved through the acquisition of face image information, data input, and face detection on it. The study selects the RV1126_RV1109 EVB platform as the embedded platform, and most embedded hardware has certain requirements for the inference framework, which makes the traditional embedded neural network framework less applicable to the device [36]. The study suggests putting platform model experiments based on Neural-Network Process Units (NPU) chips into practice. These chips must load and quantize the model format that was transformed on the computer side; the quantization is primarily fixed-point quantization of the floating-point model. Considering the limited resources of embedded devices, the study designs cross-compilation tool chains to process the formats under different architecture tools. The cross compiler compiles the program written on the PC, and the generated file is in the format accepted by the ARM architecture. The cross compilation tool chain designed for research is gcc-arm-8.3-2019.03-x86_ 64 arm linux gnueabihf, which is the system related code developed based on C++, can realize preprocessing, syntax analysis and the generation of files that can be recognized and executed by the linker. The compilation process is shown in Figure 5.

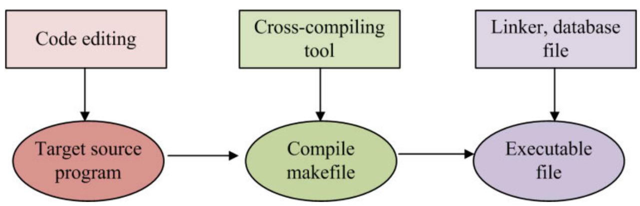

FIGURE 5. Cross-compilation process.

To improve the effectiveness of face recognition, the OpenCV computer vision library is also employed, and OpenCV is configured and compiled for use in embedded devices. The study also looks into the deployment of deep learning models on embedded hardware that have been network pruned, quantized, and lightened to make them more device-adaptable. The lightweighting process is discussed in Section 2.1 of the methodology and is not described here. Network pruning is the process of reducing the redundant information in the network model by processing the weights and structure of the network model, retaining the more important channels, and removing the less important ones, thus reducing the complexity of the network. The model is also quantized to avoid performance loss, as quantization is the mapping of the original floating point numbers to reduce their range. The two main forms of model quantization by the embedded platform are subjectively provided quantization datasets and quantization models derived from deep learning frameworks [37]. The mathematical expression for quantization is shown in equation (7).

$$r = Round(S(q - Z)) \tag{7}$$

In equation (7), q denotes a floating point 32-bit value, Z, S is the offset and scaling factor, and $Round$ denotes the mathematical function expression for rounding. When the offset is 0, the quantization shows symmetric as well as asymmetric variations. The study performs static quantization of the face recognition and detection model in terms of Sbit, i.e. quantization formulation and data calibration of the processed floating-point model. The model object is initially initialized, the input characteristics are configured, including the index position, memory size, and data type format of the input data, and then the model is used for inference and result output in order to deploy the face detection and recognition model with the embedded platform. Figure 6 shows the flow diagram of the system software.

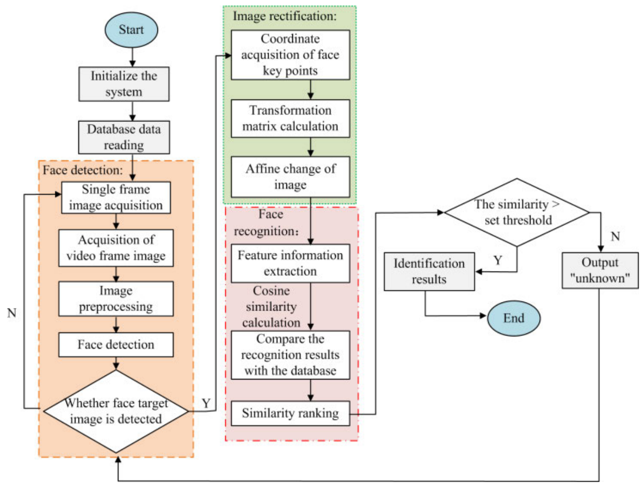

FIGURE 6. System software flow diagram of face recognition model.

In Figure 6, the embedded platform system is initialised and the face database data is read, followed by face detection, image correction, and image recognition of the image data, where the recognition part requires target framing of the detected face image and automatically locates key features such as eyes, nose tip, eyebrows, etc. To assess whether the recognition result's similarity exceeded a set, its similarity to the database was compared. The output of the recognition result will finish the identification and matching of facial features if it exceeds the predetermined threshold.

IV. ANALYSIS OF THE APPLICATIONEFFECT OF EMBEDDEDFACERECOGNITIONMODEL

The study introduces the WIDER FACE dataset with lighting factors, different face scales and expressions for testing and training, and relies on a computer-based training platform, the PyTorch deep learning framework and the Ubuntu 18.04 operating system for experimental simulations. The performance of the proposed face detection and recognition algorithm is first examined and compared with the local

Binary patterns-discrete cosine transform (LBP-DCT) based algorithm, Unconstrained-Deep Convolutional Neural Network (U-DCNN) and the Adaboost-based improved embedded face recognition algorithm are compared. Figure 7 shows the losses of the four algorithms at different numbers of iterations.

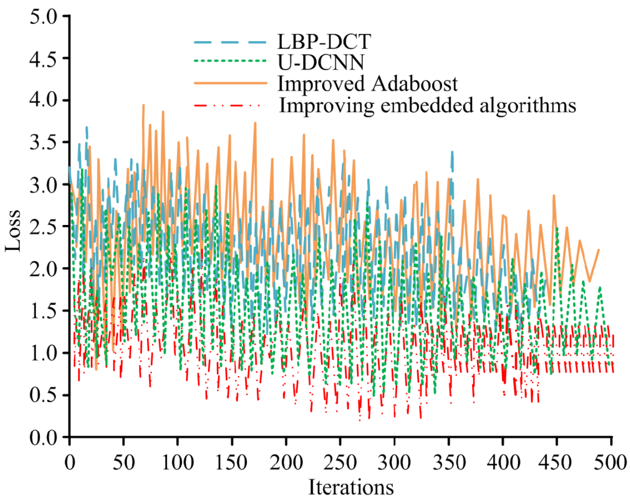

FIGURE 7. **Loss of four algorithms under different iteration times.**

The horizontal and vertical coordinates in Figure 7 indicate the number of iterations of data training and the loss value of the function, respectively, and what can be seen from the results is that the increase in the number of iterations causes the loss values of all four algorithms to vary in magnitude. The LBP-DCT algorithm, U-DCNN algorithm, improved Adaboost algorithm, and improving embedded algorithms all have loss values of 1.42, 1.73, 2.01, and 0.46, respectively, when the number of iterations is 300. The overall loss variation curve reveals that the average loss values of the four algorithms are 1.23, 1.72, 2.05, and 0.97, respectively. The loss situation of the U-DCNN algorithm is similar to that of the improved embedded algorithm, but its curve fluctuation is more obvious. The peak loss situation of the improved Adaboost algorithm basically occurs before the number of iterations is less than 200. The above results indicate that the performance stability of the U-DCNN algorithm and the improved Adaboost algorithm is poor compared to the improved embedded algorithm. The accuracy of the facial recognition process depends on the integrity of the image information extraction, and information loss and algorithm performance stability are crucial factors to take into account. The study analyzed the comparison results of algorithms under different losses. The comparison algorithms here are LBP-DCT algorithm, U-DCNN algorithm, and embedded recognition algorithm. The improved Adaboost algorithm was not considered due to its significant performance differences, and the results are shown in Figure 8.

The findings in Figure 8 show that the accuracy of face recognition by the LBP-DCT method and the U-DCNN

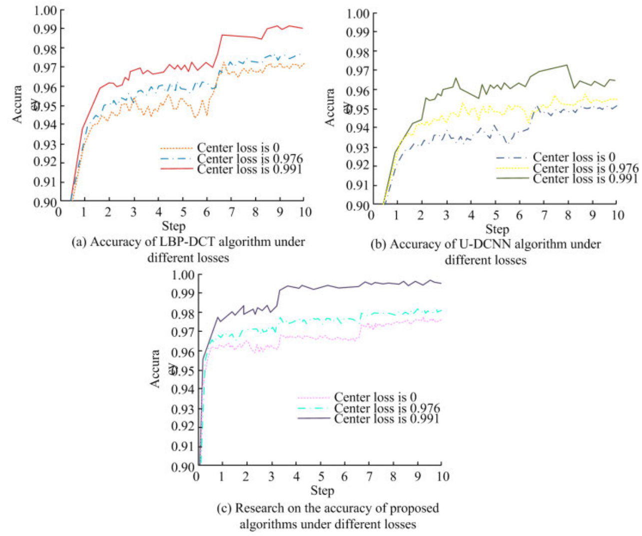

FIGURE 8. **Identification accuracy of three models under different loss values.**

algorithm will be influenced to varied degrees when there is a variation in the loss value and the variability is significant. In more precise terms, under various loss scenarios, the accurate identification curve of the LBP-DCT algorithm has a maximum numerical difference variation of 0.25, whereas the accurate identification curve of the U-DCNN algorithm has a maximum numerical difference variation of 0.27. The embedded facial recognition algorithm proposed in the study has higher accuracy and is less affected by data loss.

The recognition performance of the four algorithms was then examined in terms of Accuracy (ACC), which determines the probability that two detected images are the same person. The similarity of the face features in the dataset was calculated and the test recognition results were analyzed with the help of a ten-set cross-test. The results are shown in Figure 9.

The results in Figure 9 show that the algorithms with better overall performance in terms of overall face feature recognition accuracy, from good to poor, are the improving embedded algorithms > LBP-DCT algorithm > U-DCNN algorithm > Improved Adaboost algorithm, with ACC values of 0.924, 0.915, 0.909 and 0.894 respectively, and the improving embedded algorithms for face feature recognition were stable and less volatile.The video image feature extraction results of the proposed algorithm are analyzed, and the results are shown in Figure 10.

The results in Figure 10 show that after the application of the method adopted in the research, the fluctuation of the time node of the target feature extraction in the video image location has been significantly improved compared with that before the application, and the time consumed is basically stable at 0.75s in the later stage of multi-frame image processing. The results of Peak Signal to Noise Ratio (PSNR) show that the method adopted in the study can effectively reduce the

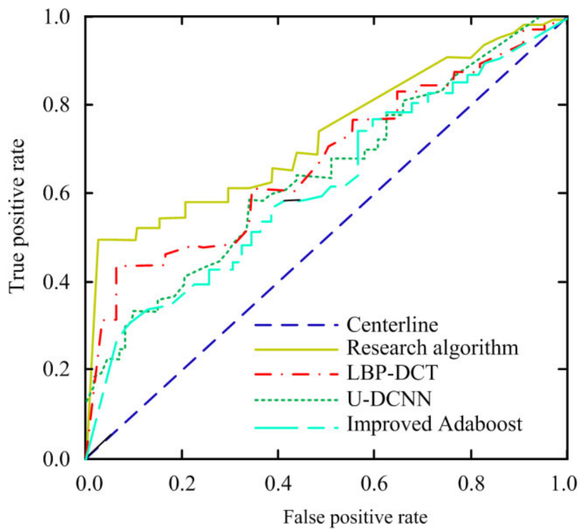

FIGURE 9. Accurate performance of four algorithms for face feature recognition.

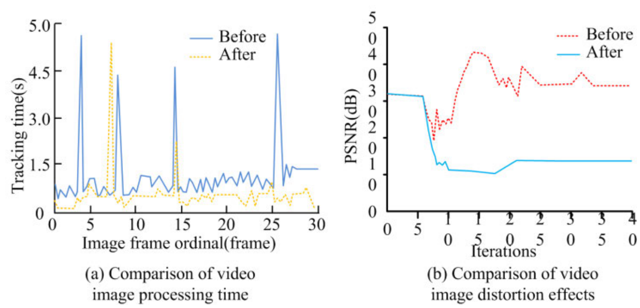

FIGURE 10. Comparison of video image processing time and distortion effect.

information distortion of image features. The value after its application is basically less than 20dB, and the improvement in image quality is more than 30% after the number of iterations exceeds 15.The algorithm was then analyzed for cosine similarity, system testing time, and memory occupation for face images of different complexity in the WIDER FACE dataset, and the dataset was classified into simple, medium, and complex according to complexity.

The results in the table show that the performance of different algorithms varies considerably in data sets of different complexity. From the perspective of cosine similarity, a larger value indicates a significant decrease in the quantization accuracy of the model. The similarity values studied under the three datasets were 0.999997, 0.999998, and 0.999998, with relatively small numerical variability. The similarity value of the Improved Adaboost algorithm is 0.999638, 0.999530, 0.999421. The similarity value of the LBP-DCT algorithm is 0.999974, 0.999969, 0.999963. The similarity values of the U-DCNN algorithm are 0.999986 and 0 999974, 0.999962. These three comparing algorithms' similarity scores all varied significantly, which makes it difficult to predict how well they would perform when executed on various datasets. At the

TABLE 1. Index test of four algorithms in face images with different complexity.

Dataset complexity	Index	Improved Adaboost algorithm	LBP-DCT algorithm	U-DCNN algorithm	Research algorithm
Simple	Cosine similarity	0. 999638	0. 999974	0.999986	0.999997
	System test time	15ms	12ms	14ms	10ms
	Memory percentage	10.69MB	10.29MB	10.14MB	7.12MB
Medium	Cosine similarity	0. 999530	0. 999969	0. 999974	0.999998
	System test time	16ms	15ms	15ms	9ms
	Memory percentage	12.54MB	11.85MB	10.38MB	8.34MB
Complex	Cosine similarity	0. 999421	0. 999963	0. 999962	0.999998
	System test time	22ms	13ms	14ms	7ms
	Memory percentage	13.06MB	11.26MB	10.54MB	8.06MB

same time, under different datasets, the minimum testing time and memory of the improving embedded algorithms are 7ms and 8.06MB, respectively, which are much smaller than the 15ms and 10.69MB of the Improved Adaboost algorithm, 12ms and 10.29MB of the LBP-DCT algorithm, and 14ms and 10.14MB of the U-DCNN algorithm. The above results indicate that the improved embedded algorithms perform well on different datasets and is less susceptible to fluctuations due to the difficulty of data information recognition. At the same time, it demonstrates that its algorithm performance has high application effectiveness, and can adapt to different data types with less runtime and memory use. The data in Table 1 shows that the similarity values of the proposed algorithm in the three data sets are 0.999997 and 0.999998 respectively, which are all higher than the similarity values of other algorithms, and the accuracy values of other algorithms are more likely to be affected by the difficulty of identifying data information, with the improved Adaboost algorithm in particular having the lowest similarity values. The suggested algorithm performed best in terms of test time and memory usage, with test times of 7 ms and 8.06 MB, respectively, and good performance in terms of stability and effectiveness. To further analyze the effectiveness of the proposed algorithm in face recognition, the application effect of the proposed face detection and recognition algorithm will be analyzed. Character images of different scales and experimental environments will be randomly selected from the test dataset, including scene information, changes in environmental factors, and face samples with scale differences. Firstly, ensure that the selected image format and clarity are consistent. Next,

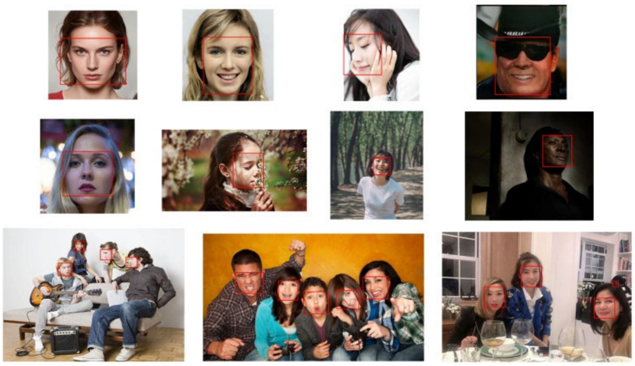

FIGURE 11. Face recognition effect of research algorithm.

compute the convolution kernel of the feature channel using MobileFaceNet's separable convolution, and combine feature information maps of various scales using the feature pyramid. In an attempt to increase the recognition of intricate image elements, the channel attention method is also implemented to limit the size of the anchor box. Determine the center coordinates in the feature pre-selection box, ascertain the position of the face picture, and adjust the relevant step sizes of the feature map to 8, 16, and 32 based on the size difference of the chosen image. The application results are shown in Figure 11.

The results in Figure 11 show that the proposed algorithm can perform face recognition and detection well, with small deviation values and effectively avoiding the effects of image scale, lighting factors, image angles, and face expressions on recognition accuracy, and has good application results.

V. CONCLUSION

Deep learning networks are increasingly in demand for deployment on embedded hardware as artificial intelligence advances, and various hardware device platforms offer various benefits and applications for data processing. The models involved in artificial intelligence networks are relatively large, and compressing the models is a common and computationally efficient approach. The most common methods are pruning and quantization. IA lightweight network architecture is suggested for the design of face recognition frameworks in order to improve the application effect of intelligent networks on embedded devices. Improvements are also made in data training, face detection, and alignment to improve the application effect of the models. The lightweight backbone network MobileNet0.25 extracts features, and the feature pyramid realizes multi-scale feature fusion, which greatly improves the pertinence of face recognition. In addition, the deployment of embedded platforms reduces the amount of computation while avoiding data loss. The average loss value (0.97) of the embedded algorithm in the experimental results is smaller than that of the LBP-DCT algorithm (1.23), U-DCNN algorithm (1.72), and improved Adaboost algorithm (2.05), and its function loss curve changes relatively smoothly. The LBP-DCT algorithm suffers from data loss when the number of iterations exceeds 450, while the U-DCNN algorithm exhibits significant curve fluctuations. The accurate identification curves of the LBP-DCT algorithm and the U-DCNN algorithm exhibit maximum numerical differences of 0.25 and 0.27 under different loss conditions. Enhance the face detection network's performance in small object detection by increasing its size and optimizing the size of its anchor boxes. The addition of cosine similarity can improve face feature matching, and the addition of the channel attention method further clarifies the representativeness of various features. In the experimental results, the order of the accuracy and effectiveness of facial feature recognition algorithms for different algorithms is: embedded algorithm (0.924)>LBP-DCT algorithm (0.915)>U-DCNN algorithm (0.909)>improved Adaboost algorithm (0.894). In the simultaneous testing of application effects, the cosine similarity value of the embedded algorithm is relatively high, with a minimum testing time and memory of 7ms and 8.06MB, respectively. It can recognize face images in a variety of scenarios and scales, and the total data detection accuracy is good and reasonably consistent, indicating that it has a high accuracy effect in face recognition applications on embedded systems. Future research and improvements should focus on enhancing the detection of multi-dimensional target images and the extraction of information categories for data feature correctness.

REFERENCES

[1] R. Ranjan, A. Bansal, J. Zheng, H. Xu, J. Gleason, B. Lu, A. Nanduri, J. Chen, C. D. Castillo, and R. Chellappa, "A fast and accurate system for face detection, identification, and verification," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 1, no. 2, pp. 82–96, Apr. 2019.

[2] D. Wofk, F. Ma, T. Yang, S. Karaman, and V. Sze, "FastDepth: Fast monocular depth estimation on embedded systems," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 6101–6108.

[3] P. Li, L. Prieto, D. Mery, and P. J. Flynn, "On low-resolution face recognition in the wild: Comparisons and new techniques," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 8, pp. 2000–2012, Aug. 2019.

[4] J. Fu, J. Liu, J. Jiang, Y. Li, Y. Bao, and H. Lu, "Scene segmentation with dual relation-aware attention network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 6, pp. 2547–2560, Jun. 2021.

[5] W. Zhang, X. Zhao, J. Morvan, and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 3, pp. 611–624, Mar. 2019.

[6] K. Juneja and C. Rana, "An extensive study on traditional-to-recent transformation on face recognition system," *Wireless Pers. Commun.*, vol. 118, no. 4, pp. 3075–3128, Jun. 2021.

[7] J. G. Cavazos, P. J. Phillips, C. D. Castillo, and A. J. O'Toole, "Accuracy comparison across face recognition algorithms: Where are we on measuring race bias?" *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 3, no. 1, pp. 101–111, Jan. 2021.

[8] D. Wang, H. Wang, J. Sun, J. Xin, and Y. Luo, "Face recognition in complex unconstrained environment with an enhanced WWN algorithm," *J. Intell. Syst.*, vol. 30, no. 1, pp. 18–39, Jul. 2020.

[9] G. Goswami, A. Agarwal, N. Ratha, R. Singh, and M. Vatsa, "Detecting and mitigating adversarial perturbations for robust face recognition," *Int. J. Comput. Vis.*, vol. 127, nos. 6–7, pp. 719–742, Mar. 2019.

[10] F. Keinert, D. Lazzaro, and S. Morigi, "A robust group-sparse representation variational method with applications to face recognition," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2785–2798, Jun. 2019.

[11] L. He, H. Li, Q. Zhang, and Z. Sun, "Dynamic feature matching for partial face recognition," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 791–802, Feb. 2019.

[12] X. Wang, S. Wang, C. Chi, S. Zhang, and T. Mei, "Loss function search for face recognition," in *Proc. Int. Conf. Mach. Learn.*, Jul. 2020, pp. 10029–10038.
[13] G. Lou and H. Shi, "Face image recognition based on convolutional neural network," *China Commun.*, vol. 17, no. 2, pp. 117–124, Feb. 2020.
[14] M. R. Rejeesh, "Interest point based face recognition using adaptive neuro fuzzy inference system," *Multimedia Tools Appl.*, vol. 78, no. 16, pp. 22691–22710, Aug. 2019.
[15] S. Ge, S. Zhao, C. Li, and J. Li, "Low-resolution face recognition in the wild via selective knowledge distillation," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 2051–2062, Apr. 2019.
[16] K. Grm, V. Štruc, A. Artiges, M. Caron, and H. K. Ekenel, "Strengths and weaknesses of deep learning models for face recognition against image degradations," *IET Biometrics*, vol. 7, no. 1, pp. 81–89, Jan. 2018.
[17] P. Annamalai, "Automatic face recognition using enhanced firefly optimization algorithm and deep belief network," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 5, pp. 19–28, Oct. 2020.
[18] Y. Ahn, S. M. Lee, H. N. Noh, W. Kim, J. Choe, K.-H. Do, and J. B. Seo, "Use of a commercially available deep learning algorithm to measure the solid portions of lung cancer manifesting as subsolid lesions at CT: Comparisons with radiologists and invasive component size at pathologic examination," *Radiology*, vol. 299, no. 1, pp. 202–210, Apr. 2021.
[19] L. Wang and A. A. Siddique, "Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology," *Meas. Control*, vol. 53, nos. 7–8, pp. 1070–1077, Aug. 2020.
[20] H. Yang and X. Han, "Face recognition attendance system based on real-time video processing," *IEEE Access*, vol. 8, pp. 159143–159150, 2020.
[21] Z. Balogh, M. Magdin, and G. Molnár, "Motion detection and face recognition using raspberry pi, as a part of, the Internet of Things," *Acta Polytechnica Hungarica*, vol. 16, no. 3, pp. 167–185, 2019.
[22] J. Zhao, S. Yan, and J. Feng, "Towards age-invariant face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 474–487, Jan. 2022.
[23] L. Zhang, L. Sun, L. Yu, X. Dong, J. Chen, W. Cai, C. Wang, and X. Ning, "ARFace: Attention-aware and regularization for face recognition with reinforcement learning," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 4, no. 1, pp. 30–42, Jan. 2022.
[24] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, "Real-time driver drowsiness detection for Android application using deep neural networks techniques," *Proc. Comput. Sci.*, vol. 130, pp. 400–407, Jan. 2018.
[25] Y. Li, J. Zeng, S. Shan, and X. Chen, "Occlusion aware facial expression recognition using CNN with attention mechanism," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2439–2450, May 2019.
[26] T. Zhou, S. Canu, and S. Ruan, "Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 1, pp. 16–27, Mar. 2021.
[27] J. W. Crampton, "Platform biometrics," *Surveill. Soc.*, vol. 17, nos. 1–2, pp. 54–62, Mar. 2019.
[28] S. Yedulapuram, R. Arabelli, K. Mahender, and C. Sidhardha, "Automatic door lock system by face recognition," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 981, no. 3, Dec. 2020, Art. no. 032036.
[29] S. A. Hussain and A. S. A. A. Balushi, "A real time face emotion classification and recognition using deep learning model," *J. Phys., Conf. Ser.*, vol. 1432, no. 1, Jan. 2020, Art. no. 012087.
[30] P. Kaur, K. Krishan, S. K. Sharma, and T. Kanchan, "Facial-recognition algorithms: A literature review," *Med., Sci. Law*, vol. 60, no. 2, pp. 131–139, 2020.
[31] R. A. R. Ahmad, M. I. Ahmad, M. N. M. Isa, and S. A. Anwar, "Face recognition using assemble of low frequency of DCT features," *Bull. Electr. Eng. Informat.*, vol. 8, no. 2, pp. 541–550, Jun. 2019.
[32] F. Zhao, J. Li, L. Zhang, Z. Li, and S.-G. Na, "Multi-view face recognition using deep neural networks," *Future Gener. Comput. Syst.*, vol. 111, pp. 375–380, Oct. 2020.
[33] A. M. Jagtap, V. Kangale, K. Unune, and P. Gosavi, "A study of LBPH, eigenface, Fisherface and Haar-like features for face recognition using OpenCV," in *Proc. Int. Conf. Intell. Sustain. Syst. (ICISS)*, Feb. 2019, pp. 219–224.
[34] S. Suwarno and K. Kevin, "Analysis of face recognition algorithm: Dlib and OpenCV," *J. Informat. Telecommun. Eng.*, vol. 4, no. 1, pp. 173–184, Jul. 2020.
[35] S. Mittal, "A survey on optimized implementation of deep learning models on the NVIDIA Jetson platform," *J. Syst. Archit.*, vol. 97, pp. 428–442, Aug. 2019.
[36] D. S. AbdELminaam, A. M. Almansori, M. Taha, and E. Badr, "A deep facial recognition system using computational intelligent algorithms," *PLoS ONE*, vol. 15, no. 12, Dec. 2020, Art. no. e0242269.
[37] M. Kosinski, "Facial recognition technology can expose political orientation from naturalistic facial images," *Sci. Rep.*, vol. 11, no. 1, pp. 1–7, Jan. 2021.

WEI LI was born in January 1979. She received the Ph.D. degree in measurement technology and instrument from the Shenyang University of Technology, in 2013. Currently, she is a Lecturer with the Shenyang University of Technology. She has published six papers and has participated in six projects. She mainly engages in the research of machine vision and biometric identification.

JINBAO SUN was born in January 1979. He is a Senior Technical Expert with the Neusoft Group. He was in charge of many large projects, such as Neusoft digital campus products and Neusoft online learning products. He mainly engages in the research of e-learning and machine vision.

JING ZHANG was born in October 1994. She received the B.Sc. degree from the North University of China, in 2017. Currently, she is a Postgraduate Student with the Shenyang University of Technology. She mainly engages in the research of machine vision and biometric identification.

BOCHENG ZHANG was born in August 1999. He received the B.Sc. degree from the Hebei University of Technology, in 2021. Currently, he is a Postgraduate Student with the Shenyang University of Technology. He mainly engages in the research of machine vision and biometric identification.