Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 4, pp. 2463–2475, DOI 10.12694/scpe.v25i4.2870

INTELLIGENT EDUCATION MANAGEMNT SYSTEM DESIGN FOR UNIVERSITIES BASED ON MTCNN FACE RECOGNITION ALGORITHM

LIN LI* AND QI ZHANG†

Abstract. Face recognition technique has made significant advancements in security and attendance, but its application in teaching management is minimal. To address the issues of insufficient teacher resources and declining educational quality, the paper designs an intelligent education management system for colleges and universities based on improved Multi-task Cascaded Convolutional Neural Networks (MTCNN) face recognition. The purpose is to achieve accurate recognition of faces through improved facial recognition technology, thereby analyzing the attendance status of students, and improving the efficiency and quality of educational resource utilization. Firstly, an improved MTCNN facial recognition technology is adopted to achieve real-time monitoring of student status and attendance in the classroom through a B/S network structure. Secondly, through cluster deployment and load balancing, system stability and response speed can be improved. The results indicated that the improved MTCNN had better facial recognition accuracy and GPU utilization than traditional systems under different occlusion conditions. When there was no occlusion, recognition accuracy was 99.4%. However, when occlusion was presented at 10%, 20%, and 30%, the accuracy dropped to 92.3%, 84.25%, and 73.4%, respectively. Additionally, when the number of concurrent users was 1000, the maximum GPU utilization rate was 75%, which was 11% lower than traditional MTCNN systems. The use of an improved MTCNN facial recognition-based intelligent education management system in universities can effectively enhance the quality of classroom teaching and monitor the status of students. Further optimization of algorithm performance is needed in subsequent research to support larger-scale concurrent user usage while reducing hardware resource consumption.

Key words: Face recognition, MTCNN, Education management system, Classroom, Intelligence

1. Introduction. With the continuous development of the economy and the expansion of educational resources, many universities are facing a rapid increase in student numbers that far exceeds the supply capacity of high-quality teacher resources, leading to an imbalance in educational quality [9]. The modern education system requires teachers not only to impart knowledge but also to pay attention to the personalized development and ability cultivation of students, which puts higher demands on the teaching staff. Although live education platforms have emerged in recent years, the quality of teaching on education platforms varies, and online classes cannot monitor students' classroom status, resulting in poor overall student performance [19]. Some parents have to attend off-campus tutorials for the sake of their children's studies, and the quality of off-campus tutorials also varies, which cannot solve the current contradiction between the number of teachers and students. In this context, the rapid development of artificial intelligence technology has brought new solutions to the education field. For instance, intelligent homework evaluation systems and educational robots have shown significant improvements in learning efficiency [5]. The current intelligent education mainly includes intelligent homework assessment and intelligent educational robots. The homework assessment software utilizes algorithms to evaluate assignments and generate learning plans based on the results. Additionally, education robots interact with students to enhance their learning abilities [20]. The main development direction of artificial intelligence is face recognition, which is currently used in security, time and attendance, robotics smartphone unlocking, etc. [8]. Among them, face recognition-based face check-in technology is available because of its advantages of convenience and no manual check-in. This paper proposes an intelligent education management system that integrates Multi-task Cascaded Convolutional Neural Networks (MTCNN) facial recognition technology with traditional education to address the issue of declining education quality and the application of facial

*School of Energy, Power and Mechanical Engineering, North China Electric Power University, Baoding, 071000, China (lzx615@126.com)

†Department of Computer, North China Electric Power University, Baoding, 071000, China (Corresponding author, Zq_superland@163.com)

recognition technology. It aims to assist teachers in managing the classroom status of students, improve the quality of classroom education, and provide effective technical support and solutions for the education field.

2. Related Works. The world is entering a new era of intelligence, which has prompted the development of artificial intelligence. One important direction of this development is face recognition, making it a hot topic for domestic and international research. Basjaruddin N C and his team have designed an attendance system that integrates MTCNN face recognition, mask detection, and body temperature reading systems. This system addressed the limitations of feature extraction and recognition of faces wearing masks during the COVID-19 epidemic. Facial recognition technology was combined with mask detection technology to recognize the face, and a temperature reading system was used to detect body temperature, allowing for contactless attendance. The experimental results showed that the accuracy of facial recognition on the original dataset was between 92% and 100% [2]. Liu J proposed a face recognition model based on the MTCNN algorithm to address the issue of poor stability in dynamic videos with multiple angles and facial expressions. This model used Proposal Network (P-Net), Refine Network (R-Net), and Output Network (O-Net) network models to extract and process facial image features, and finally output the processed facial image. Through training experiments on the Wide Face database training set, the MTCNN algorithm outperformed SVM algorithm and traditional algorithms, achieving the requirements of facial image recognition [6]. Khan and Bhat proposed a face recognition attendance system based on neural networks to achieve intelligent student attendance. However, the system has limited applicability in large-scale concurrent users and diverse classroom environments. The system extracted image features through MTCNN and used SVM for image classification to achieve computer recognition of faces. Through simulation experiments, it has been proven that the system can accurately record attendance of attendees [10].

Dang T V et al. investigated the adaptability issues of user processing capabilities and different classroom scenarios. They applied the ArcFace model architecture of MobileNet V2 to deep Convolutional Neural Networks (CNN) and utilized highly discriminative feature technology to improve facial recognition, forming an effective deep learning model for facial recognition. The recognition accuracy of this model reached 97% and the processing speed is 25FPS [4]. Ramaraj P et al. proposed a face recognition technique based on self channel attention combined with self spatial attention to address the issue of adaptability under different lighting and pose changes. The original image was translated using a Gaussian filter and facial contours were recognized using edge detection algorithms. Self-channel attention was utilized to fuse feature maps in both channel and space. The experimental results showed that this face detection technology was superior to other face recognition models [13].

Sun et al. proposed an English pedagogy model to enhance students' learning efficiency. They utilized a decision tree algorithm to extract necessary data from a large amount of information. Then, they summarized and analyzed the data and patterns through neural networks to improve English performance and assist teachers in enhancing their education [14]. Wu C et al. addressed the problem that face detection cannot achieve fast detection and the problem of ensuring accuracy. They proposed a FaceNet face detection technique based on MTCNN. The MTCNN was used for fast alignment and detection of faces, and the face was verified and recognized by an optimized loss function FaceNet model. By comparison, the model outperformed other models with 99.85% recognition accuracy [18].

Ozdemir and Ugur addressed the challenge of identifying whether a student is present in distance education and proposed a remote attendance model based on a face recognition algorithm. They used filters for image processing to detect faces and trained the model in combination with an SQL server database. This model ensured that students could actively participate in learning and improve learning efficiency. Simulation experiments proved that the accuracy exceeded 80% [11].

Tata Sutabri et al. proposed a face recognition attendance system based on deep metric learning combined with neural network of nearest neighbor algorithm to address the problems of long time and missed signatures in traditional attendance. Deep metric learning was utilized to achieve face embedding. The nearest neighbor algorithm was then used to classify student faces, enabling computer recognition of faces and automatic attendance. Attendance records were then saved [15].

Pabba et al. propose a real-time monitoring system based on CNN for problems such as the difficulty of offline education teachers to control student participation and interactivity. The student states collected by

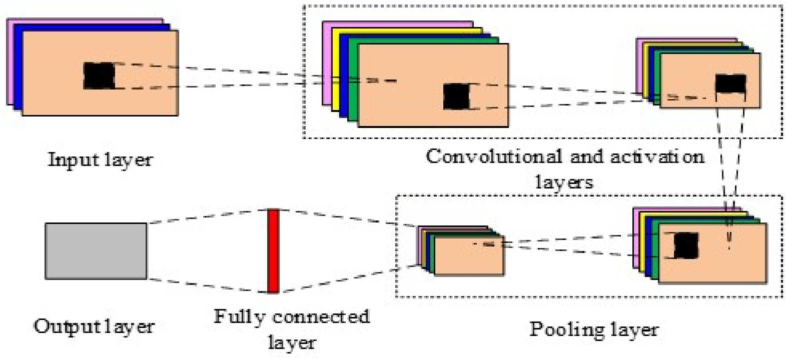

Fig. 3.1: CNN Structure

CNN were analyzed using frame-by-frame group participation estimation to automatically detect and analyze student participation and behavior. The accuracy of the model training reached 78.7% and the surveillance system was feasible [12].

In summary, many scholars have researched the use of facial recognition technique in the education industry, but there is still a need for improvement in the intelligent education management systems. Therefore, the study uses improved MTCNN face recognition combined with a new education management system to achieve an intelligent education management system, expecting to provide some help to the education business.

3. Design of an intelligent management system for education based on enhanced MTCNN facial recognition.

3.1. Improvement of MTCNN face detection algorithm. Currently, the network architecture has become more perfect, and in face recognition, researchers are commonly implementing by changing the loss function to enhance the precision of recognition [3]. CNN detection is the basis of face recognition algorithm. Input, convolutional, pooling, activation function, and fully connected layers typically make up a CNN. Fig. 3.1 indicates its structure diagram.

In Fig. 3.1, the input layer first normalizes and removes the mean of the original image, and then uses convolutional layers to capture features from the image and process them to output the dimension of the feature map. The obtained output image features are input into the pooling layer to sample the element map, and the activation function layer is used to solve nonlinear calculations. Through fully connected layers, different abstract image features are combined, and finally, the same image as the original image is output. The convolution layer serves as the foundation for the image's feature extraction. Its operation equation is shown in 3.1.

$$y_j' = f\left(\sum_{i\in M_j} y_j^{l-1} * k_{yj}^J + b_j^l\right) \tag{3.1}$$

In equation 3.1, y_l' represents the activation value of the output feature map j at the first layer after passing the activation function, M_j represents the input original map or the output feature map at the previous layer. $*$ represents the two-dimensional convolution operator. b_l' represents the bias value. $k_{ij}^{\dagger}$ represents the convolution kernel. The expression of the dimension of the convolutional output feature map is shown in 3.2.

$$\begin{cases} W_{at} = \frac{W_{in}-K+2P}{S} + 1 \\ H_{at} = \frac{H_{in}-K+2P}{s} + 1 \\ C_{ost} = C_{wn} \end{cases} \tag{3.2}$$

In equation 3.2, W_{at} denotes the output feature map width.H_{at} denotes the output feature map height.C_{ost} denotes the number of outputs.W_{in} denotes the input feature map width.H_{in} denotes the input feature map height. C_{wn} denotes the number of inputs. k denotes the convolutional kernel size. p denotes the number of boundary padding. s denotes the convolutional step size. The pooling layer can reduce the complexity and computation of the network, as shown in equation 3.3.

$$y_j^l = D\left(y_j^{l-1}\right) \tag{3.3}$$

In equation 3.3, the function $D()$ represents downsampling. The nonlinear problem is solved by the activation function, and the activation function equation is shown in equation 3.4.

$$\operatorname{Re} LU(x) = \begin{cases} x & ifx > 0 \\ 0 & ifx \leq 0 \end{cases} \tag{3.4}$$

In equation 3.4, x denotes the input value. By batch normalization process, the inputs of each layer are normalized to obtain the output normalized to a normal distribution of (0,1), and the computational equation is shown in equation 3.5.

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \xi}} \tag{3.5}$$

In equation 3.5, x_i denotes the input small batch of data.μ denotes the mean of the input values. σ^2 denotes the variance of the input values. ξ denotes the deviation. The data distribution learned in the previous layer is restored using equation 3.6.

$$y_i = \gamma \hat{x}_i + \beta \tag{3.6}$$

In equation 3.6, γ represents the stretching parameter and β represents the offset parameter. The output result is obtained by combining all the above output image features through the fully connected layer, and equation 3.7 displays the equation for the fully connected layer.

$$y = w x_i + b_j^i \tag{3.7}$$

In equation 3.7, is the weight, and the final image classification is achieved. The commonly used face detection algorithm in multi-person face recognition technology is MTCNN, which is chosen due to its multi-level network structure, including P-Net, R-Net, and O-Net. This network can recognize and correct facial information at different levels, and can also handle diverse lighting conditions in the classroom, various expressions and postures of students, and potential facial occlusion issues [1]. Firstly, after preliminary detection, the recognition results are gradually refined, and the recognition accuracy under partial occlusion is improved through key point localization. Secondly, O-Net can adapt to different sitting positions and head angles of students in the classroom based on actual situations, while maintaining high detection accuracy even when lighting conditions change. Fig. 3.2 displays the MTCNN face detection flowchart. The original image is processed by P-Net to generate a bounding box. The generated bounding box is input to R-Net, corrected by R-Net to generate a new bounding box. Then, the newly generated bounding box is corrected by O-Net to generate an exportable bounding box and key features of the face.

According to Fig. 3.2, MTCNN face detection must consider the varying distance of the face from the camera in the picture, which can result in different face sizes. Therefore, the picture needs to undergo resizing processing, as shown in equation 3.8.

$$\text{nextsize} = \text{originsize} * \left(\frac{12}{\text{min size}}\right) * \text{ factor }^n, n = \{1, 2, 3, \text{ L}, n\} \tag{3.8}$$

In equation 3.8, *originsize* denotes the initial size of the original image. *factor* denotes the scaling ratio. *nextsize* denotes the minimum value of the detected face. *nextsize* denotes the scaled image. After the

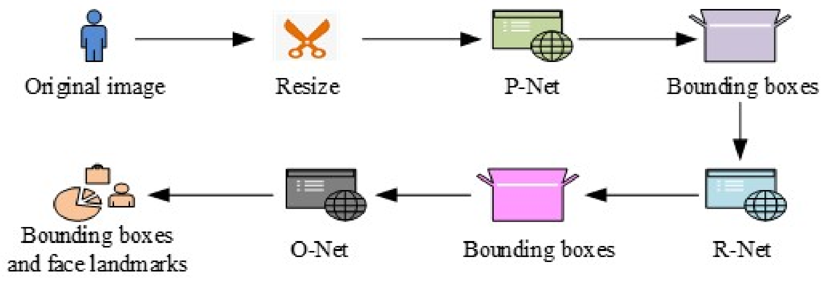

Fig. 3.2: MTCNN algorithm flowchart

cropped image is obtained, the image is classified using the loss function of cross-entropy, which is calculated as shown in equation 3.9.

$$L_i^{\text{det}} = -\left(y_i^{\text{det}} \log\left(p_i\right) + \left(1 - y_i^{\text{det}}\right)\left(1 - \log\left(p_i\right)\right)\right) \tag{3.9}$$

In equation (9), p_i denotes the predicted value. $y_1^{det} \in \{0, 1\}$ denotes the true value. To generate the bounding box of the face, by figuring out the difference between the current bounding box and the labeled bounding box, the difference is calculated as shown in equation 3.10.

$$L_i^{box} = \left\|\hat{y}_i^{box} - y_i^{box}\right\|_2^2 \tag{3.10}$$

In equation (10), $\hat{y}_i^{box}$ denotes the candidate frame of the model output. $y_i^{box} \in \mathbb{R}^4$ denotes the labeled bounding box. The face regions are corrected by the regression of the bounding boxes and merged using Non-Maximum Suppression (NMS), and the NMS formula is shown in equation 3.11.

$$G_i = \begin{cases} G_i, & \operatorname{iou}\left(M, q_i\right) < N_t \\ 0, & \operatorname{iou}\left(M, q_i\right) \geq N_t \end{cases} \tag{3.11}$$

In equation (11), G_i indicates the score of the bounding box. M indicates the box with the highest score at present. q_i indicates one of the remaining boxes. indicates the set queue. To avoid large fluctuations in the detection results, the problem is solved by the score reset function, which is shown in equation 3.12.

$$G_i = G_i e^{-\frac{iou\left(M, q_i\right)^2}{\sigma}}, \forall q_i \notin D \tag{3.12}$$

In equation (12), D denotes all bounding boxes. In the classroom environment, errors or omissions in MTCNN face detection may occur due to factors such as posture, angle, and facial occlusion by students or teachers. Facial marker localization is a key factor in generating faces. By comparing the differences between the annotation results and the model results, parameter fine-tuning is introduced in the O-Net stage, abandoning the facial alignment module. The calculation formula is shown in equation 3.13.

$$L_i^{\text{lamamar } k} = \left\|\hat{y}_i^{\text{lamdmark }} - y_i^{\text{landmark }}\right\|_2^2 \tag{3.13}$$

In equation 3.13, denotes the coordinates of feature points of the model output. denotes the coordinates of the labeled real position. Equations 3.9, 3.10, and 3.13 can be fused by an objective function, which is shown in equation 3.14.

$$\min \sum_{i=1}^{N} = \sum_{j \in \{\text{det}, box, \text{ landmark }\}} \alpha_j \beta_i^j L_i^j \tag{3.14}$$

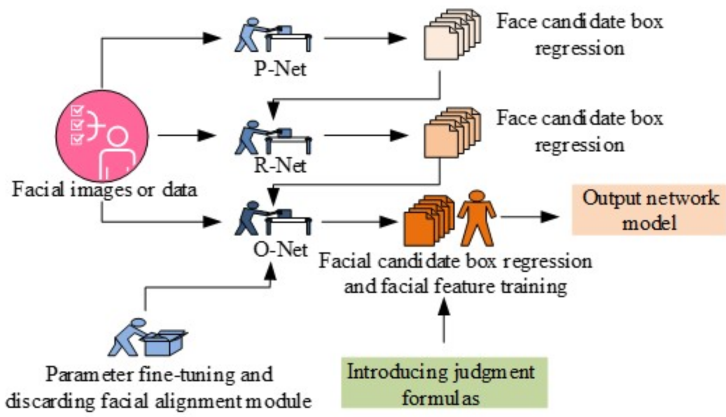

Fig. 3.3: Structure diagram of improved MTCNN algorithm

In equation (14), α_j and β_i^j denote super-parameters. After introducing the parameter fine-tuning module, the accuracy of face detection can be improved, but false detection still exists. As a result, the discriminant formula module is added out of the face feature point training frame, thus further improving the MTCNN structure. Equation 3.15 shows the discriminant formula.

$$f(p) = \begin{cases} 1 & p \geq 0.97 \\ 0 & p < 0.97 \end{cases} \tag{3.15}$$

In equation 3.15, p denotes the probability of face detection. When $f(p) = 1$, the result of detection is a face, and vice versa, When $f(p) = 0$ is not a face. Fig. 3.3 illustrates the MTCNN algorithm's enhanced structural diagram. This algorithm is based on the traditional MTCNN algorithm and adds parameter fine-tuning models at the O-Net structure. By selecting hyper-parameters, it ensures that the MTCNN model can adapt to the needs of different scenarios. Discriminant formulas are introduced in the facial feature point training box to determine whether the detected image is a human face. The selection of hyper-parameters is primarily accomplished through the following methods. First, a preliminary screening is conducted based on previous research and experimental experience. Next, a grid search method is used to systematically explore various combinations of hyper-parameters [17]. Then, the random search is used to explore valuable parameter combinations within a wider parameter space. Finally, Bayesian optimization methods are used to predict the performance of each set of parameters. After determining the hyper-parameters, the generalization ability and stability of these parameters on different subsets are verified through cross-validation.

3.2. Intelligent education management system design. The study adopts an improved MTCNN face detection algorithm to construct a face recognition model, and uses accuracy, recall, and F1 score evaluation indicators to evaluate the recognition performance of the model in occlusion or missed detection situations. The experiment was conducted on a machine containing an Intel i7 processor and 32GB of memory, operating system Windows 10. The main software used included TensorFlow 2.x and OpenCV 4.x. The study used the Megaface public dataset, which has a diverse range of facial images, including different poses, expressions, and lighting conditions [16]. In data preprocessing, the image is first converted to grayscale and resized to reduce the computational complexity of model training. Secondly, data augmentation techniques such as random rotation and cropping are used to increase the diversity of the dataset. Based on this face recognition algorithm, the intelligent education management system of colleges and universities is designed to realize the intelligence of education management. Fig. 4 illustrates the functional structure of the intelligent education management system.

Fig. 3.4 shows that the management system includes four main functions, such as classroom attendance, status monitoring, report analysis and background management. The classroom attendance module mainly

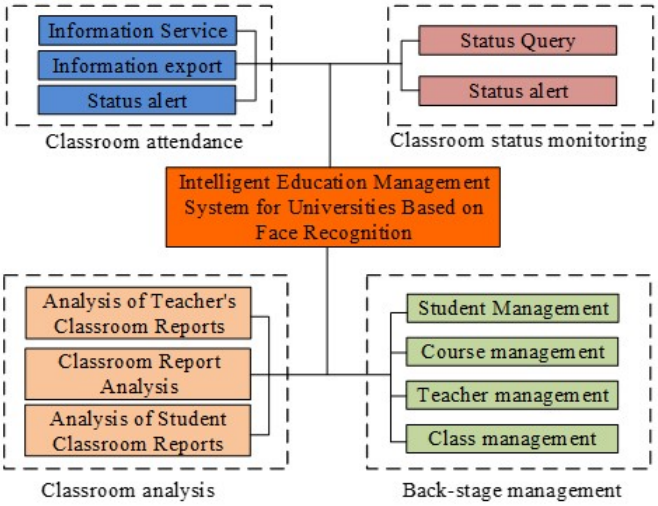

Fig. 3.4: Structure of intelligent education system in universities

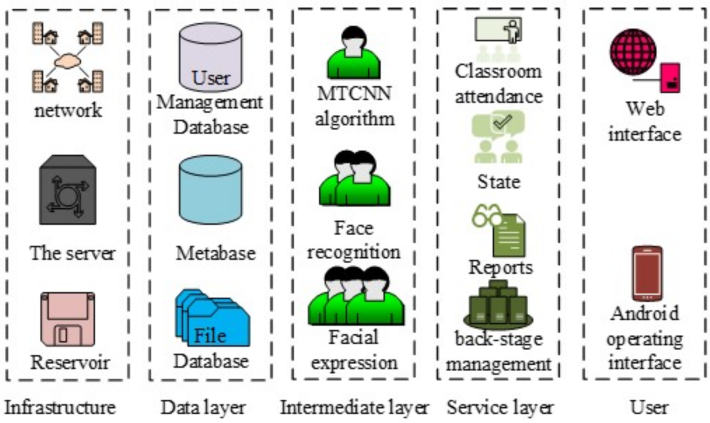

Fig. 3.5: Intelligent education management system architecture

includes attendance information inquiry, attendance information export and attendance status warning. The classroom analysis module is to analyze and save different kinds of classroom reports for each class, which mainly includes classroom, teacher and student report information. Users can send these reports to the objects they want to know. The background management module is divided into student, teacher, course and class management. The module consists of administrators and common users, wherein common users can only view, retrieve and other operations, only administrator users can modify or add content. In the design of intelligent education management systems, research is conducted on the use of B/S network structure models, as well as P-Net, R-Net, and O-Net network structures, to construct an education system that is easy to expand, manage, and maintain. The system's servers are cluster deployed while balancing the load on each server. The architecture of the intelligent education management system is shown in Fig. 3.5.

In Fig. 3.5, the intelligent education management system architecture design contains user layer, service layer, middle layer, data layer and infrastructure. The user layer is designed with two ports, Web side and Android side. Customers may sign in and use the education management system on different operating systems. The service layer is a service designed based on the division of the modules of the intelligent education

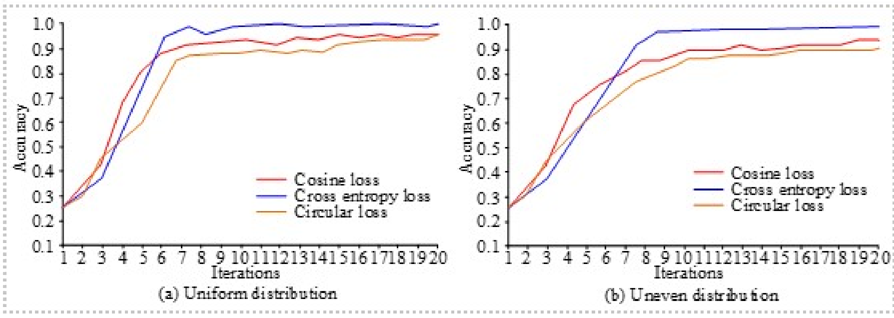

Fig. 4.1: Comparison of accuracy under different data distributions

management system, and this layer is mainly used for the Web server and APP side to handle the logic of each transaction. The middle layer comprises the enhanced MTCNN face detection algorithm, face recognition algorithm, and expression recognition algorithm. This layer enables the system to obtain facial data of both students and teachers, which is the core technology for the system to achieve intelligent face recognition. The data layer includes three parts: user management database, meta database and file database. This layer has storage function and mainly stores data files required for test data, students' and teachers' data, all algorithm model training, etc. Finally, the infrastructure includes three parts: network, server and disk. The system transmits data through the network and uses disk to store the server and data for system operation, which is the hardware support. MySQL is used to store the system's user, student and attendance information. A distributed file system is used to handle the files needed to store the system. Video streams are processed through Redis and sent to the algorithm model and data is cached. TensorFlow and OpenCV are used as the system's computational framework. SpringBoot is used to handle the low coupling between the module layers. To meet persistence operations, the database interacts with data through MyBatis. Application program interfaces communicate with each other through HTTP/HTTPS. Finally, Vue.js and Layui are used on the Web and App ends, which have the characteristic of easy scalability. Finally, nginx negative balance is used to achieve balanced optimization of the system and decrease the pressure of the intelligent education management system [7].

4. Intelligent education management system performance analysis.

4.1. Face recognition effect analysis. The study utilized publicly available WIDER FACE datasets and self-collected university classroom datasets. These datasets contained facial images captured in various lighting conditions and classroom scenes with occlusions. The experiment divided the dataset into 70% training set, 15% validation set, and 15% testing set. The performance of the model is evaluated using accuracy, recall, and F1 score to comprehensively assess the effectiveness of face detection and recognition. One of the key points of the improved MTCNN is alterations to the loss function. In the case of the same face recognition model, different loss functions are introduced to compare and analyze its accuracy. The accuracy comparison under different data distribution is shown in Fig. 4.1.

In Fig. 4.1 (a), the precision of the cross-entropy loss function reaches about 95.6% after 7 iterations of the data distribution and stabilizes between that value. Among them, the accuracy of the circular loss function tends to stabilize after 15 iterations with a value of 90.1%, and its accuracy is the lowest. In Fig. 4.1(b), the precision of the cross-entropy loss function is also the best when the data are not evenly distributed, with the accuracy reaching 93.4% and stabilizing after 9 iterations. The accuracy of the circular loss function is 84.7% and stabilizing after 17 iterations, which is the worst. It is proved that the face recognition accuracy of the quoted cross-entropy loss function is the best and has stronger generalization ability. In the same dataset, the accuracy recall of the traditional MTCNN and the improved MTCNN are compared under simple, medium and complex difficulty, respectively, as shown in Fig. 4.2.

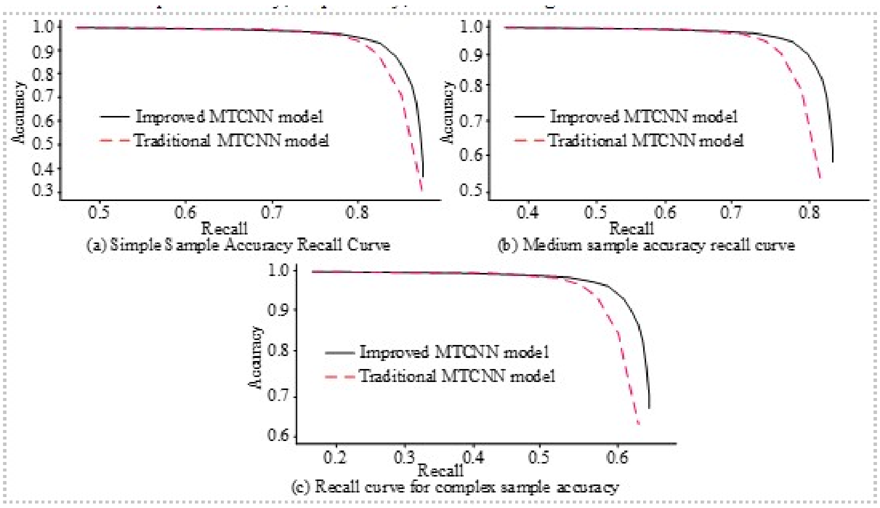

Fig. 4.2: Accuracy recall rate under different samples

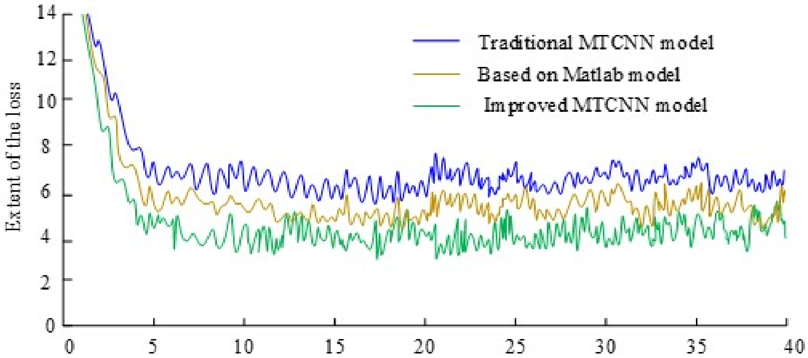

Fig. 4.3: Comparison of loss results between different models

In Fig. 4.2, the improved MTCNN outperforms the traditional MTCNN algorithm in terms of its recall rate under inter-single, medium, and complex dataset difficulties, and the accuracy is inversely related to the recall rate. In Fig. 4.2 (a), the precision of the improved MTCNN algorithm face recognition model decreases exponentially when the recall is 0.85 under the condition of simple dataset, and increases 0.08 compared to the recall of the traditional MTCNN. In Figure 4.2(b) and 4.2 (c), the accuracy of the improved MTCNN algorithm for face recognition decreases as the dataset difficulty increases, with a recall of 0.74 and 0.56, respectively. However, both still outperform the traditional MTCNN algorithm, indicating an improvement in accuracy. Further network training is performed on the algorithms, both based on the cross-entropy loss function. The comparison of the loss function results of different models is obtained as shown in Fig. 4.3.

From Fig. 4.3, the improved MTCNN face recognition system has the fastest loss function reduction. After 5 iterations, its loss function decreases to about 4 and stays up and down in the range of 4 to 6. The loss function curve of the Matlab face recognition system is similar to that of the improved MTCNN face recognition system, and the loss function remains stable between 5 and 7 after 7 iterations. The loss function of the traditional MTCNN face recognition system decreases the slowest, and the loss function tends to stabilize after 10 iterations and stays between 6 to 8. It indicates that the improved MTCNN face recognition system has the least curve fluctuation and its accuracy and stability are better than the other two models. The cross entropy loss function

Table 4.1: Comparison of recognition accuracy under different degrees of facial occlusion conditions

Algorithm type	Number of testers	Alignment rate/%	Unobstructed Recognition accuracy/%	Facial occlusion 10% recognition accuracy/%	Facial occlusion 20% recognition accuracy/%	Facial occlusion 30% recognition accuracy/%
Traditional MTCNN	1500	99.80%	97.33%	71.21%	62.30%	43.20%
Matlab	1500	99.87%	97.13%	68.60%	52.10%	31.20%
Dlib	1500	99.73%	98.67%	74.22%	60.81%	39.80%
Adaboost	1500	99.93%	96.40%	80.10%	69.20%	49.21%
Improving MTCNN	1500	100.00%	99.40%	92.30%	84.25%	73.40%
CNN-Based Approach	1500	99.85%	98.00%	98.32%	97.22%	72.32%
Ensemble Method	1500	99.90%	98.50%	98.50%	98.50%	82.22%
MobileNetV2	1500	99.78%	99.65%	97.80%	97.65%	75.89%
FaceNet	1500	99.82%	99.78%	98.20%	82.05%	78.27%

has a high penalty for misclassification, and can quickly guide the model to learn and distinguish different categories in early iterations. Therefore, the cross entropy loss function performs well on the Megaface dataset. The circle loss function, on the other hand, needs to undergo more precise adjustments and optimizations to adapt to complex facial recognition tasks.

The study compares the performance of the descent rate and stability of the loss function with the accuracy indicators of other systems. Under the same conditions, the improved MTCNN model achieves a recognition accuracy of 98.5% on the Megaface dataset, while the Matlab face recognition system achieves 97.8% and the traditional MTCNN system achieves 96.2%. This comparative result indicates that although the Matlab system is similar to the improved MTCNN in terms of stability of the loss function curve, there is still a gap in accuracy.

Table 4.1 indicates the contrasting of the recognition precision of each face recognition system in a classroom environment with different degrees of face occlusion.

Table 4.1 shows that the MobileNetV2 facial recognition system has the highest recognition accuracy of 99.65% when there is no occlusion. FaceNet follows with an accuracy of 99.78%, and the improved MTCNN has an accuracy of 99.4% under all test conditions. This is because the algorithm adopts feature learning and error correction methods when dealing with partially occluded facial features. However, Adaboost's facial recognition accuracy is 96.4%, with the worst performance. This is because the feature extraction and classification processes are highly sensitive to facial integrity and lack sufficient mechanisms to adapt or correct feature distortions caused by occlusion. The MTCNN face recognition accuracy is highest at 10%, 20%, and 30% face occlusion, with values of 92.3%, 84.25%, and 73.4%, respectively. In contrast, the Matlab-based face recognition system has the lowest recognition accuracy, with values of 68.6%, 52.1%, and 31.2%, respectively. Adaboost facial recognition has an accuracy of 96.40% in unobstructed conditions. However, its recognition rate drops to 49.21% when facial occlusion is 30%, indicating a significant decrease in performance when dealing with occlusion. Therefore, the improved MTCNN facial recognition system has better facial recognition accuracy than other systems in the presence and absence of occlusion. Even if students wear masks, the system can still effectively identify areas such as the eyes and forehead that are not covered by masks. This is because the system compensates for missing information in occluded areas by weighting key facial regions and utilizing contextual information.

4.2. Performance analysis of intelligent education system based on face recognition. This paper assesses the performance of an intelligent education administration system that utilizes enhanced MTCNN face recognition technology. Table 4.2 displays the system's environmental configuration.

In Table 4.2, the MTCNN face recognition college intelligent education management system is compared to the traditional MTCNN face recognition model and Matab face recognition model in terms of system response speed under different concurrent user conditions. The comparison of the response time of the three education management systems is shown in Fig. 4.4.

Table 4.2: System environment configuration

Name	System Configuration
GPU version	NVIDIA GTX 1080 Ti
Tensorflow version	1.12.0
Opencv	3.4.2
Python	3.6
Java	1.7
System memory, hard disk size	16GDDR, 1Thard disk
Database	MySQL
GPU operating system	Ubuntu 14.0

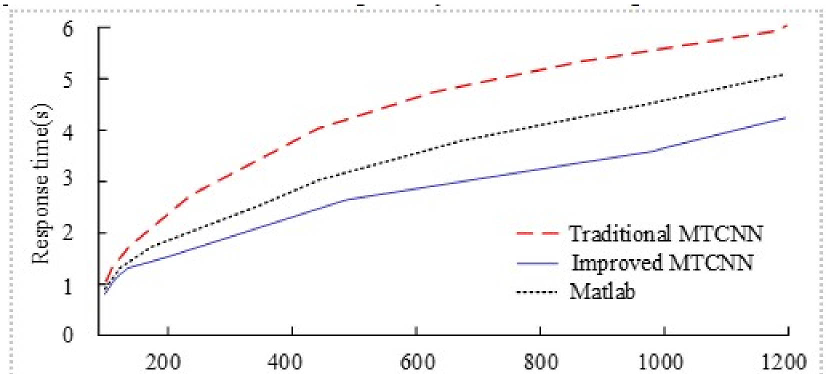

Fig. 4.4: Comparison of response times of three education management systems

In Fig. 4.4, each education management system's response time is longer as the users increases. The response time of the intelligent education management system with MTCNN face recognition is the slowest, taking about 1.48 seconds when there are 200 users. In comparison, the Matlab-based face recognition education management system takes 1.85 seconds and the traditional MTCNN face recognition education management system takes 2.3 seconds with 200 users. MTCNN's response speed is the fastest, and the traditional MTCNN's response speed is the slowest. Additionally, enhancing the MTCNN facial recognition education management system will encourage students to initiate videos and keep track of those who have not yet done so. If a student fails to open the video within the specified time, the system will automatically send a reminder to both the student and the teacher. Secondly, teachers can manually mark the attendance status of students through the system backend and judge their classroom participation based on their voice participation. Therefore, the improved system demonstrates performance stability, system architecture optimization, and efficient resource management in high concurrency environments. Fig. 4.5 illustrates a comparison of the GPU utilization rate between the traditional MTCNN face recognition education management system and the improved MTCNN face recognition education management system with varying numbers of concurrent users.

From Fig. 4.5 (a), the GPU utilization rate of both systems is almost the same when the users are 200. From Fig. 4.5 (b), when the users are 500, the GPU utilization rate of the educational management system for traditional MTCNN face recognition is greatly increased, reaching a maximum of about 62%, while for the improved MTCNN system, the maximum GPU utilization rate is 54%. In Fig. 4.5 (c), when the users increase to 1000, the GPU utilization rate of the improved MTCNN system is 75% at maximum, while the GPU utilization rate of the traditional MTCNN system reaches 86%. The improved MTCNN system saves GPU memory during runtime and can withstand more users using it simultaneously. It improves the scalability and reliability of the system while reducing the cost of optimizing resource utilization.

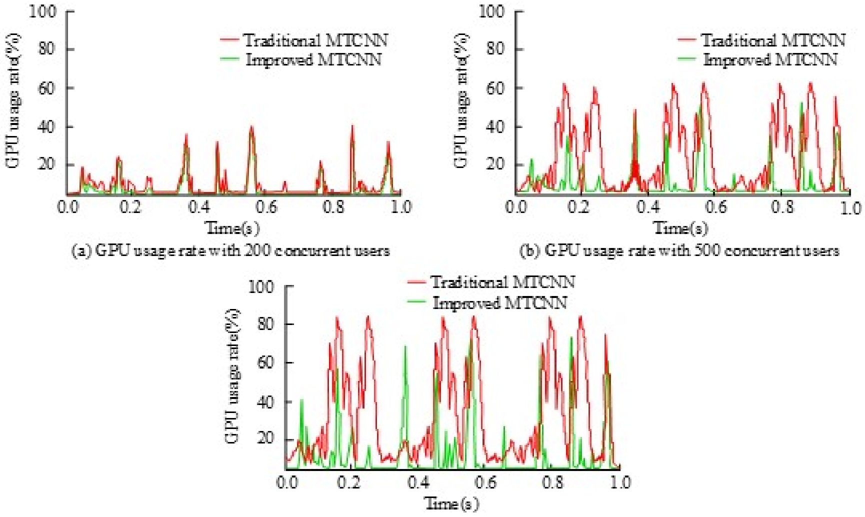

Fig. 4.5: Comparison of GPU usage between traditional MTCNN and improved MTCNN with different concurrent users

5. Conclusions. In university management, the conventional education management platform can only facilitate regular teaching management. It cannot monitor students' learning status and quality, and relies solely on teachers' in-class management. To enhance the quality of classroom instruction for kids, the study introduced parameter fine-tuning model and discriminant formula for optimization on the basis of traditional MTCNN face recognition. Moreover, the improved MTCNN face recognition was introduced into the education management system for design to realize the intelligence of education management system. The results indicated that the loss function of the improved MTCNN face recognition system decreased the fastest to 4 and kept fluctuating between 4 and 6, which was better than the other two systems. The recognition accuracy of the improved MTCNN face recognition system was 99.4%, 92.3%, 84.25% and 73.4% when there was no occlusion, 10% occlusion, 20% occlusion and 30% occlusion. The system could effectively monitor student attendance and participation in the classroom, providing accurate data support for teachers. This allows for timely adjustments to teaching strategies and classroom activities, enhancing student learning motivation and classroom interaction, ultimately improving the quality of education. The improved MTCNN facial recognition university intelligent education management system had a maximum GPU utilization rate of 75% when the number of concurrent users was 1000, which was 11% lower than the traditional MTCNN system. This fully demonstrates the significant advantages of the system in saving computing resources and supporting large-scale user management. It is evident that the system can enhance both the efficiency and quality of classroom management while also aiding teachers in promptly adjusting their teaching strategies based on students' real-time classroom performance. However, the performance of the system in facial recognition of different races, ages, and expressions has not been explored, and it is still difficult to completely avoid the problem of dataset bias, which can easily lead to recognition errors. Secondly, as the number of students using the system increases, the utilization rate of GPU resources increases, which is a key constraint on the performance of its computing resource system, making it difficult to maintain high system performance. Future research will add more datasets, including facial images of different races, age groups, and expressions, to enhance the universality of the system in different environments. By combining image processing technology to assist facial recognition, the robustness of the system in extreme occlusion situations can be improved. In addition, student learning assessment data is included in the system to build a comprehensive educational analysis system, providing teachers with more comprehensive teaching feedback and helping them understand students' learning progress and situation.

REFERENCES

[1] K. T. Atanassov, *New topological operator over intuitionistic fuzzy sets*, Journal of Computational and Cognitive Engineering, 1 (2022), pp. 94–102.

[2] N. C. Basjaruddin, E. Rakhman, Y. Sudarsa, M. B. Z. Asyikin, and S. Permana, *Attendance system with face recognition, body temperature, and use of mask using multi-task cascaded convolutional neural network (mtcnn) method*, Green Intelligent Systems and Applications, 2 (2022), pp. 71–83.

[3] Z. Chen, *Research on internet security situation awareness prediction technology based on improved rbf neural network algorithm*, Journal of Computational and Cognitive Engineering, 1 (2022), pp. 103–108.

[4] T.-V. Dang, *Smart home management system with face recognition based on arcface model in deep convolutional neural network*, Journal of Robotics and Control (JRC), 3 (2022), pp. 754–761.

[5] W.-Y. Hsu, H.-J. Hsu, Y.-Y. Wang, and T. Wang, *Multilayer-neighbor local binary pattern for facial expression recognition*, Intelligent Systems in Accounting, Finance and Management, 29 (2022), pp. 156–168.

[6] J. Liu, *Research on video image face detection and recognition technology based on improved mtcnn algorithm*, International Journal of Wireless and Mobile Computing, 22 (2022), pp. 205–212.

[7] K. Liu, M. Zhang, and M. K. Hassan, *Intelligent image recognition system for detecting abnormal features of scenic spots based on deep learning*, Journal of Intelligent & Fuzzy Systems, 39 (2020), pp. 5149–5159.

[8] N. Liu, H. Liu, and H. Liu, *Mental health diagnosis of college students based on facial recognition and neural network*, Journal of Intelligent & Fuzzy Systems, 40 (2021), pp. 7061–7072.

[9] D. T. Long, *A facial expressions recognition method using residual network architecture for online learning evaluation*, Journal of Advanced Computational Intelligence and Intelligent Informatics, 25 (2021), pp. 953–962.

[10] M. Mohd, R. Jan, and M. Shah, *Text document summarization using word embedding*, Expert Systems with Applications, 143 (2020), p. 112958.

[11] D. Ozdemir and M. E. Ugur, *Model proposal on the determination of student attendance in distance education with face recognition technology*, Turkish Online Journal of Distance Education, 22 (2021), pp. 19–32.

[12] C. Pabba and P. Kumar, *An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition*, Expert Systems, 39 (2022), p. e12839.

[13] P. Ramaraj, *A neural network in convolution with constant error carousel based long short term memory for better face recognition*, Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12 (2021), pp. 2042–2052.

[14] Z. Sun, M. Anbarasan, and D. Praveen Kumar, *Design of online intelligent english teaching platform based on artificial intelligence techniques*, Computational Intelligence, 37 (2021), pp. 1166–1180.

[15] T. S. Tata Sutabri, P. Pamungkur, A. K. Ade Kurniawan, and R. E. S. Raymond Erz Saragih, *Automatic attendance system for university student using face recognition based on deep learning*, International Journal of Machine Learning and Computing, 9 (2019), pp. 668–674.

[16] X. Wang, M. C. Fairhurst, and A. M. Canuto, *Improving multi-view facial expression recognition through two novel texture-based feature representations*, Intelligent Data Analysis, 24 (2020), pp. 1455–1476.

[17] N. Wenyan, *Construction of network open teaching platform of analytical chemistry based on facial recognition and artificial intelligence*, Journal of Intelligent & Fuzzy Systems, 40 (2021), pp. 7435–7445.

[18] C. Wu and Y. Zhang, *Mtcnn and facenet based access control system for face detection and recognition*, Automatic Control and Computer Sciences, 55 (2021), pp. 102–112.

[19] Y. Yang and X. Song, *Research on face intelligent perception technology integrating deep learning under different illumination intensities*, Journal of Computational and Cognitive Engineering, 1 (2022), pp. 32–36.

[20] M. Yu, H. Zheng, Z. Peng, J. Dong, and H. Du, *Facial expression recognition based on a multi-task global-local network*, Pattern Recognition Letters, 131 (2020), pp. 166–171.

Edited by: Mudasir Mohd
Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions
Received: Nov 17, 2023
Accepted: May 6, 2024