The Application of Face Recognition Model Based on MLBP-HOG-G Algorithm in Smart Classroom

Xiaoxia Li
College of Artificial Intelligence and Big Data, Zibo Vocational Institute, Zibo, 255000, China

***Abstract*—The development of Internet and Internet of things technology has accelerated the informatization construction of smart education. But the traditional face recognition algorithm used in smart classrooms inevitably has problems such as large amount of calculation, obvious resource and memory consumption, and poor recognition accuracy. In order to promote the informatization construction of colleges and universities and the accuracy of face recognition, a face recognition model based on multi-feature Local Binary Pattern Directional Gradient Histogram Gabor Filter algorithm is proposed. The model first extracts the binary texture image, and then carries out secondary feature extraction, dimension reduction processing and serial fusion with the gray level co-occurrence matrix feature weighting to improve the recognition accuracy. The results show that the recognition rate of the proposed method in ORL database, CMU_PIE database and Yale database can reach 95%, 94.12% and 93.33%, which is better than other algorithms. And in the comprehensive data set, the training and verification recognition accuracy of the proposed method for face recognition is basically 98% and 97.23%, which has good generalization and stability, and its cumulative error result of face key point detection is less than that of other comparison methods. The proposed method can provide new opportunities and possibilities for the application effect of face recognition, smart classroom construction and teaching development.**

***Keywords*—Multi feature local binary pattern; directional gradient histogram; Gabor filter; face recognition; smart classroom**

I. Introduction

With the development and promotion of information technology, smart classrooms utilize technologies such as artificial intelligence and big data to monitor and analyze classroom teaching in real-time, providing teaching feedback and personalized learning services for teachers and students [1]. As an important foundational technology for intelligent classrooms, student facial recognition can achieve functional statistics and analysis of student attendance, classroom performance, and other content. Wang et al. proposed a facial recognition intelligent education system based on MTCNN and FaceNet models. The results show that the accuracy of the system in both facial recognition and student emotion recognition performance is over 90% [2]. Dang T V scholar proposed an improved facial recognition model architecture based on the MobileNetv2 backbone network to achieve facial recognition. The results show that the accuracy of this depth method exceeds 95% on small datasets of original face images [3]. However, facial recognition faces many challenges in different classroom environments, such as complex classroom layouts and a large number of interactive devices that may cause changes in lighting, shadows, and reflections. Different quantities, scales, multi pose faces, face occlusion, and other factors can lead to lower detection and accuracy rates in object recognition [4]. The existing facial recognition technology is difficult to meet the needs of real-time processing and teaching recognition. For example, traditional methods such as principal component analysis and linear discrimination extract information from facial region images. The classic local feature extraction methods require a large amount of computation, and some deep learning methods are prone to losing control over recognition performance in unconstrained environments. Therefore, in response to these issues and shortcomings, a gradient oriented Gabor (MLBP-HOG-G) algorithm based on multi feature local binary pattern histogram is proposed for face recognition model. This method overcomes the shortcomings of traditional methods in complex environments. Compared with some traditional algorithms designed based on rules or fixed features, it can learn features through adaptive methods, improving adaptability to new environments and different student groups. Compared with traditional facial recognition technology, this fusion strategy achieves more comprehensive feature capture and solves problems such as lighting changes, facial occlusion, and pose changes. The face recognition model based on MLBP-HOG-G algorithm has unique advantages in feature diversity, robustness, adaptability, and performance improvement. It can provide more effective solutions and references for the application fields of intelligent classroom safety management, student behavior monitoring, and intelligent teaching.

The research mainly analyzes the application of facial recognition models in smart classrooms from four aspects. Section I is a literature review and discussion of facial recognition technology in current smart classroom applications. Section II is to design the MLBP-HOG-G algorithm to achieve smart classroom facial recognition, including feature extraction, weighted combination gray level co-occurrence matrix design, and construction of cascaded classifiers. Section III is to test and analyze the application effect of this feature recognition model. Section IV is an overview summary of the entire text.

II. Related Work

In a large-scale educational environment, student management and supervision are extremely challenging tasks. Educational institutions must effectively track students' attendance, participation, and behavior to ensure their safety and academic progress. This challenge has driven the demand for more efficient and intelligent student management

methods. Applying facial recognition technology to recognize and track individuals has become a major leap in smart classrooms, and some scholars have conducted a series of related studies on this topic. Researchers such as Niu proposed a feature fusion method with channel attention networks, aiming to fully utilize a limited number of hyperspectral samples for deep learning training. The experiment showcases that this method could markedly reduce storage space and computational overhead while still maintaining competitive accuracy and efficiency. These characteristics also indicate that this method has broad applicability on edges and mobile devices [5]. Widjaya and other researchers have proposed a random challenge response authentication method aimed at addressing the vulnerability of commercial facial recognition engines. This method is based on activity detection and is designed to protect against deception attacks, photo attacks, and video attacks. The experiment illustrates that the accuracy of this method is 99%, with an F-value of 98.99%. This study verifies the effectiveness of random challenge response authentication in resisting photo and video attacks in FR and anti-deception [6]. Scholars such as Nam presented a FR method that combines deep learning and binary patterns, for growing the accuracy of FR in high noon conditions. Experiments indicate that this method has high effectiveness and applicability when facial images are incomplete [7].

Fan et al. presented a Sprinter FR algorithm on the ground of sliding data camera measurement, aiming to solve the problems of low accuracy in facial key point recognition and noise errors in recognition. The experiment showcases that the algorithm successfully detects and recognizes six key points of the face, with a noise error of less than 1.3%, achieving the established goal and possessing practical application value [8]. MLBP, as a commonly used computer vision algorithm for FR, can extract local texture features of images. Therefore, it has wide applications in fields such as facial feature extraction, facial detection, facial expression recognition, facial authentication and recognition, and live body detection. Wang and his collaborators proposed a method for extracting texture features. This method utilizes multi-scale and multi-directional local binary patterns, aiming to classify hyperspectral images through a small number of labeled samples. The experiments indicate that this method could more markedly extract texture features and further strengthen the classification of hyperspectral images by combining it with the guidance of hyperpixel segmentation maps for decision-making [9]. Kaplan et al. proposed a multi-scale accessibility configuration file aimed at describing the multi-scale accessibility levels of various cities. Experiments have shown that there is an inherent correlation between universal accessibility at different scales and urban performance [10]. Considering that most studies have not integrated various visual cues such as facial expressions and body posture, Pabba C et al. proposed using OpenPose and Py Feat frameworks to extract multiple features and perform classification recognition under a cascaded neural network architecture. The results show that this method can effectively recognize students' facial features and behaviors, with an accuracy rate of over 90% [11]. El Mashad Y et al. used video facial recognition technology to implement smart classrooms, which can recognize individuals under different lighting conditions and facial expressions. The results indicate that this method has smaller errors and higher classification accuracy [12]. Yuan Z et al. proposed a face detection algorithm based on an improved YOLOv5, which introduces CSPDarknet53 backbone network, loss function, and self-attention mechanism modules to improve detection performance. The results show that the accuracy of this method for face detection exceeds 85%, and the detection accuracy in simple scenes exceeds 95% [13]. Aly M scholar attempted to use facial expression recognition techniques such as Residual Network with 50 layers (ResNet50), Convolutional Block Attention Module (CBAM), and Temporal Convolutional Network (TCN) to track students' classroom performance. The results indicate that this combination method is helpful in capturing facial expressions and monitoring learning behaviors [14].

Channel Attention Network Feature Fusion (Niu JY) can extract important features and reduce computational overhead, but it is difficult to adapt to high resource environments. The Random Challenge Response Authentication Activity Detection Method (Widjaya C) enhances the security of facial recognition, but it requires additional hardware support and computing resources. The face recognition method combining deep learning and binary patterns (NAM V-H) can still maintain high effectiveness and applicability in the case of incomplete facial images, but it relies heavily on data and has a high computational cost. The Sprinter facial recognition algorithm (fan y) has good noise control performance and accurate keypoint recognition, but it has a significant dependence on specified parameters. Multi scale and multi-directional local binary mode (Liguo Wang) can achieve classification of hyperspectral images, but it has fewer labeled samples. From the above content, it can be seen that using only attention mechanisms for feature fusion is difficult to ensure the comprehensiveness of information selection. Single thinking perspectives based on feature extraction (Nam v h, Liguo Wang) are inevitably affected by computational costs, resource constraints, environmental differences, and so on. The application of previous methods in facial recognition has limitations such as single feature extraction, insufficient robustness to complex scenes, high computational complexity, and storage overhead. The research proposes using the MLBP-HOG-G algorithm to recognize faces, and its multimodal combination approach can improve feature extraction ability and recognition accuracy. And this method utilizes the gray level co-occurrence matrix feature weighting method to perform secondary processing and dimensionality reduction on the extracted features, reducing the interference of noise and redundant information, effectively solving the limitations of previous research. This model not only improves the accuracy and efficiency of facial recognition, but also enhances its practical application ability in complex environments of smart classrooms, providing strong technical support for the informationization construction of universities.

III. Method Design for MLBP-HOG-G FR Model in Smart Classroom

This study focuses on facial recognition algorithms, including image processing, feature extraction, classification algorithms, and system design. The study used MLBP and

HOG to extract features, and then weighted combined grayscale co-occurrence matrix features to form MLBP-HOG-G features. It conducts facial recognition experiments through a classifier and constructs a SVM-KNN cascade classifier. Finally, it uses MATLAB GUI tools to design a facial recognition system, including identity verification functions.

A. Analysis of MLBP Based Facial Recognition Algorithm

FR is an identity recognition method achieved through computer vision technology. In the development process of smart classrooms, facial recognition technology can markedly enhance the operational efficiency of schools and decrease the workload of faculty and staff. The general process includes steps such as data collection, preprocessing, feature extraction, feature matching and storage, discrimination and decision-making, and feedback of recognition results. The recognition process is shown in Fig. 1.

In Fig. 1, the facial recognition process mainly includes three parts: facial image preprocessing, facial detection, and facial recognition. It first collects facial data and extracts discriminative features after preprocessing. Then it is matched with known features and the identity is determined on the ground of the matching results. Finally, it provides corresponding feedback on the ground of the recognition results. When describing facial recognition features, LBP has become one of the commonly used feature descriptors in the field of facial recognition due to its texture representation, invariance, dimensionality reduction, and high computational efficiency. The LBP operator was initially defined in a 3x3 pixel window, consisting of a central pixel and its 8 adjacent pixels. To represent the LBP operator, the function E can be used to represent the joint distribution function of the central pixel and adjacent pixel points. The calculation is showcased in Eq. (1).

$$E = t(g_c, g_0, ..., g_{p-1}) \quad (1)$$

In Eq. (1), g represents the pixel at the center; $g_0, g_1, ..., g_{p-1}$ are the eight surrounding pixels. By comparing the Pixel Values (PVA) of the center pixel of the window with the PVA of its surrounding 8 adjacent points, the joint distribution function of the disparity in the PVA of the center point and the PVA of the surrounding eight adjacent points can be used to describe the characteristics of the region. This study assumes that the PVA of the central pixel has little impact on the loss or impact of the texture feature information of the image, mainly affecting the brightness of the image. Therefore, the PVA of the center pixel can be ignored, and the joint distribution function of simplified texture features can be expressed as Eq. (2).

$$E \approx t(g_c - g_0, ..., g_{p-1} - g_c) \quad (2)$$

The above function describes the texture distribution of each pixel in the domain. Generally speaking, prominent texture features in texture distribution that cannot directly observe numerical features can be converted into binary features through the LBP algorithm. The local binary mode compares the grayscale values of a certain pixel in the image with neighboring pixels one by one, as shown in Fig. 2.

In Fig. 2, (a) is a 3x3 template, in which the grayscale value of the central pixel is used as the threshold. If the values of 8 pixels in the neighborhood are greater than or equal to this threshold, then the values of these pixels are set to 1, otherwise 0. Next, it starts from a starting point and sets the weights of each pixel in a clockwise direction, as shown in Fig. 2(c). Then, it converts the binary numbers around the center pixel into decimal numbers for obtaining the LBP value of the center pixel, as shown in Fig. 2(d). This process can be represented by Eq. (3).

$$LBP_{pj} = \sum_{s=1}^{8} t(p_s - p_j) \times 2^{s-1} \quad (3)$$

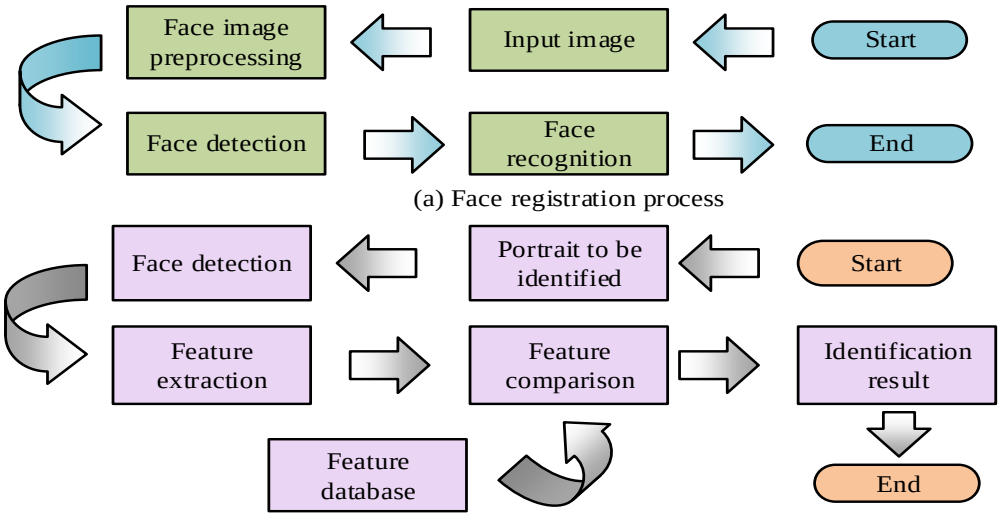
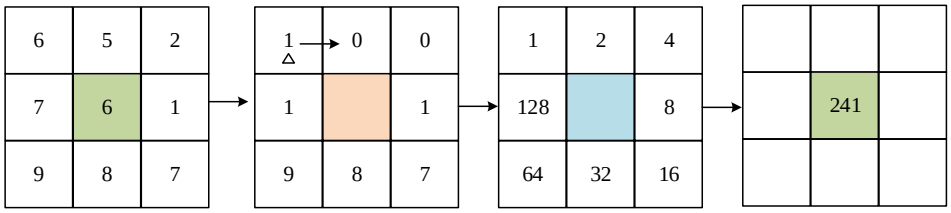


Fig. 1. FR process.



(a) 3X3 mobile template (b) Binary conversion (c) Weights (d) LBP eigenvalue

Fig. 2. LBP pixel comparison diagram.

In Eq. (3), s represents the eight nearest pixel points around the marked center pixel; p_s is the value of the pixel; p_j is the central pixel point; $t(r)$ is a symbolic function. To improve the LBP algorithm, MLBP introduced variance testing. It first calculates the variance of nine pixels within a 3x3 template to understand the fluctuations in PVA. When the variance is small and the texture is relatively smooth, MLBP uses the average of the maximum and minimum values of eight pixels around the center pixel as the threshold to prevent the loss of detail features. When the variance is large, the texture changes greatly. MLBP uses the median of nine pixels as the threshold to reduce noise interference in LBP calculation, and then recalculates the LBP code to update the texture features. This approach improves the LBP algorithm and better adapts to different image situations. It constructs a 3x3 template and calculates the variance V of the nine pixels in the template. The formula for calculating the variance is shown in Eq. (4).

$$V = \frac{1}{9}\sum_{j=1}^{9}(M - P_j)^2 \tag{4}$$

In Eq. (4), P_j represents the average value of nine pixels; P_j represents the PVA of nine templates, and the calculation is demonstrated in Eq. (5).

$$M = \frac{1}{9}\sum_{j-1}^{9} p_j \tag{5}$$

The principal component analysis method is used to analyze multivariate data, identifying the most important variables by calculating weights. In multivariate analysis, with the variables grows, the complexity of the problem grows, so reducing variables is necessary to reduce computational complexity [15-17]. In the calculation process, if N is defined as the quantity of samples and the vector dimension is M, then the sample set can be represented as N vectors $X1, X2, X3, ..., XN$. Each vector Xi represents the i-th sample. Next, the study can use these samples to calculate the covariance matrix of the training samples, and the specific formula is indicated in Eq. (6).

$$V_t = \frac{1}{N}(X - \overline{X})(X - \overline{X})^T \tag{6}$$

In Eq. (6), $\overline{X} = [\eta, \eta, ..., \eta]$, η represents the mean values of all samples. The calculation is shown in Eq. (7).

$$\eta = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{7}$$

This study calculates the eigenvalues of the covariance matrix（λi，$1 \le i \le m$) and eigenvectors（ωi，$1 \le i \le m$). These eigenvalues and eigenvectors represent the principal components of the data. For the original dataset X, the calculation is indicated in Eq. (8).

$$Y = W^T(X - \eta) \tag{8}$$

In Eq. (8), to reduce the dimension to K dimension, simply select the first K row of Y. This study used the PCA algorithm for dimensionality reduction, and then input the reduced feature vectors into a simple K-nearest neighbor classifier. The K nearest neighbor classifier counts the range in the test sample and each training sample, and determines the classification of the test sample on the ground of the labels of the K closest samples. The relevant details are showcased in Fig. 3.

In Fig. 3, the value of K is set to 5. The classifier uses template matching and experimental parameters to classify each test sample. After classification is completed, the data that is successfully matched is counted, and then the RR is counted for evaluating the performance.

B. Design of FR Algorithm Based on MLBP-HOG-G

Histogram of Oriented Gradients (HOG) is a commonly utilized algorithm for describing local texture features of images. HOG expresses local features on the ground of the direction and density distribution of gradients in the image, generates histograms through statistical gradient information, and then combines these histograms into feature vectors. This feature vector can be used for various image tasks, such as facial recognition. HOG feature extraction includes the following steps. Firstly, it performs grayscale processing on the input image and converts it into a grayscale image. Next, it uses the gamma correction method to normalize the grayscale image, which helps to decrease the interference of local shadows, lighting changes, and noise on feature extraction. The gamma correction formula is shown in Eq. (9).

$$H(x, y) = H(x, y)^{gamma} \tag{9}$$

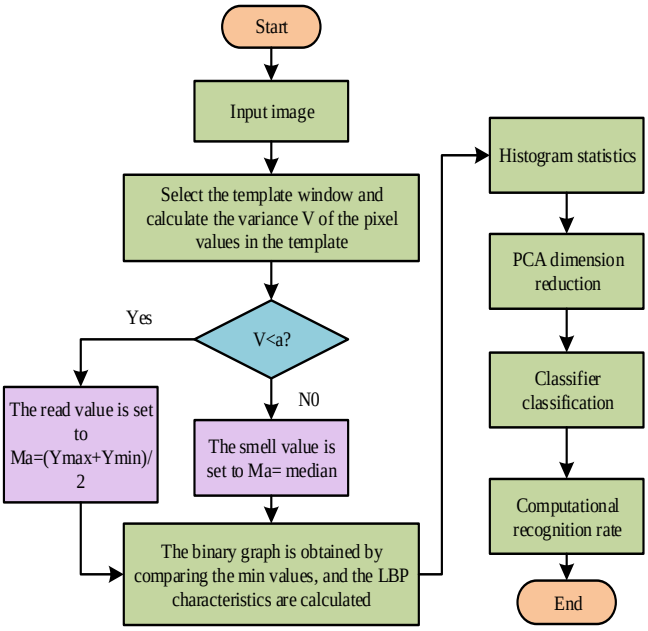


Fig. 3. Flow chart of FR algorithm on the ground of MLBP.

In equation (9), the Gamma value is set to 5. When counting the horizontal gradient $G(x)$ and vertical gradient $G(y)$ of the image, two templates [-1,0,1] and [1,0,1] were utilized for performing convolution operations on the image, respectively. This obtains the gradient direction value for each pixel position. The calculation is showcased in equation (10).

$$\begin{cases} G_x(x,y) = H(x+1,y) - H(x-1,y) \\ G_y(x,y) = H(x,y+1) - H(x,y-1) \end{cases} \quad (10)$$

In equation (10), $G_x(x,y)$ serves as the horizontal gradient at point (x,y). $G_y(x,y)$ represents the gradient in the vertical direction, and $H(x,y)$ serves as the PVA. It continues to calculate the gradient value, as shown in Eq. (11).

$$\begin{cases} G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \\ a(x,y) = \tan^{-1}\left(\dfrac{G_y(x,y)}{G_x(x,y)}\right) \end{cases} \quad (11)$$

In Eq. (11), $G(x,y)$ is the gradient amplitude of the input image at pixel (x,y). $a(x,y)$ is the gradient direction of the graph at pixel (x,y). On the ground of gradient amplitude and directional weight projection, this algorithm divides an image of 64 * 128 size into multiple 2 * 2 cells, each containing 8 * 8 pixels. By scanning the image in steps of eight pixels, the gradient values of the pixels are divided into nine directional ranges, each occupying 40°. This process calculates features on the ground of the weight projection of gradient amplitude and direction. The gradient bin averaging diagram is shown in Fig. 4.

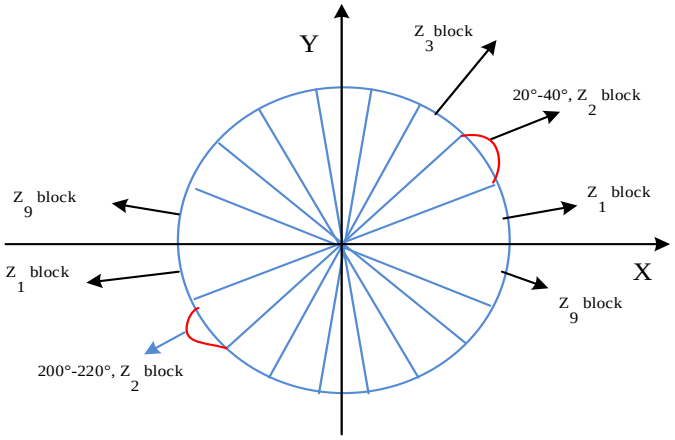


Fig. 4. Gradient bin equalization diagram.

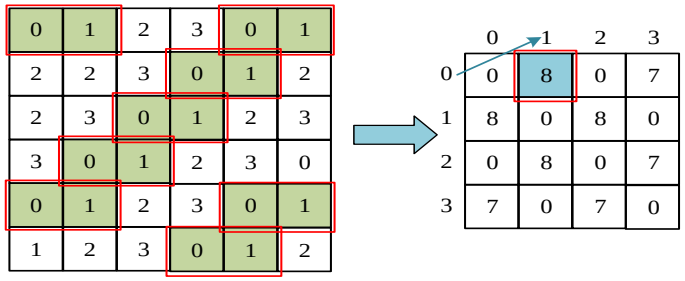
According to Fig. 4, it calculates the gradient amplitude and direction of pixels, and assign weights to the direction bin of each pixel. Each cell contains 9 features, and each block has 36 features. If 8 pixels are used as a step size, there are 7 scanning windows in the horizontal direction and 15 scanning windows in the vertical direction, so a 64 * 128 image will generate 3780 features. It normalizes the contrast of cells within each overlapping block and uses the L2 norm algorithm for normalization calculations. The normalized feature vector is represented as C, and the normalization calculation is shown in Eq. (12).

$$C \leftarrow C / \sqrt{\| C \|_2^2 + \varepsilon^2} \tag{12}$$

In Eq. (12), the function of ε is to prevent the denominator from becoming 0, as shown in Eq. (13).

$$\| C \|_2 = \sqrt{\sum_{k=1}^{n} | C_k |^2} \tag{13}$$

In Eq. (13), the initial value of K is 1. After image normalization, it extracts the feature vectors of HOG. To highlight local detail features and preserve image edge gradient features, this study proposes a secondary feature extraction algorithm. This algorithm processes LBP texture maps with directional gradient histograms and utilizes the MLBP algorithm instead of the LBP algorithm for feature description [18-20]. The grayscale co-occurrence matrix can be regarded as a matrix function that integrates information such as different directions, intervals, changes in amplitude, and speed in the image, and then presents this information in the form of a matrix, as shown in Fig. 5.



(a) Grayscale images (b) Gray co-occurrence matrix

Fig. 5. Gray co-occurrence matrix.

Fig. 5 demonstrates that in a grayscale image, this study can select two pixel points, record the number of times the combination of their values appears, and organize these records into a matrix, which is the grayscale co-occurrence matrix. To better capture the details and spatial relationships of images while preserving edge information, this study proposes the MLBP-HOG-G algorithm. This algorithm integrates MLBP, HOG, and grayscale co-occurrence matrix features together. This is to improve the effectiveness of feature extraction. This study selected the feature fusion method of serial fusion and adopted the weighted serial fusion method. The calculation method is showcased in equation (14) [21-22].

$$\bar{b} = \frac{1}{i} \sum_{j=1}^{i} b_j \tag{14}$$

In Eq. (14), $\bar{b}$ represents the mean of MLBP-HOG features; α represents the variance of MLBP-HOG features. The weighted results of MLBP-HOG-G features are shown in Eq. (15).

$$L = \frac{\alpha}{\alpha + \beta} C_1 + \frac{\beta}{\alpha + \beta} \tag{15}$$

In Eq. (15), β represents variance. The facial recognition system proposed in this study combines MLBP, HOG, and grayscale co-occurrence matrix features. The system first uses the MLBP algorithm to extract texture information from the image, and then uses the HOG algorithm to further extract features. These features are combined into a vector $C1$, and then dimensionality is reduced. Meanwhile, the Grayscale Co-occurrence Matrix features are also extracted as vector G. The entire algorithm process is shown in Fig. 6.

Fig. 6 shows that after extracting vector, and are merged and weighted to form a feature vector. This vector is input into the classifier for classifying the test samples. After classification is completed, the RR is used to evaluate the performance of the algorithm. The object of smart classroom face recognition system is teachers or managers. Its main function is to recognize and record the identity of students in the classroom through face detection and recognition, count the attendance in class, and cooperate with the classroom to complete teaching evaluation. The 1080p camera is used to collect students' classroom videos. In the face database coding link, the system will collect and store the face information photos of students in the classroom. These photos will be used to build the face feature database. The system uses Python code for unified size processing, and usually adjusts the image to a size of 160 × 160. Users are allowed to send requests to upload videos through the web. In the process of data transmission, TLS encryption protocol is used to protect data to prevent it from being intercepted by the system in the middle. After receiving the requests, the back-end server processes the received videos. This processing stage includes video pre-processing steps to ensure the adaptability and preparation of video data. Through analyzing the classroom monitoring video uploaded by the smart classroom system, the research shows the large visual screen of face recognition,

data analysis results and statistical check-in results to teachers or managers in the system to help teachers understand students' learning and attendance more objectively. Research and design the main functions of the smart classroom face recognition system include video upload, student face detection and recognition, and the display of results. In the non-functional design part of the system, we pay attention to protecting the safety of student face data, and comply with relevant data protection laws and regulations. If the data cannot be used for other purposes, we will prevent the data from being stolen or leaked. Design a login authentication mechanism based on user name and password. After the user enters the user name, enter the password in the user password box, enter the verification code in the verification code box, and then click the login button to verify the login. Different user types are assigned different permission levels, and the administrator manages the user. The user information is encrypted to prevent the potential risk of user password disclosure, so as to ensure the safety and privacy of users. When using face recognition data for statistical analysis, the data are anonymized to ensure that personal identity cannot be directly recognized through the data, and the personal privacy data are desensitized to cover up sensitive information and ensure that personal privacy will not be revealed when the information is used. At the same time, regularly review security protocols and access logs, timely find and respond to potential threats, and realize the security of personal privacy in intelligent classroom applications.

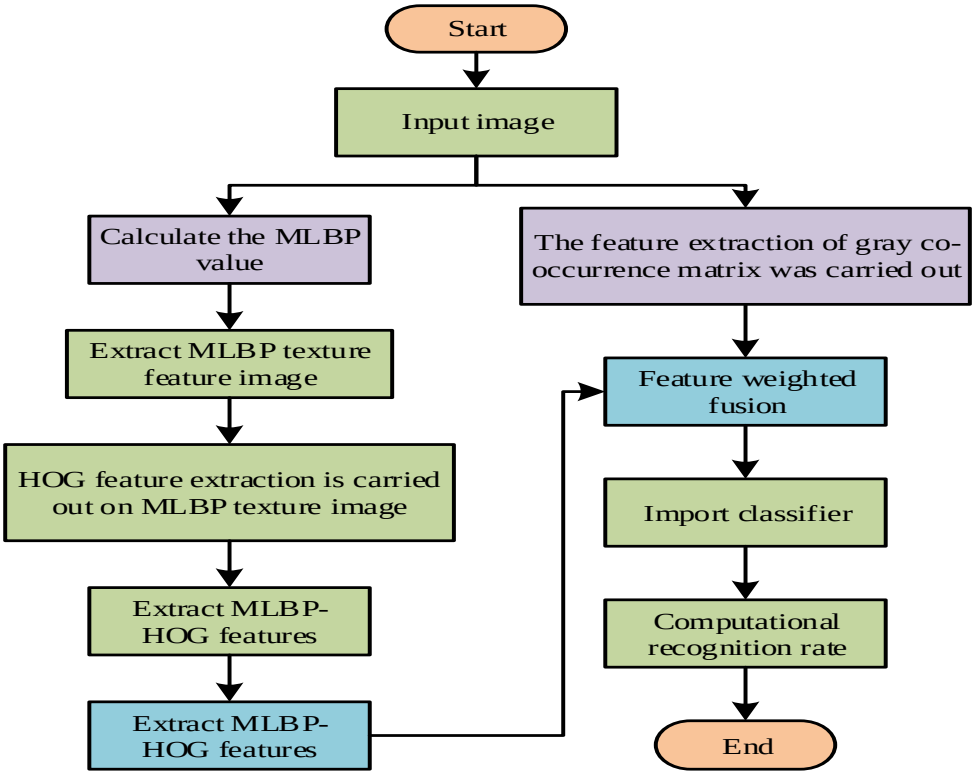


Fig. 6. Flow chart of FR algorithm based on MLBP-HOGG.

IV. Experimental Verification of Facial Recognition Model based on MLBP-HOG-G Algorithm

The experiment is conducted on a hardware platform equipped with an Intel i3-4030U processor (with a clock speed of 1.9 GHz), 8 GB of memory, and Intel HD Graphics Family GPU, running a 64 bit Windows 10 operating system. The simulation software uses MATLAB 2013 and relies on Image Processing Toolbox and Deep Learning Toolbox. During the training process, the ORL dataset takes about 2 hours, the CMU-PIE dataset takes about 12 hours, the YALE dataset takes about 1 hour, and the peak memory usage is about 4 GB. Due to limited GPU performance, it mainly relies on CPU computing. The experimental setup has a batch size of 32, an initial learning rate of 0.001, and employs an exponential decay strategy. This experiment uses three publicly available facial databases: ORL, CMU-PIE, and YALE. The ORL database contains 400 images (resolution: 92 × 112 pixels), covering different lighting, expressions, and poses; The CMU-PIE database contains 41368 images (resolution: 640 × 480 pixels), providing 13 poses, 43 lighting conditions, and various facial expressions; The YALE database contains 165 images (resolution: 320 × 243 pixels) covering different expressions and lighting conditions. Normalize and grayscale all images before the experiment, and divide them into training set, validation set, and test set in a ratio of 70%: 15%: 15%. To verify the robustness of the algorithm, salt and pepper noise (intensity: 0.1) and Gaussian white noise (mean: 0, variance: 0.01 and 0.1) were added to the ORL dataset. In addition, the training set is randomly rotated, translated, and scaled to enhance data diversity. The experimental results under Gaussian white noise attack are shown in Fig. 7.

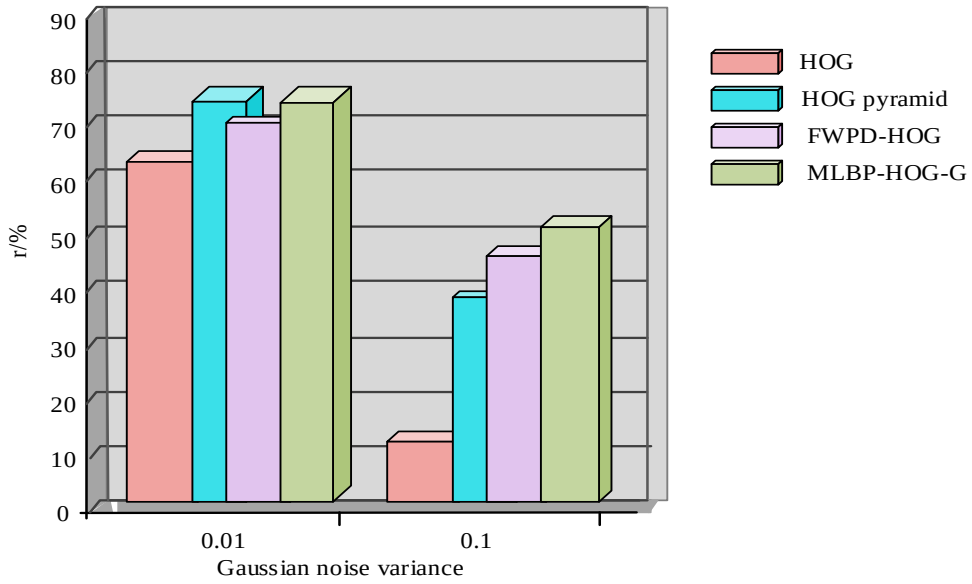


Fig. 7. The recognition rate of the proposed method is compared with the comparison method under Gaussian noise attack.

According to Fig. 7, when the Gaussian noise variance is 0.01, this research method improves the RR by 1% to 13% compared to other methods; when the variance is 0.1, it increases by 6% to 44%. Compared with the FWPD HOG method, the FWPD HOG pyramid method has a higher RR, demonstrating the advantages of multi-scale pyramid representation in combating noise. The experiment was conducted on the ORL facial database, consisting of 400 images, each with 10 images, totaling 40 groups, with an image size of 112x92. This study added Gaussian noise and salt and pepper noise for testing the algorithm. Each group of experiments will select four images as training samples, and the remaining ones as test samples. Ten repeated experiments will be conducted for calculating the average RR and compare the RRs of several algorithms in different dimensions. The outcomes are showcased in Table I.

TABLE I THE AVERAGE RECOGNITION RATE OF ORL ALGORITHMS UNDER DIFFERENT NOISES

Method	Salt-and-pepper noise (%)	Noiseless (%)	Gaussian noise (%)
WSRC	80.3	81.6	58.4
PCA-SRC	76.7	79.6	52.5
RPH-WSRC	85.0	92.5	74.1
HOG-SRC	81.2	83.2	64.3

Table I shows that when there is no noise, the RR has grew by 16.10%, 13.26%, and 11.16% compared to other algorithms. Even when different noises are introduced, the average RR of RPH-WSRC remains at the highest level, demonstrating strong anti-interference ability. Fig. 8 showcases the RR curves of each algorithm in the ORL dataset.

Fig. 8 shows that as the feature dimension increases, the RRs of various algorithms show an upward trend and eventually tend to stabilize; despite some fluctuations, this indicates that not all features contribute to classification recognition. When noise is introduced into facial images, the images are contaminated and occluded, and the RR of RPH-WSRC algorithm exceeds other algorithms, indicating that the algorithm has a certain degree of robustness against noise. For verifying the MLBP algorithm, this study designed a FR algorithm on the ground of MLBP. Considering that the original LBP may lose detailed features during feature extraction, this study proposes the MLBP algorithm to ensure the preservation of image detail features and enhance robustness during the feature extraction process. Therefore, on the ground of the MLBP facial recognition algorithm, a series of experiments were conducted for evaluating the recognition of MLBP and compared with different LBP algorithms. In the experiment, a block size of 5 * 5 was used and the dimension was reduced to 60 dimensions, which were tested in different databases. The experiment is showcased in Fig. 9.

Fig. 9 shows that the MLBP algorithm performs well in different databases. In the ORL database, the RR reached 95%, higher than 92.5% for LBP and 94.17% for ULBP. In the CMU_PIE database, the MLBP algorithm is also the best, with a RR of 94.12%, while the RRs of LBP and ULBP are 90.07% and 91.18%, respectively. In the YALE database, the RR of the MLBP algorithm is 93.33%. Although the database has the strongest variation factors, it is still higher than the 88.33% and 90% of the LBP and ULBP algorithms. In the database, the selection method for training samples is as follows: 7 images of each person are chosen from the ORL database, and 20 images of each person are chosen from the CMUPIE database; In the YALE facial database, each person selects 7 images as training samples, while the remaining images are utilized for experimental testing. Fig. 12 shows the comparison of RRs for different dimensions of MLBP-HOG features in the experiment, as well as the comparison of RRs for various dimensions of MLBP-HOG in MLBP-HOG-G.

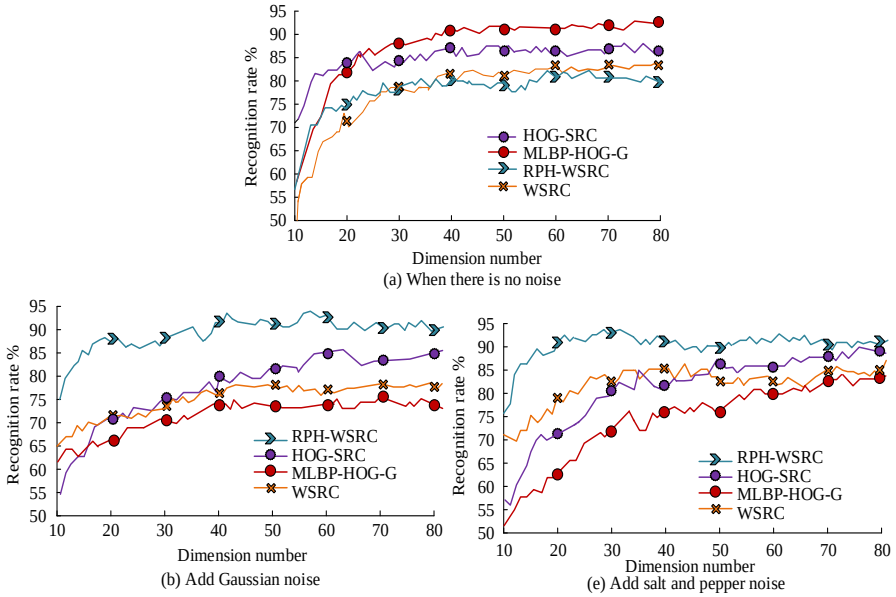


Fig. 8. Experimental recognition rate curves of each algorithm in ORL data set.

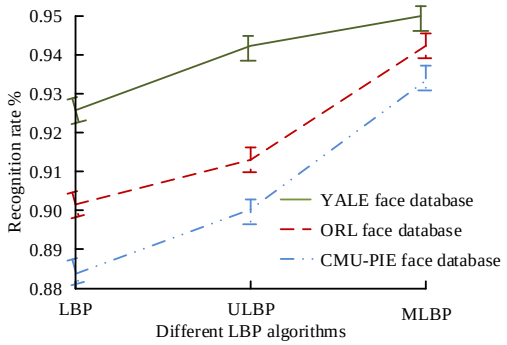


Fig. 9. Comparison of different LBP recognition rates.

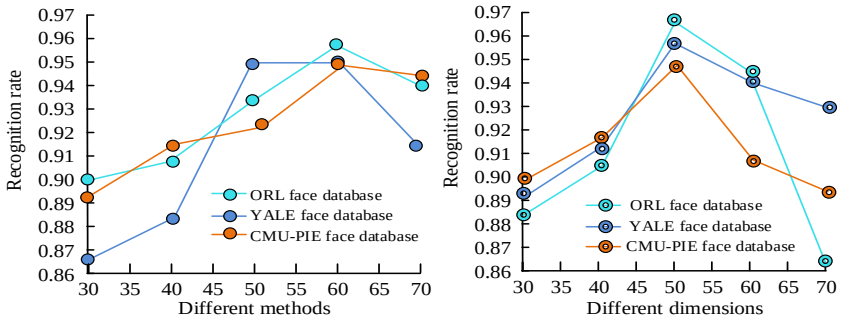


Fig. 10. Compare the recognition rate of different dimensions of MLBP-HOG feature and compare the recognition rate of different dimensions of MLBP-HOG in MLBP-HOG-G.

According to Fig. 10(a), the impact of the dimensions of MLBP-HOG features on RR in the ORL database is as follows. The highest value is 95.83% at 60 dimensions; 94.17% at 70 dimensions. The 30-50 dimensions are 90%, 90.83%, and 93.33%, respectively. In the CMU_PIE database, the dimension of MLBP-HOG features has the following impact on RR, with 60 dimensions being the highest at 94.85%. The 30-50 dimensions are 89.34%, 91.54%, and 92.28%, respectively. The 70 dimensional ratio is 94.49%. According to Fig. 10(b), the 50 dimensional features perform best in different databases. In the MLBP-HOG-G feature of the ORL database, when the dimension of the MLBP-HOG feature is 50 dimensions, the RR reaches 95.83%. In the CMU_PIE database, the RR reaches 95.22%. In the YALE database, the RR is 96.67%. Fig. 13 shows the relevant results of RRs among various methods in the experiment.

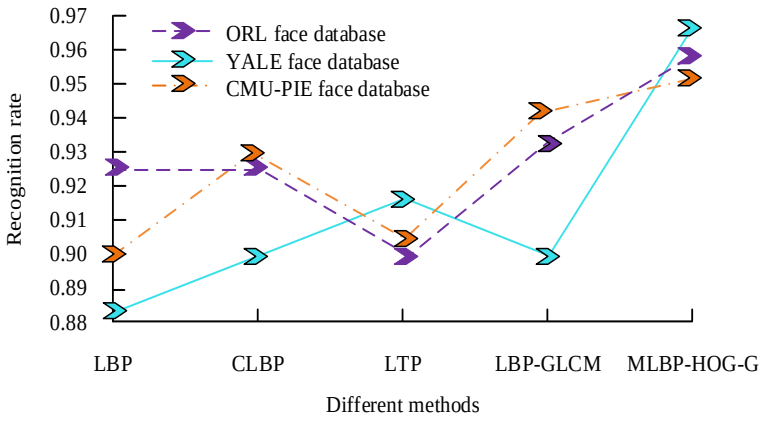
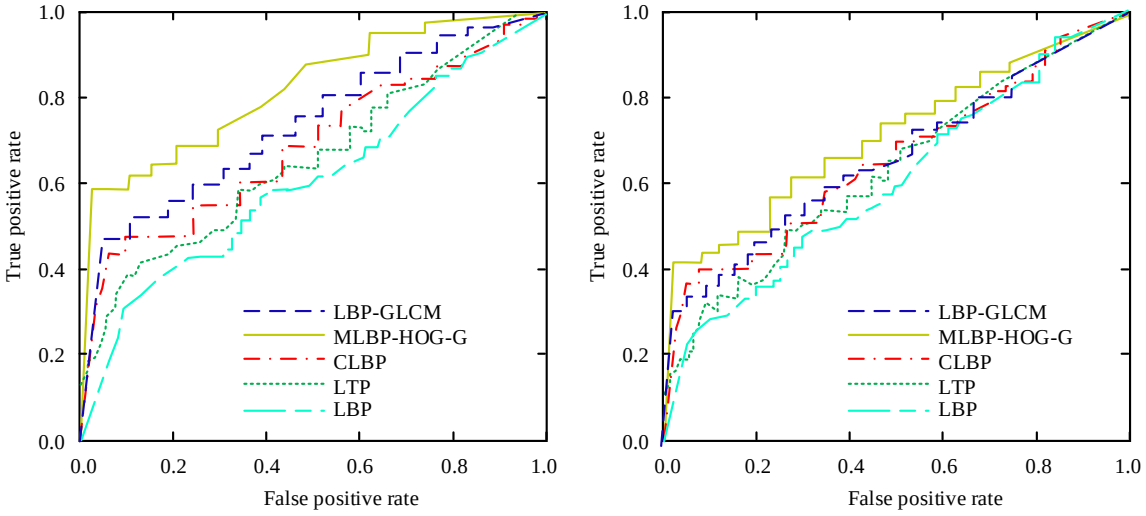


Fig. 11. Comparison of recognition rate of different methods.

According to Fig. 11, in the ORL database, the RR of MLBP-HOG-G features is the highest at 95.83%, followed by LBP+GLCM features at 93.33%. In the CMU_PIE database, MLBP-HOG-G features perform best with a RR of 95.22%. In the YALE database, the same MLBP-HOG-G feature has the highest RR of 96.67%. The method combining MLBP-HOG features and grayscale co-occurrence matrix features has shown excellent performance in FR. The face recognition results of the proposed method on different databases have been analyzed in the previous content, and further analysis and comparison have been made to further verify the performance of the proposed method. The ORL database and Yale face database is combined with and cover the student face data set designed by the research. At the same time, to further evaluate the facial recognition performance of the proposed algorithm in large datasets, the MegaFace dataset and VGGFace2 dataset were introduced for analysis. The MegaFace dataset is the largest publicly available facial recognition dataset, with one million faces and their respective bounding boxes, making it one of the largest public facial recognition datasets currently available. This facial image covers variations in age, gender, race, and facial expressions. VGG2 (9K ids/3.31M images) VGGFace2 is a dataset containing over 4.3 million facial images of more than 33000 different individuals, including facial images of different poses, ages, lighting, and backgrounds. It can be used for facial recognition tasks in complex scenarios such as age and pose changes. Fig. 12 shows the facial recognition accuracy results of different algorithms on a large dataset.

The results in Fig. 12 indicate that on the VGGFace2 dataset, the MLBP-HOG-G algorithm and LBP-GLCM algorithm have better facial feature recognition accuracy, with corresponding accuracy ACC values greater than 0.90. On the MegaFace dataset, although the sample size has been expanded and the recognition accuracy of the comparison algorithm has been affected, the MLBP-HOG-G algorithm proposed by the research institute still has good recognition accuracy, with its accuracy curve closer to the upper left corner. The mixed data set is divided into the test data set and validation data set according to the ratio of 6:4, and the recognition performance of different algorithms is compared. The detected face image is unified to a size of 160*160 and input into the face recognition model. Fig. 12 shows the face recognition training and verification accuracy of different algorithms, and the comparison algorithms are literature [23], literature [24] and literature [25].



(a) VGGFace2 dataset (b) MegaFace dataset

Fig. 12. Facial recognition accuracy results of different algorithms on large datasets.

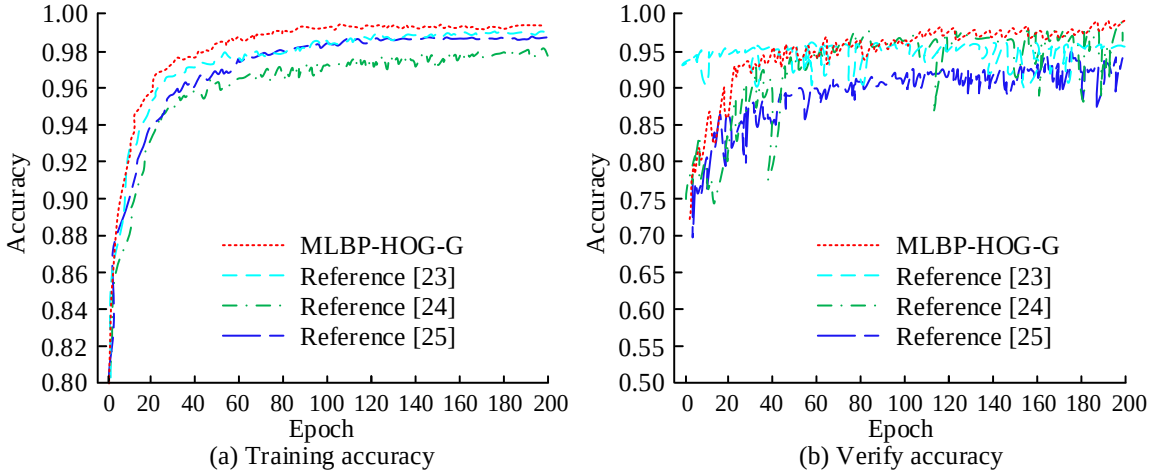


Fig. 13. Comparison of training and validation accuracy of different models.

In Fig. 13, the corresponding methods in literature [23], literature [24], and literature [25] are hog joint convolutional neural network, yolo-v4 network under the improvement of embedded components, and deep learning algorithm, respectively. In Fig. 12(a), the training accuracy curve of the MLBP-HOG-G algorithm proposed in the study has little fluctuation and is relatively stable. Its recognition accuracy in the later training batch is basically more than 98%, with a maximum training accuracy of 99.30%, and its performance is better than other comparison methods. The maximum training accuracy of literature [23], literature [24] and literature [25] are 98.95%, 98.90% and 98.00%, respectively, and there are certain fluctuations in the early stage. In Fig. 12(b), the validation accuracy of yolov3 hog is 97.23%, the generalization performance is the best, and the overall trend of the curve is relatively stable. In reference [25], feature extraction with the help of principal component and directional gradient histograms is inevitably affected by noise. This results in large fluctuations in its validation accuracy curve, presenting an unstable state with the increase of iteration times, and the maximum validation accuracy is not more than 95%. The validation accuracy curves of references [23] and [24] exceed 90%, but there are also some node fluctuations. Then the detection performance of face key points is analyzed, and the results are shown in Fig. 14.

The smaller the value of normalized mean error (NME), the better the robustness of the algorithm. Fig. 14 shows that on the training and test datasets, the mlbp-hog-g algorithm proposed in the study shows a small cumulative error result in the detection of key points in face recognition, and the overall curve change node amplitude is relatively small. The NME values of the other three comparative literatures increase with the increase of the map scale. And the fluctuation of the curve nodes is obvious, with varying degrees of deviation, and poor robustness on different datasets. After that, the performance of the proposed algorithm is compared and analyzed under different experimental conditions. The results are shown in Table II.

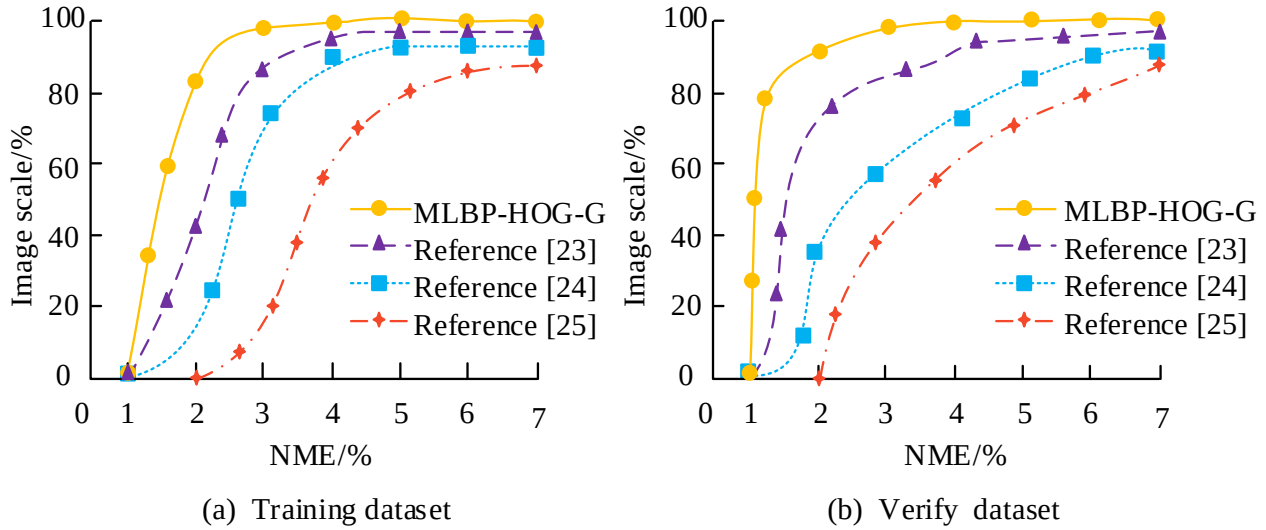


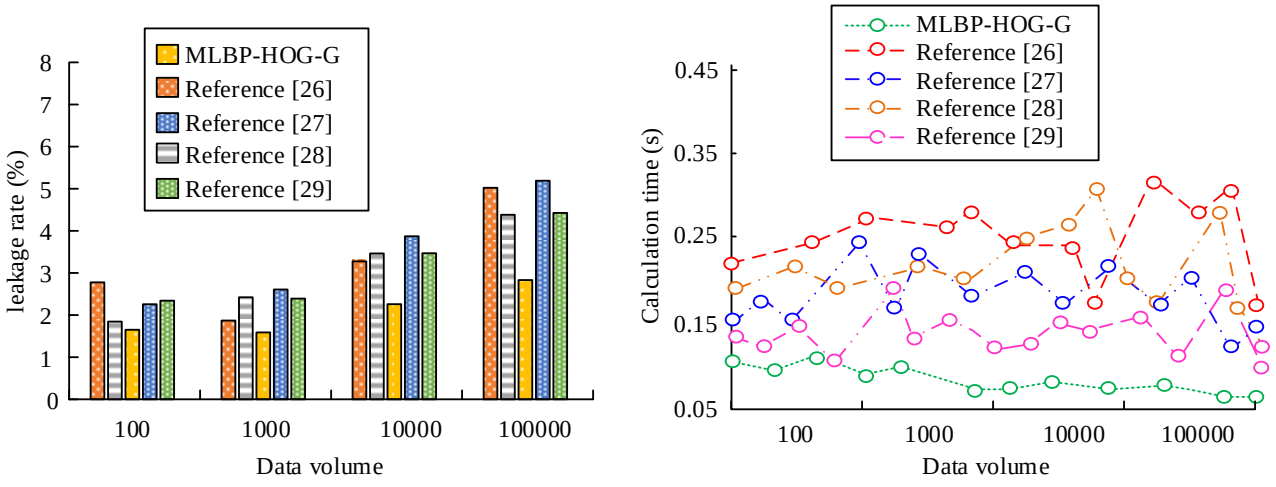
Fig. 14. Standard normalized mean error results of different comparison algorithms.

TABLE II INDEX TEST OF FOUR ALGORITHMS IN FACE IMAGES WITH DIFFERENT COMPLEXITY

Dataset Complexity	Index	MLBP-HOG-G	Reference [23]	Reference [24]	Reference [25]
Simple	mAP	98.56	90.14	89.33	92.07
	FPS (img/s)	88.19	84.97	85.13	81.01
	Detection time (s)	21.05	38.78	35.26	34.39
	Energy Efficiency (J/img)	3.2	6.4	5.2	6.7
Secondary	mAP	97.28	89.86	89.75	89.23
	FPS (img/s)	73.43	66.61	73.39	51.55
	Detection time (s)	20.12	23.06	24.17	28.33
	Energy Efficiency (J/img)	2.4	3.6	3.2	3.9
Complex	mAP	98.16	85.87	87.98	88.92
	FPS (img/s)	75.32	62.17	52.26	66.14
	Detection time (s)	24.36	30.41	31.65	29.38
	Energy Efficiency (J/img)	2.3	4.8	5.2	5.7

The indicators used in Table II include Mean Average Precision (map), Frames Per Second (FPS) and energy efficiency. The simple condition refers to the classroom face image under normal environment (no occlusion and no light change), while the medium and complex conditions mainly refer to the face image under partial occlusion and occlusion and light shadow change. Table II shows that the map values of the research algorithm under the three conditions are 98.56, 97.28 and 98.16, which are much higher than other algorithms under the same conditions. In terms of test efficiency and energy efficiency, the difference between the comparison algorithm and the research algorithm is at least more than 5img/s and 0.8j/img. In terms of running time, the running time of mlbp-hog-g algorithm in three conditions is 21.05s, 20.12s and 24.36s, respectively, which is less than other comparison algorithms. In conclusion, mlbp-hog-g algorithm has good performance in face recognition and detection, and has good adaptability under different conditions.

The use of deep learning methods to achieve classroom face recognition has become a research focus for many scholars. In order to further test the effectiveness of the MLBP-HOG-G algorithm proposed in this study, it was compared with literature [26], [27], [28], and [29], all of which were deep recognition results designed for classroom teaching. The results were analyzed from the perspectives of computational cost and recognition accuracy, as shown in Fig. 15.



(a) Leakage rate (%) (b) Calculation time (s)

Fig. 15. The computational cost and recognition accuracy results of different algorithms for facial recognition.

From Fig. 15(a), it can be seen that as the amount of data increases, the missed detection rate of the MLBP-HOG-G algorithm remains relatively low and changes steadily, with an overall missed detection rate of no more than 3%. The missed detection rates of other comparative literature are all above 2%, and their changes are more significant with the increase of data volume. From Fig. 15(b), it can be seen that as the amount of data increases, although the computation time of the MLBP-HOG-G algorithm increases, the growth rate is relatively small, and the overall computation time remains within an acceptable range, below 0.15s. However, the computation time of the remaining literature is relatively high, with the maximum computation time of literature [26] and literature [28] exceeding 0.30s. The above results indicate that the MLBP-HOG-G algorithm proposed by the research institute has high efficiency in processing large-scale data, while maintaining a low false positive rate in face recognition, it also has high computational efficiency and significant application effects.

V. Conclusion

Traditional facial recognition algorithms often require a large amount of computing resources and memory, which is a significant problem for environments with limited resources such as embedded systems. In this context, this study proposes an innovative FR method that combines MLBP-HOG features with grayscale co-occurrence matrix features. This method extracts binary texture images through the MLBP algorithm, and obtains MLBP-HOG features through secondary HOG feature extraction. The experiment showed that in the comparison of MLBP-HOG features with different dimensions and MLBP-HOG with different dimensions in MLBP-HOG-G, the highest RR was achieved in the ORL database at 60 dimensions, reaching 95.83%. In the MLBP-HOG-G features of the ORL database, when the dimension is 50 dimensions, the RR reaches 95.83%; In the CMU_PIE database, the RR is 95.22%; In the YALE database, the RR is 96.67%. The experiment indicates that the algorithm can adaptively learn and extract facial features, reducing the dependency of feature engineering.

The MLBP-HOG-G algorithm has a recognition rate of over 95% on ORL, CMU-PIE, and YALE databases, indicating its high accuracy on standard datasets. In complex situations such as changes in lighting, posture, and facial occlusion, the performance of the MLBP-HOG-G algorithm is significantly better than traditional methods, indicating its strong environmental adaptability. The reason is that through feature dimensionality reduction and serial fusion, the MLBP-HOG-G algorithm significantly reduces computational complexity and storage overhead while maintaining high recognition rates. And with the increase of data volume, the missed detection rate of MLBP-HOG-G algorithm is relatively low and stable, with an overall missed detection rate of no more than 3%, and the overall calculation time remains within an acceptable range, less than 0.15s. The detection rate of missed diagnosis in other comparative literature is above 2%. The above results indicate that the MLBP-HOG-G algorithm has high efficiency in large-scale data processing, while maintaining a low false positive rate in face recognition. It also has high computational efficiency and significant application effects. The reason is that the MLBP-HOG-G algorithm combines the multimodal features of MLBP, HOG, and gray level co-occurrence matrix, which can more comprehensively describe facial images, and adaptively learn and extract facial features, reducing reliance on artificial feature engineering. However, the use of facial recognition technology in smart classrooms involves personal data of students and teachers. Therefore, privacy and data security issues are a serious concern. In smart classrooms, facial recognition technology collects highly sensitive personal information, including facial features, attendance behavior, learning/teaching status, etc. Once these data are illegally obtained or leaked, they may seriously violate personal privacy, be used for identity theft, fraud, or other illegal activities, and harm their education and other legitimate rights and interests. To strengthen ethical data protection and usage restrictions, the purpose of data collection and use should be clearly defined. In the future, data encryption and storage security can be strengthened, such as adopting advanced data encryption technology, establishing strict data access control mechanisms, and preventing data leakage. At the same time, the scope of data use should be clearly defined, the dissemination and sharing of data should be restricted, and it should not be disclosed or sold to third parties. Enhance users' right to information and choice, ensure that the collection and use of facial recognition data comply with data protection regulations, and follow the principles of fairness, impartiality, and transparency.

In summary, the study proposes the MLBP-HOG-G algorithm, which not only provides a new approach for facial recognition technology, but also offers valuable exploration for future research to find a balance between improving recognition performance and ensuring data security. Future research should focus on developing more secure data encryption technologies, privacy protection mechanisms, and reliable data storage and transmission solutions to ensure that facial recognition technology can comply with ethical standards and protect user privacy in its widespread application in fields such as education.

References

[1] Jiang D. Research on remote monitoring method of smart classroom based on internet of things. International journal of autonomous and adaptive communications systems. 2022, 15(3): 220-234.

[2] Petrovi L, Stojanovi D, Mitrovi S, Bara D, Bogdanovi Z. Designing an extended smart classroom: An approach to game-based learning for IoT. Computer Applications in Engineering Education, 2021, 30(1): 117-132.

[3] Wang X, Cheng M, Eaton J, et al. Fake node attacks on graph convolutional networks. Journal of Computational and Cognitive Engineering, 2022, 1(4): 165-173.

[4] Oslund S, Washington C, So A, et al. Multiview Robust Adversarial Stickers for Arbitrary Objects in the Physical World. Journal of Computational and Cognitive Engineering, 2022, 1(4): 152-158.

[5] Niu J Y, Xie Z H, Li Y, Cheng S. J, Fan J. W. Scale fusion light CNN for hyperspectral face recognition with knowledge distillation and attention mechanism. Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2022, 52(6): 6181-6195.

[6] Widjaya C, Wicaksana A. Liveness Detection with Randomized Challenge-Response for Face Recognition Anti-Spoofing. International journal of innovative computing, information and control, 2023, 19(2): 419-430.

[7] Nam V H, Huong N M, Cuong P. Masked face recognition with convolutional neural networks and local binary patterns. Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2022, 22(5): 5497-5512.

[8] Fan Y. Face recognition algorithm of sprinters based on sliding data camera measurement. International Journal of Reasoning-based Intelligent Systems, 2023, 15(1): 79-85.

[9] Wang L, Shi Y, Zhang Z. Hyperspectral Image Classification Combining Improved Local Binary Mode and Superpixel-level Decision. Journal of Signal Processing, 2023, 39(1): 61-72.

[10] Kaplan N, Burg D, Omer I. Multiscale accessibility and urban performance. Environment and Planning B: Urban Analytics and City Science, 2022, 49(2): 687-703.

[11] Pabba C, Bhardwaj V, Kumar P. A visual intelligent system for students' behavior classification using body pose and facial features in a smart classroom. Multimedia Tools and Applications, 2024, 83(12): 36975-37005.

[12] El-Mashad Y, Ali H A. A new approach for smart attendance system based on improved video facial recognition technology for smart university. 2024: 77-95

[13] Yuan Z, Jiazheng Y, Hongtian L I, Hongzhe L I U, Chneg X U. Intelligent Classroom Face Detection Algorithm with Improved YOLOv5. Journal of Computer Engineering & Applications, 2024, 60(11).

[14] Aly M. Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model. Multimedia Tools and Applications, 2024: 1-40.

[15] Niu J Y, Xie Z H, Li Y, Cheng S J. Fan J. W. Scale fusion light CNN for hyperspectral face recognition with knowledge distillation and attention mechanism. Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2022, 52(6): 6181-6195.

[16] Dongbo L I, Huang L. Reweighted sparse principal component analysis algorithm and its application in face recognition. Journal of Computer Applications, 2020, 40(3):717-722.

[17] Wang S, Wang D. Grey Relational Analysis Coupled with Principal Component Analysis Method for Optimization Design of Novel Crash Box Structure.2019, 28(3):577-584.

[18] Shuang, Wang, Dengfeng, et al. Grey Relational Analysis Coupled with Principal Component Analysis Method for Optimization Design of Novel Crash Box Structure. Journal of Beijing Institute of Technology, 2019,101(3):199-206.

[19] Chen Y, Chen Y. A Network Flow Correlation Method Based on Chaos Theory and Principal Component Analysis. International Journal of Network Security, 2020, 22(2):242-249.

[20] Velilla José A, Volpe M R, Kenney G E, et al. Structural basis of colibactin activation by the ClbP peptidase. Nature chemical biology, 2023, 19(2):151-158.

[21] Wang Z, Zhan J, Duan C, Guan, X, Yang K. Vehicle detection in severe weather based on pseudo-visual search and HOG–LBP feature fusion. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2022, 236(7):1607-1618.

[22] A S L, B P R, C P M, A F. L. Less-is-Better Protection (LBP) for memory errors in k NNs classifiers. Future Generation Computer Systems, 2021, 117:401-411.

[23] Fakhar S, Baber J, Bazai S U, Marjan S, Hasinaska E, Chaudhry M U. Smart classroom monitoring using novel real-time facial expression recognition system. Applied Sciences, 2022, 12(23): 12134.

[24] Chen H, Guan J. Teacher–student behavior recognition in classroom teaching based on improved YOLO-v4 and Internet of Things technology. Electronics, 2022, 11(23): 3998.

[25] Geerthik S, Karthikeyan R, Keerthana G. Face Recognition based Automated Smart Attendance using Hybrid Machine Learning Algorithms and Computer Vision (ICAAIC). IEEE, 2024: 606-611.

[26] Trabelsi Z, Alnajjar F, Parambil M M A, et al. Real-time attention monitoring system for classroom: A deep learning approach for student's behavior recognition. Big Data and Cognitive Computing, 2023, 7(1): 48.

[27] Lasri I, Riadsolh A, Elbelkacemi M. Facial emotion recognition of deaf and hard-of-hearing students for engagement detection using deep learning. Education and Information Technologies, 2023, 28(4): 4069-4092.

[28] Gupta S, Kumar P, Tekchandani R K. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimedia Tools and Applications, 2023, 82(8): 11365-11394.

[29] Villegas-Ch W E, García-Ortiz J, Sánchez-Viteri S. Identification of emotions from facial gestures in a teaching environment with the use of machine learning techniques. IEEE Access, 2023, 11: 38010-38022.