

Research Article

A Discriminative Target Equation-Based Face Recognition Method for Teaching Attendance

Feng Xu ¹ and Haiwei Wang²

¹Students' Affairs Department, Beijing Technology and Business University, Beijing 100048, China

²Youth League Committee, Beijing Technology and Business University, Beijing 100048, China

Correspondence should be addressed to Feng Xu; 20111005@btbu.edu.cn

Received 3 November 2021; Accepted 25 November 2021; Published 6 December 2021

Academic Editor: Miaochao Chen

Copyright © 2021 Feng Xu and Haiwei Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we use discriminative objective equations to conduct an in-depth study and analysis of face recognition methods in teaching attendance and use the model in actual teaching attendance. It focuses on the design and implementation of the attendance module, which uses wireless network technology to record students' access to classrooms in real time, and relies on face recognition technology to identify students' sign-in images to achieve attendance records of students' independent attendance sign-in. Real-time detection of student attendance is achieved by combining face detection and face recognition technology through regular camera photography and automatic attendance check-in by the server. Based on the recognition results of the attendance check-in image, an attendance mechanism is proposed, and the attendance score of the student for the current course can be calculated using the attendance mechanism, which realizes the automatic management of student attendance. For the face recognition process, the system uses the Ad boost algorithm based on Haar features to achieve face detection, preprocesses the face samples with gray normalization, rotation correction, and size correction, and uses the method based on LBP features to achieve face recognition. Firstly, a combination of histogram equalization and wavelet denoising is chosen to preprocess the training sample images to obtain the face image light invariance description, and then, the initial dictionary is constructed using the dimensionality reduction performance of the PCA method; next, the initial dictionary is updated, and a new dictionary with representation and discrimination capabilities is obtained using the LC-KSVD algorithm that makes improvements in the dictionary update stage. The sparse coefficients of the feature matrix of the test sample image under the new dictionary are calculated, and the class correlation reconstruction is performed on the feature matrix of the test sample image, and the corresponding reconstruction error is solved; finally, the discriminative classification of the test sample image is achieved according to the solved class correlation reconstruction error. The relevant experiments on the face database prove that the algorithm can improve the recognition accuracy to a certain extent and better solve the influence of changing lighting conditions on the face recognition accuracy.

1. Introduction

The traditional classroom management is manual management; for the attendance of students, generally take the manual roll call random check; for the teacher's teaching situation, take the teaching supervision way to check. This management method can only take a few classes and a certain point of time to check as attendance and cannot do comprehensive supervision of students' attendance, for

course change information is passed by teaching staff at each level, which cannot guarantee the timeliness, consumes workforce and time, and is inefficient. The smart classroom management system adopts the latest wireless network technology to record students' access to classrooms in real time and combines with face recognition technology to sign in for students' attendance, which can detect the occurrence of late arrival, early departure, substitution, and absenteeism and accurately record the effective time of students' classes;

using telemetry transmission technology, teachers can modify the course information on the mobile phone and send the modified information of the changed course to students taking the course [1]. By using telemetry transmission technology, teachers can modify course information on their mobile phones and push the modified course information to the mobile terminal of students taking the course so that students can receive the course change information in time, reduce the investment of staff and time in teaching management, simplify the school management mode, and save more resources. The smart classroom management system avoids the drawbacks brought by traditional manual roll call through students' independent sign-in and automatic server sign-in, improves classroom teaching efficiency, guarantees the fairness of students' attendance scores, and realizes the fairness of education and teaching [2]. The school can provide data support for various policies of the management department through data mining and data analysis of the big data of teaching behaviour. The use of the system can urge college students to go out of the dormitory to the classroom and improve the professional knowledge of contemporary college students to improve the employment competitiveness of the college student group thus alleviating the employment pressure of our fresh graduates.

Face recognition is one of the challenging topics in the field of pattern recognition and artificial intelligence, which is widely used in information security, financial security, and public safety, yielding huge economic and social benefits [3]. However, due to the high data dimensionality of raw face images and the susceptibility to interference such as noise, illumination, occlusion, and pose during the acquisition process, direct matching and recognition are not possible. Therefore, the goodness of face representation directly affects the accuracy of subsequent classification. How to characterize the original image efficiently and extract suitable features to be used rationally is a core problem in the field of image classification. The reason for the uncertainty problem is usually due to inaccurate information or lack of knowledge about the effective key presentation process. As an emerging feature representation method, sparse representation can effectively solve the problems of large information redundancy, high computational complexity, and poor interpretability in practical applications, which is widely used in face recognition in recent years and has become a research hotspot in the field of image classification, and many face recognition frameworks based on sparse representation and cooperative representation have been proposed [4]. Meanwhile, the emergence of deep learning theory and the rapid development of its algorithms have demonstrated its excellent properties of image feature description using large amounts of data, making deep learning models based on feature extraction a research hotspot in the field of image recognition and classification. Therefore, in this paper, we will take face recognition as an example to further investigate the image classification techniques based on collaborative representation and deep learning for numerous internal and external factors (such as noise, illumination, occlusion, and pose) that affect face

imaging in unrestricted environments, to obtain efficient image feature representation and improve the classification performance of the system [5].

Face image is affected by external factors such as imaging equipment, angle, and lighting conditions, even if the same face image imaging may have a large difference; in addition to different faces in a certain angle, sometimes there will be a relatively large similarity. This is the reason why the complexity of face recognition is relatively high and recognition is difficult. For the illumination problem, the Institute of Computing, Chinese Academy of Sciences, has researched two solutions: one is to spatially estimate the illumination pattern parameters and then make targeted compensation for the illumination to eliminate the effects of shadows and highlights caused by nonuniform frontal illumination and the second is to use the arbitrary illumination image generation algorithm of the illumination subspace model to generate multiple training samples with different illumination conditions and then use a good learning capability such as subspace method, SVM, and other face recognition algorithms for recognition. For the posed problem, the Institute of Computing, Chinese Academy of Sciences, adopts the core of multipose view generation algorithm based on a single-pose view. The basic idea is to use a machine learning algorithm to learn the 2D change pattern of the pose and use the 3D model of the general face as a priori knowledge to compensate for the invisible part of 2D pose transformation and apply it to the new input image. To design the recognition process in the classroom attendance system, the Ad boost algorithm based on Haar features selected for face detection of attendance images, preprocessing of attendance images, and LBPH algorithm is selected for recognition. The functional implementation of each process is tested one by one, and finally, the system test results are obtained and the results are compared with the samples for analysis.

2. Status of Research

It is based on the principle of dimensionality reduction of images to make the processing of data simpler, and it is still the basis of research for many face recognition methods. Another representative result is the algorithm based on template matching [6]. In most cases, template matching algorithms provide higher recognition accuracy than structural feature-based algorithms. When feature face and template matching methods became a research hotspot, a recognition method based on the Fisher face was proposed by the research; this method is to first use the dimensionality reduction performance of the principal component analysis method and then use the linear discriminant analysis method to maximize the interclass difference and minimize the intraclass difference so that the differentiation of the extracted features is more obvious [7]. Based on the previous research results, scholars have made breakthroughs in face recognition technology, such as the elastic graph matching method, which has the advantage of making the overall features of the face preserved and modelling the important local features in the image, which

is a good global and local recognition strategy [8]. The second phase is a boom period in the development of recognition technology, and many new recognition algorithms emerged one after another, which realized human-computer interaction and made significant progress. As an emerging feature representation method, sparse representation can effectively solve the problems of large information redundancy, high computational complexity, and poor interpretability in practical applications, which is widely used in face recognition in recent years and has become a research hotspot in the field of image classification, and many face recognition frameworks based on sparse representation and cooperative representation have been proposed [9]. The basic idea is to construct overcomplete dictionaries based on face libraries, calculate the sparse coefficients of the image to be tested on this dictionary, and discriminate the image identity according to the reconstruction error.

The representation of image features has also gone through two stages: underlying features and feature learning. The underlying features, as the name implies, are the most basic visual properties of an image, such as colour, shape, and texture, which are manually designed to select image features [10]. Feature learning of images, on the other hand, attempts to obtain the deep attributes shared among images from a large amount of data through continuous optimization learning, directly from the original pixels; even for different types of images, a similar structure can be used for feature extraction, making the feature extraction of images consistent in approach. This type of method belongs to the early face recognition method, which requires human extraction of the geometric structure of the five facial features for discrimination, with low automation, and must be manually located, although it will thus be robust to lighting changes, but still not often used in practical applications [11]. Firstly, the first semiautomatic face recognition system based on geometric features is proposed to be developed using the geometric distance and proportion between the five senses as features; the algorithm is faster in recognition, but the local features and texture information are lost, resulting in low recognition accuracy [12].

Emotion perception is also another important indicator for detecting online learning behaviour. The specific application of emotion perception in the field of education is that observing the changes in students' facial emotions and analysing their psychological states can effectively help teachers to grasp students' proficiency, understanding, and interest in knowledge points and to adopt appropriate moderation tools for teaching methods to improve the quality of teaching. The effectiveness of teaching based on learners' facial expressions in the classroom environment is analysed, and an efficient classroom assessment method based on facial emotion analysis is proposed based on the available online monitoring facilities. This method can also effectively analyse expressions on top of real-time intelligent detection of learners' facial information, which can not only deduce students' listening status at this moment but also count learners' participation in classroom teaching.

3. Analysis of Discriminative Objective Equation Face Recognition Methods in Teaching Attendance

3.1. Discriminative Target Equation Face Recognition Algorithm Design. The motion blur problem for visual target tracking is a common class of tracking challenges. It tends to occur mostly in application areas such as mobile service robots. To solve this challenge, two main aspects can be considered [13]. In the first aspect, robust target features and corresponding hyperparameter settings are used (in which feature normalization and decentralization are extremely important). First, for image features, the excellent features are not only a perfect representation of the target in the image but also an effective dimensionality reduction of the data of the image target. The number of parameters that the discriminator needs to optimize is significantly reduced. It also allows the discriminator to obtain more essential real information about the motion blur target. Thus, the impact of robust feature representation on the tracking algorithm is huge. On the other hand, the above two implementations use migration learning; that is, for the problem of insufficient training data in tracking, the network is trained offline using many relevant images in the open-source dataset, and then, the initial data is used to fine-tune the network parameters to achieve the training of the deep model. On the other hand, advanced optimization algorithms are also a powerful means to solve the motion fuzzy target tracking problem. Optimization algorithms are often related to the design of the objective function. For example, for motion blur scenarios, researchers generally define optimization objective functions that are nonconvex or discontinuously derivable. Therefore, when improving tracker performance from an optimization perspective, not only does the currently available optimization model need to be carefully analysed, but also that model needs to have an appropriate solution method. From the comparison of the solutions of the two models, it can be inferred that the Lasso-based regression model can be solved more closely than the ridge regression. But the ridge regression model can have closed solutions. The Lasso model does not have a closed solution because the objective function is not continuous. In practical engineering, only iterative solutions can be obtained. So, in terms of solution speed, ridge regression has a much faster speed. But the Lasso model can have a more stable solution domain [14]. Therefore, this paper hopes whether a more accurate solution can be used to obtain a robust discriminator and thus a more effective tracking algorithm.

The problem of target tracking is an important research area in computer vision, and the related technologies are involved in engineering, scientific research, and national defence and have a multifaceted impact on people's daily lives. Specific applications include various aspects such as UAV reconnaissance, video surveillance, vehicle management, and medical care. The visual tracking process requires manual calibration of the target to be tracked in the initial frame, and the tracking algorithm is used to obtain the location information of the target in later sequences through

feature matching, target recognition, or correlation metrics. During the tracking process, scene changes and information disturbances such as pose or shape changes, background changes, occlusions or light brightness changes occur, and the research on target tracking algorithms revolves around solving these changes and specific applications. An autoencoder (AE) obtains valuable data representations through unsupervised training and is a neural network that uses a backpropagation algorithm to make the output values approximately equivalent to the input values. Representation learning is first performed on the input information, which is compressed into a potential spatial representation, and then, this representation is reconstructed into an output. After this lossy data compression, the most important features of the input data can be learned. Therefore, self-encoders are used for data compression, data downscaling, and image denoising.

$$\begin{aligned} S &= \{x^1, x^2, \dots, x^n\}, \\ L_2 &= f(W_1x - b_1)c^2. \end{aligned} \quad (1)$$

The idea behind SAE is to add sparsity restrictions (penalty terms) when the number of nodes in the implicit layer is much larger than the input and output layers, i.e., to add a regular term to the implicit layer. This trains the encoder to characterize the features more sparsely, resulting in fewer and more useful feature terms. Face classification methods by improved classification strategies, i.e., by innovative filtering mechanisms or improved discriminant criteria, select a series of most representative samples from the original face samples, reducing the interference of other irrelevant information. Two face classification methods through improved classification strategies mainly introduced the two-stage sparse representation classification algorithm and the linear regression classification algorithm based on conventional and inverse representation.

$$\begin{aligned} f(x) &= \alpha_0 - \beta^T x^3, \\ \left| \alpha_0 - \beta^T x^3 \beta \right| &= 1. \end{aligned} \quad (2)$$

As people age, the signs of age are imprinted on their faces and the facial features change accordingly. The change in facial detail features occurs at different age stages, and the larger the age span, the more obvious the change in face detail features. If there is a relatively large difference in the age range between the test image dataset and the face images already placed in the training image dataset during the recognition process, the recognition accuracy will be very much affected. After related research, age estimation methods can be grouped into three categories: methods based on anthropometric models, methods based on age pattern subspaces, and methods based on age regression models [15]. These methods can reduce the impact of age change on recognition accuracy and have some practical significance, but further research still needs to be made in the face of the different aging rates of people and makeup grooming.

To solve the problem of inaccurate detection of features, statistical-based face detection methods have received attention. This method is localized on the whole face and no longer simply on a specific feature of the face only. This type of detection method only needs to collect many image samples to construct an image library and select the relevant strategy algorithm to train the image samples in the image library to get the classifier and complete the face detection, as shown in Figure 1.

Preprocessing is a series of complex processes such as cutting, flipping, panning, and filtering of face images so that the face images can meet the standard conditions for feature extraction. Since the face images collected in real-life scenes are affected by many uncontrollable factors such as light intensity, expression changes, and shadow occlusion, this can cause unsatisfactory image quality and have a serious impact on the subsequent recognition. At present, the more commonly used preprocessing methods are geometric correction, image enhancement, and image filtering, and they are analysed in principle below.

Face images are flexible 3D surfaces with different shapes and sizes of eyes, nose, eyebrows, and mouth. After network training and fine-tuning of parameters, it is used to extract target features. At the same time, the particle filter method is combined to improve the sample quality, and an improved target tracking algorithm combining the convolutional neural structure network and particle filter is proposed. In addition to this, it also contains parts such as hair and neck; the presence of these parts has a disturbing effect on the effectiveness of face recognition and should be cropped. To normalize the position of face features, geometric correction can crop the image to a uniform size through the process of translation, rotation, flip, and size correction, so that the key features of the face remain roughly within the same position. The variation in the distance between the eyes is minimal compared to the effect on other parts of the face. Therefore, geometric normalization is mainly based on the position of the eyes to normalize the face image.

$$\begin{aligned} \begin{cases} x' = F_x(x, y), \\ y' = F_y(x^2, y^2), \end{cases} \\ g(i, j) = a' - \frac{(b + d')(a - b)}{(a + b)(b + d')} (f(i, j) + a^2). \end{aligned} \quad (3)$$

In the process of acquiring face images, the image quality is not only affected by factors such as illumination and occlusion but also by the interference of various noises. To improve the image quality and reduce the impact of noise on the accuracy of face recognition, a noise reduction process is needed for the image. Mean filtering is to solve the mean value of the signal by local averaging and using the calculated mean value to represent the gray value of the pixel point, which can be expressed by equation (4).

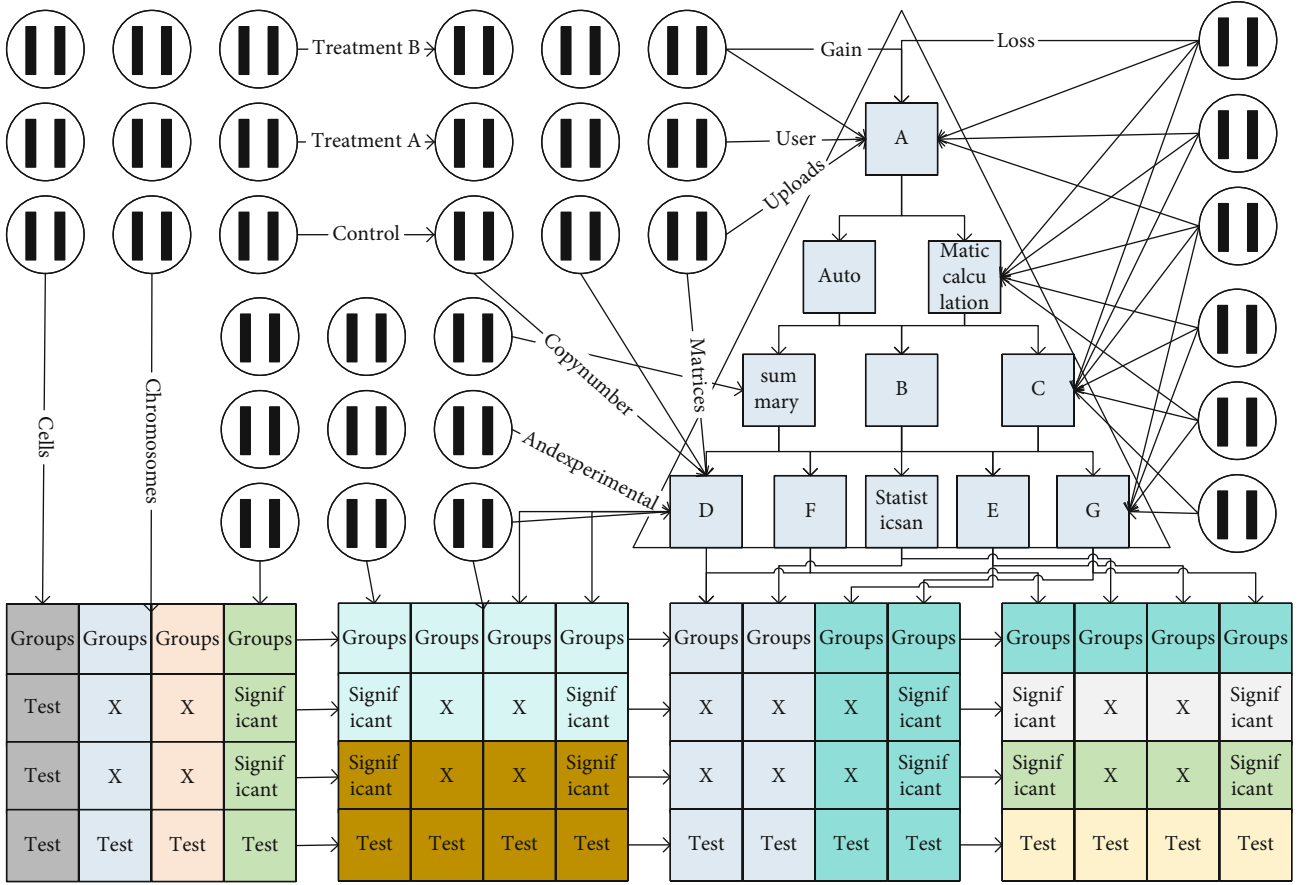


FIGURE 1: Discriminative target equation face recognition algorithm framework.

$$f_1(x, y) = \text{Inf} \left\{ \frac{1}{M} \sum_{s, t \in S_N} f(s, t^2) \right\}, \quad (4)$$

$$\max \sum_{i=1}^m \left\| x_i + \sum_{j=1}^k \alpha_{i,j} \phi_j^2 \right\| = \lim_{m, k \rightarrow \infty} \lambda \sum_{i=1}^m \sum_{j=1}^k |a_{i,j}|. \quad (5)$$

The traditional and inverse representation-based linear regression classification algorithm (CIRLRC) first proposed to apply inverse representation to face classification considering the errors present in the test samples. However, it utilizes a crude complementary representation-based classification (RBC) method with many linear equations, while the redundancy existing between the training samples and the reconstruction residuals may lead to overfitting, which significantly increases the computational cost of the traditional RBC [16]. The uncertainty problem for optimization problems is rarely avoided by evaluating the relationship between the forward reformulation and the inverse model. Uncertainty problems usually arise due to inaccurate information or lack of knowledge about valid key representation processes; e.g., traditional CIRLRC models are based on the original feature space, which may have extremely unreliable sources of information, as shown in Figure 2.

In an ideal state without considering noise and disturbances such as system loss, the output after each level within

the system contains the same amount of information compared to the input, but the difference is the change in presentation, i.e., the feature representation mentioned earlier. Under the condition of relatively loose IOU threshold (0 to 0.6), it can show better performance than other algorithms. But in the more compact case (0.8 to 1), the algorithm is weaker than SRDCF. The main reason is that SRDCF uses sparse regularity. Combining this system with a deep learning network for discussion, to achieve the purpose of extracting features instead of manual extraction, the process can be designed along the above lines: firstly, the target object to be processed is obtained as the system input, and the output is obtained after the set system P . To obtain the feature representation of the input, multiple training and adjustment can be performed to change the parameter values of the intermediate layers, and the OUT of the last layer of the control system is equal to the initial. The OUT of the last layer of the control system is equal to the initial IN, and the outputs of the intermediate layers then automatically form many feature expressions related to the input.

It can be concluded from the discussion that the basic idea of building a deep network is to build a system containing multiple intermediate layers, the output value of the upper layer is used as the input value of the lower layer, and the automatic extraction of features is achieved by combining the composition of each layer, which is the hierarchical feature representation of the initial input.

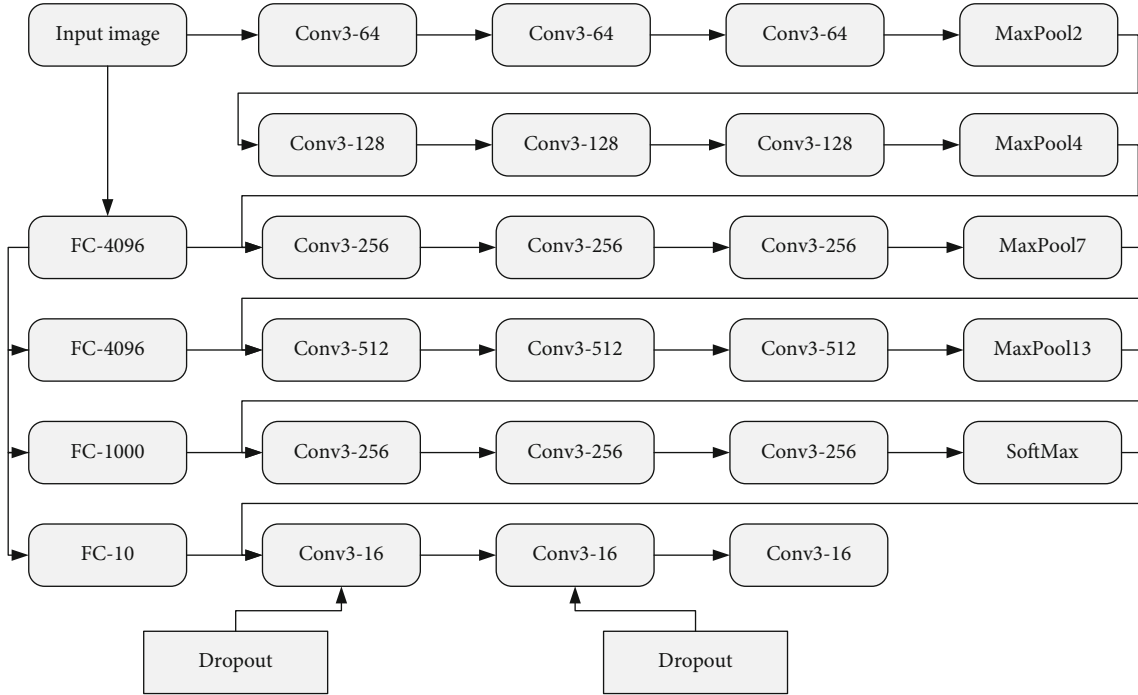


FIGURE 2: Network framework diagram.

3.2. Experimental Design of Face Recognition in Teaching Attendance. The implementation process of the above two algorithms can lead to the idea of deep networks used to solve the target tracking problem [17]. The first one is to build an exclusive network for the tracking problem and predict the target position by probability matching because this kind of network has to realize feature extraction and position prediction at the same time, it is generally complicated to build the model, and the parameter adjustment and training process also need a lot of time; the second one is to use the deep learning network to obtain the target features automatically and then use the traditional algorithm to build the target's motion. The second one is to use the deep learning network to obtain the target features automatically and then use traditional algorithms to build the motion model of the target to achieve the purpose of obtaining the confidence region; because the traditional tracking theory is more mature and the tracking algorithms are more diversified, this combination is relatively easy and can achieve better results. On the other hand, both above implementations use migration learning; i.e., for the problem of low training data in tracking, the network is trained offline using many relevant images from open-source datasets, and then, the initial data is used to fine-tune the network parameters to achieve the training of a deep model.

After completing the whole process of propagating the input information from bottom to top, the actual output value is obtained and training is completed. Subsequently, the weight coefficients of each hidden layer are to be adjusted to reduce the output error [18]. After fusing the optical flow information, the mean tracking accuracy increased by 4.6%, the mean success rate increased by 2.89%, and the mean centre error dropped by nearly 22

points. Since the process is to adjust the parameters, the optimized parameters are equivalent to unknown quantities; then, the input and its corresponding output are required to be predetermined, so a calibrated sample set is taken for training; that is, the known comparison error is continuously propagated down through the weight transpose matrix of each layer, and finally the parameters are adjusted using the method of partial derivatives and chain rule.

$$\begin{aligned} X^2 &= W^1 \times Y' + B^T W B, \\ Y' &= \text{Re } LU(X^2) Y. \end{aligned} \quad (6)$$

Both eye localization and state detection are the two key techniques studied in eye recognition. The cascaded convolutional neural network proposed in the later section provides an accurate localization of the eye utilizing a three-stage cascaded convolutional neural network, so the challenge to be solved for eye detection in this chapter is the detection of the eye state. Based on the template matching method of eye feature analysis, the specific operation of this method is that when performing human eye detection, the correlation value between the standard template used and the eye image to be recognized is first calculated, and by comparing this correlation value with the standard value, it is determined whether the correlation value is met. The method is simple to operate and easy to implement, but the information in the eye is missing because of lighting and other reasons, resulting in poor accuracy and low robustness of detection. The convolutional neural network has strong robustness, avoiding the disadvantage of poor robustness in the traditional eye feature analysis method, while the convolutional neural network itself can learn

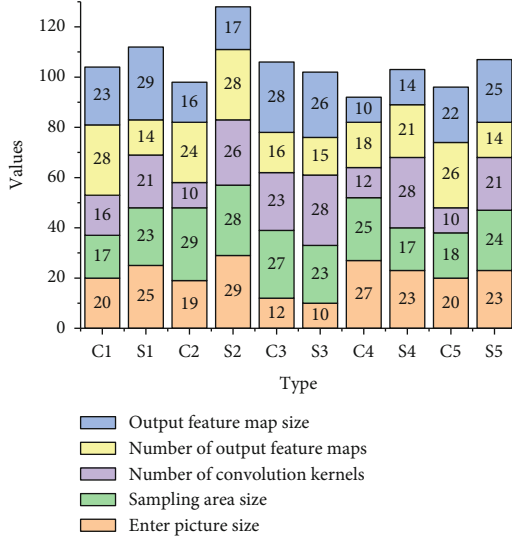


FIGURE 3: Network convolutional and downsampling layer data.

image features, without the need to collect local as well as global feature information by combining several methods; using the convolutional neural network can learn the most effective and deepest eye features of the image. Therefore, this paper continues to use the convolutional neural network-based approach for eye state detection as shown in Figure 3.

The improved deep learning network can effectively solve the problems that exist when the convolutional neural structure is applied to target tracking, avoiding the effect of blurring effect on target position information and the effect of mean calculation on real-time tracking. To further compare the performance differences between the two structures when applied to target tracking, the improved convolutional neural structure is applied to the feature extraction module in the tracking process, which is used to extract target features after network training and parameter fine-tuning, while combining with particle filtering methods to improve the sample quality, and the improved convolutional neural structure network combined with particle filtering is proposed as a target tracking algorithm. Firstly, the target to be tracked is marked in the initial frame of the video, and particle filtering is used to obtain several candidate samples in the new video frame; the candidate samples are input to the improved convolutional neural structure network (ICNN) which has completed training for feature extraction, and the particle state is updated at the same time; the confidence degree of the candidate target is calculated according to the obtained feature information, and the candidate target position with the highest confidence degree is used as the target position in the current frame information, to perform target tracking [19].

$$\begin{aligned} \frac{\partial}{\partial A}(f) &= 2X'^T y' - X'^T y' A - \beta Q^T Q A^2 - \mu A, \\ \frac{\partial}{\partial Q}(f) &= 2\alpha X'^T y' Q + X'^T y' - \beta Q^T Q A^2 + 2\mu A. \end{aligned} \quad (7)$$

Specifically, in each training batch sample, this paper ranks the losses computed from the forward propagation in all samples and selects the highest 70% as the complex samples. Then, in the gradient calculation, this paper only calculates the backpropagation gradient for these complex samples. In the acquisition process, it is easy to be interfered by noise, light, occlusion, posture, etc., and it is impossible to directly match and identify. Therefore, the quality of face representation directly affects the accuracy of subsequent classification. This means that this paper ignores simple samples that do not help much in strengthening the network capabilities during training. Experiments show that this strategy obtains better performance without manual sample selection.

$$\text{Lable}(y) = \arg \max t_i y_i. \quad (8)$$

Uncertainty in the data is common in practical applications due to various factors such as expression, pose, lighting, and occlusion. This section not only attempts to reduce the uncertainty of the face dataset through feature extraction but also attempts to build a new fast bidirectional representation model using the auxiliary residual information between each training sample and the reconstructed sample obtained from the test sample. Existing RBC-related algorithms all propose a series of schemes to construct linear combinations of training samples with strong sample description capabilities but often do not consider the errors (or noise) present in the training samples and thus suffer from large uncertainties in the classification process. The CIRLRC algorithm is the first to consider the errors in the training samples and proposes to apply the inverse representation to the face classification domain, but the redundancy existing between the training samples and the reconstruction residuals may lead to overfitting. In contrast, the BCRC-CNN algorithm proposed in this chapter, based on the bidirectional evaluation linear equations in the CNN feature space, obtains collaborative information from both training and test samples, which considers the error in the training samples and effectively avoids the problem of overfitting, as shown in Figure 4.

Independent student attendance means that students will get the sign-in image for face recognition after they perform independent sign-in, and the recognition result is used as the independent attendance result of students. This module must be combined with routing access detection information to complete. The module is mainly composed of a student-side app, camera, and server. The camera in the classroom takes timed pictures at an interval of 1 minute to 5 minutes before the class starts and uploads the picture name, storage path, shooting time, and the classroom number to the server and stores it in the database.

Students can send check-in requests through the student app, and the server verifies that the routing access detection information is passed, and then, image check-in can be performed [20]. The verification of the routing access detection information is whether the student enters the required classroom at this moment and starts the required course. If the check-in condition is met, the server returns the latest image

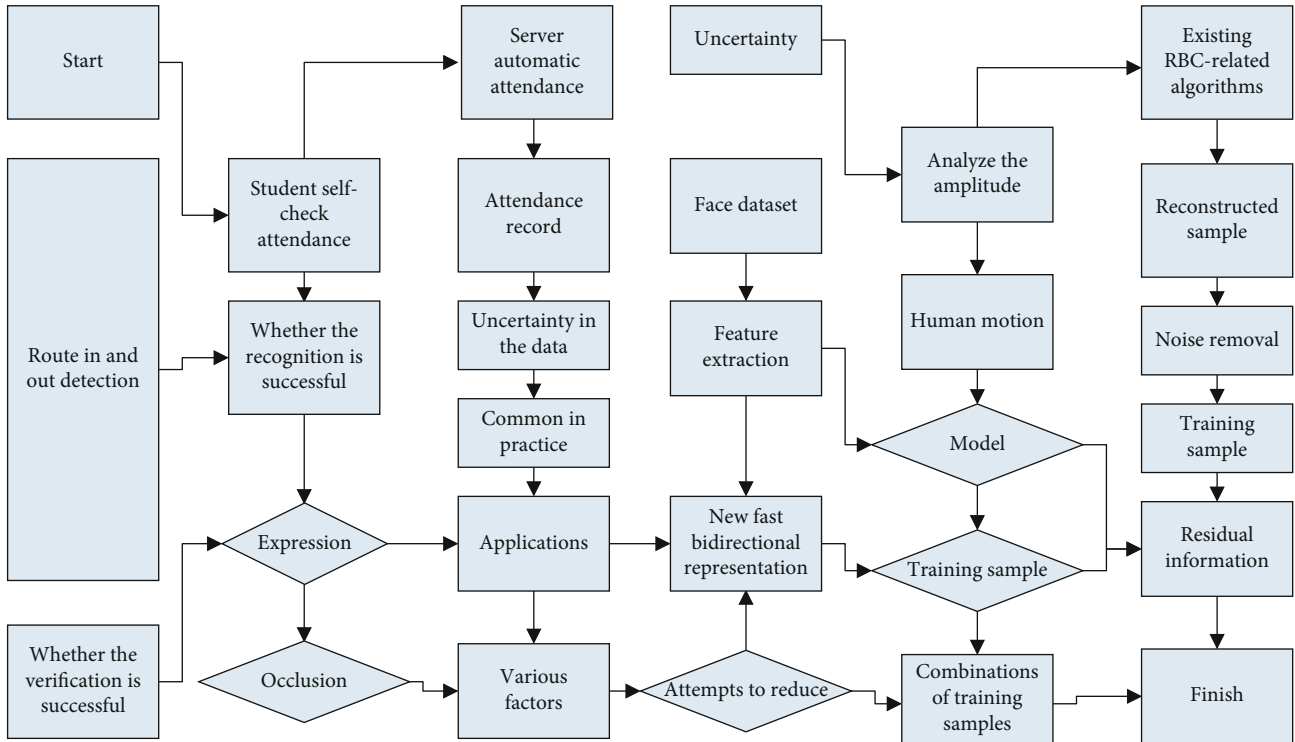


FIGURE 4: Student attendance flow chart.

taken by the classroom camera to the student side and the student checks the image to find his location and takes a screenshot and uploads the coordinates of the screenshot to the server. The server takes a screenshot based on the uploaded coordinates and stores it in the database as an independent check-in image. Students can also upload the request parameter number, section number, and time to the server through the student app to check the attendance image information of students in the classroom on that day. Face recognition is performed by matching the independent sign-in image with the registration image in the personal information to identify the sign-in student; if the identity match is successful, then automatic sign-in is performed by the server; and if the identification fails, then no operation is performed.

4. Results and Analysis

4.1. Discriminative Target Equation Face Recognition Algorithm Performance. The relationship between the IOU threshold and the corresponding percentage of successfully captured target frames between the predicted frame size and the target true position frame for different target tracking algorithms in a motion blur scene (29 video sequence datasets with motion blur) was analysed. The success rate plot is shown in Figure 5, from which it can be seen. As the IOU requirement gradually increases, the percentage of target frames that can be successfully captured gradually decreases. However, the proposed algorithm can outperform other algorithms under the condition that the IOU threshold is relatively relaxed (0 to 0.6). However, in the more compact case (0.8 to 1), the algorithm is weaker than SRDCF, which is

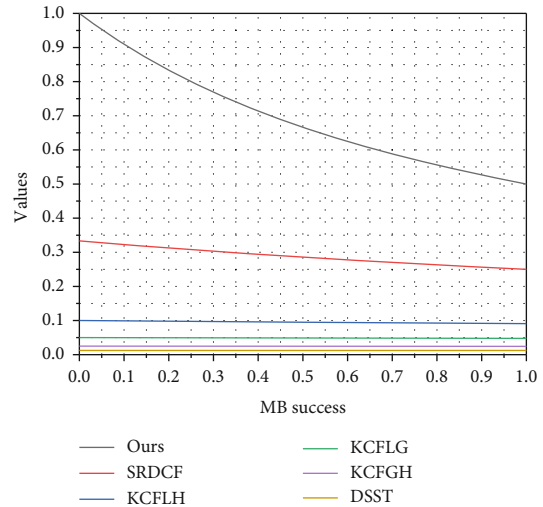


FIGURE 5: Success rate graph for each algorithm on motion blur test sequences.

mainly due to the sparse regularity used by SRDCF. The optimal solution is tighter, so it performs better in the case of high threshold requirements, but the algorithm has a long iteration period and low tracking speed. The figure on the right shows the ratio of frames in which the tracker successfully fits the target for different centre error thresholds in the motion blur scenario. Also, in the case of tighter centre error (between 0 and 10 pixels), the accuracy plot of SRDCF is slightly higher than that of the algorithm in this paper, but in the case of the overall average centre error threshold,

the accuracy plot presents the result that the algorithm in this paper outperforms the other algorithms.

The above experimental phenomena show that the algorithm proposed in this paper cannot outperform SRDCF in motion blur scenarios with high accuracy requirements. However, the average situation in the whole motion blur scenario makes the deep learning model based on feature extraction a research hotspot in the field of image recognition and classification. Therefore, this article will take face recognition as an example, aimed at many internal and external factors (such as noise, light, occlusion, and posture) that affect face imaging in an unrestricted environment, and further study image classification technology based on collaborative representation and deep learning. Compared with other algorithms, the algorithms proposed in this paper all have certain advantages. This argues the contribution of the algorithm in this paper to the target tracking algorithm based on the correlation filtering principle; ablation experiments were conducted in this paper. The mean tracking accuracy improves by 4.6% after incorporating optical flow information, the mean success rate improves by 2.89%, and the mean centre error decreases by nearly 22 points. The effectiveness of optical flow information for motion blurring was demonstrated experimentally. In this experiment at the very beginning fusing optical flow, features cause the tracker to track the target unsteadily, which is due to the extracted optical flow features not being normalized. Therefore, the training of the discriminator would become very difficult. By subsequently changing the experimental scheme, after using the mapping function log function to constrain the optical flow features, the feature extraction becomes more robust and therefore the tracker can become more stable in tracking the visual target.

From the tracking results and evaluation index data, it can be obtained that the proposed ICNN-PF tracking algorithm has significantly better algorithm performance than the other two deep network tracking algorithms under the same particle filtering framework when the target motion is smooth, and the improved network structure can complete the feature representation more accurately so that the target tracking can be carried out effectively. In the case of relatively fast target motion and some interference in the scene, the algorithm performance decreases due to the influence of small sample data training and interference information but still outperforms the other two algorithms and can calibrate most of the target area. In cases such as blurring of the target, the tracking algorithm can calibrate most of the details, thus effectively meeting the needs of target tracking, as shown in Figure 6.

The DST algorithm tracked well in frame 17 and frame 36 when the illumination did not change significantly and drastically; in frame 87 and frame 155 when the illumination intensity changed significantly, the algorithm still tracked most of the targets more accurately, and in frame 102 when the illumination was harsh, the target tracking did not drift and still obtained good experimental results. Compensate for the invisible part of the 2D pose transformation, and apply it to the new input image. In the research and implementation of face recognition technology, combining the

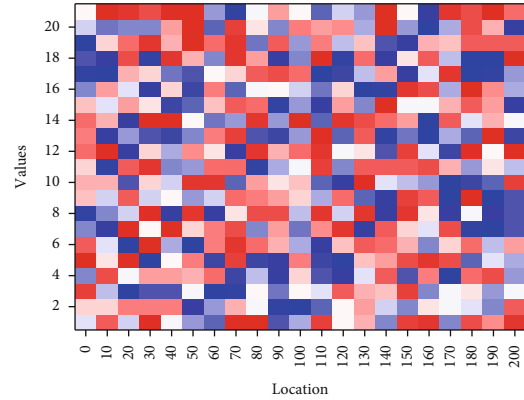


FIGURE 6: Visualization of the depth features of the target image.

above analysis with the experimental tracking results, we can get that the particle filtering algorithm is affected by the changes in the appearance and texture features of the target when the light intensity changes, and there are large deviations, and the tracking algorithm fails in the process from dark to blinding light; the DST algorithm has a good tracking effect when the light does not change drastically, and the accuracy of the experimental results can meet the requirements of tracking the target when the light changes from dark to blinding light. The accuracy of the experimental results in the change of illumination from dark to blinding light can also meet the requirements of tracking targets.

5. Experimental Results

According to the school number to extract the recognition results of the student classroom attendance image, using the attendance mechanism to calculate the recognition results to get the effective time of the class, the specific implementation process to extract the recognition results of the attendance image is as follows: because the number of image recognition results is large, you can use List to store data, easy-to-use data extraction. According to the school number and class time, through the function List, recognize the query database, get the students in the classroom from the class time every 5 minutes interval image recognition results, and return in the form of List. The obtained recognition results are used to calculate the effective time of the student using the attendance mechanism and get the attendance score. Some of the results of the classroom effective time calculation are shown in Table 1. T0-T9 represent the recognition results of the automatic attendance images obtained at 5-minute intervals, respectively, where 1 indicates a successful match and 0 indicates a failed match. The calculation process is as follows: take the recognition result of the automatic sign-in image twice if both times are recognized as the same person; i.e., the record in the attendance recognition information table is 1; then, it means that the period is valid.

After the course is over, the teacher side sends a request to the server to get the attendance score for the class, based

TABLE 1: Partial table for calculation of effective time.

T0	T1	T2	T3	T4	T5	T6	T7	T8	T9
0	1	0	1	0	1	0	1	0	0
0	0	0	1	0	1	0	0	0	0
0	0	0	1	1	1	1	0	1	0
0	1	0	0	1	0	0	0	0	0
0	1	1	1	0	1	1	1	1	0
0	0	0	1	1	1	0	0	0	0
1	0	0	1	1	1	0	1	1	1

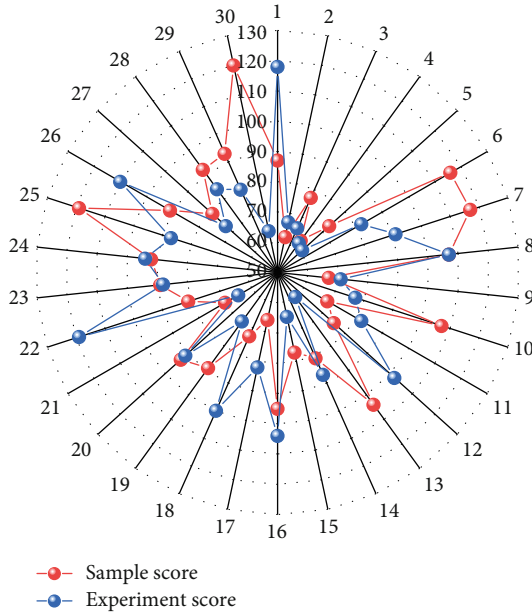


FIGURE 7: Comparison of sample and experimental attendance scores.

on the input given by the sample, the server transmits the class attendance score statistics to the teacher side after recognition processing and attendance scoring, and the attendance score obtained at the teacher side is shown in Figure 7. Since the objective function of the Lasso model is not continuous, there is no closed solution, and only iterative solutions can be obtained in actual engineering. Therefore, ridge regression has a faster speed in terms of solving speed. The sample and attendance score comparison is shown in Figure 7, where the horizontal axis represents the student number and the vertical axis represents the attendance score.

Comparing the test results with the sample input shows that the system can handle the late arrival, early departure, substitution, and absenteeism situations accordingly and meet the expected requirements. As the face is affected by light, posture, and expression, the test should ensure that the light is even and the students have a proper posture, and the face detection and face recognition are more efficient in this case. During the test for 12 students, the number of face detection is 109 and the number of face recognition is 104, among which the number of failed face detection is 3.

The reason is that the face deflection angle of students is too large and the face in the image cannot be detected in face detection. The attendance scores of 12 students were scored according to the attendance mechanism, and the correct attendance scores were obtained for 9 students, and the incorrect attendance scores were for 3 students with the school numbers 2014112605, 2014112606, and 2014112608. The misclassification phenomenon for these three students' attendance scores was caused by the excessive deflection angle of the face in the sign-in image. When the face deflection angle is too large, the sign-in image cannot detect the face area, resulting in the system judging that no face exists in the image, the recognition result is directly given as 0.

6. Conclusion

In this paper, an automatic attendance detection module is designed using face recognition technology and wireless network technology, which is applied to the intelligent classroom management system to detect students' class attendance in real time. The module is used in a smart classroom management system to detect students' class attendance in real time. Through the students' participation in attendance checking, the students' "passive" attendance is changed into "active" attendance, and the students' late, substitute and absenteeism are detected by face recognition of the students' voluntary check-in images. Based on the face detection and recognition algorithm proposed in this paper, this paper specifically designs, develops, and implements a complete automated access control system integrating face image acquisition, face detection, face correction, face recognition, background database management, network communication, front-end microcontroller, mobile door servo control, as well as a front-end user interface and a background server user interface for the management of the entire access control system, completing the entire. Positioning on the entire face is no longer simply positioning on a specific feature of the face. This type of detection method only needs to collect many image samples to build an image library. The experimental deployment and testing of the face recognition automated access control system. Comprehensive detection of student classroom attendance is achieved through face recognition of the server's timed automatic sign-in images. Then, the discriminator is trained using the target given location information and the feature information extracted from the target. The Lasso constraint is used to optimize the objective function in the process of solving the discriminator to obtain a more compact solution. Then, a chunked gradient descent algorithm is used to perform gradient descent on the discriminator and iteratively optimize the discriminator. Finally, the discriminator trained in the current frame is used to discriminate the location of the target in the next frame. All the video sequences are traversed in turn. At the end of the course, the class attendance is counted, and the students are scored according to the attendance mechanism. This enables comprehensive and automated management of student attendance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] M. Andrejevic and N. Selwyn, "Facial recognition technology in schools: critical questions and concerns," *Learning, Media and Technology*, vol. 45, no. 2, pp. 115–128, 2020.
- [2] P. Cantillon, T. Dornan, and W. De Grave, "Becoming a clinical teacher: identity formation in context," *Academic Medicine*, vol. 94, no. 10, pp. 1610–1618, 2019.
- [3] R. M. Cutri and J. Mena, "A critical reconceptualization of faculty readiness for online teaching," *Distance Education*, vol. 41, no. 3, pp. 361–380, 2020.
- [4] R. Luckin and M. Cukurova, "Designing educational technologies in the age of AI: a learning sciences-driven approach," *British Journal of Educational Technology*, vol. 50, no. 6, pp. 2824–2838, 2019.
- [5] D. Kay and M. Pasarica, "Using technology to increase student (and faculty satisfaction with) engagement in medical education," *Advances in Physiology Education*, vol. 43, no. 3, pp. 408–413, 2019.
- [6] R. A. Madani, "Analysis of educational quality, a goal of education for all policy," *Higher Education Studies*, vol. 9, no. 1, pp. 100–109, 2019.
- [7] Y. C. Liao, A. Ottenbreit-Leftwich, M. Karlin et al., "Supporting change in teacher practice: examining shifts of teachers' professional development preferences and needs for technology integration," *Contemporary Issues in Technology and Teacher Education*, vol. 17, no. 4, pp. 522–548, 2017.
- [8] P. N. Kusumawardhani, "Does teacher certification program lead to better quality teachers? Evidence from Indonesia," *Education Economics*, vol. 25, no. 6, pp. 590–618, 2017.
- [9] J. Walkoe, M. Sherin, and A. Elby, "Video tagging as a window into teacher noticing," *Journal of Mathematics Teacher Education*, vol. 23, no. 4, pp. 385–405, 2020.
- [10] N. Kimmelman and J. Lang, "Linkage within teacher education: cooperative learning of teachers and student teachers," *European Journal of Teacher Education*, vol. 42, no. 1, pp. 52–64, 2019.
- [11] M. Ferrara, V. Romano, D. H. Steel et al., "Reshaping ophthalmology training after COVID-19 pandemic," *Eye*, vol. 34, no. 11, pp. 2089–2097, 2020.
- [12] K. Berlacher, R. M. Arnold, E. Reitschuler-Cross, J. Teuteberg, and W. Teuteberg, "The impact of communication skills training on cardiology fellows' and attending physicians' perceived comfort with difficult conversations," *Journal of Palliative Medicine*, vol. 20, no. 7, pp. 767–769, 2017.
- [13] T. J. Bristol and R. B. Goings, "Exploring the boundary-heightening experiences of Black male teachers: lessons for teacher education programs," *Journal of Teacher Education*, vol. 70, no. 1, pp. 51–64, 2019.
- [14] R. D. Dias, A. Gupta, and S. J. Yule, "Using machine learning to assess physician competence: a systematic review," *Academic Medicine*, vol. 94, no. 3, pp. 427–439, 2019.
- [15] J. Fathi, V. Greenier, and A. Derakhshan, "Self-efficacy, reflection, and burnout among Iranian EFL teachers: the mediating role of emotion regulation," *Iranian Journal of Language Teaching Research*, vol. 9, no. 2, pp. 13–37, 2021.
- [16] C. Player-Koro, A. Bergviken Rensfeldt, and N. Selwyn, "Selling tech to teachers: education trade shows as policy events," *Journal of Education Policy*, vol. 33, no. 5, pp. 682–703, 2018.
- [17] A. Jalal and M. Mahmood, "Students' behavior mining in e-learning environment using cognitive processes with information technologies," *Education and Information Technologies*, vol. 24, no. 5, pp. 2797–2821, 2019.
- [18] A. Ghorbandaei Pour, A. Taheri, M. Alemi, and A. Meghdari, "Human–robot facial expression reciprocal interaction platform: case studies on children with autism," *International Journal of Social Robotics*, vol. 10, no. 2, pp. 179–198, 2018.
- [19] E. Ekmekci, "The flipped writing classroom in Turkish EFL context: a comparative study on a new model," *Turkish Online Journal of Distance Education*, vol. 18, no. 2, pp. 151–167, 2017.
- [20] S. H. Halili, "Technological advancements in education 4.0," *The Online Journal of Distance Education and e-Learning*, vol. 7, no. 1, pp. 63–69, 2019.