Behavior Recognition of College Students Based on Improved Deep Learning Algorithm

Xiaoke Ning, Jiangxi College of Foreign Studies, China*

ABSTRACT

With the vigorous development of intelligent campus construction, great changes have taken place in the development of information technology in colleges and universities from the previous digital to intelligent development. In the teaching process, the analysis of students' classroom learning has also changed from the previous manual observation to intelligent analysis. Based on this, this paper studies the behavior recognition of college students based on the improved deep learning algorithm. Based on a brief analysis of the research background of behavior recognition, the research framework of college students' behavior recognition is constructed. Finally, the authors designed an experiment to evaluate the accuracy of classroom student behavior recognition analysis. The results show that the improved recognition of college students' behavior based on deep learning algorithm can improve the recognition accuracy.

KEYWORDS

Behavior Analysis, Continuous Learning, Deep Learning Algorithm, Domain Adaptation

INTRODUCTION

Classroom observation is a more common way to evaluate teaching and learning in higher education (Byeon et al., 2021). Detecting and analyzing students' different behaviors in the classroom through digital technology can not only remind students to adjust their own behaviors, but also reflect the degree of classroom activities and help teachers improve teaching methods (Chonggao, 2021). The students' behavior and posture in class reflect not only the students' participation in learning, but also the teachers' teaching level to some extent. Various interactive analysis systems have emerged in the existing classroom behavior analysis research, such as Student-Teacher classroom teaching analysis and analysis methods based on information technology (Chen et al., 2020). All these methods are based on manual observation to record students' learning behavior (Cheng, Wei et al., 2022). However, this method wastes more time and energy and is inefficient. Teachers cannot always pay attention to every student (Cheng, Ma et al., 2022). With the development of artificial intelligence technology and machine learning algorithm, new research methods have emerged in student behavior analysis.

DOI: 10.4018/ijwltt.320647
*Corresponding Author

In recent years, some intelligent video recording and broadcasting systems have combined computer vision technology, which can reduce manual intervention and greatly improve the recognition effect and efficiency. The intelligent video recording and broadcasting system does not need manual participation in the recording process of the whole classroom teaching activities. It only needs to install one or more cameras at the fixed position of the classroom in advance and simulate the operation of manual lens switching through computational vision technology, so that the camera can automatically record classroom teaching, which greatly liberates manpower and ensures the quality of shooting. It is a major breakthrough in the development of video recording and broadcasting.

Based on this background, in this paper the author investigates student behavior recognition in college classrooms based on improved deep learning algorithms. The paper is divided into four parts. The first part is a brief introduction about student behavior analysis in college classrooms and the arrangement of this study. The second part introduces domestic and foreign algorithms about behavior analysis and deep learning algorithms and summarizes the shortcomings of current research. The third part provides a framework for student behavior recognition in college classrooms. then, it proposes an adaptive normalization algorithm based on deep learning filters distracting information and addresses the shortcomings of deep learning algorithms. Migrating behavioral features are added to the model to meet the demand for online real-time updates, while empirical memory prior knowledge is screened to learn new knowledge. The fourth part offers a simulation and analysis of the student behavior analysis model of college classroom based on the improved deep learning algorithm the author constructed in this paper; subsequently, it provides an evaluation of the algorithm performance by the accuracy of behavior analysis. The experimental results show that the proposed algorithm has better classification recognition accuracy, compared with the existing behavior analysis algorithms.

The innovation of this paper lies in the improved learning algorithm. Considering that the identification of the same type of student behavior in the classroom environment requires the extraction of higher-order spatio-temporal features, the author proposes the normalization method based on deep learning, and filters the noise information through the fusion method. The author establishes an adaptive learning model, which incorporates a domain judge and an attention learning model to achieve migration behavior analysis. Then, the researcher introduces the continuous learning mechanism of empirical memory to enable the analysis of new data for action recognition online.

Based on some shortcomings of the traditional classroom and the development of modern technology, the significance of applying deep learning technology in the field of education in this study is mainly shown in the following aspects: (1) It can timely find out students' abnormal behaviors in the classroom, supervise and manage the classroom so that the classroom can proceed normally, thus guaranteeing everyone's right to study and achieving the classroom objectives as a whole; (2) using computer technology to identify students' listening behavior is conducive to teachers' understanding of students' class status, improving teaching methods and optimizing classroom teaching and management, so as to improve the efficiency of teaching and learning and help teaching reform.

RELATED WORKS

In recent years, more artificial intelligence technologies have emerged, and more researchers have focused on artificial intelligence, machine learning, and deep learning. Both deep learning and machine learning belong to artificial intelligence, and deep learning also belongs to the category of machine learning. Recently, deep learning technology has been applied in many fields, such as face recognition, autonomous navigation, medical diagnosis, human-computer interaction, and satellite remote sensing. The rapid development of computer technology also promotes the research of educational intelligence. Although the application of artificial intelligence in the field of education has not changed the nature of education, it has provided new teaching methods for education, broken the original organizational order of education, and also provided more solutions to teaching problems.

This has led to the reform of teaching models and a huge leap in education, making education more diversified and enriched, and providing new ideas for education.

In the field of computer vision research, human behavior recognition is one of the important research directions and has wide applications in human-computer interaction and video retrieval (Cheng, Yang et al., 2022). Video-based behavior recognition techniques have gained attention in recent years, and research results have emerged. In a human action recognition research, Lu (2021) proposed a multifeature fusion human behavior recognition algorithm based on deep reinforcement learning, selected several typical human behavior datasets, constructed an attention model, and used small sample regions as model inputs. In a student classroom behavior recognition research, Dai et al., combined traditional cluster analysis algorithm and random forest algorithm to improve the traditional algorithm, constructed a network topology model, and verified the effectiveness of spatial angle features extracted based on human skeletal model (Dai et al., 2019). Zhang, et al. (2019) used deep learning techniques to learn the sample data collected from in-vehicle sensors in their study, proposed a joint data enhancement scheme, designed a new multiview convolutional neural network (CNN) model to construct a sample dataset that better matches the complex real driving environment, and developed a new multiview convolutional neural network model for training, learning, and recognition of driving behavior. Zahid et al. (2021) proposed a machine learning model to recognize human behavior by classifying 13 different crash types and using crash classification to predict crash times and developed a time prediction machine learning model to compensate the control scheme of APCID by predicting the collision time. Byeon et al. (2021) proposed a four-stream integrated CNN based on ROI, where the data consisted mainly of images and skeletons, by converting 3D skeleton sequences into pose evolution images, inputting RGB videos into 3D-CNN to extract temporal and spatial features, limiting the body ROI of RGB video into 3D-CNN and RGB video limited to ROI of hand-object interaction into 3D-CNN. Hsueh et al. (2020) proposed a recurrent neural network algorithm based on long and short term memory to implement behavioral patterns to improve the accuracy of human activity behavior recognition. Konstantinova combined PSO algorithm with KNN algorithm to obtain a joint PSO-KNN algorithm to construct a classroom student behavior recognition model based on image processing techniques and using key frame detection for feature recognition (Konstantinova et al., 2018). Jiang built a learner modeling mechanism by monitoring the free behavior of learners in math classes (Jiang et al., 2018). The teaching concepts and models in China are quite different from those in foreign countries. Although there are many researches on behavior recognition in foreign countries, there are few researches on students' behavior recognition in the classroom. The lack of research will inevitably lead to the lack of public student behavior datasets, which are the basis of relevant model training. The lack of research datasets will inevitably lead to the slow development of student behavior recognition research direction.

In summary, it can be seen that there are many studies in the field of student classroom behavior recognition, and intelligent behavior recognition is also an inevitable choice to promote the construction of digital campus. The intelligent classroom for intelligent behaviors recognition is a new personalized, digital, and intelligent learning environment built with the support of advanced information technology. It is characterized by active perception of teaching situations, automatic collection and analysis of data, adaptive push of learning resources, and diversification of learning tools. Smart classroom provides diversified and multidimensional data support for teachers' teaching decisions and students' learning opportunities and is an intelligent space for students' quality training and ability development. Expression recognition, emotion recognition, and behavior recognition are covered in recognition, which is no longer limited to manual observation and coding, and artificial intelligence techniques are beginning to be emphasized. However, most of the existing behavior recognition technologies are easily disturbed by background noise, the recognition accuracy is not high, and the detection of key points of human body is also disturbed. On the other hand, in classroom student behavior analysis, there are no strict data requirements for behavioral actions, so there may be similar characteristics of different types of actions, and existing recognition methods are only for behavioral actions with large differences. Therefore, based on the research of other scholars, this study applies deep learning technology to teaching through observation and experiment, recognizes students' behaviors and actions in class, urges students to learn, improves students' efficiency in class learning, strengthens students' self-management ability, and improves students' knowledge system.

MATERIALS AND METHODS

Design of Student Behavior Recognition Framework in College Classroom

The process of identifying students' classroom behaviors based on in-depth learning specifically includes:

1. Data collection, that is, collecting five typical classroom behaviors of research objects.
2. Data preprocessing, including image clipping and image data enhancement.
3. Taking the VGG16 network model loaded on the large-scale dataset ImageNet as the pretraining model, and fine-tuning the VGG16 network model, that is, using the training set and verification set to further optimize the network model parameters to obtain a new model suitable for students' classroom behavior recognition.
4. Inputting the test samples into the trained VGG16 network model and outputting the recognition results, that is, each test image is recognized as one of the five typical learning behaviors.

With the development of artificial intelligence technology, classroom intelligence system has also gradually entered the college campus. In addition, the combination of behavior recognition technology and classroom environment can analyze students' classroom behavior in a timely and effective manner and provide information for teachers to understand students' learning situation in time (Liang et al., 2019). On the one hand, teachers can focus more on classroom teaching and improve teaching efficiency, and, on the other hand, it is convenient for teachers to adjust the teaching schedule in time (Lu, 2021). In this study, the author designed a student behavior recognition system based on improved deep learning algorithm in their research, synthesized multiple data sources, and built corresponding models according to the requirements. Considering that the system needs user interface and back-end algorithm to work together, the author used QT GUI framework as the basic development framework.

According to the videos collected by the modern education center of the school, ithe author found that there are five typical states of students in class: Looking at the blackboard, reading, sleeping, turning around, and playing with mobile phones. These behaviors not only reflect the basic state of students, but also form the basis of complex learning activities. Identifying these typical behaviors can provide data support for subsequent automatic teaching analysis. Because of the experimental equipment, if the video of students' class behavior is directly converted into pictures, the pictures of each student's class behavior obtained are very blurred and the clarity cannot be reached. Therefore, the author invited some students to imitate the posture in class and take photos to get the data set needed for the experiment.

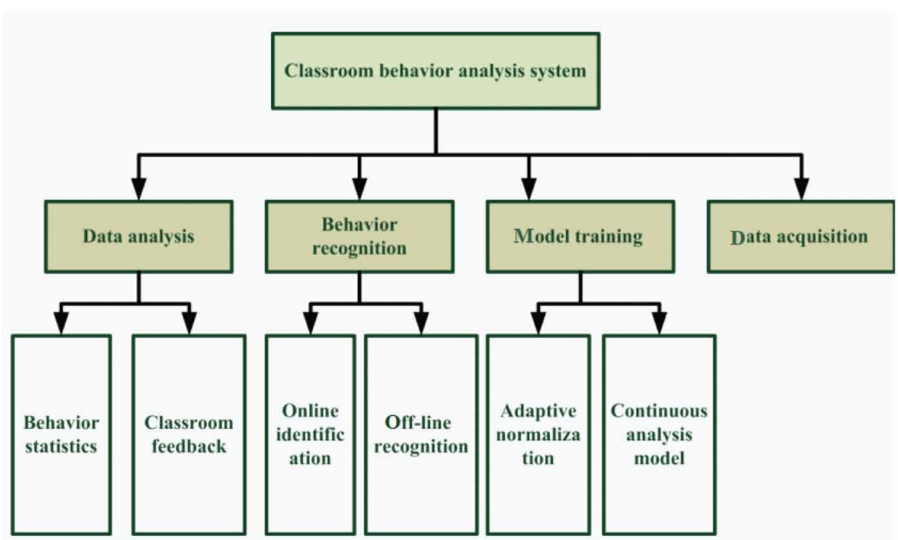
In the design of the system, the analysis of students' classroom learning is mainly realized, and feedback information for teachers is obtained. The system design should cover the skeleton action classification function to achieve the classification and recognition of data actions. Upon achieving behavior classification, the data should be capable of being saved for later use (Liu et al., 2021). The system also needs to have a data analysis and update function, which can directly visualize data in data analysis and, at the same time, can be updated to facilitate the addition of new data or modules at a later stage. The construction steps of the data visualization system include: Defining the type and structure of data, such as category, numerical or space-time, how many dimensions are there, and whether the structure is tree or graph; clear purpose; specifying how the data should be processed to clean up, organize, and finally save to the database.

By combining the behavior recognition algorithm, graphical user interface development system, and requirement analysis, the overall structure of college classroom behavior analysis is designed holistically. This structure covers a range of modules including data collection, recognition, model training, and data analysis (as illustrated in Figure 1). With the integration of these modules, the system is able to capture and analyze students' behavior in the classroom, providing valuable insights for teachers to improve teaching effectiveness.

In this system design, the model training module mainly takes the normalized processed data, extracts the key points to form action feature values, uses them as labels, and then inputs them to the deep learning classifier, while also being able to train the newly added data (Liu & Long, 2020). In the design of the recognition module, the main purpose is to take the collected classroom data information and classify and recognize it in the training model. The data display module mainly implements statistics as well as visualization of the data to facilitate teaching and learning viewing (Mabrouk & Zagrouba, 2018). Before conducting network experimental training, it is necessary to process the collected data. There are many reasons for this, such as the large size and resolution of unified data, adding some noise or reducing the brightness of the image. Noise in an image is often represented as an isolated pixel or pixel block that causes a strong visual effect. Generally, the noise signal is irrelevant to the object to be studied. It appears in the form of useless information and disrupts the observable information of the image. During the image acquisition process of two common types of image sensors, CCD and CMOS, due to the influence of sensor material properties, working environment, electronic components, and circuit structure, various noises will be introduced, such as thermal noise caused by resistance, channel thermal noise of FET, photon noise, dark current noise, and nonuniform light response noise. In this study, these methods are also applicable to data processing.

In behavior recognition, skeleton-based data analysis methods are most commonly applied. In addition to the advantages compared with other modal data, the skeleton sequence has the following three main characteristics: 1) Spatial information, there is a strong correlation between adjacent joints, so rich human structure information can be obtained within the frame; 2) temporal information, which can be used by interframes; 3) cooccurrence relationship in space-time domain, which is applicable when considering joints and bones. Therefore, many researchers use skeleton data for human behavior recognition or detection, and more and more researchers will use skeleton data. In skeleton sequence

Figure 1. Design of Behavior Recognition Framework



space modeling, RNN algorithms based on long and short-term memory are mostly used, but it is difficult to obtain feature data with strong discriminative properties (Qiao et al., 2019). In contrast, the covariance matrix can obtain high-dimensional features, which can effectively describe spatial features by using time series in the calculation of skeleton joint coordinates. The combination of the two can effectively improve the recognition accuracy, but also has its own shortcomings (Uddin et al., 2020). These models are not normalized when processing raw data, and skeletal data may lead to degradation of model performance. Traditional normalization methods are difficult to deal with dynamic change processing, so, in this study, the author proposed adaptive normalization algorithm. Min max normalization and mean normalization are suitable when the maximum and minimum values are clearly unchanged. If there are strict requirements on the processed data range, min max normalization or mean normalization should also be used. Z-score normalization can also be called standardization. The processed data show a distribution with a mean value of 0 and a standard deviation of 1. Standardization can be used when there are outliers, and the maximum and minimum values are not fixed. Standardization will change the state distribution of data but will not change the type of distribution. Z-score normalization is often used in neural networks. Nonlinear normalization is usually used in scenarios with large data differentiation. Sometimes, it is necessary to map the original values through some mathematical functions, such as logarithm and arctangent. Switchable normalization (SN) can determine an appropriate normalization operation for each normalization layer in a depth network. SN is not only easy to use, but also superior in performance. It is robust; indeed, its insensitivity to minibatch size keeps its accuracy stable under various batch size settings. Especially in visual tasks with limited batch size, such as object detection, instance segmentation, and video recognition. SN is applicable to various network structures, including CNNs and RNNs, and can solve a variety of visual tasks. The adaptive planning model is designed to normalize the skeleton sequence data according to its adaptive mean difference and standard deviation. Assuming that the skeleton sequence data are represented by X and each sequence data has d-dimensional data values, normalization is used to process the data, and the formula is expressed as follows:

$$x_j^{-i} = \left(x_j^i - u^i\right) \otimes \delta^2 \tag{1}$$

where j denotes the characteristics of the skeleton sequence data and x denotes the normalized measurement. Hadamard division can be expressed as follows:

$$x_{j,k}^{-i} = \frac{x_{j,k}^{-i} - u_k^i}{\delta_k^i} \tag{2}$$

In the global normalization process, both u and δ may not be optimal for the measurement, so it is necessary to normalize the frame series in its entirety and adjust these two parameters dynamically. The input values of each sliding window are influenced by taking the mean estimate, and the formula is the following Equation:

$$u^i = \frac{1}{T_i} i \sum\nolimits_{j=1}^{T_i} x_j^i \tag{3}$$

where u denotes the initial average value. The adaptive mean can be expressed as the following Equation:

$$a^i = W_u u^i + b_u \tag{4}$$

where a denotes the adaptive mean and W denotes the deep learning weight factor. The adaptive mean change of the skeleton sequence is expressed as the following Equation:

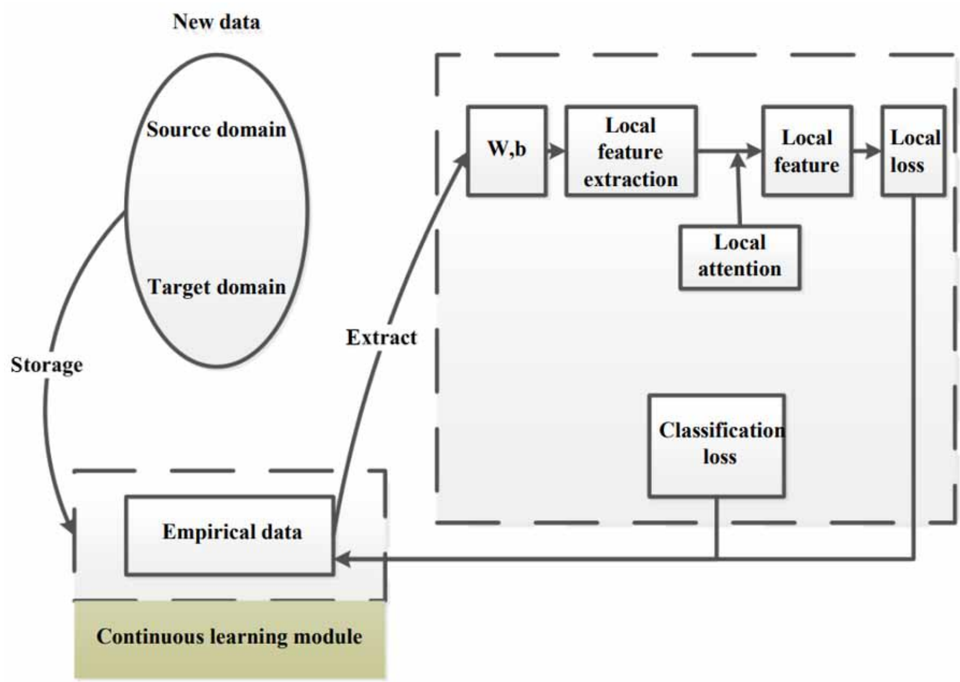
$$c_k^i = \sqrt{\frac{1}{T_i} \sum_{j=1}^{T_i} \left(x_{j,k}^i - a_k^i \right)} \tag{5}$$

Behavior Recognition Based on Improved Deep Learning Algorithms

In the field of behavior recognition, deep learning techniques have been fully developed, but many shortcomings still appear. Unlike traditional machine learning algorithms, deep learning algorithms require a large amount of data in order to ensure superior performance. When the sample dataset is very small, the simple linear model can also be better than the deep network model. Deep learning requires many datasets of different instances, from which the model learns the features to find and generates an output with probability vectors. Whether the performance of deep learning can be improved depends on the size of the dataset. The more parameters the model learns, the more data the training requires. Otherwise, problems with more dimensions and small data will lead to over fitting. In model training, the quality of data will directly affect the classification results. However, in the application of student behavior recognition in the university's Meiwei classroom, it costs much to obtain a large amount of data. Also in algorithm performance analysis, there are requirements for training samples and test samples, and the data are often difficult to meet the standard requirements. In real-time surveillance videos, traditional behavior recognition algorithms cannot be updated in a timely manner (Zhang, Huang et al., 2019). Therefore, in this study, the author proposed an adaptive model based on local attention, based on the existing general model, introducing new data, combining the concept of continuous learning, establishing empirical memory, and adjusting the results according to the real-time video. Figure 2 shows the system architecture.

Image classification is to classify images according to the semantic information contained in them. The steps of image classification are generally as follows: First, the features of the image are extracted, then the target model is obtained by training, and finally the extracted feature map is input into the target model for classification (Zahid et al., 2021). With the development of deep learning, more and more people gradually have studied the application of deep learning related technologies to image classification (Zheng et al., 2020). Deep learning image classification model is also a machine learning technology. In this framework design, the model is able to train on the source domain data to get the classification model and migrate it to the target domain (Zhang, Huang et al., 2019). Combined with the characteristics of student behavior analysis in college classrooms, the behavior data are preprocessed to extract features for the local behavior target domain (Zhao, 2021). Focusing on the local variation of each action, decomposition is performed. The local feature information is input to the discriminator, and the migrated features are reinforced by local attention to induce feature matching of datasets with different source and target domains (Zhang et al., 2021). The obtained feature values are then fed to the classifier to obtain more accurate classification. Classification is a very important method of data mining. The concept of classification is to learn a classification function or construct a classification model on the basis of existing data. This function or model can map the data records in the database to one of the given categories, so that it can be applied to data prediction. Classifier is a general term for the methods of classifying samples in data mining, including decision tree, logical regression, naive Bayes, neural network, and other algorithms. During the operation of the system, the data are continuously updated and the previous parameters are constantly changed, which, in turn, affects the behavior recognition accuracy. In order to improve recognition stability, empirical

Figure 2. Improved Continuous Learning Behavior Identification Framework



values are introduced and some of the source domain data are stored. When the model is updated, in addition to the new data, the data are filtered from the empirical data for recognition to improve the dynamic update capability. Key point identification directly regresses the probability value of each type of key point in the form of thermal map. The most obvious color in each thermal map is most likely to be the location of coordinate points. The prediction accuracy of key point position is relatively high. In addition, different key points are divided for independent recognition, and the task is divided into multiple subtasks, so the recognition difficulty is lower than that of traditional image detection and recognition. It is these two advantages that make human body key points less susceptible to background interference.

The attentional learning migration module is built on the basis of cognitive imitation and emphasizes the importance of input data. In this module design, the idea lies on a small portion of important data computationally based, and the rest of the data depend on the context, using a gradient descent approach in training data learning. Gradient descent is a kind of iterative method, which can be used to solve the least square problem (linear and nonlinear). When solving the model parameters of machine learning algorithms, that is, unconstrained optimization problems, gradient descent is one of the most commonly used methods. When solving the minimum value of the loss function, the gradient descent method can be used to iteratively solve step by step to obtain the minimum loss function and model parameter values. This theory has many applications in machine learning algorithms and is popularized in the field of computer vision. Deep learning algorithms are introduced in which are additional neural networks that assign weights by observing the input data and thus filter out important parameter information. In the field of behavior recognition, the behavior of different individuals varies greatly, so the behavioral data will not match the training samples. Also, to ensure that the data features are more specific, a local attention mechanism is used. Assuming that x is used to denote the input vector and that Z, a , and g denote the feature vector, attention vector, and element multiplication results, formally, the formula can be expressed as the following Equation:

$$\begin{aligned} a &= f_{\Phi}(x) \\ g &= a \otimes z \end{aligned} \tag{6}$$

where z denotes the output result of the neural network with another parameter. The input data may be migratory in some cases, and, in order to be able to match the input data features of different domains, discriminators are introduced, each of which is responsible for the source and target domains. The loss function can be expressed as follows:

$$L_1 = \frac{1}{nK} \sum_{k=1}^{K} \sum_{x \in D_s} L_{entropy}\left(D^k\left(G_{if}\right), d_i\right) \tag{7}$$

where L denotes the defined information entropy loss value, K denotes the number of discriminators, and d denotes the domain label corresponding to the input data. The discriminant domain determines the probability value based on the input data can be expressed as the following Equation:

$$d_i^k = D^k\left(G_{if}\left(x_i\right)\right) \tag{8}$$

The probability values range from 0 to 1. When the probability value is close to 1, it means that the data comes from the source domain and close to 0 indicates that it comes from the target domain. The model wants to be more migratory and introduces the attention mechanism, and the local attention can be expressed as follows:

$$w_i^k = 1 - h\left(d_i^k\right) \tag{9}$$

The formula for calculating h in Equation is expressed in the following Equation:

$$h(p) = -\sum_i p_i \log\left(p_i\right) \tag{10}$$

With the introduction of the local attention eigenvalues, the formula is transformed into the following Equation:

$$g_i^k = \left(1 + w_i^k\right) G_{if}\left(x_i\right) \tag{11}$$

In the module design, the source domain classification loss function cannot be neglected and the total loss function is calculated. The formula is expressed as the following Equation:

$$J_\theta = aL_s + L_a \tag{12}$$

After the calculation, it is possible to obtain greater attention weights. In designing a behavior recognition system, it is necessary to be able to ensure that the data is updated in a timely manner and that the old empirical data remains important. Continuous learning is the algorithm that continuously expands new knowledge, and, in this system, deep learning algorithms are increasingly valued as a common tool. Reinforcement learning algorithm is able to achieve a stable distribution of data through the formation of new data. However, as more and more persistent problems emerge, it is difficult to collect new data and the training strategy is less feasible, so improvements to it are needed.

In model learning, it is necessary to filter some data from empirical data. However, considering that the probability of these data appearing is very low, the sampling probability is low by just random sampling, so the sampling is performed by defining the priority, and the formula is expressed as the following Equation:

$$p_i = \left|J_i\right| + e \tag{13}$$

where e denotes the constant and J denotes the loss value. For source-domain data, the data with large loss values are retained. Large loss values indicate that the model is not strong in classification, so they are retained to facilitate later use. For empirical data, the data cannot be selected by evaluating the priority alone, which is prone to overfitting. Therefore, the sampling is changed by defining the priority, and the formula is expressed as the following Equation:

$$p_i = \frac{p_i^a}{\sum_k p_k^a} \tag{14}$$

where a denotes the hyperparameter, ranging from 0 to 1. When the value is 0, the model takes the value from experience, and when the value is 1, the data are filtered according to the priority. The loss function of the whole network can be expressed as the following Equation:

$$L_{all} = J_{new} + \delta \frac{1}{m} \sum J(x_i) \tag{15}$$

where δ denotes the hyperparameter, m denotes the number of data, and J denotes the new data loss value.

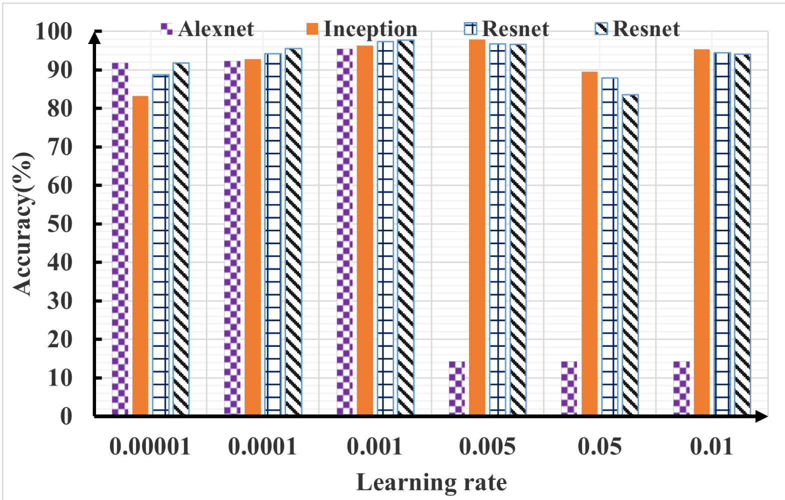
RESULT AND DISCUSSION

Improved Deep Learning Algorithm Parameter Optimization

In the whole improved deep learning algorithm, which covers data preprocessing, model selection, and parameter tuning processes, optimization of parameters is required to improve behavior recognition accuracy. In the dataset analysis, it is randomly classified according to the ratio, and it is divided into training set, test set, and validation set. The training set is divided into the original data and the expanded data. In the image preprocessing, the blurred image data are removed first, and the image data that are not relevant to the analysis of student behavior are reduced, while the data are expanded. The acquired image data are kept at 224*224 size. Some networks may stop training when the loss value is very low after 10 iterations. Under the same conditions, in the experiment, the learning rate is varied and the accuracy is measured. The learning rate was set at 0.0001~0.01, and the dataset was analyzed using images where students accounted for a relatively small number of images, and the batch_size was set at 8. The accuracy was measured under different models. Figure 3 shows the measurement results.

As the change in the data in Figure 3 show, the accuracy has not changed much, at the end, as the value of the loss function keeps increasing. When the learning rate is set in the range of 0.0001

Figure 3. Effect of Learning Rate on Classification Accuracy



to 0.001, the accuracy is continuously increasing. When the learning rate keeps increasing, the experimental accuracy starts to show a gradual decrease, and there is an inability to converge. When the learning rate was set to 0.05 or 0.0001, the loss value could not decrease or appeared to decrease slowly, so the optimal learning rate parameter was set to 0.005.

Behavior Recognition Experiment Simulation

To determine the classification accuracy of behavior recognition, the author conducted tests using the same type of actions on different datasets, carried out on the PyTorch platform. The researcher determined the effect of different empirical memory sizes on the accuracy rate.

To compare the classification accuracy under different empirical memory sizes, the size of the data set is denoted by M. The empirical memory capacity sizes are set to 0, M/2, M/4, M/8, M/16, and M/32. Figure 4 and Figure 5 show the results of the model classification accuracy measurement.

Figure 4. Test Results of Different Empirical Capacity in Office-31 Dataset

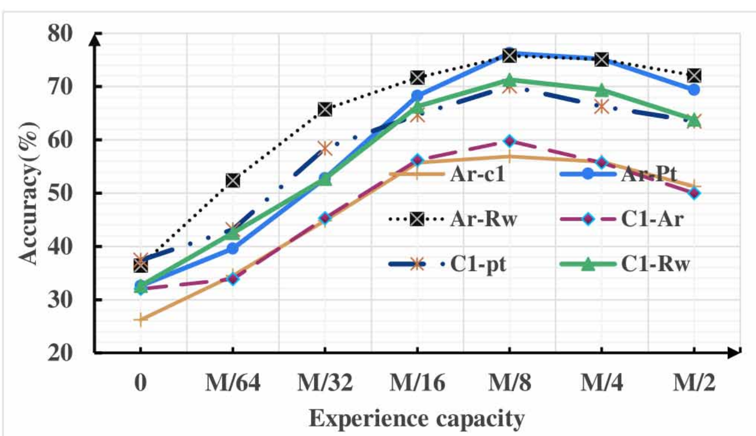
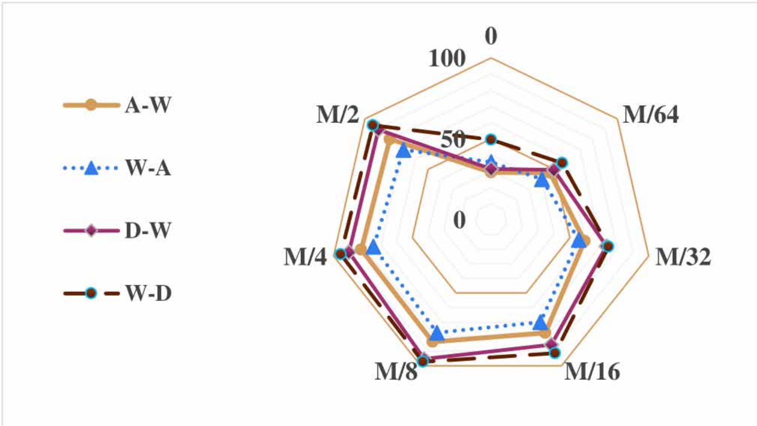


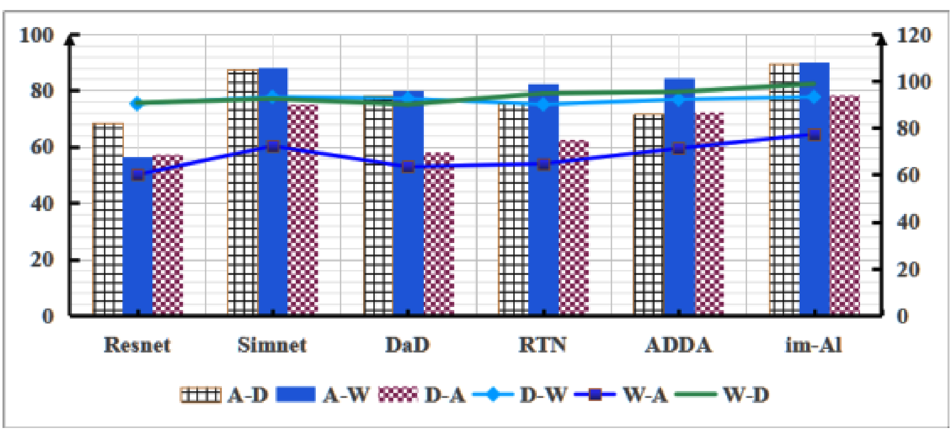
Figure 5. Test Results of Different Empirical Capacity in Office Home Dataset



Figures 5 and 6 show that, in the current network structure design, the empirical memory size is set to M/8 to achieve the highest classification accuracy. When the empirical memory capacity is less than this value, the classification accuracy increases as the capacity increases. This may be because the increased capacity enables more prior knowledge to be obtained and better differentiate behavioral actions. When the empirical size capacity is larger than this value, the accuracy gradually starts to decrease, which may be because each data screens probability values; when the capacity is too large, the probability sizes are similar to get important features and do not produce much help for classification.

The author measured the accuracy of the algorithm in this study and other learning migration algorithms on different datasets, with the empirical memory size set to M/8; Figure 6 shows the results. The variation of the data in Figure 6 evidence that the classification accuracy of the model with the addition of local attention reached the highest in the migration task. In other models, the classification accuracy of the model with the addition of local attention is basically the same as that of the MEDA algorithm. This indicates that the classification accuracy of the proposed algorithm with local attention is higher compared to other algorithms, probably because the model is able to migrate knowledge for models such as translation and rotation as well.

Figure 6. Accuracy of Different Models



The author analysed the algorithm of this study with other behavioral analysis algorithms. In the experimental simulation, the empirical memory size is set at M/8; Figure 7 shows the measurement results.

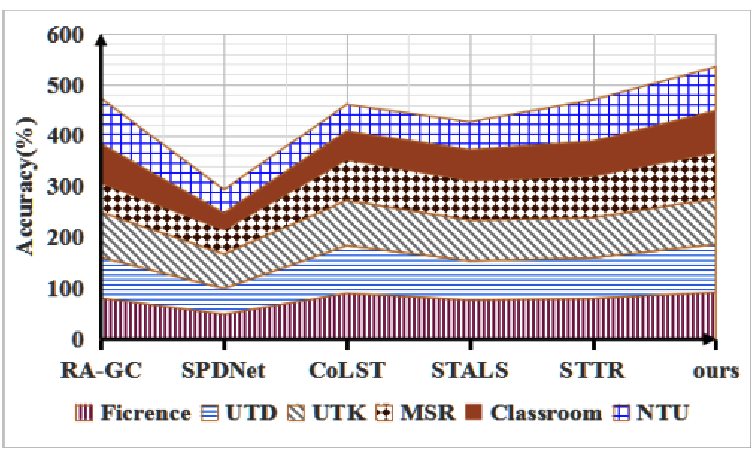
The data in the Figure 7 highlight that the proposed algorithm shows superiority on all datasets and can improve the classification accuracy by 2%-7%, compared to other behavioral analysis algorithms. In the public test dataset, the accuracy of the proposed algorithm is also improved compared with other algorithms. Compared with the current common GCN algorithms, the proposed algorithm has some advantages on all datasets.

CONCLUSION

At present, the analysis of students' classroom behavior is too dependent on the analysis of teachers, which is inefficient. In the application of intelligent algorithm, the traditional machine learning algorithm has limited modeling ability and is vulnerable to the influence of illumination, occlusion, and other factors, which cannot meet the needs of behavior recognition. Based on this, in this research the author studied the behavior analysis of college students in the classroom based on the improved deep learning algorithm, proposed a solution to data preprocessing and higher-order spatiotemporal extraction algorithm, established an adaptive learning model, introduced an attention model, and verified the effectiveness of the algorithm through simulation analysis. The results showed that the proposed algorithm is superior in all datasets, and can improve the classification accuracy by 2% - 7%, compared with other behavior analysis algorithms. Compared with other algorithms, the accuracy of the algorithm is also improved in the public test dataset. Compared with the current common GCN algorithm, this algorithm has certain advantages in all datasets. The student's classroom behavior recognition method based on the improved in-depth learning algorithm can be used to identify the typical classroom behaviors of students (including reading blackboard, reading, sleeping, turning around, and playing mobile phone), which can reflect the students' learning status in a timely and effective manner, and help teachers accurately grasp the students' classroom learning, thus helping intelligent classroom teaching.

Importantly, with regard to empirical memory selection knowledge, the author adopted the time randomization strategy, which can examine other empirical memory extraction methods to extract more valuable data and improve the performance of the algorithm. The focus of the research was on data preprocessing and model exploration, which can be considered from the improvement of the

Figure 7. Accuracy of Different Algorithms



network structure for further research. In addition, the author focused on the identification of students' classroom behavior. In this study, the data the author used were to let students pose for shooting, imitate students' classroom behavior from multiple angles, and shoot and produce a dataset. In fact, students' classroom behavior is generally based on video observation. The method of identifying students' classroom behavior the author proposed in this paper cannot be directly applied to students' classroom. The next step is to study student behavior recognition based on real student classroom video.

FUNDING STATEMENT

This work was supported by the Education Department of Guangdong Province, Guangdong higher education teaching research and reform project "Exploration of Ideological and Political Teaching Practice of Applied Undergraduate Courses -- Taking Art and Design Courses as an Example," Project No. 2020SJJXGG03, and the E-Education Department of Guangdong Province, teaching quality and teaching reform project of undergraduate universities in Guangdong Province "Teaching and Research Office of introduction to Design," Project No. 2021SJZLGC03.

ACKNOWLEDGMENT

The author would like to show sincere thanks to the technicians who contributed to this research.

REFERENCES

Byeon, Y. H., Kim, D., Lee, J., & Kwak, K. C. (2021). Body and hand-object ROI-based behavior recognition using deep learning. *Sensors (Basel)*, *21*(5), 1838. doi:10.3390/s21051838 PMID:33800776

Chen, X., Qi, L., Yang, Y., Luo, Q., Postolache, O., Tang, J., & Wu, H. (2020). Video-Based Detection Infrastructure Enhancement for Automated Ship Recognition and Behavior Analysis. *Journal of Advanced Transportation*. https://link.gale.com/apps/doc/A622268315/AONE?u=anon~d46a47ec&sid=googleScholar&xid=dd28e995

Cheng, H., Ma, P., Dong, G., Zhang, S., Wei, J., & Qin, Q. (2022). Characteristics of carboniferous volcanic reservoirs in Beisantai oilfield, Junggar Basin. *Mathematical Problems in Engineering*. 10.1155/2022/7800630

Cheng, H., Wei, J., & Cheng, Z. (2022). Study on sedimentary facies and reservoir characteristics of Paleogene sandstone in Yingmaili block, Tarim basin. *Geofluids*. 10.1155/2022/1445395

Cheng, H., Yang, D., Lu, C., Qin, Q., & Cadasse, D. (2022). Intelligent oil production stratified water injection technology. *Wireless Communications and Mobile Computing*. 10.1155/2022/3954446

Chonggao, P. (2021). Simulation of student classroom behavior recognition based on cluster analysis and random forest algorithm. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 2421–2431. doi:10.3233/JIFS-189237

Dai, C., Liu, X., Lai, J., Li, P., & Chao, H. C. (2019). Human behavior deep recognition architecture for smart city applications in the 5G environment. *IEEE Network*, *33*(5), 206–211. doi:10.1109/MNET.2019.1800310

Hsueh, Y. L., Lie, W. N., & Guo, G. Y. (2020). Human behavior recognition from multiview videos. *Information Sciences*, *517*, 275–296. doi:10.1016/j.ins.2020.01.002

Jiang, Z., Crookes, D., Green, B. D., Zhao, Y., Ma, H., Li, L., Zhang, S., Tao, D., & Zhou, H. (2018). Context-aware mouse behavior recognition using hidden markov models. *IEEE Transactions on Image Processing*, *28*(3), 1133–1148. doi:10.1109/TIP.2018.2875335 PMID:30307863

Konstantinova, M., Shedenko, K., Boyko, L., Perepelkina, O., Kazimirova, E., & Latanov, A. (2018). Gaze behavior in mixed facial emotion recognition task. *International Journal of Psychophysiology*, *131*, S129. doi:10.1016/j.ijpsycho.2018.07.348

Liang, R., Zhi, H., & Kamruzzaman, M. M. (2019). Methods of moving target detection and behavior recognition in intelligent vision monitoring. *Acta Microscopica*, *28*(4), 750–759.

Liu, H., Liu, Y., Zhang, R., & Wu, X. (2021). Student behavior recognition from heterogeneous view perception in class based on 3-D multiscale residual dense network for the analysis of case teaching. *Frontiers in Neurorobotics*, *15*, 63. doi:10.3389/fnbot.2021.675827 PMID:34393749

Liu, H., & Long, Z. (2020). An improved deep learning model for predicting stock market price time series. *Digital Signal Processing*, *102*, 102741. doi:10.1016/j.dsp.2020.102741

Lu, C. (2021). Multifeature fusion human motion behavior recognition algorithm using deep reinforcement learning. *Mobile Information Systems*. 10.1155/2021/2199930

Mabrouk, A. B., & Zagrouba, E. (2018). Abnormal behavior recognition for intelligent video surveillance systems: A review. *Expert Systems with Applications*, *91*, 480–491. doi:10.1016/j.eswa.2017.09.029

Qiao, W., Tian, W., Tian, Y., Yang, Q., Wang, Y., & Zhang, J. (2019). The forecasting of PM2. 5 using a hybrid model based on wavelet transform and an improved deep learning algorithm. *IEEE Access: Practical Innovations, Open Solutions*, *7*, 142814–142825. doi:10.1109/ACCESS.2019.2944755

Uddin, M. Z., Hassan, M. M., Alsanad, A., & Savaglio, C. (2020). A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare. *Information Fusion*, *55*, 105–115. doi:10.1016/j.inffus.2019.08.004

Zahid, F. B., Ong, Z. C., Khoo, S. Y., & Salleh, M. F. M. (2021). Inertial sensor based human behavior recognition in modal testing using machine learning approach. *Measurement Science & Technology*, *32*(11), 115905. doi:10.1088/1361-6501/ac1612

Zhang, L., Huang, J., & Liu, L. (2019). Improved deep learning network based in combination with cost-sensitive learning for early detection of ovarian cancer in color ultrasound detecting system. *Journal of Medical Systems*, *43*(8), 1–9. doi:10.1007/s10916-019-1356-8 PMID:31254110

Zhang, W., Cheng, Z., Cheng, H., Qin, Q., & Wang, M. (2021). Research of tight gas reservoir simulation technology. In *Proceedings of the IOP Conference Series: Earth and Environmental Science* (Vol. 804, No. 2, p. 022046). IOP Publishing. doi:10.1088/1755-1315/804/2/022046

Zhang, Y., Li, J., Guo, Y., Xu, C., Bao, J., & Song, Y. (2019). Vehicle driving behavior recognition based on multi-view convolutional neural network with joint data augmentation. *IEEE Transactions on Vehicular Technology*, *68*(5), 4223–4234. doi:10.1109/TVT.2019.2903110

Zhao, G. (2021). Feature recognition of human motion behavior based on depth sequence analysis. *Complexity*. 10.1155/2021/4104716

Zheng, B., Yun, D., & Liang, Y. (2020). Research on behavior recognition based on feature fusion of automatic coder and recurrent neural network. *Journal of Intelligent & Fuzzy Systems*, *39*(6), 8927–8935. doi:10.3233/JIFS-189290