Research Article

A Convolutional Neural Network (CNN) Based Approach for the Recognition and Evaluation of Classroom Teaching Behavior

Guang Li, Fangfang Liu, Yuping Wang, Yongde Guo, Liang Xiao, and Linkai Zhu

1Institute of Data Science, City University of Macau, Macau, China
2Capital Normal University High School, Beijing 100089, China
3Beijing Haidian Experimental Middle School, Beijing 100089, China
4Mathematics and Data Science in the School of Information Technology, Macao University of Science and Technology, Macau, China
5Trusted Computing and Information Assurance Laboratory, Institute of Software Chinese Academy of Sciences, Beijing 100190, China
6Institute of Data Science, City University of Macau, Macau, China

Correspondence should be addressed to Yongde Guo; wtkuok@cityu.mo

Received 11 August 2021; Revised 18 September 2021; Accepted 24 September 2021; Published 20 October 2021

Academic Editor: Mian Ahmad Jan

Copyright © 2021 Guang Li et al. 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.

To improve classroom teaching behavior recognition and evaluation accuracy, this paper proposes a new model based on deep learning. First, we obtain the classroom teaching behavior characteristic data through the SVM’s linear separable initial and determine the relationship of the characteristic sample data in the hyperplane. Then, we obtain the heterogeneous support vector of the online learning behavior characteristic sample data in the SVM’s hyperplane and complete the extraction of data with the help of convolutional neural networks. We then use a decision matrix to analyze the hierarchical process, determine the weight of classroom teaching behavior indicators, verify their consistency, and complete the evaluation by calculating the membership of evaluation factors. The experimental results show that the identification and evaluation method of classroom teaching behavior in this paper can effectively improve the identification accuracy of the classroom teaching behavior.

1. Introduction

In the realm of education, recognizing classroom teaching behavior has long been a study emphasis. Teachers in conventional classrooms assess students’ learning levels by watching their behavior and then providing teaching feedback. However, this technique is time-consuming and labor-intensive, and it cannot satisfy the demands of today’s intelligent learning environment for large-scale classroom analysis [1]. As a result, it is critical to investigate artificial intelligence technologies for automatically identifying student’s classroom behavior. This study has significant implications for enhancing classroom teaching quality and assisting educational decision-making and administration. The advent of the big data age has caused information to expand at an exponential rate. One of the key issues in big data analysis is how to use these data in a reasonable and efficient manner [2]. Artificial intelligence technologies can effectively screen data and turn abstract data into useful knowledge. Deep learning, which has recently gained popularity, offers a solution to this problem. Deep learning offers greater accuracy, can automatically extract information, is more adaptable, and is easier to develop than standard machine learning algorithms. Deep learning in education is steadily increasing as the deep learning movement grows. Teaching activities have long been the most important and fundamental activities in school education research, as well as the most important component of the school’s teaching and educating function. The classroom behavior of students will immediately reflect the instructional effect. As a result, analyzing students’ classroom behavior will aid in the improvement of teaching and...
classroom management. Students’ classroom behavior may be divided into two categories: classroom engagement and general classroom environment. Nonlearning behavior is the other. Teachers create an effective appraisal of students in the conventional classroom by watching students’ class status and assignment completion in the classroom. However, because it is impossible for teachers to monitor all pupils in a systematic and comprehensive manner, the teaching impact has not been fed back in time, prompting the development of a recording and broadcasting system. The problem has been alleviated to some extent. The conventional technique of analyzing instructor and student behavior is still a manual observation of instructional films, which is time-consuming, tedious, and inefficient [3]. At this time, society has entered a more intelligent epoch. The transition of the education sector will be unavoidable, whether it comes from the exterior organizational structure or the core internal content. The reform will be focused on intelligent education and intelligent management among them. One of the technological topics that needs to be explored in the future is how to utilize intelligent teaching to develop effective classroom evaluation and enhance teaching quality. It can assist instructors to enhance their teaching to a large amount by solving the problem of teachers finding it difficult to watch each student’s classroom behavior in detail [4].

Zhao et al. [5] proposed a multidimensional feature fusion-based intelligent analytic technique for training behavior. This method creates a multidimensional feature fusion intelligent teaching behavior analysis mode, with four core elements: “visual auditory feature-based teaching behavior analysis coding system, auditory feature recognition of teaching behavior, visual feature recognition of teaching behavior, and visual presentation of data in the teaching analysis process.” They then proposed three realistic paths: “visual feature-based,” “auditory feature-based,” and “fusion feature-based.” They preliminarily analyzed 43 classroom teaching videos, extracting the visual characteristics of teaching behavior in the time dimension, to provide a reference for teaching behavior intelligent analysis activities such as “one teacher, one pedagogy.” This approach has a strong theoretical foundation and a broad influence in practice, but it still needs to be refined. A follow-up study [6] suggested a compound feature and deep learning-based classroom teaching behavior recognition system.

To address the limitations of the preceding techniques, this study uses a convolutional neural network to identify and evaluate classroom teaching behavior in order to improve teaching quality and contribute to educational growth. In our work, we use the crowd as the main object, extract the crowd’s static information using the foreground extraction method, obtain the crowd’s dynamic information using changes in crowd movement, use the CNN model to learn two different crowd behavior characteristics, and then combine these two characteristics to analyze the spread. Population behavior analysis is influenced by the location and interval of population data extraction.

According to the experimental results, these two population characteristics can better characterize population status in the geographical dimension and population change in the temporal dimension. Data intervals and locations that are reasonable can significantly increase the expressive ability of crowd data. Finally, existing crowd behavior analysis approaches are compared to the suggested method. The suggested approach’s efficiency is confirmed by quantitative and qualitative testing findings, which demonstrate that the proposed method can provide a better confusion matrix and higher accuracy. This approach offers a high level of accuracy when it comes to teaching behavior recognition. There are still some modifications in the identified items; therefore, the recognition and assessment cannot be done simultaneously.

2. Extraction and Identification of Classroom Teaching Behavior Characteristics

2.1. Extraction of Classroom Teaching Behavior Characteristics. Teaching behavior research has progressively grown into a distinct study subject since the mid-twentieth century. It has steadily concentrated on the microlevel of classroom instruction in recent years. Classroom teaching behavior is a multilevel and intricate synthesis of subject interaction behavior in the classroom. It differs from teaching concepts and ideas in that it is explicit, dynamic, individual, intentional, and situational, with several categories [7]. It primarily consists of teaching, auxiliary teaching, and classroom management. Scholars combine several categorization bases to create a wide range of categories, as shown in Table 1.

Based on the classification, the classification of classroom teaching behavior mainly focuses on the essential elements in the classroom, including teachers, students, teaching content, and teaching media (complex media and soft media). Habermas, a social theorist, once proposed that human communication activities exist between more than two subjects with language ability and actionability, and there is a relationship between speech and external action among subjects. Classroom teaching activities are interactive activities between teachers and students in classroom scenes. To achieve common classroom goals, the sum of teaching, learning, and other related behaviors in interactive activities is collectively referred to as classroom teaching behavior. Therefore, the author divided classroom teaching behavior into teachers’ teaching behavior, students’ learning behavior, and other behaviors.

Teaching behavior is considered the behavior taken by teachers and students to achieve certain teaching objectives in the teaching process. Teachers’ classroom teaching behavior refers to all the interactions between teachers and students in the classroom and the behavior related to teaching and learning. Since teachers and students are the composite subjects of educational activities (including teaching), based on the above definitions of experts, classroom teaching behavior is the sum of teachers’ teaching, students’ learning, and other related behaviors in the classroom environment to achieve common mathematical goals [8]. Based on the above classification results, its characteristics are extracted first to realize the recognition and evaluation of classroom teaching behavior.
In the extraction of classroom behavior, SVM [7, 9] has been extensively used in the existing literature. The classroom teaching behavior feature data are regarded as a collection of finite samples in the two-dimensional space, and the data are initially obtained through SVM linear segmentation. Assuming that the sample set, $A$ of the feature data, is $\{(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)\}$, the classroom teaching behavior feature sample data appear in the hyperplane as follows:

$$
\begin{align*}
A : V^T a + B &= 0, \\
A_1 : V^T a + B &= 1, \\
\vdots & \\
A_n : V^T a + B &< 1.
\end{align*}
$$

(1)

In the formula, $v$ represents the normal vector of the characteristic data within this space, $B$ represents the location of the data in the space, and $A_1, A_n$ represent the support vectors of the two hyperplanes, respectively.

To ensure the maximum [10] of classroom teaching behavior feature data in the hyperplane, it should meet

$$
\begin{align*}
&\text{S.T. } B_i(V^T E_i + B), \\
&\text{MIN } \frac{1}{3} \left\| V \right\|^2.
\end{align*}
$$

(2)

In the formula, $E_i$ represents the Lagrange multiplier. Furthermore, the classroom teaching behavior characteristic data exist in the SVM hyperplane, and the characteristic sample data are expressed as [8]

$$
C = \frac{V^T (1 - B + 1(U))}{\left\| V \right\|},
$$

(3)

where $C$ represents the heterogeneous class support vector values of characteristic data.

Due to many factors, the characteristic data of classroom teaching behavior cannot always maintain a linear separable state. Therefore, we should also consider the nonlinear characteristics formed by some external factors [9]. The nonlinear characteristics of classroom teaching behavior characteristic data can be transformed into linear separable data by reducing the dimension of the data in the hyperplane. The nonlinear data features obtained after calculation are minimized to convert the nonlinear data features of classroom teaching behavior data into the minimum value; that is,

$$
\begin{align*}
&\text{S.T. } b_i(R^T f(v) + d) > 1, \\
i &= 1, 2 \ldots n,
\end{align*}
$$

(4)

The feature data of classroom teaching behavior are first obtained using SVM linear separability, then the relationship of feature sample data in the hyperplane is determined, the heterogeneous support vector of online learning behavior feature sample data in the SVM hyperplane is obtained, and the feature data extraction is completed.

2.2. Recognition of Classroom Teaching Behavior Based on Artificial Intelligence. The combination of AI and learning science has given rise to a new field—educational artificial intelligence (EAI). A large number of EAI systems integrating AI technology, e.g., data mining, have been applied to schools for efficient utilization of big data generated within the campus by analyzing and predicting students’ learning performance. It provides students with more personalized and diversified learning. AI technologies widely used in the field of education include image recognition, speech recognition, semantic recognition, and machine translation. Therefore, this paper uses a deep learning-based method named convolutional neural network to recognize classroom teaching behavior [11].

CNN is one of the most successful fields in deep learning algorithm applications, mainly composed of convolutional, pooling, and full connection layers. In addition, some common layers, such as the activation function layer, can significantly reduce the network training time, enhancing the generalization ability of the network and preventing the dropout layer with an overfitting problem. CNN is a hierarchical model, where first, the original data are extracted through the convolutional layer. The pooling layer then takes the dimension reduction data as the input of the lower convolution, repeating the convolution and pooling operation. The high-level semantic information in the original data is gradually transmitted from low to high to the last level. The relevant parameters are updated by calculating the difference between the actual and predicted values and then combined with the backpropagation algorithm, thus yielding a convergence model [12]. Its basic structure is shown in Figure 1.
In classroom teaching behavior recognition, the core of constructing a convolutional neural network is the convolution layer. After smoothing the input signal through the convolution kernel and filter, it is biased to extract local image features. The depth of each convolution kernel is consistent with the input data, so the convolution kernel smooths along the height and width to map a new feature map. The eigenvalues extracted in the first few layers mainly refer to low-order features such as vertical and horizontal edges. The latter layers combine the low-order features into high-order ones. These high-order features can distinguish these data and map them to the category space for classification operation. At this time, the mapped classroom teaching behavior data matrix is

\[
S = \begin{bmatrix}
S_{11} & S_{12} & \cdots & S_{1n} \\
S_{21} & S_{22} & \cdots & S_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
S_{m1} & A_{m2} & \cdots & S_{mn}
\end{bmatrix}
\]  

(5)

According to the data matrix obtained from the convolution layer, we determine the processing parameters of the pooling layer. Compared with the convolution layer, there are no parameters to learn, select the maximum value or mean value from the target region, and then form these maximum values or mean values into a new feature map. Pooling is essentially a process of downsampling to reduce the number of parameters to reduce the computational complexity of the network. To some extent, it can prevent overfitting problems [13]. Pooling is to reduce the spatial operation of the target area without changing the number of channels of input and output data. Moreover, when the data change slightly, pooling will return the same result, which enhances the robustness of the network. At this time, the recognition result is

\[
L^{n+1} = \frac{1}{q} \sum |P - S_{i} |^C Z
\]  

(6)

Input the recognition result data obtained above to the last layer of the CNN as the full connection layer, which is connected with all the feature images of the previous layer and usually converts the two-dimensional image output by convolution into a one-dimensional vector. Because the full connection layer does not retain the spatial information of the image, it cannot maintain the spatial structure and then change the high dimension into the low dimension through dimension transformation while retaining helpful information. In the CNN, the full connection layer takes the feature vector as the classifier’s input to obtain the output of the corresponding category to realize the original data’s recognition. At this time, the output classroom teaching behavior recognition results are optimized by the objective function to obtain

\[
f(z) = \frac{1}{1 + e^{-z}}
\]  

(7)

In the formula, \( f(z) \) represents the ultimately obtained identification results, and \( e \) represents the corrected error value.

In the recognition of classroom teaching behavior, we use a convolutional neural network, pooling layer, and full connection layer to get the recognition of classroom teaching behavior with the help of the objective function.

3. Research on Classroom Teaching Behavior Evaluation

Based on the above identified classroom teaching behavior, we carry out the evaluation analysis of classroom teaching behavior. Firstly, the weight of classroom teaching behavior data is calculated by constructing the confusion matrix.

The hierarchical structure can reflect the relationship between factors. Using the analytic hierarchy process to form a decision matrix is a quantitative treatment of subjective judgment based on qualitative judgment. It is an essential step from qualitative to quantitative analysis. Because at the standard level, each standard occupies a different proportion in the minds of different decision makers. To ensure the quality of the obtained index judgment, the classroom teaching evaluation questionnaire and index scoring scale were distributed to 7 experts of the speciality, who were invited to compare and score the indicators. The questionnaire results were summarized and statistically combined with the judgment given by the experts to form a matrix [14].

Now, the \( n \) comparison factors, \( X = (X_1 \ldots X_n) \), have the effect of an uncertainty factor 2. To provide real and credible data for the study, we will establish the comparison matrix by comparing the two factors to each other. Two
factors in each comparison and the full comparison results are represented by the matrix:

$$A = (a_{ij})_{n \times m}$$

(8)

We can represent the Z − X confusion matrix as

$$a_{ij} = \frac{1}{a_{ij}}$$

(9)

According to the above constructed confusion matrix, we complete the determination of the weight of classroom teaching behavior as

$$G[f] = \int H(\sum_{i=1}^{k} \varepsilon_k)qdz.$$  

(10)

In the formula, \(G[f]\) represents classroom teaching behavior weights, \(H\) stands for the number of judgments, and \(\varepsilon_k\) represents the scale factor.

According to the above determined classroom teaching behavior weight value, to achieve its evaluation, it needs to calculate its consistency as

$$w_k = \sum_{j=1}^{m} b_j c_{ij} \quad (i = 1, \ldots n).$$

(11)

The weight randomness of classroom teaching behavior data is

$$CR = \frac{\sum_{j=1}^{m} b_j CI_j}{\sum_{j=1}^{m} b_j RI_j}.$$  

(12)

Based on this, classroom evaluation is divided into two evaluation subjects to evaluate the classroom with peer teacher evaluation and expert evaluation [15]. Suppose the total collection of classroom evaluation is \(U\), peer teacher evaluation is \(U_1\), and expert evaluation is \(U_2\); then, the classroom teaching evaluation can be expressed as follows:

$$U = \{U_1, U_2\}.$$  

(13)

The set of evaluation factors in evaluation is set to

$$U_1 = \{U_{11}, U_{12}, U_{13}, U_{14}\}.$$  

(14)

At this time, the subset of the classroom teaching behavior membership is constructed as follows:

$$R_i = (r_{i1}, r_{i2}, \ldots, r_{im}),$$  

(15)

where \(R_i\) represents the evaluation indicators in the evaluation collection and \(r_{i1}\) represents the membership of the evaluation elements. The final evaluation results of the classroom teaching behavior can be shown by the following equation:

$$r_i = \frac{v_i}{p_n},$$  

(16)

where \(r_i\) represents the evaluation value, \(v_i\) represents the number of factors evaluated, and \(p_n\) represents the number of factors evaluated.

In the evaluation of the classroom teaching behavior, we first construct the confusion matrix through the hierarchical analysis method, then determine the weight of the classroom teaching behavior indicators, verify their consistency, and complete the evaluation of the classroom teaching behavior.

4. Experimental Analyses

To verify the scientific effectiveness of the proposed method, in this section, we perform various types of experiment and present the results.

4.1. Experimental Environment. In the experiment, 100 English students majoring in English in the 20th grade of a university were selected as the experimental objects to collect the classroom teaching behavior data of students in the major and identify and evaluate their classroom teaching behavior with the number of students raising their hands, the number of students answering questions, and students' positive behavior research data. The operating system is Nvidia 1050 GPU with 8 GB of memory. The number of experimental iterations is 100, and the data processing results all process the experimental data through SPSS.

4.2. Experimental Index Setting. This section compares our method in this paper with the methods in [5] and [6]. We use the accuracy of the classroom teaching behavior recognition and the evaluation error for the comparison.

4.3. Analysis of Experimental Results. By analyzing the data in Figures 2 and 3 under the same experimental conditions, the accuracy of sample classroom teaching behavior recognition is analyzed using the methods of this paper, Zhao et al. [5], and Ya-jun et al. [6]. Among them, the accuracy of our proposed method is always higher than that of the other two methods. When the number of iterations is 40, the identification accuracy of our proposed method reaches the highest value, about 98%. When the number of iterations is 80, the highest accuracy of the method in [5] is about 91%. When the number of iterations is 40, the highest accuracy of the method in [6] is about 89%. In contrast, our proposed method has the highest accuracy in identifying sample classroom teaching behavior. This is because we use a convolutional neural network to recognize classroom teaching behavior recognition.

In order to further verify the effectiveness of our method, we compared the evaluation results of sample classroom teaching behaviors and conducted experimental analysis on the results. The evaluation error is shown in Figure 4.

By analyzing Figure 4, it can be seen that there is a certain gap in the evaluation error of the evaluation results of sample classroom teaching behavior by using the methods of this paper, Zhao et al. [5], and Ya-jun et al. [6]. When the number of iterations is 20, the error of our method on the evaluation of sample classroom teaching behavior is about 1.9%, the error of Zhao et al.'s [5] method on the evaluation of sample classroom teaching behavior is about 4.1%, and the
error of Ya-jun et al.’s [6] method on the evaluation of sample classroom teaching behavior is about 5.5%. When the number of iterations is 40, the error of our method on the evaluation of sample classroom teaching behavior is about 1.5%, the error of Zhao et al.’s [5] method on the evaluation of sample classroom teaching behavior is about 3.5%, and the error of Ya-jun et al.’s [6] method on the evaluation of sample classroom teaching behavior is about 5%. When the number of iterations is 60, the error of our method on the evaluation of sample classroom teaching behavior is about 1.2%, the error of Zhao et al.’s [5] method on the evaluation of sample classroom teaching behavior is about 5%, and the error of Ya-jun et al.’s [6] method on the evaluation of sample classroom teaching behavior is about 4.1%. Similarly, when the number of iterations is 100, the error of our method on the evaluation of sample classroom teaching behavior is about 1.8%, the error of Zhao et al.’s [5] method on the evaluation of sample classroom teaching behavior is about 3%, and the error of Ya-jun et al.’s [6] method on the evaluation of sample classroom teaching behavior is about 2.8%. In contrast, the error of the sample classroom teaching behavior evaluation of the proposed method is the lowest, which verifies the effectiveness of this method.

We also experiment with different dropout levels because dropout plays such an important role in the convergence time and performance of any DL-based model [16]. The impact of the dropout is seen in Figure 5. It can be observed that when dropout is not utilized, our model outperforms the identical model. This also confirms that our dropout is required for our proposed model to work optimally.
5. Conclusion

This paper studies classroom teaching behavior recognition and proposes a novel method based on the convolutional neural network (CNN). We first obtain the classroom teaching behavior characteristic data through the SVM linear separable initial, determine the relationship of the characteristic sample data in the hyperplane, obtain the heterogeneous support vector of the online learning behavior characteristic sample data in the SVM hyperplane, and complete the extraction of the characteristic data. With the help of the convolution, pooling, and fully connected layers, we recognize classroom teaching behavior. Then, through an analytic hierarchy process, we construct the confusion matrix to determine the weight of classroom teaching behavior indicators, verify their consistency, and complete the evaluation of classroom teaching behavior by calculating the membership of evaluation factors. The experimental results show that the identification and evaluation method of classroom teaching behavior in this paper can effectively improve classroom teaching behavior identification accuracy.

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 conflicts of interest.

Acknowledgments

This study was supported by the General Project of Beijing Educational Science Planning in 2019, Research on the Transformation and Application of Archaeological Resources Based on U-M Model under the Background of Beijing-Tianjin-Hebei Region (no. CDDB19237).

References


