



Research article

A Vision sensing-based automatic evaluation method for teaching effect based on deep residual network

Meijuan Sun^{1,2,*}

¹ Kansas International School, Sias University, Zhengzhou 451100, Henan, China

² Center of Research on Sino-American Cooperation in Running Schools of Sias University, China

* **Correspondence:** Email: 12015@sias.edu.cn.

Abstract: The automatic evaluation of the teaching effect has been a technical problem for many years. Because only video frames are available for it, and the information extraction from such dynamic scenes still remains challenging. In recent years, the progress of deep learning has boosted the application of computer vision in many areas, which can provide much insight into the above issue. As a consequence, this paper proposes a vision sensing-based automatic evaluation method for teaching effects based on deep residual network (DRN). The DRN is utilized to construct a backbone network for sensing from visual features such as attending status, taking notes, playing phones, looking outside, etc. The extracted visual features are further selected as the basis for the evaluation of the teaching effect. We have also collected some realistic course images to establish a real-world dataset for the performance assessment of the proposal. The proposed method is implemented on collected datasets via computer programming-based simulation experiments, so as to obtain accuracy assessment results as measurement. The obtained results show that the proposal can well perceive typical visual features from video frames of courses and realize automatic evaluation of the teaching effect.

Keywords: residual neural network; intelligent assessment; behavior recognition; social computing

1. Introduction

University English teaching activities are a dynamic process involving many variables and influencing factors, which makes the challenging to establish an automatic evaluation scheme from the side of teachers [1]. Except for the consideration of teachers' classroom teaching [2, 3], it is also necessary to take into account the factors of students who directly participate in teaching [4]. Students participate in the full range of classroom teaching, directly feeling the teacher's teaching quality. Also, students participate in the whole process of classroom teaching and can directly feel the quality of teachers' teaching, as well as their own learning effect [5].

Digital teaching mode is one of the new classroom teaching modes promoted by the Ministry of Education, not only enriches English teaching resources, but also optimizes English teaching methods. Its re-integration of English teaching and learning is even more advantageous [6]. An information-based teaching environment was created, and this changes the traditional teaching environment [7]. It not only enriches English teaching resources and optimizes English teaching methods, but also re-integrates and gives full play to teaching advantages, creates an information-based teaching environment, changes the traditional teaching mode, meets the current requirements for “hardware facilities” in curriculum construction, and promotes the benign development of the university teaching ecosystem [8, 9]. The software construction of “golden course” is about improving the quality of teaching in the classroom. Among them, the key to ensure the quality of the course is to build a scientific and effective course quality evaluation system. At present, most domestic universities’ classroom teaching quality evaluation is coordinated by an academic affairs organization, including the construction of evaluation system and the publication of an evaluation results. Especially in a setting of evaluation indexes, many universities do not distinguish the nature of courses and use uniform evaluation indexes that include both teaching attitudes and contents and teaching methods and effects, which has the advantage of ensuring the evaluation of common problems and the disadvantage of ignoring the characteristics of each course [10, 11]. With the innovation of the education evaluation concept, the drawbacks of this teaching evaluation system gradually appear, lagging behind and having limitations, which can neither improve the quality of course teaching nor restrict the creation of high-quality courses. Therefore, it has become an important task to build an evaluation system which is in line with the new teaching concept and adapts to the new mode of teaching reform in the digital era [12, 13].

Classroom teaching quality assessment is the basic content of education quality evaluation, which should not only assess the effect of the credit hours but take note of the effect of on a time period as well. An effective teaching evaluation model needs the support of effective evaluation indexes in order to get scientific evaluation results. In recent years, academics have actively carried out research on digital classroom teaching evaluation, such as Fang Xucai [14], who proposed four dimensions of planning and preparing classroom teaching, second classroom extension and teaching responsibility and other four dimensions of constructing a foreign language teaching evaluation model; Li Zhihe et al. [15] constructed evaluation indexes of online teachers’ teaching ability, etc. However, these evaluation model studies focus on the exploration of the validity of evaluation indicators. A deepening of the research on evaluation methods is still needed.

Classroom teaching behavior analysis technology is moving toward automation and intelligence with information technology, which makes the original teaching evaluation system and method extremely challenging. The application of new technologies to normalize and scale up classroom teaching evaluation provides strong support for studying the laws of classroom teaching and exploring the essence of learning. In the learning process, emotions influence human cognition and behavior, and grasping the emotional state of learners is particularly important for future research on intelligent and personalized education. Therefore, a number of scholars have introduced emotion recognition into instructional analysis models. Han Li compared cognitive behaviors with students’ head posture and facial expression behavior to build a classroom teaching evaluation system based on face detection and expression analysis [16]. Cao et al. [17] used multimodal data to build a student learning engagement recognition model based on deep learning networks. Zhao Min et al. based on the original quality evaluation indexes, through modeling and obtaining teaching quality evaluation through deep learning network. Sun et al. [18] analyzed and studied students’ emotions based on video processing technology. There are few papers dedicated to the identification of students’ classroom behaviors. Zhou et al. [19] used a face detection, wheel. Contour detection and subject action amplitude detection were obtained from the dataset, and a Bayesian causal net was used as an

inference model to determine the characteristics of subject behavior for classroom teaching behavior recognition. Dang [20], on the other hand, described and judged actions by extracting Zernike moment features, optical flow features, and global motion direction features of actions and combining them with a plain Bayesian classifier. Zhang [12] classified and identified the action vectors by extracting features from human skeletal vectors, and then using an SVM classifier [13]. The above methods mainly use traditional machine learning methods that require a large number of manual steps and have a low accuracy rate. Liao et al. [21] successfully identified three classroom behaviors: sleeping, playing with the phone, and normal, by capturing students' classroom behaviors through a camera and extracting the target regions through background differencing into a VGG network [22]. This study provides a new idea and method for classroom behavior recognition by applying deep learning technology to classroom teaching image recognition, but the number of students recognized is small, and the recognition of a students' actions in the classroom is simple and the accuracy rate is still low.

In recent years, deep convolutional neural networks have been rapidly developing, and models such as AlexNet [22], VGGNet [23], and GoogLeNet [24–26] have been proposed one after another. However, when the number of network layers keeps deepening, the gradient explosion or gradient disappearance problem in deep neural networks during training process becomes more and more obvious. To solve this issue, He et al. proposed ResNet [27]. One of the important features of this network is the inclusion of the residual module, which successfully alleviates the network degradation problem when the network layers are too deep by adding Shortcut structures between the convolutional layers. The ResNet has been extended into many application scenarios due to its resilience and proper processing performance. Qiao et al. explored the utilization of ResNet in the scene of heart disease-related medical image processing in work [28]. On this basis, Qiao et al. also proposed a novel feature learning detection system with the use of deep learning-based vision sensing technology in the work [29]. Besides, a more advanced four-chamber semantic parsing network is also proposed in a similar area by Qiao et al. in the work [30]. To identify more students' behaviors in the classroom with higher accuracy, we applied a deep residual network to classroom behavior recognition. The classroom behavior recognition dataset is constructed from a wide range of images of live student classroom behavior. A deep residual network for this dataset over the characteristics of the residual module, which offers a fresh approach to recognizing students' classroom behaviors.

2. Methodology

The assessment of the teaching effect is implemented by perceiving various behaviors in the classrooms, with the assistance of a deep neural network. The ResNet model is selected here for this purpose. The main solution thought of this paper is to recognize some basic behaviors of classes using the ResNet model. These involved behaviors contain both positive behaviors and negative behaviors. The ResNet model is utilized to perceive these behaviors and extract these features as the final discriminative basis.

Then, the discriminative results of the teaching effect can be calculated.

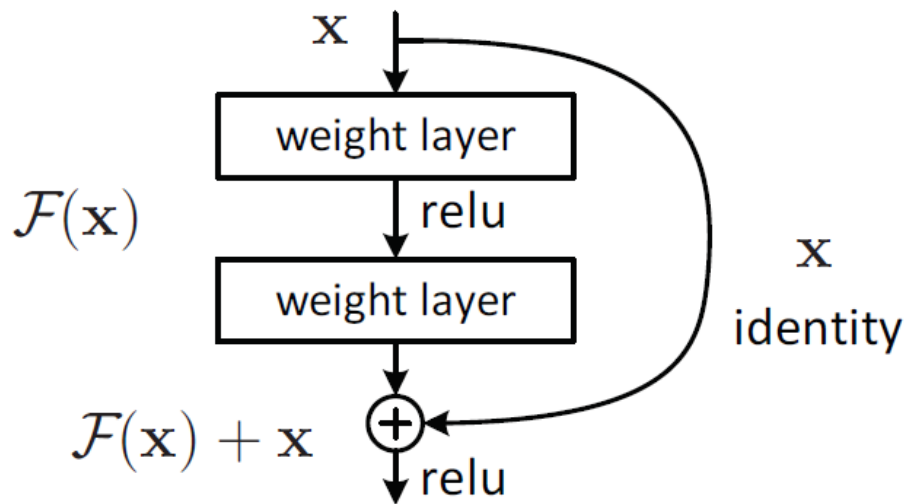


Figure 1. Structure of a residual block.

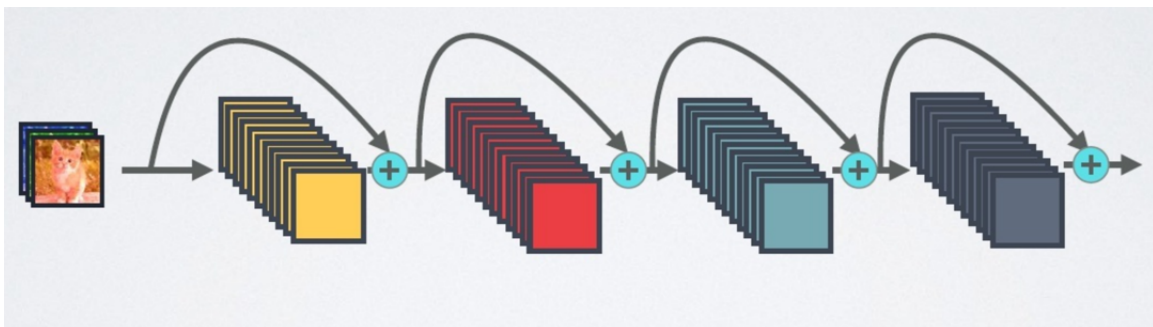


Figure 2. Proposed deep residual network structure.

2.1. Residual network

Residual network belongs to a deep convolutional neural network. As far as convolutional neural networks are concerned, the fitting ability can be enhanced by increasing as many layers of the network as possible. But, as the number of layers deepens, the training of convolutional neural networks turns out to be extremely difficult. Once more than a certain level of layers are added, there is a decrease in the recognition ability of the network [31–33]. During gradient back propagation, parameters of the network near the output layer converge fast, whereas those near the input layer converge slowly due to the deeper layers of the network.

To overcome the problem of decreasing recognition accuracy caused due to too many layers, a residual unit is introduced by the residual network. That is, a Shortcut structure is added between the convolutional layers, which leads to a residual after the objective function of the network training is made to differ from the input function, as shown in Figure 1. In traditional learning goal, the objective is to learn a mapping function, as follows:

$$y = f(x) \quad (2.1)$$

where x denotes input of learning function, and y denotes learning result. Differently, in residual

network, the expression of learning goals is:

$$y = f(x) + x \quad (2.2)$$

Its goal is to learn the residual between y and x , instead of single y . In other words, the actual output is the sum of the original output and the original input, thus transforming the fit of the network to $f(x)$ into the fit of $h(x)$. The residual item is denoted as follows:

$$f(x) = y - x \quad (2.3)$$

This structure does not add new parameters and extra computational effort, and also solves the problem of gradient dispersion in the back-propagation process of the network. The Figure 1 gives structure of a residual block, and a number of residual blocks are able to constitute the different residual neural networks.

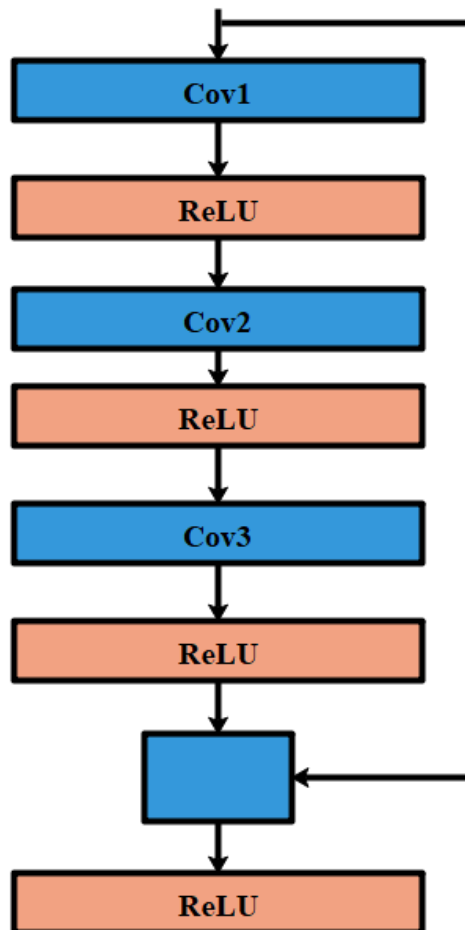


Figure 3. Proposed constant module structure.

2.2. Deep residual network

The structure of the deep residual network used to identify students' classroom behavior is shown in Figure 2. The network is made up of one convolutional layer, two ReLU layers, three pooling layers, one convolutional module, two constant modules, two fully connected layers, and finally a

classification layer, in which the convolutional layers are filled with same. The input image first passes through the convolutional layer, which contains 64 convolutional kernels of size 2×2 with an operation step of 2, activated by the ReLU activation lattice for initial feature extraction. It is followed by a convolutional module, two constant modules for deep feature extraction, and two fully connected layers of different sizes for feature dimension reduction. The output neurons of the latter fully connected layer are 6. This corresponds to the latter fully connected layer has six neurons, which corresponds to the 6 behaviors of students in the classroom, and finally, the classification results are output through the classification layer. For convolution layers and pooling layers, the scale transformation rule is represented as the following the formula:

$$\Delta_{transform} = \frac{\Delta_{initial} + 2\Delta_{pad} - \Delta_{filter}}{\Delta_{stride}} \quad (2.4)$$

where $\Delta_{transform}$ denotes scale of feature map after convolution or pooling operation, $\Delta_{initial}$ denotes scale of feature map before convolution or pooling operation, Δ_{pad} denotes scale of padding operation, Δ_{filter} denotes scale of convolution filter or pooling filter, Δ_{stride} denotes scale of the stride operation.

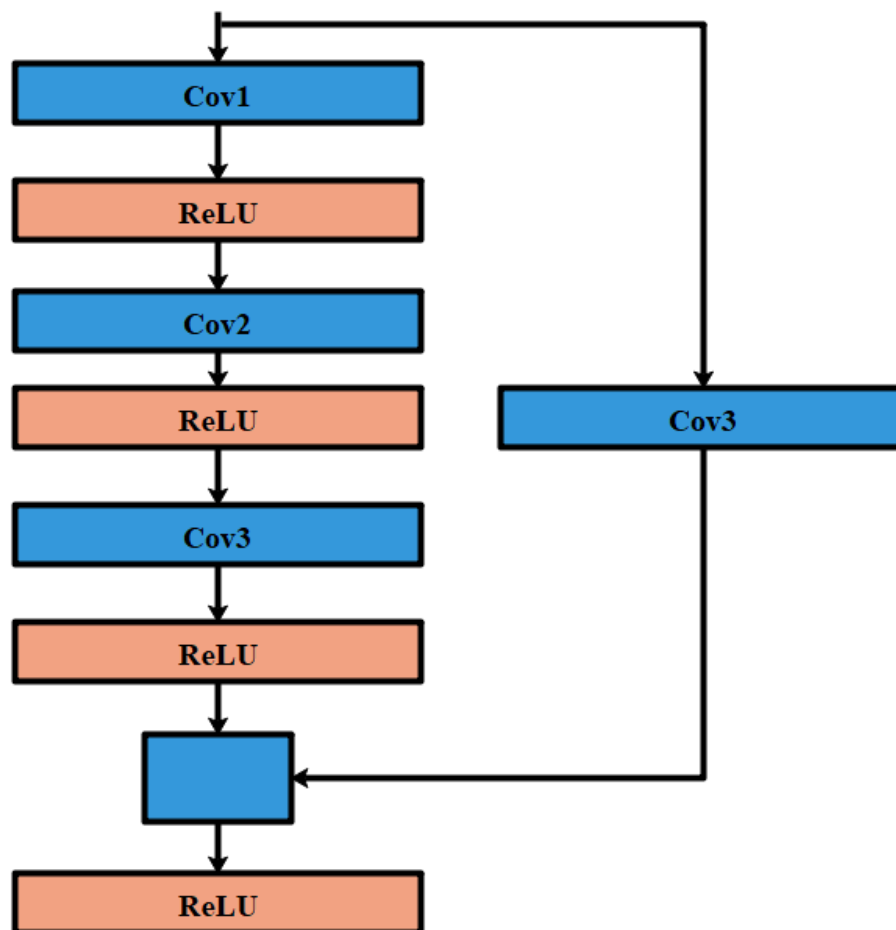


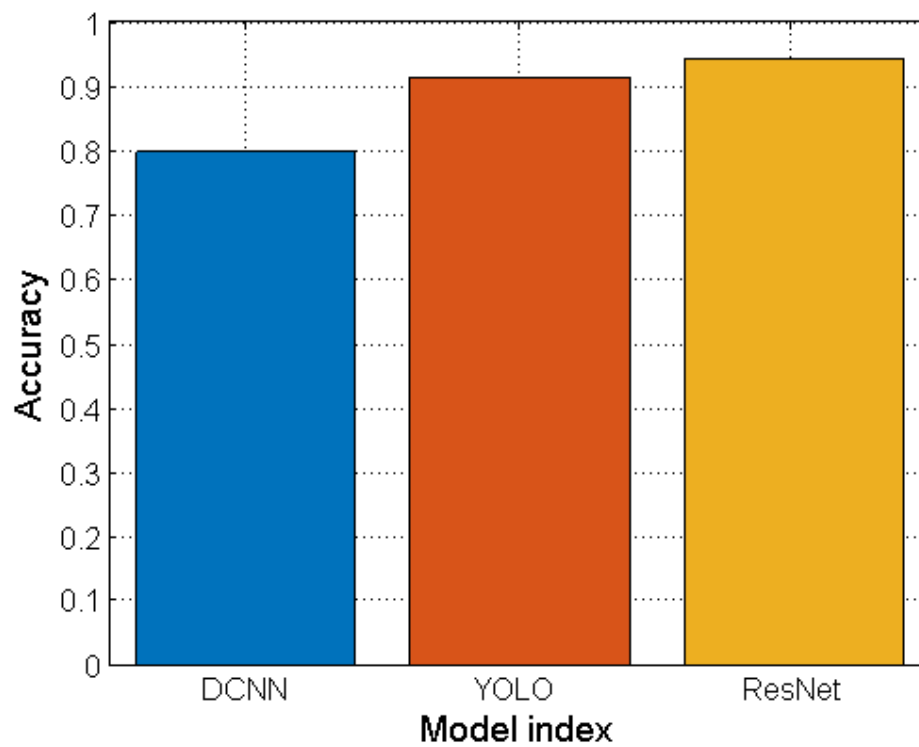
Figure 4. Convolutional module structure.



(a) Recognition towards individuals



(b) Recognition towards groups

Figure 5. Two demos of recognition results from ResNet method.**Figure 6.** The performance of classroom behavior recognition.

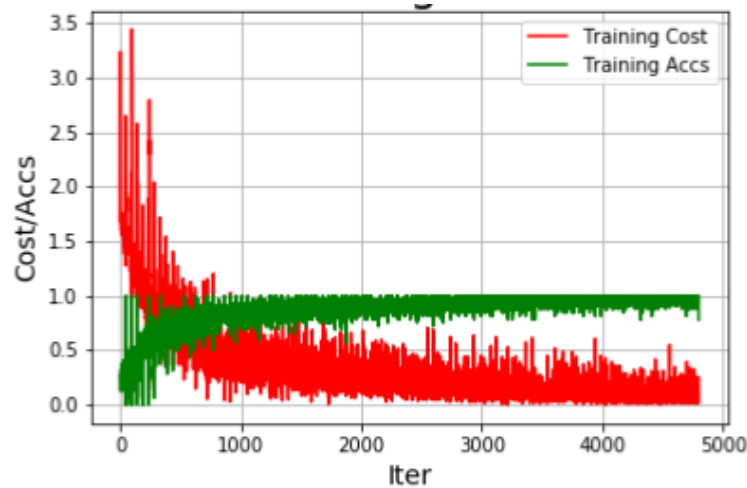


Figure 7. The cost/acc of the proposed ResNet.

The structure of the constant module in Figure 4. The network structure of the convolutional module is shown in Figure 3. The constant module combines three convolutional layers, three ReLU layers and a shortcut connection operation. The convolutional layer 1 contains 64 channels of 1×1 convolutional kernels with operation step 1, the convolutional layer 2 contains 64 channels of 3×3 convolutional kernels with operation step 1, and the convolutional layer 3 contains 256 channels of 1×1 convolutional kernels with operation step 1. The Shortcut join operation is used to add the input of the constant module with the output after three convolutional operations, which reflects the basic idea of residual network. This operation reflects the basic idea of the residual network. The following formula can be deduced to represent the transformation process of convolution operations:

$$K = \mu_1 (B_C + W_C \star C) \quad (2.5)$$

where W_C denotes the weighted parameter for this transformation, B_C denotes the bias parameter for this transformation, \star denotes convolution operation symbol, and $\mu_1(\cdot)$ is another activation function named as Sigmoid, whose representation is as follows:

$$\mu_1(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.6)$$

The convolutional module consists of four convolutional layers, three ReLU layers and a Shortcut connection operation. The convolutional layer 1 contains 64 channels of kernels with a size of 1×1 ; the convolutional layer 2 contains 64 channels of kernels with a size of 3×3 ; the convolutional layer 3 contains 256 channels of kernels with a size of 1×1 ; and the convolutional layer 4 contains 256 channels of kernels with a size of 1×1 . Compared with the constant module, it performs a convolutional operation on the network input x before the Shortcut connection operation. After some residual blocks, the initial input is transformed into representations with the abstract format. It will be transferred into the prediction result via a full connection operation structure. This process takes the form of the following equation:

$$D = \mu_2 (W_C \cdot C + B_C) \quad (2.7)$$

where W_C denotes weight parameter in this operation, B_C denotes bias parameter in this operation, and $\mu_2(\cdot)$ denotes sigmoid activation function, which is expressed as:

$$\mu_2(x) = \frac{1}{1 + e^{-x}} \quad (2.8)$$

It can be seen from the above formula that it tries to confine the output into the range of (0, 1).

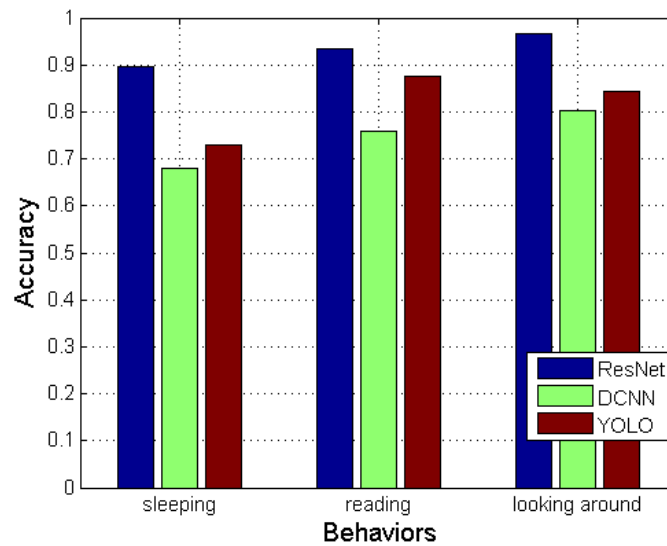


Figure 8. Recognition accuracy with different method.

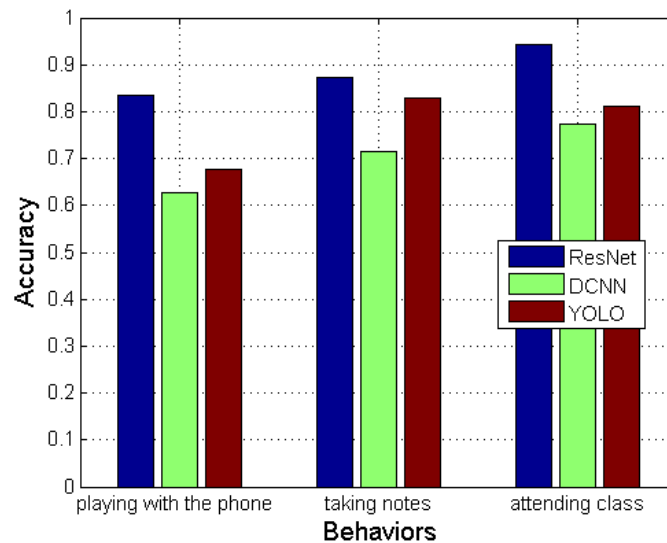


Figure 9. Recognition accuracy with different method.

2.3. Implementation of deep residual network

Convolutional neural networks have powerful fitting capabilities and are capable of learning complex mapping relationships from input to output. Even if the exact mathematical expressions from input to output are not known, the convolutional neural network can establish the mapping relationship between them more accurately by learning specific patterns from input to output. The training of convolutional neural networks is generally done by supervised training. The training process is divided into two main phases, namely the forward propagation phase and the backward

propagation phase [34, 35]. In the forward propagation phase, to improve the accuracy of the model and enable the network to converge rapidly. In this paper, the training set is first randomly disrupted, and then a fixed number of small batches of images are selected as the network input in each-iteration with the machine situation. The input is propagated forward through the constructed network architecture layer by layer, and finally, the probability of recognition of each behavior is output through the softmax classification layer.

In the back-propagation stage, the error value is first calculated using cross-entropy as the loss function, and then the error is back-propagated by the Adam optimizer to update the network weights and gradually make the loss function close to the optimal value in order to optimize the whole network. In addition, this paper adopts drop-out coding method when coding label categories, and the network the learning rate is 0.001, and uses Drop-out technique back in the fully connected layer, i.e., the neurons are deactivated randomly at each training, so as to alleviate the network overfitting and achieve the regularization effect. And the loss function is shown as follows:

$$L = \frac{1}{N} \sum_i -[y_i \cdot \log(p_i) + (1 + y_i) \cdot \log(1 - p_i)] \quad (2.9)$$

Then, the stochastic gradient descent can be utilized to solve the above objective function. Let Θ denotes the set of all the parameters to be learned in the proposed model. The learning process of Θ with use of stochastic gradient descent can be represented as the following formula:

$$\Theta^{(l+1)} \leftarrow \Theta^{(l)} - r \cdot \frac{\partial L}{\partial \Theta} \quad (2.10)$$

where l denotes the index number of iterative rounds, and r denotes the learning rate.

3. Experiment and evaluation

3.1. Experimental setup

Because classroom behavior data is not publicly available on the web, we collected the data ourselves to create a classroom behavior recognition dataset. The video is collected by a camera installed in the classroom with a resolution of 2560×1536 . The video collection includes six actions that students frequently perform in the classroom, such as taking notes, looking around, and reading books. Then the videos are collected, videos are first sampled with uniform frames and converted into images. The images were then cropped into images containing individual students and reshaped to a resolution of 128×128 , and the classroom behaviors of students in each image were labeled to obtain a total of 1020 images with labels. The original dataset was expanded by mirror-symmetric data augmentation to obtain a classroom behavior recognition dataset containing 2040 images. The number of images for each behavior is the same. The data was randomly selected for training and testing. The training set and test set had 1560 and 480 images respectively.

Table 1. Recognition accuracy of different behaviors.

Behaviors	sleeping	reading	playing phones	taking notes	attending class	looking around
Accuracy	96%	92%	96%	94%	91%	86%

Based on the data from both sides of teaching obtained by video analysis method, this paper initially designs an index system for online teaching quality evaluation, and the techniques used for each index are different. The most difficult one is the emotional changes of both teaching parties, where the

emotional identification includes the emotional interaction between teaching and learning process in addition to the original single emotional identification [36–38]. This interactive emotional analysis can better reflect students' engagement in learning and teachers' adjustment of teaching according to students' emotional changes.

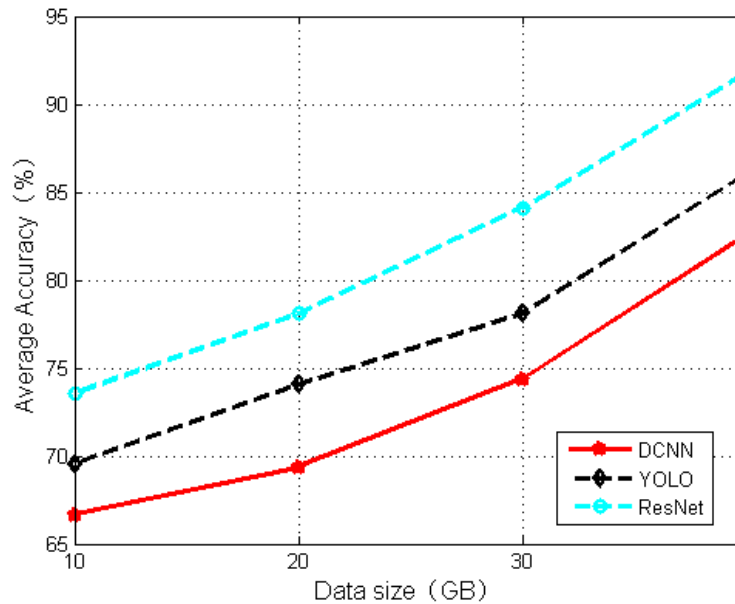


Figure 10. The effect of data set size with different method.

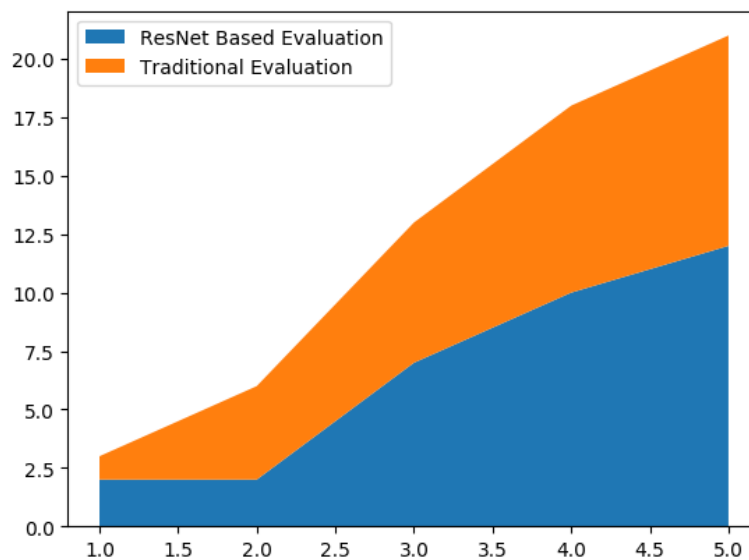


Figure 11. Assessment of satisfaction with different method.

To verify the effectiveness of the index system of the teaching quality assessment model constructed in this study, this study organized a total of 347 undergraduate students in the 2019 class of a university to evaluate 12 English teachers of the university by means of a questionnaire.

3.2. Evaluation and analysis

The ResNet is first implemented on the experimental image dataset [39]. It serves as the backbone network for visual feature extraction. Two demos for recognition results from the ResNet method are demonstrated in Figure 5. It has two subfigures, in which (a) corresponds to the recognition under scenes of individual objects, and (b) corresponds to the recognition under scenes of groups of objects. It can be observed from Figure 5 that the ResNet can well deal with scenes of both individual objects and groups of objects, and that different behaviors in classrooms can be well recognized.

To better reveal the performance superiority of the ResNet method, his paper compares the recognition performance of ResNet with DCNN and YOLO models on classroom behavior data sets. The average recognition accuracy of the three experimental methods is illustrated in Figure 6. It can be seen from this image that the utilization of ResNet as a backbone can achieve better results than the other two. And the final accuracy of the ResNet and DCNN is shown in Figure 7. From this figure, we can see that the generalization accuracy of DCNN is 89.46%, while that of ResNet is 91.91%. That is, the generalization accuracy of ResNet is higher than CNN, indicating that the addition of residual structure can enhance network performance. One of the comparison plots for the iterative update of the model recognition accuracy is shown in Figures 8 and 9. The former focuses on behaviors of sleeping, reading, and looking around, while the latter focuses on behaviors of playing with the phone, taking notes, and attending classes. Besides, the relationship between data amount size and average accuracy is shown in Figure 10. The X-axis denotes the data size amount, and the Y-axis denotes accuracy values. The graph shows that as the amount of data size increases, the average accuracy shows an upward trend. This may be attributed as the fact that more video amount is able to help train better recognition models.

The accuracy of each behavior recognized by the deep residual network is shown in Table 1, in which sleeping and reading achieved higher recognition accuracy of 97.06% and 94.12%, respectively, while the recognition accuracy of playing with the phone, taking notes, attending class, and looking around were 92.65%, 89.71%, 91.18%, and 86.76%, respectively. The recognition accuracy of looking around and class is relatively low, which is probably due to the fact that students sit in various positions in the classroom and their heads have certain biases during class, resulting in a certain similarity between the behavior of students in class and looking around in the dataset, which leads to network misidentification and reduces their recognition accuracy.

In addition, we compared the results of this paper's proposed method with those of traditional methods for assessing satisfaction with teaching quality, and a comparison graph of satisfaction is given in Figure 11. From this figure, it shows that the satisfaction of teaching quality assessment using the method proposed in this paper is more than 32% higher than that of the traditional method due to the full consideration of students' classroom interaction performance.

4. Conclusions

It takes complex operations to extract features from classroom behavioral images with traditional machine learning methods. And it has a low classification accuracy. The image features can be automatically extracted by CNN compared to traditional methods. Its end-to-end approach to training the network from input to output allows the network to recognize classroom behavior, which improves the accuracy and reduces computational complexity at the same time. However, network training becomes more complicated as the level of the network increases. It even brings the problem of network performance degradation. In this paper, we propose a deep residual network model for classroom behavior recognition by introducing the residual structure into the CNN. The results of the experiments demonstrate that this network has a better performance compared with the deep

convolutional neural network.

As has been noted by Amelio et al. [40], although the ResNet has many advantages in some specific image processing tasks, it still suffers from some rough issues such as heavy network structure. This requires high computational performance and proper hardware conditions. In addition, it is a typical peer-to-peer model that acts as a black box during processing tasks. This limits its explainability when dealing with various tasks. The two major points are our future direction after this work. It is expected to explore the better applications of the ResNet methods to make it have a lighter network structure and better explainability.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by First-class Course of Zhengzhou Sias University.

References

1. Z. Guo, K. Yu, Z. Lv, K. K. R. Choo, P. Shi, J. Rodrigues, Deep federated learning enhanced secure poi microservices for cyber-physical systems, *IEEE Wireless Commun.*, **29** (2022), 22–29. <https://doi.org/10.1109/MWC.002.2100272>
2. S. Xia, Z. Yao, G. Wu, Y. Li, Distributed offloading for cooperative intelligent transportation under heterogeneous networks, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 16701–16714. <https://doi.org/10.1109/TITS.2022.3190280>
3. Z. Guo, K. Yu, A. Jolfaei, F. Ding, N. Zhang, Fuz-spam: label smoothing-based fuzzy detection of spammers in internet of things, *IEEE Trans. Fuzzy Syst.*, **30** (2022), 4543–4554. <https://doi.org/10.1109/TFUZZ.2021.3130311>
4. L. Zhao, Z. Yin, K. Yu, X. Tang, L. Xu, Z. Guo, et al., A fuzzy logic based intelligent multi-attribute routing scheme for two-layered sdvns, *IEEE Trans. Network Serv. Manage.*, **2022** (2022). <https://doi.org/10.1109/TNSM.2022.3202741>
5. Z. Zhou, X. Dong, Z. Li, K. Yu, C. Ding, Y. Yang, Spatio-temporal feature encoding for traffic accident detection in vanet environment, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 19772–19781. <https://doi.org/10.1109/TITS.2022.3147826>
6. S. Zhang, H. Gu, K. Chi, L. Huang, K. Yu, S. Mumtaz, Drl-based partial offloading for maximizing sum computation rate of wireless powered mobile edge computing network, *IEEE Trans. Wireless Commun.*, **21** (2022), 10934–10948. <https://doi.org/10.1109/TWC.2022.3188302>
7. D. Peng, D. He, Y. Li, Z. Wang, Integrating terrestrial and satellite multibeam systems toward 6G: techniques and challenges for interference mitigation, *IEEE Wireless Commun.*, **29** (2022), 24–31. <https://doi.org/10.1109/MWC.002.00293>
8. A. Büyükkarci, M. Müldür, digital storytelling for primary school mathematics teaching: product and process evaluation, *Educ. Inf. Technol.*, **27** (2022), 5365–5396. <https://doi.org/10.1007/s10639-021-10813-8>
9. Z. Guo, K. Yu, A. K. Bashir, D. Zhang, Y. D. Al-Otaibi, M. Guizani, Deep information fusion-driven POI scheduling for mobile social networks, *IEEE Network*, **36** (2022), 210–216. <https://doi.org/10.1109/MNET.102.2100394>

10. A. Cahyadi, Hendryadi, S. Widyastuti, Suryani, Covid-19, emergency remote teaching evaluation: the case of Indonesia, *Educ. Inf. Technol.*, **27** (2022), 2165–2179. <https://doi.org/10.1007/s10639-021-10680-3>
11. Y. Lu, L. Yang, S. X. Yang, Q. Hua, A. K. Sangaiyah, T. Guo, et al., An intelligent deterministic scheduling method for ultra-low latency communication in edge enabled industrial internet of things, *IEEE Trans. Ind. Inf.*, **19** (2023), 1756–1767. <https://doi.org/10.1109/TII.2022.3186891>
12. B. Huang, K. Wang, An improved BP neural network-based quality evaluation model for Chinese international education teaching courses, in *ICCD 2022: The 6th International Conference on Compute and Data Analysis*, (2022), 122–127. <https://doi.org/10.1145/3523089.3523109>
13. Q. Zhang, K. Yu, Z. Guo, S. Garg, J. Rodrigues, M. M. Hassan, et al., Graph neural network-driven traffic forecasting for the connected internet of vehicles, *IEEE Trans. Network Sci. Eng.*, **9** (2022), 3015–3027. <https://doi.org/10.1109/TNSE.2021.3126830>
14. C. Hou, J. Ai, Y. Lin, C. Guan, J. Li, W. Zhu, Evaluation of online teaching quality based on facial expression recognition, *Future Internet*, **14** (2022). <https://doi.org/10.3390/fi14060177>
15. S. Qi, L. Liu, B. S. Kumar, A. Prathik, An english teaching quality evaluation model based on gaussian process machine learning, *Expert Syst. J. Knowl. Eng.*, **39** (2022). <https://doi.org/10.1111/exsy.12861>
16. H. Shu, English teaching effect evaluation based on data association mining, in *CIPAE 2021: 2nd International Conference on Computers, Information Processing and Advanced Education*, (2021), 1223–1226. <https://doi.org/10.1145/3456887.3457494>
17. P. Gao, VIKOR method for intuitionistic fuzzy multi-attribute group decision-making and its application to teaching quality evaluation of college english, *J. Intell. Fuzzy Syst.*, **42** (2022), 5189–5197. <https://doi.org/10.3233/JIFS-211749>
18. D. Wei, Y. Rong, H. Garg, J. Liu, An extended WASPAS approach for teaching quality evaluation based on pythagorean fuzzy reducible weighted maclaurin symmetric mean, *J. Intell. Fuzzy Syst.*, **42** (2022), 3121–3152. <https://doi.org/10.3233/JIFS-210821>
19. M. Li, Multidimensional analysis and evaluation of college english teaching quality based on an artificial intelligence model, *J. Sens.*, **2022** (2022), 1–13. <https://doi.org/10.1155/2022/1314736>
20. S. Zeng, Y. Pan, H. Jin, Online teaching quality evaluation of business statistics course utilizing fermatean fuzzy analytical hierarchy process with aggregation operator, *Systems*, **10** (2022). <https://doi.org/10.3390/systems10030063>
21. B. Feng, Dynamic analysis of college physical education teaching quality evaluation based on network under the big data, *Comput. Intell. Neurosci.*, **2021** (2021). <https://doi.org/10.1155/2021/5949167>
22. X. Xu, F. Liu, Optimization of online education and teaching evaluation system based on GA-BP neural network, *Comput. Intell. Neurosci.*, **2021** (2021). <https://doi.org/10.1155/2021/8785127>
23. J. Heo, S. Han, The mediating effect of literacy of LMS between self-evaluation online teaching effectiveness and self-directed learning readiness, *Educ. Inf. Technol.*, **26** (2021), 6097–6108. <https://doi.org/10.1007/s10639-021-10590-4>
24. R. Tárraga-Mínguez, C. S. Guerrero, P. Sanz-Cervera, Digital teaching competence evaluation of pre-service teachers in spain: a review study, *IEEE Rev. Iberoam. Technol. Aprendizaje*, **16** (2021), 70–76. <https://doi.org/10.1109/RITA.2021.3052848>

25. Y. V. Tsekhmister, T. Konovalova, B. Y. Tsekhmister, A. Agrawal, D. Ghosh, Evaluation of virtual reality technology and online teaching system for medical students in ukraine during COVID-19 pandemic, *Int. J. Emerging Technol. Learn.*, **16** (2021). <https://doi.org/10.3991/ijet.v16i23.26099>
26. Y. Wang, S. Li, B. Zhao, J. Zhang, Y. Yang, B. Li, A resnet-based approach for accurate radiographic diagnosis of knee osteoarthritis, *CAAI Trans. Intell. Technol.*, **7** (2022), 512–521. <https://doi.org/10.1049/cit2.12079>
27. Y. Wang, C. Sun, Y. Guo, A multi-attribute fuzzy evaluation model for the teaching quality of physical education in colleges and its implementation strategies, *Int. J. Emerging Technol. Learn.*, **16** (2021). <https://doi.org/10.3991/ijet.v16i02.19725>
28. S. Qiao, S. Pang, G. Luo, S. Pan, Z. Yu, T. Chen, et al., RLDS: an explainable residual learning diagnosis system for fetal congenital heart disease, *Future Gener. Comput. Syst.*, **128** (2022), 205–218. <https://doi.org/10.1016/j.future.2021.10.001>
29. S. Qiao, S. Pang, G. Luo, S. Pan, T. Chen, Z. Lv, FLDS: an intelligent feature learning detection system for visualizing medical images supporting fetal four-chamber views, *IEEE J. Biomed. Health Inf.*, **26** (2022), 4814–4825. <https://doi.org/10.1109/JBHI.2021.3091579>
30. S. Qiao, S. Pang, Y. Sun, G. Luo, W. Yin, Y. Zhao, et al., Sprechd: four-chamber semantic parsing network for recognizing fetal congenital heart disease in medical Metaverse, *IEEE J. Biomed. Health. Inf.*, (2022), 1–11. <https://doi.org/10.1109/JBHI.2022.3218577>
31. Y. Zhang, The development of an evaluation model to assess the effect of online english teaching based on fuzzy mathematics, *Int. J. Emerging Technol. Learn.*, **16** (2021). <https://doi.org/10.3991/ijet.v16i12.23325>
32. Y. Han, Evaluation of english online teaching based on remote supervision algorithms and deep learning, *J. Intell. Fuzzy Syst.*, **40** (2021), 7097–7108. <https://doi.org/10.3233/JIFS-189539>
33. H. Liang, Role of artificial intelligence algorithm for taekwondo teaching effect evaluation model, *J. Intell. Fuzzy Syst.*, **40** (2021), 3239–3250. <https://doi.org/10.3233/JIFS-189364>
34. Y. Liu, Evaluation algorithm of teaching work quality in colleges and universities based on deep denoising autoencoder network, *Mobile Inf. Syst.*, **2021** (2021). <https://doi.org/10.1155/2021/8161985>
35. G. Li, F. Liu, Y. Wang, Y. Guo, L. Xiao, L. Zhu, A convolutional neural network (CNN) based approach for the recognition and evaluation of classroom teaching behavior, *Sci. Program.*, **2021** (2021). <https://doi.org/10.1155/2021/6336773>
36. P. Liu, X. Wang, F. Teng, Online teaching quality evaluation based on multi-granularity probabilistic linguistic term sets, *J. Intell. Fuzzy Syst.*, **40** (2021), 9915–9935. <https://doi.org/10.3233/JIFS-202543>
37. Q. Wang, Research on teaching quality evaluation of college english based on the CODAS method under interval-valued intuitionistic fuzzy information, *J. Intell. Fuzzy Syst.*, **41** (2021), 1499–1508. <https://doi.org/10.3233/JIFS-210366>
38. H. Yu, Online teaching quality evaluation based on emotion recognition and improved aprioritid algorithm, *J. Intell. Fuzzy Syst.*, **40** (2021), 7037–7047. <https://doi.org/10.3233/JIFS-189534>
39. Y. Yu, English teaching ability evaluation algorithm based on big data fuzzy k-means clustering, *Advances in Intelligent Systems and Computing*, Springer, (2021), 557–564. https://doi.org/10.1007/978-3-030-69999-4_77

-
40. A. Amelio, G. Bonifazi, F. Cauteruccio, E. Corradini, M. Marchetti, D. Ursino, et al., Representation and compression of residual neural networks through a multilayer network based approach, *Expert Syst. Appl.*, **215** (2023). <https://doi.org/10.1016/j.eswa.2022.119391>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)