

RESEARCH

Open Access



CIES: Cloud-based Intelligent Evaluation Service for video homework using CNN-LSTM network

Rui Song¹, Zhiyi Xiao², Jinjiao Lin^{1,3} and Ming Liu^{1,4*}

Abstract

Video (used as a form of examination or homework) as an efficient approach for examining students' abilities is drawing increasing attention in the education field. How to assess video assignments effectively and accurately has become a significant topic in academia. This work proposes a method based on a multi-channel CNN-LSTM hybrid architecture to extract and classify image features such as students' actions and expressions, as well as audio features such as speech rates and pauses in the video assignments, and then conducts a two-category assessment of "qualified" or "unqualified". Additionally, build this system in a cloud computing environment as a Cloud-based Intelligent Evaluation Service application could provide universal service to meet the needs of multiple teaching units. The proposed method is shown to be feasible and effective through experiments.

Keywords: Video assignments, cloud computing, convolution neural networks, long short-term memory, intelligent evaluation

Introduction

Assigning students to hand in videos as a part of homework has become a new trend in the context of the continuous development of education in recent years. Normally completing assignments in form of video is based on subjective questions. Compared with examination through texts or audios, this method can more accurately evaluate through expressions, movements, and intonations, which can intuitively and accurately reflect how students understand and apply knowledge flexibly. In some subjects, such as nursing and clinical medicine, video assessment has unparalleled advantages. In some pedagogical experiments, students trained by giving video assignments have been shown to have higher average scores than those trained by traditional methods [1, 2].

As the demands of student's ability oriented investigation increased, many teachers prefer to assign homework in the form of video. However, against the background

of intelligent tutoring, video homework quantities will be considerably huge if the homework is assigned in the form of video because of the large numbers of students who attend online classes. Facing with such a large number of video homework assessments, teachers find it challenging to mark those assignments efficiently and give feedback to the students in a timely manner.

Teachers normally mark video assignments manually. Whiling marking the video assignments, they have to spend lots of time on watching the entire video. Sometimes the marking is divided into several days, whereas the standard might change with time, which might result in inefficient and non-uniform standards. Therefore, an intelligent video assignment marking method is needed in the age of intelligent tutoring.

Previous studies in automatic scoring have mainly been applied to examination or homework in paper form or vocal recording. Most of these studies were designed to examine whether the answer is correct. However, it is more important to examine through video homework whether the students have mastered the knowledge and whether the required ability has been achieved. Therefore, we need to extract features of expression, movement, and voice for indirect evaluation.

* Correspondence: 20088031@sdufe.edu.cn

¹School of Control Science and Engineering, Shandong University, Jinan 250002, China

⁴College of Electrical Engineering and Automation, Shandong University of Science and Technology, Jinan 250031, China

Full list of author information is available at the end of the article

Considering the above analysis, we propose a method based on a multi-channel hybrid network combining CNN and LSTM to automatically assess video homework presented with PPT. The method evaluates the performance of students through the image and audio data in the video, considering mainly expressions, actions and tones. We determine whether the student is familiar with the content he or she delivers, or whether the student is repeating mechanically through the tones and expressions of the students (for example, whether the gaze is always focused on the PPT screen or whether the expression is nervous). Additionally, we refer to studies about PowerPoint presentations [3, 4]. Ultimately, video assignments are graded as “qualified” or “unqualified”.

The proposed method is uploaded to the cloud, which can be used by multiple teaching units. These units are separated by permission to realize secure access control of data.

This study offers three contributions as follows:

1. An intelligent approach is proposed to assess video assignments, which is a new topic. The proposed approach can improve the efficiency of correction and contribute to promoting video assignments. Experiments on practical data samples have demonstrated that this method is feasible and potentially valuable.
2. The proposed automated assessment network can simultaneously analyse images and audio in the video and then combine the information for evaluation, which has rarely been mentioned in previous automated grading studies in education. Previous studies mainly use text, audio, images or text-audio and text-images as input. Few studies consider integrating images and audio. In the experiment, this network was shown to be more accurate than using only images or audio as input for analysing video homework.
3. This study presents CIES, a cloud computing solution to put the model into use. The processing of a large amounts of videos is difficult for personal computers. Hence, we adopt a cloud-based system. In this way, the proposed model becomes practical. Meanwhile, the promotion and application of this model are more convenient and cheaper.

The related work on automated grading are reviewed in [Related work](#) section, and the details of the proposed approach are introduced in [Methodology](#) section. In [Experiment](#) section, the experimental results are shown. Finally, we draw conclusions and discuss future research directions in [Conclusion](#) section.

Related work

As the rapid development of computer science, applications of intelligent evaluation techniques have greatly developed recently. In the education field, these applications can be categorised as text-based, audio-based, and image-based applications. Regarding text-based applications, the representative applications correct subjective questions. Various models have been proposed for different applications. The study of English writing questions was the earliest application. Project Essay Grade (PEG) was proposed in the 1960s. Another automated scoring system applied to English writing called E-rater was also put into practice on the GMAT examination in 1999 [5]. Recently, research on subjective questions in Chinese has also made progress. Liu et al. described a method of sentence similarity measurement based on simple word matching [6]. Then, other methods based on corpus-based similarity calculation have been developed. Additionally, methods aimed at other types of subjective questions, such as automatic scoring for computer programming subjective questions [7], have also been proposed. In recent years, subjective evaluation methods using machine learning techniques such as [8, 9] have also emerged.

Regarding the automatic evaluation of audio, the main applications are spoken English assessment and Mandarin assessment. Considering the spoken English assessment, many approaches have been proposed for assessing various indexes. Liang et al. proposed an algorithm to evaluate pronunciation accuracy and fluency [10]. Huang et al. proposed an approach to inspect fluency and rhythm [11]. SpeechRater, a system that can comprehensively evaluate fluency, pronunciation, rhythm, vocabulary diversity, grammatical accuracy and complexity, was introduced in the literature [12]. SpeechRater has recently been applied to automatically assess the spoken section of the TOFEL exam. For the Putonghua assessment, Zhang et al. proposed a method to test the tone error [13]. A comprehensive method to inspect phonemes, vowels, tones and rhotic accents has also been proposed [14]. China’s Putonghua proficiency test has adopted automatic scoring technology recently.

For intelligent scoring based on images, automatic scoring of pencil-filled answer sheets has been promoted. Regarding the scoring of answer sheets that are not pencil-filled but handwritten, Deng et al. proposed an approach to automatically recognize handwritten characters [15]. Xu-Yao Zhang et al. proposed a more accurate Chinese character handwriting recognition model by combining conventional methods and a deep convolutional neural network [16]. The “All Discipline Machine Scoring Technology” system launched by iFLY-TEK CO. LTD. can automatically give relatively accurate scores from scanned test paper images.

Regarding the automatic evaluation based on image and audio, Luo et al. proposed a classroom teaching evaluation system that can evaluate classroom Teaching & Learning conditions [17]. This research was novel, because it integrates image and audio to made evaluation. However, images and audios were used independently for different evaluations: images for students' learning conditions, audios for teachers' teaching conditions. Hence, this method can't be applied on video evaluation. As for other methods to evaluation, Wei et al. proposed an approach to recognize students' actions, then evaluate their learning conditions [18]. These studies can be referenced, but a proper approach for evaluating video assignments is still necessary.

In conclusion, automatic evaluation systems for certain aspects of text, audio or images have been developed. Some systems combine texts and audios (such as SpeechRater) or texts and images (such as the system launched by iFLYTEK). A few researches combine audio and images to make automatic evaluations (such as [17]), but these methods can't be directly used on the evaluation for video. Hence, an approach that simultaneously analyses images and audio to assess video is required.

We also noticed that many novel algorithms on video processing has been proposed recently. On the domain of video segmentation, a new model based on Generative Adversarial Network has been proved to be more accuracy [19]. Ke et al. propose Frame Segmentation Network, improving mAP (mean average precision) and IoU (Intersection over Union) simultaneously [20]. As for action recognition, the S-TPNet proposed in [21] has a good performance. Its accuracy is around 74% on dataset HMDB51, and around 95% on dataset UCF101. However, although existing models could process video efficiently, they are not appropriate for automatic evaluation of video homework. Therefore, proper and efficient algorithms are required.

Methodology

CIES is composed of CIES platform and CNN-LSTM network. Homework are firstly uploaded to cloud. Then they are processed and distributed to CNN-LSTM network by CIES platform. Finally, CNN-LSTM network makes automatic evaluation.

CIES platform

CIES platform is built to overcome the problem that video processing requires high-performance hardware and make the proposed model practical. CIES is a multi-tenant management system with independent services and unified control. The system includes user authentication, authorization, quota management, and resource control. User authentication is realized by KERBEROS

protocol and LDAP, which ensure only identified users can access the system. Authorization ensures only users who have been granted permission can access to assess system. Quota management and resources control limit the frequency of accessing. These 3 parts ensure the security and availability of the system jointly. The structure of CIES platform is shown in Fig. 1.

CIES reduces the high cost of hardware procurement, operation and maintenance of users because the cloud servers are built by cloud operators. Furthermore, cloud computing architecture can flexibly utilize resources such as cloud storage and parallel computing. The advanced architecture improves the operational efficiency of proposed approach.

With the parallel computing framework of the CIES, video is divided into image and audio parts, the two parts are separately partitioned into data blocks, and then pre-processed. Each data block pre-processing corresponds to one computing task, and is automatically performed on the cluster nodes. Tasks are assigned to nodes and executed, then calculation results are collected. The collected pre-processed results are fed into CNN network. Complex details such as data distribution storage, data communication, and fault-tolerant processing are handled by CIES, improving processing speed and robustness.

Overview of CNN-LSTM network

Long short-term memory (LSTM) is a special kind of recurrent neural network (RNN) that can be used to improve the vanishing gradient problems of RNN. For the samples correlated with times, the LSTM network is used to obtain the sequence characteristics of the samples while CNN extracts features to improve the system's accuracy [22]. At present, the CNN-LSTM hybrid network is widely used in video analysis. To extract and classify features (such as expressions, actions, tones, and speech rates) in the video simultaneously, we adopt a multi-channel CNN-LSTM network.

Referring to the methods for processing images proposed in [22–24], images are first processed to a proper form (discussed in [Experiment](#) section part A). We obtain the local spatial features of the video sequences through CNN's sliding windows and weight sharing and then use them as the inputs to the LSTM layer. We use LSTM to acquire the time characteristics of video data. Then, they are combined to make full use of their respective advantages.

For the processing of audio signals in videos, we apply the same CNN-LSTM network. Audio and images are input simultaneously to two different channels. The difference between audio and image processing is the pre-process (discussed in [Experiment](#) section part A).

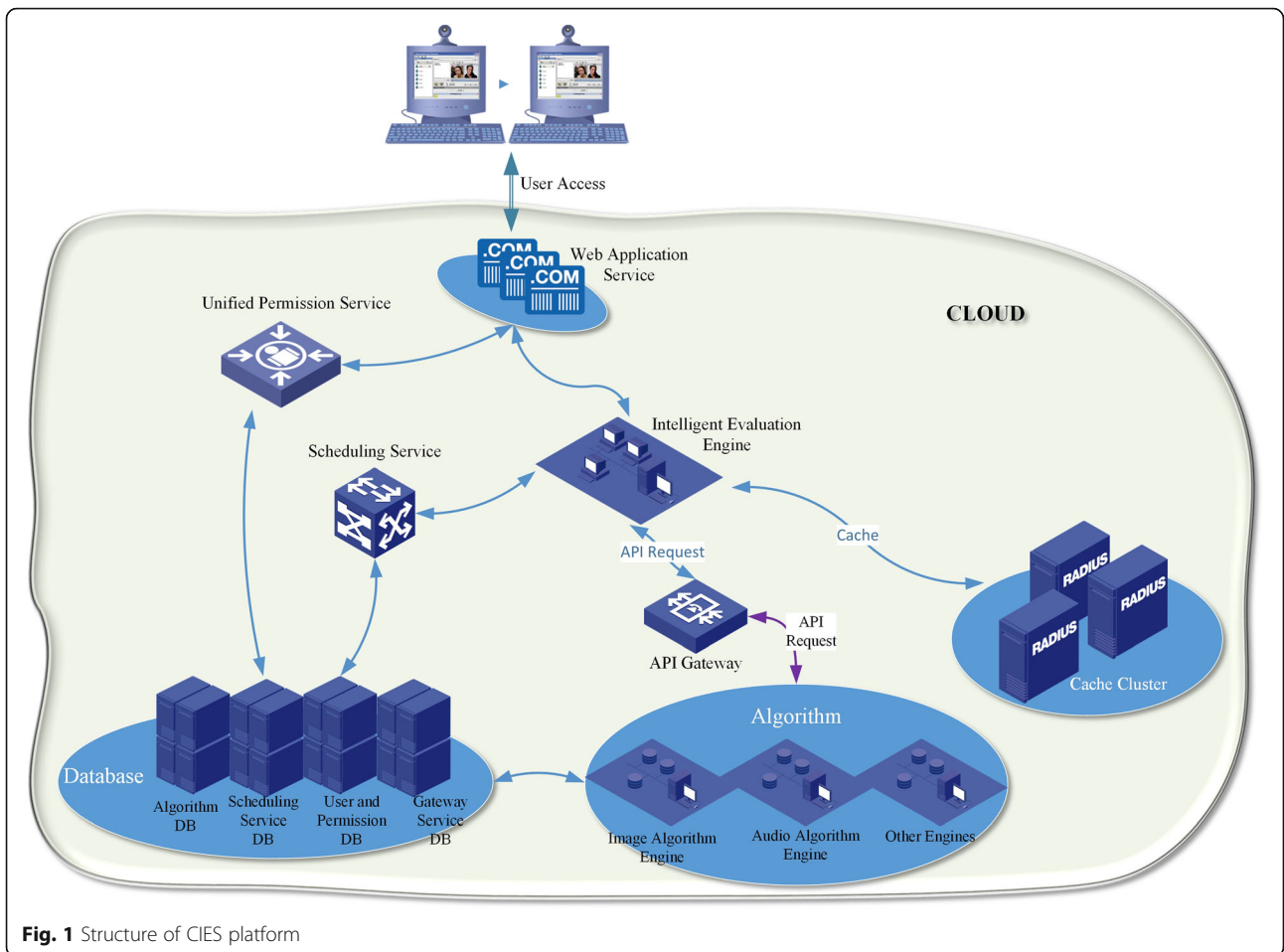


Fig. 1 Structure of CIES platform

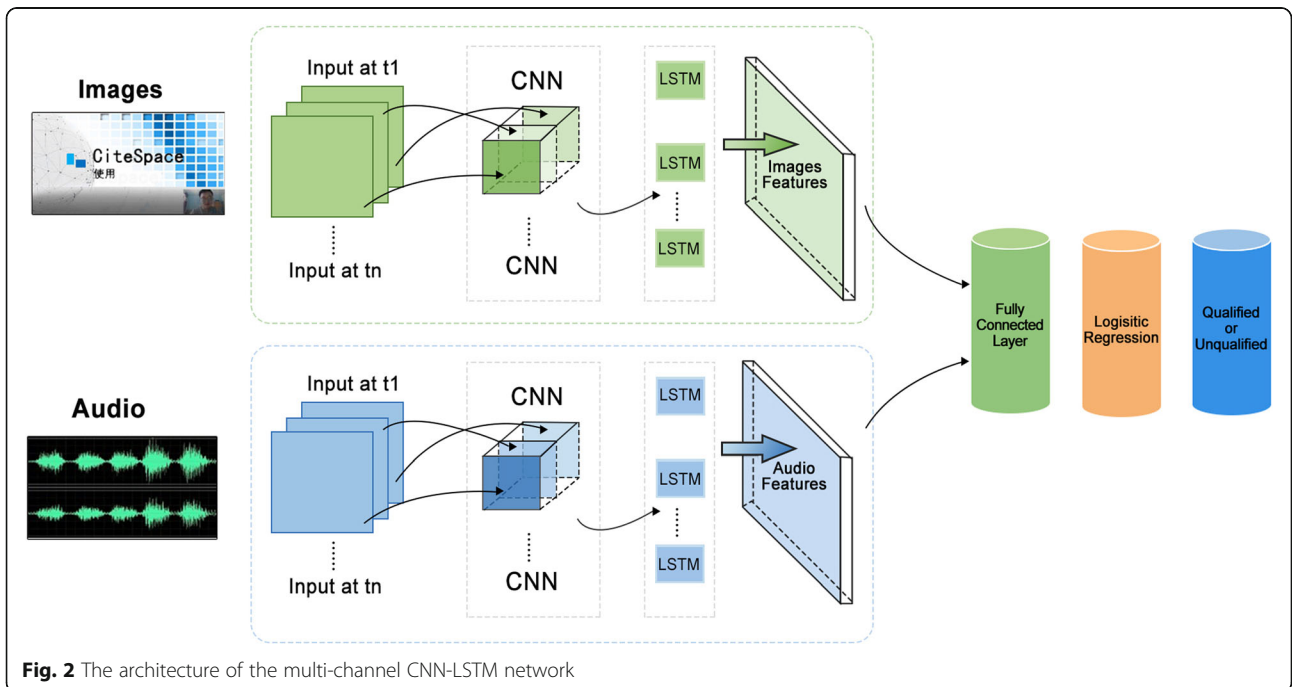


Fig. 2 The architecture of the multi-channel CNN-LSTM network

Subsequent processes are the same as the processing of images (Fig. 2).

Convolutional neural network

We adopt the deep residual network ResNet-50 as the convolutional neural network. This multi-channel CNN architecture has 2 separate input channels for audio and images, and they share the same parameter settings. All the strides for the convolutional layers are 2 [25].

In Fig. 3, CONV denotes a convolution operation. Batch Norm denotes batch normalization. ReLU denotes the ReLU activation function.

MAX POOL stands for maximum pooling

$$R_n = \max(Y_{ij})$$

R_n represents the feature matrix of the n^{th} image in the sequence image after the convolution and pooling operations. The above operations are performed on the image sequence separately, and the feature matrix of each image frame in the sequence can be represented by $R = (R_1, R_2, \dots, R_n)$.

Avg POOL denotes mean pooling. Flatten denotes the flattened layer. ID BLOCK denotes a residual block that does not change dimensions, called the identity block, and CONV BLOCK denotes the

residual block that adds dimension. Each residual block includes 3 convolution layers (Fig. 4).

Long short-term memory

The entire system contains 1 LSTM, following CNN. We adopt an LSTM structure given in [22]. An output R of the CNN's pooling layer corresponds to an LSTM input at time t, and the result of each recursive operation is a integration of all previous features and current features. At the t moment, the components of the LSTM unit are updated as follows:

$$i_t = \sigma(W_{ri}R_t + U_{hi}h_{t-1} + b_i)$$

$$f_t = \sigma(W_{rf}R_t + U_{hf}h_{t-1} + b_f)$$

$$\tilde{c}_t = \tanh(W_{rc}R_t + U_{hc}h_{t-1} + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

$$o_t = \sigma(W_{ro}R_t + U_{ho}h_{t-1} + b_o)$$

$$h_t = o_t \tanh(c_t)$$

where σ denotes the sigmoid activation function, R_t denotes the feature matrix input at time t, W_{ri} , W_{rf} , W_{rc} , and W_{ro} denote the weight matrix between the input layer to the input gate, the forget gate, the memory cell, and the output gate, respectively, U_{hi} , U_{hf} , U_{hc} and U_{ho}

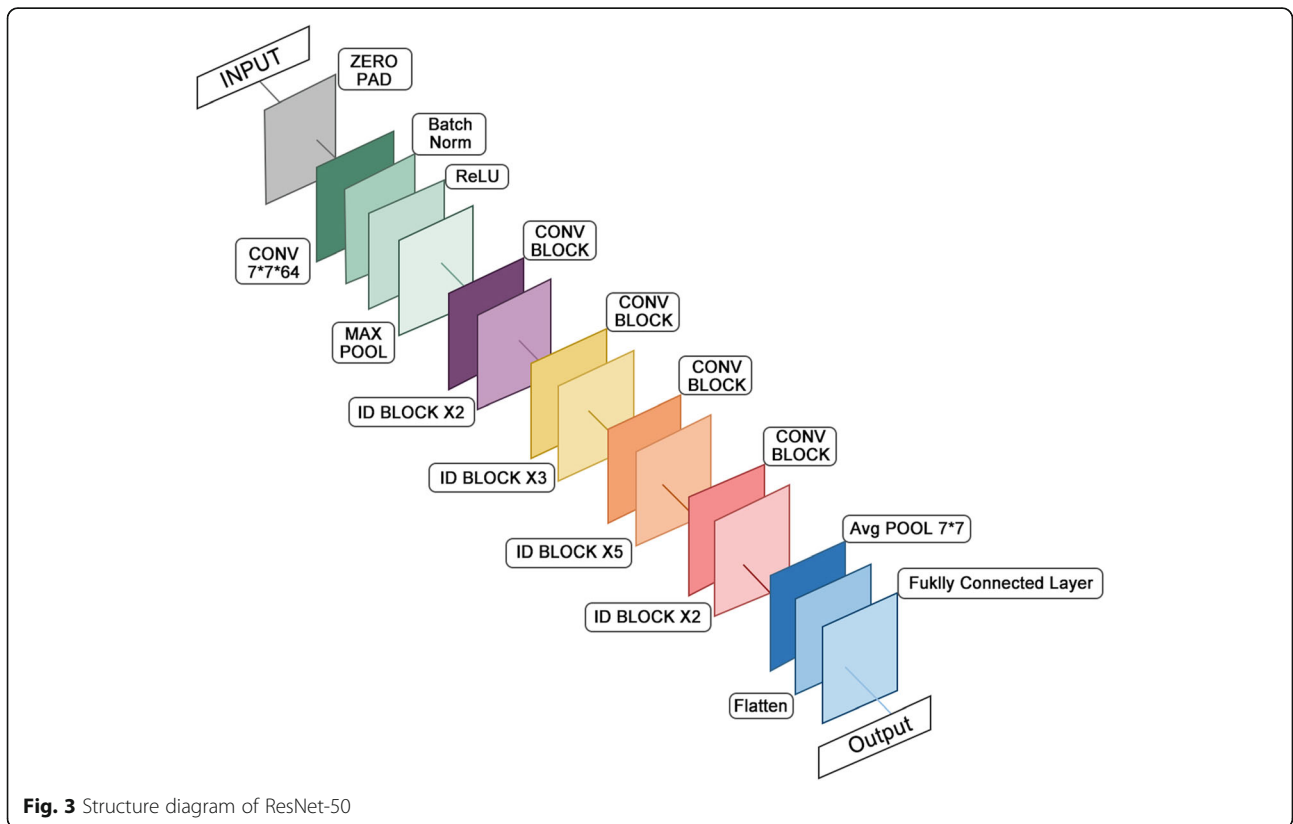
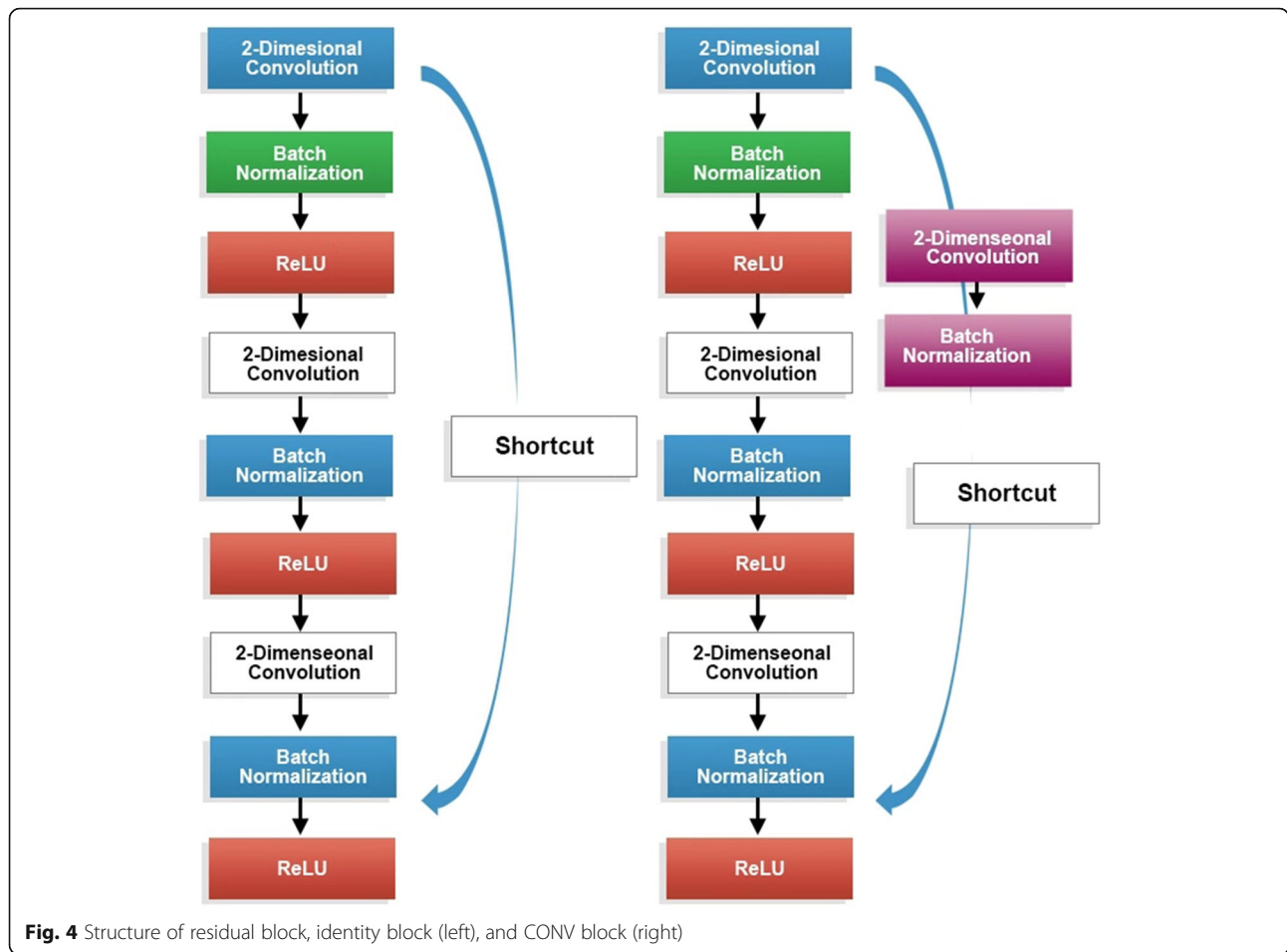


Fig. 3 Structure diagram of ResNet-50



denote the weight matrix from the hidden layer to the input gate, forget gate, memory cell and output gate, respectively, b_i , b_f , b_c , and b_o denote the offset value of the input gate, forget gate, memory cell and output gate, respectively.

Loss function

As the results are two-category classification, we adopt logistic regression in this network.

The loss function is defined as

$$L(\hat{y}-y) = -y\log(\hat{y})-(1-y)(\log(1-\hat{y}))$$

where y denotes the true classification of the sample and \hat{y} denotes the model recognition result.

Optimization

We adopt the Adam optimizer in this network, and the implementation process is as follows:

Compute the gradient at moment t :

$$g_t = \nabla_{\theta}(\theta_{t-1})$$

Update biased first moment estimate:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)g_t$$

Update biased second raw moment estimate:

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2$$

Compute bias-corrected first moment estimate:

$$\hat{m}_t = m_t / (1-\beta_1^t)$$

Compute bias-corrected second raw moment estimate:

$$\hat{v}_t = v_t / (1-\beta_2^t)$$

Update parameters:

$$\theta_t = \theta_{t-1} - \alpha * \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)$$

Experiment

Since there is no suitable existing datasets, one is constructed by our own. The raw data come from practical video homework submitted by students in the management course offered by Shandong University of Finance and Economics during 2nd semester, 2018/2019. Then

Table 1 Video duration information

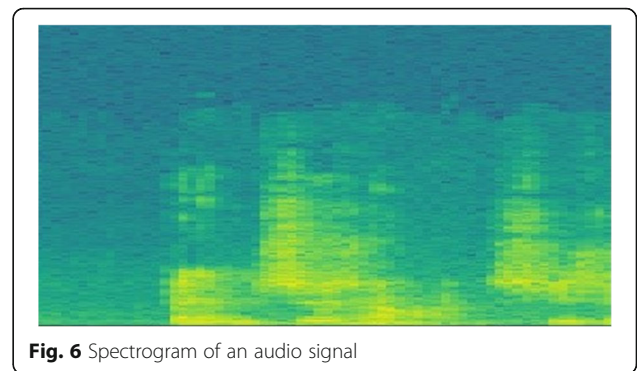
Number of video segments	61
Duration of the longest video	17:32
Duration of the shortest video	3:23
Average duration	9:14

videos are assessed by 3 experienced professors respectively, only the same assessments are adopted. These assignments account for 50% of students' final results. The dataset contains 61 students' video homework (38 qualified and 23 unqualified) submitted by 48 different students.

Pre-processing

To increase the number of samples and specification data, pre-processing is required. The students' homework contains videos of different lengths, and the video length statistics are shown in Table 1. The videos are first segmented by the same length. New video classification still belongs to the category of the original video after segmenting. After this process, the training set samples reach 9898, and the test set samples reach 122. Then 76 qualified segments and 46 unqualified segments are selected as test samples.

Images are also pre-processed. The first step of image pre-processing is frame extraction: one frame is extracted at the intervals of 3 frames. The assessment is based on features such as students' actions and expressions. However, the PPT occupies a large part of the entire image, as shown in Fig. 5, so the position-sensitive region segmentation method is used in the frame images to extract, then scale the important part to generate a fixed-size image of 224*224 pixels. Finally, the number of images extracted from qualified video is 215,730, and the number of images extracted from unqualified video is 85,230.



To pre-processing the audio signals, firstly divide each video into segments by the time interval of 1 s. Then extract the audio signals from the video. At last audio signals are subjected to spectrum analysis to obtain a spectrogram, as shown in Fig. 6. The spectrogram is also scaled into the same fixed size of 432*288 pixels. Spectrograms extracting from qualified video reach 13,500, while the ones from unqualified video is 5850.

Training and evaluation

We adopt logistic regression to make a two-category (qualified or unqualified) identification, where 1 denotes qualified, and 0 denotes unqualified. The loss function is defined as

$$L(\hat{y}-y) = -y\log(\hat{y})-(1-y)(\log(1-\hat{y}))$$

where y denotes the true classification of the sample and \hat{y} denotes the model recognition result.

The Adam optimizer is adopted in this network, where the learning rate is 0.01.

To evaluate the superiority of the approach that adopts both audio and images as input, we conducted comparison experiments between the approaches that input only

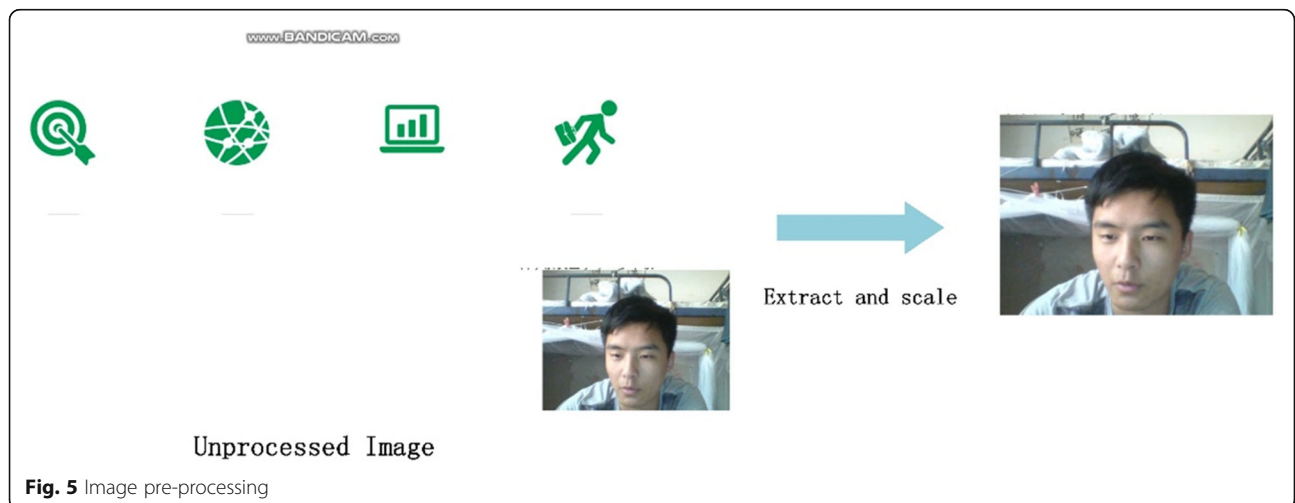


Table 2 Accuracy of different inputs

Features	Accuracy
Only image	63.9%
Only audio	68.0%
Fusing image and audio	75.4%

audio and only images. The accuracies are shown in Table 2. As shown in Table 2, the proposed approach which integrates image and audio is feasible and more accurate.

To verify the advantages of CNN-LSTM hybrid architecture, experiments between CNN-LSTM network and some typical CNN networks are conducted as following. In these experiments, all networks used same dataset which contained both images and audio. The accuracies are shown in Table 3. According to Table 3, CNN-LSTM network has the best performance.

According to the experiments, this model could preliminarily distinguish qualified and unqualified video. As the train set and test set varies, the accuracy is relatively steady. However, the accuracy is not very satisfying for some videos turn out to be too dark to recognize the expression or vague voice because of noises, discontinuous, etc. To further improve accuracy, another algorithm is required to detect and process defective videos automatically.

Conclusion

With the promotion of video homework, an efficient and accurate approach to mark these videos is in demand. However, most proposed intelligent grading studies focus on text, audio or images, rare methods can be used on video homework. Hence, we propose an approach using a multi-channel CNN-LSTM network to assess video homework intelligently. A novel method of integrating image features and audio features on the topic of intelligent evaluation of homework is also presented. This approach preliminarily classifies qualified and unqualified video homework, which has been demonstrated by experiments. In addition, CIES platform improves the computing efficiency and makes it more convenient to use the model.

However, it should also be noted that the accuracy of this approach is not extremely satisfying. The model introduced in [26, 27] is designed to make automatic evaluation (based on text) for English essay. It could reach a pretty high accuracy. In most cases, there is only ±0.25 points error between scores given by human and

Table 3 Accuracy of different networks

Network	Accuracy
GoogLeNet Inception V3	67.6%
VGG16	69.1%
CNN-LSTM	73.5%

given by this model. Similarly, the method introduced in [28] is an automated speech scoring system (based on audio). The correlation coefficient between human grading and this method is proved to be 0.97. As for the grading of programs, the method proposed in [29] has the accuracy of 94.48%. Compared with the achievements in automatic grading for other forms of homework, we simply propose a preliminary model to deal with video homework. Its accuracy can still to be improved.

Additionally, the proposed model is only aimed at two-category classification. A model that can assign more specific grades is still to be discovered in the future.

Abbreviations

CIES: Cloud-based Intelligent Evaluation Service; CNN: Convolutional Neural Networks; LSTM: Long Short-term Memory

Acknowledgements

Not applicable.

Authors' contributions

All authors have participated in conception and design, or analysis and interpretation of the data, drafting the article or revising it critically for important intellectual content, approval of the final version.

Funding

Teaching Reform Research Project of Undergraduate Colleges and Universities of Shandong Province (Z2016Z036), the Teaching Reform Research Project of Shandong University of Finance and Economics (jy2018062891470, jy201830, jy201810), Shandong Provincial Social Science Planning Research Project (18CHLJ08), Scientific Research Projects of Universities in Shandong Province (J18RA136), Youth Innovative on Science and Technology Project of Shandong Province (2019RWF013).

Availability of data and materials

Because dataset and code involve our interests, we're sorry but we can't publish dataset or source code at present. However, they are available from the corresponding author on reasonable request.

Competing interests

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue. The authors declare that they have no competing interests among authors.

Author details

¹School of Control Science and Engineering, Shandong University, Jinan 250002, China. ²College of Electromechanical Engineering, Qingdao University of Science and Technology, Qingdao 266061, China. ³School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan 250014, China. ⁴College of Electrical Engineering and Automation, Shandong University of Science and Technology, Jinan 250031, China.

Received: 19 October 2019 Accepted: 22 January 2020

References

- Jinyu T, Fang C et al (2018) Application of video work in higher vocational nursing training class [J]. *Nurs Res China* 35(1):101–120
- Chen Y, Ding S, Xu Z, Zheng H, Yang S (2019) Blockchain-Based Medical Records Secure Storage and Medical Service Framework. *J Med Syst* 43(1):5: 1–5:9
- Zhuang-lin HU, Jia DONG et al (2006) How Is Construed Multimodally — A case study of a PowerPoint presentation contest [J]. *Techn Enhanced Foreign Lang Educ* 03:3–12

4. Caixia W, Xinping Z et al (2015) Research on presentation teaching and high quality PPT making strategy [J]. *Comput Era* 07:78–80
5. Wei LIU, Zi-sen QI, Mu-xuan WANG et al (2016) Automated Assessment of Subjective Tests [J]. *J Beijing Univ Posts Telecomm (Social Sciences Edition)* 18(04):108–116
6. Gao S, Chunfeng Y et al (2004) The Application of Sentence Similarity Measurement in Automated Assessment Technology of Subjective Tests [J]. *Comput Eng Appl* 14:132–135
7. Wei-ping DING, Zhi-jin GUAN, Jian-ping CHEN et al (2007) Research and Application of Intelligent Assessment Algorithm Based on Programming Subjective Questions [J]. *Comput Technol Dev* 11:205–208
8. Alawadi S (2019) Manuel Fernández Delgado, David Mera, Senén Barro. Polynomial Kernel Discriminant Analysis for 2D visualization of classification problems. *Neural Comput & Applic* 31(8):3515–3531
9. Bhatia S, Singh R (2016) Automated Correction for Syntax Errors in Programming Assignments using Recurrent Neural Networks [J]
10. Weiqian L, Guoliang W, Jia L, Runsheng L et al (2005) Phone-based pronunciation quality assessment algorithm [J]. *J Tsinghua Univ (Science and Technology)* (01):5–8
11. Shen H, Hongyan LI, Shijin W, Jia'en L, Bo X et al (2009) Automatic assessment of speech fluency in computer aided speech grading systems [J]. *J Tsinghua Univ (Science and Technology)* 49(S1):1349–1355
12. Wang Z, Zechner K, Sun Y (2018) Monitoring the Performance of Human and Automated Scores for Spoken Responses [J]. *Lang Test* 35(1):101–120
13. Zhang Yanbin, Hu Yueping, Chu Min, Huang Cha, Liang Mangui, et al. Mandarin tone error detection [J]. *J Tsinghua Univ (Science and Technology)*, 2008(S1):683–687
14. Long Z et al (2014) Research on Automatic Evaluation Methods of Mandarin Pronunciation Quality [D]. Harbin: Harbin Institute of Technology
15. Kai D et al (2015) An Automatic Marking System Based on Handwritten Character Recognition [D]. Taiyuan: Taiyuan University of Technology
16. Zhang X-Y, Bengio Y, Liu C-L (2017) Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark [J]. *Pattern Recogn* 61(61):348–360.
17. Zu-ying L, Dan-hui Z et al (2018) Auto-evaluation of Classroom Teaching & Learning and Its Preliminary Research Findings [J]. *Mod Educ Technol* 28(08):38–44
18. Yan-tao W, Dao-ying Q, Jia-min H, Huang Y, Ya-fei S et al (2019) The Recognition of Students' Classroom Behaviors based on Deep Learning [J]. *Mod Educ Technol* 29(07):87–91
19. Gammulle H, Denman S, Sridharan S, Fookes C (2020) Fine-grained action segmentation using the semi-supervised action GAN [J]. *Pattern Recogn* 98: 107039
20. Yang K, Shen X, Qiao P, Li S, Li D, Dou Y (2019) Exploring frame segmentation networks for temporal action localization [J]. *J Visual Comm Image Representation* 61:296–302
21. Zheng Z, An G, Wu D, Ruan Q (2019) Spatial-temporal pyramid based Convolutional Neural Network for action recognition [J]. *Neurocomputing* 358:446–455
22. Bhattacharya S, Roy S, Chowdhury S (2018) A neural network-based intelligent cognitive state recognizer for confidence-based e-learning system. *Neural Comput & Applic* 29(1):205–219
23. Wu Z, Wang X, Jiang YG et al (2015) Modeling Spatial-Temporal Clues in a Hybrid Deep Learning Framework for Video Classification [J]. *IEEE Trans Multimedia* 20(11):3137–3147
24. Ferreira JP, Vieira A, Ferreira P, Crisóstomo MM (2018) A. Paulo Coimbra. Human knee joint walking pattern generation using computational intelligence techniques. *Neural Comput & Applic* 30(6):1701–1713
25. Heng W, Xia L, Xiaofang L, Xu W et al (2019) Classification of breast cancer histopathological images based on ResNet50 network [J]. *J China Univ Metrology* 30(01):72–77
26. Williamson DM (2015) A Study of the Use of the "e-rater"® Scoring Engine for the Analytical Writing Measure of the "GRE"® revised General Test [J]. *Ets Res Rep* 2014(2):1–66
27. Ramineni C, Trapani CS, Williamson DM (2015) Evaluation of e-rater® for the Praxis I® Writing Test [J]. *ETS Res Rep Series* 2015(1):1–28
28. Luo K-z, Bao-cheng H (2014) A Critical Review of Two Automated Speech Scoring Systems: Ordinate & SpeechRater [J]. *Comput Assist Foreign Lang Educ* 04:27–32
29. YueXia LIU, Zhiyao NIU, Ning WU (2016) An automatic Scoring Method of Student programs Using Multi-Feature Analysis for Massive Open Online Courses [J]. *Journal of Xi'an Jiaotong University* 50(10):64–70

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)