Machine Learning-Based Student Emotion Recognition for Business English Class

Yuxin Cui, Sheng Wang
Beijing Forestry University, Beijing, China

Ran Zhao (✉)
Beijing Institute of Economics and Management, Beijing, China
zhaoran@biem.edu.cn

Abstract—Traditional English teaching model neglects student emotions, making many tired of learning. Machine learning supports end-to-end recognition of learning emotions, such that the recognition system can adaptively adjust the learning difficulty in English classroom. With the help of machine learning, this paper presents a method to extract the facial expression features of students in business English class, and establishes a student emotion recognition model, which consists of such modules as emotion mechanism, signal acquisition, analysis and recognition, emotion understanding, emotion expression, and wearable equipment. The results show that the proposed emotion recognition model monitors the real-time emotional states of each student during English learning; upon detecting frustration or boredom, machine learning will timely switch to the contents that interest the student or easier to learn, keeping the student active in learning. The research provides an end-to-end student emotion recognition system to assist with classroom teaching, and enhance the positive emotions of students in English learning.

Keywords—Teaching model, machine learning, emotion recognition, emotion understanding, emotion expression

1 Introduction

Emotions are the psychological states and emotional responses of a person to things based on the person's subjective experience in a certain environment. They are the physiological and psychological responses of the person to the good or bad information in the environment he/she is experiencing [1]. With the improvement of the computing power of computers and the advancement of technologies such as neural network, machine translation, and behavior recognition, artificial intelligence has made positive progresses, starting from simple seeing and hearing, it now has entered the stage of human-computer interaction [2]. Machine learning is an important field of artificial intelligence. Only by truly understanding the emotions of humans can the machines make correct responds to the environment using the thinking patterns of

mankind [3, 4]. Human emotions are extremely rich in states, and the machines need to understand various signals corresponding to each emotion to truly understand these emotions [5].

Emotion recognition technology is a field related to artificial intelligence; it can help computers intelligently recognize human emotions [6]. With the continuous development of the emotion recognition technology, there are more and more researches on emotion recognition, and these researches have occupied an important position in different application fields such as human-computer interaction [7]. The primary objective of emotion recognition systems is to interpret input signals of different patterns [8, 9]. Human expressions and actions often carry more emotional information than the language. Most human emotion identification methods are performed based on the analysis of facial expressions. However, in many cases, people tend to conceal their true emotions, but they can hardly hide their body languages [10].

During the learning of business English, Emotional Intelligence (EI) plays an important role in English learning. Students with high EI can quickly relieve negative emotions such as dislike and depression during the learning process by regulating their own emotions [11, 12]. And high EI students know how to motivate themselves to have positive emotions in learning and maintain good learning status [13]. The learning of English is a very complex psychological process; the learning process involves three stages of language recognition, language comprehension, and language communication, which respectively corresponding to language information processing, storage, and extraction [14]. In terms of the defining of machine learning environment, during the learning of business English, students can learn and practice in a purposeful, planned, and organized way, develop their knowledge and ability in English recognition, understanding and communication, and use emotions to express their business English learning during this process [15, 16]. Based on machine learning and business English class, this paper proposed a student facial expression feature extraction method and constructed a student emotion recognition model to obtain the emotional states of learners of business English class.

2 Emotion Recognition Technology

2.1 Facial expression recognition

Facial expressions are an important way for humans to express their emotional states, and the emotional states of learners can be recognized by analyzing and processing their facial expressions [17]. Figure 1 shows the two-dimensional theory of emotions. In the figure, the arrow direction indicates increase of arousal values, when pleasure increases, positive emotion increases; when pleasure decreases, negative emotion appears. According to this theory, surprised and excited correspond to positive emotions; while anger, sad, hate, and fear correspond to negative emotions [18, 19]. Positive emotions promote the occurrence of cognition activities, while negative emotions hinder the cognition process [20, 21].
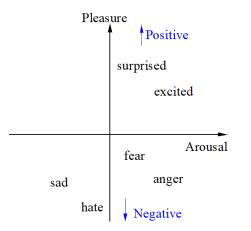
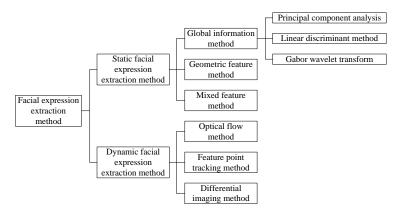**Fig. 1.** Two-dimensional theory of emotions

Fig. 2. Facial expression feature extraction methods

According to the features of different expressions, facial expression images can be divided into static image features and dynamic sequential image features. The expression recognition based on static images can extract, classify, and identify facial expression feature information; while the dynamic sequential facial expression images contain the dynamic information of the continuous changes of expressions, and they can reflect the changing process of the facial expressions [22]. Figure 2 shows the facial expression feature extraction methods. The static facial expression extraction method includes three types: global information method, geometric feature method, and mixed feature method; wherein the global information method contains the principal component analysis method, the linear discriminant method, and the Gabor wavelet transform method. The dynamic facial expression extraction method includes optical flow method, feature point tracking method, and differential imaging method. To select key features that can summarize the characteristics of facial expression changes, first, the facial movement units should be analyzed, the data of the multiple geometric shapes between facial feature points in the sequential images should be

calculated, and these data need to be mapped to figures and tables for comparative analysis [23, 24]. Then, according to the characteristics of the data changes in the images, features with more obvious changes should be selected; after that, through analysis, changes of facial feature points should be extracted to reflect the changes in facial features, so as to improve the expression recognition rate [25].

2.2 Analysis of emotion recognition algorithm

According to the research of domestic and foreign scholars, the basic process of facial expression recognition is to preprocess the human face images after obtaining the facial expression images, and then extract the facial expression features based on these processed images, and finally select a suitable classifier to realize the facial expression classification, as shown in Figure 3. A facial expression recognition system generally includes two parts: a system workbench and an inference engine. The system workbench forms a framework for mixed facial feature detection with multiple feature detection technologies applied in parallel; and the redundant information is used to define the parts in the image that do not contain the geometrical structure of human faces, and it also contains missing or highly inaccurate data [26, 27]. The inference engine can convert low-level facial geometric shapes into high-level facial behaviors, and finally into high-level weighted emotion tags [28].

Fig. 3. Facial expression recognition process

3 The Emotion Recognition Method for Learners Of Business English Class

3.1 Feature extraction method of facial expression sequential images based on fused features

Facial feature extraction is the focus of image recognition. Extracting key features that can summarize facial feature changes can improve the recognition effect of the classifier, therefore, in the entire recognition process, feature extraction is of great significance to facial expression recognition [29]. In order to solve problems such as too high feature dimensions, too much memory consumption, and information redundancy, this paper adopted the feature extraction method based on the peak frame of expression images, which can directly use the expression peak images provided by static database or manually select expression images provided by dynamic database. The changing process of facial expressions is reflected by extracting facial feature points with higher contributions to the expression changes. One method of dynamic feature extraction is to track the facial feature points of facial expression changes, during the tracking process, background information that is irrelevant to expressions

could be ignored. For the feature extraction method based on fused feature sequence, although it can use the position change information of feature points to represent the complete human facial expressions, still it's difficult to summarize the specific shape changes of each organ.

3.2 Facial expression classification method based on deep multi-kernel learning

The core of expression extraction is the extraction of features. A suitable classification algorithm should be chosen to process the labeled training set, and then the classifier model obtained after processing is used to classify and identify the remaining dataset of unknown expression categories. Multi-kernel learning is to obtain the best kernel by conducting combinatorial learning on a set of defined basic kernels. Deep learning is a machine learning method that interprets and analyzes images by simulating the mechanism of human brain. This method has a very strong learning ability, and it shows great advantages in image classification and shallow learning. Deep multi-kernel learning is composed of multiple layers of kernel functions and neural networks. The kernel function is the key component of deep multi-core learning, and its architecture is a multi-layer network architecture, and each layer has a set of kernels.

Fig. 4. An example image of the expression "surprise"

Facial expression recognition uses expression database for testing, based on the deep core-kernel learning method, expression images such as Figure 4 are selected as examples. First, by extracting expression images from each frame of human face pictures, the features of geometric shape changes in local areas of the human face picture are described, then, the Gabor features describing the textures are extracted from the peak frame of human facial expression images, and the extracted two-categories of features are mixed in series and input into the deep multi-kernel learning model for training. Finally, the trained model is applied to classify the expressions. Figure 5 shows the model recognition rate based on deep multi-kernel learning. It can

be clearly seen from the figure that, compared with the classification method based on geometric feature classification, the recognition rate of the classification method based on deep multi-kernel learning and fused features has been improved. For the expression given in Figure 4, the recognition rate of the expression "surprise" was 100%, and the second highest recognition rate was the expression "happy".
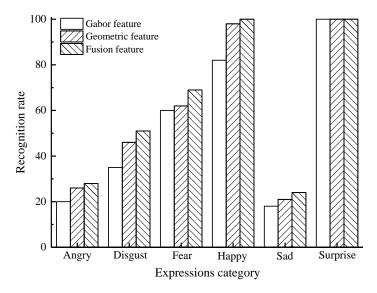

Fig. 5. Model recognition rate based on deep multi-kernel learning

4 Human-Computer Interaction Control Method Based on Machine Learning

4.1 Human-computer interaction control method based on emotion recognition

Machine learning is a general method of artificial intelligence. The use of machine learning does not require assumptions of unknown or undefined mechanisms. Typical machine learning includes data normalization, representation learning, model fitting and evaluation, etc. [30]. The emotion recognition system monitors the emotional states of students during English learning in real time. When it detects that a student is frustrated or bored, the machine learning will timely switch to the contents that interest the student or easier to learn, keeping the student active in learning. The emotion recognition system consists of four layers, namely the hardware layer, software layer, information layer and service layer; wherein the hardware layer includes the physiological signal acquisition device and the rehabilitation robot system; the software layer processes the collected physiological signals; the information layer is responsible for the analysis of personal emotion information files and emotion performance;

the service layer adjusts the corresponding learning content after receiving the emotion results identified by the software layer. Figure 6 shows the emotional state regulation mechanism. After difficulty level is initialized, if frustration is detected and the system is not at the lowest difficulty level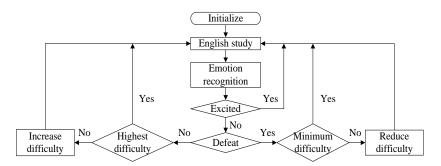, the difficulty level of English learning will be decreased; if boredom is detected and the system is not at the highest difficulty level, the difficulty level of English learning will be increased.

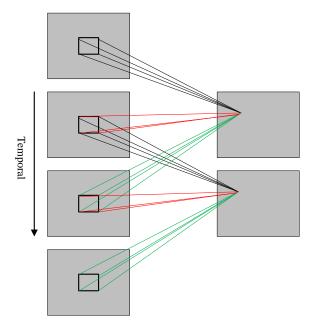**Fig. 6.** Emotional state regulation mechanism

Fig. 7. Convolutional mode of the 3D convolutional neural network

Figure 7 shows a diagram of the convolution method of a 3D convolutional neural network. The 3D convolutional neural network can solve temporal and spatial problems at the same time. One convolution kernel can convolve the images of multiple adjacent frames at a same time, and the values in the feature images are convolutions of the same position of each image in the previous layer. Except for this aspect, the

principles of the 3D convolutional neural network are the same with those of the 2D convolutional neural network [31]. During English learning, the expressions captured by the machine are single-frame static expressions, which can be analyzed by the convolutional neural network. Figures 8 and 9 are the accuracy and loss of the convolutional neural network model. By comparing the results of training and testing, we can see that the experimental results agreed well.
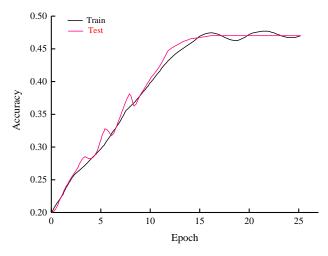

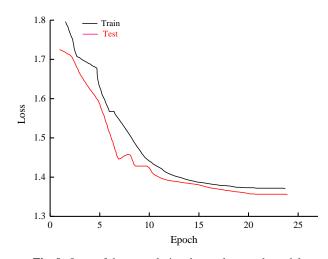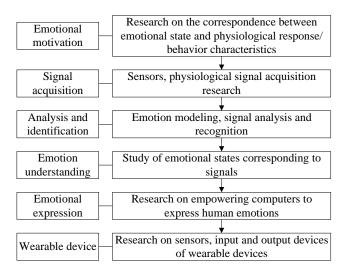
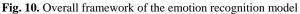**Fig. 8.** Accuracy of the convolutional neural network model

Fig. 9. Loss of the convolutional neural network model

4.2 Construction of the emotion recognition model based on machine learning

Figure 10 shows the overall framework of the emotion recognition model, including emotion mechanism, signal acquisition, analysis and recognition, emotion under-

standing, emotion expression, and wearable equipment. Emotion mechanism is the reaction of emotional states and physiological response/behavioral characteristics; emotion understanding is the expression of emotional states corresponding to the extracted emotion signals; emotion expression is the study of giving computers the human emotional expressions. Figure 11 shows the emotion recognition model. Users log in to the system by providing user information; the system acquires images through cameras during machine learning, then, it performs face detection on the original surveillance images, preprocesses the images, and recognizes expressions from the preprocessed human face images; and finally, the system can classify the emotions in the emotion recognition results.
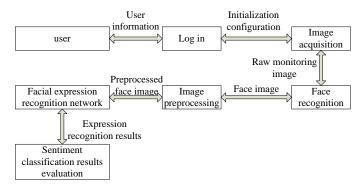

Fig. 10. Overall framework of the emotion recognition model

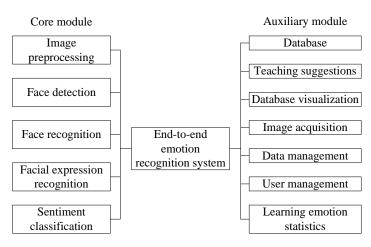**Fig. 11.** Emotion recognition model

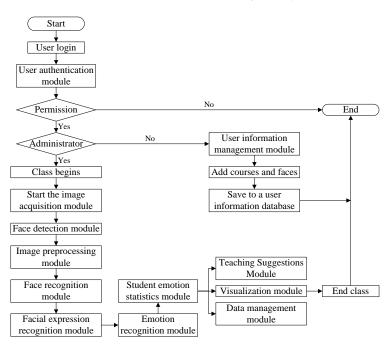**Fig. 12.** Modules of the emotion recognition system

Fig. 13. Flow of function modules of the emotion recognition system

Figure 12 shows the modules of the emotion recognition system. The system includes two main parts: the core modules and the assistant modules. The core modules include image preprocessing, face detection, face recognition, expression recognition and expression classification; and the assistant modules include the database, teaching suggestions, database visualization, image acquisition, data management, user management, and learning emotion statistics. Figure 13 shows the flow of function mod-

ules of the emotion recognition system. When a learner acquires permissions, the image acquisition module, face detection module, image preprocessing module, face recognition module and expression recognition module will turn on one by one; the emotion recognition module will analyze the learning emotions, and the learning emotion statistics module will provide suggestions for the teaching suggestion module, the visualization module, and the data management module. In the entire emotion recognition model, in order to improve the emotion recognition accuracy of the system, the images captured by the image acquisition device should be preprocessed, including the contrast, brightness and gray level of the images.

5 Conclusion

Based on machine learning and business English class, this paper proposed a feature extraction method for students' facial expression features, established a student emotion recognition model, and obtained the emotional states of learners of business English class. The conclusions drawn in this research are as follows:

1. The static facial expression extraction method includes three types: global information method, geometric feature method, and mixed feature method; the dynamic facial expression extraction methods include optical flow methods, feature point tracking method, and differential imaging method.
2. One method of dynamic feature extraction is to track the feature points of facial expression changes, during the tracking process, background information that is irrelevant to expressions could be ignored.
3. The emotion recognition system monitors the emotional states of students during English learning in real time. When it detects that a student is frustrated or bored, the machine learning will timely switch to the contents that interest the student or easier to learn, keeping the student active in learning. The emotion recognition system consists of four layers, namely the hardware layer, software layer, information layer and service layer.
4. The overall framework of the emotion recognition model includes emotion mechanism, signal acquisition, analysis and recognition, emotion understanding, emotion expression, and wearable equipment.

6 References

[1] Marui, B., Katavi-Aui, S. (2018). The word class adjective in English business magazines online. Journal of Language and Cultural Education, 6(2): 99-115. https://doi.org/10.2478/jolace-2018-0018

[2] González, J.F.E. (2019). Increasing motivation for in-class reading comprehension in a business English course at the university of costa Rica (UCR). Research in Pedagogy, 9(2): 254-265. https://doi.org/10.17810/2015.106

[3] Guo, Q. (2020). Detection of head raising rate of students in classroom based on head posture recognition. Traitement du Signal, 37(5): 823-830. https://doi.org/10. 18280/ts.370515

[4] Anthwal, S., Ganotra, D. (2020). Dynamic features based on flow-correlation and hog for recognition of discrete facial expressions. Revue d'Intelligence Artificielle, 34(5): 585-594. https://doi.org/10.18280/ria.340508

[5] Houari, H., Guerti, M. (2020). Study the influence of gender and age in recognition of emotions from algerian dialect speech. Traitement du Signal, 37(3): 413-423. https://doi.org/10.18280/ts.370308

[6] Adamson, J. (2005). Using marketing visuals for product talk in business english classes. Business Communication Quarterly, 68(3): 358-367. https://doi.org/10.1177/1080569905279020

[7] Deng, L.J. (2018). The project-based flipped learning model in business English translation course: learning, teaching and assessment. English Language Teaching, 11(9): 118-128. https://doi.org/10.5539/elt.v11n9p118

[8] Zhang, X., Tao, H., Zha, C., Xu, X., Zhao, L. (2015). A robust method for speech emotion recognition based on infinite student's -mixture model. Mathematical Problems in Engineering, 2015, 475810. https://doi.org/10.1155/2015/475810

[9] Reddy C.V.R., Reddy U.S., Kishore K.V.K. (2019). Facial emotion recognition using NLPCA and SVM, Traitement du Signal, 36(1): 13-22. https://doi.org/10.18280/ts.360102

[10] Yang, D., Alsadoon, A., Prasad, P.W.C., Singh, A.K., Elchouemi, A. (2018). An emotion recognition model based on facial recognition in virtual learning environment. Procedia Computer Science, 125: 2-10. https://doi.org/10.1016/j.procs.2017.12.003

[11] Rasheed-Karim, W. (2020). The Effect of Stressful Factors, Locus of Control and Age on Emotional Labour and Burnout among Further and Adult Education Teachers in the U.K., International Journal of Emerging Technologies in Learning, 15(24): 26-37. https://doi.org/10.3991/ijet.v15i24.19305

[12] Rasheed-Karim, W. (2020). The Influence of Policy on Emotional Labour and Burnout among Further and Adult Education Teachers in the U.K., International Journal of Emerging Technologies in Learning, 15(24): 232-241. https://doi.org/10.3991/ijet.v15i24.19307

[13] Yang, G., Jiao, R., Jiang, H., Zhang, T. (2020). Ground truth dataset for EEG-based emotion recognition with visual indication. IEEE Access, 8: 188503-188514. https://doi.org/10.1109/access.2020.3030680

[14] Hussain, S.A., Balushi, A.S.A.A. (2020). A real time face emotion classification and recognition using deep learning model. Journal of Physics: Conference Series, 1432(1): 012087. https://doi.org/10.1088/1742-6596/1432/1/012087

[15] Hayat, A.A., Shateri, K., Amini, M., Shokrpour, N. (2020). Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical students: a structural equation model. BMC Medical Education, 20(1): 76. https://doi.org/10.1186/s12909-020-01995-9 https://doi.org/10.21203/rs.2.13211/v3

[16] Meng, X., Zhang, J. (2020). Anxiety recognition of college students using a Takagi-Sugeno-Kang fuzzy system modeling method and deep features. IEEE Access, 8: 159897-159905. https://doi.org/10.1109/access.2020.3021092

[17] Li, A.H., An, L., Che, Z.H. (2020). A Facial expression recognition model based on texture and shape features. Traitement du Signal, 37(4): 627-632. https://doi.org/10.18280/ts.370411

[18] Feng, X., Wei, Y., Pan, X., Qiu, L., Ma, Y. (2020). Academic emotion classification and recognition method for large-scale online learning environment—based on a-CNN and LSTM-ATT deep learning pipeline method. International Journal of Environmental Research and Public Health, 17(6): 1941. https://doi.org/10.3390/ijerph17061941

[19] [19] Sharma, M., Khera, S.N., Sharma, P.B. (2019). Applicability of machine learning in the measurement of emotional intelligence. Annals Data Science, 6(1): 179-187. https://doi.org/10.1007/s40745-018-00185-1

[20] Ali, A., Nasir, J. A., Ahmed, M.M., Naeem, S., Anees, M.S. (2020). Machine learning based statistical analysis of emotion recognition using facial expression. RADS Journal of Biological Research and Applied Science, 11(1): 31-38. https://doi.org/10.37962/jbas.v11i1.262

[21] Parada-Cabaleiro, E., Costantini, G., Batliner, A., Schmitt, M., Schuller, Björn W. (2019). Demos: an italian emotional speech corpus: elicitation methods, machine learning, and perception. Language Resources and Evaluation, 54(2): 341-383. https://doi.org/10.1007/s10579-019-09450-y

[22] Blan, O., Moise, G., Petrescu, L., Moldoveanu, A., Moldoveanu, F. (2019). Emotion classification based on biophysical signals and machine learning techniques. Symmetry, 12(1): 21. https://doi.org/10.3390/sym12010021

[23] Christ, N.M., Elhai, J.D., Forbes, C.N., Gratz, K.L., Tull, M.T. (2021). A machine learning approach to modeling ptsd and difficulties in emotion regulation. Psychiatry Research. 297: 113712. https://doi.org/10.1016/j.psychres.2021.113712

[24] Mehta, D., Siddiqui, M.F.H., Javaid, A.Y. (2019). Recognition of emotion intensities using machine learning algorithms: a comparative study. Sensors, 19(8): 1897. https://doi.org/10.3390/s19081897

[25] Fei, D. (2019). Stochastic model for emotion contagion in social networks security based on machine learning. Safety Science, 118: 757-762. https://doi.org/10.3390/s19081897 https://doi.org/10.1016/j.ssci.2019.06.004

[26] Matsufuji, A., Sato-Shimokawara, E., Yamaguchi, T. (2020). Adaptive personalized multiple machine learning architecture for estimating human emotional states. Journal of Advanced Computational Interlligence and Intelligent Informatics, 24(5): 668-675. https://doi.org/10.20965/jaciii.2020.p0668

[27] Xue, J., Chen, J., Hu, R., Chen, C., Zhu, T. (2020). Twitter discussions and emotions about the covid-19 pandemic: machine learning approach. Journal of Medical Internet Research, 22(11): e20550. https://doi.org/10.2196/preprints.20550

[28] Chen, X., Cao, M., Wei, H., Shang, Z., Zhang, L. (2021). Patient emotion recognition in human computer interaction system based on machine learning method and interactive design theory. Journal of Medical Imaging and Health Informatics, 11(2): 307-312. https://doi.org/10.1166/jmihi.2021.3293

[29] Seng, K.P., Ang, L.M., Ooi, C.S. (2016). A combined rule-based & machine learning audio-visual emotion recognition approach. IEEE Transactions on Affective Computing. 9(1): 3-13. https://doi.org/10.1109/taffc.2016.2588488

[30] Soroush, M.Z., Maghooli, K., Setarehdan, S.K., Nasrabadi, A.M. (2018). Emotion classification through nonlinear EEG analysis using machine learning methods. International Clinical Neuroscience Journal, 5(4): 135-149. https://doi.org/10.15171/icnj.2018.26

[31] Martínez-Tejada, L.A., Maruyama, Y., Yoshimura, N., Koike, Y. (2020). Analysis of personality and EEG features in emotion recognition using machine learning techniques to classify arousal and valence labels. Machine Learning and Knowledge Extraction, 2(2): 99-124. https://doi.org/10.3390/make2020007

7   Authors

Yuxin Cui is studying in Beijing Forestry University with a Bachelor's degree in Business English. Her main research interest is Business English. Email: joeyCui0123@163.com

Ran Zhao graduated from Beijing Forestry University with a Master's degree in English translation. She is currently a full-time teacher of English at Beijing Institute of Economics and Management. Her main research interest is English teaching. Email: zhaoran@biem.edu.cn

Sheng Wang received his bachelor's degree in computer science and technology from Beijing Forestry University, in 2019. He is currently pursuing the master's degree in Computer science and technology, with a focus on few-shot learning. His research interests include machine learning, software develop, program design, remote sensing. Email: ws13121192200@bjfu.edu.cn