Development of a Multimedia-Based Psychological Education Assessment System for Higher Education Institutions

Yue Rong, Wuxi Vocational College of Science and Technology, China*

ABSTRACT

Mental health education in colleges and universities has made considerable progress, but the existing assessment model still faces challenges in terms of time overhead and rank indicators. In response, this paper proposes a new psychological education assessment model for colleges and universities, based on multimedia feature extraction techniques. The proposed model utilizes word vectorization with Word2vec and an improved transformer network, incorporating deeply separable convolutions and an LSTM network to establish long-range dependency coding. Experimental results show that the proposed model outperforms traditional feature extraction methods in terms of extraction speed, feature readability, and model efficiency. Furthermore, the study suggests that the use of reinforcement learning can improve the system's ability to capture key concepts and enhance accuracy. The proposed approach has significant implications for improving mental health education in colleges and universities and can be applied to similar professional environments.

KEYWORDS

LSTM, Multimedia Feature Extraction, Psychoeducational Assessment, Transformer

INTRODUCTION

In a broad sense, mental-health education evaluation is the use of scientific methods and means to collect objective information about the work of mental-health education in schools, to measure the achievement of its goals, and to make a realistic evaluation of its effectiveness. At present, the level of mental-health services of students varies among colleges and universities, partially due to the unsoundness of the mental-health education services evaluation system (Zainuddin et al., 2020). Colleges and universities that prioritize mental health education, with assured annual funding that increases steadily, and dedicated facilities such as full-time counseling rooms or psychological interview rooms (Chen et al., 2021). The stable faculty, consisting of both full-time and part-time positions, ensures the provision of elective courses and lectures in schools, individual counseling, and group training, which greatly promote the popularization of mental-health knowledge and scientific

DOI: 10.4018/IJWLTT.340034 *Corresponding Author

judgment and guidance for college students' psychological problems, thus promoting the improvement of students' mental health. The opposite is true for schools where mental-health education is not a priority, and all aspects need further improvement (Lamppu et al., 2021).

The establishment of a mental-health education evaluation index system is necessary for the in-depth development of mental-health education in colleges and universities and can raise people's awareness of the necessity of mental-health education (Wright et al., 2021). Creating a practical assessment index system for objective and reliable evaluation of moral education in colleges and universities can help bridge the gap in moral education and expedite its comprehensive and rapid development by standardizing and enhancing the assessment process. Meanwhile, in the current higher education environment, students' mental-health problems are becoming more and more prominent, affecting students' physical and mental health and academic performance. Therefore, the establishment of a higher education mental-health system is of great significance in safeguarding students' physical and mental health, improving their self-knowledge and self-management ability, easing psychological distress, promoting educational and teaching reform and innovation, and improving the social image and influence of universities.

The establishment of a scientific assessment index system has an important impact on the development of mental-health education activities in higher education. First, assessment indicators can measure and evaluate the effect and quality of educational activities in a more scientific way and play a key role in promoting the standardization and normalization of mental-health education activities. Second, assessment indicators can clarify the objectives and requirements of higher education mental-health education activities, providing guidance and help for their planning, implementation, and evaluation. In addition, the assessment indicators can also promote the rational allocation and utilization of mental-health education resources in higher education, promote the tilting of educational resources to the field of mental health, and enhance students' psychological quality and self-protection ability.

Since its inception in the 1980s, mental health education in China's colleges and universities has undergone several phases, including investigation and advocacy, establishment, exploration and development, and promotion of prosperity. Throughout these stages, it has experienced significant growth and advancement (Poorthuis & Van Dijk, 2021). Its rapid development is inseparable from the following factors (Cheng et al., 2021). First, the mental-health problems of college students are becoming more and more prominent, which points to the importance of mental-health education in colleges and universities, and it is imperative to accelerate its development and improve its content. Second, the quality of education has continuously deepened. China needs high-quality talents, and their cultivation is inseparable from quality education. The important role of quality psychological education is becoming more and more obvious. Therefore, mental-health education is an inevitable requirement for the development of quality education. Third, the relevant department of education administrations pay extreme attention to the development of mental-health education in colleges and universities. During the period from 1990 to 2011, the State Council approved and promulgated more than ten relevant regulations, outlines, opinions, plans, and standards. The successive issuance of these policy documents not only put forward specific requirements for mental-health education work in colleges and universities, but also provided a policy basis for its development (Goldberg & Wagner, 2019). Furthermore, the relentless efforts of mental health education researchers have made a significant contribution. The keywords "mental-health education in colleges and universities" were used to search journal articles published between 1991 and 2011 on Wanfang, China National Knowledge Infrastructure, and there were about 2,039 articles (Betsch et al., 2018). These published academic papers have made the positive results of mental-health education in colleges and universities more widely known, studied, and researched. Moreover, the growing exchange of academic and professional experiences with developed countries and regions worldwide is expected to positively influence the advancement of mental health education in Chinese universities.

Some colleges and universities carry out mental-health education only as a facade whose purpose is to cope with inspection by the relevant authorities at higher levels. At present, there are only a few teachers specializing in mental-health education, and most of them are part-time teachers. The curriculum of mental health education is still imperfect, as it is currently focused solely on the curriculum without addressing issues such as integration with subjects and other related concerns. Of course, there may be some problems that have not yet been discovered, as well as new problems that have emerged with the changing times and circumstances. Therefore, the assessment of mental-health education in colleges and universities is particularly important in order to monitor the improvement of existing problems. The evaluation process may also uncover problems that have not yet been identified for timely reform and improvement.

Regular assessment helps to adjust the direction of the work and promote the development of mental-health education in a timely manner. Research on assessment in higher education has received much attention globally, and countries are committed to conducting high-quality assessment research in order to improve the quality and effectiveness of higher education. However, there are still bottlenecks in research such as indicator selection, methodology improvement, and international cooperation. There is a need for scientifically sound assessment indicators and improved methods, as well as enhanced international exchanges, to improve the accuracy and validity of assessment so as to provide colleges and universities with better quality assurance and directions for improvement.

Cognitive load theory (CLT) postulates that human cognitive abilities are limited and instructional design should consider the amount of mental work required by learners to process information (Sweller et al., 2011). According to CLT, there are three types of cognitive load: internal load (related to the inherent difficulty of materials), external load (related to the way information is presented), and close-relationship load (related to the construction of schemata or psychological representations) (Paas et al., 2003). Multimedia learning theory (MLT) hypothesizes that learners can learn from both words and pictures, not just words, because they can use language and visual channels for information processing (Mayer, 2019). According to MLT, there are several principles for multimedia instructional design that can reduce additional loads and increase closely related loads by optimizing the use of language and visual modes. Interactive multimodal learning environment theory (IMLET) postulates that learners can benefit from interacting with multiple modes of expression, such as text, audio, video, and animation, because they can participate in active cognitive processing and self-regulation strategies, thereby improving learning outcomes (Moreno & Mayer, 2007). According to IMLET, an interactive multimodal learning environment has several characteristics that can cultivate learner engagement and motivation by providing feedback, control options, and adaptive guidance (Mazher et al., 2017). Based on these existing theoretical frameworks, this study aims to develop a new psychological education evaluation model for universities that combines multimedia feature extraction technology with cognitive load measurement methods.

The purpose of this work is to overcome the time cost and ranking indication restrictions of current models by investigating how multimedia feature-extraction technology can be applied to the construction of an evaluation system for college psychology education. The study aims to address the shortcomings of the existing mental-health education assessment model in colleges and universities by proposing a new set of assessment models combining multimedia feature-extraction techniques. The problem associated with the existing model is that it is not competitive under the time overhead and rank indicators. Therefore, the proposed model aims to reduce the time overhead by improving the existing transformer network using deeply separable convolutions and establishing long-range dependency coding to fully exploit the contextual information in the text vector using the long short-term memory (LSTM) network. The main purpose of using LSTM to improve the transformer model is to improve its modeling ability for sequence data and its ability to handle long sequence data.

Although the original transformer model has been very successful in areas such as natural language processing, it suffers from the problem of gradient vanishing or gradient explosion when dealing with long sequence data, which leads to a decrease in model performance. To overcome this problem, recurrent neural network structures such as LSTM can be used to replace the self-attention mechanism in the transformer to improve the model's ability to model sequence data. The objective of the study is to establish a more effective psychological-education assessment system for colleges and universities that can provide a competitive advantage in terms of accuracy, efficiency, and scalability. The multimedia-based psychoeducational assessment system is currently an innovative tool in mental-health education in higher education institutions. The assessment system provides intuitive and vivid assessment through various media forms such as images, video, and audio, which increases the objectivity and accuracy of the assessment results. At the same time, the system is also able to provide personalized assessment content and feedback to meet the psychological characteristics and needs of different students, providing strong support for students' mental health. Research indicates that a multimedia-based assessment system offers a scientific and effective approach to evaluating moral education in higher education institutions. Therefore, the development of a multimedia-based psychoeducational assessment system is highly significant in advancing the overall progress of moral education in higher learning institutions. The cultivation of high-quality talents required for the development of society in the new era cannot be separated from college mental-health education, and the development of mental-health education cannot be separated from the work of mental-health education assessment. Thus, studying the assessment of mental-health education in colleges and universities is also necessary to cultivate the high-quality talents needed by society.

The contribution of this thesis is mainly in the development of a multimedia-based psychoeducational assessment system. It provides a comprehensive and intuitive assessment method for higher education institutions. Through the system, educators can accurately assess students' psychological states, learning abilities, and development, providing a scientific basis for the development of individualized educational programs and support measures. This innovative assessment tool and method helps to enhance students' mental health, development potential, and learning outcomes and promotes the overall development of higher education and the implementation of quality teaching.

RELATED WORKS

Current Status of Educational Evaluation Research in Higher Education

Assessments of mental health education in colleges and universities are constantly evolving in response to changes in people's perceptions of the world and objective realities. Assessments of mental-health education are directly impacted by societal shifts, advances in science and technology, and ongoing improvements to national education regulations. Therefore, it is possible to more successfully carry out the practical activities of mental-health education assessment and increase the effectiveness of mental-health education by starting from reality and enhancing and upgrading assessment theories with a developmental perspective.

In recent years, mental-health education for college students has been carried out in depth, and foreign experts and scholars are extremely concerned about the design and development of psychological education software systems for college students in China. Researchers have conducted focused research on web-based mental-health assessment systems at different angles and levels (Van Herpen et al., 2020). The current mental-health assessment system for college students is mainly for file management, and the first stand-alone mental-health system developed in China is applied mainly in personnel file management and document management (Zawacki-Richter et al., 2019). By the 1990s, with the need for mental-health education in colleges and universities, researchers began to manage psychological profiles through a pair of mental-health systems. The development of the stand-alone version of the mental-health management system of Beijing University of Technology has led many researchers to embark on the exploration of the psychological profile system, followed by many domestic scholars and research experts to study and practice it. In addition, China's renowned Tsinghua University, Peking University, Zhejiang University, and Fudan University are all committed to improving the mental health of their students. By establishing mental health education and counseling centers, offering psychology majors, and organizing mental health weeks, these universities provide students with a range of services including personal counseling, psychological assessments, mental wellness training, and thematic lectures. These initiatives not only help students to comprehensively recognize and cope with psychological problems but also improve their mental-health awareness and skills. The continuous development of university psychological education not only helps students' personal growth and healthy development, but also creates a more positive and healthy learning and living environment for the university.

The system contains mainly the psychological building scale, the classification number, the selection of entries, the mean standard deviation, and other aspects, and the capacity statistics are better, but the system is more complex and complicated to use, and it is difficult for ordinary personnel to operate and master.

In 2004, China also developed the golden sea psychological profile system, which is more widely used (Aldowah et al., 2019). This system has more than 30 kinds of measurement scales, which is a great improvement on the operation of mental-health rating scales and occupational tests, but the statistical function is not good, and it can only do the mean and tolerate the standard deviation routine statistics.

Currently, scholars have not provided a very precise definition of the concept of mental-health education assessment in colleges and universities (Al-Balas et al., 2020). Through extensive research, the following understandings were synthesized. First, the purpose of mental-health education assessment in colleges and universities is to develop supervised mental-health education. Therefore, it should address the entire dynamic process of mental-health education development activities, such as whether the assessment object has undergone beneficial changes, whether the assessment process has been smooth, whether there are areas for improvement, and whether the assessment results have reached the expected goals. Second, the purpose of mental-health education evaluation in colleges and universities is to ensure that mental-health education activities are carried out in a normal and orderly manner and ultimately promote the healthy physical and mental development of college students. In mental-health education assessment, it is important both to focus on the process and, more importantly, to emphasize development (Carrillo & Flores, 2020). Therefore, mental-health education assessment activities should be carried out according to certain criteria, and a specific and operable assessment index system should be established. Again, the process of mental-health education evaluation in colleges and universities is to collect the real information of the assessment subjects and to use scientific and effective technical means to process and handle the information.In evaluating mental health education in colleges and universities, it is crucial not only to assess the implicit effects of educational activities and the expected goal achievement, but also to consider and anticipate additional benefits and impacts beyond these. Because there are also many uncertainties in the educational process and we may face unexpected events at any time, we cannot ignore the judgment of some potential values. Based on the above understanding, in this paper, mental-health education evaluation in colleges and universities is defined as the process of making value judgments on the process and results of the development and changes of mental-health education activities according to the criteria of mental-health education work by using a scientific evaluation index system.

Current Status of Multimedia Feature-Extraction Research

Traditional Image Feature Extraction

Traditional multimedia feature-extraction algorithms are based on information such as image grayscale mean, texture, denseness, etc. as extraction targets, among which the more mainstream methods are HOG and SIFT (Mo & Sun, 2020). HOG is the most commonly used of the traditional algorithms for extracting local texture features. The first step of the main process is image segmentation. Since the HOG extracts the local features of the image, it is not ideal for the whole image, so the image needs to be segmented first. Non-overlap or overlap segmentation can be used. The first method refers to the segmented image blocks without overlapping images between them, while in the second method image blocks have overlap with each other (Qin et al., 2021). Then the HOG features of each image block after segmentation, the amplitude, and the orientation are calculated (Zhao et al., 2019). The HOG of each image block is statistically calculated, and the corresponding feature information is obtained. Finally, a sliding window of 3x3 is used for each image block to calculate the features of each pixel point, and then they are formed into the overall HOG features.

SIFT is more stable compared to HOG and other traditional image-processing methods. The SIFT operator is scale-invariant; i.e., when the luminance or offset rotation angle of the sample changes, it can still perform feature extraction well, so SIFT is more valued in computer vision. The approximate flow of the algorithm is as follows: the scale space is first approximated by Gaussian fuzzy functions with different parameters to form a Gaussian differential scale space (Ng & Chan, 2019). By obtaining features that exist under different spaces in the DoG (Difference of Gaussians) space, the feature's unaltered character is satisfied. The commonly used method for extracting features is the Laplacian of Gaussian algorithm. However, it has the disadvantage of being computationally intensive, so it is often replaced with a Gaussian difference algorithm. The extracted feature pixels are used as the center of a circle, the gradient histogram within a certain radius is counted, and the directions of the feature points and the extreme value points are regarded as the same. The pixels in the 16x16 range with the point as the center of the circle are planned into 16 regions according to the feature direction as the coordinate direction, and the magnitude and histogram in eight different directions are calculated for each part to obtain the 128-bit feature vector.

Multimedia-Based Feature Extraction

Multimedia-based feature extraction refers to the process of learning, extracting, and representing image information features using convolutional network models (He et al., 2019). Convolutional neural networks learn the underlying feature representations by mimicking the cognitive process of neural networks in the human brain, extracting them from raw data and combining them into high-level semantic information by nonlinear or linear computation. Convolutional neural networks (CNNs) have many components, such as the number of neurons, network depth, convolution size, step size, and high or low learning rate, and these components require a large number of hyperparameters for learning. There are many convolutional networks that are widely used in various tasks in computer vision, which are described below for typical networks.

Although LeNet is the first deep convolutional network in computer vision, it can be used only for handwritten digit discrimination and cannot be applied to other image category recognition studies, which has great limitations. AlexNet was proposed due to its outstanding achievements in image recognition and image classification research and the ability to improve the effectiveness and performance of feature extraction from network models by appropriately increasing the depth of the convolutional network to learn more hyperparameters (Geng, 2021). In LeNet, only five convolutional layers are used for vector acquisition, while in AlexNet, two additional layers are added, making it easy to apply to a variety of different classes of image recognition and classification tasks (Khan et al., 2022). Although increasing the network depth can improve the generalizability of feature extraction under different resolution images, it tends to cause learning overfitting. To solve this problem, Abu-

Arqoub et al. (2020) use Hinton's idea of randomly skipping certain units during model training, which makes the model extract more robust features by forcing means. In addition, since ReLU solves part of the gradient vanishing phenomenon, using it as the activation function can accelerate the model convergence (Aslam & Curry, 2021).

Kumar et al. (2022) proposed the VGG network, which allows the number of layers of the network to be further deepened according to the modular layering principle. VGG replaced the 5x5 and 11x11 filters with small-size (3x3) filters and further demonstrated that using multiple small-size filters at the same time can generate the same perceptual field as the large size. The smaller the filter size, the fewer the parameters and the lower the computational complexity. VGG extracts a linear combination of features, adds a maximum pooling layer after the convolution layer, and maintains the same spatial resolution by filling pixels. Its improvement is evident in both image recognition and classification tasks and localization problems. However, the network still suffers from high computational cost. Although the complexity is reduced by using a small-size filter in the feature-extraction process, it still needs to learn 140 million hyperparameters.

As CNNs have risen in computer vision, more networks have been proposed for research, including the GoogLeNet network with more convolutional layers. Instead of simply stacking multiple convolutional layers together, it constructs the whole network by using the Inception V1 structure as a base, so although it is 22 layers deep, it has much fewer parameters (Islam et al., 2021). In this structure, convolutional modules of different sizes and pooling operations are used to run in parallel, and several 1x1 convolutional blocks are added in different branches. By using the Inception V1 module, the number of parameters and dimensions of the network are effectively reduced, redundant filters are reduced, and the controllable size of the network is also ensured. In the Inception V2 module, the double convolution strategy is applied to each branch to effectively reduce the complexity of the network. This is demonstrated by the fact that the module replaces the 3x3 convolution with the sum of two convolutions of 3x1 and 1x3. In the other branches, the 5x5 convolution is replaced by two serial 3x3 convolution operations, once again reducing the number of parameters in the model. Thanks to the innovative results of these two modules, GoogLeNet has been widely used in various fields of computer vision.

Even though the aforementioned models have been effectively used to extract features from text or picture data, they are weak at collecting the important information and only extract the spatial dimension. In order to overcome this, this research presents an altered transformer model that incorporates the temporal dimension and is built on an LSTM network. It is then applied to the psychoeducational evaluation system used in colleges and universities.

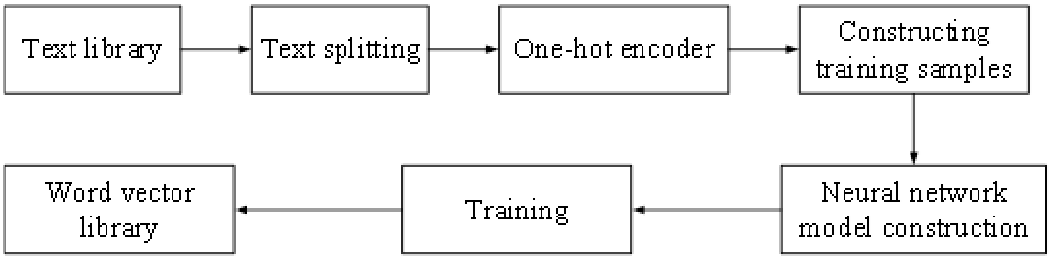
METHODOLOGY

Word2vec Word Vector Library Construction

For text-feature extraction in the field of psychological education in colleges and universities, a vector library needs to be trained using texts from the field, and the resulting word vectors are more relevant to the problem domain. The specific word vector library construction process is shown in Fig. 1.

With one-hot encoding, a matrix of lexicon dimensions is created for each word unit, which creates too much useless information when the number of words is too large. The one-hot coding vector is expressed using a sparse matrix, resulting in interword discretization, making it difficult for the model to learn useful features, and the sparse matrix is also unable to represent the semantic relationships between words. To address the issues arising from one-hot, this research used word2vec to vectorize the keywords in the university psychoeducation system. Word embeddings learned through word2vec have proven successful on a variety of downstream natural language processing tasks, and word2vec is commonly used with CBOW and skip-gram.

Figure 1. Word Vector Library Construction

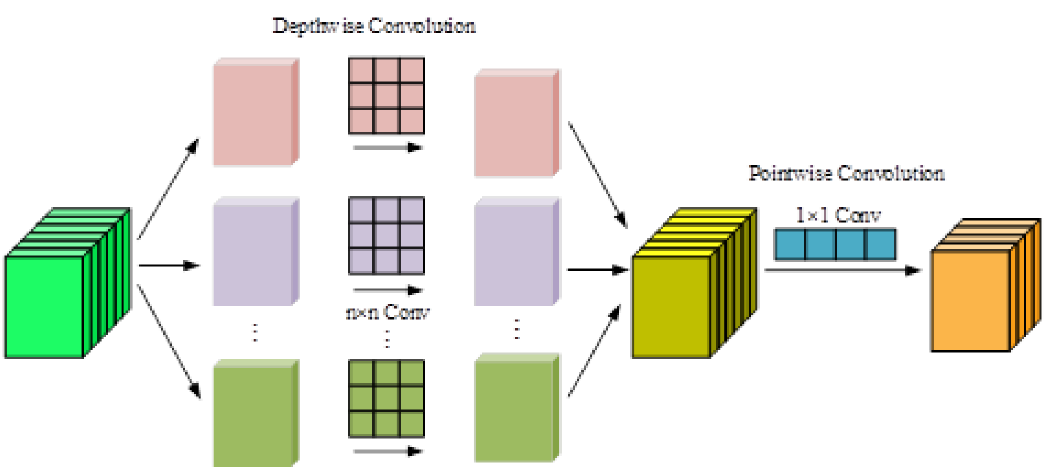


Feature Extraction

A residual network is a deep convolutional neural network architecture that is widely used for computer vision tasks. Its advantages include solving the gradient vanishing and gradient explosion problems, improving network expressiveness, allowing deeper network structures, and improving convergence speed and generalization. As a result, residual networks perform well in tasks such as image classification, target detection, and semantic segmentation, making them an important model architecture in the research field. The rich and diverse text and image data in the psychological assessment system of colleges and universities, and the extraction of richer word-level or sentence-level features by continuously stacking the number of network layers, are very likely to lead to gradient disappearance, performance degradation, and gradient dispersion in the model. Therefore, this paper utilizes a deep residual learning module to fit the residual mapping by stacking layers. In addition, the increase in the number of network layers leads to an increase in the computational complexity of the model, and the ResNeXt network can solve the problem better. Also, the ResNeXt network has the ability to handle contextual timing relationships. In addition, this paper takes into account the high-performance requirements of deep networks for hardware devices and uses depthwise separable convolution instead of convolution operation to improve the original ResNeXt and speed up the model inference by reducing the number of parameters. The depth-separable convolutional structure is shown in Fig. 2.

In Fig. 2, the psychologically relevant text vectors captured in the psychoeducational assessment system are first subjected to a convolution operation to generate M feature maps. Next, multiple

Figure 2. Deeply Separable Convolutional Neural Network Structure



convolutional blocks of convolutional kernel size $k \times k$ are used to combine the feature maps in a weighted manner along the depth direction to obtain a feature map of the input character vector in the depth space. Then convolutional filtering is performed using a 1×1 convolutional kernel during the point-by-point convolutional operation. Thus, the depth-separable convolution can be seen as a combination of depth convolution and point-by-point convolution, and the validation of this part is shown in the ablation experiments.

Psychologically relevant texts captured in psychoeducational assessment systems contain information limitations due to their varying lengths, and multiscale features are applicable to the problem of feature limitations. Therefore, in this paper, ResNet and ResNeXt networks are selected as the backbone networks to extract word-level features of psychologically relevant texts and to enhance the characterization ability of features using multi-scale features, with the structure shown in Fig. 3. The calculations are shown in (1), (2), and (3).

$$R_1(str_i) = f_{resnet}(str_i) \tag{1}$$
$$R_2(str_i) = f_{resnext}(str_i) \tag{2}$$
$$R(str_i) = R_1(str_i) + R_2(str_i) \tag{3}$$

$R_1(str_i)$ denotes the text features extracted by ResNet, $R_2(str_i)$ denotes the text features extracted by ResNeXt, and f denotes the feature-extraction function.

Feature Encoder

In recent years, transformer has been widely used in the fields of natural language processing and computer vision and has made breakthroughs. Therefore, this paper attempts to use transformer as an encoder for the depth features of psychologically relevant texts captured in the psychoeducational assessment system, with the structure shown in Fig. 4.

The transformer network consists of multiple identical stacked layers, with each encoder layer consisting of a multi-head attention mechanism and a feed-forward network connected to a residual network. Among them, the attention mechanism is used to strengthen the decision-making ability of keywords or key characters, which can be expressed as in (4).

Figure 3. Backbone Network Structure

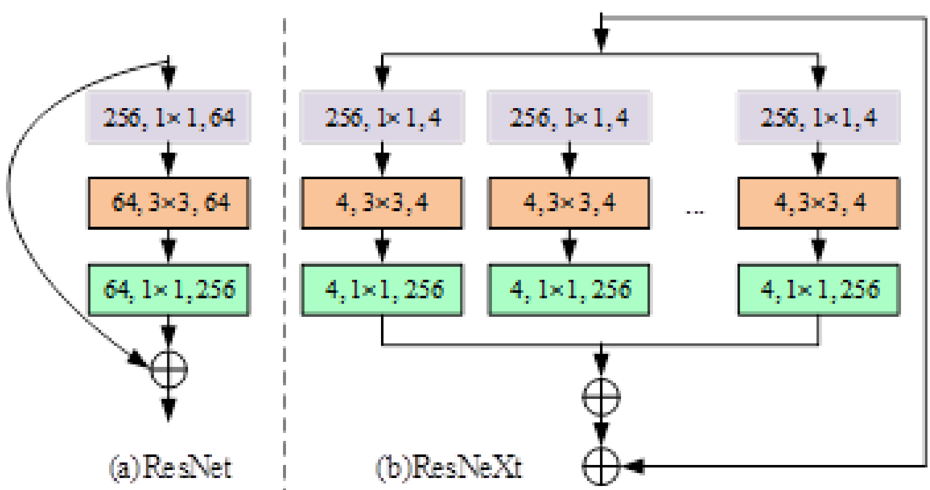
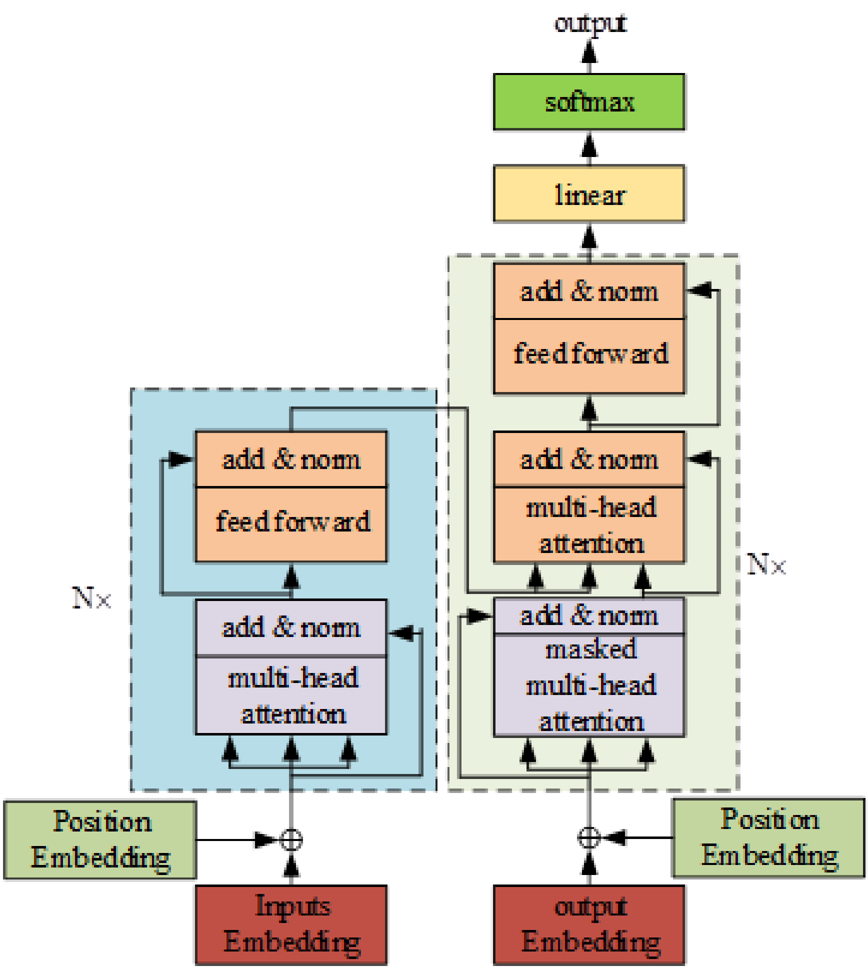


Figure 4. Transformer Model



$$attention < \boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V} >= soft\max(\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}})\boldsymbol{V} \tag{4}$$

Q, K, and *V* denote the input matrices of query, key, and value, respectively. Here, the multi-headed attention mechanism is introduced to enable the model to simultaneously focus on the depth-feature information of the extracted feature *A* at different spatial locations. By mapping the inputs *Q*, *K,* and *V* several times, the resultant values of multiple pairs of attention are obtained, expressed in (5).

$$multi-head(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = concat(head_1, \ldots, head_n)\boldsymbol{W}' \tag{5}$$

$head_i = attention(\boldsymbol{Q}\boldsymbol{W}_i^{\boldsymbol{Q}}, \boldsymbol{K}\boldsymbol{W}_i^{\boldsymbol{K}}, \boldsymbol{V}\boldsymbol{W}_i^{\boldsymbol{V}})$ denotes the output of the *i*th head, *n* is the total number of heads, and $\boldsymbol{W}_i^{\boldsymbol{Q}}$, $\boldsymbol{W}_i^{\boldsymbol{K}}$, $\boldsymbol{W}_i^{\boldsymbol{V}}$ and ***W'*** are the parameter matrices, all obtained by model learning.

In addition, although transformer network provides location encoding, only relative location encoding information is considered, which is not related to the context of psychologically relevant text, making the encoding features unable to combine the contextual information of the text depth vector and relative location encoding information. Therefore, in this paper, the transformer is first improved by drawing on the long- and short-term memory neural network LSTM; the improved model contains both sentence-level and word-level contextual information and the relative position encoding information provided by transformer. Second, inspired by the residual network, the features extracted by the backbone networks ResNet and ResNeXt are given different proportional weights and are computed as dot product with the transformer output values introduced into the LSTM to further enhance the importance of character-level and word-level features. Finally, the improved model is used to encode the psychoeducational text with deep features, and the softmax layer is used for classification detection to quickly give the determination results of the psychoeducational text to be tested.

Intensive Learning Optimization

To further improve the accuracy of the model in locating and identifying key information of psychologically relevant texts captured in the psychoeducational assessment system, reinforcement learning of strategy gradients was introduced. The baseline model is used as the agent in reinforcement learning, and the sequence of psychoeducational text subject words is used as environment. For each time step of the model, the agent corresponds to predict an outcome, and based on the prediction accuracy, the reward value is calculated and fed back to the agent.

When the model is optimized using reinforcement learning, the sequence consisting of the state values, outcomes, and rewards predicted by the psychoeducational text string at moment t are set $\tau_t = \{s_1, a_1, r_1, \cdots, s_t, a_t, r_t\}$, where s_t denotes the state of the environment at moment t, a_t denotes the predicted outcome at time t, and r_t denotes the reward value. Finally, the gradient of the loss function is calculated by using the loss and reward values, and the overall optimization of the model parameters based on the gradient values is calculated as shown in (6) and (7).

$$L(\theta) = -\frac{1}{N}\sum_{\tau_t} R(\tau_t)\log \pi_\theta(\tau_t) = -E_{\tau_t \sim \pi_\theta}[R(\tau_t)] \tag{6}$$

$$\nabla_\theta L(\theta) = -E_{\tau_t \sim \pi_\theta}[R(\tau_t)\cdot \nabla \log \pi_\theta(\tau_t)] \tag{7}$$

$\pi_\theta(\tau_t)$ is the predicted probability value of the model.

EXPERIMENT

Experimental Environment and Evaluation Index

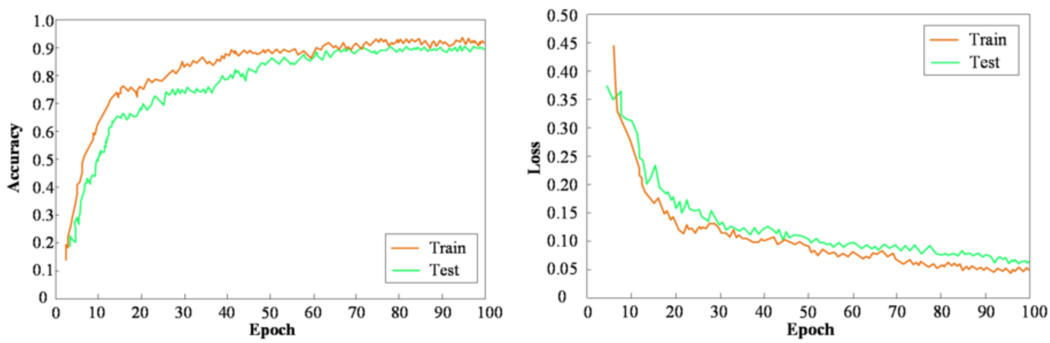
The test environment is a 64-bit Lenovo ThinkPad P15 (8-core 16-thread i9-10885H, 1TB SSD), the model uses the PyTorch deep-learning framework, the development environment is Anaconda 3.5.7, Python Version 3.7, and the graphics card is Nvidia GTX3060Ti. The optimizer uses Adam with an initial learning rate of 0.0001. In addition, dropout=0.5 and batch_size=128 are used to prevent the model from overfitting.

The correspondence curves between the training accuracy, loss, and number of iterations epoch of the model in this paper are given in Fig. 5. It can be seen that the accuracy- and loss-curve regions are smooth in the training and testing phases when epoch is 90; therefore, epoch is set to 90.

Experimental Structure and Analysis

To validate the efficacy of feature extraction in this study, three distinct informational feature extraction methods, as illustrated in Figure 6, were independently constructed based on relevant literature (method

Figure 5. Training and Testing Curve: (a) Accuracy Curve, (b) Loss Curve



1) and additional literature (method 2), the variation between the number of iterations and the extraction speed of the three different methods can be seen. When the number of iterations is 20, the extraction speed of traditional feature-extraction method 1 is 20/time and the feature extraction speed is smaller; the extraction speed of traditional feature-extraction method 2 is 28/time and the feature extraction speed is faster. The extraction speed of this paper's feature-extraction method is around 32 features/time, and its actual extraction speed is the fastest compared with the two traditional feature-extraction methods. The reason for this is that this paper, based on the original feature-extraction network, uses depth-separable convolution instead of traditional convolutional blocks to reduce the time overhead by reducing the model parameters, and the above data also verify the effectiveness of the model.

After controlling the three feature-extraction methods to process the same content dataset simultaneously, the average run time of traditional feature-extraction method 1 is around 95.6 ms, according to the run-time results in Fig. 7. Under the same experimental conditions, the three feature extraction methods were utilized to discretely process numerical data, and the runtime was measured and recorded for each method. The results of the runtime analysis are presented in Figure 7. The run time required for feature extraction is the longest. The average run time of traditional feature-extraction method 2 is around 50.9 ms. The average run time of the feature-extraction method designed in this paper is around 27.4 ms, which is the shortest run time required compared with the two traditional informative feature-extraction methods. The two points below are among the primary causes: 1) The high time overhead issue is significantly reduced when the computational complexity of the convolutional layers is substantially reduced by the use of deeply separable convolution; 2). The conventional transformer network is improved through the integration of the LSTM network, enabling enhanced long-range modeling and the capture of long-range dependencies in text data within university psychoeducational evaluation systems.

The rank values generated by the three feature-extraction methods were compared under the same data in the above experimental setting, and the detailed results are shown in Fig. 8. It can be seen that the rank values obtained by traditional feature-extraction method 1 are between 20 and 30, the rank values are small, and the readability of the extracted feature data is poor. The rank values obtained by traditional feature-extraction method 2 are between 35 and 40, the actual rank values are larger, and the actual feature data are more readable. The feature-extraction method designed in this paper obtains rank values between 50 and 55, and the feature data extracted by this method are more readable compared with those of the two traditional feature-extraction methods.

The time to train a round on the training set, the number of parameters, and the time to detect the test set were compared for each model. The thop tool in Python was used to calculate the number of parameters and floating-point operations of the model, and the time tool was used to calculate the running time of the model. The experimental results are shown in Fig. 9.

Figure 6. Comparison of Extraction Speed of Different Feature-Extraction Methods

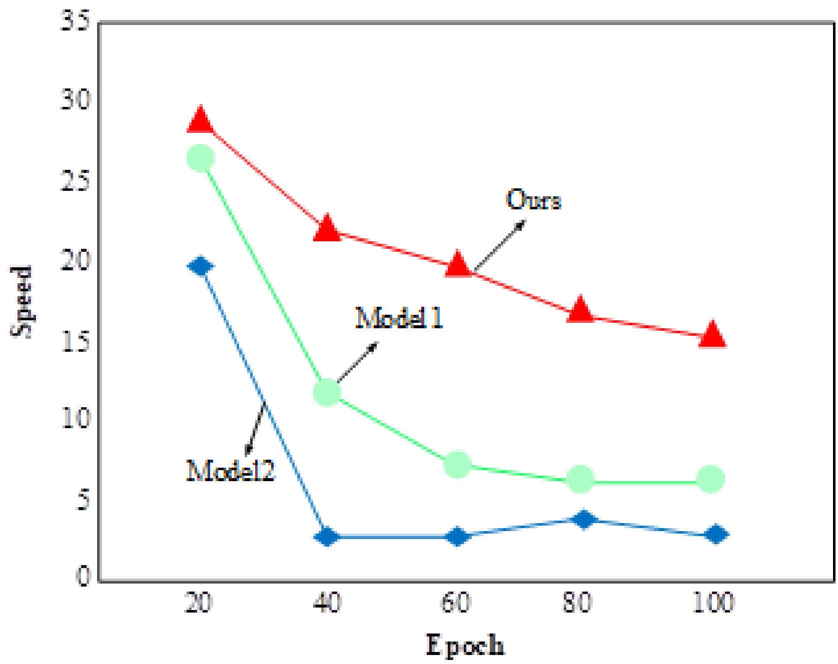
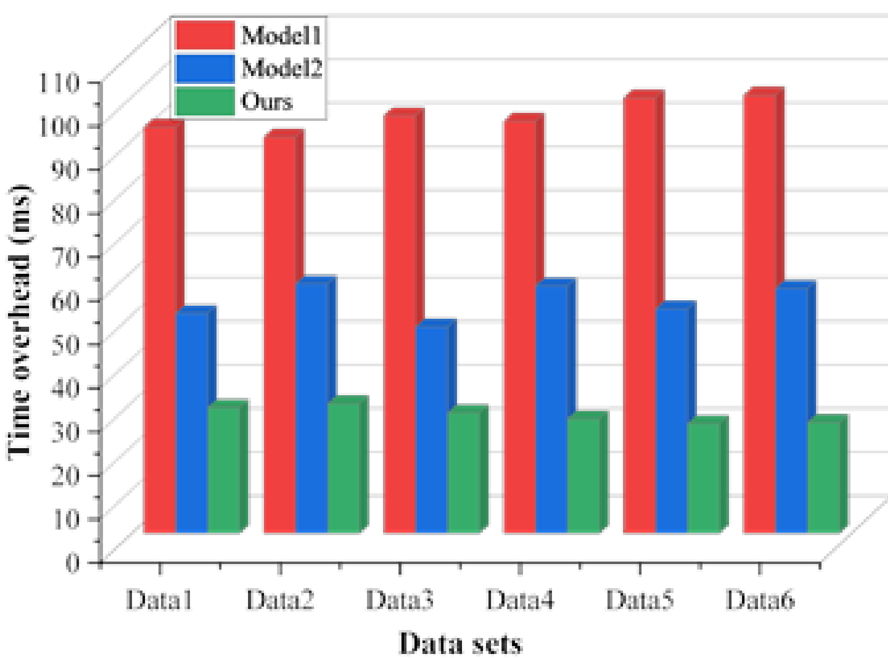
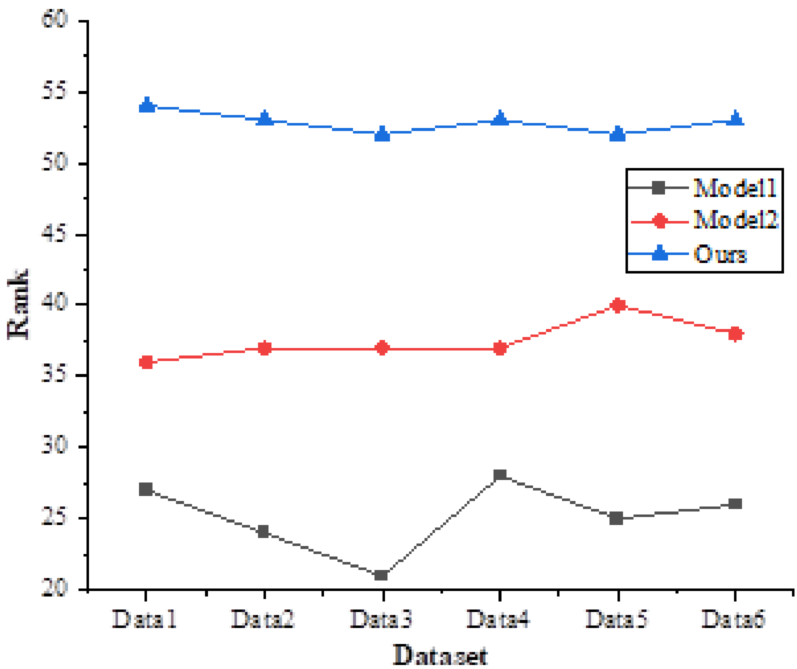


Figure 7. Comparison of Run Time of Different Feature-Extraction Methods



The advantages of the model in this paper in terms of training time, testing time, and number of model parameters are obvious. The primary reason is that, rather than using conventional convolutional

Figure 8. Rank Comparison of Different Feature-Extraction Methods



layers, this method uses depth-separable convolution. As a result, the attention mechanism included with the transformer fully utilizes the contextual encoding information of the text data, reducing overhead time and enhancing extraction quality. In addition, to explore the effect of reinforcement learning on the overall performance of the model, the following ablation experiments were conducted and the results are shown in Fig. 10. It can be seen that the use of reinforcement learning can effectively improve the ability for and accuracy of keyword capture in the university psychoeducational assessment system under the scenario where the experimental conditions and data are kept consistent. This is made possible by the fact that reinforcement learning constantly adapts and adjusts to the perceptual environment. Second, reinforcement learning has a strong in-range perception, which applies to the ability to capture keywords over long distances in the university psychoeducational assessment system, and the above data also verify the effectiveness of introducing reinforcement learning.

Reinforcement learning offers several advantages in the field of authorship modeling. First, reinforcement learning enables the model to continuously learn and improve the quality and accuracy of answering questions through automatic decision-making and optimization. Second, reinforcement learning is adaptive and flexible and can autonomously adjust its strategy when facing different inputs and scenarios to better meet user needs. In addition, reinforcement learning learns through interactions with users; by collecting user feedback and evaluations, it continuously improves the model and the ability to answer questions. Most importantly, reinforcement learning supports continuous learning and incremental improvement, allowing the model to be continuously optimized during use to adapt to user needs and changes.

CONCLUSION

Although China's college mental-health education started late, its practical work and related theoretical research have developed rapidly. However, the development of mental-health education assessment

Figure 9. Comparison of Time Complexity and Number of Parameters

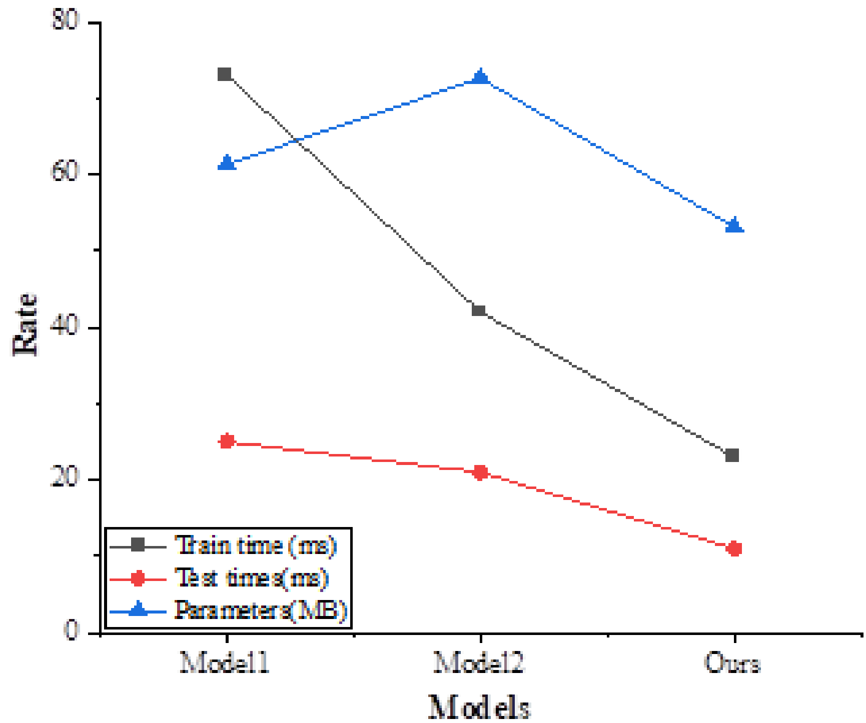
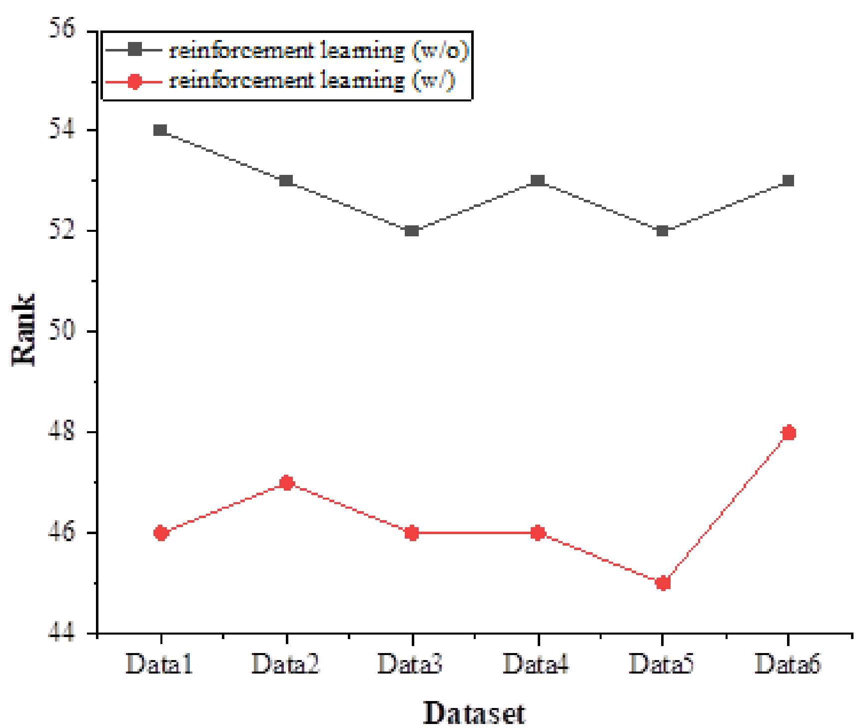


Figure 10. Intensive Learning Ablation Experiment



in higher education in China is slow, both in terms of the development of its practical work and the development of related research. The assessment of mental health education in colleges and universities serves as a guiding guarantee for achieving the set goals of mental health education. It is the basis for objective evaluation and analysis of mental-health education; it is the compass for colleges and universities to reform and adjust unsuitable mental-health education course contents, activities, or other work related to mental-health education in a timely manner. Therefore, researching mental-health education assessment, enriching and improving the theories related to the assessment of mental-health education, and promoting the scientific development of mental-health education will further promote the development of mental-health education in colleges and universities and have great theoretical and practical significance for cultivating the high-quality talents required by society in the new century. The proposed psychological-education assessment system has not been tested in real-world settings. Therefore, future research should focus on the implementation and evaluation of the system in actual educational contexts, which could provide valuable insights into its effectiveness and practicality. Additionally, this study focuses mainly on the technical aspects of the proposed system. Future research should consider incorporating theoretical perspectives and practical strategies from psychology and education to further enhance the system's impact on mental-health education in higher education institutions.

AUTHOR NOTE

The figures used to support the findings of this study are included in the article.

The authors declare that they have no conflicts of interest.

This work was not supported by any funds.

The authors sincerely thank those who have contributed to this research.

REFERENCES

Abu-Arqoub, M., Issa, G., Banna, A. E., & Saadeh, H. (2020). Interactive multimedia-based educational system for children using interactive book with augmented reality. *Journal of Computational Science*, *15*(11), 1648–1658. doi:10.3844/jcssp.2019.1648.1658

Al-Balas, M., Al-Balas, H. I., Jaber, H. M., Obeidat, K., Al-Balas, H., Aborajooh, E. A., Al-Taher, R., & Al-Balas, B. (2020). Distance learning in clinical medical education amid COVID-19 pandemic in Jordan: Current situation, challenges, and perspectives. *BMC Medical Education*, *20*(1), 341. doi:10.1186/s12909-020-02257-4

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, *37*, 13–49. doi:10.1016/j.tele.2019.01.007

Aslam, A., & Curry, E. (2021). A survey on object detection for the internet of multimedia things (IoMT) using deep learning and event-based middleware: Approaches, challenges, and future directions. *Image and Vision Computing*, *106*, 104095. doi:10.1016/j.imavis.2020.104095

Betsch, C., Schmid, P., Heinemeier, D., Korn, L., Holtmann, C., & Böhm, R. (2018). Beyond confidence: Development of a measure assessing the 5C psychological antecedents of vaccination. *PLoS One*, *13*(12), e0208601. doi:10.1371/journal.pone.0208601

Carrillo, C., & Flores, M. A. (2020). COVID-19 and teacher education: A literature review of online teaching and learning practices. *European Journal of Teacher Education*, *43*(4), 466–487. doi:10.1080/02619768.2020.1821184

Chen, D., Kong, X., & Wei, Q. (2021). Design and development of psychological virtual simulation experiment teaching system. *Computer Applications in Engineering Education*, *29*(2), 481–490. doi:10.1002/cae.22293

Cheng, A. S., Kang, Y. Q., Lim, T. S. H., & Ragen, E. S. (2021). Minimizing risk of infection when using developmental and psychological assessment tools: Looking beyond COVID-19. *The Nurse Practitioner*, *46*(6), 12–16. doi:10.1097/01.NPR.0000751816.41679.09

Geng, L. (2021). Evaluation model of college English multimedia teaching effect based on deep convolutional neural networks. *Mobile Information Systems*, *1874584*, 1–8. Advance online publication. doi:10.1155/2021/1874584

Goldberg, S. G., & Wagner, K. (2019). American Psychological Association practice guidelines for psychopharmacology: Ethical practice considerations for psychologists involving psychotropic use with children and adolescents. *Journal of Clinical Psychology*, *75*(3), 344–363. doi:10.1002/jclp.22705

He, Q., He, B., Zhang, Y., & Fang, H. (2019). Multimedia based fast face recognition algorithm of speed up robust features. *Multimedia Tools and Applications*, *78*(17), 24035–24045. doi:10.1007/s11042-019-7209-0

Islam, N., Altamimi, M., Haseeb, K., & Siraj, M. (2021). Secure and sustainable predictive framework for IoT-based multimedia services using machine learning. *Sustainability (Basel)*, *13*(23), 13128. doi:10.3390/su132313128

Khan, A. A., Shaikh, A. A., Cheikhrouhou, O., Laghari, A. A., Rashid, M., Shafiq, M., & Hamam, H. (2022). IMG-forensics: Multimedia-enabled information hiding investigation using convolutional neural network. *IET Image Processing*, *16*(11), 2854–2862. doi:10.1049/ipr2.12272

Kumar, V., Dhingra, G., Saxena, N., & Malhotra, R. (2022). Machine learning based analysis of learner-centric teaching of Punjabi grammar with multimedia tools in rural Indian environment. *Multimedia Tools and Applications*, *81*(28), 40775–40792. doi:10.1007/s11042-022-12898-w

Lamppu, P. J., Laakkonen, M. L., Finne-Soveri, H., Kautiainen, H., Laurila, J. V., & Pitkälä, K. H. (2021). Training staff in long-term care facilities—Effects on residents' symptoms, psychological well-being, and proxy satisfaction. *Journal of Pain and Symptom Management*, *62*(4), e4–e12. doi:10.1016/j.jpainsymman.2021.03.020

Mayer, R. E. (2019). *Multimedia learning* (3rd ed.). Cambridge University Press.

Mazher, M., Aziz, A. A., Malik, A. S., & Amin, H. U. (2017). An EEG-based cognitive load assessment in multimedia learning using feature extraction and partial directed coherence. *IEEE Access : Practical Innovations, Open Solutions*, *5*, 14819–14829. doi:10.1109/ACCESS.2017.2731784

Mo, C., & Sun, W. (2020). Point-by-point feature extraction of artificial intelligence images based on the Internet of Things. *Computer Communications*, *159*, 1–8. doi:10.1016/j.comcom.2020.05.015

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*(3), 309–326. doi:10.1007/s10648-007-9047-2

Ng, S., & Chan, M. (2019). Effect of neighbourhood size selection in SOM-based image feature extraction. *International Journal of Machine Learning and Computing*, *9*(2), 195–200. doi:10.18178/ijmlc.2019.9.2.786

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. doi:10.1207/S15326985EP3801_1

Poorthuis, A. M. G., & Van Dijk, A. (2021). Online study-aids to stimulate effective learning in an undergraduate psychological assessment course. *Psychology Learning & Teaching*, *20*(2), 236–249. doi:10.1177/1475725720964761

Qin, P., Cai, Y., Liu, J., Fan, P., & Sun, M. (2021). Multilayer feature extraction network for military ship detection from high-resolution optical remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 11058–11069. doi:10.1109/JSTARS.2021.3123080

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory* (Vol. 1). Springer Science & Business Media. doi:10.1007/978-1-4419-8126-4

Van Herpen, S. G., Meeuwisse, M., Hofman, W. H. A., & Severiens, S. E. (2020). A head start in higher education: The effect of a transition intervention on interaction, sense of belonging, and academic performance. *Studies in Higher Education*, *45*(4), 862–877. doi:10.1080/03075079.2019.1572088

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, *89*, 98–110. doi:10.1016/j.chb.2018.07.027

Wright, A. J., Chávez, L., Edelstein, B. A., Grus, C. L., Krishnamurthy, R., Lieb, R., Mihura, J. L., Pincus, A. L., & Wilson, M. (2021). Education and training guidelines for psychological assessment in health service psychology. *The American Psychologist*, *76*(5), 794–801. doi:10.1037/amp0000742

Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education*, *145*, 103729. doi:10.1016/j.compedu.2019.103729

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. doi:10.1186/s41239-019-0171-0

Zhao, X., Wei, H., Wang, H., Zhu, T., & Zhang, K. (2019). 3D-CNN-based feature extraction of ground-based cloud images for direct normal irradiance prediction. *Solar Energy*, *181*, 510–518. doi:10.1016/j.solener.2019.01.096

Yue Rong was born in Jiangsu, China, in 1982. From 2001 to 2005, she studied in Jiangnan University and received her bachelor's degree in 2005. After graduation, she has been working in Wuxi Vocational College of Science and Technology. From 2008 to 2010, she studied in Suzhou University making use of weekdays and received her Master's degree in 2010. She has published more than ten papers. Her research interests are included ideological and political education for college students. Email: 3101080@wxsc.edu.cn.