Research Article

Xiaoran Li*

Application of emotion recognition technology in psychological counseling for college students

Abstract: In traditional psychological counseling, limitations of patients' subjective feelings and expressive abilities, as well as the varying levels of professional skills and conversation skills of counselors, make it difficult to properly solve psychological problems. Emotional recognition technology can help psychological counselors better understand the inner world of patients, thereby providing more targeted psychological counseling services for college students. 2D convolutional neural network-long short-term memory (2DCNN-LSTM) and 1D convolutional neural network-long short-term memory network (1DCNN-LSTM) were used for emotion recognition. Wavelet packet coefficient sequences were utilized to analyze multimodal data; multi-core convolutional neural network models were applied for multimodal feature layer fusion, and LSTM recurrent neural networks were used to effectively identify valence and college student emotions. The experimental results showed that under multimodal recognition, the average recognition accuracy of 2DCNN-LSTM for fourth-year university students was 18.7% higher than that of 1DCNN-LSTM for fourth-year university students. Compared with 1DCNN-LSTM, 2DCNN-LSTM can achieve better recognition accuracy. The research results can help psychological counselors better understand the emotional state and needs of students and provide more personalized and targeted counseling services.

Keywords: emotion recognition, convolutional neural network, psychological counseling, long short-term memory, multimodal identification

1 Introduction

Over the past few years, anxiety, depression, low self-esteem, interpersonal sensitivity, and other psychological health problems have occurred frequently among college students, and some of them even have suicidal thoughts, which has a very severe negative impact on individuals, families, and society. As a result, it is valuable to find an effective way to identify mental health issues in students. Emotion recognition technology is a technology that can automatically detect and recognize human emotional states. Its principle is to analyze and process various information such as dialogue speech, text, and physiological data to obtain the emotional state of the person. In psychological counseling for college students, emotion recognition technology can provide effective auxiliary tools for counselors, further improving the effectiveness and quality of counseling.

The increase in mental health issues among college students has attracted widespread attention from both academia and society. Learning about the mental health needs of students of color is a growing priority for colleges and universities across the country. Lipson Sarah Ketchen aimed to learn about the mental health of students of color, including the prevalence of mental health problems and treatment use. Samples consisted of

---

* **Corresponding author: Xiaoran Li,** School of Psychology, Inner Mongolia Normal University, Hohhot 010022, China,
e-mail: lixiaoran1971@163.com

43,375 undergraduate and graduate students from 60 institutions, and data were analyzed at the individual level using bivariate and multivariate models, which showed that only 20% of students with significant mental health problems received treatment, and attitudes related to mental health treatment varied widely [1]. Schwitzer Alan et al. investigated the psychological counseling situation of college students and found that approximately 10% of college students seek psychological health counseling, with many students unable to persist without support. Students who went to counseling centers and continued to receive advice were more likely to improve their mental health. Although the proportion of university psychological counseling centers reporting service to students has increased, those with marginalized identities still have not received sufficient attention [2]. Banks Brea reviewed a university's response to the call for student psychological counseling and specifically explored how the program was implemented, as well as the associated costs and benefits [3]. Novella Jocelyn et al. compared the effects of online synchronous video counseling and face-to-face psychological counseling and used a solution-centered short therapy to consult college students with mild to moderate anxiety. The findings revealed significant changes in participants' scores on the generalized and social anxiety subscales in both study conditions, with no significant difference in validity between the two modalities [4]. Psychological counseling services have become increasingly important on university campuses, but students are often unwilling or hesitant to express their inner emotions. Therefore, how to better identify the psychological state of college students has become an important issue.

Emotional recognition technology can be used to monitor and analyze the emotional state of college students, helping to identify potential mental health issues. Khare Smith and Bajaj proposed using convolutional neural networks (CNNs) to automatically extract and classify emotional features. He used time–frequency representation to convert the filtered electroencephalogram (EEG) signal into an image and evaluated its performance by measuring accuracy and precision. The accuracy of CNN was 93.01% [5]. New human–computer interaction research attempts to consider emotional states to provide better human–computer interfaces. Abdullah et al. reviewed multimodal signal emotion recognition techniques based on deep learning and compared their applications on the basis of existing research. Multimodal emotional computing systems and single-modal solutions were studied together because they provide higher classification accuracy, which encourages research to better understand the physiological signals and emotional awareness issues of the scientific status quo [6]. EEG signals can be utilized to record ongoing neuronal activity in the brain to obtain information about human emotional states. Gupta et al.'s aim was to study the channel specificity of EEG signals comprehensively and to provide an effective way of recognizing emotions based on the flexible resolved wavelet transform, which decomposed EEG signals into different subband signals. Compared with the existing ways, this way exhibited better performance in human emotion classification [7]. Neuroscience research has revealed the differences in emotional expression between the left and right brains of humans. Inspired by this, Li et al. proposed a new bi-hemisphere difference model to learn the difference information between two hemispheres, to enhance the emotion recognition of EEG. He adversarially induced the entire feature learning module to generate emotion related but domain invariant feature representations, thereby further promoting the effectiveness and superiority of the bi-hemisphere difference model in solving EEG emotion recognition problems [8]. The application of emotion recognition technology in psychological counseling for college students is expected to provide better support, intervene in mental health issues in advance, and help college students better cope with challenges.

The 2DCNN-LSTM model is a deep learning model that combines CNN and LSTM to analyze and recognize the physiological signals and emotions of college students. Among them, 2DCNN was utilized to extract spatial and temporal features of physiological signals; LSTM was used to establish temporal dependencies of signal sequences, and 2DCNN was used to extract spatial features of texts, which can help models understand the relationships between words. For speech data, 2DCNN was used to extract features of the sound spectrum, which was input into the LSTM layer. The features of text or speech data were input into the 2DCNN layer, and then the output of the 2DCNN was connected to the LSTM layer. LSTM considers temporal dependencies in text or speech data and learns emotional expression patterns in text or speech data. Through reasonable data processing, feature extraction, and model training, it can help identify and understand the emotional expressions of college students, which can be applied in schools, social media, or mental health fields.

2 Current status of mental health among college students

2.1 Data collection

In this article, the data collection of 2,000 college students in a certain city is implemented through the survey on the status of mental health, with a total of 1,995 questionnaires recovered and 1,977 valid questionnaires. The data from the valid questionnaires are summarized, generalized, and analyzed to provide a more comprehensive understanding of the college students' psychological status.

Basic information about the survey respondents is displayed in Table 1.

Table 1: Basic information of survey subjects

Category		Number of students	Percentage
Gender	Male	1,082	54.7
	Female	895	45.3
Grade	Grade 1	621	31.4
	Grade 2	431	21.8
	Grade 3	524	26.5
	Grade 4	401	20.3

As shown in Table 1, among the survey subjects, the proportion of males is relatively high, at 54.7%, and the proportion of females is 45.3%.

College students are the successors of the great rejuvenation of the Chinese nation and an important group in promoting social development. The mental health status of college students directly affects their personal growth, success, and family happiness, as well as the safety, stability, and harmony of the campus. At the same time, it is closely related to the fate of the nation and the development of society. Adverse events caused by psychological abnormalities among college students are often seen in the media. Most unsafe issues on campus occur due to psychological abnormalities among students, and their mental health is not optimistic, seriously affecting campus safety, family harmony, and social stability [9,10].

Students with psychological issues are depicted in Table 2.

Table 2: Situation of students with psychological problems

Category		Number of students	Percentage
Psychological issues	No psychological issues	1,422	71.9
	Have psychological issues	555	28.1
Gender	Male	221	39.8
	Female	334	60.2
Grade	Grade 1	129	23.24
	Grade 2	102	18.38
	Grade 3	119	21.44
	Grade 4	205	36.94

As shown in Table 2, the count of students with psychological issues is 555 with 28.1%; the count of female students with psychological issues is higher with 60.2%; the count of students with psychological issues in the 4th year of the university is the highest out of the 4 years with 36.94%.

The 4 years of university are a crucial period for the growth of college students, as well as an important process for their psychological development. Contemporary college students bear important responsibilities

for national construction and development, and healthy psychological qualities and mental state are key factors in pursuing personal development [11,12]. However, the country is in a stage of social transformation, with social problems such as fierce competition and difficult employment emerging. At the same time, the amount of college students is constantly increasing, and the pressure they face in terms of learning, interpersonal relationships, economy, and employment is also increasing. The psychological pressure of college students is also increasing accordingly [13,14]. The conditions caused by psychological disorders are commonly seen in the media. Some students are isolated, introverted, and silent, while others have emotional disorders, anxiety, and depression. Some even seek liberation by ending their own or others' lives [15,16].

2.2 Data filtering

In addition to focusing on teaching quality, the psychological health of students cannot be ignored in various universities. To provide psychological assistance to students, many schools not only offer psychological courses but also set up psychological counseling institutions and assign professional staff. However, due to widespread discrimination against mental illness, many students with mental health problems do not actively seek help. Every year when new students enter the school, psychological tests are conducted, and some students conceal their answers. Even when there are many questionnaire questions, there is a phenomenon of random answers, which cannot reflect the true mental health status of students [17,18]. Therefore, both counselors and teachers pay close attention to the behavior of students. Counselors promptly report students with mental health issues through the Student Worker Online System, based on feedback from their classmates, parents, teachers, and school hospitals. Experts at the psychological center assess the severity level of students based on reported events and observations of their micro expressions [19,20].

The psychological level includes three levels: mild, moderate, and heavy. The psychological level of students is illustrated in Figure 1 (the horizontal axis indicates mild, moderate, and heavy, and the vertical axis indicates the number of students).
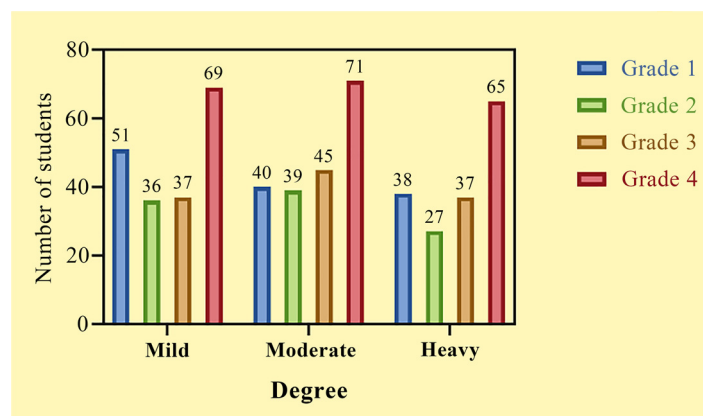

Figure 1: Student's psychological level.

As shown in Figure 1, the number of people with heavy psychological problems in grades 1, 2, 3, and 4 is 38, 27, 37, and 65, respectively.

College students' mental health has a key role to play in their own growth and success. Having a healthy psychology is an effective guarantee for college students to pursue education, communicate with others, and work in the future. The down-to-earth progress in learning and the pursuit of truth and pragmatism in life cannot be separated from a healthy and positive psychological state [21,22]. A good psychological state can not

only guide college students not to lose heart and not be discouraged when facing setbacks and challenges but also actively adapt to adversity. It can promote the cultivation of various abilities of college students, which is a solid foundation for their growth and success. On the contrary, unhealthy psychology leads to college students neglecting their studies and even going astray, causing serious harm to their personal growth and success.

2.3 Data analysis

The academic pressure of college students is directly proportional to their mental health. When their learning tasks increase, their psychological problems such as irritability and depression worsen, leading to various psychological disorders and serious harm to their mental health [23,24]. For individuals, mental health issues can have negative impacts, weaken their adaptability in society, and even pose a serious threat to their physical health. For families, mental health issues can cause psychological burden and also increase their financial burden.

The situation of students who need psychological counseling is described in Table 3.

Table 3: Situation of students who need psychological counseling

Degree of need	Grade 1	Grade 2	Grade 3	Grade 4
Very need	51	43	50	99
Comparative need	40	32	42	73
Indifferent	36	15	17	20
Not need	2	12	10	13

As shown in Table 3, it can be seen that the majority of students are in great need of psychological counseling. The number of students in grades 1, 2, 3, and 4 who are in very need of psychological counseling is 51, 43, 50, and 99, respectively.

Overall, college students' mental health is becoming more serious, which may be related to the increasingly fierce competition and increased pressure on their learning and employment. This also indicates that although universities are currently placing increasing emphasis on college students' mental health, it is not yet in-depth enough and falls far short of the needs of college students. The focus of the mental health work for college students should not be on how to diagnose and treat mental illnesses, but on preventive and developmental research. The psychological intervention model should be constructed from the perspective of adaptation and development and should be carried out through multiple aspects, levels, channels, and means to truly play a greater role in school mental health education in higher education [25].

3 Psychological analysis based on emotion recognition technology

3.1 Text emotion recognition

Textual emotion recognition is based on natural language processing, whose main task is to take feature extraction, analysis, induction, and deduction operations on social media comments, especially textual comments containing emotional states [26,27]. Text emotion recognition is shown in Figure 2.

As illustrated in Figure 2, the primary operation of text emotion recognition is to preprocess the text, then extract features, and finally perform emotion recognition.

The specific ways to preprocess text include word segmentation, part of speech tagging, syntactic and semantic analysis, etc., followed by text feature extraction and fed into a classifier for emotion classification
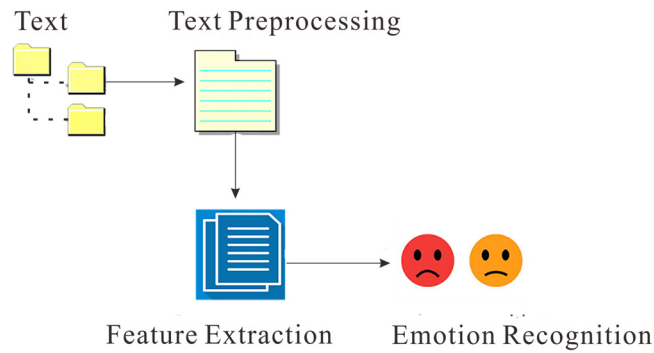
Figure 2: Text emotion recognition.

[28,29]. Text feature extraction and representation is the process of quantifying word vectors extracted from text to make it easy for computers to recognize. There are many popular algorithms for extracting statistical features at this stage. For instance, one-hot coding, which is also one bit effective coding, encodes the categorical values as a binary representation. Although it is said to be able to handle discrete values and expand the features for the purpose of handling words well, there are some prominent drawbacks such as not considering the order between words, and so on.

Term frequency-inverse document frequency is a technique used in text analysis and information retrieval to evaluate the importance of a word in a document set. Term frequency-inverse document frequency (TF-IDF) is a simple and fast text representation method that considers word frequency. TF-IDF is essentially a weighted technique for suppressing noise, which calculates the importance of a character or word in the entire corpus by the number of times it appears in the text and the frequency of documents appearing in the entire corpus. The TF-IDF calculation formula is presented in formula (1):

$$T_{\mathrm{F}} - I_{\mathrm{DF}} = \frac{T_{\mathrm{F}}}{I_{\mathrm{DF}}}. \tag{1}$$

Among them, I_{DF} is the inverse document frequency of the weight parameter and T_{F} is the word frequency.

3.2 Speech emotion recognition

Speech emotion recognition is a technology used to analyze and understand the speaker's speech signal to determine the emotional state or content it expresses. The information conveyed by speech includes semantic information and acoustic information, and acoustic features contain a lot of key information that can be used as auxiliary semantic information. Different data in the audio modality include different emotional information, which has a certain impact on the findings of the final classification of emotion recognition [30,31]. The Mel spectrum is a feature proposed based on human auditory perception, and the Mel cepstral coefficient is nonlinearly related to frequency. The Mel spectrum in speech recognition is chosen to be borrowed for the purpose of better characterizing the energy distribution laws.

To standardize data, the purpose of standardization is to standardize each dimension of features to a specific interval range, so that the data distribution is not too scattered, and to transform dimensional expressions into dimensionless expressions. Standard deviation normalization is a data preprocessing method used to convert data into a standard normal distribution with a mean of 0 and a standard deviation of 1. Standard deviation standardization is the standardization method, and the standardized data conform to the standard normal distribution. The standardized calculation method is shown in formula (2):

$$a_{\mathrm{normal}} = \frac{a - \mu}{\alpha}, \tag{2}$$

α is the standard deviation of all data samples and μ is the mean of all sample data. The standardized data are more conducive to accelerating speed and improving recognition accuracy.

Power spectral density (PSD) is the distribution of signal power at each frequency point expressed in terms of the concept of density, a measure of the mean square value of a random variable, and the average power measure per unit frequency. The PSD is a real number and nonnegative. The PSD of a signal can be used to describe the energy characteristics of the signal as a function of frequency, while also normalizing data at different frequency resolutions, eliminating the influence of resolution.

If function x_t represents the autocorrelation function of the signal, its PSD X can be written as follows:

$$X = \int_{-\infty}^{+\infty} x_t e^{-i} d_t. \tag{3}$$

In the process of collecting speech signals, they are generally stored in a discrete form, so when extracting features from speech signals, only discrete data can be analyzed and processed. Formula (3) is for correlation processing of continuous signals and is not applicable to discrete signals. Therefore, formula (4) is used to process discrete signals:

$$X_m = \sum_{n=0}^{N-1} x_n e^{-i}. \tag{4}$$

Among them, x_n is a discrete value. Extracting the PSD characteristics of a signal can analyze the patterns and structures of the signal in the frequency domain and thus identify the signal. To solve the problem of some signal features not being obvious in the time domain, PSD can be utilized to extract the frequency domain features of the signal.

3.3 1DCNN-LSTM model

When processing longer signals, CNN cannot recognize longer physiological signals. When the sequence is long, using CNN makes it difficult for the gradients behind the sequence to propagate back to the previous sequence, resulting in the problem of vanishing gradients and sometimes even exploding gradients. To solve this problem, 1DCNN-LSTM emerged. The 1DCNN-LSTM model combines these two types of neural networks, typically on the input of text sequences. 1DCNN is used to extract local features, and LSTM is used to integrate sequence information, allowing the model to simultaneously consider both local features and contextual information in the text, which helps with more accurate emotion recognition.
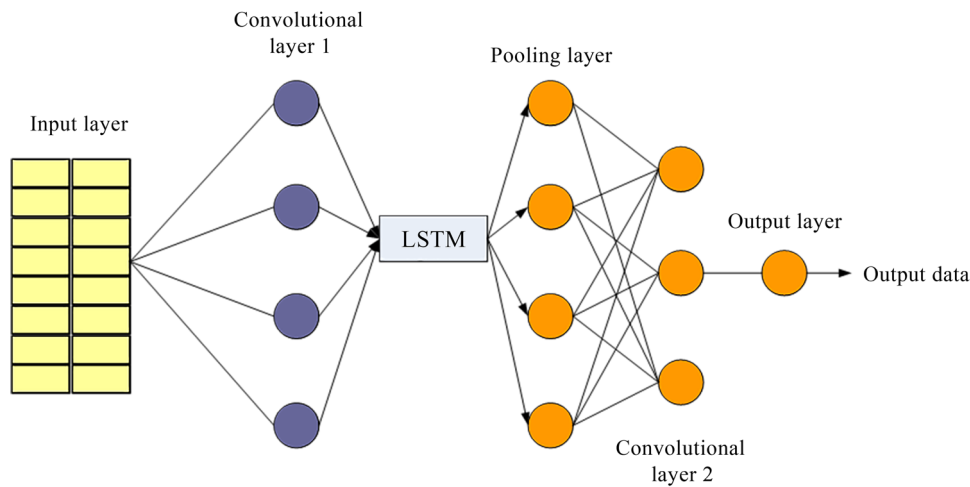

Figure 3: Schematic diagram of 1DCNN-LSTM.

1DCNN-LSTM is a deep learning model used for emotion recognition, typically used to process text, speech, or time series data. The schematic diagram of 1DCNN-LSTM is illustrated in Figure 3.

As shown in Figure 3, 1DCNN-LSTM includes input layers, convolutional layers, pooling layers, LSTM, output layers, and output data.

LSTM has a memory unit that can store and access previous information, allowing the model to handle tasks that require memory and reference of historical information. LSTM is a special type of neural network primarily designed to address related issues that arise during training and effectively model the context of physiological signals. LSTM can enable the network to automatically learn and adjust the attention to the input of extended contextual features. This method can improve data quality, optimize feature vectors, and effectively highlight key emotional information. By calculating input gates, forget gates, cell states, and output gates, high-level emotional-related features are obtained.

The main inputs in the input gate are x_t and h_{t-1}. Among them, x_t represents the current input of the cell, and h_{t-1} represents the previous cell state. These parameters form a weight matrix that determines how much new information is added:

$$i_t = \sigma W_i * h_{t-1} \times x_t + b_i. \tag{5}$$

The forgetting gate mainly determines how much old information new cells need to be discarded and unwanted information.

The output gate contains the current input, the previous hidden layer state, and the current cell state. These parameters form a weight matrix that determines which information is output, as shown in formula (6):

$$h_t = o_t \times C_t. \tag{6}$$

LSTM is also unable to correctly learn and process certain very long or continuous time series, which are not prepartitioned into appropriate training subsequences with clearly defined starting and ending points. The problem is that a continuous input stream may ultimately lead to an unlimited increase in the internal values of the unit, even if the repetitive nature of the problem indicates that they should be reset occasionally.

3.4 2DCNN-LSTM model

Different modalities of data, such as text and speech, contain different representation information. Research on multimodal emotion recognition requires the fusion of information from different modalities to achieve the synergistic effect of information from each modality [32,33]. The core idea behind the LSTM architecture is the storage unit that can maintain its state over time, and the nonlinear gating unit can adjust the speed of information flow into and out of the unit. LSTM is now applied to many learning problems, whose scale and properties are completely different from the original problems.

2DCNN-LSTM is a combination of 2DCNN and LSTM. Its core idea is to use two-dimensional CNNs to extract local features of text and speech and use short-term memory learning methods to obtain the temporal and contextual relationships of text and speech.

2DCNN is a deep neural network primarily used for processing high-dimensional data. In the field of emotion recognition, text data are generally represented as vectors. 2DCNN extracts local features from vocabulary information, such as phrases and sentence structures, through convolutional operations. Due to the diverse manifestations of emotions, these local features are crucial for emotion classification, as different emotions are often associated with different expressions. The convolutional layer formula is expressed as follows:

$$F(i, j, k) = f\left(\frac{I}{W} \times k\right) + b. \tag{7}$$

Among them, $f\left(\frac{I}{W} \times k\right)$ represents the activation function; (i, j, k) represents input data; and b represents the bias term. Convolutional layers are utilized to extract features of input signals.

LSTM is suitable for processing time series data. In terms of emotion recognition, LSTM can assist models in better understanding the tense of text, such as the order of words, contextual relationships, etc. This is the key to distinguishing different types of emotions, as the same word has different emotional meanings in different contexts.

2DCNN-LSTM generally includes a complete connected layer, which effectively fuses the features of 2DCNN-LSTM to obtain emotion recognition results based on deep neural networks. Activation functions are often utilized to output a probability distribution for each emotion type. The output layer formula is expressed as follows:

$$Y = f \times W_H + b. \tag{8}$$

The 2DCNN-LSTM model can be applied to emotional analysis of social media, public opinion monitoring, emotional analysis of customer service, and mental health. It can help businesses understand the level of customer satisfaction or help healthcare workers determine the emotions of patients.

4 Emotional recognition effect

Given the fast growth in the fields of big data and machine learning, emotion recognition is being accepted by more and more people. Among them, deep learning is a relatively mature emotion recognition method. The experiment compared the application effects of 1DCNN-LSTM and 2DCNN-LSTM deep neural networks in college student emotion recognition.

4.1 Single modal recognition accuracy

4.1.1 Comparison of text emotion recognition

The experimental data used data from students with psychological problems mentioned earlier. The actual psychological problems of students are presented in Table 4.

Table 4: Actual psychological problems of students (multiple choice)

Psychological issues and average	Grade 1	Grade 2	Grade 3	Grade 4
Depressed	47	62	50	83
Anxiety	50	47	64	57
Anger	54	55	47	60
Fear	55	75	51	72
Lonely	41	63	67	61
Disappointed	60	55	61	75
Average	51	60	57	68

As displayed in Table 4, the average number of actual psychological problems among students in grades 1, 2, 3, and 4 was 51, 60, 57, and 68, respectively.

Only text emotion recognition was performed on students, and they were asked to write down their emotions. The emotion recognition results of 1DCNN-LSTM and 2DCNN-LSTM are illustrated in Figure 4 (the horizontal axis shows different psychological states, and the vertical axis shows the number of errors).
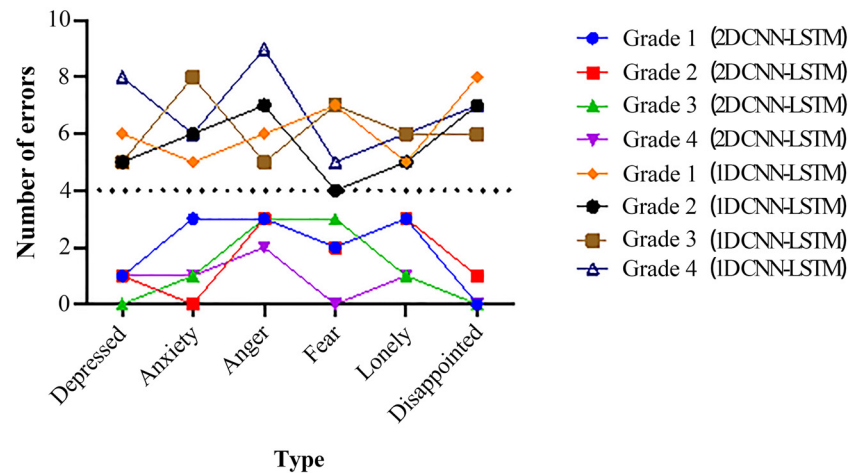
Figure 4: Emotion recognition performance of 1DCNN-LSTM and 2DCNN-LSTM.

As illustrated in Figure 4, the recognition error of 2DCNN-LSTM for the number of students with different psychological states was within 4, while the recognition error of 1DCNN-LSTM for the number of students with different psychological states was mostly above 4.

The recognition error of 2DCNN-LSTM for the number of students with different psychological states was smaller than that of 1DCNN-LSTM for the number of students with different psychological states. 2DCNN-LSTM can extract more features, better capture text, and improve recognition accuracy.

2DCNN-LSTM can process multiple types of data simultaneously, but 1DCNN-LSTM can only handle local data, so it requires a high amount of data. When the sample size is insufficient, the role of the 1DCNN-LSTM model cannot be demonstrated.

4.1.2 Comparison of speech emotion recognition

The average recognition accuracy of 1DCNN-LSTM and 2DCNN-LSTM for speech emotion recognition of students is illustrated in Figure 5 (where the horizontal axis is different grades and the vertical axis is accuracy).
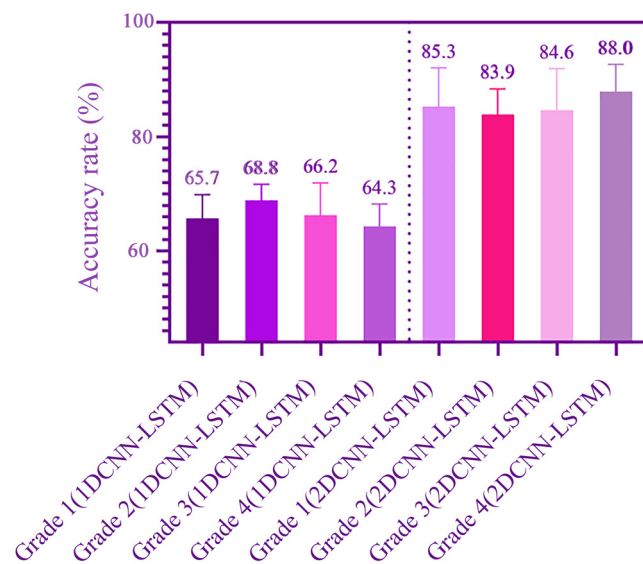

Figure 5: Accuracy of speech emotion recognition using different methods.

As presented in Figure 5, the left side of Figure 5 shows that the average recognition accuracy of 1DCNN-LSTM for different psychological states was 65.7, 68.8, 66.2, and 64.3%, respectively; the right side of Figure 5 shows that the average recognition accuracy of 2DCNN-LSTM for different psychological states was 85.3, 83.9, 84.6, and 88.0%, respectively.

It can be seen that in speech emotion recognition, the average recognition rate of 2DCNN-LSTM was relatively high, while the average recognition rate of 1DCNN-LSTM was relatively low.

4.2 Multimodal recognition accuracy

1DCNN-LSTM was used to simultaneously recognize students' text and speech, that is, to perform multimodal recognition. The situation is shown in Table 5.

Table 5: Recognition accuracy of 1DCNN-LSTM under multimodal conditions (%)

Psychological issues and average	Grade 1	Grade 2	Grade 3	Grade 4
Depressed	84.12	77.43	78.21	75.72
Anxiety	74.24	76.18	83.50	72.90
Anger	71.54	77.06	76.81	82.61
Fear	80.30	83.00	72.77	80.42
Lonely	75.40	83.18	79.88	78.58
Disappointed	71.11	76.92	77.53	77.09
Average	76.12	78.96	78.12	77.89

As presented in Table 5, the average recognition accuracy of 1DCNN-LSTM for first-, second-, third-, and fourth-grade students under multimodal conditions was 76.12, 78.96, 78.12, and 77.89%, respectively.

The recognition accuracy of 2DCNN-LSTM under multimodal conditions is presented in Table 6.

Table 6: Recognition accuracy of 2DCNN-LSTM under multimodal conditions (%)

Psychological issues and average	Grade 1	Grade 2	Grade 3	Grade 4
Depressed	88.06	90.56	90.22	85.73
Anxiety	90.20	99.39	94.72	91.42
Anger	89.34	93.33	93.21	98.01
Fear	97.73	97.34	91.86	98.56
Lonely	88.82	98.06	88.85	85.27
Disappointed	85.46	99.22	98.00	95.86
Average	89.94	96.32	92.81	92.48

In multimodal conditions, the average recognition accuracy of 1DCNN-LSTM and 2DCNN-LSTM for fourth-year university students was 77.89% (data in Table 5) and 92.48%, respectively. The average recognition accuracy of 2DCNN-LSTM for fourth-year university students was 18.7% $\left(\frac{92.48 - 77.89}{77.89} = 18.7\% \right)$ higher than that of 1DCNN-LSTM for the fourth-year university students.

Combining Tables 5 and 6, it can be seen that the recognition accuracy of 2DCNN-LSTM was generally higher than that of 1DCNN-LSTM.

For the emotion recognition task of college students, the 2DCNN-LSTM model can more accurately capture multimodal information compared to the 1DCNN-LSTM model and therefore performs better in processing multimodal data [34,35]. However, it should be noted that in practical applications, it is necessary to choose a suitable model structure based on the characteristics of the scene and data and undergo strict testing and verification to ensure the performance of the model. The psychological counseling model based on 2DCNN-LSTM usually combines multimodal data (such as speech and text) for emotional analysis and counseling support. This model can more comprehensively capture the emotional state of the counselor, thereby providing more accurate counseling support. By analyzing the voice emotions of the counselor, their emotional state can be better understood, and personalized and targeted counseling advice can be provided.

4.3 Psychological counseling effectiveness

Traditional psychological counseling often relies on manual evaluation and is easily influenced by personal cognition and experience. Emotion recognition technology utilizes objective calculation methods to analyze and identify the emotions of consultants, reducing the influence of subjective factors and achieving more objective and less biased emotional evaluations. Emotional recognition can provide real-time analysis of the emotional state of college students and provide information to psychological counseling institutions. Through this approach, psychological counselors can timely understand the psychological changes of students and develop corresponding psychological counseling strategies based on this, providing personalized psychological support for students.

After 2 months of traditional psychological counseling (subjective judgment by psychological counselors) and psychological counseling based on emotion recognition technology, the improvement of students' psychological problems is presented in Figure 6 (the horizontal axis depicts grade, and the vertical axis depicts the number of students).
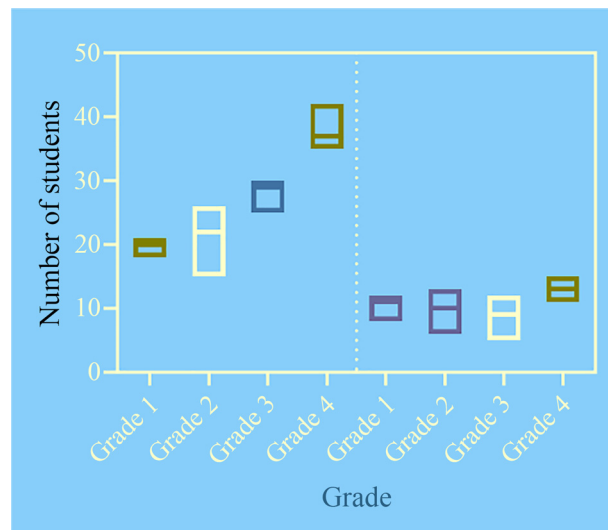

Figure 6: Improvement of psychological problems among students under different psychological counseling.

As illustrated in Figure 6, the left side of Figure 6 shows the improvement of psychological problems among students after traditional psychological counseling, which has decreased compared to the number of students who did not receive psychological counseling at the beginning. However, there are still situations where students in their fourth year of university still rank above 30. The right side of Figure 6 shows the

improvement of psychological problems among students after psychological counseling based on emotion recognition technology. The number of students with psychological problems is within 20.

It indicates that the improvement effect of psychological counseling based on emotion recognition technology on students' psychological problems is better than that of traditional psychological counseling.

5 Discussion

Figure 4 shows that the recognition error of 2DCNN-LSTM is smaller. 2DCNN-LSTM uses 2DCNN, which can effectively perform spatial analysis on speech signals. It can fully utilize the multiple time–domain characteristics of speech signals, reflecting both the local characteristics of speech signals and the overall characteristics. This helps people better understand the emotional expression in speech, especially when emotional characteristics are associated with the spectral pattern of speech. LSTM can capture temporal dependencies when processing sequence data.

In 2DCNN-LSTM, LSTM can extract richer speech information from the features of 2DCNN, thereby helping people better understand emotional changes. 2DCNN-LSTM has shown good fitting performance in certain emotion recognition problems due to its multiple parameters and high complexity. This also means that it is easier to discover hidden patterns from the data, thereby improving the accuracy of recognition.

Figure 6 shows that psychological counseling based on emotion recognition technology has a better effect. Psychological counseling based on emotion recognition technology may have better effects in emotional analysis and counseling support compared to traditional psychological counseling. This is because emotion recognition technology can provide richer and more comprehensive emotional information, enabling a deeper understanding of the emotional state of the counselor. However, the specific choice of richness still needs to be comprehensively considered in practical application scenarios. In addition, any form of psychological counseling requires the cooperation and guidance of professional psychological counselors, and deep learning models are only auxiliary tools and do not replace human professional judgment and intervention.

Emotion recognition technology can analyze the emotional state of students through language, speech, and other information. In this way, psychological counselors can understand the emotional needs and troubles of the counselor and provide appropriate support and advice for each case. Traditional psychological counseling cannot accurately capture subtle emotional changes and contextual information. Emotional recognition technology can provide certain protection for the privacy of consultants. Compared with traditional psychological counseling, this method does not require much background information or face-to-face communication, allowing counselors to conduct counseling in a relatively anonymous environment, thereby reducing potential awkwardness, privacy leaks, and other issues that may arise during the counseling process.

6 Conclusions

As society develops and people's awareness of mental health increases, college students' mental health problems are gradually receiving attention. As a special group, college students often have mental health issues that involve their learning, life, career planning, and other aspects, which often require psychological counseling to solve. However, traditional psychological counseling methods have many problems, such as the professional ability and communication skills of counselors and the language expression ability of patients themselves. Therefore, using technological means to assist psychological counseling can effectively improve the effectiveness of counseling, and the application of emotion recognition technology in college student psychological counseling has broad prospects for application. In future research, continuous optimization of the algorithm, expansion of the data collection range, and systematic application evaluation are needed to further enhance the application effect and applicability of the technology. Meanwhile, the cross research between technology and psychology needs to be strengthened to explore more applications of emotion

recognition technology in the field of psychology, so as to make a greater contribution to the health and happiness of human beings.

References

[1]   Lipson SK, Kern A, Eisenberg D, Breland-Noble AM. Mental health disparities among college students of color. J Adolesc Health. 2018;63(3):348–56. doi: 10.1016/j.jadohealth.2018.04.014.

[2]   Schwitzer Alan M, Moss CB, Pribesh SL, St. John DJ, Burnett DD, Thompson LH, et al. Students with mental health needs: College counseling experiences and academic success. J Coll Stud Dev. 2018;59(1):3–20. doi: 10.1353/csd.2018.0001.

[3]   Banks Brea M. Meet them where they are: An outreach model to address university counseling center disparities. J Coll Stud Psychother. 2020;34(3):240–51. doi: 10.1080/87568225.2019.1595805.

[4]   Novella Jocelyn K, Ng K-M, Samuolis J. A comparison of online and in-person counseling outcomes using solution-focused brief therapy for college students with anxiety. J Am Coll Health. 2022;70(4):1161–8. doi: 10.1080/07448481.2020.1786101.

[5]   Khare Smith K, Bajaj V. Time–frequency representation and convolutional neural network-based emotion recognition. IEEE Trans Neural Netw Learn Syst. 2020;32(7):2901–9. doi: 10.1109/TNNLS.2020.3008938.

[6]   Abdullah SMSA, Ameen Ameen SY, Sadeeq MAM, Zeebaree S. Multimodal emotion recognition using deep learning. J Appl Sci Technol Trends. 2021;2(2):52–8. doi: 10.38094/jastt20291.

[7]   Gupta V, Chopda MD, Pachori RB. Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. IEEE Sens J. 2018;19(6):2266–74. doi: 10.1109/JSEN.2018.2883497.

[8]   Li Y, Wang L, Zheng W, Zong Y, Qi L, Cui Z, et al. A novel bi-hemispheric discrepancy model for EEG emotion recognition. IEEE Trans Cognit Dev Syst. 2020;13(2):354–67. doi: 10.1109/TCDS.2020.2999337.

[9]   Choi N-Y, Miller MJ. Social class, classism, stigma, and college students' attitudes toward counseling. Couns Psychol. 2018;46(6):761–85. doi: 10.1177/0011000018796789.

[10]   Gibbons S, Trette-McLean T, Crandall AA, Bingham JL, Garn CL, Cox JC. Undergraduate students survey their peers on mental health: Perspectives and strategies for improving college counseling center outreach. J Am Coll Health. 2019;67(6):580–91.

[11]   Karaman MA, Lerma E, Vela JC, Watson JC. Predictors of academic stress among college students. J Coll Couns. 2019;22(1):41–55. doi: 10.1002/jocc.12113.

[12]   Lipson SK, Lattie EG, Eisenberg D. Increased rates of mental health service utilization by US college students: 10-year population-level trends (2007–2017). Psychiatr Serv. 2019;70(1):60–3. doi: 10.1176/appi.ps.201800332.

[13]   Cadaret MC, Bennett SR. College students' reported financial stress and its relationship to psychological distress. J Coll Couns. 2019;22(3):225–39. doi: 10.1002/jocc.12139.

[14]   Selvaraj PR, Bhat CS. Predicting the mental health of college students with psychological capital. J Ment Health. 2018;27(3):279–87. doi: 10.1080/09638237.2018.1469738.

[15]   Jones PJ, Park SY, Tyler Lefevor G. Contemporary college student anxiety: The role of academic distress, financial stress, and support. J Coll Couns. 2018;21(3):252–64. doi: 10.1002/jocc.12107.

[16]   Levin ME, Hicks ET, Krafft J. Pilot evaluation of the stop, breathe & think mindfulness app for student clients on a college counseling center waitlist. J Am Coll Health. 2022;70(1):165–73. doi: 10.1080/07448481.2020.1728281.

[17]   Cross TL, Cross JR, Mammadov S, Ward TJ, Neumeister KS, Andersen L. Psychological heterogeneity among honors college students. J Educ Gifted. 2018;41(3):242–72. doi: 10.1177/0162353218781875.

[18]   Ratts MJ, Greenleaf AT. Counselor–advocate–scholar model: Changing the dominant discourse in counseling. J Multicultural Couns Dev. 2018;46(2):78–96. doi: 10.1002/jmcd.12094.

[19] Airdrie JN, Langley K, Thapar A, van Goozen SHM. Facial emotion recognition and eye gaze in attention-deficit/hyperactivity disorder with and without comorbid conduct disorder. J Am Acad Child Adolesc Psychiatry. 2018;57(8):561–70. doi: 10.1016/j.jaac. 2018.04.016.

[20] House LA, Neal C, Kolb J. Supporting the mental health needs of first generation college students. J Coll Stud Psychother. 2020;34(2):157–67. doi: 10.1080/87568225.2019.1578940.

[21] Browning BR, McDermott RC, Scaffa ME, Booth NR, Carr NT. Character strengths and first-year college students' academic persistence attitudes: An integrative model. Couns Psychol. 2018;46(5):608–31. doi: 10.1177/0011000018786950.

[22] Cohen KA, Graham AK, Lattie EG. Aligning students and counseling centers on student mental health needs and treatment resources. J Am Coll Health. 2022;70(3):724–32. doi: 10.1080/07448481.2020.1762611.

[23] Baams L, De Luca SM, Brownson C. Use of mental health services among college students by sexual orientation. LGBT Health. 2018;5(7):421–30. doi: 10.1089/lgbt.2017.0225.

[24] Scheel MJ, Stabb SD, Cohn TJ, Duan C, Sauer EM. Counseling psychology model training program. Couns Psychol. 2018;46(1):6–49. doi: 10.1177/0011000018755512.

[25] DeBlaere C, Singh AA, Wilcox MM, Cokley KO, Delgado-Romero EA, Scalise DA, et al. Social justice in counseling psychology: Then, now, and looking forward. Couns Psychol. 2019;47(6):938–62. doi: 10.1177/001100001989328.

[26] Zhang K, Li Y, Wang J, Cambria E, Li X. Real-time video emotion recognition based on reinforcement learning and domain knowledge. IEEE Trans Circuits Syst Video Technol. 2021;32(3):1034–47. doi: 10.1109/TCSVT.2021.3072412.

[27] Zhang T, Zheng W, Cui Z, Zong Y, Li Y. Spatial–temporal recurrent neural network for emotion recognition. IEEE Trans Cybern. 2018;49(3):839–47. doi: 10.1109/TCYB.2017.2788081.

[28] Kim BH, Jo S. Deep physiological affect network for the recognition of human emotions. IEEE Trans Affect Comput. 2018;11(2):230–43. doi: 10.1109/TAFFC.2018.2790939.

[29] Saxena A, Khanna A, Gupta D. Emotion recognition and detection methods: A comprehensive survey. J Artif Intell Syst. 2020;2(1):53–79. doi: 10.33969/AIS.2020.21005.

[30] Kollias D, Zafeiriou S. Exploiting multi-cnn features in cnn-rnn based dimensional emotion recognition on the omg in-the-wild dataset. IEEE Trans Affect Comput. 2020;12(3):595–606. doi: 10.1109/TAFFC.2020.3014171.

[31] Li J, Qiu S, Shen Y-Y, Liu C-L, He H. Multisource transfer learning for cross-subject EEG emotion recognition. IEEE Trans Cybern. 2019;50(7):3281–93. doi: 10.1109/TCYB.2019.2904052.

[32] Huang H, Xie Q, Pan J, He Y, Wen Z, Yu R, et al. An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness. IEEE Trans Affect Comput. 2019;12(4):832–42. doi: 10.1109/TAFFC.2019.2901456.

[33] Li P, Liu H, Si Y, Li C, Li F, Zhu X, et al. EEG based emotion recognition by combining functional connectivity network and local activations. IEEE Trans Biomed Eng. 2019;66(10):2869–81. doi: 10.1109/TBME.2019.2897651.

[34] Li Y, Zheng W, Wang L, Zong Y, Cui Z. From regional to global brain: A novel hierarchical spatial-temporal neural network model for EEG emotion recognition. IEEE Trans Affect Comput. 2019;13(2):568–78. doi: 10.1109/TAFFC.2019.2922912.

[35] Liu W, Qiu J-L, Zheng W-L, Lu B-L. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. IEEE Trans Cognit Dev Syst. 2021;14(2):715–29. doi: 10.1109/TCDS.2021.3071170.