Rev.int.med.cienc.act.fís.deporte - vol. 24 - número 96 - ISSN: 1577-0354

Li B et al. (2024) SCREENING AND PREVENTION OF EARLY MENTAL ILLNESS IN ATHLETES BASED ON BIOMEDICAL DATA ANALYSIS AND PATTERN RECOGNITION TECHNIQUES. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (96) pp. 146-165
DOI: https://doi.org/10.15366/rimcafd2024.96.009

ORIGINAL

SCREENING AND PREVENTION OF EARLY MENTAL ILLNESS IN ATHLETES BASED ON BIOMEDICAL DATA ANALYSIS AND PATTERN RECOGNITION TECHNIQUES

Baoqiang Li[1,2*], Gong Zhang[1,2], Yi Niu[1,2], Zhongju Liu[1,2]

[1] Xi'an Polytechnic University, Xi'an, 710048, China
[2] Xi'an Jiaotong University, Xi'an, 710049, China
E-mail: 17802902929@163.com

Recibido 07 de octubre de 2023 **Received** October 07, 2023
Aceptado 07 de puede de 2024 **Accepted** May 07, 2024

ABSTRACT

In recent years, the mental health problems of athletes have become more and more serious, and due to the complexity and invisibility of mental illnesses, it is difficult to detect and intervene in the early stage of mental illnesses due to a number of reasons. With the explosion of biomedical data and the rapid development of pattern recognition technology, the problem of recognizing mental illnesses can be solved by using artificial intelligence related technology. This paper proposes a mental illness recognition model based on Roberta Text CNN. Based on Roberta Text CNN model, a text classification model is built for two types of labels, namely, mental illness and suicidal tendency, and a comparison experiment is carried out with several mainstream models. The experimental results show that the proposed method achieves competitive results. In addition, this paper also introduces the EEG signals to be organized into dimensionally consistent matrices using Mel's inverted spectral coefficients for scale change, and then inputs them into the convolutional neural network as image features for further extraction of high-dimensional features and serial fusion of EEG features and then inputs them into the long and short-term memory network, and the application of the pattern recognition technology provides a two-branch mental illness recognition model, which further enhances the capability of screening and preventing early mental illnesses in athletes. The application of pattern recognition technology provides a two-branch mental illness identification model, which further enhances the ability of screening and prevention of early mental illness in athletes.

KEYWORDS: Biomedical Data Analysis; Pattern Recognition; LSTM; Deep Learning

1. INTRODUCTION

In recent years, due to rapid economic development, accelerated pace of life and other reasons, the number of patients with mental illness has increased year by year. According to the World Health Organization (WHO), mental disorders affect about 1 billion people, and someone loses their life to suicide every 40 seconds. After the COVID-19 epidemic, the number of patients with depression increased by 53 million, and anxiety disorders and major depression increased by 26% and 28% respectively (Park & Kim, 2020; Santomauro et al., 2021; Twenge & Joiner, 2020). The pressure of the epidemic has brought huge pressure on the diagnosis and treatment of depression. Relevant data shows that emotional distress such as depression and anxiety ranks first among various health concerns. More than 90% of respondents believe that they have psychological problems, and about 50% of teenagers say that they suffer from emotional problems (Jiayao, 2022). 14.8% of athletes are at risk of depression to varying degrees, which is higher than that of adults. Behind these figures, a grim fact is revealed: teenagers are experiencing a "psychological crisis", the mental health problems of teenagers need to be paid more attention to, and the mental health education system needs to be improved urgently (Groffik et al., 2020; Osman, 2020). The traditional mental health diagnosis method mainly requires patients to go to mental hospitals for consultation. In fact, many teenagers bury a lot of negative emotions in their hearts, often ignore their own mental health status, and do not take the initiative to seek medical treatment. In addition, families and schools It is difficult to make professional judgments about the mental health status of teenagers, and they do not have the professional mental medical knowledge to provide auxiliary treatment to patients. Therefore, the condition is easily delayed, leading to irreversible serious consequences (Alexandre et al., 2018; Bonci et al., 2008).

In the field of artificial intelligence (Chen, Han, et al., 2023; Chen, Li, et al., 2023), natural language processing (NLP) is an important research content. Through designed algorithms, machines can automatically parse and generate natural language. Natural language processing mainly includes two core tasks, namely the understanding of natural language text and the generation of natural language text (Li & Cao, 2023). Natural language understanding (NLU) refers to converting the natural language used by users to interact with computers into structured data, that is, understanding the intent of the natural language input by the user and extracting key information in the text. Natural language generation (NLG) refers to converting structured data into human-understandable text output. Natural language processing technology is widely used in applications such as machine translation and dialogue systems. Among them, dialogue systems have broad application prospects and have been regarded as a research hotspot in industry and academia in recent years.

Natural language understanding is an important module in the dialogue system, and in the natural language understanding module, intent recognition is a crucial step. The higher the accuracy of intent recognition, the better the overall performance of the dialogue system. Although the mental illness recognition studied in this article is not traditional intent recognition, it essentially extracts semantic information from text, which is usually solved by researchers as a text classification task. Since the psychologically relevant feature information in dialogue text is relatively obscure, it is difficult for traditional rule-based and machine learning-based methods to extract deep semantic information, so this topic uses the currently popular deep learning method for research. Compared with traditional methods, deep learning (Ameer et al., 2022; Cho et al., 2019; Ning et al., 2023) does not require manual construction of features, but uses a large amount of corpus training to mine text features. Common deep learning networks include recurrent neural networks (RNN), long short-term memory networks (LSTM), bidirectional long and short neural networks (BiLSTM), convolutional neural networks (CNN), etc. These neural networks have good results in the field of intent recognition (Glick & Applbaum, 2010; Shoeibi et al., 2021). With the rise of pre-training research, pre-training models are widely used in text classification tasks, such as the currently popular BERT, ELMO, ENRIE, etc. This article mainly studies the field of mental illness identification under mainstream deep models. It uses pre-trained language models to vectorize text, and then uses text classifiers to complete mental illness identification (Xu & Li, 2019).

Currently, knowledge graphs are used in various fields. Knowledge graphs can store knowledge in an associative and structured manner, thereby better expressing the complex relationships between knowledge. In the dialogue system, it is difficult to obtain enough information to identify the user's psychological condition in a single round of dialogue, and due to the stigmatized psychology of the patient, the patient is often unwilling to disclose or express more information. At this time, it is necessary to use the dialogue system to conduct multiple rounds of questioning on the user, and induce the user to state more content through multiple questions, thereby improving the effect of identifying mental illness. The knowledge graph can be applied to the dialogue system, and the knowledge graph can be used to drive multi-round dialogue strategies to complete mental illness identification tasks and mental medical resource recommendation. Therefore, this article studies the introduction of these two technologies into the field of athlete mental health, and builds a mental illness identification and mental medical resource recommendation model. The main contributions are as follows:

(1) In this paper, we address the problem of identifying athlete mental illnesses in a mental health dialogue system by applying a text classification

model to initially identify the presence of mental illnesses in the user, and then make a specialized diagnosis of the illnesses based on a domain knowledge graph.

(2) In this paper, the EEG signals are introduced to be organized into dimensionally consistent matrices using Mel's inverted spectral coefficients for scale change, and then they are inputted into convolutional neural networks as image features for further serial fusion of extracted high-dimensional features with the EEG features and then inputted into long and short-term memory networks, which further enhances the ability of screening and prevention of athletes' early mental illnesses.

2. Related Works

2.1 Text CNN

(1) The CNN network is an expanded implementation of the feed-forward neural network (MLP), also known as a multilayer perceptual machine (MLP). MLP operates on the premise that neurons within a specific region of its surroundings can influence the responses of other neurons. In general, convolutional neural networks are composed of the following layers: convolutional, pooling, input, and entirely connected. While convolutional neural networks are predominantly employed in computer vision due to their inherent properties, they have also demonstrated promising outcomes in the domain of natural language processing. Kim from New York University introduced the Text CNN model in 2019. This marks the inaugural application of a convolutional kernel neural network to the task of text categorization (Khare et al., 2021).

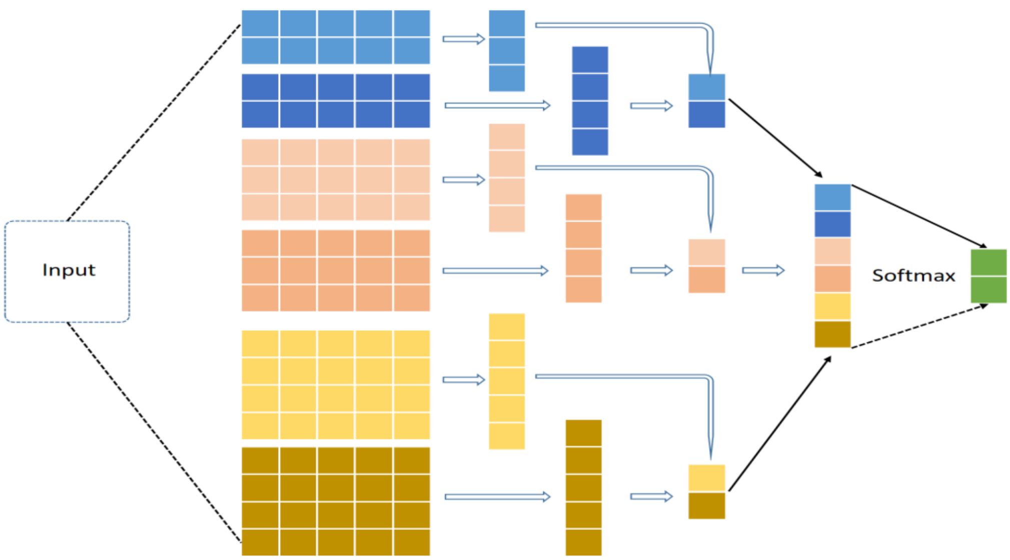

Figure 1: Schematic representation of the Text CNN model structure.

The architecture of Text CNN, as seen in Figure 1, consists of an input layer that comprises an embedding matrix with dimensions n*k. Here, n represents the number of words in the sentence, while k denotes the dimensionality of the word vectors. The Text CNN model utilizes pre-trained word vectors as its input. Each word in the dataset is represented by a word vector, allowing each sentence to be represented as an embedding matrix. Each row in the embedding matrix corresponds to a k-dimensional word vector. Each element in a row of the embedding matrix represents a word vector with k dimensions. The convolutional layer operates similarly to a conventional convolutional layer in a CNN, where the width of the convolutional kernel is equivalent to the size of the word vector k. Hence, the convolutional kernel just requires vertical shifting, with the word serving as the lowest unit of displacement. The equation for calculating the feature a_i obtained from the convolutional layer is as follows:

$$a_i = f(W \cdot X_{i:i+h-1} + b) \qquad (1)$$

Where X is the input text, W is the convolution kernel, b is the bias, $X_{i:i+h-1}$ is the sliding window and f is the ReLu activation function. The pooling layer, on the other hand, compresses the size of the parameter matrix by extracting the maximum value in the eigenvectors, and then picks up the fully-connected layer, then:

$$Y = \mathrm{softmax}(W \cdot (Z \circ R) + b) \qquad (2)$$

2.2 Knowledge graph

There are two main methods for the construction of knowledge graph, namely, top-down construction method and bottom-up construction method. The top-down construction method is usually used for domain-based knowledge graph, which needs to analyse the domain knowledge in advance to construct the ontology library model, and then complete the construction of the knowledge graph according to the constraints of the ontology library. The data source of this method is usually structured data or semi-structured data. Bottom-up construction method is usually used for generalized knowledge graph, which extracts information from existing public data, and at the same time integrates the data with high credibility into the knowledge graph to continuously enrich the knowledge graph, and the bottom-up method does not have strict requirements on the boundary of knowledge, and does not need to construct an ontology library. The data source of this method is generally unstructured data. As can be seen from the previous section, there are two main stages in the construction of the knowledge graph, namely, knowledge extraction and knowledge fusion. (1) Knowledge Extraction Knowledge extraction refers to the extraction of knowledge elements such as entities and entity relationships from unstructured or semi-structured data. In general, knowledge extraction can be divided into entity extraction, attribute extraction, and relationship extraction. Entity extraction, which can also be called Named Entity Recognition (NER), refers to the automatic identification of named entity elements from data, mainly extracting place names, organization names, times, events, amounts, etc. from text data. Commonly used methods for entity extraction include Simple Bayes, K-neighbour algorithm, support vector machine, and deep learning based methods. Attribute extraction refers to extracting the attribute information of a specified entity. Commonly used attribute extraction methods include rule-based slot filling method, sequence labelling method based on deep learning, etc (Marriwala & Chaudhary, 2023; Vázquez-Abad et al., 2020). Relationship extraction refers to the automatic identification of the target relationships of entities from data, and the methods for relationship extraction include rule-based and template-based methods, machine learning-based methods, and deep learning-based methods. Each of these methods has its own characteristics, and different methods can be selected according to the application requirements. Among them, the rule- and template-based methods need to formulate semantic rule templates in advance, and then obtain inter-entity relationships through template matching. Template matching is suitable for small-scale domains, but the portability is low and the model must be reconstructed in other domains. Machine learning-based approaches treat the relationship extraction task as a text classification task, which requires manually labelling the data, designing effective features, and predicting relationships by training a classification model. Deep learning-based approaches are the most popular relationship extraction methods, which are similarly treated as text categorization tasks, and which are capable of supervised learning and automatic relationship extraction. (2) Knowledge Fusion Knowledge fusion refers to the integration of knowledge from different sources into a single knowledge base, which results in a comprehensive and accurate description of the entities and makes the knowledge base more expressive and logical. Knowledge fusion requires the use of entity disambiguation techniques; entity disambiguation is the solution to the problem of the same entity having different names or not being of the same type. Commonly used methods for knowledge fusion are machine learning based methods and deep learning based methods. Machine learning based methods contain SVM, decision trees, integrated learning, etc. (Rajendran et al., 2023). Deep learning based methods use neural networks which are better able to handle large scale data.

3. Methodology

3.1 Roberta-Text CNN

3.1.1 Model Selection

In the text classification task, the feature representation of text will directly affect the effect of text classification, given the current rapid development of pre-training models, this paper chooses the combination of pre-training models and classifiers to complete the text classification task. Pre-training models can be divided into two types: static word vector and dynamic word vector models. Common static word vector models include Word2Vec, Fast Text, Glove, etc. Although this type of static word vector model has a simple structure, it is not capable of solving the problem of multiple meanings of words. Another kind of dynamic word vector models can learn context-specific word feature representations, and the more popular dynamic word vector models are ELMO, BERT, Roberta, etc. Among them, ELMO is only a stack of bi-directional long and short-term memory networks, while BERT is a bi-directional pre-training model integrating the Transformer structure and the attention mechanism, so BERT has a better feature extraction ability than ELMO. The Roberta model is an optimization of the BERT model, and there is no difference in the structure of the Roberta model and the BERT model, but in terms of the training strategy, the Roberta model has made a number of optimizations, and the main optimizations are in the improvement of the masking method, the addition of the pre-training task, the increase of the training corpus, the tuning of the hyper-parameters, and the optimization of the (compressed) model structure, etc. In summary, this paper chooses the model that has better features than the ELMO model. In summary, this paper chooses the Roberta model as the pre-training model, which has better feature expression ability. Roberta model can learn the feature information of the text very well, but due to its small parameter changes, there is an overfitting problem, and Roberta can only learn the word features. In order to avoid the above problems, this paper chooses Text CNN classifier to extract the word feature information of different lengths, and obtains a better text classification effect by fusing the advantages of Roberta and Text CNN.

3.1.2 Construction of mental illness recognition model based on Roberta-Text CNN

This study aims to develop a mental illness identification model by integrating the transfer learning capabilities of large-scale pre-training models and the feature extraction capabilities of convolutional neural networks. To achieve this, the researchers propose a model called Roberta-Text CNN, which combines the Roberta pre-training model with the Text CNN classifier. This approach allows for the extraction of multi-dimensional feature information for improved identification and classification of mental illnesses. Figure 2 illustrates the structure of the Roberta-Text CNN model.

The input text sequence E is typically generated using one-hot encoding or random encoding. Subsequently, the Encoder component in the bidirectional Transformers of the Roberta pre-training model processes the input sequence, incorporating contextual information from the text. This processing results in the output sequence T. The input for the Text CNN classifier model is denoted as T.

T undergoes a series of processing steps, including convolution, pooling, fully connected layers, and soft max processing in the output layer. This process results in the categorization of T into a label category based on the highest likelihood.

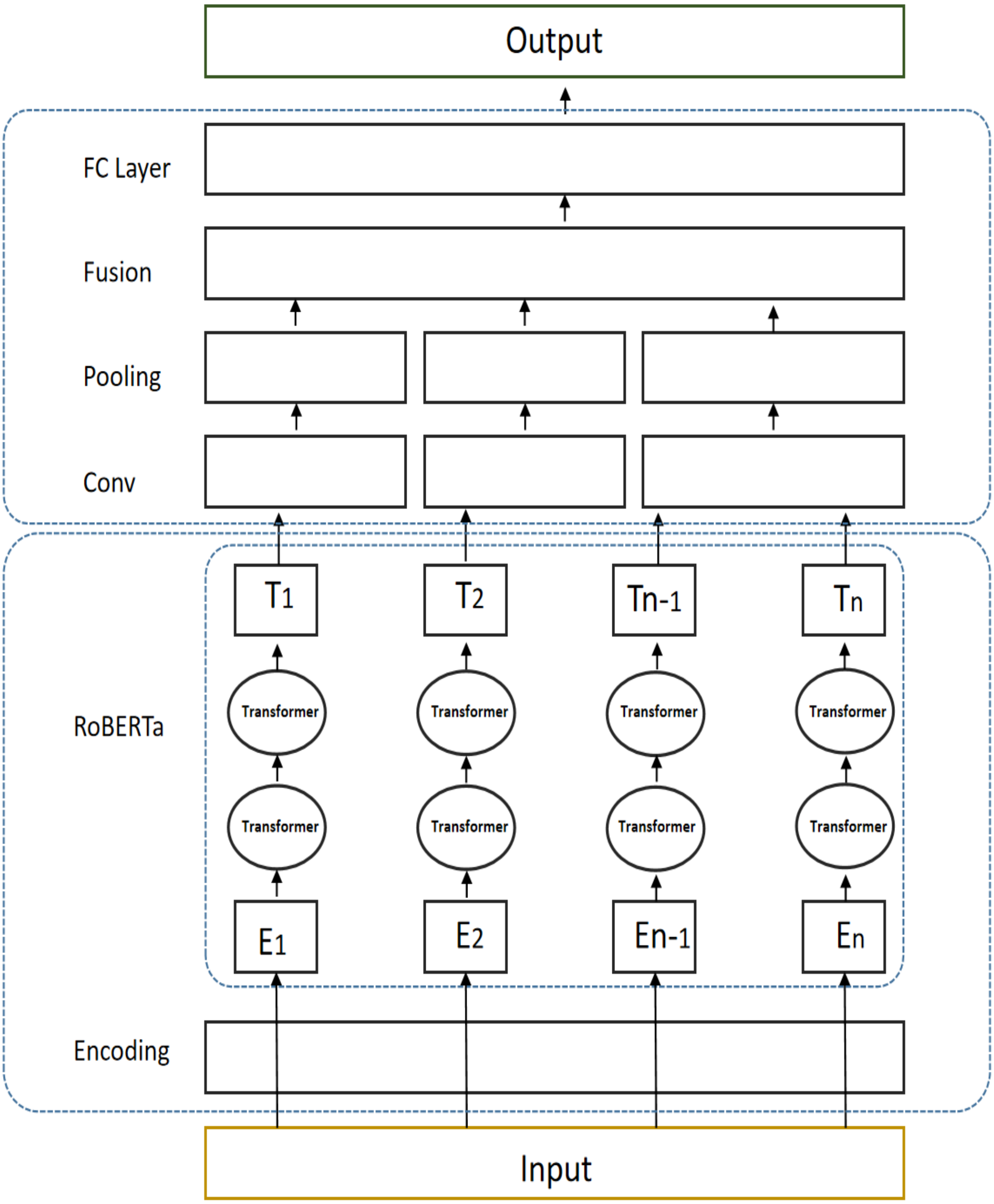

Figure 2: Schematic representation of Roberta-Text CNN model structure.

3.2 EEG signal recognition based on pattern recognition

3.2.1 EEG signals and depressive disorders

Researchers first recorded spontaneous EEG phenomena in rabbit brains in 1875. Later, Professor Berg from Germany recorded the EEG of the cerebral cortex of patients with head trauma by inserting electrodes, and later found that it could also be recorded by attaching electrodes to the outside of the scalp. Since its development, it has been widely used in the fields of medicine and scientific research through the use of specialized EEG collection caps for data recording. Figure 3 shows the international EEG cap electrode placement standards.

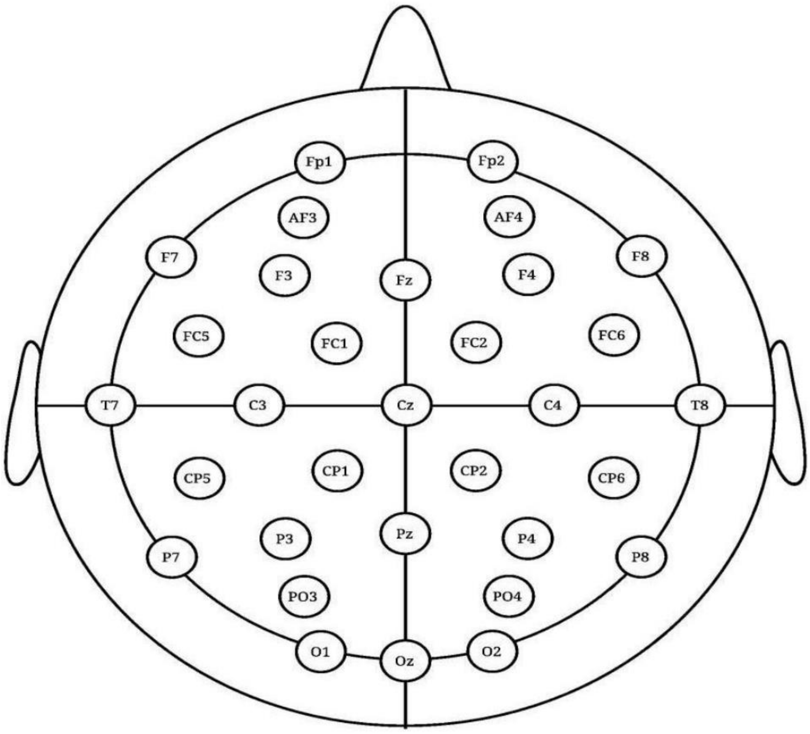

Figure 3: Electrode Placement Standards for the International 10-20 System.

The human brain is spatially divided into a left hemisphere and a right hemisphere. The left hemisphere controls human behavior, and the right hemisphere controls human imagination, spatial thinking and other feelings. According to different functions, the brain can be divided into four functional areas: frontal lobe, parietal lobe, occipital lobe and temporal lobe. Among them, the frontal lobe is responsible for people's emotions and controls people's planned behavior; the parietal lobe is responsible for human touch and limb movement functions, and the part connected with the occipital bone is also responsible for people's speech and speech understanding; the occipital lobe is mainly responsible for visual functions; and the temporal lobe is mainly responsible for visual functions.

The lobes are located on both sides of the human brain, level with the ears, and control hearing and short-term memory functions. The EEG signal itself is generated by the nerves inside the human brain and is transmitted to the scalp through multi-layered structures in the brain such as the cortex, allowing us to non-invasively detect human brain activity on the scalp. This physiological structure makes the EEG signal have the following characteristics: (1) The signal amplitude is weak. Generally, the amplitude of the EEG signal collected outside the scalp is about 50 microvolts, and its upper and lower fluctuation range ranges from a few microvolts to a hundred microvolts. Therefore, when collecting brain electricity non-invasively, the signal is very easily interfered by instruments or other physiological signals from the human body (most commonly myoelectricity and electrooculography).

It is generally believed that signals exceeding ± 100 microvolts are noise signals and can be directly removed by filtering. (2) Strong randomness and non-stationary signals. The brain itself has complex functions, and its internal laws have not yet been fully discovered. Various factors such as differences between individuals and the impact of environmental changes on the brain's processing of information lead to strong randomness in the signal and diverse statistical characteristics of the signal. Characteristics of non-stationary signals. (3) Nonlinear characteristics.

This is caused by the characteristics of organisms themselves. The ability of biological cells to self-regulate and adapt to environmental changes will inevitably affect the generation and propagation of bioelectricity. This change causes brain electricity to exhibit nonlinear characteristics in time and space. Therefore, analysis based on nonlinear characteristics also has important value in the field of EEG research. The EEG frequency range is roughly 0.5~50Hz. After summary and analysis by researchers, it can be divided into five frequency bands, as shown in Figure 4.

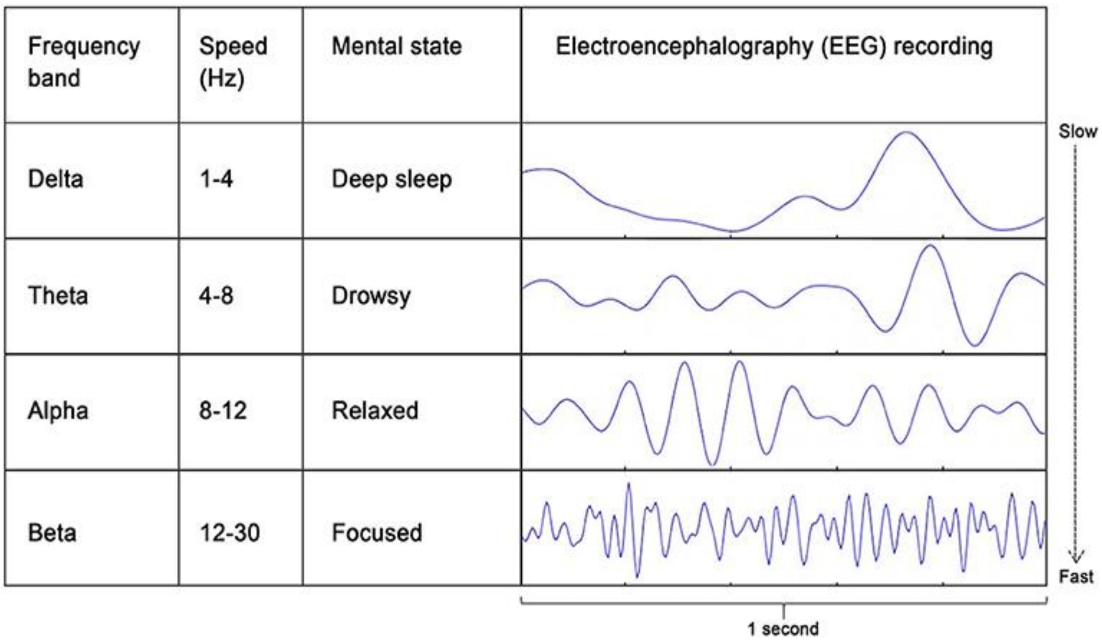

Figure 4: Schematic diagram of EEG waveforms in each frequency band.

Since the generation of depressive disorders may be related to the transmission of signals between the cerebral cortex, and EEG is essentially the sum of potential changes generated during the electrophysiological activity of individual neurons in the cerebral cortex, abnormal changes in the brain can be quantitatively explored through EEG signal changes, which will provide a reliable research basis for the use of EEG to explore the mechanisms of the occurrence of depressive disorders and such psychiatric disorders, resulting in a wide range of attention to this direction of research.

3.2.2 EEG feature extraction

First, the linear features are extracted and the equation for calculating the mean value of the EEG signal is as follows:

$$Mean = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (3)$$

The standard deviation of the EEG signal is then calculated by the following equation:

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)} \quad (4)$$

where μ_x is the mean value of the current signal. The equation for calculating the first order differential mean is as follows:

$$MeanFD = \frac{1}{n-1}\sum_{i=1}^{n-1}|x_{i+1} - x_i| \quad (5)$$

And the equations for calculating the second order difference mean are as follows:

$$MeanSOD = \frac{1}{n-2}\sum_{i=1}^{n-2}|x_{i+2} - x_i| \quad (6)$$

Secondly, we extract the nonlinear features, there are many methods to calculate the fractal dimension, Higuchi's method is widely used because of its fast computation speed and low complexity, and its computational steps are as follows: construct the discrete time series:

$$\{x(i), i = 1,2,\cdots,n\} \quad (7)$$

Reconstructed as the following time series:

$$x_m^k = \left\{x[m], x[m+k],\cdots, x\left[m + int\left(\frac{N-m}{k}\right) \times k\right]\right\} \quad (8)$$

where m is the time starting point, k is the time interval, and $int(\ )$ denotes rounding down. The length of the new sequence is as follows:

$$L_m(k) = \frac{1}{k}\left\{\sum_{i=1}^{int\left(\frac{N-m}{k}\right)} \left| \begin{matrix} x(m+ik) - \\ x(m+(i-1)k) \end{matrix} \right|\right\} \times \frac{N-1}{int\left(\frac{N-m}{k}\right)} \quad (9)$$

Then the average length of the sequence is expressed as:

$$L(k) = \frac{1}{k}\sum_{m=1}^{k} L_m(k) \quad (10)$$

The estimated Higuchi fractal dimension can be calculated as follows:

$$D = \frac{ln(L(k))}{-\, ln(k)} \quad (11)$$

C0 Complexity is a nonlinear feature used to characterize the degree of irregularity of a sequence, and can be used to measure the proportion of irregular parts of a time series. The average amplitude of the power spectrum of the sequence itself is first calculated, and this value is used as a cut-off line, the part of the power spectrum above this value is kept unchanged, and the part below this value is set to zero, and the changed signal is Fourier inverted to obtain a completely new time series.(Jegede, 2016)

The sum of the squares of the new time series and the original sequence is the value of the overall randomized part, and this value is compared with the sum of the squares of the original sequence, and this result is the C0 complexity of the original sequence. Generally speaking, the larger the value of C0, the larger the percentage of irregularities in the EEG signal, and the greater the complexity of the signal, the less regular it is.

3.3 LSTM

The RNN model was proposed at the end of the last century and is now widely used in various research fields. Recurrent neural networks can connect previous information to the current task and can better utilize the temporal information of data than other neural network models. However, this advantage also has certain flaws - the problem of long-term dependencies.

That is, historical information can be used when the distance between the relevant information and the need for the information is close, but the learning ability for longer time information is insufficient. In order to solve the long-term dependency problem of RNN models, H. Sepp and S. Jurgen proposed a new neural network model in 1997 - Long-short-term Memory Network (LSTM). The network consists of three "gate" structures, and its structural frame is shown in Figure 5. Its specific workflow is as follows.

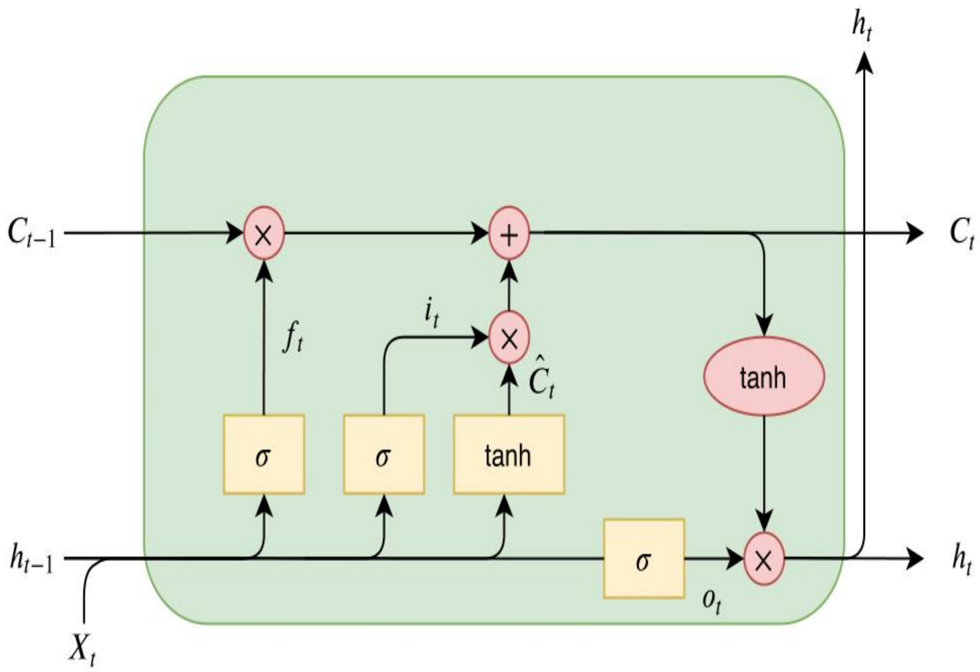

Figure 5: Schematic representation of the LSTM cell structure.

First the network has to decide which useless information is to be deleted. "The Oblivion Gate reads h_{t-1} and x_t, and outputs a value between 0 and 1 for the previous state C_{t-1}, with 1 meaning "fully retained" and 0 meaning "Completely discarded".

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (12)$$

where σ denotes the Sigmod function. The "input gate" then determines what information is added to the network, with the sigmoid layer determining what information is updated and the tanh layer generating alternative alternatives to the update.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (13)$$

$$\hat{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (14)$$

Update the current state C_t:

$$C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t \quad (15)$$

The final output is determined by an "output gate". A sigmoid layer determines which parts to output, then it is processed through a tanh layer and multiplied by the output of the sigmoid layer to get the final result.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (16)$$

$$h_t = o_t \times tanh(C_t) \quad (17)$$

where W is the weight matrix for each gate and b is the offset matrix for each gate.

3.4 Fusion Model

As shown in Figure 6, the deep neural network model proposed in this article fuses EEG and text features. First, consider the output features of Roberta-Text CNN as the first branch network, then fuse the extracted high-dimensional features with the additionally extracted EEG frontal 7-channel serial fusion features, and extract the time domain features again through the LSTM network. Finally, input into the fully connected layer. The fully connected layer here consists of two layers. Since the amount of data in this article's experiment is not very large, in order to prevent premature overfitting from affecting network training, we added a signal loss (dropout) layer to the two fully connected layers. It is hoped to alleviate the over-fitting problem and also reduce the amount of calculation.

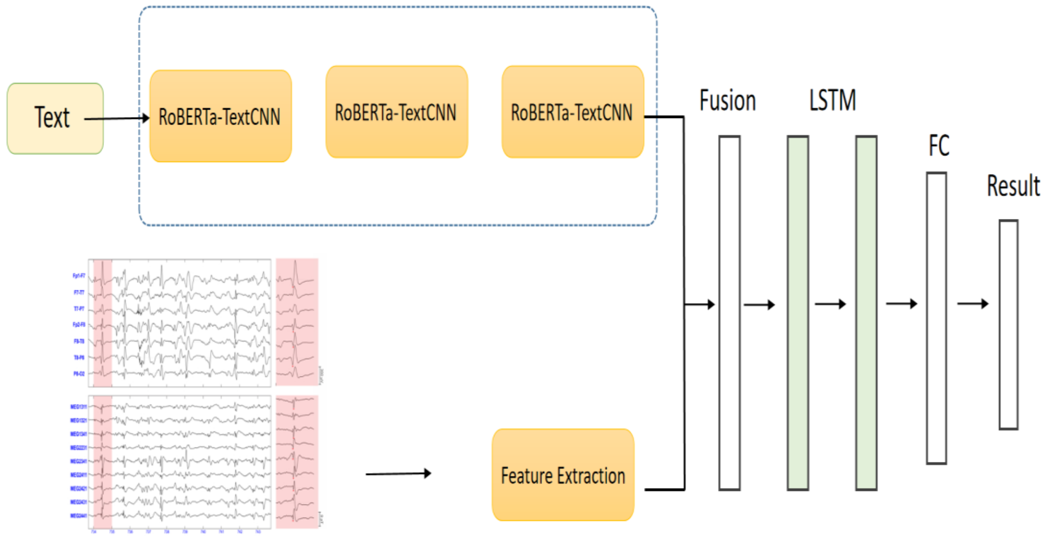

Figure 6: Schematic diagram of the two-branch fusion model.

4. Experiment and Results

4.1 Datasets

(1) The text data in this article uses the efaqa-corpus-zh data set as the sample data for the experiment, and two classification models, Roberta-Text CNN-S1 and Roberta-Text CNN-S2, were established on the labels of mental illness and suicidal tendencies respectively. The experiment in this chapter first uses the Roberta pre-training model to represent the psychological dialogue text as a sentence vector, and then inputs the training data into the Text CNN

model for training, adjusts the parameters of the model, optimizes the classification model, selects evaluation indicators, and finally analyzes and verifies the experimental results.

Feasibility of the model in identifying neighborhoods of mental illness. (2)There are a total of 38 subjects in the EEG data. After processing, there are 912 feature maps in total. Since the experimental data is not very large, this article does not set up a verification set data and directly randomly divides the experimental data into two groups. We randomly selected all the data of 4 subjects from the control group and the experimental group as the test set, and the remaining subjects' data constituted the training set.

4.2 Experimental setup

The experimental setup consists of two parts, the experimental environment setup used during the experiment and the hyperparameter setup of the model used, as shown in Tables 1 and 2.

Table 1: Experimental environment setup.

TYPE	PARAMETERS
OS	Linux
GPU	Nvidia GeForce 2080Ti
RAM	16G
PYTHON	3.6
TRANSFORMERS	3.0.2
PYTORCH	1.4.0
CUDA	9.2

Table 2: Hyper parameter settings

HYPERPARAMETER	VALUE
MAX_LEN	310
BATCH_SIZE	32
DROPOUT	0.5
OPTIMIZER	Adam
LEARNING RATE	1e-4
EPOCH_NUM	50

4.3 Experimental results and analysis

In the model training phase, the time when training is completed is determined by the accuracy and loss value. When the training accuracy is greater and the loss value is smaller, it indicates that the training effect of the model is better. In the initial stage of model training, the accuracy will significantly increase with the increase of epoch value, and the loss value will significantly decrease with the increase of epoch. When epoch exceeds a certain value, the accuracy rate and loss value of model training will tend to Stabilize. Therefore, epoch not only affects the accuracy of the experiment, but also affects the adequacy of model training and training efficiency. This chapter conducts experiments with different training rounds on mental illness S1 and suicidal tendency S2, and draws an easy-to-observe curve chart based on the accuracy and loss value results of the experiment. Figure 7 is the Accuracy Epoch curve of the Roberta-Text CNN model experiment on mental illness labels, and Figure 8 is the Loss-Epoch curve. As can be seen from the figure, when epoch reaches 10, the model training tends to be stable, the accuracy reaches 0.901, and the loss value reaches 0.4.

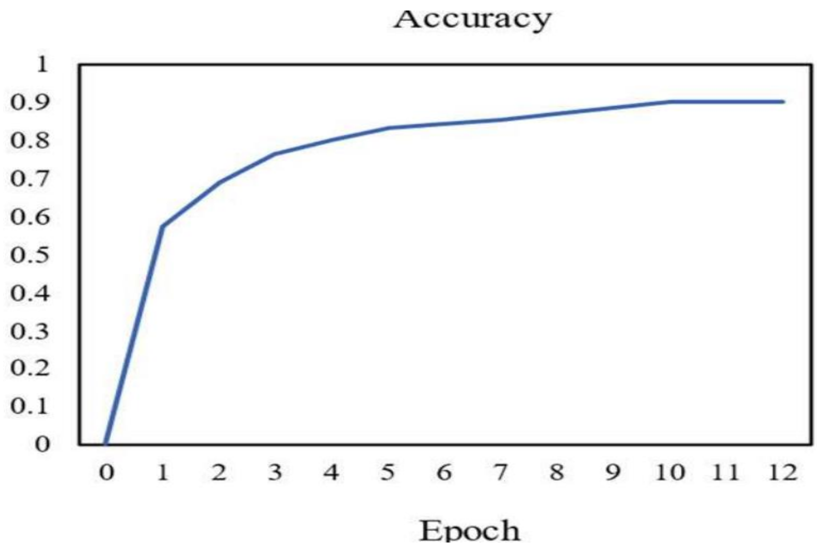

Figure 7: S1-Accuracy-Epoch.

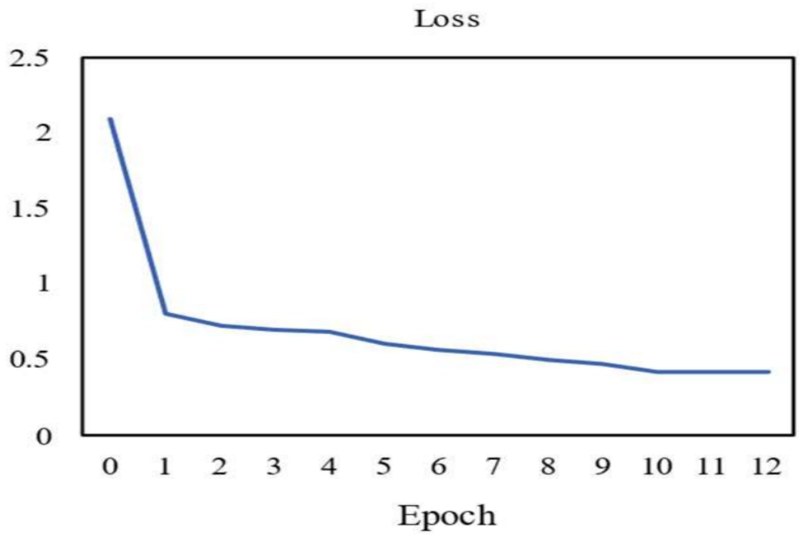

Figure 8: S1-Accuracy-Loss.

When the Roberta-Text CNN model was used to classify mental diseases on the data set, it achieved good results in 8 categories of mental diseases, and the overall accuracy of the model was 0.8892. The precision, recall, and f1-scorce under the calculation rules of micro-average and weighted average are shown in Table 3.

Table 3: Experimental results of Roberta-Text CNN-S1.

MODEL	PRECISION	RECALL	F1-SCORE
MACRO AVG	0.8496	0.8295	0.8395
WEIGHTED AVG	0.8611	0.8893	0.8733

When the Roberta-Text CNN model was used to classify suicidal tendencies on the data set, all six suicidal tendencies categories achieved good results, and the overall accuracy of the model was 0.9711. The precision, recall, and f1-score under the calculation rules of micro-average and weighted average are shown in Table 4.

Table 4: Experimental results of Roberta-Text CNN-S2.

MODEL	PRECISION	RECALL	F1-SCORE
MACRO AVG	0.9535	0.9455	0.9490
WEIGHTED AVG	0.9656	0.9711	0.9682

After using EEG data, Figure 9 shows the average distribution of the first-order difference mean characteristics of different groups of people. It is obviously found that the characteristics of the prefrontal area of people with depressive disorders are higher than those of normal people in this area, which is reflected on electrodes E27, E23, E9, E2 and E123. At the posterior frontal E119 electrode, there was a higher value in people with depressive disorders. At the same time, in the middle of the occipital lobe E59 and E66 areas, people with depression affected by these two electrodes have significantly darker colours in this area than normal people.

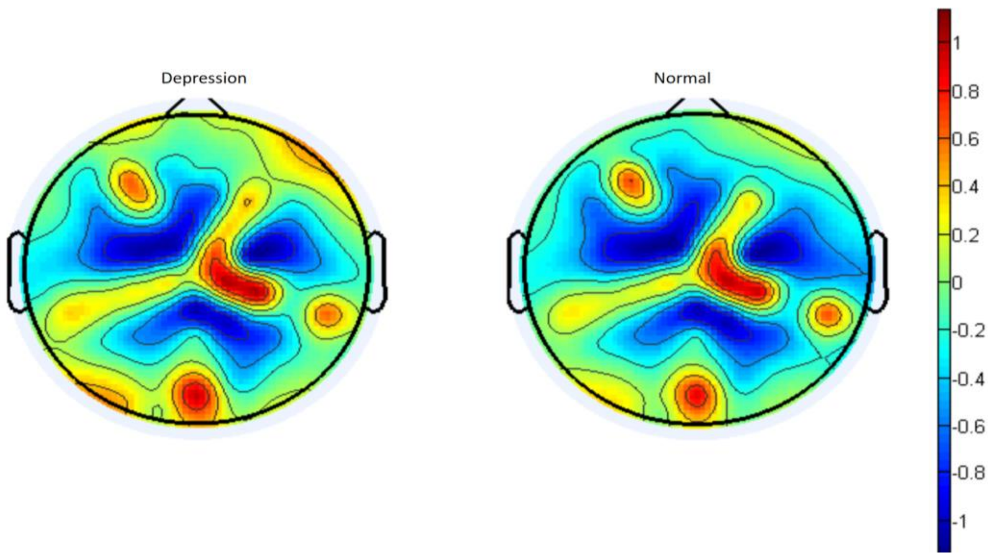

Figure 9: Depression identification results using EEG data.

5. Conclusion

In this paper, we propose a mental illness recognition model based on Roberta-Text CNN. Based on Roberta Text CNN model, we build text classification models for two types of labels, namely, mental illness and suicidal tendency, on the efaqa-corpus-zh psychological dataset, and carry out comparative experiments with several mainstream models. The experimental results show that the Roberta-Text CNN model is better than other models, and the accuracy of the model reaches 0.903 and 0.978 on the two classification tasks of mental illness and suicidal tendency, respectively. The experiments demonstrate the effectiveness and feasibility of the Roberta-Text CNN model in the field of mental illness recognition. In addition, this paper also introduces the EEG signals using the Mel's inverted spectral coefficient for the scale change to organize them into a dimensionally consistent matrix, which is then inputted into the convolutional neural network as image features for further extraction of high-dimensional features and serial fusion of EEG features and then inputted into the long and short-term memory network. The application of the pattern recognition technology provides a two-branch mental illness recognition model, which further enhances the effectiveness and feasibility of the screening and prevention of early mental illnesses in athletes. The application of pattern recognition technology provides a two-branch mental illness identification model, which further enhances the ability of screening and prevention of early mental illness in athletes.

Reference

Alexandre, C. M.-A., Ricardo, V. d.-S., Daniel, A.-B., Dumitriu, Z.-S., & Salvador, J. d.-T. A. A. (2018). Lessons Learned After 366 Thermoablated Veins. *Vascular & Endovascular Review*, *1*.

Ameer, I., Arif, M., Sidorov, G., Gòmez-Adorno, H., & Gelbukh, A. (2022). Mental illness classification on social media texts using deep learning and transfer learning. *arXiv preprint arXiv:2207.01012*.

Bonci, C. M., Bonci, L. J., Granger, L. R., Johnson, C. L., Malina, R. M., Milne, L. W., Ryan, R. R., & Vanderbunt, E. M. (2008). National athletic trainers' association position statement: preventing, detecting, and managing disordered eating in athletes. *Journal of athletic training*, *43*(1), 80-108.

Chen, J., Han, P., Zhang, Y., You, T., & Zheng, P. (2023). Scheduling energy consumption-constrained workflows in heterogeneous multi-processor embedded systems. *Journal of Systems Architecture*, *142*, 102938.

Chen, J., Li, T., Zhang, Y., You, T., Lu, Y., Tiwari, P., & Kumar, N. (2023). Global-and-local attention-based reinforcement learning for cooperative behaviour control of multiple UAVs. *IEEE Transactions on Vehicular Technology*.

Cho, G., Yim, J., Choi, Y., Ko, J., & Lee, S.-H. (2019). Review of machine learning algorithms for diagnosing mental illness. *Psychiatry* *investigation*, *16*(4), 262.

Glick, D., & Applbaum, K. (2010). Dangerous noncompliance: a narrative analysis of a CNN special investigation of mental illness. *Anthropology & Medicine*, *17*(2), 229-244.

Groffik, D., Mitáš, J., Jakubec, L., Svozil, Z., & Frömel, K. (2020). Adolescents' physical activity in education systems varying in the number of weekly physical education lessons. *Research quarterly for exercise and sport*, *91*(4), 551-561.

Jegede, C. O. (2016). The indigenous medical knowledge systems, perceptions and treatment of mental illness among the Yoruba of Nigeria. *Studies in Sociology of Science*, *7*(5), 12-20.

Jiayao, C. (2022). Evaluation of Causes and Impacts of Emotional Pressure Among Teenagers. 2021 International Conference on Social Development and Media Communication (SDMC 2021),

Khare, S. K., Bajaj, V., & Acharya, U. R. (2021). SPWVD-CNN for automated detection of schizophrenia patients using EEG signals. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-9.

Li, Y., & Cao, J. (2023). Adaptive binary particle swarm optimization for WSN node optimal deployment algorithm. *IECE Transactions on Internet of Things*, *1*(1), 1-8.

Marriwala, N., & Chaudhary, D. (2023). A hybrid model for depression detection using deep learning. *Measurement: Sensors*, *25*, 100587.

Ning, X., Wang, Y., Tian, W., & Liu, L. (2023). A Biomimetic Covering Learning Method Based on Principle of Homology Continuity. *IECE Transactions on Neural Computing*, *1*(1), 1-13.

Osman, M. E. (2020). Global impact of COVID-19 on education systems: the emergency remote teaching at Sultan Qaboos University. *Journal of education for teaching*, *46*(4), 463-471.

Park, S.-C., & Kim, D. (2020). The centrality of depression and anxiety symptoms in major depressive disorder determined using a network analysis. *Journal of affective disorders*, *271*, 19-26.

Rajendran, S., Gandhi, R., Smruthi, S., Chaudhari, S., & Kumar, S. (2023). Diagnosis of Mental Illness Using Deep Learning: A Survey. In *Artificial Intelligence for Societal Issues* (pp. 223-244). Springer.

Santomauro, D. F., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., Abbafati, C., Adolph, C., Amlag, J. O., & Aravkin, A. Y. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*, *398*(10312), 1700-1712.

Shoeibi, A., Sadeghi, D., Moridian, P., Ghassemi, N., Heras, J., Alizadehsani, R., Khadem, A., Kong, Y., Nahavandi, S., & Zhang, Y.-D. (2021). Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models. *Frontiers in neuroinformatics*, *15*, 777977.

Twenge, J. M., & Joiner, T. E. (2020). US Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. *Depression and anxiety*, *37*(10), 954-956.

Vázquez-Abad, F. J., Bernabel, S., Dufresne, D., Sood, R., Ward, T., & Amen, D. (2020). Deep learning for mental illness detection using brain SPECT imaging. Medical Imaging and Computer-Aided Diagnosis: Proceeding of 2020 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2020),

Xu, Y., & Li, Y. (2019). The Fate of the Superior Mesenteric Artery in Fenestrated Endovascular Repair of Complex Abdominal Aortic Aneurysms. *Vascular & Endovascular Review*, *2*. https://doi.org/https://doi.org/10.15420/ver.2019.1.1