Research on Stress Reduction Model Based on Transformer

Xin Xu[1], Yikun Zhao[1*], Ruhao Zhang[1], and Tingting Xu[1]
[1] School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications
Nanjing, Jiangsu 210003 China
[e-mail: 15036373520@163.com]
[*]Corresponding author: Yikun Zhao

Received June 1, 2022; revised September 19, 2022; accepted November 22, 2022; published December 31, 2022

Abstract

People are constantly exposed to stress and anxiety environment, which could contribute to a variety of psychological and physical health problems. Therefore, it is particularly important to identify psychological stress in time and to find a feasible and universal method of stress reduction. This research investigated the influence of different music, such as relaxation music and natural rhythm music, on stress relief based on Electroencephalogram signals. Mental arithmetic test was implemented to create a stressful environment. 23 participants performed the mental arithmetic test with and without music respectively, while their Electroencephalogram signal was recorded. The effect of music on stress relief was verified through stress test questionnaires, including Trait Anxiety Inventory (STAI-6) and Self-Stress Assessment. There was a significant change in the stress test questionnaire values with and without music according to paired t-test ($p<0.01$). Furthermore, a model based on Transformer for stress level classification from Electroencephalogram signal was proposed. Experimental results showed that the method of listening to relaxation music and natural rhythm music achieved the effect of reducing psychological stress and the proposed model yielded a promising accuracy in classifying the Electroencephalogram signal of mental stress.

Keywords: EEG, mental stress, music, self-attention, Transformer

This research was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX21_0752.

http://doi.org/10.3837/tiis.2022.12.009 **ISSN : 1976-7277**

1. Introduction

Life maintains a complex dynamic equilibrium, called homeostasis, which is constantly under shock from adverse pressures, internal or external, real or perceived [1]. The human body's physiology and psychology will respond accordingly to such changes, which is called stress when balance, equilibrium, or harmony is threatened [2]. When people face stress, the human body regulates through the complex physiological and behavioral central nervous system, but if the stress is excessive or the duration is too long, it can facilitate the development of diseases. Researches show that chronic exposure to stressful circumstances leads to a certain degree of atrophy of the hippocampus, leading to lapse of memory, in addition to increase susceptibility of people exposed to the virus to upper respiratory infections [3].

There are many stress management techniques, such as meditation, yoga, sleep, deep breathing, watching movies, massage, music, sports and playing with friends, which help relieve stress. Music listening is increasingly recognized as an effective stress management method. Regular music listening has potentially positive effects on stress relief [4]. The experimental results from previous studies have showed that music is able to relieve stress [5]. Calming music can reduce tension and have a relaxing effect [6]. Music that triggers autonomous sensory meridian response (ASMR) and relaxing sounds can lower people's stress levels [5]. An analysis of 22 experimental studies demonstrated that listening to music could significantly reduce stress-induced arousal [7]. Research has also showed the effectiveness of listening to music on improving alpha activity and hence relaxation [8].

There are two test methods, one is the stress perception scale, the other is to measure stress through physiological changes. Physiologically, pressure can be measured by skin conductance level galvanic skin response (GSR) [9], Electroencephalogram (EEG) [10, 11], blood pressure (BP) [12], electromyography [13], Blood Volume Pulse (BVP), heart rate variability (HRV) [14], skin temperature (ST), Respiration, blink rate and pupil dilation. Common stress perception scales include Cohens's Perceived Stress Scale (PSS) [15], holmes rahe stress inventory [16], state-trait anxiety inventory (STAI) [17], NASA Task Load Index rating scale (NASA-TLX), Stress Response Inventory (SRI) [18] and Dundee Stress State Questionnaire. Most of the existing studies employed psychological tests to measure stress, such as scales and questionnaires, which have a lot of subjective factors and take a long time. Besides, this type of stress assessment method requires well-trained personnel to process the assessment, and the physiological indicators are not easy to detect.

Pressure monitoring by EEG has the advantages of safety, portability, low cost and short time consumption. EEG data showed significant differences, even when the subjects were performing the same tasks under the same conditions, due to different brain conditions in the subjects, including levels of stress, mood, concentration, and fatigue [19]. Brain is the central part of stress generation and its spectral features are closely related to different psychophysiological conditions. As a result, researches believes that the power of EEG bands, such as alpha, beta, delta, theta and gamma, is the most credible indicators when detecting and analyzing stress [20, 21]. Power spectrum has been widely applied in the analysis of stress [22]. Alpha and beta waves are markers of pressure levels [23], and alpha band asymmetry and relative gamma power can be used to assess pressure [24].

Recently, machine learning, especially deep-learning based stress analysis using EEG signals has attracted significant attention. In [25], multi-layer perceptron (MLP), Naive Bayes (NB), and support vector machine (SVM) were used for stress classification and an accuracy of 92.85% and 64.28% is achieved in the two categories and the three categories of stress. In the study of [26], long-term stress was classified using support vector machine based on resting state EEG signal. The alpha asymmetry was calculated, which was used as feature for classification with a classification accuracy of 85.20%. In [27], authors developed an EEG-based pressure recognition system by convolutional deep learning neural network (deep CNN) and Fully Connected Deep Neural Network and with an average recognition accuracy of 86.62%. In [28], it was showed that a multiclass two-layer Long Short Term Memory (LSTM) Recurrent Neural Network (RNN) classifier had good effect on stress recognition and with an accuracy of 93.27%.

The stress state occurs over a period of time, rather than a transient response, and there may be a connection between pulses that occur within a short time period. It would be nice if the model used for stress state identification also takes into account the past under the circumstances. CNNs are local networks decided by the kernel size and respective step size, and LSTMs may be unable to consider this long-term dependency because of the forgetting factor. However, the attention mechanism used in the transformer is a mechanism to freely select contextual information which is a reference [29]. The key of attention mechanism in Transformer models dependencies, and it is not constrained by long distances in the sequence [30]. On the other hand, the convolution operation destroys the spatial properties of EEG signals due to the mixed multiplication and addition. Research suggests that it may be more beneficial to employ the multi-head attention on spatial ordinates to include representational similarities between regions [31]. The network architecture Transformer, completely applies the attention mechanism to map the global dependencies between inputs and outputs, avoiding recursion. It is entirely based on the attention mechanism, and self-attention associates different positions of a sequence to compute a representation of that sequence. Intrinsically, it is a mechanism for measuring similarity [29].

Transformer was first used in encoder-decoder in machine translation fields to compress all necessary information into a vector of length X. It requires significantly less training time and it has more parallelism [32]. The model has proven to be a simple and extensible framework, and is extensively applied in Natural Language Processing fields, for example machine translation, question answering system, text summarization and speech recognition. In addition, the great success in natural language processing has been explored in the computer vision. It is becoming a more general framework for learning sequence data.

The study mainly explores the stress-relieving effects of different music by analyzing of EEG and to verify the function of music on stress relief through stress test questionnaires. A model based on Transformer is proposed to classify EEG signals at different levels of the mental stress. The remainder of this paper is structured as follows. First of all, Section 2 presents the specific methodology, which discusses methodologies for data acquisition, signal processing and classification. Section 3 discusses the experimental results of stress classification in response to different music as well as the statistical analysis. Finally, Section 4 of this paper makes a conclusion.

2. Method

Fig. 1 illustrates the overview of a model based on Transformer for stress level classification from EEG signal. As the first step, the pressure experiment was carried out in the laboratory. EEG signals are recorded using the EEG signal acquisition equipment during the experiment to obtain the original stress experiment data. Then, as the second step, denoising the raw data by MATLAB to reduce noises and artifacts of the EEG signal to improve signal quality. The preprocessed data were normalized and divided to form a multidimensional matrix, which was used as the input of the psychological stress classifier model. As the last step, the model based on Transformer was used to classify data set.

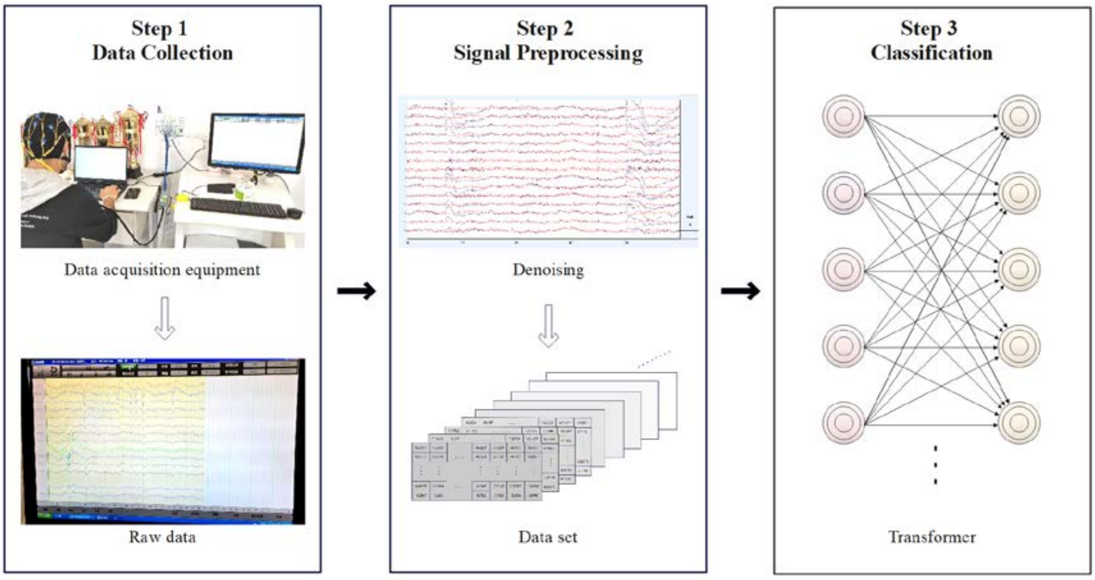

Fig. 1. The overview of a model based on Transformer for stress level classification from EEG signal

2.1 Experiment Procedure

The EEG signals were collected from 23 healthy right-handed subjects recruited from the university (age ranged: 18-24, 6 females and 17 males). Participants were asked to get enough sleep the day before the experiment and were forbidden from drinking alcohol, coffee, tea and smoking. The experimental procedure was introduced to the subject and informed consent was obtained.

The EEG signals were recorded using the EEG cap of GREENTEK and Neuroscan 64 lead electroencephalograph device on Scan 4.5 acquisition software platform. The sampling frequency is 512 Hz. Before data recording, conductive paste was applied to the subject's EEG cap electrodes while the impedance was monitored to keep it below 10kΩ. The electrodes were placed using the international 10-20 system, and the schematic diagram is demonstrated in **Fig. 2**.

The experiment was designed on E-Prime. Cognitive stimulation methods include visual stimulation, auditory stimulation and so on. Research shows that more than 70% of human perception of the external environment comes from vision system [33]. In this experiment, mental stress was induced by mental arithmetic test, and the data was collected with the aid of stress test questionnaire. First, the experiment was introduced to the subjects to help them get familiar with the experimental environment in the introductory section. Second, the subjects solved sample problems in training section to familiarize them with the experimental process before the formal experiment began. The final part was the experiment section after the break. The introductory section, the training section and the resting section prevent the subjects from being prone to wrong answers and high pressure at the beginning because they are nervous when they enter an unfamiliar environment, and having high accuracy and low stress in later experiments due to adaptation to the environment.

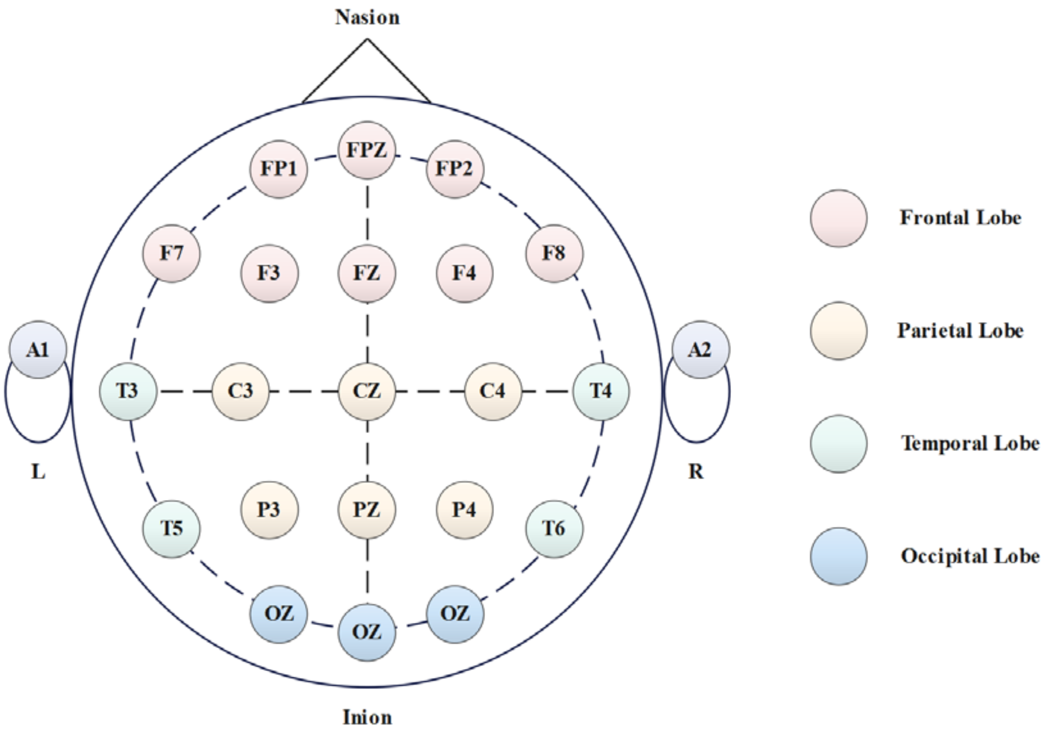

Fig. 2. 10-20 system

Experiment section was divided into four parts, task1, task2, task3 and task4. The subjects were asked to remain in a resting state in task1. The subjects were required to complete the mental arithmetic test without music in task 2. In the mental arithmetic test, they were required to complete the task quickly and accurately. Every arithmetic question commenced with a fixation sign (+). In the mental arithmetic test, mixed operations of addition, subtraction, multiplication and division were performed such as $39 + 12 \times 4$. The questions were randomly generated and had a numerical value from 1 to 100.

The feedback displayed “correct”, “incorrect” or “no response” depending on correct response, incorrect response or no response to the question. The time limit was 20% less than the time limit in the training section. By completing the task within a limited time in this way, the subjects were guided to generate time pressure. Task 2 and task 3 required participants to finish mental arithmetic test with music I and music II respectively. Music II is relaxing music (Weightless, by Marconi Union), and music II is natural rhyming music (ocean, forest, rainy day, flowing water, summer night), which were selected by the subjects themselves in order to get better decompression effect.

Immediately after completion of each task, the subjects were required to complete the Trait Anxiety Scale (STAI-6) and the Self-assessment of Stress, to provide feedback based on their mental stress level just now. Subjects rated the intensity of the stress state from 0 to 9, with 0 representing no stress and 9 representing high stress in the Self-assessment of Stress. Subjects were asked to complete 12 questions, with higher scores indicating greater stress in the Trait Anxiety Scale (STAI-6). The mental stress level experienced by the subject in different tasks was verified by the stress level determined in the self-assessment questionnaire.

Fig. 3 shows the detailed experimental process.

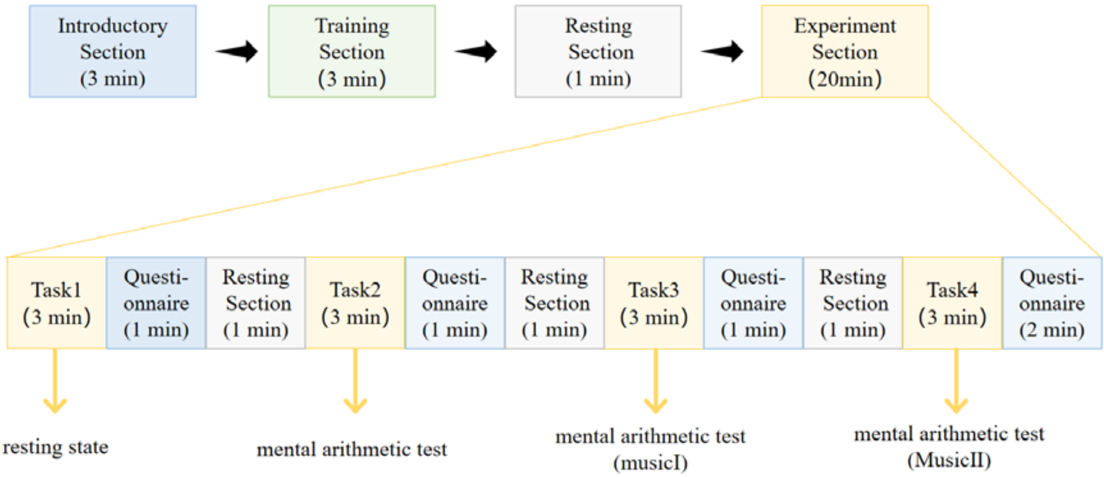

Fig. 3. Experimental process

2.2 Signal Preprocessing

EEG signals are contaminated with noises and artefacts that are not originated from the brain. Since cardiac activity, eye and muscle activity signals, and power line noise are also recorded by EEG equipment, which reduces the quality of the data, it is necessary to use denoising methods to remove these noises.

A total of 19 electrodes were selected: F3, F4, F7, F8, Fp1, Fp2, Fz in the frontal lobe, C3, C4, Cz in the central lobe, P3, P4, Pz in the parietal lobe region, T3, T4, T5, T6 in the temporal lobe and O1, O2 in the occipital region. EEG preprocessing and labeling were carried out by EEGLAB toolbox of MATLAB. First of all, the EEG signal was bandpass filtered with cutoff frequencies of 0.5 Hz and 30 Hz. afterwards, the artefacts were removed by independent components analysis technique (ICA). Electro-oculogram (EOG), electromyography (EMG) and electrocatdiogram (ECG) artifacts were separated from EEG signals by ICA, and the putative source signals were separated from the brain to obtain pure induced EEG signals.

EEG signals from 23 students were collected. Each experiment includes four tasks, each task was 180s, and we took 150s data in the middle for processing. Each 1s segment is taken as a sample, which is a 19*128 matrix. The input of the classification model is a matrix of data from all subjects. Before dividing the data set, we randomly scrambled the samples of each subject, and then normalized them. Dividing 80% of the data into the training set and the rest into the test set. The division of training sets and test sets was done randomly and there was no overlap between them.

2.3 Model Based on Transformer

A stress classification model based on Transformer is presented in this research to make it suitable for EEG signal, which is based entirely on attention mechanisms and completely abandons recursion and convolution. Model architecture is shown in **Fig. 4**. The model follows this overall architecture using embedding layer, positional encoding, Transformer Encoder and fully connected layer. And Transformer Encoder consists of N blocks with the same structure.

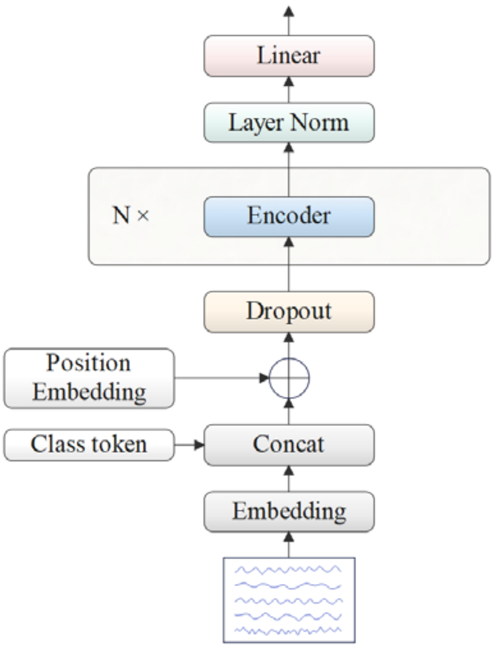

Fig. 4. Model architecture

The first part is the embedding layer. The EEG data of each channel is regarded as a patch, which are input to embedding layer to obtain the corresponding token embedding.

The second part is positional encoding. First of all, generating the class token, which is used for classification, referring to the bert network. Among them, class token are a learnable parameters, and the dimension is the same as that of the token.

There is no recursion and convolution used in this model, and it is completely replaced by the attention mechanism. Therefore, the model does not have the ability to learn sequence information like RNN. For leveraging the order and spatial structure of EEG signals, it is necessary to actively provide sequence information to the model to help it learn the position information. Therefore, the model injects some information about the relative or absolute position of the tokens in the EEG signals, that is, the "positional encodings" is added to input embeddings. The positional encodings and embeddings have the same dimension d_{model}, so that they can be added together [32]. Gaussian embedder is used in this paper and the definition of Gaussian embedder is

$$\psi(t,\ x) = \exp\left(- \frac{\left\| t - x \right\|^{2}}{2\sigma^{2}} \right) \tag{1}$$

Where σ is the standard deviation. The Gaussian embedder is also approximately bandlimited like the square embedder. However, the Gaussian embedder has a higher upper bound for the stable rank that can be controlled by σ.

Then, the token embedding and the positional encoding are added to force the subsequent linear layers to learn the temporal relationship.

Next is a dropout layer and then input the class token, token and position embedding are input into the transformer encoder together.

The third part is Transformer Encoder. Structure diagram of Transformer Encoder is demonstrated in **Fig. 5**.

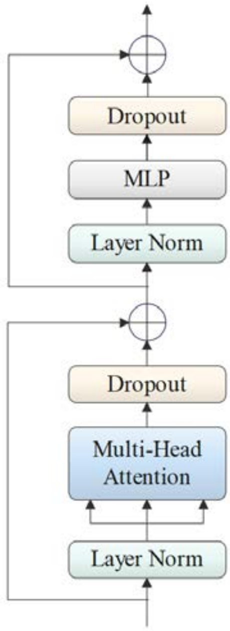

Fig. 5. Structure diagram of Transformer Encoder

And then there are N blocks with the same structure. Each block contains two sublayers, which can be represented by a formula. Where Sublayer (x) represent the output through the sublayer. The first sublayer includes layer norm, muliti-head attention, dropout layer and residual connection. The second sublayer includes layer norm, feed-forward network, dropout layer and residual connection. The residual connection is used to connect the input and output together. They need to be the same as the input and output, so the dimension of each layer is set as a consistent d_{model}. This is different from CNN, which generally reduces the spatial dimension and increases the channel dimension. There are only two parameters, N and d_{model}.

$$x + Sublayer\left(x\right) \tag{2}$$

The layer norm in Transformer Encoder is designed to prevent overfitting. The formula is as follows.

$$y = \frac{x - E[x]}{\sqrt{Var[x] + \epsilon}} * \gamma + \beta \tag{3}$$

Where γ and β are learnable parameters.

The next layer is a multi-head attention which is the kernel. Where each head independently generates a different query Q, keys K and values V, where they are all tensors. Multi-head attention transforms the query, key and value by learning different projection matrices, and then sends the transformed query, key and value to the attention layer in parallel. Last, the outputs that attention are combined and transformed by a linear projection that can be learned to produce the output. Specifically, the output is the weighted sum of the valve, so the output dimension and the value dimension are consistent. The weight of each value is calculated from the similarity between the key and the query corresponding to the value. As the query changes, the weights will be different, and the output will be different. That is, a set of tensors query Q, keys K and values V go through different fully connected layers, then through the attention layer, and then splicing the output of the attention layer, and finally passing through a fully connected layer. Parallel attention layers of h = 4 are used in this work.

$$MultiHead\left(Q,\ K,\ V\right) \ = \ Concat\left(head_1,\ \ldots,\ head_h\right)W^O \tag{4}$$

$$head_i = Attention\left(QW_i^Q,\ KW_i^K,\ VW_i^V\right) \tag{5}$$

Where the projections are parameter matrices $W_i^Q \in \mathbb{R}^{d_{model} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{model} \times d_k}$, $W_i^V \in \mathbb{R}^{d_{model} \times d_v}$ and $W^o \in \mathbb{R}^{hd_v \times d_{model}}$.

The multi-head attention pay close attention to information in diverse representation subspaces at different locations, thus realizing the average calculation with one attention. Multi-head attention achieves a global attention extraction and expands the model's ability to focus on different positions. Apart from this, the model has excellent parallel processing ability. **Fig. 6** illustrates the structure diagram of multi-head attention.

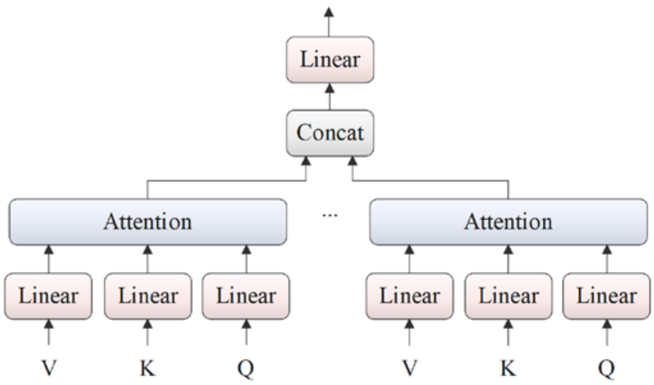

Fig. 6. Structure diagram of multi-head attention

Multi-head attention is composed consists of several attentions, that is, independently computed multiple times. An attention function implements mapping a query and a set of key-value pairs to an output, which is calculated as a weighted sum of the values. The weight assigned to each value is calculated by a compatibility function of the query with the corresponding key [31]. Namely, the self-attention layer uses the query Q and the keys K to construct the correlation coefficient, then multipy the values V are by the correlation coefficient, and finally sum it up, which is input to the next layer.

The cognitive behavior of the human brain does not occur in an instantaneous response. CNN is difficult to model long sequences. If the sampling points are far apart, complex operations are required to connect them. And the convolution operation destroys the spatial properties of EEG signals. However, if attention mechanism is used, all data can be seen at the same time. The attention mechanism is a mechanism to freely select contextual information, which is a mechanism for measuring similarity.

The following formula shows the attention mechanism. Among them, queries and keys are of equal length, which is d_k, and the length of values is d_v. Finishing the inner product of each query and key, as the similarity, and then dividing the length of the vector $\sqrt{d_k}$. Then geting weights using softmax, and the sum of the weights are 1. Because matrix multiplication is used, it can be computed in parallel. In contrast, RNN is calculated step by step, so it is difficult to parallelize the operation, which takes a long time. And if the information is relatively long, the earlier information may be lost, or may occupy a large amount of memory

$$Attention(Q,\ K,\ V) = soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{6}$$

The structure of the MLP (Multilayer Perceptron) layer in Transformer Encoder is shown in Fig. 7. MLP layer includes a fully connected layer, an activation layer, a dropout layer, a fully connected layer, and a dropout layer. Among them, the parameters are different at each layer.

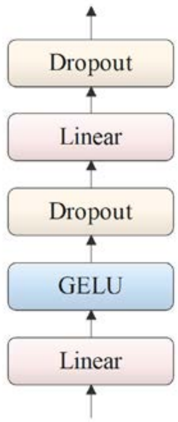

Fig. 7. Structure diagram of MLP

$$FFN(x) = max(0,\ xW_1 + b)W_2 + b_2 \tag{7}$$

The activation function selects GELU (gaussian error linear units), which could avoid the problem of gradient disappearance. The formula is as follows.

$$\begin{aligned} GELU(x) &= x * \Phi(x) \\ &= x * \int_{-\infty}^{x} \frac{e^{-\frac{(X-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}\, dX \end{aligned} \tag{8}$$

Where Φ (x) represents the cumulative probability distribution of the Gaussian distribution.

The fourth part is the fully connected layer. The output of the class token is extracted, and the classification result is obtained through the fully connected layer.

Cross entropy was used as the loss function for the model in this research and Adam optimizer was used to train the model. The cross entropy loss function formula is as follows.

$$H(p(x),q(x)) = -\sum_i p(x_i)\log(q(x_i)) \tag{9}$$

Where p (x) is the true distribution and q (x) is the predicted distribution.

The formula for the Adam optimizer is as follows.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{10}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{11}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{12}$$

$$W_{t+1} = W_t - \frac{\eta}{\sqrt{\hat{v}_t}+\epsilon}\hat{m}_t \tag{13}$$

Where (10) represents the estimation of the first-order moment of the gradient. Equation (11) represents the estimation of the second-order moment of the gradient. Equation (12) represents the correction of the estimation of the first-order and second-order moment estimation. Equation (13) represents the dynamic constraint on the learning rate. β is the learning rate, t is the number of iterations, and v is the momentum variable.

3. Result and Discussion

3.1 Statistical Analysis

Statistical analysis was carried out for group I and group II respectively. Paired t-test is performed respectively based on the Self-assessment of Stress and Trait Anxiety Scale before and after listening to the music track. The results shown in **Table 1**.

The mean, standard deviation and p-values based on the Self-assessment of Stress and Trait Anxiety Scale are shown in the table. As you can see from the table, the stress value of listening to music I was 81.75% and 90.08% of the before, respectively. Similarly, the stress of listening to music II was 75.18% and 88.63%. It can be seen from the experimental results that both relaxing music and natural rhythm music have a certain decompression effect.

The significance level selected in this research is 0.05 in the paired t-test. The p-values show that there was significant change in stress before and after listening to music. In other words, subjects experienced a significant reduction in stress by listening to Music I and Music II.

Table 2 shows the average response time and response accuracy of all subjects under different tasks. The results show that subjects achieved shorter average response time and higher response accuracy by listening to Music I and Music II compared with the results from no music condition. This is because the higher the stress, the more difficult it is for people to concentrate when performing a task, which leads to significantly increased task completion time. This result is consistent with previous literatures that stress levels can be reduced by listening to music.

Table 1. Differences among the subjects of music I group and music II group based on Trait Anxiety Scale and Self-assessment of Stress scores before and after listening music tracks by paired t-test

		music I group		music II group	
		Self-assessment of Stress	Trait Anxiety Scale	Self-assessment of Stress	Trait Anxiety Scale
Mean	Before	5.96	29.83	5.96	29.83
	After	4.87	26.87	4.48	26.43
Standard Deviation	Before	1.40	5.56	1.40	5.56
	After	1.42	5.37	1.90	6.45
p-values		0.006	0.041	0.004	0.018

Table 2. The average response time and response accuracy of different tasks

	Task 2 (without music)	Task 3 (music Ⅰ)	Task 4 (music Ⅱ)
Average Response Time	4.42	3.69	4.38
Response Accuracy	72.42%	74.18%	73.81%

3.2 Experimental Results

Table 3 displays the classification accuracy using model based on Transformer. 92.70% accuracy was obtained in the binary classification of task 1 and task 2 in the experimental dataset using the transformer-based model. And we obtained 62.05% and 64.74% accuracy in the binary classification of tasks 2 and 3, and 2 and 4, respectively.

We further validated the effectiveness of the proposed model on a public dataset 'EEG During Mental Arithmetic Tasks' [34]. The database includes EEG signals of subjects before and during the mental arithmetic tasks. 81.09% and 74.67% accuracy were achieved based on group B and group G data from the dataset, respectively. This shows that the model has also achieved good classification results on other EEG datasets. Therefore, the Transformer model for pressure classification using EEG signal is effective. That is, Model based on Transformer has good applicability for EEG signal classification tasks.

The EEGNet network is used to compare the performance of the model based on Transformer. The EEGNet network is a small convolutional neural network for EEG-based brain-computer interfaces. It can be seen that there is a certain gap between the classification performance of model based on Transformer and EEGNet network on different datasets. The model performs moderately on datasets collected in the laboratory, but performs better on public datasets with higher data quality.

Table 3. The classification accuracy using model based on Transformer

	Experimental Dataset			Common Dataset	
	Task1&Task2	Task2&Task3	Task2&Task4	Group "B"	Group "G"
Model based on Transformer	92.70%	62.05%	64.74%	81.09%	74.67%
EEGNet	97.05%	68.85%	68.85%	77.10%	71.66%

4. Conclusion

The study aims to make a thorough inquiry of the influence of different music on the human brain in stressful environment. Mental arithmetic test is adopted to produce the stressful environment where subjects are asked to complete tasks first without music and then with music. Based on two stress test questionnaires, Trait Anxiety Inventory (STAI-6) and Self-Stress Assessment, both the relaxing music and the natural rhythm music produced a positive effect in the stressful situations. Furthermore, a Transformer-based model was proposed for EEG classification, which achieved promising performance on different datasets in the stress classification task. In summary, we show that EEG signal and the Transformer-based model combined can be adopted as an effective method to identify and monitor the stress in the human being. And listening to music, such as relaxing music and music with a natural rhythm, is a convenient and effective way for stress relief.

Appendix

Stress Scale

The questionnaire includes three sections. Please choose the option that best suits you according to your actual feeling in the experiment just now. Don't spend too much time thinking about a question.

1. Self-assessment of Stress

Please tick the corresponding number. (The higher the score, the greater the pressure.1: No psychological pressure; 9: A great deal of pressure)

- The pressure I felt in the experiment just now is:
  1 2 3 4 5 6 7 8 9

2.1 STAI-S-6

Please tick the corresponding number. (1: Not at all; 2: Some; 3: Moderate; 4: Very noticeable)

• I feel calm.	1	2	3	4
• I am tense.	1	2	3	4
• I feel at easy.	1	2	3	4
• I feel nervous.	1	2	3	4
• I feel rested.	1	2	3	4
• I am worried.	1	2	3	4

2.2 STAI-T-6

Please tick the corresponding number. (1: Almost never; 2: Sometimes; 3: Often; 4: Almost always)

• I am calm, cool and collected.	1	2	3	4
• I worry too much over something that really doesn't matter.	1	2	3	4
• I feel secure.	1	2	3	4
• I get in a state of tension or turmoil as I think over my recent concerns and interests.	1	2	3	4
• I feel nervous and restless.	1	2	3	4
• I make decisions easily.	1	2	3	4

References

[1] E. Charmandari, C. Tsigos, and G. J. A. R. P. Chrousos, "Endocrinology of the stress response," *Annual review of physiology*, vol. 67, pp. 259-284, 2005. Article (CrossRef Link)

[2] D. Goldstein, K. Pacak, I. J. S. m. g. Kopin, and n. advances, "Nonspecificity versus primitive specificity of stress response," *Stress: molecular genetic and neurobiological advances*, vol. 1, pp. 3-20, 1996.

[3] M. E. J. B. Kemeny, Behavior, and Immunity, "An interdisciplinary research model to investigate psychosocial cofactors in disease: Application to HIV-1 pathogenesis," *Brain, Behavior, and Immunity*, vol. 17, no. 1, pp. 62-72, 2003. Article (CrossRef Link)

[4] Nisar, H., & Hong, S. J., "Study of cognitive flexibility at different stress levels with background music," in *Proc. of 2017 IEEE Life Sciences Conference (LSC)*, Dec. 2017. Article (CrossRef Link)

[5] G. M. Sandstrom and F. A. Russo, "Music Hath Charms: The Effects of Valence and Arousal on Recovery Following an Acute Stressor," *Music and Medicine*, vol. 2, no. 3, pp. 137-143, 2010. Article (CrossRef Link)

[6] Nawaz, R., Ng, J. T., Nisar, H., & Voon, Y. V., "Can background music help to relieve stress? An EEG analysis," in *Proc. of 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Sep. 2019. Article (CrossRef Link)

[7] C. L. J. J. o. m. t. Pelletier, "The effect of music on decreasing arousal due to stress: A meta-analysis," *Journal of music therapy*, vol. 41, no. 3, pp. 192-214, 2004. Article (CrossRef Link)

[8] R. Nawaz, H. Nisar, and Y. V. Voon, "The Effect of Music on Human Brain; Frequency Domain and Time Series Analysis Using Electroencephalogram," *Ieee Access*, vol. 6, pp. 45191-45205, 2018. Article (CrossRef Link)

[9] S. C. Jacobs et al., "Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research," *American heart journal*, vol. 128, no. 6, pp. 1170-1177, 1994. Article (CrossRef Link)

[10] Sulaiman, N., Taib, M. N., Lias, S., Murat, Z. H., Aris, S. A. M., Mustafa, M., & Rashid, N. A., "Development of EEG-based stress index," in *Proc. of 2012 International Conference on Biomedical Engineering (ICoBE)*, Feb. 2012. Article (CrossRef Link)

[11] Hosseini, S. A., Khalilzadeh, M. A., Naghibi-Sistani, M. B., & Niazmand, V., "Higher order spectra analysis of EEG signals in emotional stress states," in *Proc. of 2010 Second international conference on information technology and computer science*, July 2010. Article (CrossRef Link)

[12] S. S. Hassellund, A. Flaa, L. Sandvik, S. E. Kjeldsen, and M. J. H. Rostrup, "Long-term stability of cardiovascular and catecholamine responses to stress tests: an 18-year follow-up study," *Hypertension*, vol. 55, no. 1, pp. 131-136, 2010. Article (CrossRef Link)

[13] Hosseini, S. A., Khalilzadeh, M. A., Naghibi-Sistani, M. B., & Niazmand, V., "Higher order spectra analysis of EEG signals in emotional stress states," in *Proc. of 2010 Second international conference on information technology and computer science*, July 2010. Article (CrossRef Link)

[14] S. Seo, Y. Gil, and J. Lee, "The relation between affective style of stressor on EEG asymmetry and stress scale during multimodal task," in *Proc. of Third 2008 International Conference on Convergence and Hybrid Information Technology*, Nov. 2008. Article (CrossRef Link)

[15] S. Cohen, T. Kamarck, R. J. J. o. h. Mermelstein, and s. behavior, "A global measure of perceived stress," *Journal of Health and Social Behavior*, vol. 24, no. 4, pp. 385-396, 1983. Article (CrossRef Link)

[16] T. H. Holmes and R. H. J. J. o. p. r. Rahe, "The social readjustment rating scale," *Journal of Psychosomatic Research*, vol. 11(2), pp. 213–218, 1967. Article (CrossRef Link)

[17] C. D. J. T. C. e. o. p. Spielberger, "State‐Trait anxiety inventory," *The Corsini Encyclopedia of Psychology*, pp. 1-1, 2010.

[18] J. B. J. A. o. g. p. Williams, "A structured interview guide for the Hamilton Depression Rating Scale," *Archives of general psychiatry*, vol. 45, no. 8, pp. 742-747, 1988. Article (CrossRef Link)

[19] S. Shin, S.-J. Jang, D. Lee, U. Park, J.-H. J. K. T. o. I. Kim, and I. Systems, "Brainwave-based mood classification using regularized common spatial pattern filter," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 2, pp. 807-824, 2016. (CrossRef Link)

[20] E. Perez-Valero, M. A. Vaquero-Blasco, M. A. Lopez-Gordo, and C. J. F. i. C. N. Morillas, "Quantitative Assessment of Stress Through EEG During a Virtual Reality Stress-Relax Session," *Frontiers in Computational Neuroscience*, p. 64, 2021. Article (CrossRef Link)
[21] E. T. Attar, V. Balasubramanian, E. Subasi, and M. Kaya, "Stress Analysis Based on Simultaneous Heart Rate Variability and EEG Monitoring," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, p. 2700607, 2021. Article (CrossRef Link)
[22] Hosseini, S. A., Khalilzadeh, M. A., Naghibi-Sistani, M. B., & Niazmand, V., "Higher order spectra analysis of EEG signals in emotional stress states," in *Proc. of 2010 Second international conference on information technology and computer science*, July 2010. Article (CrossRef Link)
[23] J. F. Alonso, S. Romero, M. R. Ballester, R. M. Antonijoan, and M. A. Mananas, "Stress assessment based on EEG univariate features and functional connectivity measures," *Physiological measurement*, vol. 36, no. 7, pp. 1351-1365, Jul 2015. Article (CrossRef Link)
[24] C. W. Quaedflieg et al., "The validity of individual frontal alpha asymmetry EEG neurofeedback," *Social cognitive and affective neuroscience*, vol. 11, no. 1, pp. 33-43, 2016. Article (CrossRef Link)
[25] A. Arsalan, M. Majid, A. R. Butt, S. M. J. I. j. o. b. Anwar, and h. informatics, "Classification of perceived mental stress using a commercially available EEG headband," *IEEE journal of biomedical and health informatics*, vol. 23, no. 6, pp. 2257-2264, 2019. Article (CrossRef Link)
[26] S. M. U. Saeed, S. M. Anwar, H. Khalid, M. Majid, and U. J. S. Bagci, "EEG based classification of long-term stress using psychological labeling," *Sensors*, vol. 20, no. 7, p. 1886, 2020. Article (CrossRef Link)
[27] H. Jebelli, M. M. Khalili, and S. Lee, "Mobile EEG-based workers' stress recognition by applying deep neural network," *Advances in informatics and computing in civil and construction engineering: Springer*, pp. 173-180, 2019. Article (CrossRef Link)
[28] A. Sundaresan, B. Penchina, S. Cheong, V. Grace, A. Valero-Cabré, and A. J. B. I. Martel, "Evaluating deep learning EEG-based mental stress classification in adolescents with autism for breathing entrainment BCI," *Brain Informatics*, vol. 8, no. 1, pp. 1-12, 2021. Article (CrossRef Link)
[29] Y. Peng, S. Tian, L. Yu, Y. Lv, R. J. K. T. o. I. Wang, and I. Systems, "Malicious URL recognition and detection using attention-based CNN-LSTM," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 11, pp. 5580-5593, 2019. Article (CrossRef Link)
[30] Arjun, A., Rajpoot, A. S., & Panicker, M. R., "Introducing attention mechanism for eeg signals: Emotion recognition with vision transformers," in *Proc. of 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Nov. 2021. Article (CrossRef Link)
[31] S. Bagchi and D. R. J. a. p. a. Bathula, "EEG-ConvTransformer for Single-Trial EEG based Visual Stimuli Classification," *Pattern Recognition*, vol. 129, 2022. Article (CrossRef Link)
[32] A. Vaswani et al., "Attention is all you need," in *Proc. of 31st Conference on neural information processing systems*, pp. 5998-6008, 2017. Article (CrossRef Link)
[33] Wang, Shuai, et al., "Human Short-Long Term Cognitive Memory Mechanism for Visual Monitoring in IoT-Assisted Smart Cities," *IEEE Internet of Things Journal*, vol.9, no.10, pp. 7128-7139, 2022. Article (CrossRef Link)
[34] K. KYUNGWON, "Preprocessed EEG During Mental Arithmetic Tasks," 2020.

Xin Xu received the master's degree from Nanjing University. He is currently pursuing the Ph.D. degree with the Nanjing University of Science and Technology. He joined the Nanjing University of Posts and Telecommunications, in August 2000. He was promoted to an Associate Professor, in May 2012. He has been a Visiting Scholar with Hong Kong Polytechnic University, since October 2018. He has published about 30 articles. He holds three patents. His research interests include signal and information processing, artificial intelligence, bioelectrical signal analysis and processing and their applications.

Yikun Zhao is currently pursuing the master's degree in the Nanjing University of Posts and Telecommunications (NJUPT). Her main research areas are signal and information processing and the analysis of EEG signals.

Ruhao Zhang is a postgraduate student in Nanjing University of Posts and Telecommunications.

Tingting Xu is a lecturer at the School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications.