Stress and anxiety detection: deep learning and higher order statistic approach

Vaishali M. Joshi¹, Deepthi D. Kulkarni¹, Nilesh J. Uke²
¹Department of Electronics and Telecommunication Engineering, KJEI’s Trinity Academy of Engineering, Pune, India
²Department of Computer Engineering, KJEI’s Trinity Academy of Engineering, Pune, India

ABSTRACT

Today's teenagers are dealing with anxiety and stress. Anxiety, depression, and suicide rates have increased in recent years because of increased social rivalry. The research is focused on detecting anxiety in students due to exam pressure to reduce the potential harm to a person's wellness. Research is performed on databases for anxious states based on psychological stimulation (DASPS) and our own database. The measured signal is divided into sub-bands that correspond to the electroencephalogram (EEG) rhythms using the Butterworth sixth-order order filter. In higher dimensional space, the nonlinearities of each sub-band signal are analyzed using higher order statistics third-order cumulants (TOC). We have classified stress and anxiety using the support vector machine (SVM), K-nearest neighbor (K-NN), and deep learning bidirectional long short-term memory (BiLSTM) network. In comparison to previous techniques, the proposed system's performance using BiLSTM is quite good. The best accuracy in this analysis was 87% on the DASPS database and 98% on the own database. Finally, subjects with high stress levels had more gamma activity than subjects with little stress. This could be an important attribute in the classification of stress.

Keywords: BiLSTM, EEG, Higher order statistics, K-NN, SVM, TOC

This is an open access article under the CC BY-SA license.

Corresponding Author:
Vaishali M. Joshi
Department of Electronics and Telecommunication Engineering, KJEI’s Trinity Academy of Engineering
Pune, India
Email: joshivaishali7@gmail.com

1. INTRODUCTION

The most prevalent mental health condition among teenagers is anxiety. Anxiety is a big contributor to bad health, according to research [1]. Numerous chronic diseases are significantly increased by prolonged anxiety. In extreme circumstances, it could lead to severe behaviours like self-harm and suicide. Stress, anxiety, and despair all have a direct impact on the suicide rate among young people. The general mental health of teenagers can be improved by quickly identifying the factors that are most responsible for their psychological problems. Today's teenagers must enhance their will power to understand their full potential and significance in life. Recognizing anxiety has become essential for the early diagnosis of anxious people and the use of mental health therapeutic techniques.

Previous study on emotions used audio signal, video, signal, text processing. Because this behaviour and approach can be faked, now the emphasis is on identifying emotions from physiological signals such as electrocardiograms, electromyograms, galvanic skin responses, respiration rates, and electroencephalograms [2]–[4]. The electroencephalogram (EEG) device is used to assess brain signals because it can identify electrical activity in the brain which is independent on external behaviour. The system involves feature extraction from the EEG data and the development of precise predictive models for emotion recognition to detect stress and anxiety [5], [6]. Adnan et al. [7] studied the stress levels and brainwave balance index of university students at

Journal homepage: http://ijeecs.iaescore.com
the beginning and end of a semester. The perceived stress scale questionnaire developed by Sheldon Cohen was used to quantify stress levels while the EEG signals were captured using g-Mobilab from Guger Technologies. 29 undergraduate students from University Teknologi MARA volunteered to participate. The findings indicate that while stress levels are higher near the conclusion of the semester, there is a tendency for a more balanced brainwave. Rajendran et al. [8] created a database to describe stress levels in students before and after exams. EEG signals were collected from 14 subjects utilizing Enobio device (neuroelectrics) in two separate test conditions: 12 minutes prior to the test and 3 minutes after the test. The relative band energies of three brain waves were considered: theta, alpha, and beta. According to the statistical data, the respective sub-band energies were substantially lower after examination than before examination. Ahuja and Banga [9] used a dataset produced at Jaypee Institute of Information Technology that included data from 206 students to detect stress. Four classification techniques were evaluated using sensitivity, specificity, and accuracy criteria: random forest, Naive Bayes, support vector machine (SVM), and K-nearest neighbor (K-NN). Researchers revealed that SVM performs more successfully than the other three approaches since it classifies data geometrically. Meng and Zhang [10] studied college student stress using their own database as well as publicly available databases. The Takagi-Sugeno-Kang (TSK) fuzzy system and deep features are proposed in this work as a method for automatically detecting anxiety of college students. Deep features can be extracted from input data using convolutional neural networks. Finally, to produce the final recognition result, features are categorized using the TSK fuzzy approach.

While reviewing previous research work, it was noticed that there is a need for additional work on databases and analyses. In the proposed work, we established a database based on stress and anxiety emotions using single channel and employed higher-order statistical feature extraction method to obtain relevant features that produce superior results when leveraging the deep learning network. The results of the experiments, which were conducted using both pre-existing and newly developed databases, highlight the technique's effectiveness in identifying anxiety.

2. METHOD AND MATERIALS

Figure 1 depicts a system block diagram. The underlying mechanisms include database acquisition, band selection, feature extraction, feature selection, and emotion categorization. Theta, alpha, beta, and gamma bands are separated using a six order Butterworth band pass filter. The feature extraction phase's target is to improve the classifier's accuracy by extracting the most valuable features of the signal. To recognize stress and anxiety during the classification phase, we used three machine algorithms: SVM, K-NN, and bidirectional long short-term memory (BiLSTM).

![Figure 1. Block diagram of proposed system](image)

2.1. Data acquisition

2.1.1. DASPS: a database for anxious states based on a psychological stimulation

In the experiment, 23 healthy volunteers without mental conditions actively participated. There are 13 women and 10 men, with a mean age of 30. They chose six instances in which individuals reported the most anxiety, as: loss (68%), family concerns (64%), financial issues (54%), deadline (46%), witnessing a fatal accident (45%), and mistreatment (40%). Face-to-face psychological elicitation carried out by a psychotherapist in a competent manner is used to stimulate anxiety. A wireless EEG headset equipped with the emotiv excess post-exercise oxygen consumption (EPOC) 14 channels was utilized to capture EEG data with two mastoids
positioned in line with the international 10-20 system. Each recording took six minutes, each trial taking one minute. Impedance was capped at 7 k while the recorded EEG signals were recorded at 128 Hz [11], [12]. The raw data was subjected to a 4-45 Hz Finite impulse response pass-band filter. Electrooculography (EOG) and electromyography (EMG) artifacts were removed using the automatic artifact removal for EEGLab tool.

2.1.2. Own database

Neurosky Mindwave Mobile 2 brain-computer interface (BCI) EEG device is used to collect data from human brain wave activity. It has a Bluetooth connection that enables wireless use and allows pairing with cell phones and laptops. To compute the electrical voltage of the brain, single channel mindwave EEG device consists of a forehead electrode (Fp1), and A1, which act as ground. To remove noise from 50 Hz power lines, EEG data were notch filtered. The raw EEG data was then passed through a Butterworth pass-band filter to obtain a frequency in the 4-44 Hz range.

a) Subject

Fifty students, 50% boys and girls, voluntarily participated in this research. The participants’ ages ranged from 15 to 20 years old, with a 17.5 years average. Students with normal or corrected eyesight took part in the experiment. All participants were given adequate instructions and required to sign consent forms prior to the experiment starting.

b) Experimental protocol

EEG data was collected from each student for about five minutes before and after the exam. The exam will last about 1 hour. Before the start of the examination the Neurosky Mindwave Mobile recorded their EEG data about five minutes. The student was instructed to either study the material that will be covered in the test or to recall the topics using books or notes prior to the exam. Following the one-hour test, a further 2-minute EEG signal recording was made.

2.2. Feature extraction

In building an EEG-based BCI system, feature extraction is essential. Feature extraction is the technique of capturing helpful information from pre-processed EEG signals. A feature is a quantity that reflects the difference between classes. Extraction of relevant features is an important aspect of any algorithm; otherwise, the classifier may perform poorly [13], [14]. Fast fourier transform, wavelet entropy, eigen vector, power density, Hjorth parameters, higher order statistics, and differential entropy are a few of the features that have been employed in the literature [15]-[19]. Higher order statistics (HOS) can distinguish variation from linearity in a signal. Because most physiological signals are non-Gaussian in nature, it may be beneficial to assess them with HOS rather than auto-correlations and power spectra. They have been applied to the modeling of time-series data [20]. The effectiveness of an EEG-based emotion identification system was evaluated using four bands of EEG data.

2.2.1. Third order cumulant

Extremely nonlinear and stagnant behaviour is revealed by the EEG signals. For analysis, EEG data with short epochs are useful; hence Hanning window has been used to obtain 1-second length small signal segments with 50% overlap. HOS- third order cumulant of a signal is its triple coherence. The existence of nonlinearity in non-stationary brain signals is demonstrated by the function of two lag factors and their harmonics. It gives a nonlinear technique for signal analysis in higher dimensions. Let \( z[k] \) is a stationary yth order random process. The yth-order moments can be used to determine the yth-order cumulant of \( z[k] \), \( k = 0, \pm 1, \ldots \) as (1) [18], [21]:

\[
C_y^z[\delta 1, \delta 2, \delta 3 \ldots \ldots , \delta_{y-1}] = M_y^z[\delta 1, \delta 2, \delta 3 \ldots \ldots , \delta_{y-1}] - \mu_y^z[\delta 1, \delta 2, \delta 3 \ldots \ldots , \delta_{y-1}] \tag{1}
\]

where \( M_y^z \) and \( \mu_y^z \) denotes the yth order moment function of \( z[k] \), and Gaussian random process, respectively. The yth order cumulants are just a function of the \((y-1)\) lags \([\delta 1, \delta 2, \delta 3 \ldots \ldots , \delta_{y-1}]\). The cumulant may be defined as follows for orders \( y = 1, 2, \) and 3:

\[
M_1^z[\delta 1, \delta 2, \delta 3 \ldots \ldots , \delta_{y-1}] = E\{z(k)z(k+\delta 1)z(k+\delta_{y-1})\} \tag{2}
\]

where \( E \) is the expectation operator.

First order cumulants-

\[
C_1^z = M_1^z \tag{3}
\]
Second order cumulant-

\[ C_2^2[\delta_1] = M_2^2[\delta_1] - (M_1^2)^2 \]  

(4)

Third order cumulant (TOC)-

\[ C_3^2[\delta_1, \delta_2] = M_3^2[\delta_1, \delta_2] - M_2^2(M_2^2[\delta_1] + M_2^2[\delta_2] + M_2^2[\delta_2 - \delta_1]) + 2(M_1^2)^3 \]  

(5)

The dynamics of a non-stationary signal can be shown more effectively by the TOC due to following reasons:
- A random variable’s TOC is equal to zero when it has a symmetric probability density function (PDF), which enables the analysis of non-Gaussian signals when there is also a Gaussian signal present.
- Signals with non-minimum phases and phase pair signals can be analysed using the TOC due to its infinite differential nature, convex structure, and scale invariance.

After feature extraction, the feature matrix values are normalised via the z-score normalisation method.

2.3. Feature selection: differential evolution

There are several approaches for reducing the dimension of the data. The group of features that are most important from the retrieved EEG features can be chosen using differential evolution (DE) method [22], [23]. DE is a stochastic optimization method that has lately gained popularity for its applicability to continual search issues. The DE methodology is used in this research to draw out the pertinent data from the ToC coefficients. The simplicity that basic DE only requires three control settings to be adjusted is its principal benefit. The trial vector generation method and the hyperparameter selection have a significant impact on DE's pursuance in each optimization task. To achieve good optimization, first choose a trial vector generation approach and then tweak the hyperparameter for the optimizer. Because the approach doesn't need gradient information, the optimization problem doesn't have to be differentiable. A population of candidate solutions (individuals) is maintained by the algorithm while it searches the design space, and new solutions are produced by merging existing ones in accordance with a predetermined method. The candidates with the best objective values are preserved in the algorithm's subsequent iteration so that each candidate's new objective value can be improved and subsequently become a member of the population; otherwise, the candidate's new objective value is eliminated. The cycle is repeated until a predetermined termination threshold is met.

2.4. Classification

In this study, SVM, K-NN, and BiLSTM network were employed to investigate EEG-based anxiety and stress identification. These classifiers were likely chosen based on their distinct characteristics and capabilities. SVM and K-NN are traditional machine learning models, while BiLSTM represents a deep learning approach, particularly effective for tasks involving temporal dependencies and sequences, such as EEG signal analysis for identifying anxiety and stress. The comparative analysis of these classifiers could provide insights into their respective strengths and performance in the context of EEG-based emotion identification. The following is an overview of these classifiers.

2.4.1. SVM

SVM is a well-known supervised algorithm for learning that can be used to solve regression and classification problems. However, it is most commonly employed in machine learning for classification problem [24]–[26]. In order to efficiently add new data points to the proper category in n-dimensional space, the SVM method seeks to identify the optimum decision boundary. A hyperplane is the ideal type of boundary. SVM selects an extreme vectors and points to assist in the hyperplane's formation. SVM techniques employ a set of mathematical operations known as the kernel. Data is fed into the kernel, which subsequently converts it to the desired format. Different SVM algorithms employ different kernel functions. These functionalities come in a variety of forms. For example, linear, nonlinear, polynomial, sigmoid, and radial basis function (RBF). RBF kernel functions are the most popular kind. Since the entire x-axis exhibits a confined and finite response. The kernel functions return the inner product between two locations in an optimal feature space. As a result, even in very high-dimensional environments, defining a notion of similarity can be done at a low computing cost.

2.4.2. K-NN

The machine learning technique K-NN makes use of supervised learning [27]–[29]. The K-NN model predicts similarity between new and prior data points and assigns the new point to the category most similar to the existing examples. The K-NN method uses similarity to update all current data and classify new data points.
2.4.3. Deep learning algorithm-BiLSTM network

Many models have been created to accurately fit and analyze the dataset while dealing with complex challenges. Recurrent neural networks (RNNs), a type of neural network topology, have developed to suit this requirement. In an RNN, a neuron's output can directly affect itself at the next timestamp. RNN memory is constrained, thus it can only retain knowledge from earlier stages. The sequence advances, it’s prone to gradient explosion and disappearance. To overcome this issue LSTM is introduced [30]. Long short-term memory (LSTM) is used to handle time sequences. Its distinctive advantage is that memory cells are utilized to replace hidden layer nodes, the issue of gradient disappearing and gradient explosion has been effectively resolved. The model gains knowledge of the EEG signal's temporal information after being fed continuous time sequences. In LSTM network consists of three control gates, input (I_t), output (O_t), forget gate (F_t), and memory cell (C_t) respectively [31]. To remember and process more data using two LSTM units, Bi-LSTM effectively combines bidirectional features with gating design. The obtained features from the EEG data are sent into the network through the input sequence layer, as shown in Figure 2. The number of features used determines the size of the sequence input layer. The BiLSTM layer examines and interprets this data. The BiLSTM layer picks up on the long-term bidirectional dependencies of the EEG data. The fully connected layer receives this data. The link between locations is learned and stored during training using the weighted matrices. The fully connected layer is the same size as classes. To classify the labelled output, the softmax layer applies the softmax function. The LSTM cell remembers the long-term associations between the sequential data’s time steps.

![Figure 2. Architecture of bidirectional LSTM](image)

3. EXPERIMENTS AND RESULT

3.1. Experiment setup

In this study, we employed the SVM, the K-NN, and the BiLSTM network to recognize anxiety and stress. SVM, K-NN, and BiLSTM network classifiers employ the normalized TOC features as inputs. Results from the proposed strategy have been verified using a variety of indicators, including accuracy (Ax), sensitivity (Sen), and specificity (Spe). Confusion matrix parameters are utilized to represent these measurements as true positive (tp), true negative (tn), false positive (fp), and false negative (fn). To test how precise the recognition was, few different models were experimented. In the SVM experiment, the different kernel function like linear, polynomial, and radial basis were employed. The value of k is altered for K-NN from K=1 to 5. We tested numerous models to confirm the accuracy of the BiLSTM network. BiLSTM has a hidden unit number of 25 to 250 and a hidden layer number of 1 to 2. Through a process of observations and experimentation, the parameters are chosen. The best-performing metrics were used as the basis for the classification procedure. For K-NN; K=3, the RBF for SVM, and 175 hidden units, single hidden layer, adaptive moment estimation optimizer, and learning rate of 0.001 was used to train the BiLSTM. Using this hyper parameter, we were able to achieve the best results.
3.2. Results

In order to analyse the stress levels of the students throughout the exam period, this study used two experimental conditions: before the exam and after the exam. Using TOC and three classifiers SVM, K-NN, and BiLSTM, anxiety and stress are examined. A 5-fold cross-validation strategy is used so that the complete data set can be used to train and test the classifiers. Here, classification performance is assessed using the classification accuracy rate across all experiments using 5-fold cross-validation. It is important to mention that the results obtained by employing TOC with the BiLSTM network are superior to those obtained by other classifiers. Furthermore, it has been noted that the gamma band can achieve up to 98% accuracy using own database and 87% accuracy using DASPS database and provides results that are more accurate as shown in Figures 3 and 4. According to the findings of our experiment, the students’ gamma band energy was higher before the exam than it was after, suggesting that they were under a lot of stress. To compare performance with the established method, we included the findings of relevant study in Table 1.

Figure 3. The classification performance measures with BiLSTM, SVM, and K-NN classifier on own database

Figure 4. The classification performance measures with BiLSTM, SVM, and K-NN classifier on DASPS database
The proposed approach outperforms the other current techniques as shown in Table 1 in terms of classification accuracy. We offer HOS-based features in our approach to boost classification accuracy. The BiLSTM trained on the relevant features retrieved from the HOS, which accurately depict the underlying signal fluctuations that contribute to emotion recognition.

### Table 1. Performance comparison with state of the art techniques

<table>
<thead>
<tr>
<th>Article</th>
<th>EEG device</th>
<th>Number of subject</th>
<th>Number of electrodes</th>
<th>Features extraction methods</th>
<th>Classification methods</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nagar and Sethia [25]</td>
<td>NeuroSky</td>
<td>53</td>
<td>1</td>
<td>PSD</td>
<td>K-NN</td>
<td>74.43</td>
</tr>
<tr>
<td>Alfred and Chia [32]</td>
<td>Mind Wave</td>
<td>25</td>
<td>1</td>
<td>DCT</td>
<td>K-NN</td>
<td>72</td>
</tr>
<tr>
<td>Blanco et al. [33]</td>
<td>Emotiv EPOC</td>
<td>18</td>
<td>14</td>
<td>RMS</td>
<td>Logistic regression</td>
<td>70.72</td>
</tr>
<tr>
<td>Katsigiannis and Ramzan [34]</td>
<td>Emotiv EPOC</td>
<td>23</td>
<td>14</td>
<td>PSD</td>
<td>SVM-RBF</td>
<td>87</td>
</tr>
<tr>
<td>Liu et al. [35]</td>
<td>Emotiv EPOC</td>
<td>30</td>
<td>14</td>
<td>PSD</td>
<td>SVM-RBF</td>
<td>92.2</td>
</tr>
<tr>
<td>Baghdadi et al. [11]</td>
<td>Emotiv EPOC</td>
<td>23</td>
<td>14</td>
<td>Hjorth, PSD</td>
<td>Stacked sparse autoencoder</td>
<td>83.50</td>
</tr>
<tr>
<td>Proposed (own database)</td>
<td>NeuroSky MindWave</td>
<td>50</td>
<td>1</td>
<td>TOC</td>
<td>BiLSTM</td>
<td>98</td>
</tr>
<tr>
<td>Proposed (DASPS database)</td>
<td>Emotiv EPOC</td>
<td>23</td>
<td>14</td>
<td>TOC</td>
<td>BiLSTM</td>
<td>87</td>
</tr>
</tbody>
</table>

The proposed approach outperforms the other current techniques as shown in Table 1 in terms of classification accuracy. We offer HOS-based features in our approach to boost classification accuracy. The BiLSTM trained on the relevant features retrieved from the HOS, which accurately depict the underlying signal fluctuations that contribute to emotion recognition.

## 4. CONCLUSION

Identification of mental stress and anxiety is an essential issue for the young generation. To avoid its harmful effects, it is necessary to diagnose and detect mental stress. This research revealed leveraging EEG signals to detect mental stress instantly and automatically. According to our research, students' gamma band energy was higher before the exam than it was after, indicating that they were under a lot of stress. To prevent students from having suicidal thoughts during exam season, our research indicates that stress analysis is significant. Under appropriate counselling and support, students stress levels significantly reduced.

This research utilizes TOC to extract features and stress and anxiety was classified using the SVM, KNN, and BiLSTM methods. The best accuracy was reached for as 98% on own database and 87% on DASPS database using BiLSTM classifier. Additionally, it has been noted that the gamma band is a key band for emotion recognition. Single channel EEG devices are inexpensive and simple to use, producing the same performance as multichannel EEG devices. The standard approach of employing full-channel EEG signals has considerable disadvantages because it necessitates the management of extra and undesirable data, which leads to increased memory requirements and consequently costly hardware requirements. The research introduced a breakthrough approach to neuroscience, demonstrating that a single-channel EEG carries enough information for emotion recognition.

## REFERENCES


BIographies of authors

Vaishali M. Joshi is Associate Professor at KJIE’s Trinity Academy of Engineering, Pune in E&TC Engineering Department. She holds a Ph.D. degree in Electronics Engineering. Her research areas are image/signal processing, VLSI, AI-ML, and pattern recognition. She has filed several patents. Her research interests include image/signal processing, biometrics, medical image and analysis, and pattern recognition. She published three papers in SCI indexed journal. She has 23 years of teaching and research experience in the field of VLSI, communication, and AI-ML. She can be contacted at email: joshivaishali7@gmail.com.

Stress and anxiety detection: deep learning and higher order statistic approach (Vaishali M. Joshi)

Deepthi D. Kulkarni awarded ME degree from MAE, Alandi under SPPU in 2012. Currently pursuing Ph.D. from G.H. Raisoni college Pune. She is working as assistant professor and HOD in KJEI’s Trinity Academy of Engineering, Pune and have completed 15 yrs of teaching experience. Her research areas include biomedical image processing, and signal processing. She can be contacted at email: deepthi.me2009@gmail.com.

Nilesh J. Uke received the B.E. degree in Computer Science and Engineering from Amaravati University, India, in 1995, and the M.E. from Bharathi Vidhyapeeth in 2005 and Ph.D. degrees in Computer Science, from Swami Ramanand Thirth Marathwada University, Nanded, India, in 2014. He is currently Principal and Professor at KJEI’s Trinity Academy of Engineering, Pune, Maharashtra. His current research interests include visual computing, artificial intelligence, human computer interface, and multimedia. He is member of IEEE, ACM and Life Member of the Indian Society for Technical Education (ISTE) and Computer Society of India (CSI). He can be contacted at email: nilesh.ukc@gmail.com.