Stress Detection using EEG Signal Based on Fast Walsh Hadamard transform and Voting Classifier

Navdeep Shakya
Madhav Institute of Technology and Science
RAHUL DUBEY (rahul@mitsgwalior.in)
Madhav Institute of Technology and Science https://orcid.org/0000-0002-9763-8005
Laxmi Shrivastava
Madhav Institute of Technology and Science

Research Article

Keywords: Brainwaves, Mental Stress, EEG Signal, FWHT, Voting Classifier.

Posted Date: September 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-782483/v1

Noname manuscript No.
(will be inserted by the editor)

Stress detection using EEG signal based on Fast Walsh Hadamard transform and Voting Classifier

Navdeep Shakya · Rahul Dubey · Laxmi Shrivastava

Received: 02 Aug 2021 / Accepted: date

Abstract Mental stress is currently a significant concern, especially among the young. Stress adversely affects the overall performance of people's work, and in certain cases, can even cause serious health issues. Everyone experiences stress in life. A unique way to identify and classify stress levels based on Electroencephalogram (EEG) is proposed in this manuscript. In this work, fast Walsh Hadamard transform is used to generate all frequencies which exist in the EEG signals. The range of alpha, beta, gamma, and delta from index value is calculated in subsequent stage. Principal component analysis (PCA) is applied for the feature dimensional reduction which is followed by the standard scaler. The PSD vector has been calculated for healthy and unhealthy EEG signal groups using the Welch method. The PSD vector is used an input to the voting classifier which is the combination of the k-NN and logistic regression classifier. The experimental results found that the proposed method provides better results when compared to the existing methods in terms of Accuracy (Acc) and Mean Square Error (MSE). The proposed method achieves a highest classification accuracy of 94.22%

Keywords Brainwaves · Mental Stress · EEG Signal · FWHT · Voting Classifier.

1 Introduction

Mental stress is the human body's response to stressful measures that are encountered due to physical or psychosocial circumstances. It affects people

*Corresponding author
Rahul Dubey
Department of Electronics Engineering
Madhav Institute of Technology and Science
Gwalior, India
E-mail: rahul@mitsgwalior.in

of various ages, genders, and occupations around the world. Work challenges, increased strain, increased everyday activities that people experience play an important role in the rise in stress. Mental stress is proved to be the primary cause behind several illnesses like Cardiovascular disease, depression, anxiety, & post-traumatic stress disorder (PTSD). To prevent disease and decrease the possibility of health issues such as clinical brain damage, it is necessary to detect stress as early as possible [1-3]. Meditation has been practiced in India for centuries, and it stems from an ancient spiritual tradition. It has gained great public acceptance for both therapeutic & self-development purposes. It has a profound influence on both the body and the mind.

Electroencephalography (EEG) is a measurement technique for studying brain electrical activity by analyzing the electrical activity of the cerebral cortex. There are two types of approaches existed to obtain brain activities named as invasive and non-invasive approach. Measurements utilizing electrodes connected to the scalp use the summed activity of brain cells to create brainwaves on an EEG [4-5]. EEG is a non-invasive tool that provides a look into the brain's physiology and behavior with deep complexity. The amplitude of an EEG impulse is highly nonlinear & nonstationary, varying considerably between people. It is a highly valuable diagnostic tool for neurological diseases such as epilepsy, insomnia.

Developing technology known as Machine Learning (ML) used to classify the objects to a distinguish classes. ML finds application in many areas not limited to speech recognition, image pattern recognition, web search, spam mail filtering, etc [6]. With the help of ML, one can more accurately forecast and identify given conditions to gauge the severity of the illness. In medicine, considerable use of machine learning has recently been observed. At some point, artifactual components need to be classified as automatic components for their general use. Machine learning approaches are used to identify and enhance classification accuracy with ML algorithms [7-8].

In [9], a single electrode EEG device was used to design a brain mapping-based stress detection system with electrodes mapping multiple stress points. Classifiers are provided with a new feature grouping of band power ratios of alpha, beta, delta, & theta bands, which is then interpreted as a novel way of presenting the information. SVM & KNN ML algorithms used to examine the actual performance. The stress score obtained from the response of the PSS-14 questionnaire is used to select the training set for the target class. The KNN algorithm achieves the highest average classification accuracy of 74.43%.

The researcher in [10] developed and suggested that the EEG is utilized in Stress Detection and meditation application to measure stress levels. Authors used FFT for feature extraction and KNN for classification purpose. The features shown to be most appropriate for stress detection & meditation include the Delta, Theta, Alpha, & Beta waves. The experiments show that the k = 3 in kNN yields the best classification accuracy of about 80%.

This research [11] is to classify the EEG band to search for a consistent abrupt change in the band that may be caused by Om mantra meditation. Twenty-three inexperienced meditators had their brains scanned before and after they chanted Om for 30 minutes. In the case of the EEG, the stationary wavelet transform is applied to produce five new EEG frequency bands. A radial basis kernel SVM classifier is used to identify the band. Significant alterations were found in the delta band, representing the brain in deep sleep, as shown in the findings. By this, it is apparent that Om meditation offers the sensation of deep sleep. The study might be beneficial to get a new perspective on meditation.

Some researchers worked on the determination of stress using Heart rate, EGG signals etc. The work in [12] investigate the growth of a stress recognition system based on heart rate. The proposed method in [12] integrates HRV data calculated and compared from time and frequency domain analyses with a kNN algorithm to be applied to classify HRV features. Using the kNN method, the experimental value includes the comparison of stress measurements obtained from time and frequency domain analyses yield the highest classification accuracy of 79.17%.

The classification of EEG signals into relaxed and stress EEG based on the sub-band power ratio is presented in [13]. Sub-band power ratio is used as a feature set to classify the EEG signals. Support Vector Machine (SVM) classifier with various kernel function and K-Nearest Neighbor (KNN) classifier has been used to classify the EEG signals into two respective classes. It is found that the SVM algorithm outperforms the logistic regression & KNN algorithms for classifying signals with 92.11% accuracy. When EEG signal classification is attempted with the SVM method, Accuracy scores of 92% are obtained in the 3-class problems correspondingly.

The main motivation behind this research is the use of voting classifier in the stress detection using EEG signal processing. The voting classifier used in this research is a combination of logistic regression (LR) and kNN. The proposed method achieved a highest classification accuracy of 94.22%. The proposed algorithm has been tested on a dataset publicly available at [14-15]. The dataset is divided into two group named healthy and unhealthy EEG signals. Each group consists of a 36 signals.

2 Methodology

This manuscript has proposed a fusion mechanism to overcome the gaps in reaearch and to identify stress in individuals. The first step is the use of an EEG to capture brain signals, & further processing of the signal would be done to create all frequency from these signals using fast Walsh Hadamard transform (FWHT). The FFT is used to calculate the range of alpha, beta, gamma and delta from the index value. Standard scaler is performed on the obtained data to apply principal component analysis (PCA) for selecting the features and calculate the PSD vector for 1st and 2nd group using the Welch method. The power spectral density (PSD) is the distribution of signal power across frequencies. The Welch algorithm [16] is a nonparametric method to estimate the PSD and it makes the frequency spectrum smoother than the raw FWHT output. The obtained PSD vector is given as an input to the voting classifier to classify the EEG signals into healthy and unhealthy EEG.

2.1 Fast Walsh–Hadamard Transform (FWHT)

The Hadamard ordered FWHT is an effective method to calculate WHT in computer mathematics. Computational complexity is a naive implementation of the WHT for $n = 2^m$ would require an $O(n^2)$. Adding or subtracting one or more numbers is all that is required for the FWHTh.

This FWHT is distributed & conquer algorithm that splits a WHT of size n in 2 small WHTs of size $n/2$ [17]. This method defines the $2^m \times 2^m$ Hadamard matrix as a recursive function of H_m.

$$H_m = \begin{Bmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{Bmatrix} \tag{1}$$

It is possible to aggregate the $1/\sqrt(2)$ normalization factors for every phase or omit them altogether. The Walsh-Hadamard transformation is also called the sequence oriented. FWHT is computed using the above equation 1, and then the outputs are reorganized. From decomposing the Hadamard transform matrix as $H_m = A_m$, where A is the m^{th} root of H_m, a simple fast non-recursive deployment of the Walsh-Hadamard transform emerges [18].

2.2 Principal Component Analysis (PCA)

For dimension reduction, PCA is applied to the original EEG. To use the highest variance requirement is similar to feature extraction, in which the principal component is being used as a new feature rather than the dependent data. PCA [19] is a useful transform algorithm which can convert source data into feature data in a linear, smaller-dimensional manner, with the feature data's variance at its maximum. To reduce the cost of computing, a higher dimension feature vector can be translated into lower dimension PC space via transformation. PCA is a method for discovering patterns in data & present them in a way that demonstrates their similarities and differences. PCA is a valuable statistical tool for evaluating data that has found application in areas like mental stress detection because it is problematic to find patterns when the data is in high dimensions. Another major benefit of PCA is that it allows you to detect trends in information & compress it by decreasing no. of dimensions without sacrificing much data. The decision to apply PCA was made since most EEG data contains a lot of redundant information that isn't useful for diagnostic purposes.

Suppose $x = (x_1, x_2, \ldots, x_M)'$ are $M \times 1$ vectors, x_i is a row vector of N points.

- Step 1

$$\overline{x_i} = \frac{\sum_{i=1}^{n} x_i}{M}, i = 1, 2, 3, ..., N \tag{2}$$

- Step 2 Detract mean

$$\Phi_i = x_i - \overline{x_i} \tag{3}$$

- Step 3 Calculate the covariance matrix

$$C = \frac{\sum_{K=1}^{M} \Phi_k \Phi_k^T}{M} \tag{4}$$

- Step 4 Calculate the eigenvalues of covariance matrix

$$C = \lambda_1, \lambda_2 \ldots \lambda_N \tag{5}$$

- Step 5 Calculate eigenvectors of covariance matrix

$$C = u_1, u_2, \ldots, u_M \tag{6}$$

- Step 6 Get a relationship of PCA & original variables

$$y^{'} = U^T \times Phi_i for K < M \tag{7}$$

The first few basis vectors with the most energy will hold most of the information, and combining this information with PCA compression may enable you to retrieve the original signal. A linear-algebra decomposition procedure generates the eigenvectors and related eigenvalues that are tied to particular distinct sets [20]. The first issue is how to select PC, or how to select K. The criteria utilized in this research are as follows:

$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{N} \lambda_i} > Threshold \tag{8}$$

where λ_i is the eigenvalues of the covariance matrix.

The second challenge is figuring out where on the electrodes the epileptic signal appears and how to relate the principal components to the initial EEG signal. When a principal component's factor loading is applied, the problem is solved. Factor loading is defined by several standard formulas:

$$\rho_(y_k, x_m) = \frac{COV(y_k, x_m)}{\sqrt{Var(y_k)}\sqrt{Var(x_m)}} \tag{9}$$

Here, y_k & x_m are the principal component, and the original signal respectively, and $u_k m$ is the corresponding coefficient vector.

2.3 Logistic Regression

The second model will be constructed. Whenever y is a categorical variable, that model seeks to fit a regression curve, $y = f(x)$. In logistic regression, statistical techniques are used to predict binary outcomes ($y = 0 or 1$). A linear learning algorithm is a logistic regression. The probability of an event occurring is used to make logistic regression predictions. The LR algorithm [21] uses the sigmoid function to map each data point. An S-shaped curve is produced using the ordinary logistic function. The sigmoid function is written below.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{10}$$

2.4 k-Nearest Neighbour (kNN)

The first model is called kNN. the test set is a subset of the training set. To find the k nearest training set vectors for each row of the test set, a majority vote is utilized, and no changes are being made to the classification, and ties are randomly broken whenever they occur. In the event of a tie for the kth nearest vector, all candidates are additionally included with voting.
A majority voting method is used by the k-NN algorithm [22]. After training on a set of real-world data, the model is then used to create predictions about new data. The training data set is evaluated to find the k-nearest records. Based on the value of the target attribute of the closest records, a prediction is made for the new record.
The basic NN algorithm makes classification predictions or regression forecasts for an arbitrary instance. To this purpose, the NN algorithm identifies a training instance that is closest to the arbitrary instance. Then, the NN algorithm returns the class label or target function value of the training instance as the predicted class label or target function value for the arbitrary instance. The KNN algorithm expands this process by using a specified no. $k \leq 1$ of the closest training examples instead of using only one instance. Typical values range from 1 to several dozens. The output depends on whether you use the KNN algorithm for classification or regression.

- The forecasted class label in the KNN classification is established by voting for the nearest neighbors, that is, the majority class label in the set of the selected k instances is provided.
- In KNN regression, the average value of the target function values of the nearest neighbors is returned as the predicted value.

By using a specified number $k \geq 1$, you can control the tradeoff between overfitting prevention and resolution. Preventing overfitting may be critical when dealing with noisy data. For example, to acquire different predictions for similar cases, one must make a decision.

2.5 Voting Classifier

This is a merge-based meta-classifier for machine learning models that implement majority voting for predictions. Hard & soft voting are utilized in classifiers that employ two different approaches. With hard voting, the aggregator randomly selects the class prediction that appears most frequently among the base models, so this class prediction becomes the final forecast. The Predict proba technique should be present in soft voting base models. Other base models, as they mix the predictions of several models, give superior overall outcomes. according to the developed framework, K- Nearest Neighbor and Logistic Regression classifier have been ensembled in hard voting. When the shuffling is complete, it passes training data and data samples to the KNN & LR models. Each model uses a different method for calculating individual predictions and combines these results using an aggregator, which provides the final forecast [20].

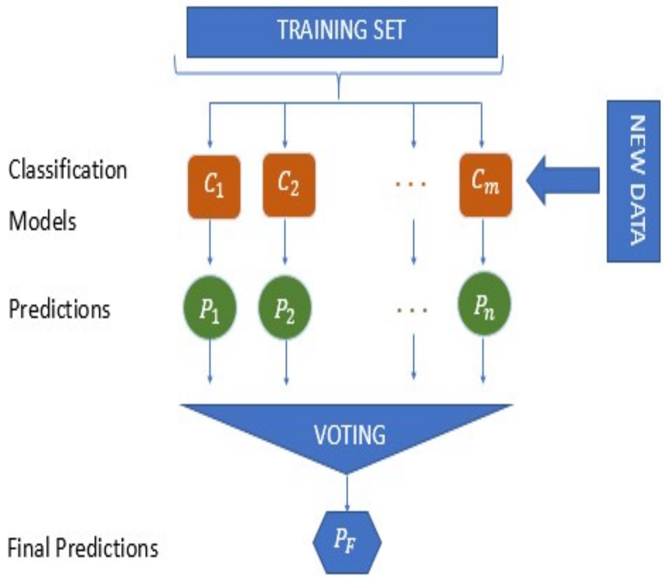

Fig. 1 Architecture of a proposed voting classifier

The block diagram of the proposed methodology is shown in Fig. 2.

2.6 Dataset Description

During and before performing mental arithmetic tasks, the database holds EEG recordings of the test subjects. Every subject has two separate files named before mental arithmatic task and after mental arithmatic task. These files are

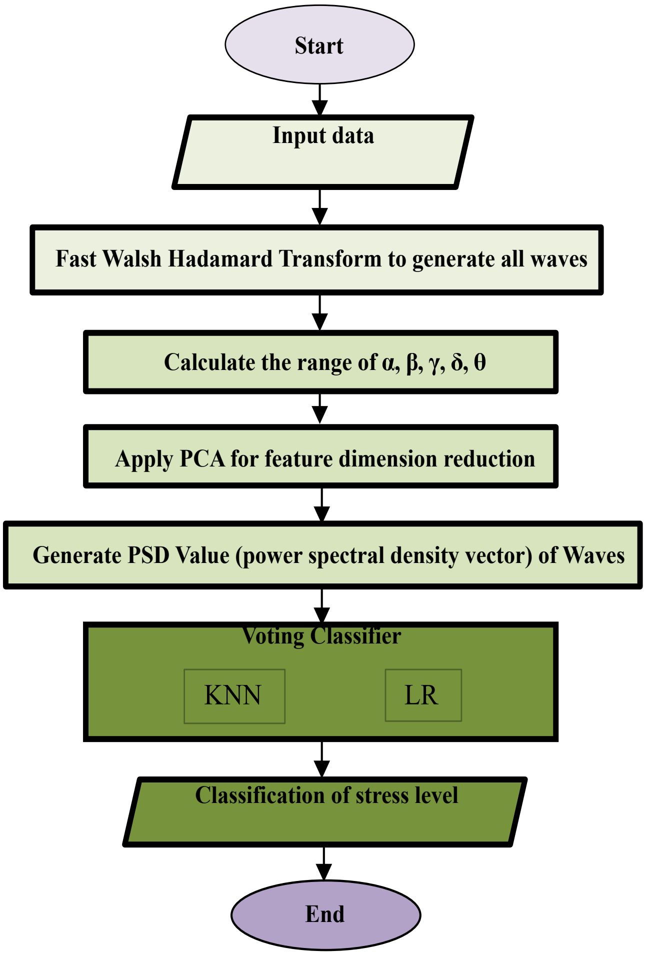

Fig. 2 Block diagram of the proposed method

named in the database with _1 and _2 suffix. There are 72 files in these group with 36 sample files in each groups. Every subject is placed into two categories in this research:

- First group contains 24 subjects who perform a fair quality count in which the mean number of operations every four minutes is 21 with a standard deviation of 7.4.
- Second group contains 12 subjects who were operated with low-quality counts, as the mean number of operations performed per four minutes is 7. However, the SD is 3.6.

3 Results and Discussion

This proposed work has experimented using Jupyter notebook in python programming. The dataset was used here for EEG signal waves that include delta, beta, gamma, alpha, and theta waves. The efficiency of the proposed voting classifier is calculated with experimental & test results. The results are discussed in details in this section.

3.1 Brainwaves Information

There are around 100 billion neurons in the human brain. Through the response that involves between these neurons, these electric currents can be captured and interpreted as brain waves via an EEG. These include the following five frequency ranges of brain waves discussed below.

Theta, delta, alpha, gamma and beta are the used in the experiment for analysis. The delta wave can calculate the depth of the sleep, and its frequency (1-4Hz). Delta wave can be identifying the slow-wave sleep using EEG, in slow-wave sleep brain waves are very slow so this is called dreamless sleep, and Dreams occurred very often. Nightmares occur during this sleep but, we are not able to recall the dreams. The following rates are decreasing during this sleep BP, respiratory rate, and BMR. Theta wave can be used to know the functions of the brain, which means that the difficult task of the brain and it's associated with the weakness level. The frequency is about (4-7Hz), Theta is connected with all-around cerebral processing such as memory conceal and cognitive workload, it is also calculating the tired level of humans. Alpha denotes our mind-released state, and it records the relaxation of the brain whenever we closed our eyes, we turn into a calm state at that time alpha wave take over, and it is related to shyness and attention, the frequency of alpha is (7-12Hz). Beta waves with frequencies of (12-30Hz). It can notice the body movements, such as limp movement fore limp (hand) hind limp (leg), this increases in beta also perceptible as we notice bodily movements of other peoples. Human brains imitate the movements of their limbs and indicating the mirror neuron system. The Gamma waves, typically the gamma frequency is (¿30Hz to 40Hz). Gamma waves can give information about our sensory inputs. These waves are similar to the REM (rapid eye movement). All the instances were used to assess at a point during this process and the remaining instances were utilized to the training of the classifier [20].

Fig. 3 and Fig. 4 shows the brainwaves for first and second group respectively. Fig.3 shows the waves before FWHT and Fig.4 shows brainwaves after applying FWHT.

In a scatter plot, points are used to indicate 2 distinct numerical values for two different variables. When generating an X and Y axis on a graph, the position of each dot on the horizontally and vertically axis shows values for a specific data point. Scatter plots are commonly used to discover patterns and

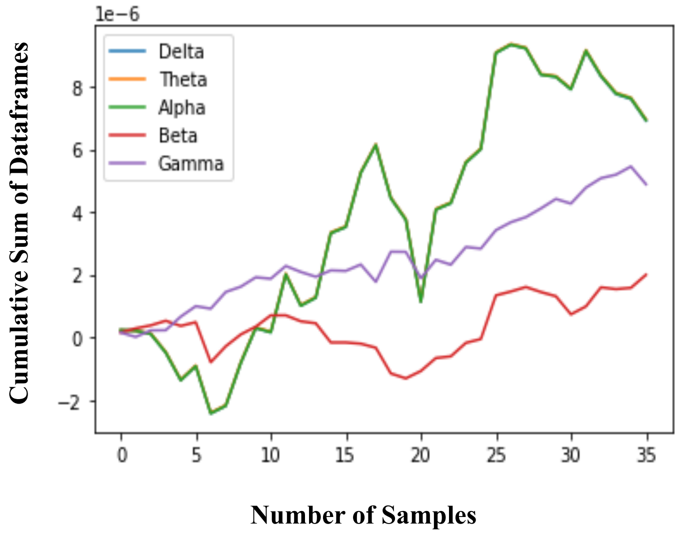

Fig. 3 Plot of waves of first group

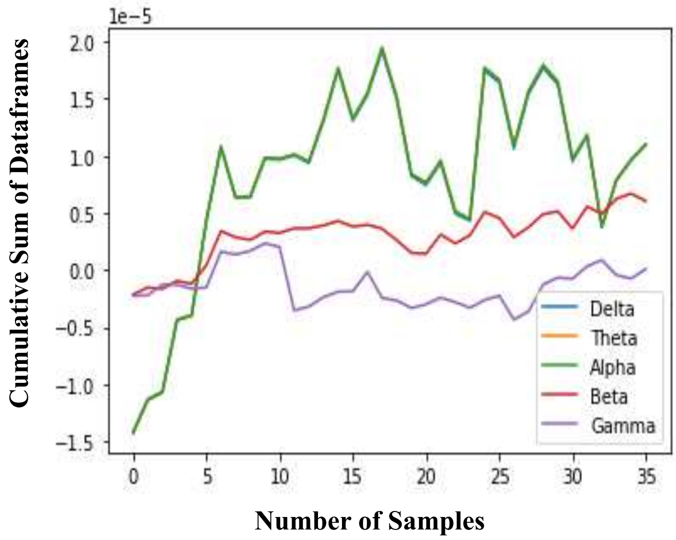

Fig. 4 Plot of waves of second group

correlations among variables. Fig. 5 represents a scatter plot for both (1st and 2nd) groups for cumulatice sum of dataframes.

A histogram is a bar graph-like description of data, which classifies results into ranges on the x-axis into columns. A representation of data distributions may be seen by the y-axis, which displays the number of repetitions within every column. Fig. 6 shows the histogram of 1^st group of data. Similarly, Fig. 7 shows the histogram of the 2^nd group of data.

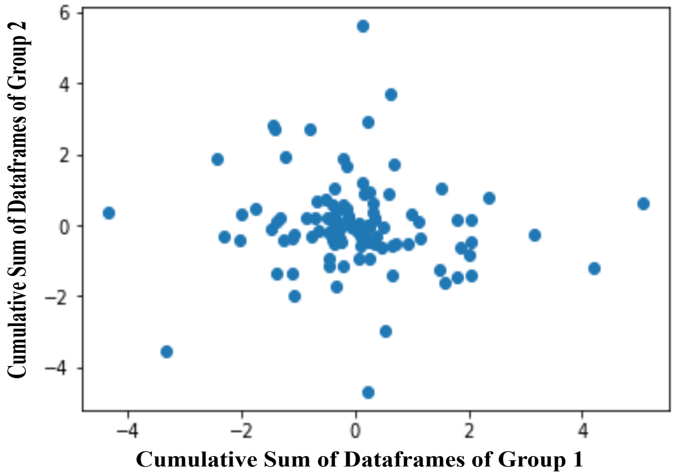

Fig. 5 Plot of waves of second group

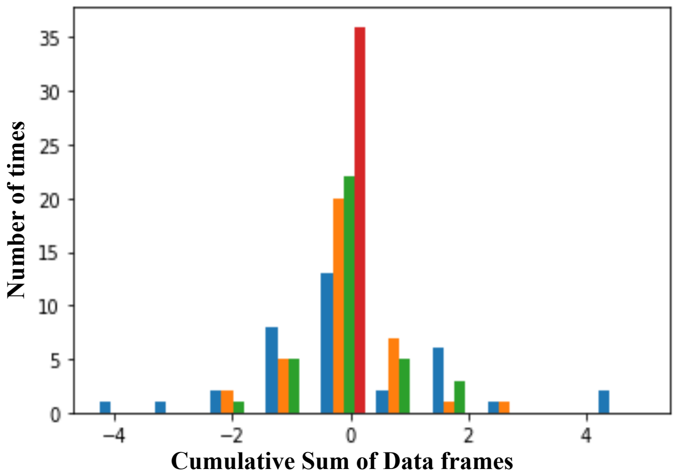

Fig. 6 Histogram plot of Cumulative sum of data frame of group 1

Table 1 is the tabular representation of the results obtained by both approaches on the collected EEG signal dataset. The performance of the proposed novel fusion mechanism has been measured in terms of MSE and Accuracy as parameters for evaluating purposes. The comparative study between the proposed methodology and existing methods is also shown graphically in Fig. 8 and Fig. 9 for MSE and accuracy respectively.

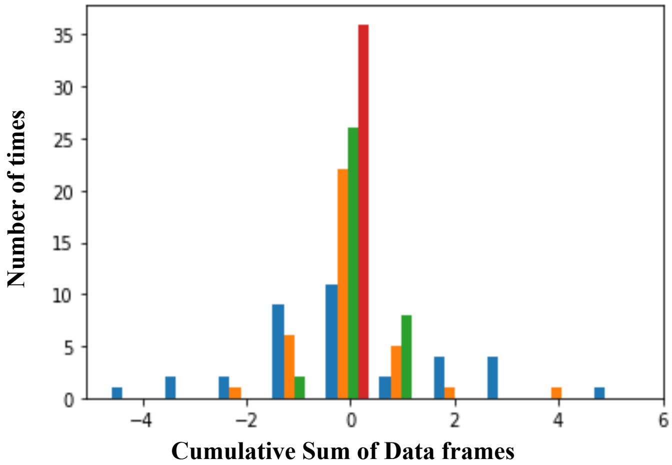

Fig. 7 Histogram plot of Cumulative sum of data frame of group 2

Table 1 Comparison of proposed method with the existing methodologies

Methods	MSE(%)	Accuracy(%)
KNN classifier and Heart rate based method [12]	20.83	79.17
SVM based method [13]	7.89	92.11
KNN classifier based approach [9]	25.57	74.43
EGG signal based approach [10]	20.00	80.00
Proposed method	5.78	94.22

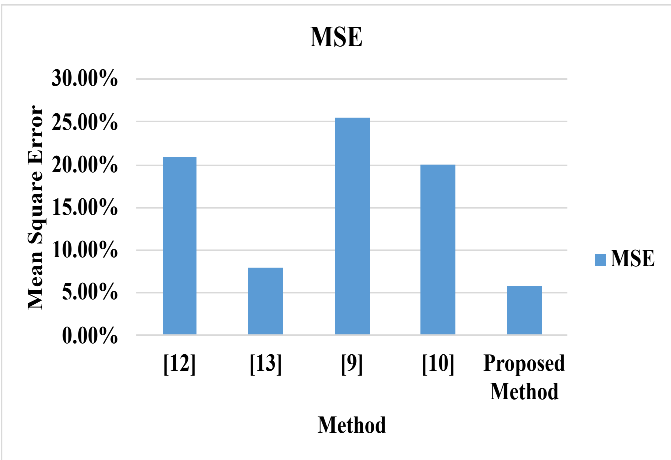

Fig. 8 Comparison between the MSE of the existing methodologies with the proposed method

4 Conclusion and Future work

In education & industry, stress detection is critical for determining instructional efficiency, improving education, and reducing the risk of human errors

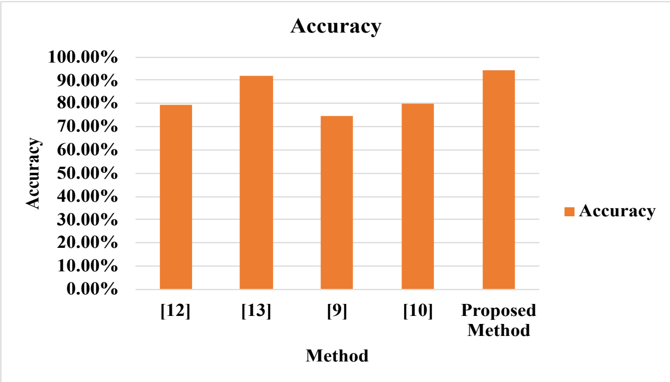

Fig. 9 Comparison between the accuract of the existing methodologies with the proposed method

caused by workers' stressful situations. In this work, we proposed a novel fusion mechanism for stress recognition based on EEG brain waves. We analyzed the principal component analysis (PCA) for selecting the features and calculated the PSD vector. At last, a voting classifier based on the k-NN and logistic regression classifier has been used for the detection of stress using the extracted features. The simulation has done using EEG signals and revealed that the proposed fusion mechanism achieved 94.22% accuracy and 5.78 MSE value. From the overall analysis, we can say that the proposed approach outperformed in comparison to the baseline k-NN approach for stress detection.

We will try to collect real time EEG data from respondents in the future and establish a database for identifying stress based on our recorded data. Most of the future effort is focused on developing an even more optimal technique for use in a variety of stress-detection tasks. Once the medical diagnoses have been established, they can be dealt with more correctly and effectively. So, future projects will have room for growth when it comes to identifying and responding to the stress revealed.

5 Data Sharing Statement

We have not recorded or generated any data for the expeirment. We have tested the proposed algorithm on the dataset publicly available at [14-15].

6 Declaration

This work has been carried out at Madhav Institute of Technology and Science, Gwalior. The funding to the same is done through Innovative Research Scheme-2020 by Madhav Institute of Technology and Science, Gwalior, India.

The authors declare that they have no conflict of interest.

The experiment work is done on a well established and explored dataset which is publicly available at [14-15].

The simulation has been done on Jupyter notebook for Python. The code will be available on request.

All authors are contributed equally in the research work, and manuscript as well.

References

[1] Suja Sreeith Panicker and Prakasam Gayathri. A survey of machine learning techniques in physiology based mental stress detection systems. Biocybernetics and Biomedical Engineering, 39(2): 444-469, 2019.

[2] Fares Al-Shargie, Tong Boon Tang, Nasreen Badruddin, and Masashi Kiguchi. Towards multilevel mental stress assessment using svm with ecoc: an eeg approach. Medical & biological engineering & computing,56(1):125-136, 2018.

[3] Prajakta Fulpatil and Yugandhara Meshram. Analysis of eeg signals with the effect of meditation. International Journal of Engineering Research & Technology (IJERT), 3(6), 2014.

[4] Omneya Attallah. An effective mental stress state detection and evaluation system using minimum number of frontal brain electrodes. Diagnostics, 10(5):292, 2020.

[5] Shamla Mantri, Vipul Patil, and Rachana Mitkar. EEG based emotional distress analysis—a survey. International Journal of Engineering Research and Development, 4(6):24-28, 2012.

[6] Sana Yasin, Syed Asad Hussain, Sinem Aslan, Imran Raza, Muhammad Muzammel, and Alice Othmani. Eeg based major depressive disorder and bipolar disorder detection using neural networks: A review. Computer Methods and Programs in Biomedicine, page 106007, 2021.

[7] Shaibal Barua and Shahina Begum. A review on machine learning algorithms in handling EEG artifacts. In The Swedish AI Society (SAIS) Workshop SAIS, 14, 22-23 May 2014, Stockholm, Sweden, 2014.

[8] Yang Si. Machine learning applications for electroencephalograph signals in epilepsy: a quick review. Acta Epileptologica, 2(1):1-7, 2020.

[9] Xiaotong Gu, Zehong Cao, Alireza Jolfaei, Peng Xu, Dongrui Wu, Tzyy Ping Jung, and Chin-Teng Lin. Eeg-based brain-computer interfaces (BCIS): A survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications. IEEE/ACM transactions on computational biology and bioinformatics, 2021.

[10] Md Mohsin Sarker Raihan, Abdullah Bin Shams, and Rahat Bin Preo. Multi-class electrogastrogram (egg) signal classification using machine learning

algorithms. In 2020 23rd International Conference on Computer and Information Technology (ICCIT), pages 1-6. IEEE, 2020.

[11] Bhavna P Harne, Yogini Bobade, RS Dhekekar, and Anil Hiwale. SVM classification of eeg signal to analyze the effect of om mantra meditation on the brain. In 2019 IEEE 16th India Council International Conference (INDICON), pages 1-4. IEEE, 2019.

[12] Prima Dewi Purnamasari, Ramdhaidtri Martmis, and Rizky Ramadian Wijaya. Stress detection application based on heart rate variability (HRV) and k-nearest neighbor (kNN). In 2019 International Conference on Electrical Engineering and Computer Science (ICECOS), pages 271-276.IEEE, 2019.

[13] Taruv Harshita Priya, P Mahalakshmi, VPS Naidu, and M Srinivas. Stress detection from eeg using power ratio. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), pages 1-6. IEEE, 2020.

[14] Igor Zyma, Sergii Tukaev, Ivan Seleznov, Ken Kiyono, Anton Popov, Mariia Chernykh, and Oleksii Shpenkov. Electroencephalograms during mental arithmetic task performance. Data, 4(1):14, 2019.

[15] Goldberger, A. L., L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. "Components of a new research resource for complex physiologic signals." PhysioBank, PhysioToolkit, and Physionet (2000).

[16] Mohammad Hossein Same, Gabriel Gandubert, Gabriel Gleeton, Preslav Ivanov, and R Landry. Simplified welch algorithm for spectrum monitoring. Applied Sciences, 11(1):86, 2021.

[17] Bernard J. Fino and V. Ralph Algazi. Unified matrix treatment of the fast walsh-hadamard transform. IEEE Transactions on Computers, 25(11):1142-1146, 1976.

[18] Rao K Yarlagadda and John E Hershey. Hadamard matrix analysis and synthesis: with applications to communications and signal/image processing, volume 383. Springer Science & Business Media, 2012.

[19] Chunmei Wang, Junzhong Zou, Jian Zhang, Min Wang, and Rubin Wang. Feature extraction and recognition of epileptiform activity in EEG by combining pca with apen. Cognitive neurodynamics, 4(3):233-240, 2010.

[20] Saloni Kumari, Deepika Kumar, and Mamta Mittal. An ensemble approach for classi

cation and prediction of diabetes mellitus using soft voting classifier. International Journal of Cognitive Computing in Engineering, 2:40-46, 2021.

[21] P Umar Sathic Ali and Dr C Jothi Ventakeswaran. Improved evidence theoretic kNN classifier based on theory of evidence. International Journal of Computer Applications, 15(5):37-41, 2011.