Voting Classifier based Model for Mental Stress Detection and Classification Using EEG Signal

Navdeep Shakya, Rahul Dubey, Laxmi Shrivastava
rahul@mitsgwalior.in
Department of Electronics Engineering, Madhav Institute of Technology & Science, Gwalior, India

Summary
According to World Health Origination (WHO), stress is a major problem of human beings which have a large effect on physical and mental health. The state of emotional tension from adverse or demanding circumstances is called stress. It can be experienced by each person in regular lifestyle due to job, some family problems, and other personal issues. Some kind of stress is important for completing the task but a lot of stress causes harm to human health. Hence, nowadays, identification of stress levels is important. This paper proposes one of the simplest methodology for the detection of stress by the analysis of EEG signal. Fast fourier transform (FFT) is used in the proposed method to generate the power spectral density (PSD) vector. PSD vector is used as an input to the voting classifier. Voting classifier is a combination of K-nearest neighbour (KNN), and random forest (RF). The proposed method achieves the highest classification accuracy of 88%.
Keywords:
Mental Stress, EEG, KNN, Random Forest, Voting classifier.

1. Introduction

The human body's response to unusual mental conditions is considered as stress [1]. In healthy people, there is a balance between parasympathetic and sympathetic arms of the autonomous nervous system. When an abnormal condition is experienced, a flight response is invoked. Stress leads to a persistent feeling of energy efficiency and depression.

Stress is classified into two categories namely, routine stress and chronic stress. Routine stress induces inefficiency and makes it a social cause while chronic stress adversely affect the anxiety and bipolar diseases. It is also well proved fact that physical health is also affected by mental stress [2]. At the present, the main cause of most medical conditions is mental stress. These comprise strokes, heart attacks, depression, nervousness, PTSD (Post-Traumatic Stress Disorders) and immunological disorders [3-4]. Counselling is a conventional method to detect stress in human but this suffers from a disadvantage that it requires a person to be able to speak frankly. Stress can also affect the activity and structure of the brain. The study of brain activity is necessary for the early detection of stress to avoid disease, and to decrease the risk of brain damage [5].

EEG is one of the most accurate measurement of electrical activity in the brain. The voltage changes in the brain neurons are used to calculate the voltage variation caused by the ionic current. EEG is used for identifying a stroke, a focal brain disorder, and tumors as a first-line method [6]. EEG will track the long-term sleep stage or epilepsy near the bed of the patient's bio signals [6]. During various phases of life, EEG may detect brain changes without distracting a patient, e.g., EEG Sleep Analysis [1]. EEG signal is classified into two types; healthy EEG signal and unhealthy EEG signal. Abnormal activity of the brain can be determine using unhealthy EEG signal. Classification of EEG signals into respective classes of healthy and unhealthy EEG has already been explored by various researchers in the literature [7].

Sub-band power ratio is used to classify the EEG signals into relaxed and stress classes of EEG signals [8]. Artificial neural network based approaches has been used in next stage to classify the EEG signals into two respective classes. 10-fold cross-validation method was utilized for the performance evaluation of the proposed features using support vector machine (SVM) and K-nearest neighbor. In [9]; author proposed that beta brain waves are highly recommended to detect the stress using EEG signals. Authors consider various stress level to detect. Average intra-subject classification accuracy of 85.6% on stress vs. rest, 71.2% on two stress and rest levels, and 58.4% on three stress and rest levels were obtained using Linear discriminant classifier with 10-fold cross-validation.

In [10]; the researchers examine the use of EEG for predicting people's stress levels without manual intervention and recording through BCI (Brain-Computer Interface). The efficiency of SVM and KNN Machine learning algorithms is evaluated. A training set, target class is allocated by a stress value derived from the PSS-14 (Full form of PSS) questionnaire response. Using the KNN algorithm, the maximum average classification accuracy is 74.43%. Author confirmed that band power ratios of different bands are correlated to the stress level of subjects from measured EEG signals in the front part of the brain.

In [11]; Signals coming from the human brain frontal lobe are used for stress measurement. The brain waves of the thirty subjects are recorded with five different levels of stress: relaxed, stressed less, stressed moderately, stressed higher, and stressed alarmingly. Preprocessing of the recorded EEG signal is followed by the feature extraction.

Manuscript received November 5, 2024
Manuscript revised November 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.11.12

Feed forward neural network is used to identify the stress in the brain. FFT was utilized in the proposed method in [12] for feature extraction. KNN was used for classifying and detecting stress. The proposed features were most appropriate for stress detection and mediation applications. Theta, delta, beta, and alpha waves were used with the best value for K=3 in KNN with 80% classification accuracy [12].

The application of IoT technologies in the healthcare sector with the cloud to assess the physical and mental activity of the brain using an EEG headset is proposed in [13-14]. Traditional EEG headsets depend upon multi-electrode invasive EEG, which stresses and concerns patients with the use of such headsets. The proposed architecture model will gather data through car racing competition using a new, convenient, easy-wear EEG headset. Such data would be forwarded to Cloud to analyze and display brainwave levels to control robotic cars to reach a certain direction of difficulty. This concept makes it simple to use and flexible for many changes and developments in the complexity of the traditional mental assessment using the Neuro-sky Mind-wave EEG headset. The main objective of the proposed method in [15-16] is to incorporate the concept of a brain race for the diagnosis and assessment of mental status as a new transition from an evaluation of entertainment and brain activity in various scenarios.

2. Methodology

The purpose of this paper is to use the auxiliary classifier, a more efficient voting classifier to classify the EEG signals into healthy and unhealthy EEG signal. The used auxiliary classifier is based on the KNN and RF classifier to increase the accuracy of the current model for a greater number of datasets. A computer-aided diagnostic tool is highly required to help people to diagnose stress and to suggest proactive action to minimize stress. In the proposed method; FFT is used to generate the periodogram of input EEG signal. PSD vector value from periodogram has been calculated and used as an input to the proposed voting classifier. The block diagram of the proposed methodology is shown in Fig. 1.

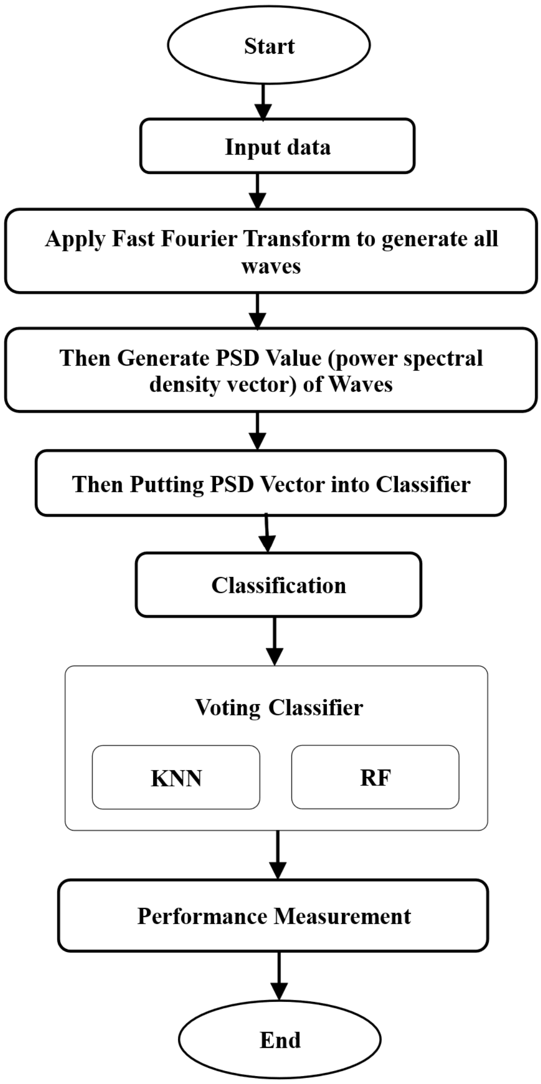

Fig 1. Flow chart of the proposed method

2.1 Fast Fourier Transform (FFT**)**

FFT is used to transform the signal from time domain to frequency domain [15]. The FFT is called the computationally efficient implementation of the DFT method, described in a set of N samples x_n as follows:

$$X_k = \sum_{n=1}^{\infty} x_n * e^{-j2\pi kn/N}, k = 0,1,2, \ldots\ldots, N-1 \quad (1)$$

Where X_k is a complex number known as an absolute value (frequency magnitude or modulus). X_k is resulting sequence defined as follows: positive frequencies $0 \leq f < F_s/2$ correspond to values $0 \leq n \leq (N-1)/2$, while negative frequencies $-F_s/2 < f < 0$ correspond to $(N+1)/2 \leq n \leq (N-1)$. F_s denotes sampling rate. Results achieved are considered voice signal frequency spectrum [17].

2.2 Power Spectral Density (PSD)

The power spectral density (PSD) demonstrates the intensity of the energy variations as a frequency function. In other words, it demonstrates the strong frequency variations and the weaknesses of frequencies. The PSD unit consists of energy by frequency (width) which you can get through the integration of PSD in a particular frequency range. PSD is computed directly using the FFT or calculating autocorrelation function and then transformed. In this work, the PSD vector generation is done using the Welch periodogram method for group 1 and group 2. The Welch algorithm [18] is a non-parametric way of estimating PSD which making the frequency spectrum smoother than the raw FFT flow.

2.3 Random Forest

Leo Breiman from California University first suggested the random forest in 2001. It consists of a large number of basic classifiers (decision trees) that are fully independent of each other. A test sample for this classification will be determined depends upon voting outcomes of every single classification and class label for this sample [19-20].

The key steps to building up the random forest classifier are as follows:

1. Specifies a no. of elements of every feature subset "M".
2. New feature subset h_k is created from the entire feature set at random based on the M value. h_k is independent by another subset in a sequence of $h_1, h_2, \dots\dots, h_k$.
3. Training the data set for each training set group with a feature sub-set to construct a decision tree. Every single classifier can be h (x, h_k) (where x represents inputs).
4. Select new h_k also repeat the above process until feature subsets have been traveled. A random forest classifier is done.

2.4 k-Nearest Neighbour

KNN method is also called a non-parametric method finds application in mathematical problem since the beginning of the 1970s. The basic theory behind KNN is that set of k samples are found nearest to unknown samples in the calibration dataset (for example, depends upon distance functions). The average response variables are calculated from these k samples by measuring the label of unknown samples. As a consequence, it can be concluded that the efficiency of KNN is dependent on the value of k. This means that k is the primary tuning parameter in KNN [19].

2.5 Voting classifier

A voting classifier also called auxiliary classifier is a machine learning based approach. This classifier is based on the maximum likelihood of the output from the chosen class, and trains on an ensemble of various examples. Rather than creating diverse special models for every one of them, we construct a solitary model which prepares and predicts execution dependent on the aggregate vote larger part of the exhibition gatherings. Random forest and KNN (typically varying types) and basic statistics (such as the average) are introduced in our proposed projects. The architecture of the proposed voting classifier is shown in Fig. 2.

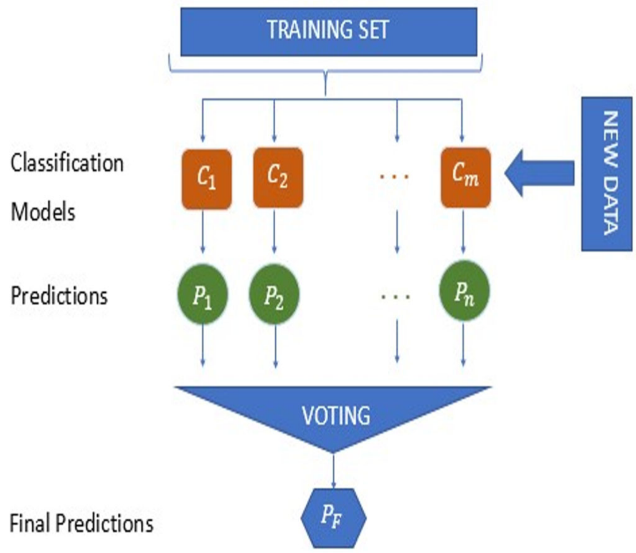

Fig 2. Architecture of a proposed voting classifier

3. Results & Discussion

This proposed methodology has been implemented using Jupyter notebook in python programming. The dataset used in the proposed method has EEG signal waves that include delta, beta, gamma, alpha, and theta waves. The test results and efficiency of the proposed voting classifier is discussed in this section. In the present study, the experiment has been conducted in two groups. There are 72 files in these groups. These 72 files have been arranged in healthy and unhealthy EEG signal with 36 signals of each class. We have applied FFT on both groups to generate PSD vector for delta, beta, gamma, alpha, and theta waves for these groups. In the last stage, an auxiliary voting classifier has been used in the experiment. The experiment performs better computational efficiency when used with the value of K=3.

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 [21].

Fig. 3 and Fig. 4 are the representation of EEG bands after applying the Fast Fourier Transform (FFT) for the 1st and 2nd groups. It shows values of EEG bands that are Theta, Delta, Beta, Alpha and Gamma.

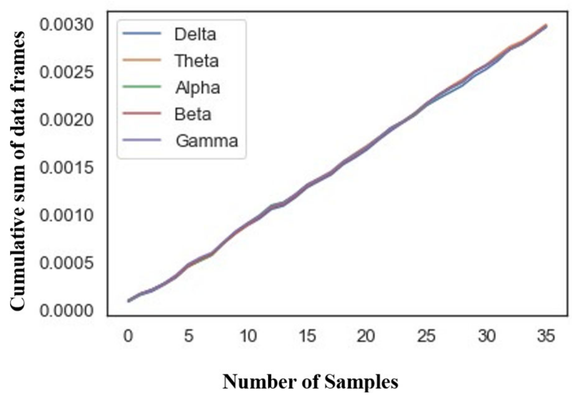

Fig 3. Plot of Cumulative sum of data frame of group 1 against samples

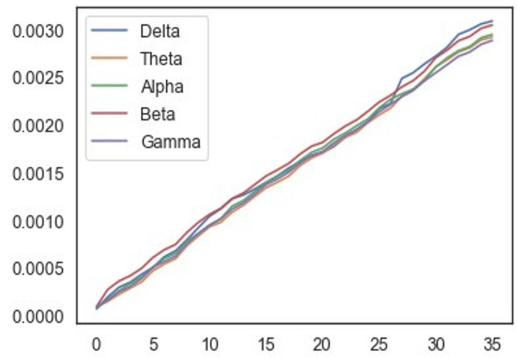

Fig 4. Plot of Cumulative sum of data frame of group 2 against samples

Fig. 5 represents a scatter plot of cumulative sum of data frames for both groups after removing the gamma waves because it will not affect the accuracy of our model. Fig. 6 and Fig. 7 shows the histogram plot of the cumulative sum of data frames for the 1st and 2nd group respectively.

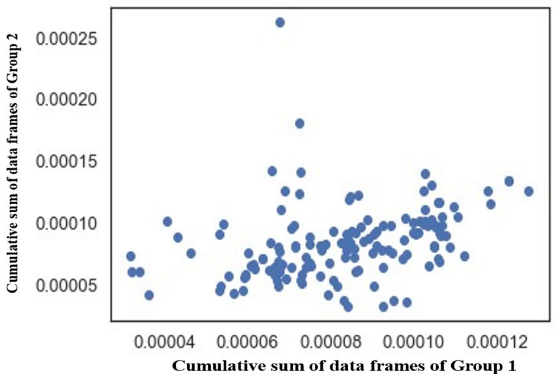

Fig 5. Scatter plot of cumulative sum of data frame of group 1 against group 2

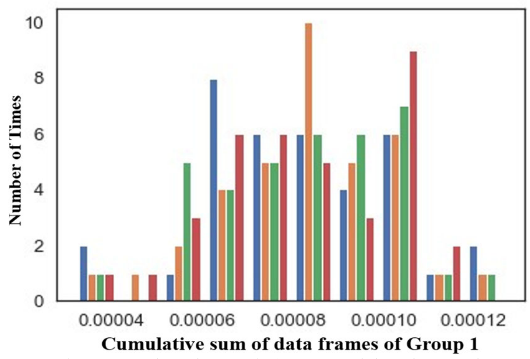

Fig 6. Histogram plot of Cumulative sum of data frame of group 1

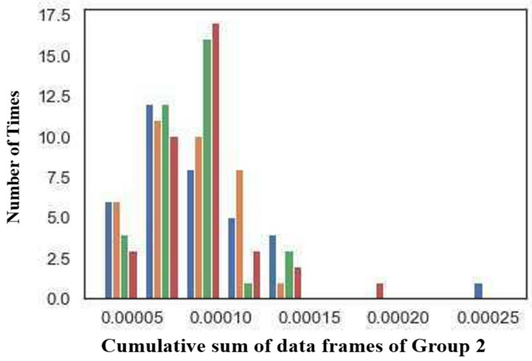

Fig 7. Histogram plot of Cumulative sum of data frame of group 2

Comparative analysis of results of the proposed method with the existing method. Mean square error (MSE) and accuracy is used as a performance parameter of the proposed method. It is proved from the Table 1 that the proposed method performs better in terms of performance parameter with respect to existing methodology. 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.

Table 1. Comparison of proposed method with the existing methodologies

Method	MSE (%)	Accuracy (%)
SVM and KNN classifier-based method [9]	14.40	85.60
Brain mapping based method [10]	25.77	74.33
KNN classifier-based approach [12]]	19.95	80.04
Proposed Method	11.41	88.58

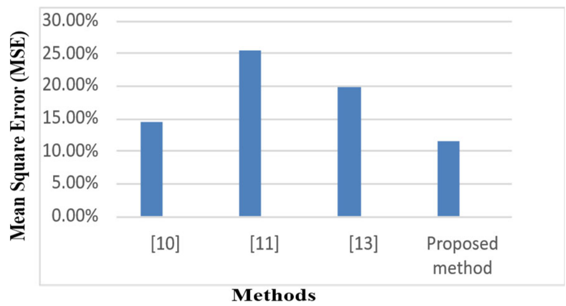

Fig 8. Comparison between the MSE of the existing methodologies with the proposed method

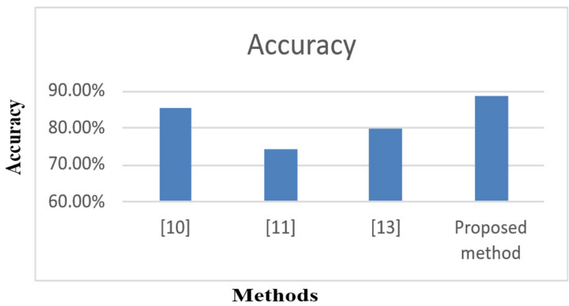

Fig 9. Comparison between the accuract of the existing methodologies with the proposed method

4. Conclusion

The work present in this paper proposed a method to detect and classify stress using EEG signals. It is based on the features extracted from the EEG waves Delta, Alpha, Beta, Theta, and Gamma. A voting classifier model is used in the proposed method to successfully detect stress at different levels. EEG signal is subjected to FFT to generate periodogram waves which used to generate the PSD vector. PSD vector is used as an input to the proposed auxiliary voting classifier. The voting classifier used in the experiment is based on KNN and RF. Experimental results have found that the proposed voting classifier model is better than the previous KNN classifier model in terms of accuracy and MSE. It has achieved 88.58% accuracy and 11.42% mean squared error. The main contribution of the paper is the use of voting classifier to detect and classify stress. In future work, we can use other biomedical signals such as ECG, EMG etc. to detect the stress using the same methodology. Moreover, intervention and sensitivity analysis methods on EEG signals can also be added to determine the stress level.

References

[1] S. M. U. Saeed, S. M. Anwar, H. Khalid, M. Majid, U. Bagci, EEG based classification of long-term stress using psychological labeling, Sensors 20 (2020) 1886.

[2] M. S. Kalas, B. Momin, Stress detection and reduction using eeg signals, in: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), IEEE, 2016, pp. 471-475.

[3] O. Attallah, An e_ective mental stress state detection and evaluation system using minimum number of frontal brain electrodes, Diagnostics 10 (2020) 292.

[4] A. R. Gaurav, V. Kumar, Eeg-metric based mental stress detection, Netw Biol 8 (2018) 25-34.

[5] S. W. Lazar, C. E. Kerr, R. H. Wasserman, J. R. Gray, D. N. Greve, M. T. Treadway, M. McGarvey, B. T. Quinn, J. A. Dusek, H. Benson, et al., Meditation experience is associated with increased cortical thickness, Neuroreport 16 (2005) 1893.

[6] M. N. B. Patil, R. Mirajkar, S. Patil, P. Patil, A method for detection and reduction of stress using eeg, International Research Journal of Engineering and Technology (IRJET) 4 (2017) 1528-1604.

[7] P. Youse_ Zowj, Topics on Multiresolution Signal Processing and Bayesian Modeling with Applications in Bioinformatics, Ph.D. thesis, Georgia Institute of Technology, 2021.

[8] T. H. Priya, P. Mahalakshmi, V. Naidu, M. Srinivas, Stress detection from eeg using power ratio, in: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), IEEE, 2020, pp. 1-6.

[9] Y. Zhang, Q. Wang, Z. Y. Chin, K. K. Ang, Investigating different stress-relief methods using electroencephalogram (eeg), in: 2020 42nd Annual International Conference of the

IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 2999-3002.
[10] X. Gu, Z. Cao, A. Jolfaei, P. Xu, D. Wu, T.-P. Jung, C.-T. 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).
[11] P. Mukherjee, A. Roy, Detection of stress in human brain, 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP) (2019) 1-6.
[12] P. D. Purnamasari, A. Fernandya, Real time eeg-based stress detection and meditation application with k-nearest neighbor, in: 2019 IEEE R10 Humanitarian Technology Conference (R10-HTC)(47129), 2019, pp. 49-54. doi:10.1109/R10-HTC47129.2019.9042488.
[13] A. Mansour, H. T. Ouda, On the road to a comparative car racing eeg-based signals for mental and physical brain activity evaluation, in: 2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON), 2019, pp. 43-48. doi:10.1109/IEMECONX.2019.8877037.
[14] J. Mladenovi_c, Standardization of protocol design for user training in eeg-based brain{computer interface, Journal of Neural Engineering 18 (2021) 011003.
[15] J. L. M. Kumar, M. Rashid, R. M. Musa, M. A. M. Razman, N. Sulaiman, R. Jailani, A. P. A. Majeed, An evaluation of different fast fourier transform transfer learning pipelines for the classification of wink-based eeg signals, MEKATRONIKA 2 (2020) 1-7.
[16] R. S. Alkhawaldeh, Arabic (indian) digit and written recognition using recurrent transfer deep architecture, Soft Computing 25 (2021) 3131-3141.
[17] M. H. Same, G. Gandubert, G. Gleeton, P. Ivanov, R. Landry, Simplified welch algorithm for spectrum monitoring, Applied Sciences 11 (2021)
[18] A. Parmar, R. Katariya, V. Patel, A review on random forest: An ensemble classifier, in: J. Hemanth, X. Fernando, P. Lafata, Z. Baig (Eds.), International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018, Springer International Publishing, Cham, 2019, pp. 758-763.
[19] N. Zaabar, S. Niculescu, M. Mihoubi, Assessment of combining convolutional neural networks and object based image analysis to land cover classification using sentinel 2 satellite imagery (tenes region, algeria), International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 43 (2021) 383-389.
[20] D. Shon, K. Im, J.-H. Park, D.-S. Lim, B. Jang, J.-M. Kim, Emotional stress state detection using genetic algorithm-based feature selection on eeg signals, International Journal of environmental research and public health 15 (2018) 2461.

Navdeep Shakya is the research scholar in the Department of Electronics Engineering, Madhav Institute of Technology and Science. He has done Bachelor of Engineering from <> in Electronics and Communication Engineering. He is pursuing Master of Engineering in Communication, Control, and Networking from Madhav Institute of Technology and Science, Gwalior. His research interest is mainly in the field of Signal processing and Biomedical engineering.

Dr. Rahul Dubey is currently associated with Madhav Institute of Technology and Science, Gwalior. He is working as an assistant Professor in the Department of Electronics Engineering. He has more than 10 years of experience in the field of academics and research. His research interest is in the field of Machine fault diagnosis, Signals and System, Biomedical Signal processing.

Dr. Laxmi Shrivastava is currently associated with Madhav Institute of Technology and Science, Gwalior. She is working as an Associate Professor in the Department of Electronics Engineering. He has more than 25 years of experience in the field of academics and research. His research interest is in the field of Wireless Sensor Networks, Biomedical Signal processing.