

CALR: Classroom Attention Level Recognition System For Monitoring The Status Of Students' Thoughts During A Lecture

Sakib Hasan¹, Fatima-Tuz-Johairah², Anika Fahmida³, Tanmoy Pias⁴ and David Eisenberg^{*5}

University of Asia Pacific and Rajshahi University of Engineering and Technology, Bangladesh & Virginia Tech, USA
{sakib.hasan¹, 16101008² and tsr⁴} @uap-bd.edu, anikafahmidahossain29@gmail.com³, davide77@vt.edu^{*5}

Abstract—Have you ever thought about how much your students are concentrating on your lecture? Many of your students may have pretended to be paying attention in class, and it may have been difficult to know which students were attentive and which were not. Machine Learning-based human attention level recognition has gained interest recently, but the attention level of a group of people (such as a classroom) has not received much interest as of yet. This paper proposes a cost-effective, hassle-free Classroom Attention Level Recognition system (CALR) to accurately identify the cumulative attention level of a classroom. This system will not only serve the purpose of helping a teacher monitor her lecture, but also will act as a quality assurance system for an educational institution. Single-channel, time-frequency scalp-EEG signals are recorded from every student by using a lightweight headset during a lecture, which was constructed by using the Bitalino EEG Sensor board, and then used to classify the cumulative attention level of the classroom among Focused, Distracted or Neutral states at any given instance. These EEG signals are pre-processed to remove noise, such as muscle movement. Statistical coefficients (i.e. mean, standard deviation, skewness, kurtosis, and entropy) and statistical wavelet transformation are used to extract meaningful features from the EEG signal. We mainly used two multi-scale Wavelet Packet Statistics (WPS) and multi-scale Wavelet Packet Energy Statistics (WPES) to generate the feature vector that trained our state-of-the-art model to recognize human attention level for a specific person. Our system thus captures a number of human attention levels at a given time, and a total is calculated to determine the percentage of the class that is actually focused on the lecture, and which portion of the class is distracted.

Index Terms—EEG(Electroencephalography), Attention Level, EEG Decomposition, Deep Learning etc.

I. INTRODUCTION

Studying brain signals requires great effort from researchers. Over the past decade not only have we seen the use of Machine Learning (ML) and Deep Learning (DL) to analyze EEG (electroencephalograph) signals, but have also seen their implementation in identifying diseases and emotional states of the human brain. *Jalilifard* [2] used ML algorithms to classify human emotional states. *Lin* [5] addressed the changes in emotional states while listening to music using ML, and *Peng et. al.* [6] used manifold extreme learning machines to detect human emotional states using EEG.

We can also see *Leracitano et. al* [1] described a ML system to identify brain states to determine whether subjects had

Alzheimer's disease, Mild Cognitive Impairment or healthy control mental states. Meanwhile, *Bhardwaj et. al.* [9] depicted a comparative analysis of ML and DL algorithms, while classifying the cognitive states during a driver's fatigue using EEG wave-streams. Another informative study was done by *Jeevan* [4] to attempt using more advanced algorithms, such as LSTM-RNN in emotional recognition. Many studies in this domain are designed to detect diseases or classify emotional states of the human brain can be seen in [7], [8], [10], [11].

Although a lot of studies took place in the domain of emotional classification or disease identification using EEG signals, almost no investigation was entertained in recognizing the attention levels of the brain. Whatever a person does results in a change in the electron flow across billions of neurons in the brain. With changing attention levels, a significant change in the EEG signals should be found. Hence, a particularly high number of attention-related signals can be found in students.

In this work, a community approach is proposed for a Classroom Attention Level Recognition System. Our method for identifying individual human attention levels uses a complex hybrid model, such that CNN-LSTM with Brain-EEG is utilized as the backbone.

II. ELECTROENCEPHALOGRAPHY(EEG): INS AND OUTS

Electroencephalography captures the wave signals generated by the electrical signals flowing across an accumulation of 100 billion neurons in the human brain.

A. Composition of a raw EEG

In the human sensory system [12], the electrical sign of a solitary neuron is too small to ever be estimated by an anode on the scalp. Thus, its activity originates from the summation of the electrical voltage fluctuations of numerous neurons on the cathode. In the event that the electrical field is estimated, the result is called an EEG. Some of the time, a collection of neurons fire synchronously, such that the captured signal becomes more powerful and increasingly synchronous. This is the reason why these signals are viewed as brainwaves and can be described in frequency sub-bands (Fig. 1).

A raw EEG signal is comprised of frequency bands: Gamma(>30Hz), Beta(13-30Hz), Alpha(8-12 Hz), Theta(4-8Hz), and Delta(<4Hz). However, due to unwanted muscle movements or external influences, a raw EEG captured from

*Corresponding Author : davide77@vt.edu

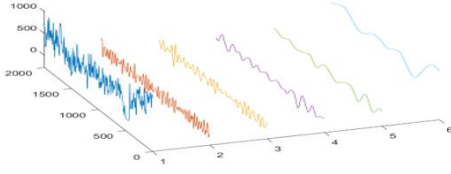


Fig. 1. 1-Raw, 2-gamma, 3-beta, 4-alpha, 5-theta, 6-delta. The length of the signal is 2000 and the z-axis represents the amplitude of the signals.

the human face contains significant interference from noise. Thus, to acquire the sub-bands from EEG, we need to refine the raw signal first [3].

B. Acquire the Sub-Bands

Although a raw approach was described well in [3] [4], in our experiment the raw EEG was decomposed into sub-bands using a 1-D wavelet decomposition function in Matlab. The signal was decomposed into 8 levels where each level represents different features. The detailed coefficient of levels 1 to 4 represents noise. So, these four bands were removed from the analysis. The rest of the detail coefficients of wavelet decomposition levels 5 to 8 and the co-efficient approximation of level 8 represent desired sub-bands namely- $\alpha, \beta, \gamma, \theta, \delta$ (Fig. 1).

III. HUMAN ATTENTION LEVEL WITH EEG & CAL

What does Human Attention Level actual mean? Put simply, it means to recognize whether a person is attentive or not to a particular task. In a classroom, the attention level is simply Attentive or Inattentive. We found the EEG signals from these two types of students, at any particular time, varies from one another. In this work, we are addressing the average state of attention among students, i.e. Classroom Attention Level(CAL). By consensus, we found that a student can be in one of four cognitive states while attending a lecture in their classroom:

- **Completely Observant** refers to students focusing solely on the lecture materials discussed.
- **Completely Non-observant** refers to students who are not at all interested in the lecture, and doing something else during the lecture.
- **Pseudo Observant** refers to those students who pretend to be attentive in class with a constant stare, but s/he is not actually paying attention.
- **Moderately Non-observant** refers to students who are trying to maintain his/her focus on the class, but fail to do so either because of surrounding influences or difficulty catching up with the speed of the lecture.

IV. PROPOSED WORK

A lot of classification mechanisms have been previously devised [1], [2], [4], [9], [10] to classify either human emotions, human sleep stages, or diseases etc. At this point, no one has concentrated on this particular target audience, which is “Classroom Students”. Human Attention Recognition

using EEG signals is found to be similar to Human Emotion Classification from EEG.

In this paper, a Classroom Attention Level Recognition (CALR) system, which enables a teacher to attain EEG raw signals from individual students, is proposed. This system, after performing an individual analysis for attention recognition, will provide an average Classroom attention level as an output to the teacher. All of these processes will be performed without the students’ knowledge, so that they cannot cheat. Since we do not want the students to know our process, wireless transmission technology to capture the EEG signals [12] must be used.

V. ARCHITECTURE OF THE PROPOSED SYSTEM

Our proposed architecture is a community-based implementation of an individual human attention level recognition model, which has been developed by us with a view towards building software. We are excited to state that we have achieved an accuracy of 89% in identifying the attention level of one person using CNN-LSTM [13]. Our aim is to capture the raw EEG signals from numerous students at random instances. Our system aims to capture those signals, process them, identify the attention level, and produce a final output for the whole class.

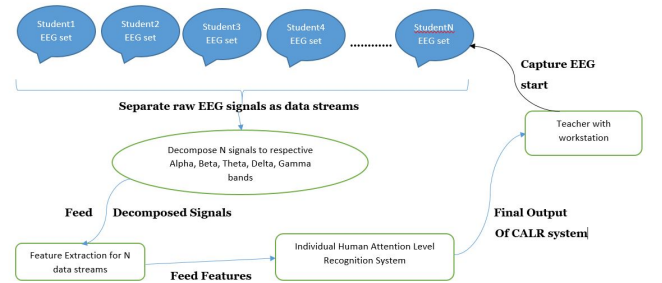


Fig. 2. Architecture of Proposed System

The system designed will deliver the expected result if it runs through the following steps successfully:

- Setup of the wireless EEG device for students.
- Capturing of the raw EEG signals from students wearing the device, with the EEG device activated for at least about ten seconds, with those students unaware that their data is being attained.
- Processing the raw EEG signals(fed wirelessly) and identifying each student’s the attention level.
- Finally, the percentage of attentive students is calculated to recognize the average attention level of the whole class.

VI. SYSTEM SETUP & FULL EXPERIMENTATION

A. Device Selection & Setup for Signal Acquisition

Muse2, Brain-sense, Neurosky, or E-motiv are a series of expensive devices that can wirelessly transmit a raw EEG signal captured from one person [12]. To minimize the costs of implementation, we built a new device using the Bitalino

main board. A BT (Bluetooth) module was attached to the motherboard of the device for wireless transmission, and resulted in a less expensive product. A device, when placed on a headband, can be strapped to each student.

B. Application on the Teacher's Workstation

A running application on the teacher's workstation is necessary for our system. The front-end teacher's display simply features a PUSH button and a GUI for the final output. All of the main works take place at the back-end of the system. The application enables the faculty to capture the individual raw EEG signals, keeping the students completely unaware, while it pre-processes the signals for feature extraction. The application recognizes if an individual student is attentive or not, through our Machine Learning-based Human Attention Recognition model. Finally, the model outputs the average classroom attention level. In the next subsections, we will discuss in detail our developed model and how it recognizes an individual EEG signal's attention level with an accuracy of 89%.

C. Individual Signal Pre-processing

While capturing the EEG data, we noted the start and end of the attentive time frame. After a few tests, we observed that the time frame typically continued for one to two seconds. The individual's attention states were categorized into four classes: neutral, attentive, happy, and boring. The attention state was then separated from the crude EEG data. The sampling rate was 1000Hz and the segment length was two seconds, which consisted of 2,000 data points. Every session (120 seconds) was segmented into 20 slots (consisting of 2,000 data points), which represent an active attention zone, which has been explained in more detail in a previous paper [13]. We took samples from twenty different students, twice each, which resulted in forty raw signals, with each raw signal's capture time being 480 seconds.

D. Feature Extraction

After splitting up the raw signals into our desired segments, we broke down each of them into corresponding sub-bands. Fig. 1. demonstrates the decomposition of a raw EEG segment in 3-D. II-B states how we developed our required sub-bands, which borrowed from a similarly designed previous approach [13].

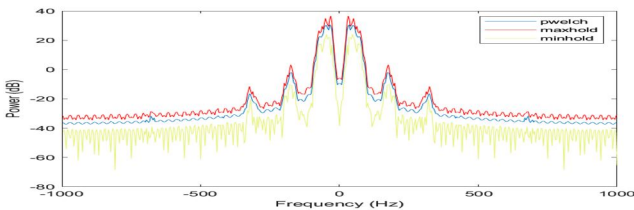


Fig. 3. Welch's Power Spectral Density Estimate

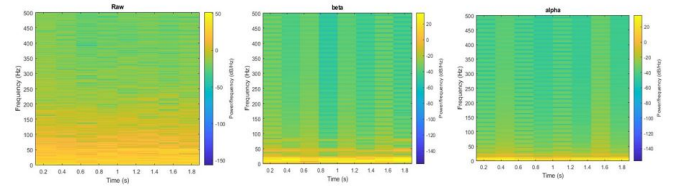


Fig. 4. Spectrogram of Raw, Beta and Alpha signal

For feature extraction, we applied two functions on those sub-signals. One is *Welch's Power Spectral Density Estimate (WPSDE)* and another function is *Spectrogram* using a short-time Fourier transform (STFT). The resulting feature estimation could be visualized from Fig. 3 & 4. Although, Spectrogram can be found out just with a simple equation like,

$$S(t, w) = |STFT(t, w)|^2 \quad (1)$$

Here, w is the width of the window for traversing the input signal $s(t)$ where t is the time instances. Side by side, the power spectral density can be generalized to discrete time variables x_n . As above we can consider a finite window of $1 \leq n \leq N$ with the signal sampled at discrete times $x_n = x(n\Delta t)$ for a total measurement period $T = N\Delta t$. Then a single estimate of the PSD can be obtained through summation rather than integration:

$$\tilde{S}_{xx} = \frac{(\Delta t)^2}{T} \left| \sum_{n=1}^N x_n e^{-i\omega n \Delta t} \right|^2 \quad (2)$$

Here, ω is the window size for windowing, and N is the total no. of data points in the truncated signal \mathbf{X} . x_n is the discrete time points from the truncated signal from the original raw EEG signal and then converted to frequency domain by STFT(short time fourier transform), and T being the total time period that signal is taken for.

E. Recognizing Individual Attention Level

Our backbone model is a Complex Deep Learning model developed with CNN LSTM algorithms [13]. The LSTM is used for sequence classification, while the visual data-set consists of Spectrogram of raw EEG, beta and alpha signal. The developed novel deep learning model used CNN and LSTM as a backbone, and acquired an astounding accuracy of 89%, which could differentiate among Neutral, Attentive, Happy or Sad levels. In previous research [14], [15] a similar artificial neural network was used for signal classification.

F. Final Output

In the VI-C to VI-E sections, we discussed the backbone of our proposed system, the individual recognition of attention levels from a single raw EEG signal.

Suppose we took 'N' students' EEG signals at any particular moment, and fed them into our system. In that case, our application would receive a separate number(N) of data streams as input. Passing these data streams through our model, we

would obtain an attention level for each of those streams. For CALR system, we will refer to a stream as **Attentive**, if it is recognized as 'Attentive' by our model, and **Inattentive** otherwise. From N students, we will refer to 'X' as the no. of students that are paying attention, and 'Y' as the no. of students who are not. Thus, our final output from the system will be determined by two metrics, P and Q, as follows:

$$\text{AttentivePercentage}, P = X/N * 100\% \quad (3)$$

$$\text{InattentivePercentage}, Q = Y/N * 100\% \quad (4)$$

At the end of all these processes, the system will provide feedback to the teacher and display the percentage of attentive and inattentive students.

VII. VIRTUAL SETUP SIMULATION

As we are yet to build the application, a real world simulation using the Python platform was run for five different scenarios, to discover if the system delivers an approximate result. In a particular scenario, we took ten separate EEG streams (N1 to N10) that we derived from our individual attention recognition data-set. We passed the separate EEG streams through our model after decomposing them to their separate sub-bands, and calculated final output with the help of equations (3) and (4), and achieved the following results.

Scenarios	No. of raw EEG streams(Total 14)				Expected P, Q	Achieved P, Q
	Attentive	Neutral	Happy	Bored		
1	10	1	2	1	71%, 29%	71%, 29%
2	8	2	2	3	58%, 42%	42%, 58%
3	0	4	3	8	0%, 100%	8%, 92%
4	5	3	3	3	36%, 64%	42%, 57%
5	13	0	0	1	92%, 8%	100%, 0%

Fig. 5. Virtual Simulation Output

The fluctuations between the expected output and the achieved output is due to the 89% accuracy of our individual attention level model; however, Scenario (1) states the potential of our developed model, with an increased number of instances in the data-set and an increased number of features for the the deep learning based model.

VIII. CONCLUSION

In our individual attention recognition model, we have already achieved 89% accuracy with samples from twenty participants and two types of features. We believe that with a larger data-set, which would serve as a next-step for this research and a subsequent paper, our model will achieve an even more significant result. Later work will also add features for separating the percentages of every attention level possible on our workstation application.

ACKNOWLEDGMENT

This research is supported financially by the Institute for Energy, Environment, Research and Development(IEERD), University of Asia Pacific located in Dhaka, Bangladesh. Therefore, we would like to express our heartfelt gratitude to them for the financial support we received for the project.

REFERENCES

- [1] Ieracitano, Cosimo, Nadia Mammone, Alessia Bramanti, Silvia Marino, Amir Hussain, and Francesco Carlo Morabito. "A Time-Frequency based Machine Learning System for Brain States Classification via EEG Signal Processing." In 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. IEEE, 2019.
- [2] Jalilifard, Amir, Ednaldo Brigante Pizzolato, and Md Kafiul Islam. "Emotion classification using single-channel scalp-EEG recording." In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 845-849. IEEE, 2016.
- [3] C.Jaganathan, A.Amudhavalli, T.Janani, M. Dhanalakshmi, Nirmala Madian. "Automated algorithm for extracting Alpha, Beta, Delta, Theta of a human EEG." In 2015 International Journal of Science, Engineering and Technology Research (IJSETR) (Vol. 4, Issue 4)
- [4] Jeevan, Reddy Koya, SP Venu Madhava Rao, Pothunoori Shiva Kumar, and Malyala Srivikas. "EEG-based emotion recognition using LSTM-RNN machine learning algorithm." In 2019 1st International Conference on Innovations in Information and Communication Technology (ICI-ICT), pp. 1-4. IEEE, 2019.
- [5] Lin, Yuan-Pin, Chi-Hong Wang, Tzyy-Ping Jung, Tien-Lin Wu, Shyh-Kang Jeng, Jeng-Ren Duann, and Jyh-Horng Chen. "EEG-based emotion recognition in music listening." IEEE Transactions on Biomedical Engineering 57, no. 7 (2010): 1798-1806.
- [6] Peng, Yong, Jia-Yi Zhu, Wei-Long Zheng, and Bao-Liang Lu. "EEG-based emotion recognition with manifold regularized extreme learning machine." In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 974-977. IEEE, 2014.
- [7] Liao, Chung-Yen, and Rung-Ching Chen. "Using Eeg Brainwaves And Deep Learning Method For Learning Status Classification." In 2018 International Conference on Machine Learning and Cybernetics (ICMLC), vol. 2, pp. 527-532. IEEE, 2018.
- [8] Bashivan, Pouya, Irina Rish, Mohammed Yeasin, and Noel Codella. "Learning representations from EEG with deep recurrent-convolutional neural networks." arXiv preprint arXiv:1511.06448 (2015).
- [9] Bhardwaj, Rahul, Swathy Parameswaran, and Venkatesh Balasubramanian. "Performance Comparison of Machine Learning and Deep Learning While Classifying Driver's Cognitive State." In 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), pp. 89-93. IEEE, 2018.
- [10] Amin, Hafeez Ullah, Wajid Mumtaz, Ahmad Rauf Subhani, Mohamad Naufal Mohamad Saad, and Aamir Saeed Malik. "Classification of EEG signals based on pattern recognition approach." Frontiers in computational neuroscience 11 (2017): 103.
- [11] Subasi, Abdulhamit. "EEG signal classification using wavelet feature extraction and a mixture of expert model." Expert Systems with Applications 32, no. 4 (2007): 1084-1093.
- [12] Kher, Rahul K., and Rathang U. Shah. "Wireless EEG Signal Acquisition and Device Control." In researchgate.com(2016)
- [13] Hassan, Reshad, Sakib Hasan, Md. Jubaer Hasan, Md. Rafat Jamader, David Eisenberg, Tanmoy Sarker Pias. "Machine Learning Based Human Attention Recognition From Brain-EEG Signals." In 2nd IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability 2020(IEEE ECBIOS 2020), "Unpublished".
- [14] Pias, Tanmoy Sarker, David Eisenberg, and Muhammad Aminul Islam. "Vehicle Recognition Via Sensor Data From Smart Devices." In 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 96-99. IEEE, 2019.
- [15] Pias, Tanmoy Sarker, Raihan Kabir, David Eisenberg, Nadeem Ahmed, and Md Rashedul Islam. "Gender Recognition by Monitoring Walking Patterns via Smartwatch Sensors." In 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 220-223. IEEE, 2019.