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

Sakib Hasan[1], Fatima-Tuz-Johairah[2], Anika Fahmida[3], Tanmoy Pias[4] and David Eisenberg[*5]
University of Asia Pacific and Rajshahi University of Engineering and Technology, Bangladesh & Virginia Tech, USA
{sakib.hasan[1], 16101008[2] and tsr[4] } @uap-bd.edu, anikafahmidahossain29@gmail.com[3], davide77@vt.edu[*5]

Abstract*—Have you ever thought about how much your students are currently concentrated on during your lecture? Obviously you have and whatever the students were thinking of, many of them faked their attention to the class after you have wanted to know if they were understanding the lecture. Machine Learning based human attention level recognition has grown much interest these days for a single person, but the attention level of a particular group of people(such as a Classroom) has not gotten that much interest as of yet. This paper proposes a cost-effective hassle free Classroom Attention Level Recognition system- CALR using Human attention level recognition system to accurately identifying the cumulative attention level of a classroom. This system will not only serve the purpose of helping a teacher of how his or her lecture is going thus far but also will act as a quality measuring system to ensure quality education in any educational institution. Single-channel time-frequency scalp-EEG signals are recorded off of every student by using our 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 transform 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 which trained our state-of-the-art model to recognise human attention level for a particular human. Our system thus captures a number of humans' attention levels at a given time and an overall is calculated from those values to subsequently telling us what percentage of the class 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 the Brain Signals has been a popular one in this rapidly advancing world. But it requires tireless efforts from the researcher point of view. Over the past decade not only have we seen the use of Machine Learning(ML) and Deep Learning(DL) to analyse EEG(electroencephalograph) signals but also have seen their implementation in finding out different diseases and emotional states of human brain. In [2] *Jalilifard* used ML algorithms to classify human emotional states, *Lin* addressed the changes in emotional states using ML while listening to music in [5] and in [6] *Peng et. al.* used manifold extreme learning machines to detect the human emotional states using EEG.

We can also see *Ieracitano et. al* described in [1] a ML system to identify the brain states to identify separately the Alzheimer's disease, Mild Cognitive Impairment or Healthy Control mental states; Meanwhile, *Bhardwa et. al.* in [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 et. al.* in [4] to attempt using more advanced algorithms like LSTM-RNN in emotion recognition. A lot many studies in this domain to detect diseases or classifying emotional states of human brain can be seen in [7], [8], [10] and [11].

Although a lot of studies took place in the domain of emotion classification or disease identification using EEG signal, 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 that forms the brain. Thus, it can be stated that with changing Attention levels a significant change in the EEG signals should be found. And the highest number of attention related signals can obviously be found in one of the most vital demographics- the Students.

In this work, a community implementation is proposed for Classroom Attention Level Recognition System where our developed model for identifying individual human attention level using complex hybrid model with CNN-LSTM with Brain-EEG is used as the backbone.

II. ElectroEncephaloGraphy(EEG): Ins and Outs

Electroencephalography captures the wave signals generated by the electric signals flowing across an accumulation of 100 billion neuron cells in human brain.

A. Composition of a raw EEG

[12] states, in the human sensory system the electrical sign of a solitary neuron is too little to ever be estimated by an anode on the scalp.Thus, the measured activity originates from the summation of the electric voltage fluctuations of numerous neurons on the region of the cathode. In the event that the electric field is estimated, the result is called an EEG. Some of the times, a collection of neurons fire all the more

*Corresponding Author : davide77@vt@edu

synchronously reflecting immediately in the captured signal as it turns more powerful and increasingly synchronous. This is the reason these signals are viewed as brainwaves and can be described in some frequency sub-bands(Fig. 1).

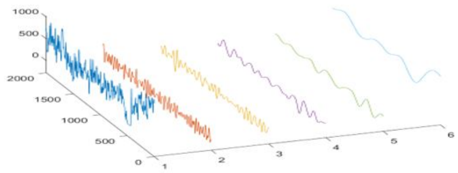

Fig. 1. 1-Raw, 2-gamma, 3-beta, 4-alpha, 5-theta, 6-delta. The length of the signal is 2000 and the z-axis represents the amplitude of the signals.

So, a raw EEG signal is comprised of frequency bands: Gamma($>$30Hz), Beta(13-30Hz), Alpha(8-12 Hz), Theta(4-8Hz), and Delta($<$4Hz). But due to the unwanted muscle movements or external influences, a raw EEG captured from any human faces interference from a lot of Noises. 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

So, what actually Human Attention Level means? In simple words, its to recognize if a person is attentive or not to a particular task. In a classroom the attention level is simply- Attentive and Inattentive. And 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 the students of any particular class i.e. Classroom Attention Level(CAL). By consensus, we found that a student can be in one of the four cognitive states states while attending a lecture in the classroom.

- **Completely Observant** states that a student is focusing solely on the lecture materials discussed.
- **Completely Non-observant** refers to those students who are not at all interested in the lecture and doing something else during the lecture.
- **Pseudo Observant** are those students who pretends to be attentive to the board or the teacher with a constant stare but s/he is not.
- **Moderately Non-observant** refers to the students who are trying to maintain his/her focus on the class, but fails to do so for either surroundings' influences or difficulty catching up with the speed of the lecture.

IV. Proposed Work

A lot of classification mechanisms were devised in [1], [2], [4], [9], [10] to classify either human emotions or human sleep stages or driver's cognitive states etc. But nobody has concentrated as of yet on this particular target audience which is 'Students in Classroom'. Looking in closely, we will find the similarity of Human Attention Recognition using EEG signals with Human Emotion Classification from EEG.

In this paper, a Classroom Attention Level Recognition(CALR) system which will enable a teacher to fetch EEG raw signals from individual students is proposed. The system, after performing the individual analysis and attention recognition, will provide an average Classroom attention level as output to the teacher. It is to mention that all these processes will be performed without the students knowing so that they cannot cheat. As we don't want the students to know of our process, wireless transmission technology to capture the EEG signals at any moment has to be used. In [12], *Kher et. al.* described in detail and with ease, the process of capturing a wireless EEG signal via Bluetooth modules.

V. Architecture of the Proposed System

The proposed one is a community based implementation of the individual human attention level recognition model developed by us with a view to building a software. We are excited to state that we have achieved an accuracy of 89% in terms of identifying the attention level of one human being using CNN-LSTM [13]. This model is the backbone of the system of this work, because we aim to capture the raw EEG signals from numerous(if possible all) students at random instances. Our system aims to capture those separate signals, process them, identify the attention level and produce a final output for the whole class.

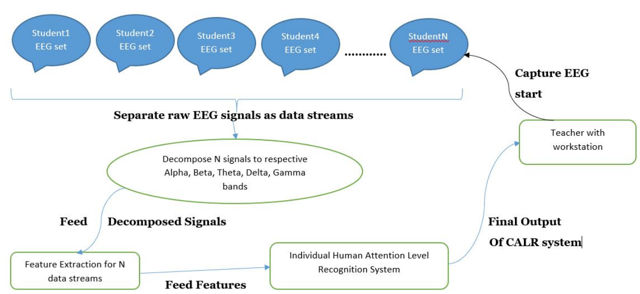

Fig. 2. Architecture of Proposed System

So, the system designed will be delivering the expected result if it runs through the following steps successfully-

- Wireless EEG capturing device setup for students.
- Capturing raw EEG signals of the students with the device put on for a short period of time say 10 seconds while the students should be completely unaware of the time period of capture through the app.

- Processing the raw EEG signals separately, identifying the attention level of the individual signals. The individual signals should be fed via wireless connection(BLUETOOTH) to the faculty workstation.
- Finally, calculating the percentage of attentive students to recognizing the average attention level of the whole classroom.

VI. System Setup & Full Experimentation

A. Device Selection & Setup for Signal Acquisition

Muse2 or Brain-sense or Neurosky or E-motiv are the existing highly expensive devices that can wirelessly transmit a raw EEG signal captured from one human [12].

But to minimize the costs(if necessary), we built up a new device using the Bitalino main board. A BT(Bluetooth) module was attached to the motherboard of the device for wireless transmission which resulted in a much cheaper product. These devices placed on a headband is strapped to one student per.

B. Application on the Teacher's Workstation

A running application on the teacher's workstation is necessary for our system. The application simply features a PUSH button and a GUI for the final output. All the main works take place at the back-end of the system. The application enables a faculty for capturing the individual raw EEG signals keeping the students completely unaware, pre-processes the signals for feature extraction, recognizing if an individual is attentive or not through our developed Machine Learning based Human Attention Recognition model, and finally providing output stating the average classroom attention level.

In the next subsections of VI we will be discussing in detail about our developed model to recognize an individual EEG signal's attention level which secures an accuracy of 89%.

C. Individual Signal Pre-processing

While capturing the EEG data, we noted down the start and end of the attentive time frame. After a few tests, we observed that that period typically went on for 1-2 seconds. We needed to group an individual's states into 4 classes-neutral, attentive, happy, and boring. The attention state was physically sectioned from the crude EEG. The sampling rate being 1000Hz, the segment length is 2 seconds which consists of 2000 data points. Every session (120 seconds) is segmented into 20 slots (consists of 2000 data points), which represent an active attention zone. For better understanding refer to [13]. We took samples from 20 different students two times each resulting in 40 raw signals, where each raw signal capture time was 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. In [13], the similar approach has been used.

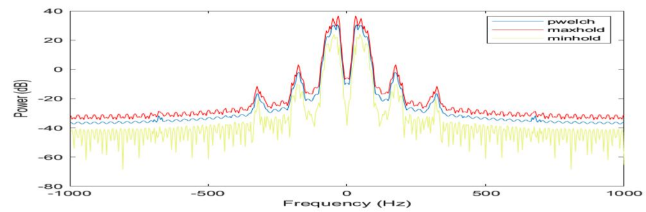

Fig. 3. Welch's Power Spectral Density Estimate

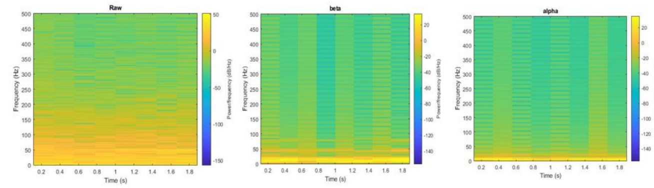

Fig. 4. Spectrogram of Raw, Beta and Alpha signal

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.

For finging out *WPSDE* & *Spectrogram*, some basic statistical features were first estimated from our data segments using *Sliding Window Technique*. In specimen the data segment was 2000 samples in length, and we estimated **Kurotsis, Skewness, Mean, Standarad Deviation** with a window of size 100 Samples. Each slide leaves behind 10 samples from the previous window.

E. Recognizing Individual Attention Level

Our backbone model is a Complex Deep Learning model developed with CNN & LSTM algorithms which is described with explanation in [13]. The LSTM is used for sequence classification, on the other hand the visual data-set consists of Spectrogram of raw EEG, beta and alpha signal. The developed novel deep learning model using CNN and LSTM as backbone acquired an astounding accuracy of 89% which can differentiate among *Neutral, Attentive, Happy or Sad* levels. In paper [14], b15 an 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, which is individual recognition of attention levels from a single raw EEG signal.

Suppose, we took 'N' students' EEG signal at any particular moment and they are fed to our system. Thus, our application will receive separate N number of data streams as input. Now, passing these N data streams through our model we will get an attention level against each of those streams as discussed before. For our CALR system, we will consider a stream **Attentive** if it is recognized as 'Attentive' by our model and 'Inattentive' if anything else. Let's say out of N students, 'X' students are paying attention to the lecture and 'Y' students are not. so, our final output from the system will be determined by two metrics P and Q,

$$AttentivePercentage, P = X/N * 100\% \quad (1)$$

$$InattentivePercentage, Q = Y/N * 100\% \quad (2)$$

So at the end of all these processes, the system will provide a 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 were run for five different scenarios to find out if the system delivers an approximate result. In a particular scenario, we took 10 separate EEG streams(N1 to N10) that we derived from our data-set used for individual attention recognition. We passed the separate EEG streams through our model after decomposing them to their separate sub-bands, and calculating final output with the help of equations (1) and (2), we 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 that you see from the expected output in the achieved output is due to the 89% accuracy of individual attention level recognizing model we have devised;however, Scenario (1) states the potential of our developed model which we can achieve by fine tuning our deep learning based model with increased number of instances in the data-set and increased number of features for the model's input as well.

VIII. Conclusion

Although we could not implement the total system for unavoidable circumstances, we can observe the potential of the system if implemented properly from VII. In our individual attention recognition model, we have already achieved 89% accuracy with samples from only 20 participants and features of 2 kinds. We believe with a larger data-set we can fine-tune our model to its maximum potential. With that target in mind, we will focus on implementing the proposed system in this paper in near future. For advanced works, we also want to add feature for separating percentage of all the 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 (ICIICT), 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 Sarkar, 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 Sarkar, 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.