Using Deep Learning for Assessment of Workers' Stress and Overload

S. Eskandar[a] and S. Razavi[b]

[a] Ph.D. Candidate, Department of Civil Engineering, McMaster University, Canada
[b]Associate Professor, Department of Civil Engineering, McMaster University, Canada
E-mail: eskandah@mcmaster.ca, razavi@mcmaster.ca

Abstract –

Spotting indications of unsafe human behaviour, a leading cause of an accident, is critical in providing a safe workspace. Among factors affecting human behaviour, stress and overload are the most significant ones, where limited knowledge exists on their underlying causes. Tracing physiological signs caused by stress and overload might be a feasible approach in detecting a specific neurological status leading to unsafe human behaviours, such as disobeying safety rules, standards, and instructions. In this paper, we present a deep learning technique to recognize distinctive neurological status by assessing physiological signals such as temperature and heart rate. An open database of non-EEG physiological signals was used to train and test the model. The database includes electrodermal activity, temperature, acceleration, heart rate, and arterial oxygen level signals of 20 healthy subjects through relaxation, physical, emotional and cognitive activities. A robust automated pattern recognition method, using deep learning, was used to predict and identify stress and overload. The experimental results indicate that the model can detect neurological status with higher accuracy than the traditional classification-based methods.

Keywords –

Deep Learning; Pattern Recognition; Stress; Overload; Safety

1 Introduction

Despite the ongoing safety studies and policy recommendations in construction, the extent of injuries is still significant. Based on the Canada Work Injury, Disease, and Fatality Statistics from Association of Workers' Compensation Boards of Canada (AWCBC/ACATC 2018)[1], each year, around 200 people die in construction sites. There are about 28000 time-loss injuries in construction-related accidents. The number of Canada's fatalities in 2018 shows that the construction industry has the highest number of fatalities among all sectors and accounts for almost 20 percent of the reported fatalities. So, improving safety and discovering the leading causes of accidents are still considered significant contributions to the construction industry.

In Eskandar et al. 2019 [2], three major categories of social, physiological, and cognitive human factors that influence the safe behaviour in construction were studied to guide future research around improving safety in construction. From a physiological perspective, stress and overload presented to have a high-level association with unsafe behaviour among construction workers, which motivated many researchers to detect stress in the work setting.

Stress can be described as the response of the body to the pressures on the human nervous system [3], which have been measured through subjective tests and questionnaires to collect individual responses [4], [5] . Moreover, stress can be measured through variations in human physiological features such as heartbeats, body temperature, and respiration [6]–[8]. With the advances in wearable biosensors and real-time data collection, many researchers focused on the physiological effects of stress on the human body to build a stress detection model. Several supervised learning algorithms were employed to detect stress and overload by detecting patterns of stored physiological signals during experiments.

Below, we first look at the stress and overload of construction workers and relevant physiological information that could guide their identification. Then a deep supervised learning model is proposed to classify different stress-inducing neurological states by processing physiological signals.

2 Stress and Overload

Stress (including physical and psychological) is among the contributing factors that lead to unsafe behaviour [9], [10]. Examples of stress factors in construction setting were presented as (1) physical stressors like noise, vibration, lighting, boredom, fatigue, cold or heat, and (2) social psychological stressors like fear, uncertainty, anxiety, mental overload, and time pressure [10]–[13]. Overload as an essential stressor that affects human behaviour was selected to focus on due to the nature of manual work in a construction environment that causes workers to exceed their capacity of handling the job.

Tracking symptoms of stress and high mental overload with multiple physiological features give us insight into a human's neurological state. Most physiological measurements come from a network of sensors in which become easier to collect in an unobtrusive real-time manner. Currently, the human body's vital signs can simply be recorded through wearable biosensors and health gadgets (e.g., smartwatch, earbuds, headset). Many researchers used sensors to measure specific physiological conditions to study factors that affect individual neurological statuses such as stress [14], sleep deprivation [15], fatigue [16], and social aspects [17].

Among different measuring methods that could reflect stress and mental overload, Electroencephalogram (EEG) sensors have been commonly applied in many studies [18]–[20]. EEG is a valuable source in identifying brain activities to measure electrical activities of the brain by electrodes positioned on the scalp, and they capture neurons in the brain by electric potentials. However, there are limitations in applications during physical activities as these signals are sensitive to face and body movements (e.g. eye blinks), which makes them impractical for application in construction safety [20]–[22]. Considering the EEG limitations, viable biosensors that could detect and reflect the stress in construction were presented in [14] as; Photoplethysmography (PPG), Electrodermal Activity (EDA), and peripheral skin temperature (ST), that are sensitive to extrinsic and intrinsic artifacts which require extensive filtering.
In addition to the feasibility of data collection in construction, data should reflect the sign of stress and overload. In an open-source database from a Birjandtalab et al. [23], that was conducted on subjects while confronted by stress and overload; seven different non-EEG physiological signals were collected during the experiment. In the current study, the above-noted database (Non-EEG dataset) was used as a source of information to study and train a stress and overload detection model. This database provides us with useful insights over the physiological features of individuals while confronted with overload and stress. This paper is distinct from previous studies by focusing on the physiological impact of mental stress and protecting the model against biases other than the ultimate goal, such as; not including 3-axis acceleration, which leads the model to movement recognition.

3 Non-EEG dataset

A non-EEG physiological signal from [23] includes acceleration (Ax, Ay, Az), electrodermal activity (EDA), temperature, heart rate, and arterial oxygen level (SpO2) signals of 20 healthy subjects during relaxation, physical, emotional and cognitive activities. This dataset includes individual responses while facing different stress-inducing activities. The experimental procedure includes: (1) Five minutes of relaxation, (2) Physical stress by walking on a treadmill at 1 mile/hr. For two minutes and jogging at 3 miles/hr for two minutes, (3) Relaxation, (4) Cognitive stress by counting backwards by sevens from 4285, and then performing Stroop test while alerted by a buzzer, (5) Relaxation, (6) Emotional stress by watching clips from a horror movie, (7) Relaxation.

Figure 4 displays a time-series for the Subject1 during the experiment collected by wrist-worn biosensors. The data file was in the WFDB (WaveForm DataBase) format, which can be read using its associated software package [24]. Records were labelled through ".atr" annotations file format (i.e. red stars on the a_x signal), which indicate a change in the activity of a human subject (i.e., moving from one step to another in the experiment steps).

The data files contain two records per person; one is recorded 3D acceleration, temperature, and electrodermal activity (EDA) with a frequency equal to eight (Figure 1). Another measurement technology recorded heart rate (HR) and arterial oxygen level (SpO2) with a frequency of one reading per second (Figure 2) [23].

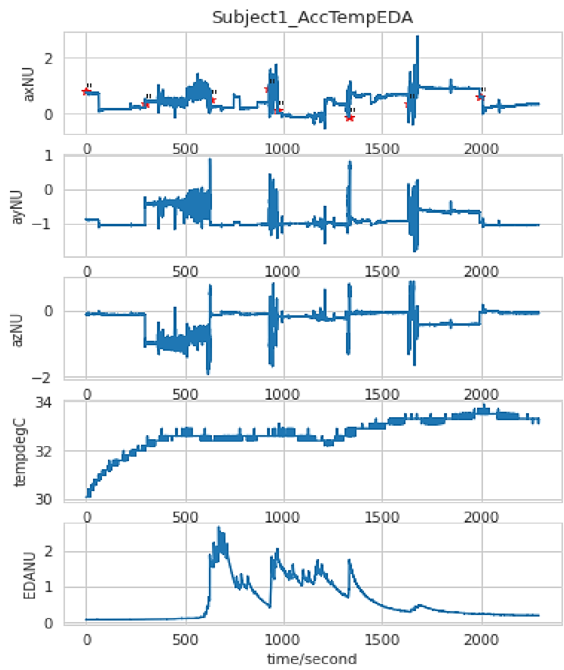

Figure 1. Acceleration (Ax, Ay, Ax), Temperature, and EDA recorded time-series for Subject1.

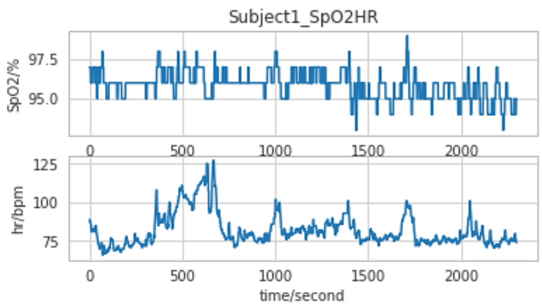

Figure 2. SpO2 and Heart rate recorded time-series for Subject1

3.1 Preprocessing

3.1.1 Aligning

Before analyzing and preprocessing the recorded data, we need to align data of different frequencies from different devices. There are two strategies for resampling and aligning: (1) Upsampling the lower frequency by repeating or interpolating data between reading samples, (2) Downsampling the higher frequency data by replacing extra readings with mean or median. The following graph (Figure 3) shows the difference between Downsampling (with median) and original reading for an accelerometer-x signal. Whether we need to maintain a precision of the higher frequency or not, we can align data of different frequencies. Downsampling signals might result in loss of data, and upsampling the lower frequency reading was chosen to align readings from two separate measuring devices (Figure 4).

Figure 3. Accelerometer-x down sampled (with median) side-by-side to the original high-frequency

3.1.2 Feature scaling

As it is visible in Figure 4, recorded signals are from different range and amounts. So it is essential to scale all data and perform feature scaling before any processing. Mainly, in classification problems, the majority of algorithms perform based on the distances, scaling the features before processing is necessary.

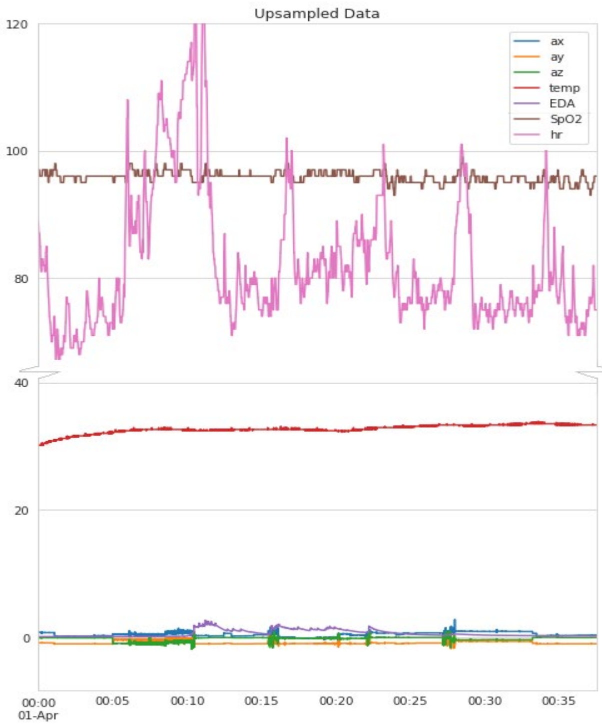

Figure 4. Aligned signals using the upsampling method (with interpolation) for SpO2 and Heart rate

3.1.3 Feature selection

In the preprocessing stage, it is essential to study correlations between different features. In Figure 5, the Pearson r correlations coefficient matrix has been shown to measure the degree of the relationship between linearly related variables to indicate whether two variables are strongly dependent or independent as a part of preprocessing. Based on the Pearson correlation coefficient and Cohen's standard, there is a significant association between acceleration signals (Ax, Ay, and Az). However, temperature and electrodermal activities (EDA) are independent.

Figure 5. Correlation matrix calculated using the Pearson method

Processing the correlation between features and eliminating those attributes that are unrelated enables a robust feature selection. Feature selection is a critical part of any machine learning pipeline, which leads to the accuracy in the models. Also, having perfectly correlated features increase a chance that the model performance obstructed by Multicollinearity (i.e. when one part can linearly be predicted from the others with a high degree of accuracy, in this case, Ay and Az are negatively correlated). High accuracy cannot be achieved without applying different methods of feature selection, such as Pearson or Spearman correlation matrix, Chi-squared, or Recursive feature elimination. Here, based on the Pearson correlation, features that are highly correlated decrease the performance of the model.

In the current research, in addition to the highly correlated features, it is required to remove three signals of accelerators. Acceleration in different directions is beneficial for activity and movement detection, not the stress and mental overload, which is the focus of this study. Especially in a construction setting due to the physically demanding nature of work, eliminating data regarding movement and activities protects the model to biases, and it focuses on the ultimate goal of stress and mental overload detection.

Since the recorded signals are continuous reading over time, slices of data during a window of time were selected as a separate entry to the model. The optimal window size was detected by calculating the accuracy of the model for different window sizes. Moreover, in classification models, it is essential to have balanced classes for training a model. So, a similar number of windows representing each class were selected as the input to the model.

3.1.4 Feature extraction

After the feature selection stage, a feature extraction tool is needed to provide the training stage with more information regarding data distribution. For this matter, a convolutional neural network (CNN) layer was added at the beginning of the model pipeline.

4 Model

For time-series pattern recognition and classifying different neurological statuses (i.e., relaxation, physical, emotional and cognitive stress), long short-term memory (LSTM) algorithm was selected for training purposes. Long short-term memory (LSTM) is a form of recurrent neural network (RNN) in the field of deep learning. LSTM has feedback connections in addition to the standard feedforward processing. These features enable LSTM to process entire sequences of data and make it accessible in time series data. At the beginning of the model pipeline, convolutional neural network CNN extracts features from signals, and it prepares input for the pattern recognition stage, which was conducted through sequential LSTM layers. Three different LSTM layers were added to the model to give more depth into the calculation, which provides a model with a better chance of prediction. Then, two fully connected layers at the final stage of training prepared the processed data for classification. Different classifiers can be applied to the final stage, namely: Bayes classifier, Hidden Markov Models, Random Forest Classifier, and Ensemble algorithm classifier (meta-algorithms that combine several methods into one model to decrease bias and variance and improve predictions). Adam classifier was chosen for the proposed model as the best fit.

There are multiple hyperparameters (e.g., number of units in each layer, number of dense layers, type of classifier, and number of epochs) in the proposed model, which requires a hyperparameter tuning for selecting the best match for the model. By defining a search area and training a model for several combinations, the best combination was chosen for training a model. For instance, by applying different window sizes, a window of 20 continuous reading of the signal was selected as an optimum number for the proposed model.

For training and testing processes, twenty percent of data were kept unseen for testing and from remaining samples, twenty percent was allocated to the cross-validation for backward propagation.

Figure 6 depicts the categorical accuracy and validation loss during the training session.

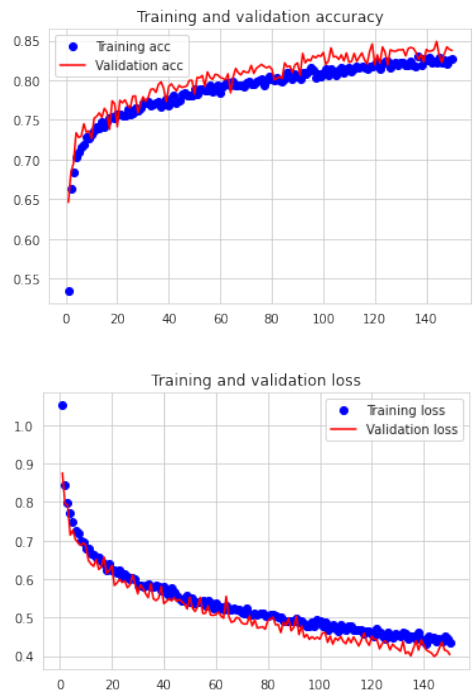

Figure 6. Training and validation accuracy and loss

Figure 7 presents the proposed model classification results that achieved an overall 85 percent accuracy for the test data set. Based on the f1-score (i.e. is a balanced amount of precision and recall) in the following report, we can conclude that the trained model is robust toward detecting physical stress and has more limitations in cognitive stress detection.

Classification Report

	precision	recall	f1-score	support
CognitiveStress	0.71	0.90	0.79	487
EmotionalStress	0.89	0.81	0.85	510
PhysicalStress	0.94	0.94	0.94	506
Relax	0.89	0.75	0.81	481
accuracy			0.85	1984
macro avg	0.86	0.85	0.85	1984
weighted avg	0.86	0.85	0.85	1984

Figure 7. Model classification report

In Figure 8, the confusion matrix for four different classes represents the actual classes in vertical and predicted classes in a horizontal direction.

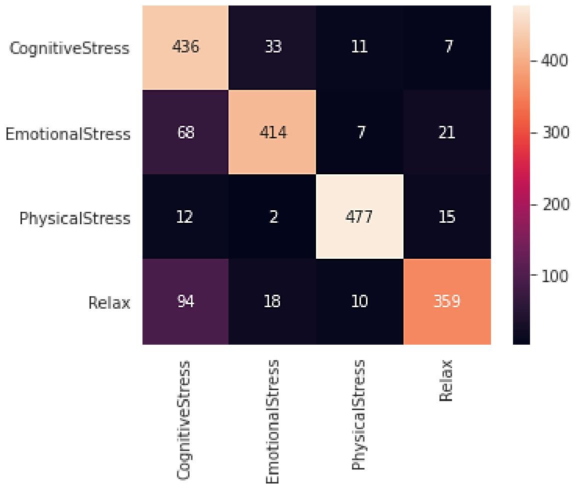

Figure 8. Confusion matrix

5 Results discussion and conclusions

In this research, efforts were taken to detect stress and mental overload, not only by employing deep learning techniques but by knowing the inputs of the model and removing preconceptions of the experimental study. The presented trained model in this study achieved 85 percent accuracy over unseen samples, which is a distinct improvement compared to using traditional methods of manual feature extraction (e.g., calculating the median, range, and standard deviation over the time series), combined with classifier algorithms.

Based on the confusion matrix, the trained model has difficulty recalling the Relax state, which needs more improvement in the data-gathering stage for future studies. Inputs to the proposed model were heart rates, temperatures, electrodermal activities, and arterial oxygen levels of subjects during an experiment, in which Ax, Ay, and Az accelerations were removed to enable the model to detect mental overload and stress instead of activities and movement detection. In addition to the importance of input to the model, the level of personal capacity while confronted with stressors has to be considered as individuals have different capacities under pressure.

Furthermore, for detecting stress and overload, multiple sub-classes should be considered to represent different levels of neurological status. For this matter, the severity of the neurological state uncovers, and only the higher level of stress in each category is considered hazardous.

References

[1] AWCBC/ACATC, "Association of Workers' Compensation Boards of Canada. 2016-2018 National Work Injury, Disease and Fatality Statistics," 2018. Accessed: Apr. 29, 2020. [Online]. Available: http://awcbc.org/?page_id=14#fatalities.

[2] S. Eskandar, J. Wang, and S. Razavi, "A review of social, physiological, and cognitive factors affecting construction safety," *Proc. 36th Int. Symp. Autom. Robot. Constr. ISARC 2019*, no. Isarc, pp. 317–323, 2019, doi: 10.22260/isarc2019/0043.

[3] S. Cohen, R. Kessler, and L. Gordon, *Measuring stress: A guide for health and social scientists*. 1997.

[4] H. Wittchen, P. B.-T. B. J. of Psychiatry, and undefined 1998, "Screening for anxiety disorders: Sensitivity and specificity of the Anxiety Screening Questionnaire (ASQ–15)," *cambridge.org*, Accessed: Jun. 26, 2020. [Online]. Available: https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/screening-for-anxiety-disorders/E786FEDE4C5C6EE0C2BD7DAE1754CC52.

[5] S. Reiss, R. A. Peterson, D. M. Gursky, and R. J. McNally, "Anxiety sensitivity, anxiety frequency and the prediction of fearfulness," *Behav. Res. Ther.*, vol. 24, no. 1, pp. 1–8, Jan. 1986, doi: 10.1016/0005-7967(86)90143-9.

[6] O. D. Kothgassner *et al.*, "Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment," *Comput. Human Behav.*, vol. 62, pp. 124–135,

Sep. 2016, doi: 10.1016/j.chb.2016.03.081.

[7] G. M. Harari, N. D. Lane, R. Wang, B. S. Crosier, A. T. Campbell, and S. D. Gosling, "Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges.," *Perspect. Psychol. Sci.*, vol. 11, no. 6, pp. 838–854, Nov. 2016, doi: 10.1177/1745691616650285.

[8] T. G. M. Vrijkotte, L. J. P. Van Doornen, and E. J. C. De Geus, "Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability," *Hypertension*, vol. 35, no. 4, pp. 880–886, 2000, doi: 10.1161/01.HYP.35.4.880.

[9] M. Y. Leung, Y. S. Chan, and K. W. Yuen, "Impacts of stressors and stress on the injury incidents of construction workers in Hong Kong," *J. Constr. Eng. Manag.*, vol. 136, no. 10, pp. 1093–1103, 2010, doi: 10.1061/(ASCE)CO.1943-7862.0000216.

[10] M. Y. Leung, Q. Liang, and P. Olomolaiye, "Impact of Job Stressors and Stress on the Safety Behavior and Accidents of Construction Workers," *J. Manag. Eng.*, vol. 32, no. 1, pp. 1–10, 2016, doi: 10.1061/(ASCE)ME.1943-5479.0000373.

[11] R. M. Choudhry, D. Fang, and S. Mohamed, "The nature of safety culture: A survey of the state-of-the-art," *Saf. Sci.*, vol. 45, no. 10, pp. 993–1012, 2007, doi: 10.1016/j.ssci.2006.09.003.

[12] R. R. Langdon and S. Sawang, "Construction Workers' Well-Being: What Leads to Depression, Anxiety, and Stress?," *J. Constr. Eng. Manag.*, vol. 144, no. 2, pp. 1–15, 2018, doi: 10.1061/(ASCE)CO.1943-7862.0001406.

[13] L. M. Goldenhar, L. J. Williams, and N. G. Swanson, "Modelling relationships between job stressors and injury and near-miss outcomes for construction labourers," *Work Stress*, vol. 17, no. 3, pp. 218–240, 2003, doi: 10.1080/02678370310001616144.

[14] H. Jebelli, B. Choi, and S. H. Lee, "Application of Wearable Biosensors to Construction Sites. I: Assessing Workers' Stress," *J. Constr. Eng. Manag.*, vol. 145, no. 12, 2019, doi: 10.1061/(ASCE)CO.1943-7862.0001729.

[15] R. Powell and A. Copping, "Sleep deprivation and its consequences in construction workers," *J. Constr. Eng. Manag.*, vol. 136, no. 10, pp. 1086–1092, 2010, doi: 10.1061/(ASCE)CO.1943-7862.0000211.

[16] T. S. Abdelhamid and J. G. Everett, "Ironworkers: Physiological demands during construction work," *Proc. Constr. Congr. VI Build. Together a Better Tomorrow an Increasingly Complex World*, vol. 278, no. October, pp. 631–639, 2000, doi: 10.1061/40475(278)68.

[17] R. M. Choudhry, D. Fang, and H. Lingard, "Measuring safety climate of a construction company," *J. Constr. Eng. Manag.*, vol. 135, no. 9, pp. 890–899, 2009, doi: 10.1061/(ASCE)CO.1943-7862.0000063.

[18] H. Jebelli, S. Hwang, and S. Lee, "EEG-based workers' stress recognition at construction sites," *Autom. Constr.*, vol. 93, pp. 315–324, Sep. 2018, doi: 10.1016/j.autcon.2018.05.027.

[19] J. Chen, J. E. Taylor, and S. Comu, "Assessing Task Mental Workload in Construction Projects: A Novel Electroencephalography Approach," *J. Constr. Eng. Manag.*, vol. 143, no. 8, p. 04017053, Aug. 2017, doi: 10.1061/(ASCE)CO.1943-7862.0001345.

[20] D. Wang, J. Chen, D. Zhao, F. Dai, C. Zheng, and X. Wu, "Monitoring workers' attention and vigilance in construction activities through a wireless and wearable electroencephalography system," *Autom. Constr.*, vol. 82, pp. 122–137, Oct. 2017, doi: 10.1016/j.autcon.2017.02.001.

[21] E. Lew, R. Chavarriaga, S. Silvoni, and J. del R. Millán, "Detection of self-paced reaching movement intention from EEG signals," *Front. Neuroeng.*, no. JULY, Jul. 2012, doi: 10.3389/fneng.2012.00013.

[22] M. Teplan, "Fundamentals of EEG measurement," *Meas. Sci. Rev.*, vol. 2, no. 2, pp. 1–11, 2002, Accessed: May 20, 2020. [Online]. Available: http://www.edumed.org.br/cursos/neurociencia/MethodsEEGMeasurement.pdf.

[23] J. Birjandtalab, D. Cogan, M. B. Pouyan, and M. Nourani, "A non-EEG biosignals dataset for assessment and visualization of neurological status," in *IEEE Workshop on Signal Processing Systems, SiPS: Design and Implementation*, Oct. 2016, pp. 110–114, doi: 10.1109/SiPS.2016.27.

[24] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.