A Proposal for Identification in using EEG Through Deep Neural Network

Abstract

Emergence of any technology brings new opportunities beside new challenges and internet is not exception in this manner. One of crucial challenges is faced identification of user, because identification methods have been available nowadays is easy to crack or some of them (for instance biometric identification methods) is not applicable with disabled individuals. One potential candidate to perform this task is using EEG. This kind of identification robust against cracking. Furthermore, it is applicable by all humans, even disabled people. In this study, a new method for identification using Deep Neural Networks (DNNs) is introduced. This method beside robustness against cracking, has reasonable accuracy comparing other biometric identification methods.

Keywords: Identification, Deep Neural Networks, biometric, EEG

Highlights:

- Identifying of individuals using EEG via DDNs
- Applicable by all people, even disabled ones and robustness against attackers
- Reasonable accuracy rate comparing other biometric identification methods.

Introduction

Two decades ago, the world has changed into a digital environment where everyone has their own unique digital identification. These identifications can take the form of conventional methods such as passwords and ID cards, but these forms can be easily bypassed. As a result, a different type of identifier was created based on a person's behavior or personal characteristics, known as biometrics. This includes identification through face recognition, voice recognition, fingerprint recognition, and iris recognition. The use of biometric systems for personal identification has recently gained attention from security research communities. Security systems are a ongoing challenge for society, and the use of personal identification systems is one of the primary tools for security. However, the widespread use of biometrics for personal identification has created a new challenge called spoofing, which poses a significant threat to security systems. Spoofing involves using methods to bypass the security of biometric systems and allow unauthorized access. There have already been several instances of spoofing attacks on biometric systems. Examples of such attacks include: the use of a printed photo to trick face recognition systems on three laptops, 2D face spoofing, and 3D mask attacks on face recognition systems. Fingerprint scanning has been bypassed using fake fingers made of gummy material, and a finger-vein commercial system has been spoofed using a piece of paper. An iris recognition system has been bypassed by placing an eyeball in front of the iris scanner. Voice recognition has also been spoofed by playing a recorded voice in front of a speaker recognition system. Due to these vulnerabilities, new biometric identification systems are needed that use invisible characteristics to eliminate external threats. One such system could be developed using brain signal electroencephalogram (EEG) for authentication. In recent years, since EEG signals captured and recorded accurately; researchers tend to utilize EEG signals in different areas. One of these areas is identification. In next section, some recent paper and their novelty are presented.

Related Work

In recent years, since EEG signals captured and recorded accurately; researchers tend to utilize EEG signals in different areas. One of these areas is identification. In [3], authors proposed identification method based on Support Vector Machine (SVM) in different sub-bands (Alpha, Low-Beta, High-Beta, and Gamma) and in order to reduce number channel, they applied Genetic Algorithm. They found that High-Beta and Gamma sub-bands have best performance with 98.17% and 98.5% accuracy, respectively. Also, these performances are reached by using 9 and 10 electrodes for High-Beta and Gamma sub-bands, respectively. Additionally, authors [1] presented a technique for identifying individuals using Visual Evoked Potential (VEP) and energy features from the gamma band of EEG signals. This method was tested on a significant number of subjects and achieved a high level of accuracy. The results showed that brain electrical activity has great potential as a biometric, as indicated by the successful analysis and simulations. In [1], authers employed the Binary Flower Pollination Algorithm (BFPA) to identify the most relevant EEG channels for verifying individuals. The researchers used a standard EEG dataset that focused on motor activity, movement, and imagination [2], and utilized autoregressive models with various orders for feature extraction. With the help of the Optimum-Path Forest (OPF) classifier, they achieved recognition rates of around 86% while reducing the number of EEG channels in half. Alyasseri et al. [6] introduced a new technique for identifying users based on EEG signals. The method employed a multi-objective Flower Pollination Algorithm and the Wavelet Transform (MOFPA-WT) to extract features from EEG signals, including various EEG energy information from sub-bands. The MOFPA-WT method extracts multiple time-domain features. The performance of the method was evaluated using accuracy, sensitivity, specificity, false acceptance rate, and F-score and compared to some advanced techniques using various criteria, with promising results.

Material and Model

Dataset

For this study, EEG Motor Movement/ Imagery dataset is used which contains over 1500 one and two-minute recording. The dataset obtained from 109 volunteers. Subjects performed different motor/imagery tasks while EEG of the subject were recorded by utilizing 64-channel BCI2000 system. The used data format is EDF+ which contain 64 EEG signal with sample rate of 160 Hz. In Figure .1, the configuration and placement of each EEG node is shown.

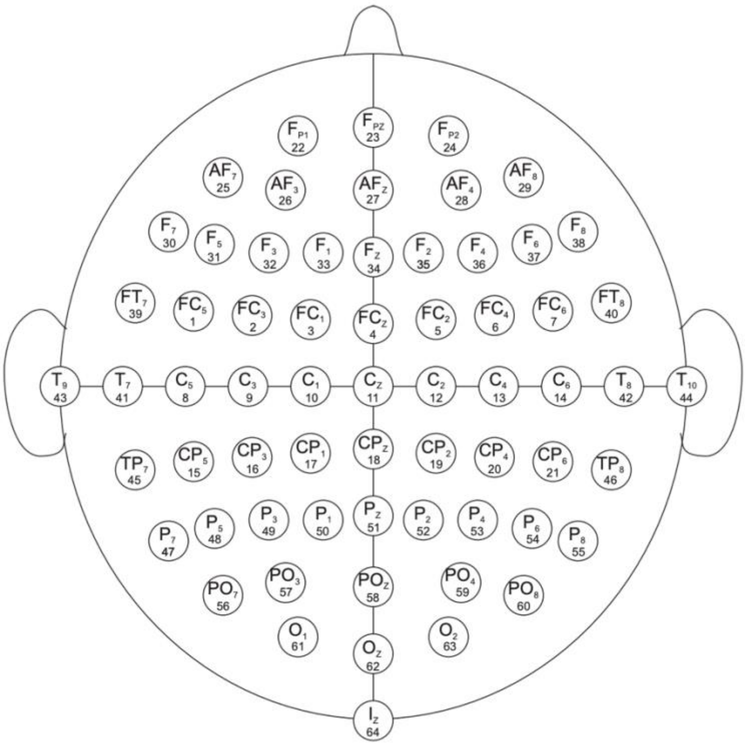

Figure 1: Configuration of each node of EEG electrode in Head

Each subject performed 14 experimental runs: two one-minute baseline runs (one with eyes open, one with eyes closed), three two-minute runs of each of the following tasks:

1. A target appears on either the left or the right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes.
2. A target appears on either the left or the right side of the screen. The subject imagines opening and closing the corresponding fist until the target disappears. Then the subject relaxes.
3. A target appears on either the top or the bottom of the screen. The subject opens and closes either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes.
4. A target appears on either the top or the bottom of the screen. The subject imagines opening and closing either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes.

In summary, the experimental runs were:

1. Baseline, eyes open
2. Baseline, eyes closed
3. Task 1 (open and close left or right fist)
4. Task 2 (imagine opening and closing left or right fist)
5. Task 3 (open and close both fists or both feet)
6. Task 4 (imagine opening and closing both fists or both feet)
7. Task 1

8. Task 2
9. Task 3
10. Task 4
11. Task 1
12. Task 2
13. Task 3
14. Task 4

During task runs, the target is shown for 4 seconds, and the subject have to perform motor/imaginary movement, subsequently, the subject relaxes for 4 seconds. This process repeated for 15 times (15*(4 s + 4 s) = 120s). In addition, the subject relax for 60 seconds during baseline runs.

Preparation

Before EEG signals can be used as input for regression or classification task, it must be prepared. The preparation procedure mostly involves reading and transforming EDF+ file, removing artifacts, sampling, feature extraction and feature reduction(optional).

For preparation procedure, MATLAB program is used in this study. After reading and transforming EDF+ file to MAT-file format, baseline of each signal is removed by subtracting fitted degree five polynomial for each one second window of correspond signal. In Fig. 2,

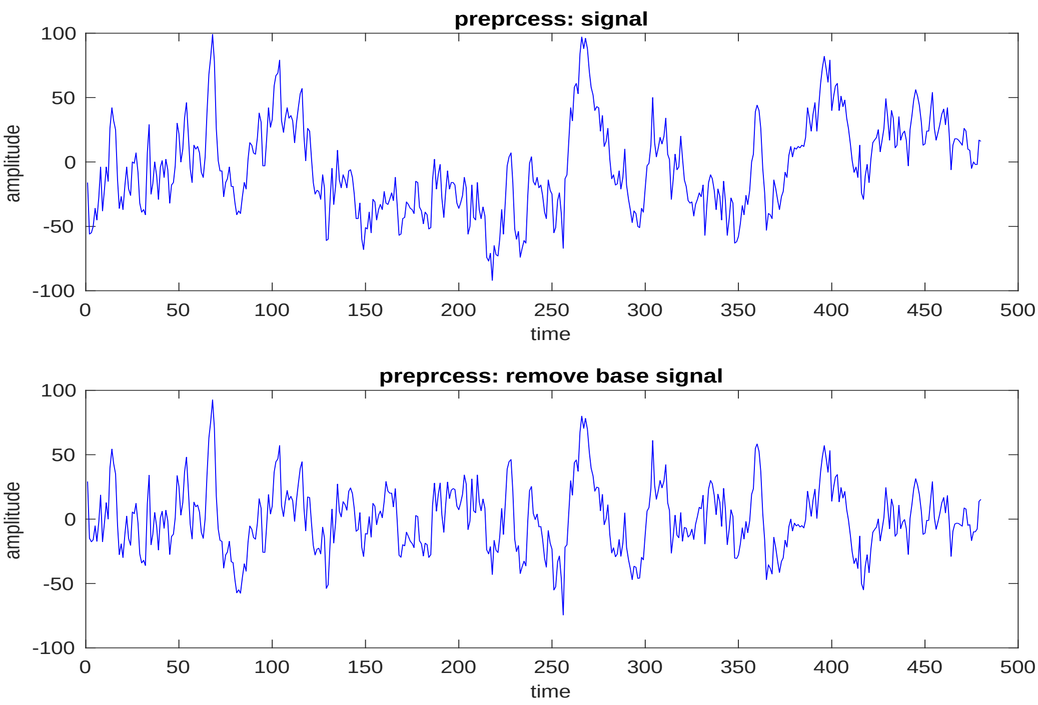

Figure 2

one channel original signal, fitted degree five polynomial using original signal and subtracted signal are showed.

In the next place, low-pass, high-pass, and notch filter is applied to subtracted signal. Since EEG signals is mostly active between 0.5 Hz and 60 Hz, therefore 0.5 Hz and 60 Hz are selected as cut-off points for high-pass and low-pass filters, respectively. In addition, notch filter is used to minimize power lines. In Figs. 3 and 4, an EEG signal before and after applying filters is illustrated.

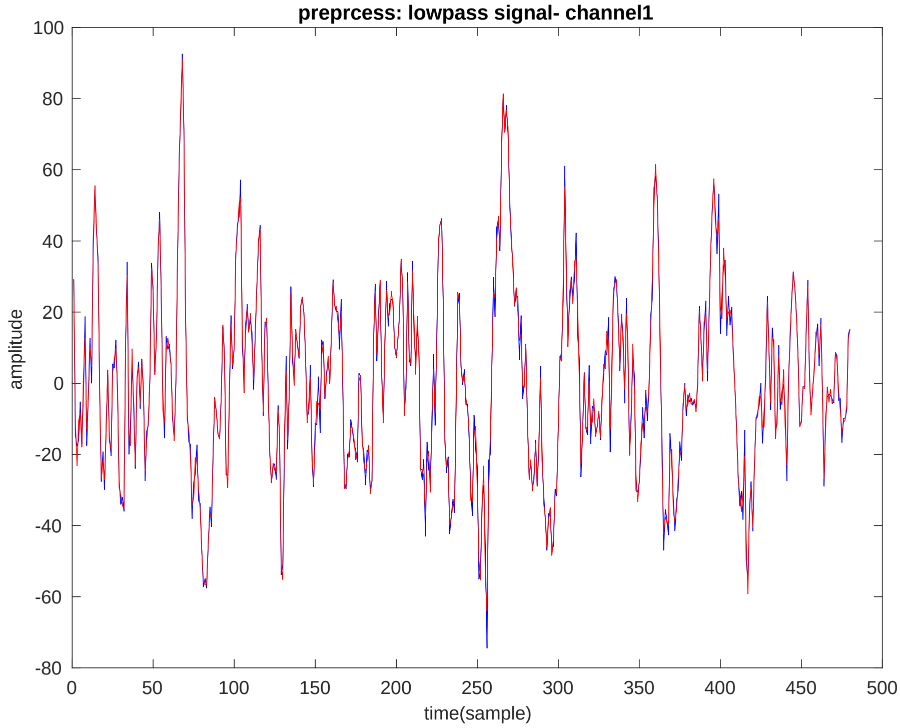

Figure 3

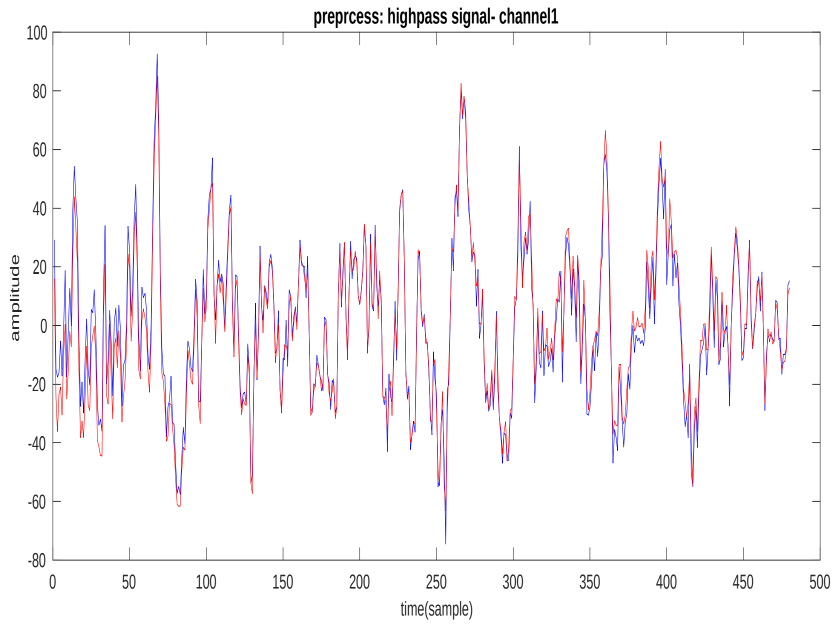

Figure 4

Subsequently, Common Average Reference (CAR) filter is applied. CAR filter can minimize common source artifacts. In Fig. 5, an EEG signal before and after applying CAR filter is demonstrated.

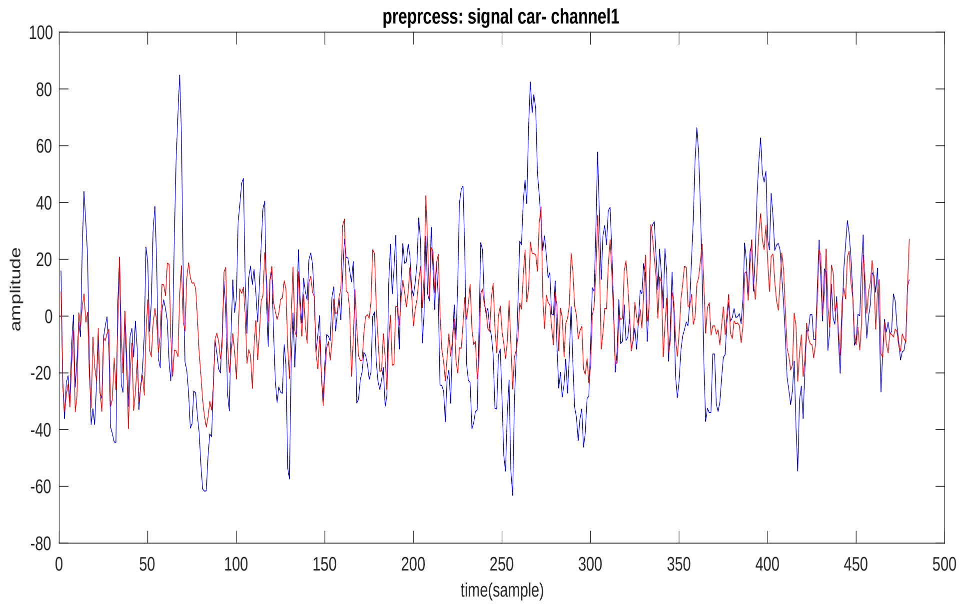

Figure 5

In last step, Savitzky-Golay filter is applied to smoothing signals with order four and frame length of seven which is presented in Fig. 6.

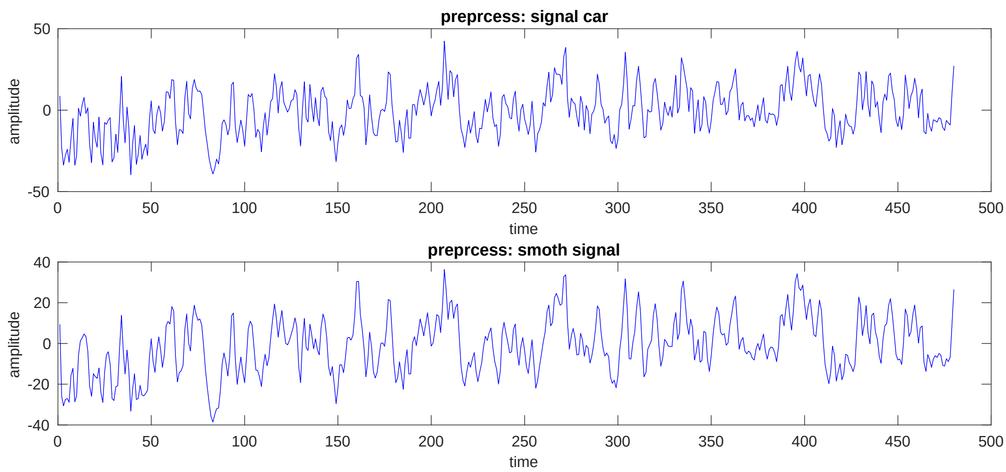

Figure 6

Since 5 individuals in dataset do not complete all tasks, therefore, for sampling, only the relaxed times is used with time window of 4 seconds. Each 4-second window is count as a sample.

$$\#samples = 104 \times (12 \times 15 + 2 \times 15) = 21840 \quad (1)$$

In order to extract feature from each sample, variance and mean of each channel are calculated with two different window size without overlap, one second and four second. Since each sample has 64 channel and from each channel, 10 feature is extracted, therefore, each sample has 640 features. Subsequently, each of features is normalized using standard deviation method.
Finally, number of features is reduced from 640 to 380 using PCA with explained variance ratio of 99%.

Model and Results

For classification task, deep neural network with 9 dense layers is utilized. Also, each dense layer follows three regulation method, drop out, L2, and batch normalization. Drop out and L2 rate set to 40% and 0.03, respectively. In addition, first 8 layer neurons unit and activation function are 700 and relu, respectively. However, last layer neurons unit are 104 because of number of subjects or class data set have. Furthermore, last layer activation function is softmax since this layer is performing classification task. The architecture of neural network is demonstrated in Fig. 7.

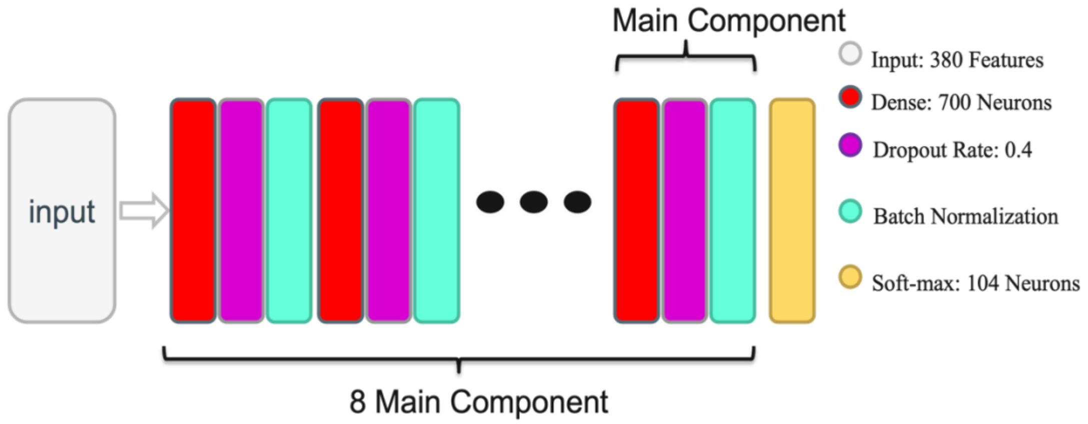

Figure 7

To achieve robust classification, the model trained 16 times, each time with different starting weights. Furthermore, majority voting method is utilized between 16 models' prediction to achieve final prediction. Also, Stochastic Gradient Descent with Warm Restarts (SGDR)[5] method is employed for setting learning rate. Training parameters of model is as follow.

- Train/Val/Test: 90%,5%,5%
- Epochs: 1270
- Batch size: 512
- Regularizer: L2 (penalty rate= 0.03)
- Dropout: 40%
- Optimizer: Adam (lr=0.0005, beta1=0.9, beta2= 0.99)
- Loss = categorical cross entropy

As it shown in Table 1, the results of 4 different subjects and Weighed results of 104 subjects. From last row, It is vivid that the model perform accurately and robust since precision, recall, and F1-score are over 98%.

Subject	Precision	Recall	F1-socre	Support
1	0.92	1	0.96	11
12	0.89	1	0.94	8
14	1	0.88	0.93	8
72	1	0.91	0.95	11
Weighed results of 104 subjects	0.99	0.98	0.98	1093

Conclusion

In this paper, a robust identification model is presented. Robustness of model is achieved in two-fold. First, using 4s waiting samples between tasks, since these data prone to more artifacts and real saturation. Therefore, identifying these samples are challenging and a model capable of identify these sample are more robust. Second, by training 16 diverse model and using majority voting, performance and robustness of model increased significantly. Also, precision, recall, and F1-score of the proposed model are over 98%. It is worth to mention the accuracy of fingerprint identification is 98.7. In addition, since the feature used in this paper are mean and standard deviation of signal, therefore, the computation cost of method is lower than other methods. However, using 64 channel EEG is main drawback of this method. Therefore, as future work author will work on channel reduction.

Bibliography

[1]R. Palaniappan, D.P. Mandic Biometrics from brain electrical activity: a machine learning approach IEEE Trans. Pattern Anal. Mach. Intell., 29 (4) (2007)

[2]Alyasseri, Z. A. A., Khader, A. T., Al-Betar, M. A., & Alomari, O. A. (2020). Person identification using EEG channel selection with hybrid flower pollination algorithm. *Pattern Recognition*, *105*, 107393.

[3]Albasri, A., Abdali-Mohammadi, F., & Fathi, A. (2019). EEG electrode selection for person identification thru a genetic-algorithm method. *Journal of medical systems*, *43*(9), 1-12.

[4]Fraschini, M., Hillebrand, A., Demuru, M., Didaci, L., & Marcialis, G. L. (2014). An EEG-based biometric system using eigenvector centrality in resting state brain networks. *IEEE Signal Processing Letters*, *22*(6), 666-670.

[5]Loshchilov, I., & Hutter, F. (2016). Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*. (ICLR 2017 conference paper)

[6]Alyasseri, Z. A. A., Alomari, O. A., Makhadmeh, S. N., Mirjalili, S., Al-Betar, M. A., Abdullah, S., ... & Abasi, A. K. (2022). EEG Channel Selection for Person Identification Using Binary Grey Wolf Optimizer. *IEEE Access*, *10*, 10500-10513.

[7]Das, B. B., Ram, S. K., Pati, B., Panigrahi, C. R., Babu, K. S., & Mohapatra, R. K. (2021). SVM and Ensemble-SVM in EEG-based Person Identification. In *Progress in Advanced Computing and Intelligent Engineering* (pp. 137-146). Springer, Singapore.

[8]Adhikary, S., Jain, K., Saha, B., & Koner, C. (2021, October). Fuzzy Logic on Long Short-Term Memory for Smart Person-Identification System through Electroencephalogram. In *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)* (pp. 609-613). IEEE.

[9]Zhao, H., Chen, Y., Pei, W., Chen, H., & Wang, Y. (2021). Towards online applications of EEG biometrics using visual evoked potentials. *Expert Systems with Applications*, *177*, 114961.

[10]Alyasseri, Z. A. A., Al-Betar, M. A., Awadallah, M. A., Makhadmeh, S. N., Alomari, O. A., Abasi, A. K., & Doush, I. A. (2021, September). EEG Feature Fusion for Person Identification Using Efficient Machine Learning Approach. In *2021 Palestinian International Conference on Information and Communication Technology (PICICT)* (pp. 97-102). IEEE.

[11]Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.