Application of LSTM Algorithm to Assist Diagnosis of Epilepsy Based on Electroencephalogram (EEG) Signals

1st Sutrisno Ibrahim
Electrical Engineering Department
Sebelas Maret University
Surakarta, Indonesia
sutrisno@staff.uns.ac.id

2nd Kaleb Nathan Zebua
Electrical Engineering Department
Sebelas Maret University
Surakarta, Indonesia
kaleb.nathan@student.uns.ac.id

3rd Faisal Rahutomo
Electrical Engineering Department
Sebelas Maret University
Surakarta, Indonesia
faisal_r@staff.uns.ac.id

4th Muhammad Alif Rizky Naufal
Electrical Engineering Department
Sebelas Maret University
Surakarta, Indonesia
rizkynaufal@student.uns.ac.id

*Corresponding author: sutrisno@staff.uns.ac.id
Received: March 10, 2025; **Accepted:** April 21, 2025

***Abstract*— Epilepsy is a common disease that affects the brain's ability and has the potential to destroy the quality of life of sufferers. Diagnosis of epilepsy can be done by clinical testing and by using the electroencephalography (EEG) method. This research aims to apply artificial intelligence to improve the effectiveness and accuracy of EEG signal analysis. Epilepsy diagnosis is done automatically based on trained EEG signal files. This application can be done by applying the Long-Short Term Memory (LSTM) machine learning algorithm for recognizing patterns from brain signals that lead to epilepsy. The development was carried out using the EEG signal dataset from the University of Bonn which consists of 5 data sets. The detection process consists of the stages of data loading, augmentation, filtering, training, and classification. The developed system will be loaded into a GUI to facilitate users. The result of this research is a machine learning model with Long Short-Term Memory (LSTM) algorithm that has an accuracy rate of 91%, validation accuracy of 94% and loss of 0.2. Compared to other machine learning approaches such as SVM, KNN, and ANN, the proposed method achieves higher accuracy without the need for explicit feature extraction, highlighting its effectiveness in time-series signal classification. The model evaluation results show that this research is successful in assisting the detection of epilepsy using EEG signals with a high level of accuracy and efficiency.**

***Keywords*— Brain, Electroencephalography, Machine, Learning, LSTM, Algorithm, Epileptic.**

I. INTRODUCTION

Epilepsy is one of the most common neurological/brain disorders in the world [1]. The World Health Organization (WHO) along with several related organizations have made Epilepsy an alarming Public Health issue. According to the World Health Organization (WHO), approximately 50 million people worldwide have epilepsy, with more than 80% of epilepsy cases occurring in developing countries [2]. The disease has a direct influence on a person's quality of life and even has the potential to seriously impact their daily life.

Epilepsy itself is a neurological disease that can be suffered by all ages which is a manifestation of neurological/brain function disorders accompanied by characteristic symptoms in the form of recurrent seizures due to excess electrical charges that are released in neurons [3]. However, epilepsy can be diagnosed by relying on medical checkups and electroencephalography (EEG) examinations. EEG is a non-invasive and affordable method that aims to assess neurological function. EEG works by measuring the electrical activity of neurons in the brain with electrodes placed on the scalp [4].

However, EEG signals are often variable and need to be enhanced. Machine Learning technology is one area that enables the use of computers to process and analyze EEG signal data quickly and accurately. Thus, the implementation of Machine Learning in EEG signal processing can be a valuable tool in supporting the diagnosis and detection of epilepsy disease.

Many previous studies have been conducted to improve the efficiency and accuracy of applying Machine Learning for EEG. However, there are still many challenges that need to be overcome, such as reducing false readings, to developing algorithms/models that can adapt to multiple brain signals from different patients.

Various algorithms contained in Machine Learning can be used in this application for EEG, such as K-Nearest Neighbours (KNN), Long-Short Term Memory (LSTM), Artificial Neural Networks (ANN), to Transfer Learning [5]. In this study, the LSTM method is specifically evaluated using the University of Bonn dataset, which consists of time series EEG data. The LSTM method itself is one of the Neural Network algorithms and is categorized as a type of Recurrent Neural Network (RNN) [6].

This method is chosen due to its ability to process complex time series data by utilizing information from previous time steps to generate new outputs [7]. Through the experiments conducted in this study, evaluation variables for the LSTM method will be obtained and can later serve as a basis for comparison in the development of an epilepsy diagnosis system. Accordingly, the findings are expected to improve the precision and efficiency of epilepsy detection, enabling faster and more reliable diagnosis.

II. METHODS

A. Dataset

The dataset used in this study is obtained from the University of Bonn [8], as shown in Figure 1 below. This dataset consists of 5 data sets, labeled from set A to set E, with each set containing 100 EEG recording results with a duration of approximately 23.6 seconds. Sets A and B are EEG data taken from healthy individuals, where EEG analysis was conducted with eyes open in Set A and eyes closed in Set B. Meanwhile, Sets C, D, and E are EEG signals taken from individuals with epilepsy.

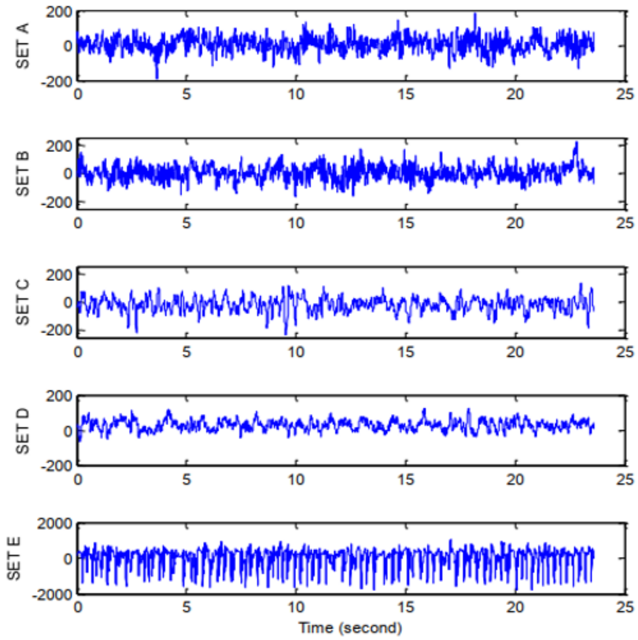

Fig. 1. EEG Time Series Signals from the Bonn Dataset (Sets A to E)

B. System Overview

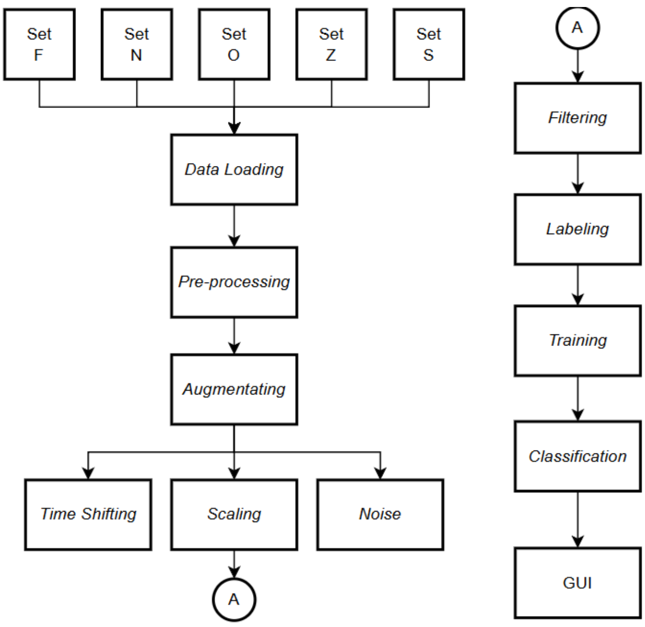

Fig. 2. Block Diagram of the Proposed Classification System using LSTM

The loading and data selection process in this study utilizes EEG datasets from two different patient groups, namely healthy patients and patients with epilepsy which are divided into 5 datasets A, B, C, D, and E. The datasets used are available in .txt format with time series type where each file consists of 4096 lines of EEG signal data.

The next stage of this epilepsy detection system is data preprocessing which involves various steps to achieve data that can be used, processed, and accepted by the system. Data preprocessing itself is the initial stage in data processing which aims to clean and prepare raw data so that it is ready for further processing.

The next stage is augmentation which serves to add variety to the data for training and also prevent overfitting when the training process is run so as to produce a good machine learning model. There are three augmentation segments that are applied to EEG signals, especially patient data suffering from epilepsy, namely Time Shifting, Adding Noise, and Scaling.

Then followed by a filtering process on the EEG signal to remove unwanted frequency components in the EEG signal and improve signal quality before proceeding to the labeling and training process.

The next step is to label the data that has gone through the augmentation and filtering process. The data will be created into a variable named 'eeg_signals'. This process is done by declaring a 'labels' array with a length (shape) equal to the amount of data owned. Then perform labeling for data that has temporary labelling 'healthy_signals' by labeling 0 as healthy patients, and the remaining data outside of temporary labelling 'healthy_signals' with label 1 as patients with epilepsy.

The next step is to train the machine learning model for epilepsy detection based on the pre-processed EEG signal. This process includes selecting the type of machine learning model used, setting parameters and layering on the model, and the training process based on the data that has been selected.

C. Classification Using LSTM

In this research, the machine learning model chosen is the Long Short-Term Memory (LSTM) neural network. LSTM was chosen because this model can handle time series data such as EEG signals well, and has the ability to understand and model temporal patterns and complex brain properties. This is in accordance with the need to identify temporal patterns and brain properties that may be signs of epileptic seizures in EEG signals.

```
model = Sequential()
model.add(LSTM(128, return_sequences=True, input_shape=(train_data.shape[1], 1)))
model.add(Dropout(0.3))
model.add(LSTM(64))
model.add(Dropout(0.3))
model.add(Dense(1, activation='sigmoid'))
```

Fig. 3. Two Layer LSTM Model Architecture with Dropout Regularization

The LSTM model used has two LSTM layers with 128 and 64 unit values, respectively. The selection of 128 units as the first choice is done to produce a model that is complex enough to learn new and complicated patterns from EEG time series data.

The LSTM architecture contains a substantial number of trainable parameters, which can increase the risk of overfitting [9]. To address this challenge, regularization methods such as dropout are commonly applied. In this study, a dropout layer with a rate of 0.3 is introduced after each LSTM layer to mitigate overfitting and enhance the model’s ability to generalize to new, unseen data.

The dropout rate of 0.3 was selected based on its frequent use in prior research, where it has demonstrated effectiveness in maintaining a balance between performance and overfitting prevention [10].

The output layer uses a sigmoid activation function because the desired output is a binary value, namely healthy (0) or epilepsy (1).

```
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

history = model.fit(train_data, train_labels,
                    epochs=10,
                    batch_size=32,
                    validation_data=(test_data, test_labels),
                    verbose=1,
                    callbacks = [myCallback()])
```

Fig. 4. Model Compilation Setup

The training process of the LSTM model for this epilepsy detection system is carried out by considering various parameters used to ensure the training process produces a reliable model. In the training process of this model, the number of epochs or training iterations of 10 times is used to provide sufficient time to learn from the data given without causing overfitting due to the training process that is too long / much. Then to speed up the training process, a batch size of 32 is used where the model weights will be updated every 32 samples trained.

The parameters used at the model compile stage include the Binary Cossentropy loss function because it is suitable for binary classification problems or classification between (0) and (1), where the model will predict the probability for each class of data. This function helps to assess the distance between the model's prediction and the actual label, where a lower loss value indicates better prediction results. Then using Adam's optimizer because it has the ability to adaptively adjust the learning rate or learning speed of the model so that it helps in the training process.

D. *Evaluation*

The categorization process's output simply serves as a "hint" or recommendation for the neurologist or other medical professional to use when making the final diagnosis. Whether or not this recommendation is accurate, the expert may provide the system with feedback. In order for the system to function better, it should be able to modify its attributes in response to this input. We propose to update the baseline in order to carry out the adaptation in this study. Some potential outcomes of the categorization include the following.

True Positive (TP): Epileptic data classified as epileptic

True Negative (TN): Healthy data classified as healthy

False Positive (FP): Epileptic data classified as healthy

False Negative (FN): Healthy data classified as epileptic

To conduct a proper and thorough evaluation of system testing, a test evaluation parameter is used by calculating the *sensitivity*, *specificity*, and *accuracy* values. The sensitivity and specificity parameters, and accuracy are statistical measurement parameters that are commonly used to assess the performance of tests carried out against existing or owned standards [36].

$$Sensitivity = \frac{TP}{TP+FN} \quad (1)$$

$$Specificity = \frac{TP}{TP+FN} \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

III. Result and Discussion

The evaluation was performed using the 'classification_report' function from the scikit-learn library to display the classification results with various evaluation parameters such as precision, recall, f1-score, and support for each class, as well as the average value for all data classes. The results of this evaluation allow for further improvement if the resulting figures are insufficient.

```
              precision    recall  f1-score   support

           0       0.89      0.92      0.91        37
           1       0.96      0.95      0.96        83

    accuracy                           0.94       120
   macro avg       0.93      0.94      0.93       120
weighted avg       0.94      0.94      0.94       120
```

Fig. 5. Classification Result

Based on EEG signal data, the model evaluation findings demonstrate high performance in distinguishing between epilepsy and non-epilepsy. The majority of the training and testing datasets were successfully identified, as seen by the classification report's 94% accuracy rate in Fig. 5 above. Both classifications' precision and recall values—0 for non-epilepsy and 1 for epilepsy—are comparatively high.

For the non-epilepsy class (0), the model achieves a precision of 89% and a recall of 92%, indicating high accuracy in identifying non-epilepsy data and minimizing errors in classifying it as epilepsy. For the epilepsy class (1), the model achieves a precision of 96% and a recall of 95%, showing that it accurately identifies epilepsy data with minimal errors in classifying it as non-epilepsy.

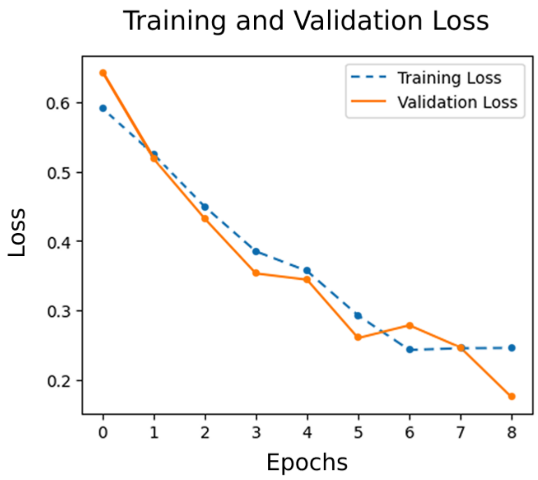

Fig. 6. Training and Validation Loss Curves

Fig. 6 above is a graph of loss training and loss validation for the model that finished training. Based on the figure above, it can be seen that the training loss figure is greater than the validation loss at the last epoch and ends with a fairly low training loss and validation loss value. This shows the ability of the model to produce good performance in predicting new data.

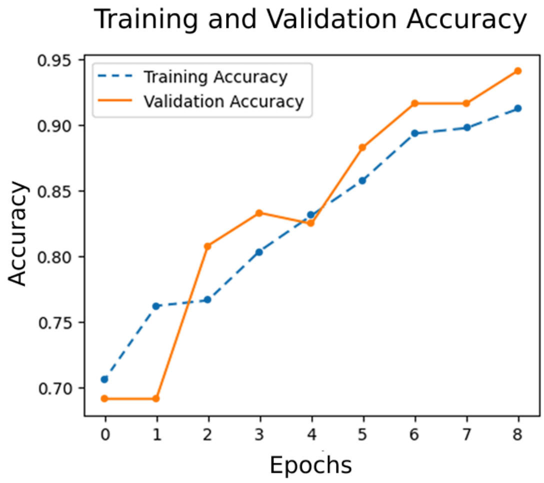

Fig. 7. Training and Validation Accuracy Curves

Fig. 7 above is a is the accuracy graph for training and validation of the model. The periodic increase in accuracy in training shows that the model is able to find the best solution when deployed. While the steady increase in accuracy in validation shows the ability of the model to prevent overfitting. So it can be seen that the model has the ability to provide good performance in predicting new data.

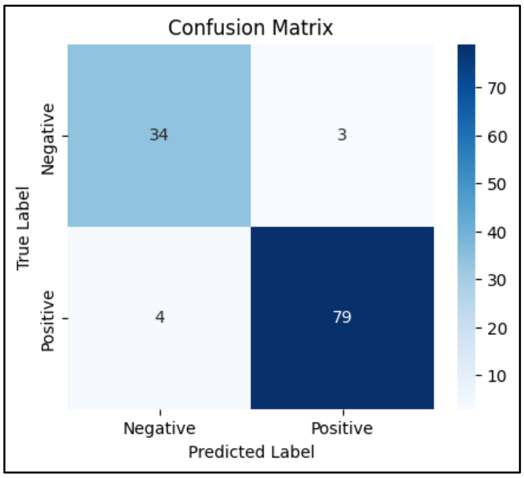

Fig. 8. Classification Results with Confusion Matrix

Based on the resulting confusion matrix, the epilepsy detection model developed has good performance in identifying two types of data divided into two classes, namely epilepsy (positive) and non-epilepsy (negative). With a total of 79 true positives and 34 true negatives, it has shown that the model successfully classifies almost all samples accurately.

TABLE I. MODEL TEST RESULT

	True Positive	*True Negative*	*False Positive*	*False Negative*
Set Z (A)	0	24	0	1
Set O (B)	0	24	0	1
Set N (C)	21	0	4	0
Set F (D)	23	0	2	0
Set S (E)	24	0	1	0
Total	68	48	7	2

Based on the tests that have been carried out, it is known that the model has a fairly high performance in classifying epileptic EEG data with non-epileptic ones. The resulting sensitivity is 97.1% which proves that almost all new epilepsy data is detected by the system as epilepsy correctly. While the specificity value is 87.2% which shows that almost all non-epileptic data is successfully detected by the system as non-epileptic correctly.

Then for testing accuracy, it is obtained at 92.8% which is quite high. This accuracy value has a similar figure to the accuracy in model evaluation of 94% which proves that the model has satisfactory performance both in the testing process, and validation to the actual testing stage using new data that has not been recognized by the model.

Based on the review by Farooq et al. [11], several previous studies have utilized the EEG dataset from the University of Bonn. A selection of those studies is summarized in Table II to provide a comparative analysis of different classification models.

TABLE II. MODEL COMPARISON

Ref	*Approaches*	*Feature Extraction*	*Accuracy*
Chen et al. [12]	SVM	DWT	86.83%
Zeiler & Fergus [13]	KNN	Time-frequency	85%
Rabcan et al. [14]	ANN	Time-frequency	85%
This Study	LSTM	-	92.8%

As shown in the table, the LSTM approach proposed in this study achieves a higher classification accuracy (92.8%) compared to other commonly used methods such as SVM (86.83%), KNN (85%), and ANN (85%), even without the use of explicit feature extraction techniques. This suggests that LSTM is more capable of capturing temporal patterns in EEG signals and is well-suited for time-series data analysis.

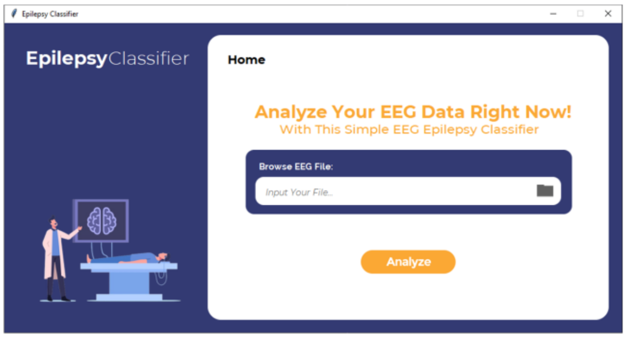

Fig. 9. GUI for Inputting EEG Data for Analysis

On the first page, the home page, a display will be designed using several elements such as images and text to welcome users. In addition, an element will also be designed to enter the dataset file to be analyzed along with a button to start the analysis process.

Fig. 10. GUI for Displaying Analysis Results

This second page serves to display the results page of the GUI that has been designed. On the page, there are several elements displayed such as the signal graph that has been processed, the writing of the detection results by the machine learning model, and also the finish button element to analyze other data.

IV. Conclusion

Based on the presentation of the material in the previous chapter, there are conclusions that can be drawn that this research produces a machine learning model that is able to accurately detect epilepsy and non-epilepsy data. This research produced a model with an accuracy rate of 94% and a loss of 0.2 and an accuracy of 92.8% in tests conducted on new data.

When compared to various studies using similar approaches, this research has a relatively high and good accuracy rate. Moreover, it is equipped with a GUI that can facilitate system access and increase efficiency and effectiveness.

For future research, it is strongly recommended to implement the system directly in real-world environments, along with the integration of more diverse datasets collected in real time, particularly multichannel EEG data. This would enable the model to learn from more complex and representative signals, thereby potentially improving its overall accuracy and performance.

Acknowledgment

Author would like to thank Mr. Ir. Sutrisno, S.T., M.Sc., Ph.D as project supervisor and also to Dr. Eng. Faisal Rahutomo, S.T., M.Kom. as advisor in the development of this artificial intelligence based epileptic diagnose research.

References

[1] C. E. Stafstrom and L. Carmant, "Seizures and epilepsy: an overview for neuroscientists.," Cold Spring Harb. Perspect. Med., vol. 5, no. 6, Jun. 2015, doi: 10.1101/cshperspect.a022426.

[2] H. Anwar, Q. U. Khan, N. Nadeem, I. Pervaiz, M. Ali, and F. F. Cheema, "Epileptic seizures.," Discov. (Craiova, Rom., vol. 8, no. 2, p. e110, Jun. 2020, doi: 10.15190/d.2020.7.

[3] A. Arianda, "Peran Usia Awitan Kejang dalam Epilepsi Intraktabel pada Pasien Epilepsi," J. Penelit. Perawat Prof. Vol. 3, vol. 3, no. 3, pp. 471–475, 2021.

[4] G. A. Light et al., "Electroencephalography (EEG) and event-related potentials (ERPs) with human participants.," Curr. Protoc. Neurosci., vol. Chapter 6, p. Unit 6.25.1-24, Jul. 2010, doi: 10.1002/0471142301.ns0625s52.

[5] Y. Sugianela, Q. L. Sutino, and D. Herumurti, "Eeg Classification for Epilepsy Based on Wavelet Packet Decomposition and Random Forest," J. Ilmu Komput. dan Inf., vol. 11, no. 1, p. 27, 2018, doi: 10.21609/jiki.v11i1.549.

[6] A. A. Ningrum, I. Syarif, A. I. Gunawan, E. Satriyanto, and R. Muchtar, "Algoritma Deep Learning-LSTM untuk Memprediksi Umur Transformator," J. Teknol. Inf. dan Ilmu Komput., vol. 8, no. 3, 2021, doi: 10.25126/jtiik.2021834587.

[7] A. S. Bayangkari Karno, "Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) Dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python.," Ultim. InfoSys J. Ilmu Sist. Inf., vol. 11, no. 1, 2020, doi: 10.31937/si.v9i1.1223.

[8] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top., vol. 64, no. 6, p. 8, 2001, doi: 10.1103/PhysRevE.64.061907.

[9] I. Basyar, Adiwijaya, and D. T. Murdiansyah, "Email spam classification using gated recurrent unit and long short-term memory," J. Comput. Sci., vol. 16, no. 4, 2020, doi: 10.3844/JCSSP.2020.559.567.

[10] G. Cheng, V. Peddinti, D. Povey, V. Manohar, S. Khudanpur, and Y. Yan, "An exploration of dropout with LSTMs," in Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2017. doi: 10.21437/Interspeech.2017-129.

[11] M. S. Farooq, A. Zulfiqar, and S. Riaz, "Epileptic Seizure Detection Using Machine Learning: Taxonomy, Opportunities, and Challenges," 2023. doi: 10.3390/diagnostics13061058.

[12] D. Chen, S. Wan, J. Xiang, and F. S. Bao, "A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG," PLoS One, vol. 12, no. 3, 2017, doi: 10.1371/journal.pone.0173138.

[13] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks BT - Computer Vision–ECCV 2014," in European Conference on Computer Vision (ECCV), vol. 8689, no. Chapter 53, 2014.

[14] J. Rabcan, V. Levashenko, E. Zaitseva, and M. Kvassay, "EEG Signal Classification Based on Fuzzy Classifiers," IEEE Trans. Ind. Informatics, vol. 18, no. 2, 2022, doi: 10.1109/TII.2021.3084352.