ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

Epileptic Seizure Classification Based on Energy Compaction of Different Transformation Technique and Machine Learning Classifiers

Kusumika Krori Dutta, Dr. Premila Manohar, Falak Naaz, and Shwethaa Rajagopalan

Abstract - Worldwide around 70 million people suffers with epilepsy and to diagnose seizure type from Electro-Encephalogram (EEG) signal, many mathematical models and methods have been proposed. Correctness in diagnosis is essential to enhance the efficacy of treatment. In this paper, an incipient pre-processing technique, utilizing energy compaction property of various transformations like Fast Fourier Transforms (FFT), Discrete Cosine Transforms (DCT), Hilbert Transforms (HT) is used and relegation of multi-class was done utilizing different machine learning algorithms like decision tree, k nearest neighbours(KNN) Discriminant Analysis, Naive Bayes Classifier, Support Vector Machine (SVM) using MATLAB.

Index Terms - EEG, Epileptic seizure, FFT, DCT, Hilbert transforms, KNN,SVM, boosted tree, Q Discriminant, TUH EEG.

INTRODUCTION

Epilepsy is a very old and one of the most common neurological disorders worldwide[1][2]. In a clinical set up, EEG is used to diagnose its presence and type. This is a tedious and error prone method when coupled with an absence of proper epileptic-attack history. Seizures may occur at any frequency and intensity. Temporary symptoms such as a loss of awareness or consciousness, and disturbances of movement, sensation (including vision, hearing and taste), mood or other cognitive functions may occur (as shown in Fig 1). Physical injuries like fractures and bruising may also occur. Each of these symptoms indicate a different type of epileptic seizure [3]-[6]. The knowledge on experiences that a patient undergoes and the actions carried out by them during the seizure plays a key role in correctly diagnosing the type of epilepsy/epileptic seizure. A person suffering from epilepsy usually does not remember the events that took place during the attack. Since the occurrence of seizures is unpredictable, it is not always guaranteed that another person who is aware of their condition, or who can reliably convey information on the episode, is present. Without this information, it is highly impossible for a clinician to diagnose the type of epilepsy/epileptic seizure. Seizures are the main symptoms of epilepsy. They may be focal, generalized, complex partial, tonic-clonic, etc. Focal seizures affect only a certain part of the brain, while generalized seizures affect the entire brain.

To aid diagnostics for identification of the type of seizure, from EEG signals, researchers have developed varied algorithms and mathematical models for different pre-processing techniques with various transforms[7]. With the advancement of machine learning methods, many algorithms have been developed using statistical, frequency domain and nonlinear parameters to detect epileptic seizures [8]-[15].

FIGURE 1
TEMPORARY SYMPTOMS DURING EPILEPTIC EPISODE

Through this research work, multi- class epileptic seizure classification is tried out with pre-processing using energy compaction property of various different transformations like FFT, DCT and HT techniques and then classified using different Machine Learning techniques. In second segment of this paper gives description of dataset used, followed by different pre-processing techniques based on number of features has been discussed. Third segment deals with pre-processing of data and the fourth segment discusses different machine learning methodologies and implementation, using MATLAB [16]. This is followed by result and conclusion.

DATASET

The Temple University Corpus (TUH EEG Corpus)\cite{b17} is a huge collection of clinical EEG data collected across more than a decade. The data is stored as an EDF (European Data Format) file, and consists of 24 to 36 channels of signal data, sampled at 250 Hz using 16 bits per sample. The number of sessions per year varies from approximately 1,000-2,500 (with the exception of years 2000-2002, and 2005, in which limited numbers of complete reports were found in the various electronic medical record archives. The data chosen follows the average reference configuration (AR), while the annotations follow the TCP channel configuration. Each EDF file is accompanied by an anonymized report given by a neurologist, including their findings and analysis. It the seizure event. Files were taken from version 1.5.0, released in March 2019. The corpus consists of 10 types of seizures: Focal Non-Specific Seizure (FNSZ), Generalized Non-Specific (GNSZ), Simple Partial Seizure (SPSZ), Complex Partial Seizure (CPSZ), Absence Seizure (ABSZ), Tonic Seizure (TNSZ), Clonic Seizure (CNSZ), Tonic Clonic Seizure (TCSZ), ATSZ (Atonic Seizure) and MYSZ (Myoclonic Seizure). A four-class classification problem is being analyzed- using the Generalized (GNSZ), Focal Non-Specific (FNSZ), Tonic (TNSZ) and Complex Partial (CPSZ) classes-with 49, 47, 50 and 59 recordings. respectively. The EEG Corpus also contains basic information of the patients such as age, gender, medication and history. From the collected data, the total seizure duration extracted is 10344.2938 seconds. The age of the patients ranged from 27 to 91 years.EDF file is extracted from the dataset and with the help of seizure start time and stop time given in excel file, the exact signal at the time of seizure occurrence has been extracted. In this paper mainly four type of seizure classification is considered as shown in Table 1.

TABLE1
DESCRIPTION OF SEIZURE TYPES SELECTED FOR STUDY

Code	Seizure	Indication	Description
GNSZ	Generalized Non-Specific Seizure	Electrographic	Generalized seizures which cannot be further classified into one of the groups below
FNSZ	Focal Non-Specific Seizure	Electrographic	Focal seizures which cannot be specified by its type
TNSZ	Tonic Seizure	Electrographic/Clinic al	Stiffening of body during seizure (EEG effects disappear)
CPSZ	Complex Partial Seizure	Electrographic/Clinic al	Partial seizures during unconsciousness. Type specified by clinical signs only.

PRE-PROCESSING OF DATA

The EDF files were downloaded from the TUSZ corpus for seizures selected above. With the help of the start and stop times of seizures provided by the corpus, each signal has been segmented to only those areas that represented the seizure event which was then individually converted to CSV format. Each row in these CSV files represented a channel and the columns were the samples. Before feeding these CSV file data to any model, they had to be pre-processed. Owing to the fact that the last few channels in the signals generally correspond to background signals and that the data had differing numbers of channels across files, a common value of 26 channels was chosen. The files were then labelled as 0, 1, 2 and 3 for FNSZ, GNSZ, TNSZ and CPSZ respectively. In addition to this, around 79 columns were found to be empty and were therefore removed. The lengths of the samples were found to be of unequal. In order to equalize the lengths while avoiding loss of frequency information, the signals were repeated to a length of 16384 samples. An additional label column was added for the seizure types resulting in a 5251x16385 matrix . Final data was then put into a matrix of size 5251x16385 with the last column for labels.

I. *Energy aspects of FFT approach to Extract Feature*

Fast Fourier Transform provides efficient transform algorithm. It is complex valued transform where the most of the energy is concentrated at the corner of transformed 2D signal[7][18][21]. It is used to extract features from signals in the form of frequencies present as given in equations (1) and (2) which represents forward and inverse FFT for 2D signal respectively, where T is the FFT coefficient matrix. Equations (3) and (4) shows matrix form of inverse and forward FFT respectively.

$$F(x,y) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m,n) e^{-j2\pi (x\frac{m}{N}+y\frac{n}{N})} F(x,y) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m,n) e^{-j2\pi (x\frac{m}{N}+y\frac{n}{N})} \quad (1)$$

$$f(m,n) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} F(x,y) e^{j2\pi (x\frac{m}{N}+y\frac{n}{N})} f(m,n) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} F(x,y) e^{j2\pi (x\frac{m}{N}+y\frac{n}{N})} \quad (2)$$

$$[f]_{N\times N} = [T^{*T}]_{N\times N}[F]_{N\times N}[T^{*}]_{N\times N}[f]_{N\times N} = [T^{*T}]_{N\times N}[F]_{N\times N}[T^{*}]_{N\times N} \quad (3)$$

$$[F]_{N\times N} = [T^{*}]_{N\times N}[f]_{N\times N}[T^{*T}]_{N\times N}[F]_{N\times N} = [T^{*}]_{N\times N}[f]_{N\times N}[T^{*T}]_{N\times N} \quad (4)$$

The power of different frequencies is also obtained, which is used to filter actual signal frequencies from the noise frequencies. The data from the matrix was extracted row-wise and the corresponding FFT was taken. With the frequency components, the single-sided power spectrum of signal for various frequencies was calculated. The findpeaks function in MATLAB was used to find the spikes in the power spectrum. Of these, frequencies having 8,10 and 15 highest peaks in power were extracted as features of the signal. As the last column of the matrix is the target variable (for seizure type), it was taken as a categorical column.

II. Energy compact Discrete Cosine Transform approach to Extract Feature

DCT is a real valued discrete sinusoidal unitary transform which is the most popular because of its energy compaction [7][18]. The 2D DCT forward F(u,v) and inverse f(x,y) formula is shown in equations (5), (6) and (7) respectively, Here f(x,y) is the time domain representation of 2D signal and F(u,v) is the frequency domain counter part of it. T is the DCT coefficient matrix. Equations (8) and (9) shows matrix form of inverse and forward DCT respectively.

Though DCT and FFT offer the same time for implementation, but, DCT gives higher energy compaction for the same 2D signal.

$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y) \quad if\ u=0\ and\ v=0 \tag{5}$$

$$F(u,v) = \frac{2}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)\cos[\frac{(2x+1)u\pi}{2N}]\cos[\frac{(2y+1)v\pi}{2N}]$$

If u≠0 and v≠0 (6)

$$f(x,y) = \frac{2}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} F(u,v)\cos[\frac{(2x+1)u\pi}{2N}]\cos[\frac{(2y+1)v\pi}{2N}] \tag{7}$$

Matrix Form

$$[f]_{N\times N} = [T^T]_{N\times N}[F]_{N\times N}[T]_{N\times N}[f]_{N\times N} = [T^T]_{N\times N}[F]_{N\times N}[T]_{N\times N} \tag{8}$$

$$[F]_{N\times N} = [T]_{N\times N}[f]_{N\times N}[T^T]_{N\times N}[F]_{N\times N} = [T]_{N\times N}[f]_{N\times N}[T^T]_{N\times N} \tag{9}$$

III. Energy aspects of Hilbert Transform to Extract Feature

The relation between Discrete Time Fourier Transform (DTFT) and Hilbert transform (HT) is F(w) in DTFT is [$-isgn-isgn$ (w)F(w)] in case of HT,[19-21].

Where |$-isgn(w)-isgn(w)$|=1 except for w=0 and since F(w) does not contain any impulse at the origin, so

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}|-isgn(w)\mathrm{F(w)}|^2dw\,\frac{1}{2\pi}\int_{-\infty}^{\infty}|-isgn(w)\mathrm{F(w)}|^2dw = \frac{1}{2\pi}\int_{-\infty}^{\infty}|\mathrm{F(w)}|^2dw\frac{1}{2\pi}\int_{-\infty}^{\infty}|\mathrm{F(w)}|^2dw \tag{10}$$

According to equation (10) it has been observed that the energy aspects of of DTFT and Hilbert transform is same. Equation (11) and (12) shows discrete Hilbert transform formula

$$H[K] = -isgn\left(\frac{N}{2}-k\right)sgn(k)H[K] = -isgn\left(\frac{N}{2}-k\right)sgn(k) \tag{11}$$

$$h[n] = \frac{2}{N}\sum_{k=1}^{\frac{N-1}{2}}\sin(\frac{2\pi kn}{N})h[n] = \frac{2}{N}\sum_{k=1}^{\frac{N-1}{2}}\sin(\frac{2\pi kn}{N}) \tag{12}$$

These features were then collected into a 5251x11 matrix and given as input to the MATLAB Classification Learner App. The cross-fold validation was set to 5 and all available models were trained for the input data. Training was done for both unshuffled and shuffled matrices for all 3 sets of features.

MECHINE LEARNING CLASSIFIERS

Some of the machine learning classifiers used in this paper is briefly explained as follows:

IV. Decision Trees:

Decision trees are a supervised predictive modelling tool for both classification and regression tasks. They form a tree-like graph with each node represents a particular non-linear decision-making surface, leaves represent the output and edges represent the decision of the nodes. A parameter called Gini Index in equation (13) is used analogous to a cost function that is used to evaluate the splits in the dataset and entropy is shown in equation (14).

$$Gini = 1 - \sum_{i=1}^{c}(p_i)^2 Gini = 1 - \sum_{i=1}^{c}(p_i)^2 \tag{13}$$

$$\text{Entropy(S)} = -P_{yes}\log_2 P_{yes} - P_{no}\log_2 P_{no} \tag{14}$$

Where p_i is the probability of an event i of state S and c is the number of classes.

V. Discriminant Analysis

Discriminant Function Analysis is a dimensionality reduction technique (follows equation (15)) used in supervised classification problems. They mainly work by projecting higher dimensional features onto lower dimensional space. The most common type is

Linear Discriminant Analysis (LDA) used when the data is linearly separable. The criterion used for Linear Discriminant Analysis for optimal clustering is to find transformation G^T such that

$$\max_G \, trace\,((G^T S_\omega G)^{-1}(G^T S_b G)) \max_G \, trace\,((G^T S_\omega G)^{-1}(G^T S_b G)) \qquad (15)$$

Where S_ω is within-class scatter matrix and S_b is between-class matrix.

Nearest neighbour Classifiers:

Nearest Neighbour Classifiers[22] achieve consistently high performance in supervised pattern recognition as it does not assume distributions of the training examples a priori. A test sample is classified by calculating distances to the nearest training samples and assigning the class of the majority to the test. The distance between two points is generally calculated as the Euclidean Distance as in equation (16).

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (16)$$

Where $(x_1, y_1)\,(x_1, y_1)$ and $(x_2, y_2)\,(x_2, y_2)$ are the points in consideration.

VI. Ensemble Classifiers:

Ensemble Classifiers improve the outcomes of a machine learning model by combining several models, drawing on the principle of Wisdom of the Crowd. A collective prediction of the models is taken through a voting process to determine the final output. The advantage of these classifiers is that they can work with models with low accuracies. There are two types of voting – hard voting and soft voting. Hard voting takes the majority votes into consideration and is given by equations (17) and (18).

$$C = \mathrm{mode}(c_i) C = \mathrm{mode}(c_i) \qquad (17)$$

Where c_i is the predictions of various classifiers and C is the aggregated prediction. Soft voting takes the weighted average of the probabilities of the predictions for the final outcome as

$$C_j = \frac{\sum_{i=1}^{N} w_i P(c_{ij})}{N} C_j = \frac{\sum_{i=1}^{N} w_i P(c_{ij})}{N} \qquad (18)$$

Where $P(c_{ij})$ is the probability of classifier i classifying the sample into class j. N is the total number of classifiers.

VII. Support vector machine

Support Vector Machines (SVM), a supervised learning algorithm that can efficiently map features into higher dimensions by designing a hyperplane. The distance between the hyperplane and classes (known as margin) is maximized while making the hyperplane lie in the centre of the margin. SVM Hyperplane equation in 'M' dimensions is given by equation (19)

$$y = \omega^T \phi(x) + b y = \omega^T \phi(x) + b \qquad (19)$$

Where, $\phi(x)\,\phi(x)$ is the fixed feature-space transformation, b is the bias parameter and ω^T is the weight vector.

VIII. *Naïve bayes classifier:*

Naïve-Bayes Classifiers are supervised learning probabilistic classification models that assume that the predictor variables are independent of each other. It is based on the fundamental Bayes' Theorem as shown in equation (20). If xi represents the predictors and y gives the class of a particular sample with maximum probability, then\

$$y = \mathrm{argmax}_y \, P(y) \prod_{i=1}^{n} P(x_i|y) \qquad y = \mathrm{argmax}_y \, P(y) \prod_{i=1}^{n} P(x_i|y) \qquad (20)$$

Where n is the total number of classes.

RESULTS AND DISCUSSION

For each of the transformation, pre-processing of the data is performed for different number of features. The network vs accuracy bar chart using FFT with 20 and 30 features, DCT with 50 and 80 features, HT with 50 and 60 features is shown in Fig. 2 to Fig. 7. For full study, cross validation with 5 folders is used for all the classifiers. It can be observed that for FFT with 20 features the highest accuracy of 80.4% is achieved with optimizable KNN method whereas for FFT with 30 features, Bagged Trees classifier gives 85% accuracy. DCT with both 50 features and 80 features gives a maximum accuracy of 83.3%, through medium Gaussian SVM classifier. The Hilbert Transform with 50 features and 60 features, with medium Gaussian SVM classifier gives an accuracy of 76.7%.

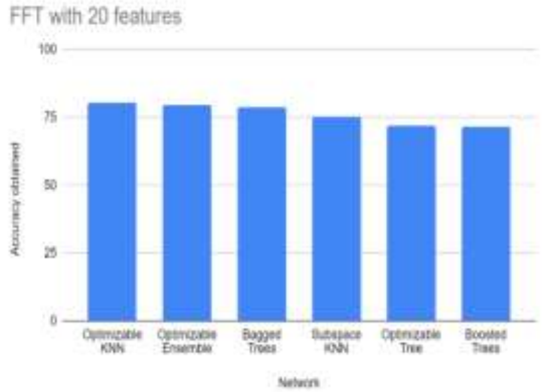

FIGURE 2
NETWORKS VS ACCURACY WITH FFT WITH 20 FEATURES

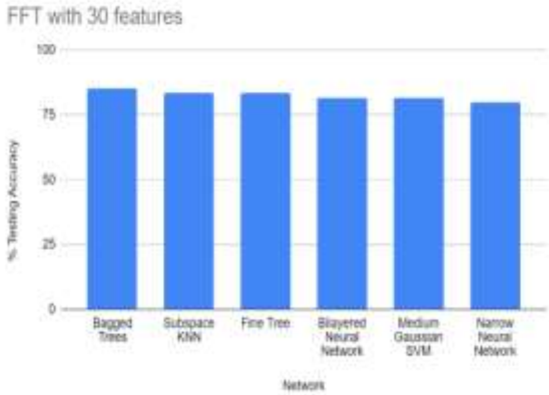

FIGURE 3
NETWORKS VS ACCURACY WITH FFT WITH 30 FEATURES

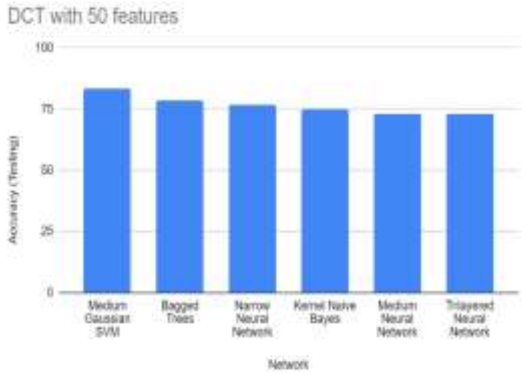

FIGURE 4
NETWORKS VS ACCURACY USING DCT WITH 50 FEATURES

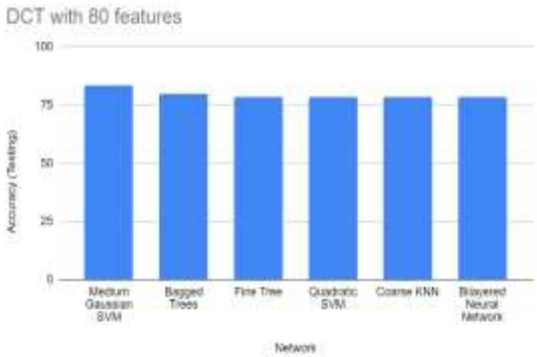

FIGURE 5
NETWORKS VS ACCURACY USING DCT WITH 80 FEATURES

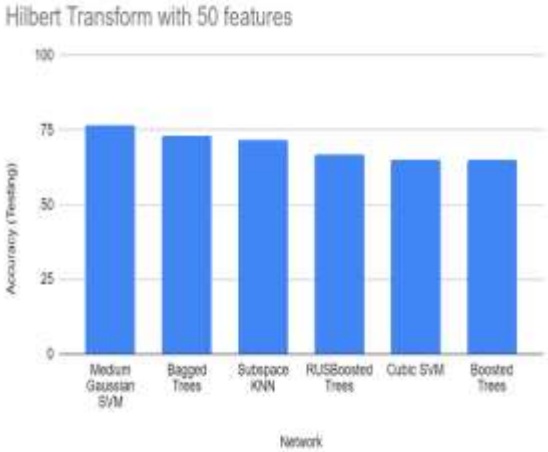

FIGURE 6
NETWORKS VS ACCURACY USING HT WITH 50 FEATURES

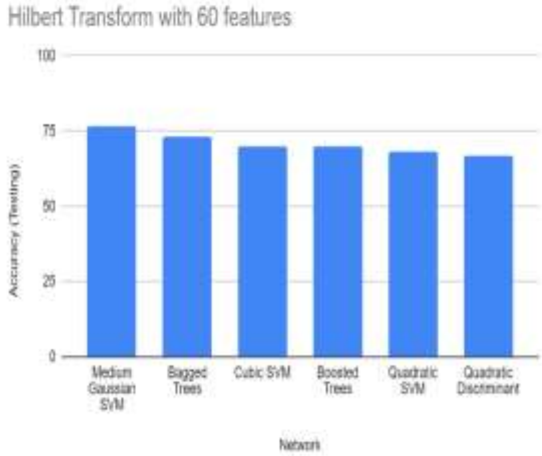

FIGURE 7
NETWORKS VS ACCURACY USING HT WITH 60 FEATURES

CONCLUSION

In this paper energy compaction property of DFT, DCT and HT is used for feature extraction and classification carried out with many machine learning classifiers with different number of features, for each of the transformations viz: FFT, DCT and HT. A comparison of the results indicates that medium gaussian SVM classifier gives better accuracy in all types of transformations with different features.

In future, dimensionality reduction method will be planned before extracting the energy of each EEG samples

.

REFERENCES

[1] Wang, L.; Xue, W.; Li, Y.; Luo, M.; Huang, J.; Cui, W.; Huang, C. Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis. Entropy 2017, 19, 222. https://doi.org/10.3390/e19060222

[2] Alotaiby, T.N., Alshebeili, S.A., Alshawi, T. et al. EEG seizure detection and prediction algorithms: a survey. EURASIP J. Adv. Signal Process. 2014, 183 (2014). https://doi.org/10.1186/1687-6180-2014-183

[3] Vetrikani, R., and T. Christy Bobby. 2017. "Diagnosis of Epilepsy — A Systematic Review." In 2017 Third International Conference on Biosignals, Images and Instrumentation (ICBSII), 1–5. Chennai: IEEE. https://doi.org/10.1109/ICBSII.2017.8082300.

[4] WHO, International League against Epilepsy, and International Bureau of Epilepsy, eds. 2005. Atlas: Epilepsy Care in the World. Geneva: WHO Press, World Health Organization.

[5] Zhang, Zisheng, and Keshab K. Parhi. 2016. "Low-Complexity Seizure Prediction From IEEG/SEEG Using Spectral Power and Ratios of Spectral Power." IEEE Transactions on Biomedical Circuits and Systems 10 (3): 693–706. https://doi.org/10.1109/TBCAS.2015.2477264.

[6] Kusumika K.D. Epilepsy is a curse: Myth or Reality

[7] An article published by InnoHEALTH magazine digital team, October 19, 2020,

[8] https://innohealthmagazine.com/2020/issues/epilepsy-is-a-curse-myth-or-reality/

[9] D. Ramesh Varma, K. Pavan Raju, M. V. R. V Prasad, Texture Synthesis using Energy Compaction Property of Different Transforms, International Journal of Computer Applications (0975 – 8887) Volume 167 – No.8, June 2017.

[10] S. Roy, U. Asif, J. Tang and S. Harrer, "Seizure Type Classification Using EEG Signals and Machine Learning: Setting a Benchmark," 2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), 2020, pp. 1-6, doi: 10.1109/SPMB50085.2020.9353642.

[11] E. M. Imah and A. Widodo, "A comparative study of machine learning algorithms for epileptic seizure classification on EEG signals," 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2017, pp. 401-408, doi: 10.1109/ICACSIS.2017.8355065.

[12] Takahashi H, Emami A, Shinozaki T, Kunii N, Matsuo T, Kawai K. Convolutional neural network with autoencoder-assisted multiclass labelling for seizure detection based on scalp electroencephalography. Comput Biol Med. 2020 Oct;125:104016. doi: 10.1016/j.compbiomed.2020.104016. Epub 2020 Sep 26. PMID: 33022521.

[13] Application of machine learning to epileptic seizure onset detection and treatment-

[14] https://dspace.mit.edu/handle/1721.1/54669

[15] C. Chiang, N. Chang, T. Chen, H. Chen and L. Chen, "Seizure prediction based on classification of EEG synchronization patterns with on-line retraining and post-processing scheme," 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp. 7564-7569, doi: 10.1109/IEMBS.2011.6091865.

[16] Kusumika Krori Dutta, Kavya V., Sunny Arokia Swamy,"Removal of Muscle Artifacts from EEG Based on Ensemble Empirical Mode Decomposition and classification of Seizure using Machine Learning Techniques",IEEE- International Conference on Inventive Computing and Informatics (ICICI 2017) 21st Nov'17,IEEE Xplore Compliant - Part Number: CFP17L34-ART, ISBN: 978-1-5386-4031-9.

[17] Kusumika Krori Dutta, Dr. Arun Shashidhar, "Application of Machine learning Techniques in Electro-encephalography Signals", chapter 3 of book titled " Brain and Behavior Computing", 1st edition published by CRC press (c)2021 Taylor and Francis group LLC ISBN: 978-1-003-09288-9(ebk) page no. 61-84, May 2021

[18] O. Smart and M. Chen, "Semi-automated patient-specific scalp EEG seizure detection with unsupervised machine learning," 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2015, pp. 1-7, doi: 10.1109/CIBCB.2015.7300286.

[19] Sundar, Aditya & Das, Chinmay. (2015). MATLAB analysis of EEG signals for diagnosis of epileptic seizures. 10.13140/RG.2.1.1367.4088

[20] TUH-EEG Seizure Corpus- Shah, V., von Weltin, E., Lopez. S., McHugh, J., Veloso, L.,Golmohammadi, M., Obeid, I., and Picone, J. (2018). The Temple University Hospital Seizure Detection Corpus. Frontiers in Neuroinformatics. 12:83. doi: 10.3389/fninf.2018.00083

[21] Cooley, James W., and John W. Tukey. 1965. "An Algorithm for the Machine Calculation of Complex Fourier Series." Mathematics of Computation 19 (90): 297–301.

[22] Mark S. Nixon, Alberto S. Aguado, Images, sampling, and frequency domain processing as a book chapter in Feature Extraction & Image Processing for Computer Vision (Third Edition), 2012.

[23] Mathias Johansson, The Hilbert transform , Master Thesis Mathematics/Applied Mathematics, 2012.

[24] Bruns, Andreas. 2004. "Fourier-, Hilbert- and Wavelet-Based Signal Analysis: Are They Really Different Approaches?" Journal of Neuroscience Methods 137 (2): 321–32.

[25] https://doi.org/10.1016/j.jneumeth.2004.03.002.

[26] Altman, N. S. 1992. "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression." The American Statistician 46 (3): 175–85. https://doi.org/10.1080/00031305.1992.10475879.