Journal of Education for Pure Science- University of Thi-Qar
Vol.13, No.2 (June., 2023)
Website: jceps.utq.edu.iq Email: jceps@eps.utq.edu.iq

DOI: http://doi.org/10.32792/utq.jceps.10.01.01

Neonatal Seizure Detection based on EEG signal by Ensemble Method

Sura S.Rasheed [1] **Firas S.Miften[2]**

[1, 2,] Department of Computer Science, College of Education for Pure Sciences , University of Thi-Qar, Iraq

Received 17/4/2023, Accepted 17/5/2023, Published 11/6/2023

Abstract:

Objective in the newborn critical care unit, seizures are one of the most frequent medical emergencies. They are identified by visually electroencephalography (EEG) results assessment, and neuroscientific professionals are responsible for their care. This method takes a long time, and the results are inconsistent. In order to solve this issue, a system for identifying epileptic seizures from EEG data collected from healthy individuals and epileptic patients is presented in this research. This system is based on a study of EEG data utilizing the discrete wavelet transform (DWT) and both linear and nonlinear classifiers.

Methods: Neonatal human EEG recordings are included in the dataset used in our method. A 19-channel EEG equipment was used to record the brainwave activity of 79 term newborns referred to the NICU at the Helsinki University Hospital. Five patients with structural seizure and non-seizure annotations were chosen for the pilot study from these databases. Three experts independently identified the presence of seizures and non-seizures in the EEGs using visual interpretation, while statistical features were retrieved using the k-mean Algorithm and RS was utilized to identify seizures**.**

Result: To improve the model's performance, (statistical and jorth and zero crossing) features were extracted. K-means' feature selection using RS, and obtain high accuracy by all channels WHICH IS 99% and 100% by C4.

Conclusion: A reliable, autonomous epileptic seizure detection system that can be used in real time to enhance healthcare and life quality is seizure detection based on DWT statistical attributes and RS.

Keywords: Electroencephalograms (EEG), seizure, discrete wavelet transform (DWT), Random Subspace Ensemble (RS).

1-Introduction:

The cortical and subcortical neurons' unchecked activity is what leads to seizures [1, 2]. With a recorded diagnosis rate of 1.5–3.5 per 1000 newborns, they are the most frequent neonatal neurological emergency [3] Several studies have demonstrated that seizures impair the neurodevelopment and worsen basic injuries like ischemia and bleeding that already occur [4, 5].Additionally, because newborn seizures are known to have a wide range of clinical manifestations and are challenging to distinguish from common neonatal movements, they rank among the most dangerous clinical issues in the world [6]. The sole technology now used to identify neonatal seizures is electroencephalography (EEG) ,This enables neurophysiologists to monitor and recording brain impulses occurring in the patient's brain. Since it is widely acknowledged that EEG is helpful in detecting seizures and distinguishing infants who should receive neuroprotective medication, newborn EEG has become necessary [7]. However, implementing multichannel EEG in neonatal intensive care units (NICUs) can be difficult as neurophysiologists with specialized knowledge in correctly interpreting its results may not always be readily available. Neurology laboratory technicians are taught how to safely handle and set the EEG system's recording electrodes on the patient's scalp. Medical personnel have used amplitude integrated EEG (aEEG), a less sophisticated form of EEG, at NICUs in healthcare facilities all around the world [8]. Although an aEEG is useful for tracking brain activity, Its accuracy tends to fluctuate.. Rennie et al. tested 40 babies in 2004 and 851 in 2007, and they found that their accuracy was, respectively, 38-55% [9] and 12-38% [10]. Evans et al. Using an aEEG is likely to overdiagnose seizures, according to a recent study that looked at the sensitivity and accuracy of seizure diagnosis in 44 neonates. [11]. Several automatic neonatal seizure detection techniques based on machine learning (ML) have been presented to support clinical seizure detection [12, 13]. Due to the time-consuming nature of attaching multichannel EEG electrodes, the developed algorithms only handle one EEG channel. [14].

2-RELATED WORK

Researchers have conducted significant attempts to enhance seizure diagnosis in [15]), using 10 features (Time-Domain Features, Entropy Based Features), 10 different FS algorithms were used in the study to minimize the classification cost. , (SVM and KNN) was utilized and obtain an accuracy of (98.8%), in [16] using e frequency domain features, use power spectral analysis to extract features per channel, a random decision forest was used for selecting a limited number of channels, a KNN classifier technique and obtain an accuracy of (80.87%), in [17] using (Time-frequency signal features ,Time-frequency statistical features, Time-frequency image), feature fusion approaches was used to extract features from multi-channel , and used sequential forward selection (SFS),sequential backward selection (SBS), sequential forward floating selection (SFFS) was used to feature selection , (SVM ,ANN, random forest (RF)) and obtain an accuracy of 85.70% , in [18]), using time- domain statistical features, using random forest with cost matrix to solve the problem of imbalance while detecting the seizure during long hour EEG recordings, obtain an accuracy of 98%, in [19] used 55 feature from each channel of (frequency domain, time domain, and information theory based characteristics of the signals, CNN AND SVM used and obtain performance (0.970,0.925), in [20] , using line length feature and Burst detection algorithm was utilized and obtain an accuracy of (84%), in [21] utilized the median of the time-frequency correlations from the TFD, amplitude envelope, frequency spectrum features, and SNLEO features (using adaptive segmentation with the smoothed non-linear energy operator). SVM is utilized, with a 95%

accuracy, in [22] used PCA Features that extracted by PCA, used one channel, logistic regression (LR), dense tree (DT), 2D support vector machine (2D SVM), and cosine k-nearest neighbor (KNN)), and obtain an accuracy of 91% by KNN.

3-Methodology of proposed methods

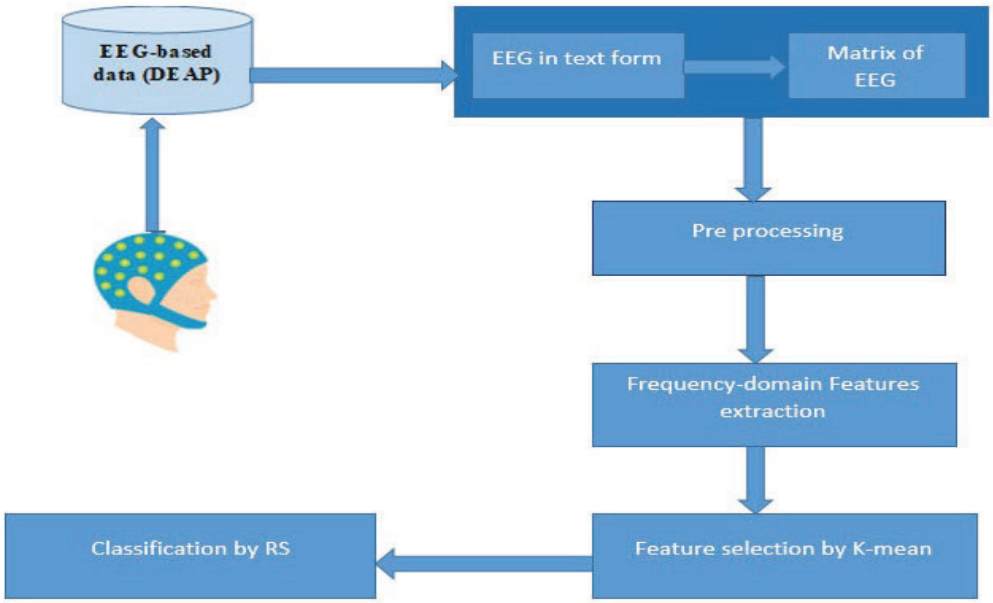

Figure1. proposed model for seizure detection from EEGs

The findings of the work on the RS algorithm were examined in order to identify seizures in their early phases. Electrodes inserted in the EEG cap were used to record EEG signals, which were then used as input into the system under consideration. The system determines whether the patient has seizures or not. By providing quicker and better findings, this tool can assist medical professionals in the laborious task of seizure prediction.

3.1-Data set:

Based on the Helsinki Dataset, which is available to the general public. This dataset [23] is composed of multi-channel EEG recordings from 79 neonates hospitalized to the Neonatal Intensive Care Unit (NICU) at the Helsinki University Hospital between 2010 and 2014. The majority of the recordings were made within a week of birth and are from 42 male and 37 female newborns. NicoletOne EEG amplifier (sampling frequency of 256 Hz) and EEG caps (Sintered Ag/AgCl electrodes) were used to record the signals. The electrode placement for the 19 electrodes followed the international 10-20 standard, with a recording reference electrode inserted at the midline. Maximum 21 channels, each recorded at 256 Hz, make up the dataset. With an IQR of 64 to 96 minutes, The average signal lasted 74 minutes. There is also access to the neonates' clinical data. Three medical experts who each annotated an average of 460 seizure occurrences on the recordings were used. If a seizure lasted more than 10 seconds, the recordings were marked as seizure. One such annotation indicating the location of a seizure is shown in Figure 2. A

spike in the image signifies the seizure. The specialists came to the conclusion that 22 newborns were seizure-free and 39 neonates had seizures.

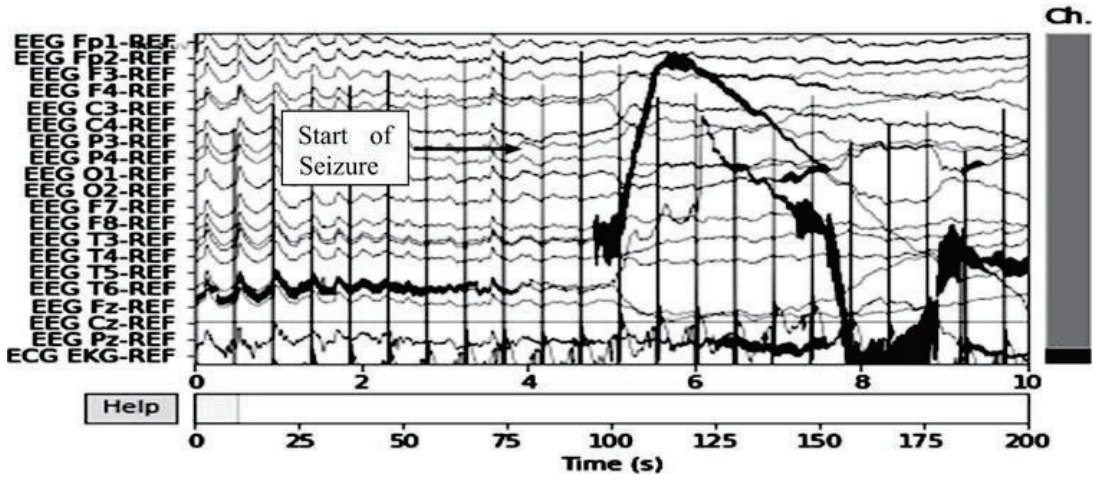

Figure 2. Annotation showing the start of seizure in EEG signals

3.1.1- dataset segmentation:

By sampling the data into 256 equal data segments, convert the signal data into a 2-D array. This facilitates analysis and provides crucial information, such as seizure detection,

3.1.2-selecting and designating patients:

Five patients in all, with seizure and non-seizure notes, were assigned in this research.as shown in table 1 to all channels in order to address the issue of imbalanced data. Three specialists independently evaluated the EEGs by visual observation for the presence of seizures and non-seizures.

Table 1. Patient ID and annotations for seizures

Patient ID	Time periods (sec) for Seizure annotation	Time periods (sec) for Non-seizure annotation
4	1059-1909	(428-609)
5	260-880, 993-1447, 1683-2254, 2262-2505, 2663-3082, 3117-3807	(1-238)
9	267-972	976-1573, 1600-2030
13	801-1296	1297-5506
14	1177-1521, 1945-2406	62-254

3.2- Pre-processing

A filter known as a band pass only allows signals that fall within a specific frequency band or passband range to pass through it, In this study, the fourth order Butterworth band pass filter was employed. This kind of filter was chosen because, in contrast to others, it exhibits a linear response. The frequency cutoff ranged from 4 to 40 Hz. Gamma frequency ranges within the range of 40 Hz are required for the analytical process and to prevent power line interference, the delta frequency (below 4 Hz) is excluded since it is viewed as noise (50 to 60 Hz).

3.3- Feature extraction

The EEG signals are divided into five bands (Gamma, Beta, Alpha, Theta, and Delta) using DWT in the feature extraction step, as shown in figure 3. Statistical measures (Arithmetic mean, Maximum value, Minimum value, Standard deviation, Variance, Approximate entropy, Median, Skewness, Kurtosis, Range, and Mode) are derived for each band, as well as the number of zero crossings (ZC) and Hjorth parameters.

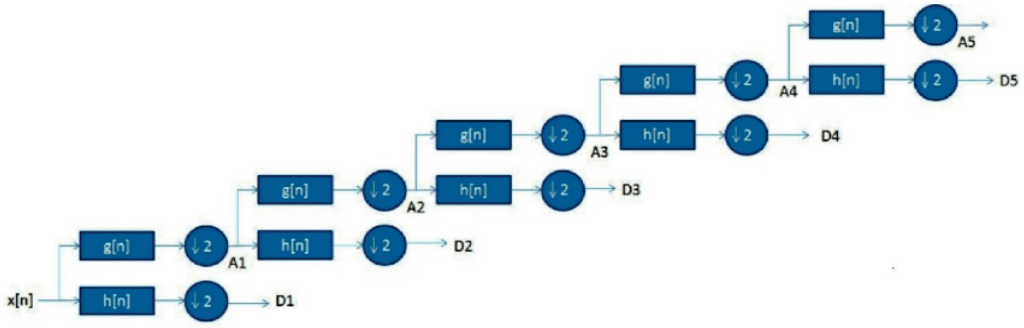

Figure 3. Five level wavelet decomposition of the EEG signal

Statistical measures: The statistical characteristics of EEG signals are collected at this stage. Table 2 offers succinct explanations of the qualities.

Table 2. Brief descriptions of the statistical features

Feature name	Formula	Explanation
Minimum value (MinV)	: MinV=min[X_n]	(1) whereX_n,n=1,2,3…n is a time series, n is the number of data points
Maximum value(MaxV)	: MaxV=max[X_n]	(2) AM is the mean of the sample
Standard deviation(SD	: $SD = \sqrt{\sum_{n=1}^{N}(x_n - AM)\frac{2}{N-1}}$	(3)
Arithmetic Mean (AM	: AM=$\frac{1}{N}\sum_{n=1}^{N} x_n$	(4)
Variance (V)	: V=$\sum_{n=1}^{N}(x_n - AM)\frac{2}{N-1}$	(5)
Skewness (S)	: S=$\sum_{n=1}^{N}(x_n - AM)\frac{3}{(N-1)SD^3}$	(6)
Kurthosis (K)	: K=$\sum_{n=1}^{N}(x_n - AM)\frac{4}{(N-1)SD^4}$	(7)

Median (MN)	**: MN=$\left(\frac{N+1}{2}\right)^{th}$**	**(8) If the number of values is odd then**
	: MN=$\frac{\left(\frac{N}{2}\right)^{th} value+\left(\frac{N}{2}+1\right)^{th} value}{2}$	**(9) If number of values is even then (where N= number of items)**

Number of zero crossings (ZC)

Zero-crossing is a time-based feature that is frequently utilized in signal processing, mathematics, electronics, and image processing. It indicates the number of zero crossings produced within a section. When there is a difference in the signals across samples, zero crossing happens. This feature may be used as a measure to quantify the noise rate in the signal. [24]. Researchers have found that throughout brain activity and various stages of sleep, the number of zero crossings in EEG signals changes [25].ZC is defined as ($x_{n-1} < 0$ and $x_n > 0$) or ($x_{n-1} > 0$ and $x_n < 0$)or ($x_{n-1} \neq 0$ and $x_n = 0$).

Hjorth parameters

Features that are frequently employed in the analysis of EEG signals include the Hjorth parameters (i.e., activity, mobility, and complexity) [26]. The first and second derivatives of signals are used to compute the Hjorth parameters, which are presented in Table 3.

Table 3. Parameters of Hjorth.

Feature name	Activity (HA)	Mobility (HM)	Complexity (HC)
Equation	HA=σ_0^2	HM=σ_1/σ_0	HC=$\sqrt{(\sigma_2/\sigma_1)^2-(\sigma_1/\sigma_0)^2}$

3.4-Feature selection

Different frequency bands in DWT correspond to distinct EEG rhythms since EEG is frequently characterized in terms of rhythmic activity. Seizure and non-seizure EEG segments in a particular EEG dataset may significantly differ in rhythmic activity in a particular frequency range (s). Similar to frequency bands, some features in an EEG dataset may be able to distinguish between seizure and non-seizure EEG segments, whereas other features merely produced data redundancy. For a few important features, we applied the k-mean clustering algorithm.

3.4.1-The K-means clustering algorithm

A feature selection approach based on k-mean clustering was suggested. The suggested model chose features based on Euclidean distance using k-means clustering. The k-means cluster was discovered to have a great potential for selecting the features with the greatest impact and removing the subpar ones [27]. The suggested model using k-means for feature selection is displayed in Figure 4. The following is a summary of the primary phases in feature selection using k-means:

Input: EEGs data

Output: Output: we got k features with a strong correlation, Following the elimination of redundant features

1. Select the number of groups K

2. Initialize the centroids by mixing the feature set first and then randomly selecting K feature points for the centroids.

3. Repeat until there is no longer any change in a middle point points.

4. We assigned S_i feature point to the nearest cluster judged by the sum of the Euclidean distance from the centroids of the cluster

$$A_d = \sum \sqrt{(S_i - N_k)^2} \qquad (1)$$

5. Recalculated the center pointer N_k for each cluster, to reflect the new tasks.

$$N_k = \frac{\sum_{i=0}^{n} X_i}{n} \qquad (2)$$

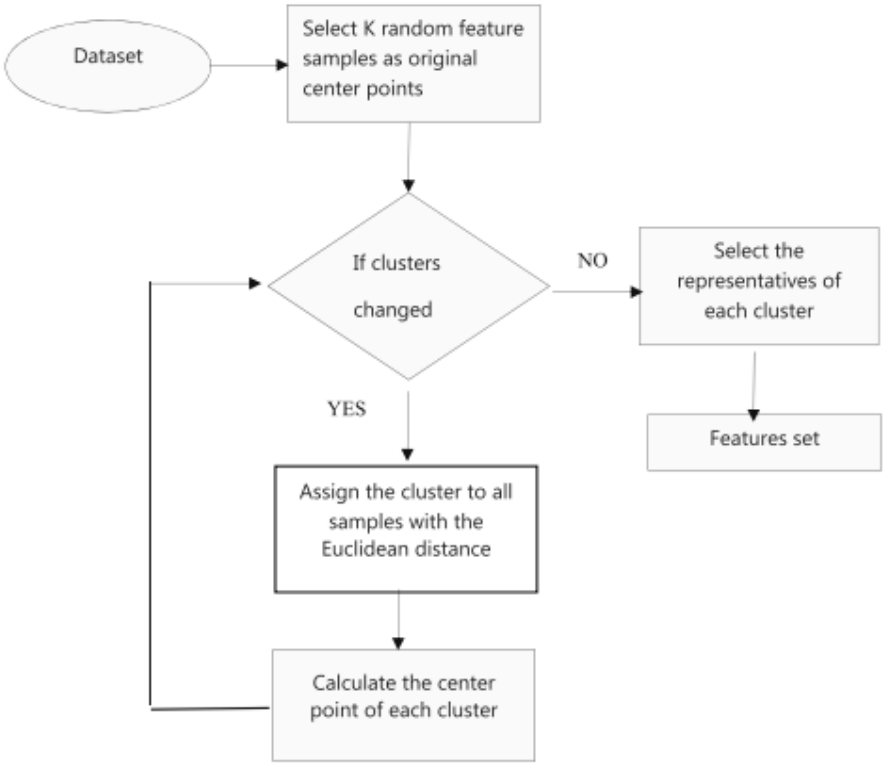

Figure 4. The flow diagram of k-means feature selection algorithm

4-Classification

The classifiers are used using the statistical features that were obtained via DWT. The classifier's job is to find abnormalities associated with epilepsy in EEG data using linear and non-linear mathematical

methods. To identify epileptic seizures in an individual's EEG data set and combine features produced from DWT with the data set, the classifier RS was employed in this study.

4.1-Random Subspace Ensemble

Random Subspace Ensemble is a machine learning approach which gathers the forecasts from several decision trees trained on various subsets of columns in the training dataset [28]. Randomly produced subsets of features are kept in subspace. The basic classifier (DT) is subsequently trained on each of the subsets in the subspace to create other classifiers. By gathering these classifiers, ensemble classifier E is created. The number of features in a subspace M and the number of classifiers (T), the RS with DT can be summarized by the following steps: 1) generate a subset of M features from the entire feature set for T times; 2) apply these features to the DT classifier and obtain T classification results; 3) produce the final classification map by combining the T predictions using a majority voting rule. The hyper-parameters of the classifiers are selected so that the best test accuracy is selected among the obtained results. M is chosen from data set feature size of {0.1, 0.2, 0.5} and the number of classifiers in the ensemble (T) is chosen from {50, 10, 20} as Algorithm.

Algorithm 3.1: Random Subspace Ensemble
Input: Training set S, T: number of classifiers, Dimension of the subspaces M
utput: Ensemble E
Begin *Step1: E ⇐ 0* *Step2: for* $i = 1$ *to* T *do* *Step3:* $S_i \Leftarrow$ *Select Random Subspace (S,*M*)* *Step4:* $C_i \Leftarrow$ *Construct Classifier (*S_i*)* *Step5:* $E \Leftarrow E \cup \{C_i\}$ *Step6: End*

K-fold cross validation gives the classification accuracy meaning [29]. Phases of the test and training process make up K-fold cross validation. The training data set is randomly divided into k segments. One K-1 part is utilized for the test set and the other part is used for training. K times are completed in this manner. The outputs from each round are added up, and the model's performance is assessed. In this study, k fold=10 was employed. Figure 5 [29] illustrates the k-fold validation scheme.

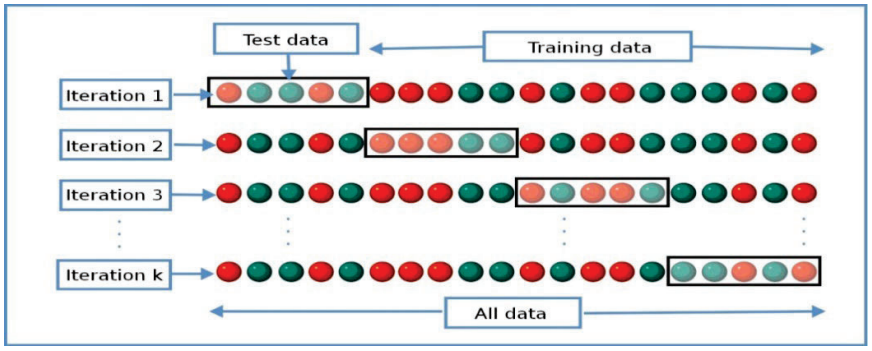

Figure 5. k_fold validation scheme

5-Performance Evaluation

Sensitivity, Specificity, Precision, Accuracy, and F-Measure performance criteria were utilized in this study to evaluate the efficacy of the classification process. These are the fundamental definitions that were utilized to determine the parameters [28].

True positive (TP) = Number of cases classified as having seizures and those who really encountered seizures.

False positive (FP) = Number of cases that were classified as not having seizures but really had seizures.

True negative (TN) = Number of cases that were classified as not having seizures even though they didn't actually have any.

False negative (FN) = The number of instances that were classified as having seizures even when they did not occur.

Accuracy (Acc): The capability to discriminate between patients and healthy cases with accuracy. Equation 9 illustrates the determination of accuracy. [31]

$$\boldsymbol{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (9)$$

Precision (P): A test's sensitivity refers to its ability to correctly recognize patient cases. In patient situations, the true positive rate is computed. to determine the sensitivity. Equation 10 explains the calculation process. [30]

$$P = \frac{TP}{TP + FN} \qquad (10)$$

Specificity (S): The ability of a test to correctly identify people without the disease . Therefore, in healthy circumstances, It is important to identify the true negative ratio.. It is denoted mathematically by equation 11[30].

$$S = \frac{TN}{TN + FP} \qquad (11)$$

Sensitivity (TPR): Displays the percentage of positive results from the TP test. It is denoted mathematically by equation 15.

$$TPR = \frac{TP}{TP + FN} \qquad (12)$$

F-measure is the weighted harmonic mean of its precision and recall

$$F - measure = 2 \times \frac{Precision * recall}{Precision + recall} \qquad (13)$$

6-Experimental results

The performance for each individual channel in this experiment was given in table 4. Each channel's performance was based on Frequency domain features chosen by k-mean and fed to Random Subspace Ensemble (RS). High accuracy gave by all channels which is 99% and 100% by C4. Table 5 explain contrasting past research with our approach, figure 6 show Comparison with state-of-the-art researches.

Table 4. Performance of each channel based on features selected by k-mean was fed to Random Subspace Ensemble

Channel label	ACC	SEN	PREC	F_M	SPE
Fp1	99.48	0.99	1	0.99	1
Fp2	99.83	1	0.99	0.99	0.99
F3	99.80	1	1	1	1
F4	99.20	1	0.98	0.99	0.98
F7	99.68	0.99	0.99	0.99	0.99
F8	99.78	0.99	0.99	0.99	0.99
FZ	99.16	0.99	0.99	0.99	0.99
C3	99.68	1	0.99	0.99	0.99
C4	100	1	1	1	1
CZ	99.04	0.99	0.99	0.99	0.99
T3	99.90	1	0.99	0.99	0.99
T5	99.50	0.99	0.98	0.99	0.98
T4	99.80	0.99	1	0.99	1
T6	99.91	1	1	1	1
P3	99.91	0.99	1	0.99	0.99
P4	99.72	0.99	1	0.99	0.99
PZ	99.67	1	0.99	0.98	0.99
O1	99.50	0.99	0.99	0.98	0.99
O2	99.92	0.99	0.99	0.99	1

Table 5. Contrasting past research with our approach.

Authers	Classifier	Feature	ACC
Merve Açıkoğlu[16]	SVM KNN	Time-Domain	%97,5 %98,8
Dmitry Yu. Isaev[20]	CNN SVM	Time-domain, Frequency domain	%97 %92
KaroliinaT Tapani[22]	SVM	Time-domain, Frequency domain	%98
Proposed method	RS	Frequency domain	%100

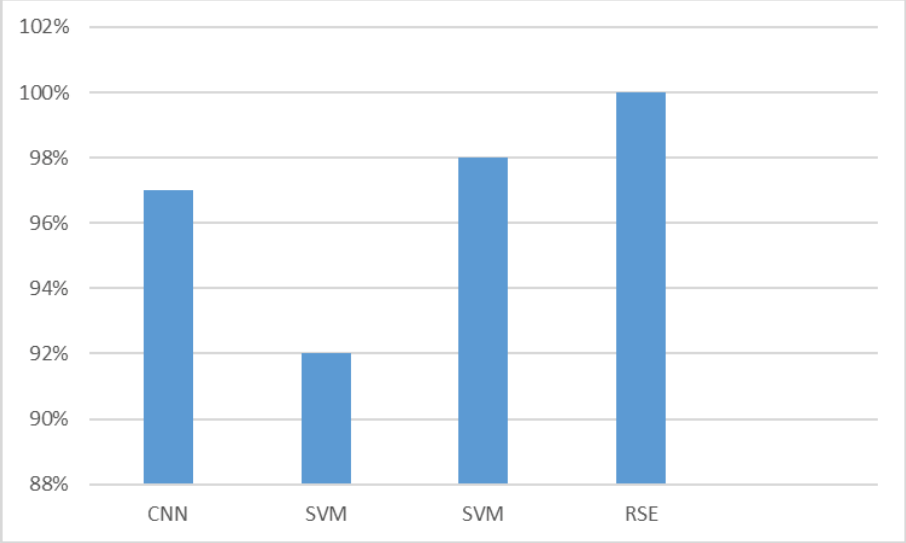

Fig 6. Comparison with state-of-the-art researches

7-Discussion

In our study, reducing the data dimensions was the only necessity for feature extraction with statistical features, which effectively increased the algorithm's performance. The developed ML architecture used RS to perform classification with the test dataset, and it received an F1 score of 100%. The outcomes were quite encouraging, despite the fact that we were unable to compare our procedure to that of other patients using the same data (since conformational labeling had not yet been carried out). A small percentage of cases were not recognized, according to the detection algorithm's sensitivity and

specificity, in particular. The goal of this form of labeling was to increase the algorithm's sensitivity. The EEG report did not indicate any significant differences in the channels to train the model to recognize seizures, even for those designated as seizure, therefore K-fold cross validation was employed to eliminate the imbalance among the datasets to achieve optimal performance [32]. On the basis of the electrode's phase variations, the EEG signals and attributes were found to occur in the channel. [33] after investigating the entire channel (as shown in Table 1). As a result, the EEG system may not be able to reflect such fluctuations sufficiently for the doctor to recognize them from one another if the distance from the area where the seizure-related neuronal potential is generated is considerable enough. The magnitude of this restriction is determined by the magnetic and volume conduction characteristics of the electrical (neuronal) signal, implying that the strength of the distinctive signal attribute could only be partially implicit in reality. In ML techniques, such as feature selection using the k-mean or RS Algorithm, such latent features are not insignificant or trivial. Thus, we made the assumption that a data augmentation effect might occur, such as the extraction of statistical features, the addition of noise to the signal, or the flipping of the signal from the data's point of view [34, 35]. The generalized efficiency of the test data as well as the training data would have been improved by the operational performance of our model. As an alternative, we looked at the potential for data augmentation, which can happen when extracting statistical characteristics, contaminating the signal with undesirable noise, or deviating it from a data-centric perspective [35,36]. The operational performance of our model had to improve both the overall performance of the test data and the training data. Furthermore, we showed that seizure annotations inside other channels can be used to classify EEG signals by using an ML model trained with seizure- and non-seizure labeled data of a chosen channel. Therefore, under such cases, it is not necessary to change the design for training stage. These results indicate that any channel might be utilized to monitor a neonate's neural activity for seizure episodes over extended periods of time using our ML architecture. Our single EEG seizure diagnosis methodology may be used to accurately recognize and medicate newborn seizures when the 10-20 EEG system is difficult to directly employ or continuous brain function monitoring is not practical. By immediately detecting segments with probable anomalies that need additional attention, it may also help to enhance neurological prognosis and shorten the time spent examining the EEG system. This allowed us to include multichannel data into our single-channel seizure detection system, enabling a qualified expert to examine these segments. Because our study was experimental, it was possible to construct a single EEG channel seizure detection method using only a little amount of data that could be utilized to carefully detect and distinguish the labels. It is crucial to apply it to enormous volumes of internal and external data from various contexts in order to validate and enhance the general seizure detection performance while taking maturational changes, pathological

processes, and drug effects into consideration. Future deep learning models can be created by pre-processing EEG data that has been collected for roughly 15 years at the Severance Hospital. In particular, age stratification during pre-processing enables the creation of specialized ML models that take developmental stages changes into account. EEG data can be associated with specific medicines, and machine learning using particular EEG data will produce better performance in circumstances where drugs are present. Additionally, The ML model ensemble can enhance the general seizure detection models' efficiency. These techniques may be used together to create an algorithm that is more precise and adaptable.

8- Conclusion

In this study, we provide a machine learning-based seizure detection system, with the RS algorithm showing the most promise for practical clinical use in the identification of neonatal seizures. Additionally, it enhanced classification performance by utilizing a significant amount of both internal and external hospital data. We also considered how high-risk patients whose brain activity has to be routinely monitored for anomalies like seizures may utilize this algorithm in primary care and private settings, such as the patient's home.

9- Future work

During the period of work on this study, it is found many ideas which represent the future work. The study suggests the following points:

- Using the proposed system with other methods of feature extraction.
- Using the proposed system with other methods to reduce the dimensions of the dataset.
- Trying the proposed system with other classifiers.

REFERENCES

[1] **Rosen JB, Hamerman E, Sitcoske M, Glowa JR, Schulkin JJBN(1996) Hyperexcitability**: exaggerated fear-potentiated startle produced by partial amygdala kindling. 110:43

[2] **Buzsaki G, Ponomareff G, Bayardo F, Ruiz R (1989) Gage FJN**.Neuronal activity in the subcortically denervated hippocampus: achronic model for epilepsy. 28:527–538

[3] **Boylan G, Burgoyne L, Moore C, O'Flaherty B (2010) RennieJJAp**. An international survey of EEG use in the neonatal intensive care unit. 99:1150–1155

[4] **Glass HC, Glidden D, Jeremy RJ, Barkovich AJ, Ferriero DM(2009) Miller SPJTJop.** Clinical neonatal seizures are independently associated with outcome in infants at risk for hypoxic-ischemic brain injury. 155:318–32

[5] **Wirrell EC, Armstrong EA, Osman LD, Yager JYJPr (2001)Prolonged** seizures exacerbate perinatal hypoxic-ischemic brain damage. 50: 445

[6] **Malone A, Anthony Ryan C, Fitzgerald A, Burgoyne L, Connolly S(2009) Boylan GBJE.** Interobserver agreement in neonatal seizure identification. 50:2097–2101

[7] **Azzopardi D, Strohm B, Edwards AD, Halliday H, Juszczak E,Levene M, Thoresen M, Whitelaw A (2009) Brocklehurst PJAoDiC-F, Edition N.** Treatment of asphyxiated newborns with moderate hypothermia in routine clinical practice: how cooling is managed in the UK outside a clinical trial. 94: F260–F264

[8] **Toet MC (2009) Lemmers PMJEhd.** Brain monitoring in neonates.85:77–84

[9] **Rennie J, Chorley G, Boylan G, Pressler R, Nguyen Y (2004)Hooper RJAoDiC-F, Edition N.** Non-expert use of the cerebral function monitor for neonatal seizure detection. 89:F37–F40

[10] **Shellhaas RA, Soaita AI (2007) Clancy RRJP.** Sensitivity of amplitude-integrated electroencephalography for neonatal seizure detection. 120:770–777

[11] **Evans E, Koh S, Lerner J, Sankar R (2010) Garg MJAoDiC-F,Edition N.** Accuracy of amplitude integrated EEG in a neonatal cohort. 95:F169–F173

[12] **Boubchir L, Daachi B, Pangracious V (2017)** A review of feature extraction for EEG epileptic seizure detection and classification.2017 40th International Conference on Telecommunications and Signal Processing (TSP). IEEE, pp 456-460

[13] **Ansari AH, Cherian PJ, Caicedo A, Naulaers G, De Vos M, VanHuffel SJIjons (2019)** Neonatal seizure detection using deep convolutional neural networks. 29: 1850011

[14] **Lu Y, Ma Y, Chen C, Wang YJT (2018) Care H.** Classification of single-channel EEG signals for epileptic seizures detection based on hybrid features. 26:337–346

[15] **M. Açıkoğlu and S. A. Tuncer,** "Incorporating feature selection methods into a machine learning-based neonatal seizure diagnosis," Med. Hypotheses, vol. 135, p. 109464, 2020.

[16] **J. Birjandtalab, M. B. Pouyan, D. Cogan, M. Nourani, and J. Harvey**, "Automated seizure detection using limited-channel EEG and non-linear dimension reduction," Comput. Biol. Med., vol. 82, pp. 49–58, 2017.

[17] B. **Boashash and S. Ouelha,** "Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study," Knowledge-Based Syst., vol. 106, pp. 38–50, 2016.

[18] **M. K. Siddiqui, X. Huang, R. Morales-Menendez, N. Hussain, and K. Khatoon,** "Machine learning based novel cost-sensitive seizure detection classifier for imbalanced EEG data sets," Int. J. Interact. Des. Manuf., vol. 14, no. 4, pp. 1491–1509, 2020.

[19**] D. Y. Isaev et al.,** "Attention-based network for weak labels in neonatal seizure detection," Proc. Mach. Learn. Res., vol. 126, p. 479, 2020.

[20**] N. Koolen et al.,** "Line length as a robust method to detect high-activity events: automated burst detection in premature EEG recordings," Clin. Neurophysiol., vol. 125, no. 10, pp. 1985–1994, 2014.

[21] K. **T. Tapani, S. Vanhatalo, and N. J. Stevenson,** "Time-varying EEG correlations improve automated neonatal seizure detection," Int. J. Neural Syst., vol. 29, no. 04, p. 1850030, 2019.

[22] S. **Ryu et al.,** "Pilot study of a single-channel EEG seizure detection algorithm using machine learning," Child's Nerv. Syst., vol. 37, no. 7, pp. 2239–2244, 2021.

[23] Stevenson **N, Tapani K, Lauronen L, Vanhatalo SJSd (2019)** A dataset of neonatal EEG recordings with seizure annotations. 6: 190039

[24] **Smith, J. R., Karacan, I., and Yang, M.,** Automated EEG analysis with microcomputers. Sleep 1(4):435–443, 1979.

[25] **Quyen, M. L. V., Martinerie, J., Baulac, M., and Varela, F.,** Anticipating epileptic seizures in real time by a non-linear analysis of similarity between EEG recordings. Neuroreport 10:2149–215, 1999

[26] **Hjorth, B.**, Time domain descriptors and their relation to a particular model for generation of EEG activity. In: CEAN – Computerized EEG analysis, Stuttgart: Gustav Fischer Verlag. pp. 3–8, 1975

[27] **Kumar KR, Iyapparaja M, Niveditha VR, Magesh S, Magesh G, Marappan S.** Monitoring and analysis of the recovery rate of Covid-19 positive cases to prevent dangerous stage using IoT and sensors. Int J Pervas Comput Commun. 2020.

[28] **T. K. Ho,** "The random subspace method for constructing decision forests," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 8, pp. 832–844, 1998

[29] **Esen H, Ozgen F, Esen M, Sengur A.** Modelling of a new solar air heater through least-squares support vector machines. Exp Syst Appl. 2009;36(7):10673-10682.

[30] **M. Xu, D.P. Papageorgiou, S.Z. Abidi, M. Dao, H. Zhao, G.E. Karniadakis**, A deep convolutional neural net- work for classification of red blood cells in sickle cell anemia. PLOS Computational Biology, 13 (10):1–27. 2017.

[31] S. **Toraman, M. Girgin, B. Ustundag, I. Turkoglu**, Classification Of The Likelihood Of Colon Cancer With Machine Learning Techniques Using FTIR Signals Obtained From Plasma. Turk J Elec Eng & Comp Sci. 2018.

[32] **S.A. Tuncer, B. Akilotu, S. Toraman**, A deep learning-based decision support system for diagnosis of OSAS using PTT signals. Medical Hypothesis, Vol. 127, pp: 15-22,2019

[33] **Rahman MM, Davis DNJIJoML, Computing (2013)** Addressing the class imbalance problem in medical datasets. 3: 224

[34] **Lachaux JP, Rodriguez E, Martinerie J (1999) Varela FJJHbm**. Measuring phase synchrony in brain signals. 8:194–208

[35] **Van Dyk DA (2001) Meng X-LJJoC,** Statistics G.The art of data augmentation.10:1–50

[36] **Cubuk ED, Zoph B, Shlens J, Le QV (2020)** Randaugment: practical automated data augmentation with a reduced search space. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp 702-703.