Machine Learning and Stereoelectroencephalographic Feature Extraction for Brain Tissue Classification

Pedro Henrique Peres Morais Lopes [*]
Mariana Mulinari Pinheiro Machado [*] **Alina Voda** [*]
Gildas Besançon [*] **Philippe Kahane** [**] **Olivier David** [***]

[*] *Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France (e-mail:{Pedro-Henrique.Peres-Morais-Lopes, Mariana.Mulinari-Pinheiro-Machado,alina.voda,gildas.besancon}@grenoble-inp.fr)*
[**] *Univ. Grenoble Alpes, CHU Grenoble Alpes, Grenoble Institut des Neurosciences, GIN, 38000 Grenoble, France (e-mail: philippe.kahane@univ-grenoble-alpes.fr)*
[***] *Univ. Grenoble Alpes, Grenoble Institut des Neurosciences, GIN, 38000 Grenoble, France; Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France (e-mail: olivier.david@inserm.fr)*

Abstract: Tissue classification of white or gray matter is a necessary information in the study of brain connectivity. Currently this classification is made by the coregistration of the implanted electrodes in the Magnetic Resonance Imaging (MRI) of the patient. This process is complex and therefore is not always carried out, and is limited by the image resolution and by the accuracy of the coregistration. This paper studies the performance of machine learning (ML) algorithms used with features extracted from Stereo-Electroencephalogram (SEEG) signals recorded from three epileptic patients, for electrode contact classification, to serve as a decision support for specialists and researchers. The features are based on epileptic detection, and are extracted from both time and frequency domain. Accuracy, Area Under Curve and F_1-Score are evaluated for each ML algorithm, and feature importance is assessed by feature permutation. Satisfactory results were achieved, with a maximum of 79% accuracy in group separation for patient specific classification, and 74% in inter-patient classification, indicating high potential in ML techniques for brain tissue classification.

Keywords: Brain tissue classification, Stereo-Electroencephalogram (SEEG) feature extraction, Machine Learning, Neural Networks, Decision support system.

1. INTRODUCTION

Stereoencephalography (SEEG) investigation is a common procedure performed in pharmacoresistant epileptic patients. These SEEG recordings done via electrodes inserted in the brain are a fundamental tool for epilepsy research to determine the cortical zones responsible for epileptic seizures before the resective surgery.

In order to better interpret these recordings, it is important to know exactly the type of matter (gray or white) in which the electrodes are inserted in. Brain tissue is usually classified by the coregistration of the implanted electrodes in the Magnetic Resonance Imaging (MRI) of the patient (as in Wang et al. (2013)). This process is complex, therefore it is not always carried out. Even for the cases in which the coregistration process is done, it can be biased by the positioning of the electrodes in the image and the image resolution and coregistration errors.

To the best of our knowledge, there does not seem to exist a lot of research regarding brain tissue classification with Machine Learning (ML) algorithms using features extracted from SEEG recordings. Mercier et al. (2017) commented on the differences in signals recorded in grey and white matter in order to study different signal referencing techniques, but they did not go as far as studying tissue classification. Pinheiro Machado et al. (2021), proposed a method for tissue classification using the SEEG signals from electrode pairs, based on non parametric identification techniques. They presented a brief study on single electrode tissue classification that achieved around 60% accuracy only. Our goal in this article is to provide a more in depth study of single electrode brain tissue classification, relying on features commonly used in epilepsy detection, as well as ML techniques instead of classical identification, in order to improve classification results.

Artificial Intelligence techniques are increasingly being used to handle brain signals, since they have high-dimensional spatial-temporal features, that are difficult to

process with conventional statistical methods. In epilepsy research, these techniques are used to detect and predict epileptic seizures, or even detect the patient's seizure onset zone. Shoeb et al. (2009) used spectral energy to automatically detect seizure activity using support vector machine (SVM). Fergus et al. (2015) compared the performance of different seizure detection ML methods, and found higher accuracy using K-Nearest-Neighbours (KNN) and SVM, and indicated the strong discriminant potential of median frequency and root-means-squares (RMS) of EEG bands. Grinenko et al. (2017) detected the epileptogenic zone, differentiating between time frequency patterns of the seizure onset zone and its propagation to its surroundings with SVM. More advanced deep learning algorithms also prove to be efficient in the epilepsy studies, like Akut (2019) who used Convolutional Neural Networks on wavelet transform coefficients for seizure detection.

In this work, we study the performance of supervised ML algorithms on the classification of electrode contact on white or grey matter, using features extracted from the baseline SEEG recordings in both time and frequency domains. Performance of each algorithm is compared based on accuracy (ACC), Area Under Curve (AUC) and F_1-Score. Feature importance is assessed via feature permutation, giving information of valuable characteristics from SEEG recordings for this classification problem.

This paper is organized as follows : the data acquisition process is presented in section 2, followed by the feature extraction methods, in section 3. Section 4 gives a brief description of the ML algorithms used, section 5 shows the evaluation metrics, classification procedure and feature importance assessment method. Classification results are shown in section 6, and finally section 7 concludes this paper.

2. EXPERIMENTAL SETUP AND DATA

The SEEG signals were recorded using depth penetration electrodes, with a distant point in white matter as reference, and a sampling frequency of $f_s = 1000Hz$. A protocol for data re-use for research purposes was validated by the Inserm Ethical Committee (INSERM IRB 14-18). Electrodes were implanted one week prior to the recording sessions, and were labeled by experts as having contact with white or grey matter based on the MRI of the patient as described in Deman et al. (2018). Recording setup can be seen in figure 1.

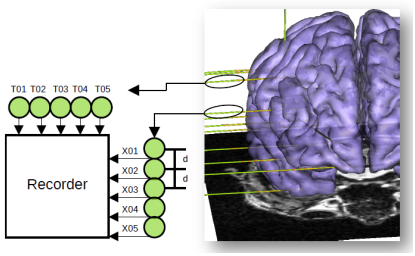

Fig. 1. SEEG Recording Experiment setup

Recordings of three different patients are used in this work. Table 1 shows the number of electrodes connected to brain matter and the number of recording sessions for each of them.

Table 1. Available data from each patient

	N° of Electrodes	N° of Recording Sessions
Patient 1	171	150
Patient 2	146	51
Patient 3	181	152

A total of 27968 signals measured by 498 different electrodes from 353 different SEEG recordings sessions were available. Since the duration of recordings varied for patients and sessions, we chose to keep the first 4.044 seconds of the signal, ensuring that every signal corresponded to recordings at rest. A sampling rate of 1000 samples per second was used, resulting in signals of size $N = 4044$. Out of the available signals, 2083 were classified as "bad channels" by experts. These may consist of non-neuronal signals from disconnected electrodes, the malfunctioning of sensors or parasitic electrical activity (Tuyisenge et al. (2018)). Features were not extracted from these channels. Missing features were replaced by their median values over all recording sessions, for the corresponding electrode, of each patient.

3. FEATURE EXTRACTION

The features used for the classification problem were extracted directly from the raw SEEG signal, and are the same commonly used in epilepsy research. Boonyakitanont et al. (2020) provides a range of possible features to be used, in time, frequency, and time-frequency domains. A total of 23 of these features were used, extracted in only time and frequency domains. Feature extraction is made on each of the recording sessions of each patient, excluding the ones classified as bad channels. To better use the available data, the median value of each predictor over all recording sessions for a patient will be assigned to the corresponding electrode. In this way the impact of possible outliers is reduced.

3.1 Time Domain Features

The first set of features consists of statistical parameters, like the mean, mean absolute value (MAV), variance, skewness (third moment order describing data asymmetry) and kurtosis (fourth moment order describing tailedness of the distribution). First and third quartiles (Q_1 and Q_3) are also used, quantifying the data denseness, as well as the interquartile range, defined as:

$$IQR = Q_1 - Q_3. \tag{1}$$

Coefficient of variation (CV) was also used, explaining the dispersion of data in relation to its mean:

$$CV = \frac{\sigma}{\mu} \tag{2}$$

with σ being the standard deviation and μ the mean value of the SEEG recordings. Energy, average power, and root means square (RMS) power of the entire baseline signal were computed and used as features. Nonlinear

energy (NE) can expand the concept of energy, including indefinite terms of shifted and lagged sequences:

$$NE = \sum_{i=1}^{N-2} x^2(i) - x(i+1) \cdot x(i-1) \tag{3}$$

where x is the SEEG signal and N the number of samples in this signal. Line length (L) was also calculated, as the sum of differences between subsequent samples:

$$L = \sum_{i=1}^{N-1} |x(i) - x(i-1)| \tag{4}$$

Boonyakitanont et al. (2020) lists numerous different entropy measurements. In this work, only the Shannon entropy is used. It reflects the uncertainty in random processes, and is defined as:

$$ShEn = -\sum_{i} p_i log(p_i) \tag{5}$$

where p_i is the probability of occurrence of each value in the SEEG recording. Hurst Exponent (H) is an index of long-term memory in time series. It quantifies the tendency of a time series to regress to a long term mean. For self-similar time-series, H is also related to fractal dimension. Its estimation was made by the linear regression of the logarithmic plot of variance of detail versus level of wavelet decomposition of the signal, as proposed by Flandrin (1992).

Finally, the number of Zero Crossings (ZC) and number of Slope Sign Changes (SSC) were computed. These were shown by Sharmila and Geethanjali (2018) to have good performance on epilepsy seizure detection when coupled with ML techniques.

3.2 Frequency Domain Features

Spectral density of the SEEG baseline was estimated using a modified periodogram, as in (6):

$$\hat{P}(f) = \frac{T}{N} \left| \sum_{n=0}^{N-1} h_n x_n e^{-j2\pi f T n} \right|^2 \tag{6}$$

where $\hat{P}(f)$ is the spectral power estimate, $T = 0.001$ the sampling period, and h_n a window function. Here, a Hamming window of size N was used. Only five frequency-domain features were used: the RMS power of the δ (0.5 - 4 Hz), θ (4 - 8 Hz), α (8 - 13 Hz), β (13 - 30 Hz) and γ (30 - 80 Hz) frequency bands. These bands are very commonly used when analysing EEG signals (Frauscher et al. (2018)).

4. MACHINE LEARNING ALGORITHMS

This section presents a brief description all of the machine learning algorithms used in this work. Parameter setting was based on the highest accuracy result in a small subset of 30% of the available data.

4.1 Decision Tree

Decision tree algorithm starts off with a root node, containing all the data. Successive binary partitions are performed, until a terminal node is reached, where there are 10 or fewer samples. The most frequent class in a terminal node is the one that will be assigned to the new sample. These partitions are based on impurities criteria, like deviance, entropy or the Gini Index (Hastie et al. (2009)). In this work, the last one was selected.

4.2 Random Forest

Bagging (Bootstrap Aggregation) predictors is a method of generating multiple versions predictors to get an aggregated predictor. When predicting a class, the aggregation does a plurality vote, where the new sample is assigned to the most voted class. This technique can reduce the variance of a single classifier, and improve prediction accuracy (Hastie et al. (2009)). The Random Forest algorithm consists in bagging n random decisions trees, thus creating a forest. Each of these trees gives a classification, and the forest chooses the classification having the largest number of votes over all the bagged trees. For this work, a number of $n = 150$ trees was used, since this gave the best accuracy result on the used subset.

4.3 K-Nearest Neighbours

The principle behind the K-Nearest Neighbour method is to find the K closest in distance training samples to the sample to be labeled. It is possible to choose any distance metric, the euclidean distance being the most common one. The label most present in the K closest neighbours will be the one assigned to the new sample (Hastie et al. (2009)). For this study, euclidean distance was used, with a number of neighbours $K = 11$.

4.4 Support Vector Machine

When working with binary classes, a good approach is to use Support Vector Machines (SVM). This consists of constructing hyperplane decision boundaries that try to separate the data as well as possible. SVMs are capable of producing nonlinear boundaries by constructing a linear boundary in a large, transformed version of the feature space. Although this seems computationally expensive, it is possible to represent the transformed feature vectors $h(x_i)$, involving input features via inner products $\langle h(x), h(x') \rangle$. A kernel function can be specified to represent this inner product $K(x, x') = \langle h(x), h(x') \rangle$(Hastie et al. (2009)). Three different kernel functions were tested: polynomial, radial-basis and linear. The last one was kept since it provided the best results.

4.5 Naive Bayes

Naive Bayes are probabilistic classifiers that applies Bayes theorem, with the assumption that features are independent given a class. Its decision rule assigns to the new sample a class with the maximum a posteriori likelihood. Although this assumption is usually not true, it drastically simplifies the estimation, and it often outperforms other alternatives (Hastie et al. (2009)).

4.6 Artificial Neural Network

Artificial neural networks are composed of a single input layer, with the same number of nodes as the number

of features, a single output layer, and any number of intermediary layers, often called hidden layers. Each of these layers consists of one or more neurons. Fitting the network to our data corresponds to adjusting each node's weights to minimize a loss function, with an optimization algorithm, the most popular one being the stochastic gradient descent (SGD). In this paper, the training of the network used SGD with momentum, minimizing the misclassification rate.

The network used here consists of a single fully connected hidden layer, with 32 neurons, as represented in figure 2, where $x_1, x_2 ... x_{23}$, are the extracted features. This structure was achieved via trial and error. Output nodes give the probability of a sample belonging to each class. The class with the highest probability is assigned to the sample.

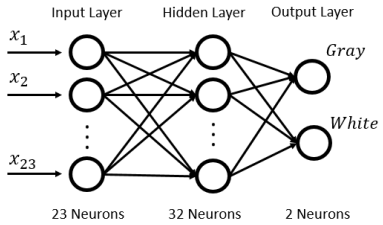

Fig. 2. Neural network representation

5. CLASSIFICATION EVALUATION

Here the evaluation method of the algorithms are described, as well as the evaluation metrics used. A description on the method used to assess feature importance, and the permutation importance, is given.

5.1 Evaluation Metrics

For the performance evaluation of each model, three metrics were used : Accuracy, area under curve (AUC) of the receiver operating characteristics (ROC) and F1-Score. Accuracy is a ratio between correctly predicted observations, both true positives (TP) and true negatives (TN) to the total observations, with FN and FP being false negative and false positive predictions.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

The F-Measure provides a way of combining precision (P) and recall (R), precision being the ration of correctly predicted observations to the total positive predicted observations, and recall the ratio of correctly predicted observations to all of predictions of a certain class:

$$\begin{aligned} P &= \frac{TP}{TP + FP} \\ R &= \frac{TP}{TP + FN} \\ F &= \frac{(\beta^2 + 1)PR}{\beta^2 * P + R} \end{aligned} \tag{8}$$

where β is the relative importance given to recall over precision. F_1-Score is defined with $\beta = 1$ as a harmonic mean between precision and recall (Chinchor (1992)). Values of F_1-Score closer to 1 indicate a good balance of class prediction, meaning that no class will be favored by the algorithm's prediction.

$$F1 = \frac{2PR}{P + R} \tag{9}$$

ROC is a visual representation of the performance of a binary classifier, showing the true positive rate (TPR) versus False Negative Rate (FNR). The AUC of a classifier, calculated as the area under the ROC curve, is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance (Fawcett (2006)).

5.2 Classification Procedure

The classification problem consists of correctly predicting the electrode contact between white or grey matter. The dataset consists of 498 samples from all patients, and 23 available features. This data must be divided into training and validation subsets. After each algorithm learns from the training set, it will assign a class (white or grey) to the samples on the validation set. To verify the influence of the training dataset, and generalisation of these algorithms, Leave-One-Out K-Fold Cross Validation with $K = 4$ was used. This technique consists in randomly splitting the features and labels into K different subsets and using $K-1$ for the training of the model, and the remaining one for validation. This is repeated until every subset was used for both training and validation. Final evaluation consists on the mean of the metrics over K folds. This entire process is then repeated $n = 20$ times, to further verify the underlying performance of the different methods. This procedure is illustrated in figure 3.

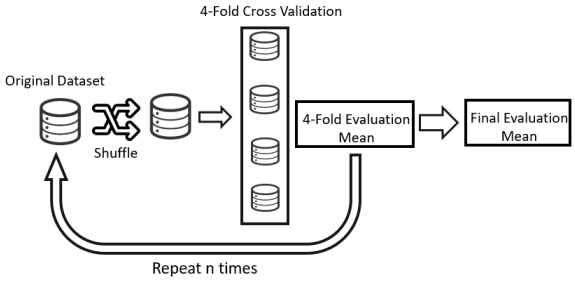

Fig. 3. Classification procedure setup

By the end of this procedure, the means and standard deviations of n repetitions of the cross-validation method of all the evaluation metrics described in the previous subsection are computed.

5.3 Feature Permutation Importance

Feature importance estimation via feature permutation is a model inspection technique that randomly shuffles one feature at a time, and records the overall increase in misclassification rate (1-ACC). Since this directly breaks the relationship between the predictor and the classes, a drop in performance indicates a dependency of the model in the shuffled feature (L.Breiman (2001)). The original

error of the model e_{orig} is stored, and for each permutation a new error is calculated e_{perm}. The difference:

$$FI = e_j^{perm} - e_j^{orig} \tag{10}$$

is the estimate of feature importance (FI) for feature j. In the case of bagged models, such as random forests, the difference between errors is calculated for each individual learner, and the FI in this situation is:

$$FI = \frac{\overline{d_j}}{\sigma_j} \tag{11}$$

with $\overline{d_l}$ as the mean of the differences between permuted and original error over all learners, and σ_j the standard deviation. Features with the highest increase in error are considered the most important. The increase in classification error will be stored for each iteration of the classification procedure, and importance will be analysed visually in section 6.1.

6. RESULTS

Feature extraction and classification were made using a commercial software package (MATLAB 9.9.0 The Mathworks Inc., Natick, MA, R2020b).Patient specific and inter-patient performance were analysed, following the classification procedure as described in subsection 5.2, with a $K = 4$ number of folds. This means that splits of 75% for training and 25% of validation data were used. This consisted of a ratio of 128 channels for training and 43 for validation for Patient 1, 110 and 36 for Patient 2, 134 and 47 for Patient 3, and 374 and 124 for the inter-patient case. Evaluation metrics will be seen in the form of $mean \pm std$.

Table 2. Performance for Patient 1

Algorithm	Accuracy (%)	AUC	F_1-Score
Decision Tree	69.40 ± 3.41	0.71 ± 0.04	0.69 ± 0.04
Random Forest	74.42 ± 2.05	0.81 ± 0.01	0.74 ± 0.03
KNN	71.69 ± 1.60	0.79 ± 0.02	0.72 ± 0.02
SVM	75.12 ± 2.82	0.82 ± 0.02	0.75 ± 0.03
Naive Bayes	67.78 ± 1.84	0.76 ± 0.02	0.68 ± 0.03
Neural Network	71.67 ± 2.84	0.80 ± 0.03	0.72 ± 0.03

Table 3. Performance for Patient 2

Algorithm	Accuracy (%)	AUC	F_1-Score
Decision Tree	69.22 ± 2.21	0.70 ± 0.02	0.69 ± 0.02
Random Forest	75.44 ± 3.02	0.85 ± 0.02	0.75 ± 0.03
KNN	70.31 ± 2.71	0.79 ± 0.02	0.70 ± 0.03
SVM	79.63 ± 2.01	0.84 ± 0.01	0.79 ± 0.02
Naive Bayes	64.70 ± 1.08	0.71 ± 0.02	0.63 ± 0.04
Neural Network	77.42 ± 3.28	0.83 ± 0.02	0.76 ± 0.03

Table 4. Performance for Patient 3

Algorithm	Accuracy(%)	AUC	F_1-Score
Decision Tree	65.44 ± 3.01	0.65 ± 0.03	0.61 ± 0.05
Random Forest	70.31 ± 2.17	0.77 ± 0.02	0.64 ± 0.06
KNN	69.62 ± 2.26	0.75 ± 0.02	0.63 ± 0.05
SVM	69.56 ± 1.69	0.75 ± 0.02	0.64 ± 0.04
NB	65.15 ± 1.00	0.73 ± 0.01	0.64 ± 0.02
Neural Network	69.84 ± 2.43	0.75 ± 0.02	0.58 ± 0.04

Tables 2 through 4 show the patient specific performance of each algorithm. We notice that Random Forests, SVM and Neural Networks outperforms the others in most cases, achieving satisfactory results in all metrics, specially for patients 1 and 2. KNN also displayed the same behaviour, but did not perform as well as the other for the second patient. The overall performance is lower for patient 3. This may indicate that this classification or feature extraction method is patient sensitive.

Overall, SVM had better performance in both accuracy and F_1-score, reaching 79.63% and 0.79 respectively, while Random Forest had better AUC results, reaching up to 0.85. The overall F_1-Score indicates that there's a good balance between precision and recall, so no class is greatly favored by any algorithm prediction.

Table 5 shows the inter-patient results. In this case, Random Forest outperforms every other in the three evaluation metrics. The standard deviations are lower than in the patient specific scenario, likely due to the greater number of samples.

Table 5. Inter-Patient Performance

Algorithm	Accuracy (%)	AUC	F_1-Score
Decision Tree	66.34 ± 2.11	0.68 ± 0.03	0.66 ± 0.03
Random Forest	74.03 ± 1.02	0.81 ± 0.01	0.73 ± 0.02
KNN	69.86 ± 1.12	0.76 ± 0.01	0.69 ± 0.02
SVM	72.03 ± 0.68	0.79 ± 0.01	0.71 ± 0.02
Naive Bayes	56.90 ± 0.81	0.69 ± 0.02	0.55 ± 0.05
Neural Network	69.69 ± 1.70	0.77 ± 0.01	0.66 ± 0.02

Overall, Random Forest, SVM and Neural Network algorithms presented satisfactory results in both inter-patient and patient specific cases. No significant difference was noted in regards to each algorithm execution time. It is important to notice that every algorithm presented low standard deviation for both accuracy and AUC, less than 3% and 0.03 for the most part. This indicates that these algorithms are robust to the training data, and may have the ability to generalise well.

6.1 Feature Importance

In order to find what features contributed the most to brain tissue classification, feature importance was assessed in the inter-patient scenario, with the Random Forest model, using (11). This algorithm was selected because it performed the best in this scenario. This can provide information on features that may be discriminant of the electrode contact tissue in general and not for each specific patient, using this specific algorithm.

Figure 4 shows the 10 features that contributed the most to the classification, with FI in the form of $mean \pm std$ in the y-axis.

RMS power from δ and γ frequency bands are the most important features. Some statistical features such as IQR, VAR and Q_3 also seem to provide valuable information for this classification problem. Energy, average and RMS power from the totality of the baseline signal also displayed this behaviour. The same is observed with the Hurst exponent, indicating that fractal analysis of the SEEG signal can also be of use on electrode contact classification. Standard deviation over the repetitions of the classification procedure is rather high, FI rankings of features shown in figure 4 may vary depending on the number of repetitions.

There is not a great disparity between the FI values of the ten most important features, going from 0.78 of the most

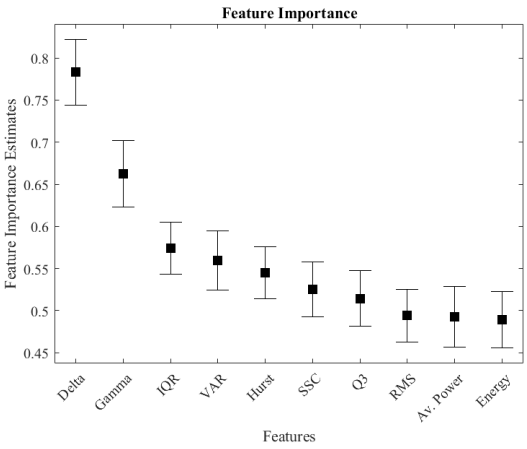

Fig. 4. Feature Permutation Importance

important feature (RMS power from δ band) to 0.48 to the tenth most important (Energy). This means that the contribution of most features is still important to achieve this level of performance. Kurtosis, skewness and RMS power of frequency band β had FI values of around 0.25. Line length proved to be the least important, with a FI value of 0.1. All the remaining features had a FI value of above 0.4.

7. CONCLUSION

In this paper an electrode contact classification method was proposed, that reaches up to 79% accuracy in patient specific and 74% in inter-patient applications, that could be used as a decision support for specialists in the field. This performance was achieved with a rather small dataset of only 3 patients. Further tests with a larger number of patients would be required for validation of the method, and it is expected that the performance would increase.

This method relies solely on features extracted from the baseline of SEEG recordings. Common features used in epilepsy detection proved to have some discriminatory power of cortical tissue as-well. Feature importance analysis showed that δ and γ frequency bands behaviour were the most important features for this classification, showing consistent higher values of FI than other features. Further studies regarding frequency and fractal analysis of SEEG recordings may provide even more discriminant information and increase correct prediction rate.

REFERENCES

Akut, R. (2019). Wavelet based deep learning approach for epilepsy detection. *Health Information Sci Syst.*, 7(8).

Boonyakitanont, P., Lek-Uthai, A., Chomtho, K., and Songsiri, J. (2020). A review of feature extraction and performance evaluation in epileptic seizure detection using eeg. *Biomedical Signal Processing and Control*, 57, 101702.

Chinchor, N. (1992). MUC-4 evaluation metrics. In *Fourth Message Uunderstanding Conference (MUC-4): Proceedings of a Conference Held in McLean, Virginia, June 16-18, 1992*.

Deman, P., Bhattacharjee, M., Tadel, F., Job, A.S., Rivière, D., Cointepas, Y., Kahane, P., and David, O. (2018). Intranat electrodes: A free database and visualization software for intracranial electroencephalographic data processed for case and group studies. *Frontiers in Neuroinformatics*, 12, 40.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8), 861–874. ROC Analysis in Pattern Recognition.

Fergus, P., Hignett, D., Hussain, A., Al-Jumeily, D., and Abdel-Aziz, K. (2015). Automatic epileptic seizure detection using scalp eeg and advanced artificial intelligence techniques. *Biomed Res Int.*, 2015.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2), 910–917.

Frauscher, B., Ellenrieder, N.V., Zelmann, R., Doležalová, I., Minotti, L., Olivier, A., Hall, J., Hoffmann, D., Nguyen, D.K., Kahane, P., and et al. (2018). Atlas of the normal intracranial electroencephalogram: neurophysiological awake activity in different cortical areas. *Brain*, 141(4), 1130–1144.

Grinenko, O., Li, J., Mosher, J.C., Wang, I.Z., Bulacio, J.C., Gonzalez-Martinez, J., Nair, D., Najm, I., Leahy, R.M., Chauvel, P., and et al. (2017). A fingerprint of the epileptogenic zone in human epilepsies. *Brain*, 2017, 117–131.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer, New York, 2nd edition.

L.Breiman (2001). Random forests. *Machine Learning*, 45, 5–32.

Mercier, M.R., Bickel, S., Megevand, P., Groppe, D.M., Schroeder, C.E., Mehta, A.D., and Lado, F.A. (2017). Evaluation of cortical local field potential diffusion in stereotactic electro-encephalography recordings: A glimpse on white matter signal. *NeuroImage*, 147, 219 – 232.

Pinheiro Machado, M.M., Voda, A., Besançon, G., Becq, G., and David, O. (2021). Frequency-domain identification of stereoelectroencephalographic transfer functions for brain tissue classification. 19th IFAC Symposium, SYSID 2021.

Sharmila, A. and Geethanjali, P. (2018). Effect of filtering with time domain features for the detection of epileptic seizure from eeg signals. *Journal of Medical Engineering & Technology*, 42(3), 217–227.

Shoeb, A., Carlson, D., Panken, E., and Denison, T. (2009). A micropower support vector machine based seizure detection architecture for embedded medical devices. *2009 Annual International Conference of the IEEE*, 4202–4205.

Tuyisenge, V., Trebaul, L., Bhattacharjee, M., Chanteloup-Forêt, B., Saubat-Guigui, C., Mîndruţă, I., Rheims, S., Maillard, L., Kahane, P., Taussig, D., and et al. (2018). Automatic bad channel detection in intracranial electroencephalographic recordings using ensemble machine learning. *Clinical Neurophysiology*, 129(3), 548–554.

Wang, P.T., King, C.E., Shaw, S.J., Millett, D.E., Liu, C.Y., Chui, L.A., Nenadic, Z., and Do, A.H. (2013). A co-registration approach for electrocorticogram electrode localization using post-implantation mri and ct of the head. *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*.