

Received 22 April 2025, accepted 18 June 2025, date of publication 26 June 2025, date of current version 3 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3583461



RESEARCH ARTICLE

Identification of Relevant ECG Features for Epileptic Seizure Prediction Using Interpretable Machine Learning

AZRA ABTAHI^{ID}^{1,2,3}, (Member, IEEE), PHILIPPE RYVLIN^{ID}⁴,
AND AMIR AMINIFAR^{ID}³, (Senior Member, IEEE)

¹Department of Computer Science and Media Technology, Malmö University, 205 06 Malmö, Sweden

²Sustainable Digitalisation Research Centre, Malmö University, 205 06 Malmö, Sweden

³Department of Electrical and Information Technology, Lund University, 221 00 Lund, Sweden

⁴Department of Clinical Neurosciences, Lausanne University Hospital, 1005 Lausanne, Switzerland

Corresponding author: Azra Abtahi (azra.abtahi-fahliani@mau.se)

This research has been partially supported by the Autonomous Systems and Software Program (WASP).

ABSTRACT Epileptic seizure prediction holds the potential to enhance the quality of life for individuals with epilepsy by enabling the possibility of timely administration of medication and first aid, as well as preventing subsequent accidents. In this paper, we consider the well-established Heart Rate Variability (HRV) and Lorenz features, and augment them with the electrocardiogram (ECG) multifractality features for the first time for seizure prediction. Our experimental results demonstrate that incorporating multifractality features significantly enhances epileptic seizure prediction, with a 7.5% improvement over using only HRV features and a 6.9% improvement over using both HRV and Lorenz features. We also investigate the significance and impact of features in a seizure prediction Machine Learning (ML) model utilizing ECG signals, aiming to shed light on the intricate relationship between cardiac function and epileptic seizures. We employ SHAP (SHapley Additive exPlanations), an interpretability framework, to interpret the prediction patterns. Based on our analysis, multifractality features are among the most important features in seizure prediction, capturing patterns that are not captured by the HRV and Lorenz features.

INDEX TERMS Electrocardiogram (ECG), epilepsy, interpretability, explainable machine learning, multifractality, seizure prediction, SHAP value.

I. INTRODUCTION

Epilepsy affects more than 50 million people worldwide, according to the World Health Organization (WHO) [1]. It stands as the second neurological cause of years of potential life lost, primarily due to seizure-triggered accidents and sudden unexpected death in epilepsy (SUDEP) [2]. In spite of progress in anti-epileptic drugs, one-third of epilepsy patients still experience seizures, which is classified as pharmacoresistant epilepsy [3]. The prediction of epileptic seizures can improve epileptic patients' lives by enabling timely drug administration, prompt first aid, and the prevention of seizure-related accidents.

The associate editor coordinating the review of this manuscript and approving it for publication was Domenico Rosaci^{ID}.

In this work, we consider epileptic seizure prediction based on the cardiac function and electrocardiogram (ECG) signal. We exploit the well-established Heart Rate Variability (HRV) features that were previously used for seizure prediction [4], [5], [6], alongside ECG Lorenz features mainly used for seizure detection [7], [8], [9], [10], [11]. We further augment this feature set with the multifractality features of the ECG signal [12], [13], [14]. Multifractality denotes the intricate, self-similar patterns inherent in the electrical activity of phenomena, manifesting irregularities across diverse scales. It measures the extent to which the local regularity of a signal varies in time, offering a unique lens to study complex physiological signals. Previous studies have established that the ECG signal embodies multifractal characteristics [12],

suggesting that abnormalities in heart dynamics could potentially be discerned through multifractal analysis of ECG signals [15], [16], [17], [18]. However, ECG multifractality has not been used for epileptic seizure prediction to date. The HRV and Lorenz features, together with the multifractality features, are then fed to a Random Forest classifier [19] to train a seizure prediction model based on the EPILEPSIAE dataset [20], to obtain the seizure prediction model.

Next, we employ a state-of-the-art interpretability framework to investigate the importance and impact of features in ECG-based seizure prediction, for the first time to the best of our knowledge. Interpretability is particularly important in the health/medical domain, where understanding how a Machine Learning (ML) model arrives at its decisions is just as important as the decisions themselves. Our exploration aims to illuminate the intricate relationship between cardiac function and epileptic seizures through the utilization of a metric known as SHAP (SHapley Additive exPlanations) value [21], [22]. Leveraging the SHAP value, we aim to identify the most crucial features and understand the impact of each feature on the results. Specifically, we investigate whether a feature significantly influences the model output as well as whether this feature contributes positively (favoring the positive class, “pre-ictal”) or negatively (favoring the negative class, “inter-ictal”) to the classification outcome. Our study indicates that not only are multifractality features very relevant in the context of seizure prediction, but they also capture patterns that are not captured by the HRV and Lorenz features.

Our main contributions are summarized below:

- We propose an epileptic seizure prediction solution based on the state-of-the-art HRV and Lorenz features, augmented with multifractality features. Multifractality features have not been previously employed in seizure prediction. We show that combining HRV, Lorenz, and multifractality features results in a 7.5% and 6.9% improvement in epileptic seizure prediction compared to using HRV features alone and both HRV and Lorenz features, respectively, when evaluated on the EPILEPSIAE dataset [20].
- We exploit SHAP interpretability framework [21], [22] to identify the most relevant features for epileptic seizure prediction and the impact of features on the prediction outcome. Overall, our results suggest that multifractality features capture patterns that are not captured by the HRV and Lorenz features.

II. STATE OF THE ART

The overwhelming majority of the state-of-the-art studies on epileptic seizure prediction have been focused on using electroencephalogram (EEG) [23], [24], [25], [26], [27]. Among these studies, several integrated ECG with EEG to enhance performance [28], [29], [30], [31], [32], [33], [34], [35], [36], [37]. However, obtaining the EEG signal and working with it poses considerable challenges. This

signal exhibits a high sensitivity to noise, artifacts, and interferences. Furthermore, the full EEG devices are (socially) stigmatizing and inconvenient for the patients to be used in ambulatory settings [38], [39]. The ECG signal, on the other hand, is less sensitive to movements and noises and can be acquired using more convenient wearable devices. Hence, daily seizure prediction based on ECG holds potential for ambulatory monitoring.

In [40], the authors demonstrate changes in ECG HRV features before the onset of epileptic seizures. The ECG data in this work was collected from five patients and the HRV features were calculated. The analysis revealed that frequency HRV features, such as LF and LF/HF, exhibited changes at least one minute before the onset of seizures in all recorded episodes. Another study by [41] investigated the ECG signal in the 5 minutes preceding ictal events in subjects with Frontal Lobe Epilepsy. The authors compared time domain, frequency domain, and non-linear HRV features between normal and epileptic subjects, providing valuable insights into the differences.

By integrating HRV analysis with an anomaly monitoring technique, Fujiwara et al. [4] were the pioneers in introducing an epileptic seizure prediction method based on ECG signal. In [42], linear features in the time and frequency domains of the HRV signal were utilized for the prediction of seizures. In [5], a patient-specific approach to predict seizures using ECG features is proposed. Finally, in [6], a prototype of a wearable system for epileptic seizure prediction was developed based on the HRV features.

Despite the previous studies on seizure prediction based on the ECG signal, the interpretability of the ECG-based seizure prediction models has not been investigated previously. Interpretability has been considered for seizure prediction/detection in several studies, but only considering the EEG signal. In this context, the studies in [43], [44], [45], and [46] proposed partially interpretable Deep Neural Networks (DNN) for epileptic seizure detection based on the EEG signal. In [47], the authors developed a DNN model for seizure detection based on EEG signal and used visualization/interpretation methods to explore how the kernels of the first layer contribute to the final decision and highlight ictal features in the input EEG. In [48], an interpretable algorithm for seizure detection was proposed. This algorithm relies on morphological patterns observed in EEG signals during seizures, providing a personalized approach for the majority of epileptic patients. In [49], an interpretable self-supervised seizure detection framework was proposed to detect seizures based on seizure/non-seizure signatures. Similarly, in [50], an interpretable ML framework for seizure detection based on EEG was developed.

In the realm of seizure prediction, in [51], evolutionary seizure prediction model has been proposed to identify the most effective feature set, while automatically searching for the pre-ictal period. However, the proposed approach in [51] is also based on the EEG signal. Therefore, despite several valuable studies in the state of the art, investigating

interpretability in seizure prediction based on the ECG signal remains unexplored to date.

III. SEIZURE PREDICTION AND INTERPRETABILITY

In this study, we focus on seizure prediction based on the ECG signal. Particularly, we consider the HRV features and Lorenz features, augmented with the multifractality features of the ECG signal. We use Random Forest classifier due to its competitive performance, while requiring minimal hyperparameter tuning—typically only adjusting the number of trees, which already comes with a robust default value. This enables us to delve into analyzing the intricate relationship between cardiac function and epileptic seizures without any major concern about the impact of the hyperparameters on our conclusions.¹ Next, we use the SHAP value and interpret the proposed model to determine the most relevant features in prediction, their impacts on the prediction, and the most contributing features for a new model.

Below, we first delve into a more detailed discussion of the ECG features incorporated into our model for seizure prediction; then, we discuss the SHAP value for the model interpretation.

A. FEATURES FOR THE SEIZURE PREDICTION

Now, we introduce the feature sets exploited in this research.

- **HRV Features:** To extract the HRV features, we consider the RR intervals, i.e., the time series capturing the time duration between consecutive R peaks on an ECG signal, indicating the duration of the cardiac cycle. The HRV features [4], [6] are, then, extracted as follows:

- **Mean_RR:** the mean of the RR intervals;
- **Std_RR:** the standard deviation of the RR intervals;
- **RMSD:** the root mean square of differences of adjacent RR intervals;
- **pNN50:** the number of pairs of adjacent RR intervals with a difference of more than 50 ms divided by the total number of the RR intervals;
- **Tot_pow:** the total power of the RR intervals, i.e., the variance of RR intervals;
- **LF:** the power of the RR intervals in the low-frequency band (0.04-0.15 Hz) normalized by the total power of the RR intervals;
- **HF:** the power of the RR intervals in the high-frequency band (0.15-0.40 Hz) normalized by the total power of the RR intervals;
- **LF_HF:** the ratio of LF to HF.

- **Lorenz Features:** Lorenz plot (or Poincaré plot) illustrates each RR interval time length (I_k) versus the

following RR interval time length (I_{k+1}). The following features are extracted from the Lorenz plot [7], [8]:

- **sd1:** the standard deviations for the transverse direction (vertical to the $I_k = I_{k+1}$ line) that is a measure of short-term variability;
- **sd2:** the standard deviations for the longitudinal direction (parallel to the $I_k = I_{k+1}$ line) that is a measure of long-term variability;
- **Trav_L:** the transverse length of the Lorenz plot approximated as $4 \times sd1$;
- **Long_L:** the longitudinal length of the Lorenz plot approximated as $4 \times sd2$;
- **CSI:** Cardiac Sympathetic Index (CSI) calculated as $\frac{\text{Long_L}}{\text{Trav_L}}$;
- **Mod_CSI:** Modified CSI calculated as $\frac{\text{Long_L}^2}{\text{Trav_L}}$;
- **CVI:** Cardiac Vagal Index (CVI), calculated as $\log_{10}(\text{Trav_L} \times \text{Long_L})$;
- **HR_diff:** Heart Rate differential method defined as $\sum_{k=2}^K (I_{k+1} - I_{k-1})$ where $K = 30$.

These features were previously used for epileptic seizure detection but not for seizure prediction.

- **Multifractality Features:** Multifractality measures the extent to which the local regularity of a signal varies in time. Several studies have demonstrated that the ECG signal exhibits multifractal characteristics [12]. It has also been shown that the multifractality in the ECG signal can be reduced for patients diagnosed with heart diseases [15], [16], [17], [18]. However, ECG multifractality has not been used for epileptic seizure prediction yet.

As multifractality features, we consider the singularity spectrum and the corresponding Hölder exponents of the ECG signals [12], [13], [14].

We define the partition function $Z_q(a)$ as the sum of the q -th powers of the local maxima of the modulus of the wavelet transform coefficients at scale a [12], [53]. For small scales, the partition function exhibits a power-law scaling:

$$Z_q(a) \sim a^{\tau(q)},$$

where $\tau(q)$ is the scaling exponent. The singularity spectrum $D(h)$ and Hölder exponents h are derived from $\tau(q)$ through a Legendre transform [12], [14]:

$$h = \frac{d\tau(q)}{dq}, \quad D(h) = qh - \tau(q).$$

These features provide insight into the multifractal nature of the signal. Hölder Exponent h describes the local regularity or smoothness of the signal at specific points, and singularity Spectrum $D(h)$ quantifies the fractal dimension of the subsets of the signal characterized by the same Hölder exponent h . Although both metrics are related to complexity, they capture different aspects of it. The Hölder exponent reflects pointwise complexity, indicating how abruptly or smoothly the

¹Note that, in this work, we mainly focus on feature-based ML techniques and abstain from utilizing DNNs because our primary aim is to investigate and interpret the impact of the well-established state-of-the-art ECG features on seizure prediction performance, while DNNs are closed-box end-to-end models that extract their own ad-hoc features with limited clinical relevance. The neural network [52] is only included as a baseline for comparison with the state-of-the-art in terms of performance.

signal changes at a given point. In contrast, the singularity spectrum provides a global perspective on complexity, describing how different local regularities are distributed across the entire signal. A broader singularity spectrum suggests greater heterogeneity in signal structure, while a narrower spectrum indicates more uniformity.

By analyzing the singularity spectrum $D(h)$ and the corresponding Hölder exponents, we can capture the overall structural complexity and the localized variations in the ECG signals, which may vary significantly in pathological states.

In this paper, the singularity spectrum and Hölder exponents are estimated for the linearly-spaced moment parameter q within the range -5 to $+5$. Hence, we have the following multifractality features:

- 11 singularity spectrum features (Dh1, Dh2, ..., Dh11), each represents a singularity strength at a particular moment. For ECG signals, a broader range of singularity spectrum values may reflect greater complexity and adaptability in the heart's electrical activity, whereas a narrower range could indicate reduced complexity potentially associated with pathological conditions.
- 11 corresponding Hölder exponents (h1, h2, ..., h11), describing the local regularity of the signal at different moments. As the index increases, the Hölder exponents provide information about increasingly localized behaviors. Higher values of Hölder exponents across the ECG signal indicate smoother and more regular patterns of heart electrical activity, whereas lower values reflect more erratic or irregular electrical behavior.

B. SHAP VALUE FOR MODEL INTERPRETATION

For interpreting an ML model, we consider the well-established framework called SHAP [21], [22]. The concept of SHAP values originates from the following fundamental game theory question: *How should we divide up the payoff among the players with different skills in a coalition?* To answer this question, we look into the expected marginal contribution of each player over all possible coalitions. In practice, given a coalition, the marginal contribution of each player is the difference in payoff as a result of this player joining this coalition. Then, the marginal contributions of each player should be averaged over all possible sets in which the players could have joined. This methodology is now adapted to assess the significance of the features in ML models.

The SHAP value, in this context, represents the average contribution of a feature value to the model's output, considering all possible feature sets in which it could be integrated.

For a given sample \mathbf{x} , the SHAP value for feature i is calculated as [21], [54]:

$$\phi_i(\mathbf{x}) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (f(S \cup \{i\}) - f(S)) \quad (1)$$

where N is the set of all features, S is a subset of features excluding feature i , $|S|$ is the size of the subset S , $|N|$ is the total number of features, $f(S)$ is the model prediction using only the features in subset S , and $f(S \cup \{i\})$ is the model prediction when feature i is added to subset S . Hence, $(f(S \cup \{i\}) - f(S))$ is the marginal contribution of the feature i to the prediction considering the subset of S . The term $\frac{|S|!(|N|-|S|-1)!}{|N|!}$ is the weight given to the marginal contribution, ensuring that the contributions are averaged over all possible subsets of features.

Hence, the SHAP value for a feature is an indication of the importance of that feature. Moreover, the SHAP value provides us with insights into the impact of each feature on the model output. Specifically, we can explore whether and when a feature significantly impacts the output and whether it contributes positively (favoring the positive class) or negatively (favoring the negative class) to the output. In other words, utilizing the SHAP value enables us to potentially pinpoint the feature values that are highly probable to result in a specific classification outcome for a feature with a significant impact on the output. The exact calculation of the SHAP values for large tree-based models, such as large Random Forest, is, however, computationally intractable (it is an NP-hard problem) [55]. Therefore, several efficient algorithms have been proposed to calculate tree-based SHAP values, running in linear time [56].

While SHAP values provide insight about the relevance of each feature, the correlation among the features renders identifying the most relevant features a more challenging problem. To address this challenge, we investigate the contribution of each feature, considering the correlation among the features. Concretely, to identify the most contributing features, we use an iterative feature selection approach based on SHAP values. To this end, an initial model is first trained using all available features. Then, the importance of each feature in contributing to the model's performance is assessed by SHAP value, and based on this evaluation, the least important feature is eliminated from the feature set. This process is performed iteratively, with one feature being eliminated at each iteration until the desired number of features in the set is reached. The proposed scheme allows us to rank features based on their contribution, considering the correlation among the features.

IV. EXPERIMENTAL SETUP

A. DATASET AND DATA PREPARATION

We have used the ECG data of the public EPILEPSIAE dataset [20], which is one of the largest epilepsy dataset manually annotated by medical experts for seizure detection

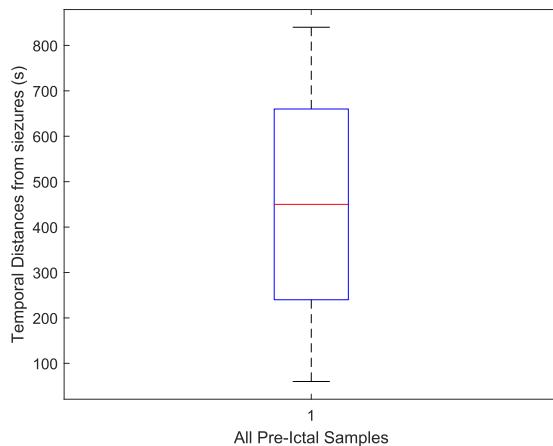


FIGURE 1. Boxplot of temporal distances from seizures.

and prediction. The recordings are conducted in a routine clinical setting. Hence, various non-seizure activities and artifacts such as head/body movement, chewing, blinking, early stages of sleep, and electrode pops/movement may be present. The dataset encompasses Complex Partial (CP), Simple Partial (SP), Secondarily Generalized Seizures (SG), and Unclassified Seizures (UC).

The EPILEPSIAE ECG data is collected from 30 patients comprising 4603 hours of recordings, segmented into one-hour files containing 277 seizures. The recordings are acquired at a sampling rate of 256 Hz with 16-bit resolution. The number of seizures per patient varies between 5 and 23, with an average of 9.23 seizures per patient. The average duration of seizures is 75.81 seconds. Additionally, the total recording duration per patient ranges from 92.90 to 266.36 hours, with an average of 153.43 hours. In this research, we aim to develop an universal model for all subjects.

We pre-process the data to obtain the pre-ictal and inter-ictal signals for the purpose of seizure prediction. For the pre-ictal signal, we first select one-hour recordings containing seizures, in which the seizure onsets occur at least 1 hour and 15 minutes after the previous seizure. From these recordings, we then choose pre-ictal signals from 15 minutes to 1 minute before seizure onset. These pre-ictal signals are further divided into 60-second windows with a 30-second overlap, from which the final pre-ictal windowed signals are selected. Hence, all windowed signals are between 14 and 1 minutes before seizure onset. A time frame of more than 1 minute is deemed sufficient to stop certain activities (such as walking, running, or driving) to reduce seizure-induced accidents or administer drugs to mitigate the impact and severity of seizures. Fig. 1 shows the boxplot of the temporal distances of the pre-ictal windowed samples from seizures, illustrating their distribution.

For the inter-ictal (signals far from seizures), we also take 14 minutes of the one-hour recordings, which are at least 1 hour away from the previous and next seizures. Then,

we window the obtained inter-ictal signals with a 60-second window. We consider 30 seconds overlap in the windowing process.

We employ the Pan-Tompkins algorithm to detect R-peaks in the ECG windowed signals. If the Pan-Tompkins algorithm is only able to detect very few R-peaks in a 1-minute window, this is due to poor signal quality or artifacts. In this work, we adopt a conservative perspective to retain the majority of the data and only exclude the very poor-quality segments with less than 7 R-peaks per minute.

To split the data into training, validation, and test sets, we randomly group all windowed signals derived from each one-hour file into a single set—either training, validation, or test—ensuring that overlapping windows are confined within one set. This approach prevents any overlap across the different sets. Then, we balance the data to have the same number of pre-ictal and inter-ictal windowed signals. Notably, this results in a total of 52,710 pre-ictal samples and 52,710 inter-ictal samples. We use 70% of the balanced windowed signals to train the model, 15% for the validation, and 15% to test the model. We utilize the validation set to compute the SHAP values and generate the SHAP value plots.

B. EVALUATION METRICS

For the evaluation of the model, we employ the following five measures:

- Sensitivity = $\frac{TP}{TP+FN}$,
- Specificity = $\frac{TN}{TN+FP}$,
- Geo-Mean = $\sqrt{\text{Sensitivity} \times \text{Specificity}}$,
- False Alarm Rate = $\frac{FP_{5min}}{TP+TN+FP+FN}$,
- Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$,

where the pre-ictal data samples are regarded as P (Positive), and the inter-ictal samples are regarded as N (Negative). TP is the number of correctly classified pre-ictal samples. TN is the number of correctly classified inter-ictal samples. FP and FN denote the number of false positives and false negatives, respectively. FP_{5min} is defined as the number of false classifications of inter-ictal samples when there are no samples classified as pre-ictal within a 5-minute time window to these inter-ictal samples. We consider Geo-Mean here, as the geometric mean is the only correct average of normalized measurements [58]. A high Geo-Mean value reflects that both Specificity and Sensitivity are high, indicating a high-quality prediction. Conversely, a low Geo-Mean reflects low values for Specificity, Sensitivity, or both, which is undesirable.

C. IMPLEMENTATION DETAILS

In this study, we trained, tested, and interpreted our prediction model in Python, leveraging the SHAP package for interpretation. All experiments conducted in this study, except for the real-time implementation (see Section V-D), were performed on a system with an 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz, 2304 Mhz, 8 Core(s), 16 Logical Processor(s), and a physical memory (RAM) capacity of 16.0 GB.

TABLE 1. Specificity, Sensitivity, Geo-Mean, False Alarm Rate, and Accuracy for different models.

Features	Specificity (%)	Sensitivity (%)	Geo-Mean (%)	False Alarm Rate (%)	Accuracy (%)
Multifractality	81.0 ± 7.2	76.0 ± 5.0	78.3 ± 4.0	2.6 ± 0.4	78.5 ± 3.9
Lorenz [7], [8]	69.4 ± 2.9	56.7 ± 4.2	62.7 ± 2.2	4.5 ± 0.3	63.1 ± 2.1
HRV [4], [6]	82.9 ± 7.0	74.1 ± 5.1	78.2 ± 4.2	2.5 ± 0.5	78.5 ± 4.1
Extended HRV [4], [6], [28]	82.6 ± 4.1	75.1 ± 5.0	78.7 ± 3.3	2.7 ± 0.4	78.9 ± 3.2
AnEp and NRRi [57]	77.7 ± 5.3	64.6 ± 5.1	70.7 ± 3.5	3.3 ± 0.4	71.1 ± 3.4
HRV and Lorenz	83.2 ± 7.9	74.8 ± 5.0	78.8 ± 4.2	2.5 ± 0.5	79.0 ± 4.1
Lorenz and Multifractality	85.8 ± 7.0	83.6 ± 4.1	84.6 ± 3.8	2.1 ± 0.4	84.7 ± 3.6
HRV and Multifractality	86.7 ± 7.0	84.4 ± 4.0	85.3 ± 3.8	2.0 ± 0.4	85.5 ± 3.6
Proposed Model: HRV, Lorenz, and Multifractality	87.1 ± 7.0	84.6 ± 4.1	85.7 ± 3.9	1.9 ± 0.4	85.9 ± 3.7
Neural Networks Baseline [52]	85.9 ± 11.0	88.9 ± 6.9	87.0 ± 7.2	1.1 ± 0.9	87.5 ± 6.6

We repeat each experiment 100 times to obtain Sensitivity, Specificity, Geo-Mean, False Alarm Rate, Accuracy, and SHAP values, and report the mean and standard deviation values.

V. EVALUATION

A. PERFORMANCE EVALUATION OF SEIZURE PREDICTION

In this section, we evaluate the performance of our proposed seizure prediction model, which integrates multifractality features with state-of-the-art HRV [4], [6] and Lorenz [7], [8] features. We compare our approach against the trained Random Forest model using the state-of-the-art features: (i) HRV [4], [6], (ii) extended HRV [4], [6], [28], (iii) Lorenz [7], [8], (iv) AnEp and NRRi [57], and (v) both HRV and Lorenz. We also compare our approach against the state-of-the-art neural networks [52].

Table 1 presents the results of these evaluations. As it can be seen, using only the proposed multifractality features, we can predict the seizures with a Geo-Mean of 78.3%. More importantly, incorporating the multifractality features with HRV and Lorenz features in our model leads to a Geo-Mean of 85.7%, which is 15%, 7.5%, and 6.9% higher than the corresponding Geo-Mean values obtained using AnEp and NRRi, HRV features, and both HRV and Lorenz features, respectively.

Previous studies [28] have proposed extending the standard HRV features exploited in [4], [6] by adding 15 more statistical HRV features. However, our results show that this extension does not have a considerable impact on seizure prediction. Therefore, we use only the HRV features from [4], [6] in our proposed model and refer to them simply as HRV features. As previously mentioned, the results are also compared against the state-of-the-art neural network: a Residual 1-Dimensional Convolutional Neural Network (Res1DCNN) proposed for seizure *detection* [52]. Our proposed approach offers interpretability by focusing on clinically relevant features, while neural networks rely on ad-hoc features. Interestingly, while offering interpretability and despite its significantly lower complexity compared to complex state-of-the-art neural networks that use 60-second

TABLE 2. Wilcoxon signed-rank test results showing statistical significance of performance differences between the Proposed model and models with state-of-the-art features.

Features	W Statistic	P-Value ≤ 0.01
HRV [4], [6]	4950.0	Yes
Lorenz [7], [8]	4950.0	Yes
HRV and Lorenz	4950.0	Yes
AnEp and NRRi [57]	4947.0	Yes

TABLE 3. Performance of the proposed model across different seizure types.

	CP	SP	SG	UC
Accuracy (%)	85.4 ± 3.1	87.4 ± 3.5	87.2 ± 3.9	86.8 ± 3.8

ECG samples with lengths of 15,360 as inputs, our model achieves high performance without any major performance degradation. Overall, Table 1 demonstrates that our model delivers high performance in epileptic seizure prediction while offering interpretability.

In Table 2, the results of the *Wilcoxon Signed-Rank Test* [59], [60] are presented to evaluate the statistical significance of performance differences between our model and the trained Random Forest model using the state-of-the-art features. A p-value below the threshold of 0.01 (1% significance level) indicates that the observed differences are unlikely to have occurred by random chance and are therefore considered statistically significant. Table 2, shows that our model demonstrates statistically significant improvements over all comparison models, trained models exploiting: (i) HRV, (ii) Lorenz, (iii) HRV and Lorenz, and (iv) AnEp and NRRi. The comparisons yield p-values less than 0.01, indicating high confidence in our model's superior performance. These results support the robustness and effectiveness of the proposed model, highlighting its ability to outperform its counterparts across all tests performed.

Next, we demonstrate the performance of our proposed model for different epilepsy types: SP, SG, CP and UC.

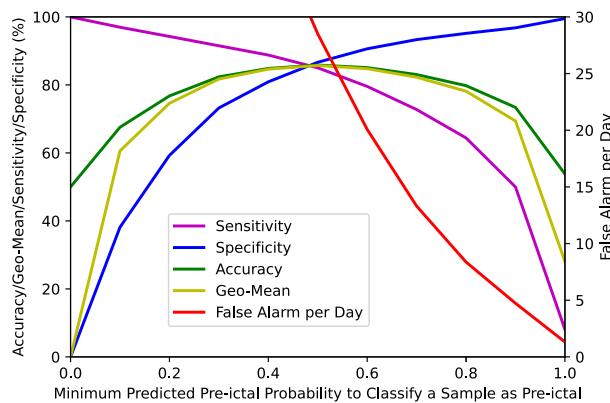


FIGURE 2. Accuracy, Geo-Mean, Sensitivity, Specificity and False Alarm per Day versus different minimum predicted pre-ictal probability to classify a sample as pre-ictal.

Table 3 highlights the Accuracy of the model for each seizure type. As shown, the model demonstrates strong performance, achieving an Accuracy of 85.4% or higher across all four groups.

The Random Forest algorithm classifies a sample based on its predicted probabilities for each class. This algorithm provides predicted class probabilities through two distinct approaches. In the first approach, the predicted class probability in a decision tree is the fraction of samples of the same class in a leaf, and the predicted class probabilities of an input sample in a Random Forest are computed as the mean predicted class probabilities of all trees. In the second approach, the predicted class probabilities can be obtained by considering the fraction of trees that vote for a particular class. All the reported measures in Table 1 are achieved by considering the first approach, where the predicted pre-ictal probability of more than 0.5 results in classifying a sample as pre-ictal.

Fig. 2 shows the Accuracy, Geo-Mean, Sensitivity, Specificity, and False Alarm per Day ($24 \times 60 \times$ False Alarm Rate) for our proposed model versus different minimum predicted pre-ictal probability to classify a sample as pre-ictal. As it can be seen from Fig. 2, if we opt to have equal or less than 5 False Alarms per Day, 0.89 is selected for the minimum predicted pre-ictal probability to classify a sample as pre-ictal. Then, Geo-Mean, Accuracy, Sensitivity, and Specificity will approximately be 67%, 71%, 52%, and 96%, respectively. That is, if we limit the number of False Alarms to equal to or less than 5 per day, we are still able to predict more than half of the seizures (i.e., Sensitivity of 52%).

B. MODEL INTERPRETATION

Next, we investigate the interpretability of our proposed scheme using the SHAP framework [21], [22]. The bar plot of mean absolute SHAP values for our model is shown in Fig. 3. To generate this summary plot, we consider 100 runs with

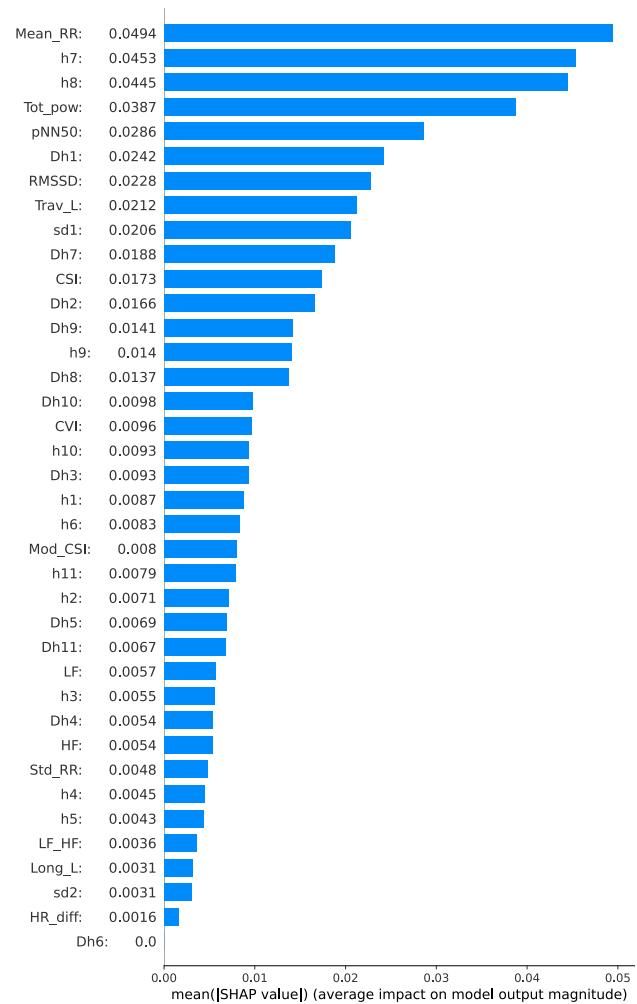


FIGURE 3. Bar plot of mean absolute SHAP values for the proposed model.

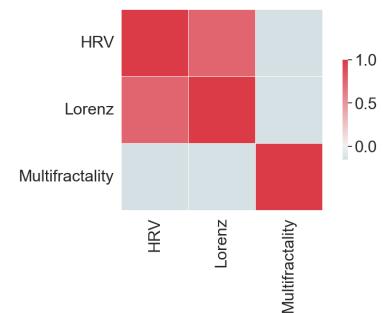


FIGURE 4. Heat-Map depicting the correlation among HRV, Lorenz, and multifractality feature groups.

random train and validation sets. The absolute SHAP values from each run are then collected and averaged. In this plot, the importance of the features decreases from top to bottom. This importance is determined by calculating the mean absolute value of the SHAP values over all the feature values.

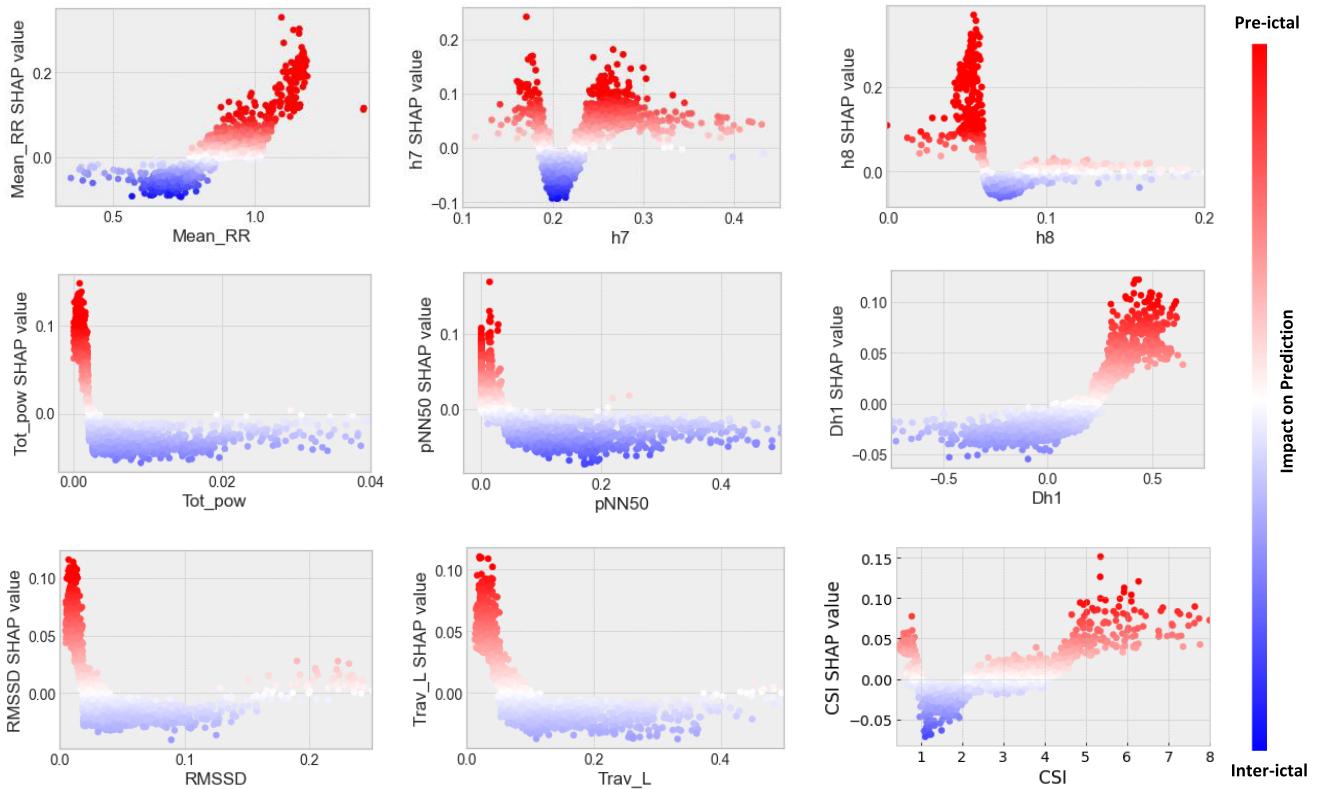


FIGURE 5. SHAP dependence plots of the most important features. Red color indicates the positive SHAP values, which support the pre-ictal classification of the sample, while blue indicates negative SHAP values, favoring the inter-ictal classification of the sample.

According to Fig. 3, “Mean_RR,” “h7,” “h8,” “Tot_pow,” and “pNN50” are respectively the most important features in this model as they are the first features on the list of the summary plots.

For deeper insights into the impact of a single feature on the output of a machine learning model, SHAP dependence plots can be employed. In Fig. 5, SHAP dependence plots are presented for the most important features based on their mean absolute SHAP values. Note that these plots belong to a specific model that we aim to interpret. According to this figure, we notice that for “Mean_RR” values more than 1.05, the SHAP values are consistently positive and also not negligible, implying that samples with “Mean_RR” values more than 1.05 are likely to be classified as pre-ictal samples. To validate this claim, we calculate the probability of being classified as pre-ictal by our model for all samples of the validation set (P_v) and test set (P_t) with “Mean_RR” values more than 1.05. These probabilities are $P_v = 1$ and $P_t = 0.99$, respectively.

Applying the same rationale, we conclude that samples with “h7” values more than 0.27 and less than 0.29 are likely to be classified as pre-ictal samples. The probability of being classified as pre-ictal by our model for these samples for the validation set is $P_v = 0.95$ and for the test set is $P_t = 0.92$. Likewise, we can conclude that the prediction in this specific model is likely to be pre-ictal for samples with at least one of

the following constraints: “h8” values more than 0.045 and less than 0.55 ($P_v = 1$ and $P_t = 0.98$), “Tot_pow” value less than 0.001 ($P_v = 1$ and $P_t = 0.99$), “Dh1” values more than 0.45 ($P_v = 0.92$ and $P_t = 0.98$), “RMSSD” values less than 0.01 ($P_v = 1$ and $P_t = 0.99$), “Trav_L” values less than 0.03 ($P_v = 1$ and $P_t = 0.99$), and “CSI” values more than 6 ($P_v = 1$ and $P_t = 0.88$).

Observing both positive and negative or positive/negative and zero SHAP values within specific feature ranges suggests that relying solely on the corresponding feature with a value in these specific ranges is insufficient for accurate predictions. Furthermore, if the SHAP values of a feature hover around zero for certain feature value intervals, it indicates that within these intervals, the feature has minimal impact on the predicted outcome.

C. FEATURES CONTRIBUTIONS CONSIDERING CORRELATION

Next, we investigate the correlation among the features considered in this work. Fig. 4 presents the Heat-Map for the feature groups, depicting the correlation among HRV, Lorenz, and multifractality feature groups. The correlation between two groups is determined by summing the correlations between each feature from the first group and all feature from the second group, divided by the product of the squares of the sums of the correlations among the features within

TABLE 4. Specificity, Sensitivity, Geo-Mean, False Alarm Rate, and Accuracy for models with 5 selected features based on iterative SHAP, SHAP, RFE, and InFS.

Technique	Features	Specificity (%)	Sensitivity(%)	Geo-Mean (%)	False Alarm Rate (%)	Accuracy(%)
Iterative SHAP	Mean_RR, h7, h8, pNN50 & RMSSD	86.6 ± 6.9	81.7 ± 4.4	84.0 ± 3.9	2.1 ± 0.4	84.1 ± 3.7
SHAP [61]	Mean_RR, h7, h8, Tot_pow & pNN50	85.4 ± 6.7	81.9 ± 3.9	83.5 ± 3.7	2.4 ± 0.4	83.7 ± 3.6
RFE [62]	pNN50, Mean_RR, Dh1, h6 & LF	80.6 ± 7.0	71.0 ± 4.9	75.5 ± 4.0	2.6 ± 0.4	75.8 ± 3.9
InfFS [63]	Mean_RR, Std_RR, Tot_pow, LF & HF	80.3 ± 4.8	70.1 ± 4.9	74.9 ± 3.4	2.9 ± 0.4	75.2 ± 3.3

each group. As observed in this figure, HRV and Lorenz feature groups are highly correlated (their correlation is 0.76), while the multifractality features are not correlated with the other two groups of features. Indeed, multifractality features capture patterns on the ECG signal, while the HRV and Lorenz features capture patterns over the RR-intervals time series.

Finally, here, we assess the contribution of each feature, considering the correlation among the features. This is done by iteratively eliminating the least important feature based on SHAP value from the feature set, to account for the correlation among the features, referred to as Iterative SHAP. Table 4 shows Sensitivity, Specificity, Geo-Mean, False Alarm Rate, and Accuracy for the models with only 5 selected features based on the Iterative SHAP approach, SHAP (i.e., top-5 features based on SHAP values) [61], Recursive Feature Elimination (RFE) [62], and Infinite Feature Selection (InFS), i.e., a graph-based feature filtering approach [63]. The features are listed in descending order of their contributions, i.e., the first feature has the highest contribution. This experiment serves to highlight the contribution of individual features and is not intended to suggest that five features are sufficient for peak performance. As shown, the best-performing approaches have two multifractality features among their most contributing features, highlighting the significant impact of the multifractality features.

D. REAL-TIME IMPLEMENTATION

To assess the feasibility of the proposed approach for real-time implementation, we conducted experiments on a 64 MHz Arm® Cortex®-M4F processor (with an FPU), equipped with 1 MB of Flash memory and 256 kB of RAM. In our implementation, we calculated the time required for feature extraction and model inference. The Random Forest model in our approach, which includes 100 trees and operates on 38 features, achieves an average inference time of 527 μ s. In comparison, the state-of-the-art Random Forest model with 16 HRV and Lorenz features achieves an average inference time of 458 μ s. These times are considerably low and have a minimal impact on the overall inference time. The majority of the computational time in our system is consumed

by the feature extraction process, which is the most time-intensive step.

Based on the experimental results, the average time required to calculate HRV Features, Lorenz Features, and Multifractality Features is 28 ms, 26 ms, and 282 ms, respectively, and the average feature set calculation time for our approach is 328 ms for each 1 min signal (we only need to find the RR intervals once). These results underscore the computational efficiency of the feature groups and show that the feature calculations are feasible for real-time processing, demonstrating their suitability for practical embedded applications.

VI. CONCLUSION

In this paper, we have considered the multifractality features of the ECG signal alongside Lorenz features and typical HRV features in epileptic seizures prediction. We have shown that incorporating the ECG multifractality features with state-of-the-art features results in a notable enhancement in epileptic seizure prediction performance, i.e., 7.5% and 6.9% increase in Geo-Mean compared to using HRV features and both HRV and Lorenz features, respectively. Furthermore, we investigated the significance and impact of each feature employed in the prediction model, aiming to illuminate the intricate relationship between cardiac function and epileptic seizures. To achieve this, we utilized an interpretability framework known as SHAP. According to the experimental results, among the three most significant features, two are multifractality features and one is an HRV feature.

Future research could explore multi-modal frameworks that combine ECG and electrodermal activity features to potentially enhance prediction accuracy in specialized cases and assess the significance and impact of these combined features on epileptic seizure prediction. However, it is essential to carefully weigh the trade-offs in complexity, computational efficiency, and practical implementation. Additionally, false alarms remain a persistent challenge in seizure detection/prediction systems. While our work represents a step forward in addressing this issue, we recognize that further research is necessary to achieve more significant improvements in reducing false alarms.

ACKNOWLEDGMENT

The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at C3SE, partially funded by the Swedish Research Council (VR) through grant agreement no. 2022-06725.

REFERENCES

- [1] WHO. (2024). *Epilepsy*. [Online]. Available: <https://www.who.int/health-topics/epilepsy>
- [2] D. J. Thurman, D. C. Hesdorffer, and J. A. French, "Sudden unexpected death in epilepsy: Assessing the public health burden," *Epilepsia*, vol. 55, no. 10, pp. 1479–1485, Oct. 2014.
- [3] W. Löscher, H. Potschka, S. M. Sisodiya, and A. Vezzani, "Drug resistance in epilepsy: Clinical impact, potential mechanisms, and new innovative treatment options," *Pharmacol. Rev.*, vol. 72, no. 3, pp. 606–638, Jul. 2020.
- [4] K. Fujiwara, M. Miyajima, T. Yamakawa, E. Abe, Y. Suzuki, Y. Sawada, M. Kano, T. Maehara, K. Ohta, T. Sasai-Sakuma, T. Sasano, M. Matsuura, and E. Matsushima, "Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 6, pp. 1321–1332, Jun. 2016.
- [5] L. Billeci, D. Marino, L. Insana, G. Vatti, and M. Varanini, "Patient-specific seizure prediction based on heart rate variability and recurrence quantification analysis," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0204339.
- [6] T. Yamakawa, M. Miyajima, K. Fujiwara, M. Kano, Y. Suzuki, Y. Watanabe, S. Watanabe, T. Hoshida, M. Inaji, and T. Maehara, "Wearable epileptic seizure prediction system with machine-learning-based anomaly detection of heart rate variability," *Sensors*, vol. 20, no. 14, p. 3987, Jul. 2020.
- [7] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen, "Using Lorenz plot and cardiac sympathetic index of heart rate variability for detecting seizures for patients with epilepsy," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4563–4566.
- [8] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen, "Detection of epileptic seizures with a modified heart rate variability algorithm based on Lorenz plot," *Seizure*, vol. 24, pp. 1–7, Jan. 2015.
- [9] F. Forooghifar, A. Aminifar, L. Cammoun, I. Wisniewski, C. Ciomas, P. Ryvlin, and D. Atienza, "A self-aware epilepsy monitoring system for real-time epileptic seizure detection," *Mobile Netw. Appl.*, vol. 27, no. 2, pp. 677–690, Apr. 2022.
- [10] F. Forooghifar, A. Aminifar, and D. Atienza, "Resource-aware distributed epilepsy monitoring using self-awareness from edge to cloud," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1338–1350, Dec. 2019.
- [11] F. Forooghifar, A. Aminifar, T. Teijeiro, A. Aminifar, J. Jeppesen, S. Beniczky, and D. Atienza, "Self-aware anomaly-detection for epilepsy monitoring on low-power wearable electrocardiographic devices," in *Proc. IEEE 3rd Int. Conf. Artif. Intell. Circuits Syst. (AICAS)*, Jun. 2021, pp. 1–4.
- [12] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, "Multifractality in human heartbeat dynamics," *Nature*, vol. 399, no. 6735, pp. 461–465, Jun. 1999.
- [13] L. Fang and R. L. Thews, *Wavelets in Physics*. Singapore: World Scientific, 1998.
- [14] A. Chhabra and R. V. Jensen, "Direct determination of the $f(\alpha)$ singularity spectrum," *Phys. Rev. Lett.*, vol. 62, no. 12, pp. 1327–1330, Mar. 1989.
- [15] S. M. Shekhatkar, Y. Kotriwar, K. P. Harikrishnan, and G. Ambika, "Detecting abnormality in heart dynamics from multifractal analysis of ECG signals," *Sci. Rep.*, vol. 7, no. 1, p. 15127, Nov. 2017.
- [16] A. M. Aguilar-Molina, F. Angulo-Brown, and A. Muñoz-Diosdado, "Multifractal spectrum curvature of RR tachograms of healthy people and patients with congestive heart failure, a new tool to assess health conditions," *Entropy*, vol. 21, no. 6, p. 581, Jun. 2019.
- [17] S. Li, "Multifractal detrended fluctuation analysis of congestive heart failure disease based on constructed heartbeat sequence," *IEEE Access*, vol. 8, pp. 205244–205249, 2020.
- [18] M. Mangalam, A. Sadri, J. Hayano, E. Watanabe, K. Kiyono, and D. G. Kelty-Stephen, "Multifractal foundations of biomarker discovery for heart disease and stroke," *Sci. Rep.*, vol. 13, no. 1, p. 18316, Oct. 2023.
- [19] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Jan. 2001.
- [20] M. Ihle, H. Feldwisch-Drentrup, C. Teixeira, A. Witon, B. Schelter, J. Timmer, and A. Schulze-Bonhage, "EPILEPSIAE: A European epilepsy database," *Comput. Methods Programs Biomed.*, vol. 106, no. 3, pp. 127–138, Jun. 2012.
- [21] S. Lundberg and S. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2017, pp. 4765–4774.
- [22] M. Sundararajan and A. Najmi, "The many Shapley values for model explanation," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2019, pp. 9269–9278.
- [23] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 804–813, Oct. 2019.
- [24] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky, "Classification of patterns of EEG synchronization for seizure prediction," *Clin. Neurophysiol.*, vol. 120, no. 11, pp. 1927–1940, Nov. 2009.
- [25] L. Kuhlmann, K. Lehnhertz, M. P. Richardson, B. Schelter, and H. P. Zaveri, "Seizure prediction—Ready for a new era," *Nature Rev. Neurol.*, vol. 14, no. 10, pp. 618–630, Oct. 2018.
- [26] U. R. Acharya, Y. Hagiwara, and H. Adeli, "Automated seizure prediction," *Epilepsy Behav.*, vol. 88, pp. 251–261, Nov. 2018.
- [27] K. Gadhoumi, J.-M. Lina, F. Mormann, and J. Gotman, "Seizure prediction for therapeutic devices: A review," *J. Neurosci. Methods*, vol. 260, pp. 270–282, Feb. 2016.
- [28] D. Zambrana-Vinaroz, J. M. Vicente-Samper, J. Manrique-Cordoba, and J. M. Sabater-Navarro, "Wearable epileptic seizure prediction system based on machine learning techniques using ECG, PPG and EEG signals," *Sensors*, vol. 22, no. 23, p. 9372, Dec. 2022.
- [29] W. Xiong, E. S. Nurse, E. Lambert, M. J. Cook, and T. Kameneva, "Seizure forecasting using long-term electroencephalography and electrocardiogram data," *Int. J. Neural Syst.*, vol. 31, no. 9, Sep. 2021, Art. no. 2150039.
- [30] B. Büyükkakir, F. Elmaz, and A. Y. Mutlu, "Hilbert vibration decomposition-based epileptic seizure prediction with neural network," *Comput. Biol. Med.*, vol. 119, Apr. 2020, Art. no. 103665.
- [31] K. Hoyos-Osorio, J. Castañeda-González, and G. Daza-Santacoloma, "Automatic epileptic seizure prediction based on scalp EEG and ECG signals," in *Proc. XXI Symp. Signal Process., Images Artif. Vis. (STSIVA)*, Aug. 2016, pp. 1–7.
- [32] M. Valderrama, S. Nikolopoulos, C. Adam, V. Navarro, and M. Le Van Quyen, "Patient-specific seizure prediction using a multi-feature and multi-modal EEG-ECG classification," in *Proc. XII Mediterr. Conf. Med. Biol. Eng. Comput.*, Jan. 2010, pp. 77–80.
- [33] F. Asharindavida, M. Shamim Hossain, A. Thacham, H. Khammari, I. Ahmed, F. Alraddady, and M. Masud, "A forecasting tool for prediction of epileptic seizures using a machine learning approach," *Concurrency Comput. Pract. Exper.*, vol. 32, no. 1, p. 5111, Jan. 2020.
- [34] L. Billeci, A. Tonacci, M. Varanini, P. Detti, G. Z. M. de Lara, and G. Vatti, "Epileptic seizures prediction based on the combination of EEG and ECG for the application in a wearable device," in *Proc. IEEE 23rd Int. Symp. Consum. Technol. (ISCT)*, Jun. 2019, pp. 28–33.
- [35] A. Popov, O. Panichev, Y. Karplyuk, Y. Smirnov, S. Zaunseder, and V. Kharytonov, "Heart beat-to-beat intervals classification for epileptic seizure prediction," in *Proc. Signal Process. Symp. (SPSympo)*, Sep. 2017, pp. 1–4.
- [36] Y. Yang, X. Qin, H. Wen, F. Li, and X. Lin, "Patient-specific approach using data fusion and adversarial training for epileptic seizure prediction," *Frontiers Comput. Neurosci.*, vol. 17, May 2023, Art. no. 1172987.
- [37] W. Phomsiricharoenphant, S. Ongwattanakul, and Y. Wongswat, "The preliminary study of EEG and ECG for epileptic seizure prediction based on Hilbert Huang transform," in *Proc. 7th Biomed. Eng. Int. Conf.*, Nov. 2014, pp. 1–4.
- [38] A. Van De Vel, K. Cuppens, B. Bonroy, M. Milosevic, K. Jansen, S. Van Huffel, B. Vanrumste, L. Lagae, and B. Ceulemans, "Non-EEG seizure detection systems and potential SUDEP prevention: State of the art," *Seizure*, vol. 41, pp. 141–153, Oct. 2016.
- [39] C. Hoppe, M. Feldmann, B. Blachut, R. Surges, C. E. Elger, and C. Helmstaedter, "Novel techniques for automated seizure registration: Patients' wants and needs," *Epilepsy Behav.*, vol. 52, pp. 1–7, Nov. 2015.
- [40] H. Hashimoto, K. Fujiwara, Y. Suzuki, M. Miyajima, T. Yamakawa, M. Kano, T. Maehara, K. Ohta, T. Sasano, M. Matsuura, and E. Matsushima, "Heart rate variability features for epilepsy seizure prediction," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, Oct. 2013, pp. 1–4.

- [41] A. Ghosh, A. Sarkar, T. Das, and P. Basak, "Pre-ictal epileptic seizure prediction based on ECG signal analysis," in *Proc. 2nd Int. Conf. Converg. Technol. (I2CT)*, Apr. 2017, pp. 920–925.
- [42] M. K. Moridani and H. Farhadi, "Heart rate variability as a biomarker for epilepsy seizure prediction," *Bratislava Med. J.*, vol. 118, no. 1, pp. 3–8, 2017.
- [43] A. H. Thomas, A. Aminifar, and D. Atienza, "Noise-resilient and interpretable epileptic seizure detection," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Oct. 2020, pp. 1–5.
- [44] Y. Statsenko, V. Babushkin, T. Talako, T. Kurbatova, D. Smetanina, G. L. Simiyu, T. Habuza, F. Ismail, T. M. Almansoori, K. N. V. Gorkom, M. Szűcs, A. Hassan, and M. Ljubisavljevic, "Automatic detection and classification of epileptic seizures from EEG data: Finding optimal acquisition settings and testing interpretable machine learning approach," *Biomedicines*, vol. 11, no. 9, p. 2370, Aug. 2023.
- [45] A. Einizade, S. Nasiri, M. Mozafari, S. H. Sardouie, and G. D. Clifford, "Explainable automated seizure detection using attentive deep multi-view networks," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104076.
- [46] X. Zhao, N. Yoshida, T. Ueda, H. Sugano, and T. Tanaka, "Epileptic seizure detection by using interpretable machine learning models," *J. Neural Eng.*, vol. 20, no. 1, Feb. 2023, Art. no. 015002.
- [47] V. Gabeff, T. Teijeiro, M. Zapater, L. Cammoun, S. Rheims, P. Ryvlin, and D. Atienza, "Interpreting deep learning models for epileptic seizure detection on EEG signals," *Artif. Intell. Med.*, vol. 117, May 2021, Art. no. 102084.
- [48] D. Sopic, T. Teijeiro, D. Atienza, A. Aminifar, and P. Ryvlin, "Personalized seizure signature: An interpretable approach to false alarm reduction for long-term epileptic seizure detection," *Epilepsia*, vol. 64, no. S4, pp. S23–S33, Dec. 2023.
- [49] B. Huang, R. Zanetti, A. Abtahi, D. Atienza, and A. Aminifar, "EpilepsyNet: Interpretable self-supervised seizure detection for low-power wearable systems," in *Proc. IEEE 5th Int. Conf. Artif. Intell. Circuits Syst. (AICAS)*, Jun. 2023, pp. 1–5.
- [50] I. Al-Hussaini and C. S. Mitchell, "SeizFt: Interpretable machine learning for seizure detection using wearables," *Bioengineering*, vol. 10, no. 8, p. 918, Aug. 2023.
- [51] M. Pinto, T. Coelho, A. Leal, F. Lopes, A. Dourado, P. Martins, and C. Teixeira, "Interpretable EEG seizure prediction using a multi-objective evolutionary algorithm," *Sci. Rep.*, vol. 12, no. 1, pp. 1–15, Mar. 2022.
- [52] S. Bagheralimi, T. Teijeiro, D. Atienza, and A. Aminifar, "Personalized real-time federated learning for epileptic seizure detection," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 2, pp. 898–909, Feb. 2022.
- [53] A. Arneodo, E. Bacry, and J.-F. Muzy, "The thermodynamics of fractals revisited with wavelets," *Phys. A, Stat. Mech. Appl.*, vol. 213, nos. 1–2, pp. 232–275, Jan. 1995.
- [54] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent individualized feature attribution for tree ensembles," 2018, *arXiv:1802.03888*.
- [55] Y. Matsui and T. Matsui, "NP-completeness for calculating power indices of weighted majority games," *Theor. Comput. Sci.*, vol. 263, nos. 1–2, pp. 305–310, Jul. 2001.
- [56] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "From local explanations to global understanding with explainable AI for trees," *Nature Mach. Intell.*, vol. 2, no. 1, pp. 56–67, Jan. 2020.
- [57] M. Ben Mbarek, I. Assali, S. Hamdi, A. Ben Abdallah, O. David, M. Aissi, M. Carrère, and M. H. Bedoui, "Automatic and manual prediction of epileptic seizures based on ECG," *Signal, Image Video Process.*, vol. 18, no. 5, pp. 4175–4190, Jul. 2024.
- [58] P. J. Fleming and J. J. Wallace, "How not to lie with statistics: The correct way to summarize benchmark results," *Commun. ACM*, vol. 29, no. 3, pp. 218–221, Mar. 1986.
- [59] W. J. Conover, *Practical Nonparametric Statistics*, vol. 350. Hoboken, NJ, USA: Wiley, 1999.
- [60] F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in Statistics: Methodology and Distribution*. Cham, Switzerland: Springer, 1992, pp. 196–202.
- [61] C. van Zyl, X. Ye, and R. Naidoo, "Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of grad-CAM and SHAP," *Appl. Energy*, vol. 353, Jan. 2024, Art. no. 122079.
- [62] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.
- [63] G. Roffo, S. Melzi, U. Castellani, A. Vinciarelli, and M. Cristani, "Infinite feature selection: A graph-based feature filtering approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4396–4410, Dec. 2021.



AZRA ABTAHI (Member, IEEE) received the Ph.D. degree in electrical engineering from the Sharif University of Technology (SUT), Tehran, Iran, in 2018. She is currently an Associate Senior Lecturer with the Department of Computer Science and Media Technology, Malm University, Sweden. During her doctoral studies, she participated in a year-long sabbatical program with Queen's University, ON, Canada, in 2016. Following her Ph.D., she was a Postdoctoral Fellow with SUT and Sweden's Lund University. She has worked on various areas of machine learning and signal processing. Her main research interest includes these areas, with a special emphasis on interpretable machine learning techniques tailored for biomedical applications.



PHILIPPE RYVLIN is currently Professor of neurology, Chair of the Department of Clinical Neurosciences and of the Neuropsychology and Neurorehabilitation Service with the University Hospital of Lausanne (CHUV), Switzerland, as well as an Affiliated Professor of multimodal epilepsy surgery evaluation with the University of Copenhagen, Denmark. He coordinates the Medical Informatic Platform and the Human Intracerebral EEG Platform of European Flagship Human Brain Project and European Research Infrastructure EBRAINS. He is the author or co-author of more than 350 scientific publications. His current research interests include digital health technologies applied to large scale data federation in clinical neurosciences, human intracerebral EEG research, seizure detection using wearable devices and understanding, predicting and preventing sudden unexpected death in epilepsy (SUDEP).



AMIR AMINIFAR (Senior Member, IEEE) received the Ph.D. degree from Swedish National Computer Science Graduate School (CUGS), Sweden, in 2016. From 2016 to 2020, he held a scientist position with Swiss Federal Institute of Technology (EPFL), Switzerland. He is currently an Associate Senior Lecturer with the Department of Electrical and Information Technology, Lund University, Sweden, and the Director of the Intelligent Systems Laboratory. His research interests include centered around machine learning in biomedical applications and health informatics. He is also an Associate Editor of *ACM Computing Surveys* and *IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS*.