Assessing the Effectiveness of Heart Rate Variability as A Diagnostic Tool for Brain Injuries in Infants

K. Rezaei[1,2] K. Yu[1,2] S. R. Mathieson[2] A. Flynn[2] G. Lightbody[1,2] G.B. Boylan[2,3] and W. P. Marnane[1,2]
[1]Electrical and Electronic Engineering, School of Engineering, University College Cork, Cork, Ireland
[2]INFANT Research Centre, University College Cork, Cork, Ireland
[3]Pediatrics and Child Health, University College Cork, Cork, Ireland

***Abstract*—Hypoxic-Ischemic Encephalopathy (HIE), marked by cerebral oxygen deprivation, prompts exploration beyond the electroencephalogram (EEG) modality. This study investigates heart rate variability (HRV) to assess its potential for seizure detection and HIE grading for neonates. This study utilizes two annotated datasets from real-world clinical settings. Heart Rate (HR) is calculated from the Electrocardiogram (ECG) signal, which are then denoised and segmented. Sixteen time and frequency domain features are extracted from each HR segment. Employing Random Forest (RF), Support Vector Machine (SVM), and Isolation Forest (IF) classifiers, the investigation addresses the detection of seizure and nonseizure segments in ECG, alongside categorizing HIE severity into two mild and normal or moderate and severe grades. While the patient-independent evaluation of the seizure detection system reveals promising outcomes for specific cases, there is a requirement for further refinement in this aspect and exploration into the correlation between HR and EEG, considering the modest AUC of 68.54 percent gained across the entire dataset. In contrast, the HIE grading results present a more promising scenario, attaining an AUC of 77.13 percent. This emphasizes the efficacy of the HIE grading system as a significant diagnostic tool, suggesting its potential for broader clinical applications.**

***Index Terms*—Electrocardiogram, Heart Rate Variability, Hypoxic Ischemic Encephalopathy grading, Seizure, Time and Frequency Domain Features, Random Forest Classifier**

I. INTRODUCTION

Newborn infants who have suffered a brain injury such as hypoxic-ischaemic encephalopathy (HIE) and stroke, very often develop seizures. Lack of treatment due to a failure in determining the severity of HIE and detecting seizures can cause further brain injury and in severe cases, death. The gold standard for detecting seizures and determining HIE grade is through continuous electroencephalogram (cEEG) monitoring. However neonatal EEG requires specialist interpretation and the expertise required may not be available 24/7 [1]. This has prompted a significant amount of research on automated seizure detection and HIE grading [2], [3].

In addition to the problem of cEEG interpretation, introducing neonatal EEG monitoring is challenging, with the expense of the equipment and the difficulty in recording the EEG of newborn infants. Thus, the need of other methods to help identify those most at risk. One of the physiological signals which is easily accessible to the clinicians and can routinely and continuously be monitored in newborn infants is the Electrocardiogram (ECG). In addition, ECG can be obtained in remote hospitals with a lower resource setting. Valuable diagnostic information can be obtained, by analysing and processing of the ECG signal. Studying Heart Rate Variability (HRV), extracted from the HR signal may enhance the potential use of the ECG recording as a clinical tool aiding diagnosis and treatment of HIE and seizures. However, not enough studies have been conducted using only HRV to assess seizure detection and HIE grading in infants.

For neonatal seizure detection using heart rate, mixed results have been achieved. Greene et al. [4] obtained a sensitivity of 54.6% and specificity of 77.3% using a patient independent linear discriminant classifier with 41 features evaluated on a dataset of 101h recordings from 8 neonates, all having seizures. Doyle et al. [5] utilised a diverse set of 62 time and frequency domain features which were extracted from a dataset of 207.86h recordings from 14 neonates with seizure events. A SVM was evaluated using a Leave-One-Out (LOO) cross-validation technique, which resulted in a sensitivity of 60% and a specificity of 60%. The results of each neonate were checked individually, which showed a notable difference between the best and worst performing patients. Overall, the system performed well only in 2 out of 14 babies, but poorly when applied on the entire group. More recently Olmi et al. [6] obtained a sensitivity of 47% and a specificity of 67% utilising 18 features from epochs of 180s, fed to a support vector machine (SVM) classifier. The database comprised of 45.76h of recordings from 51 neonates, with 22 having seizures. For HIE grading, Ahmed et al. [7] conducted the classification for two classes of HIE grades based on 54 one-hour recordings. They utilized super vectors created from seven long-term statistical features extracted the HR. The classification methods employed included SVM, Gaussian Mixture Models, and supervector-SVM. The most successful outcome was achieved using supervector-SVM, with an Area Under the Curve (AUC) of 81%. Matic et al. [8] evaluated HRV through continuous monitoring of EEG and ECG in two groups of newborns, categorized according to the severity of HIE. The findings indicate an elevation in HRV parameters within the severe HIE group.

A reason for the varied results obtained for neonatal seizure detection using heart rate may in part be found from the analysis carried out by Goulding [9]. Statistical analyses

This work was supported by Science Foundation Ireland (SFI) (19/FFP/6782).

were performed on seven HRV features extracted from the ECG recordings of 26 babies with HIE who had seizures to observe changes in HRV during seizures. The results indicate a trending towards a significant difference among three features between baseline and seizure events. In Goulding et al. [10] the same seven features were demonstrated to show a statistically significant difference between different HIE grades. Results demonstrate that six out of seven features were significantly different between mild and severe HIE groups, while all seven features were significantly different between mild and moderate HIE groups.

In this study, the effectiveness of this clinical finding is assessed from the machine learning point of view. Sixteen features including the mentioned seven features, are extracted from the HR to be explored. Classification was applied on all features and also separately on the seven features for seizure detection and HIE grading into two classes of mild/normal or moderate/severe. This grading boundary is used to determine if Therapeutic Hypothermia (TH) which is the process of body or brain cooling is employed. TH has been found to reduce the risk of death or lifelong disability. [11]

II. DATASET

In this work, two different datasets are employed. First dataset used for the purpose of HIE grading consist of 120 neonates (66 severe/moderate and 54 mild/normal) with a recording duration of about 1 hour from the ANSeR dataset [12]. Second dataset used for seizure detection consists of 14 full term neonates with seizures, secondary to HIE and each recording duration of about 24 to 60 hours. All signals were recorded in the neonatal intensive care unit of Cork University Maternity Hospital and annotated by a neurophysiologist. HRV software (HRV Analysis, University College Cork, Cork) was used to identify R-peaks within the ECG signal and produce the NN interval.

III. METHOD

A. Preprocessing and Feature Extraction

Initially, the Hampel filter was applied to remove outliers, preparing the NN interval data for subsequent feature extraction.

The average resting heart for adults lies in the range 60-100 BPM while in a newborn it is higher. While studies suggest a length of five minutes for detection of very lower frequency components, [13], an epoch length of two minutes is a good trade-off between the appropriate window length for HR data and length in which to accurately capture seizure events and abnormalities [5]. A 50% overlap is introduced to reduce information loss at the boundaries of each segment, allowing a better representation of signal patterns.

Since the ECG recordings were corrupted by artifacts, noise sections were discarded during this phase. In the context of seizure detection, 28,485 epochs were employed, with 655 epochs being excluded due to noise. For HIE grading, 7,493 epochs were utilized, and 39 epochs were removed due to artifacts.

Sixteen time and frequency domain features were extracted from each epoch. It is necessary to resample the NN sequence at uniform intervals when examining the HR time series in the frequency domain. A spline interpolation technique is employed, which involves fitting low-degree polynomials to small subsets of the values. The seven features previously investigated [9], [10], include three statistical measures in the time domain of the Mean of the NN interval (muNN), the Standard Deviation of NN Series (SDNN), and the Triangular Interpolation of the NN (TINN) (in our analysis we have replaced TINN with Triangular Index (Tri) representing the ratio between the number of NNI and the maximum of the NNI histogram as TINN provides a measure of the overall width of the histogram and may not capture the specific contributions of individual NNIs to the heart rate variability). Four Frequency domain parameters are determined by estimating the power spectral density (PSD) of the resampled NN series using Welch's method. The four features and their frequency bands are Very Low Frequency (VLF) power (0.01-0.04 Hz), Low Frequency (LF) power (0.04- 0.2 Hz), High Frequency (HF) power (0.2- 2 Hz) along with the the LF/HF ratio. The additional ten features employed include Root Mean Square of Successive Standard Deviation (RMSSD), Skewness and Kurtosis of the NN distribution and Nonlinear features Approximate Entropy (ApEn) to quantify regularity and complexity of a stationary signal, Poincaré parameters SD1, SD2, SD2/SD1 ratio to quantify self-similarity [14] and the modified Cardiac Sympathetic Index (CSI) and Cardiac Vagal Index (CVI) [15].

B. Classification

SVM with Radial Basis Function (RBF) kernel is a supervised learning method separating classes by finding the optimal hyperplane. Random Forest (RF) and Isolation Forest (IF) are tree-based classification algorithm, RF using labelled data for training, whereas IF, is an unsupervised decision-tree-based method developed for anomaly detection. To optimize classifier performance, the hyperparameters of all classifiers such as Gama and C for SVM and number of trees, maximum depth of the tree, minimum number of samples required to be at a leaf node, maximum number of features at each split and etc. for RF and IF, were tuned using randomized or grid search of N-fold Cross Validation (CV) based on the classification and data. Prediction results were compared with expert-created annotations derived from EEG signals.

For seizure detection RF, SVM and IF were employed. Since the IF algorithm detects outliers in the data, it is only trained on the nonseizure segments of the data with the purpose of detecting the nonseizure segments, including those containing noise, aiming to find any distinctive pattern in the nonseizure portions of the signal. Consequently, the Hampel filter was omitted from the preprocessing of IF input data. Given the considerable class imbalance, with the number of seizure epochs significantly less than nonseizure epochs (approximately 15 times less), a Synthetic Minority Oversampling Technique (SMOTE) was applied to oversample the minority class during the seizure detection process [16]. The epochs were subsequently categorized into seizure and baseline for the seizure detection task.

For the purpose of HIE grading, RF and SVM were used. epochs were classified into mild/normal or moderate/severe HIE. To identify grade of each baby, the majority of epochs labels was considered.

C. Evaluation

To evaluate performance a patient-independent cross validation was employed. Classification performance is reported based on sensitivity, specificity and Area Under the Curve (AUC). Since the annotations are determined by the EEG signal and the labels approximated for each segmented epoch, changes both prior to and following each epoch should be considered. Hence, a post-processing step involving a moving average filter is employed on the predicted class probabilities of the model for seizure detection. In addition, Impurity-based feature importance and Kernel Density Estimate (KDE) were performed for feature analysis to observe the significance and importance of the extracted features in this study.

IV. RESULTS

For seizure detection a Leave-One-Out (LOO) cross validation was employed. The RF classifier employing the seven features (Referred as RF2 in tables I and II) gave an AUC of 65.57%. The RF classifier employing 16 features achieved the highest AUC of 68.54% and the results generally improved after post-processing with the highest AUC of 73.8% for the SVM classifier. Additionally high specificity of 90.62 in the IF algorithm indicates its strong capability in identifying the nonseizure segments; however, low sensitivity and AUC undermine confidence in this method. Regarding HIE grading, a 10-fold cross validation, each fold containing 12 patients was conducted and as shown in table II RF similarly boasts the highest AUC albeit with a higher percentage of 77.13%.

TABLE I: Seizure detection results of applied method.

Evaluation			
Classifier	Sensitivity(%)	Specificity(%)	AUC(%)
Random Forest	50.91	78.58	68.54
Random Forest2	56.80	69.31	65.57
Support Vector Machine	57.39	78.95	68.11
Isolation Forest	11.79	90.62	53.16
Evaluation After Post Processing			
Classifier	Sensitivity(%)	Specificity(%)	AUC(%)
Random Forest	46.23	85.32	69.57
Random Forest2	53.03	74.00	67.16
Support Vector Machine	45.77	88.60	73.80
Isolation Forest	3.8	94.52	53.08

Since the RF results were the most prominent, their impurity-based feature importance, which is a key metric denoting the mean and standard deviation of accumulation of the impurity decrease within each tree, was computed. The Y-axis values on the feature importance graph indicate the significance assigned to each feature by the Random Forest

TABLE II: Hypoxic-Ischemic Encephalopathy grading results of applied method.

Classifier	Sensitivity(%)	Specificity(%)	AUC(%)
Random Forest	75.38	59.62	77.13
Random Forest2	75.38	61.54	76.24
Support Vector Machine	63.08	65.38	72.66

model. These values are normalized to sum to one, making it a proportion of the overall feature importance indicate the contribution of each feature to the overall predictive performance of the model. As illustrated in Fig. 1, the highest value of feature importance is assigned to LF Power for seizure detection, followed by the LF/HF ratio, muNN, ApEn, CVI and SD ratio. Notably, during both baseline and seizure events, LF Power exhibited the lowest P-value of 0.055 among the seven features previously studied by Goulding [9], establishing its statistical significance. However, three of the additional features of this study, ApEn, CVI and SD ratio, are among the six most effective features in the classification process which is the reason for a higher AUC value for RF compared to RF2 in table I). In the context of HIE grading, muNN takes precedence, followed by CSI and LF/HF ratio as depicted in Fig. 2. Both muNN and LF/HF ratio emerged as features with significant importance across different HIE severity based on the Goulding et al. [10] study involving the seven features.

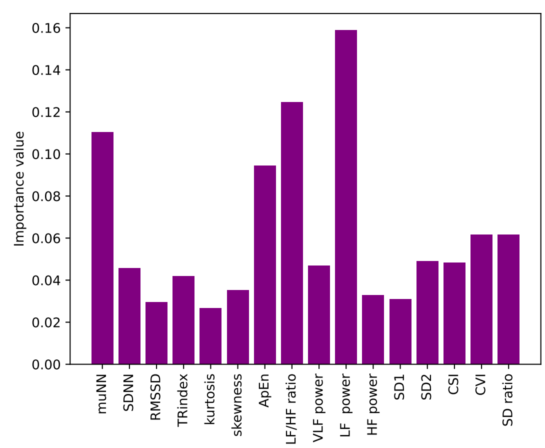

Fig. 1: Feature importance of Random Forest classifier regarding seizure detection.

The KDE distribution plot visually illustrates the distribution of data across different classes. Fig. 3 illustrates the KDE of the two features of significant importance in the context of seizure detection. Fig. 4 presents plots corresponding to two most effective features in the HIE grading phase.

V. CONCLUSION

The experimental results underscore the significance of specific features, including muNN, LF power, and LF/HF ratio, in identifying seizures and HIE severity. Notably, these effective features, derived directly from time and frequency information,

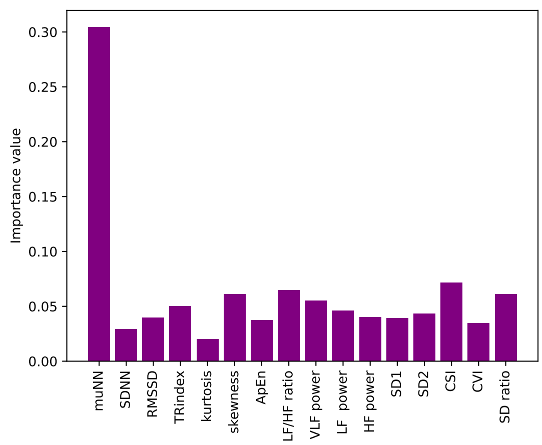

Fig. 2: Feature importance of Random Forest classifier regarding Hypoxic-Ischemic Encephalopathy grading.

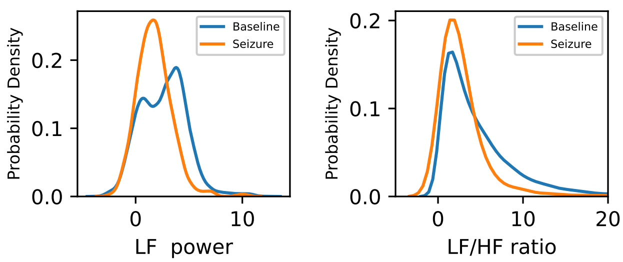

Fig. 3: Kernel Density Estimate distribution of seizure detection.

outdo ones such as kurtosis, which describe the distribution of NN intervals. However, their precision in seizure detection may be limited, as evidenced by an AUC of 68.54%. Comparing to other state-of-the-art approaches mentioned earlier, an overall evaluation results of higher than 80% using HRV features was not achieved despite using different methods and databases. This suggests the lack of reliability on this system. Additionally, as the annotation based on EEG is the only mean used to label seizures, it is worthwhile to determine the precise corresponding time of effect of brain seizure on ECG signal.

Higher sensitivity, specificity and an AUC of 77.13% underscores HRV's substantial potential for deeper exploration in automatic HIE grading, aligning with prior research, though

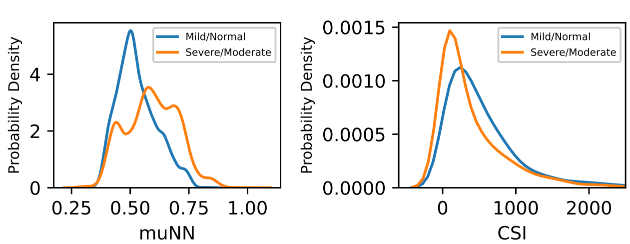

Fig. 4: Kernel Density Estimate distribution of Hypoxic-Ischemic Encephalopathy grading.

additional investigation is warranted. This marks a positive stride in assisting physicians as a valuable tool detecting HIE grades, alleviating the need for constant expert availability. Additionally, relying on the ECG signal of the baby, physicians can promptly initiate cooling procedures upon an escalating HIE grade, mitigating potential complications for infants.

References

[1] R. Dilena, F. Raviglione, G. Cantalupo, D. M. Cordelli, P. De Liso, M. Di Capua, R. Falsaperla, F. Ferrari, M. Fumagalli, S. Lori, *et al.*, "Consensus protocol for EEG and amplitude-integrated EEG assessment and monitoring in neonates," *Clinical Neurophysiology*, vol. 132, no. 4, pp. 886–903, 2021.

[2] I. Bersani, F. Piersigilli, D. Gazzolo, F. Campi, I. Savarese, A. Dotta, P. P. Tamborrino, C. Auriti, and C. Di Mambro, "Heart rate variability as possible marker of brain damage in neonates with hypoxic ischemic encephalopathy: a systematic review," *European Journal of Pediatrics*, vol. 180, pp. 1335–1345, 2021.

[3] B. Olmi, L. Frassineti, A. Lanata, and C. Manfredi, "Automatic detection of epileptic seizures in neonatal intensive care units through EEG, ECG and video recordings: A survey," *IEEE Access*, vol. 9, pp. 138174–138191, 2021.

[4] B. R. Greene, P. de Chazal, G. B. Boylan, S. Connolly, and R. B. Reilly, "Electrocardiogram based neonatal seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 4, pp. 673–682, 2007.

[5] O. Doyle, A. Temko, W. Marnane, G. Lightbody, and G. Boylan, "Heart rate based automatic seizure detection in the newborn," *Medical engineering & physics*, vol. 32, no. 8, pp. 829–839, 2010.

[6] B. Olmi, C. Manfredi, L. Frassineti, C. Dani, S. Lori, G. Bertini, C. Cossu, M. Bastianelli, S. Gabbanini, and A. Lanatà, "Heart rate variability analysis for seizure detection in neonatal intensive care units," *Bioengineering*, vol. 9, no. 4, p. 165, 2022.

[7] R. Ahmed, A. Temko, W. P. Marnane, G. Boylan, and G. Lightbody, "Classification of hypoxic-ischemic encephalopathy using long term heart rate variability based features," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2355–2358, IEEE, 2015.

[8] V. Matić, P. J. Cherian, D. Widjaja, K. Jansen, G. Naulaers, S. Van Huffel, and M. De Vos, "Heart rate variability in newborns with hypoxic brain injury," in *Oxygen Transport to Tissue XXXV*, pp. 43–48, Springer, 2013.

[9] R. M. D. Goulding, *Heart rate variability and electroencephalography in Infants with hypoxic ischaemic encephalopathy*. PhD thesis, University College Cork, 2017.

[10] R. M. Goulding, N. J. Stevenson, D. M. Murray, V. Livingstone, P. M. Filan, and G. B. Boylan, "Heart rate variability in hypoxic ischemic encephalopathy: correlation with eeg grade and 2-y neurodevelopmental outcome," *Pediatric research*, vol. 77, no. 5, pp. 681–687, 2015.

[11] S. Shankaran, "Therapeutic hypothermia for neonatal encephalopathy," *Current treatment options in neurology*, vol. 14, pp. 608–619, 2012.

[12] A. M. Pavel, S. R. Mathieson, V. Livingstone, J. M. O'Toole, R. M. Pressler, L. S. de Vries, J. M. Rennie, S. Mitra, E. M. Dempsey, D. M. Murray, *et al.*, "Heart rate variability analysis for the prediction of eeg grade in infants with hypoxic ischaemic encephalopathy within the first 12 h of birth," *Frontiers in pediatrics*, vol. 10, p. 1016211, 2023.

[13] G. Volpes, C. Barà, A. Busacca, S. Stivala, M. Javorka, L. Faes, and R. Pernice, "Feasibility of ultra-short-term analysis of heart rate and systolic arterial pressure variability at rest and during stress via time-domain and entropy-based measures," *Sensors*, vol. 22, no. 23, p. 9149, 2022.

[14] M. Toichi, T. Sugiura, T. Murai, and A. Sengoku, "A new method of assessing cardiac autonomic function and its comparison with spectral analysis and coefficient of variation of r–r interval," *Journal of the autonomic nervous system*, vol. 62, no. 1-2, pp. 79–84, 1997.

[15] 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, 2015.

[16] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance," *Information Sciences*, vol. 505, pp. 32–64, 2019.