

Grading Method for Hypoxic-Ischemic Encephalopathy Based on Neonatal EEG

Jingmin Guo¹, Xiu Cheng¹ and Duanpo Wu^{2, 3, *}

Abstract: The grading of hypoxic-ischemic encephalopathy (HIE) contributes to the clinical decision making for neonates with HIE. In this paper, an automated grading method based on electroencephalogram (EEG) data is proposed to describe the severity of HIE infants, namely mild asphyxia, moderate asphyxia and severe asphyxia. The automated grading method is based on a multi-class support vector machine (SVM) classifier, and the input features of SVM classifier include long-term features which are extracted by decomposing the EEG data into different 64 s epoch data and short-term features which are extracted by segmenting the 64 s epoch data into 8 s epoch data with 4 s overlap. Of note, the correlation coefficient and asymmetry extracted in this paper have obvious discriminating capability in HIE infants classification. The experimental results show that the proposed method can achieve the classification accuracy of 78.3%, 75.8% and 87.0% of the mild asphyxia group, moderate asphyxia group and severe asphyxia group, respectively. Moreover, the overall accuracy and kappa used to evaluate the performance of the proposed method can reach 79.5% and 0.69, respectively.

Keywords: Hypoxic-ischemic encephalopathy, electroencephalograph, neonate, support vector machine.

1 Introduction

Hypoxic-ischemic encephalopathy (HIE) is a type of brain dysfunction that occurs when an infant's brain can't receive enough oxygen and blood. It is a major cause of morbidity and mortality in neonates throughout the world with an incidence of 0.1% to 0.8% in developed countries and up to 2.6% in developing countries [Shah (2010); Awal, Lai, Azemi et al. (2016)]. HIE not only threatens the life of neonates, but also is an important factor to bring serious sequelae to neonates, which seriously affects the quality of neonates. Therefore, early diagnosis and prognosis assessment of HIE are particularly important. Optional tests

¹Fujian Provincial Maternity and Children's Hospital, Affiliated Hospital of Fujian Medical University, Fuzhou, China.

²School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, China.

³Hangzhou Neuro Science and Technology Co., Ltd., Hangzhou, China.

*Corresponding Author: Duanpo Wu. Email: wudianpo@hdu.edu.cn.

Received: 23 May 2019; Accepted: 03 July 2019.

to confirm HIE may include electrocardiogram (ECG), electroencephalograph (EEG) and evoked potential tests. In the aforementioned types of tests for diagnosis of HIE, EEG is a relatively successful and non-invasive tool for early identification of infants at risk. By observing EEG, neurologists can assess brain activity including the symmetry of activity over the two hemispheres, synchrony of activity that occurs in bursts, overall continuity, amplitude of the activity, presence of sleep wake cycling (SWC) and the response of EEG to external stimuli [Murray, Boylan, Ryan et al. (2009)]. In fact, the classification of abnormal grades of EEG can provide objective evidence for clinical diagnostic, and to guide treatment of HIE. In general, grading of HIE with EEG can be done by assessment of discontinuity, amplitude, asynchrony and presence of SWC [Mariani, Scelsa, Pogliani et al. (2008); Briatore, Ferrari, Pomero et al. (2013)], where discontinuity is the most commonly used parameter to assess the severity of HIE and it is most useful in severe grades [Matić, Cherian, Jansen et al. (2016); Dunne, Wertheim, Clarke et al. (2017); Awal, Lai, Azemi et al. (2016)]. However, grading HIE using EEG is difficult, time-consuming, and requires the presence of a highly qualified neurophysiologist. Hence, an automated grading method for HIE could be of great convenience to medical staff or assist the non-expert's assessment of cerebral activity.

The main motivation for developing an automated model for HIE grading is to assist the clinical decision making for neonates with HIE. The recent studies of HIE grading focus on developing quantitative features. Tich et al. [Tich, Derambure, Lamblin et al. (2018)] have assessed the severity of cerebral disability after perinatal hypoxia and predicted long-term prognosis by calculating the minimal/maximal amplitude index, the burst suppression rate, the spectral power and the spectral edge frequency of EEG. The results show that quantitative EEG markers are associated with the EEG grades of severity. Bourel-Ponchel et al. [Emilie, Marie-Dominique and Florence (2018)] have also confirmed that conventional EEG is still the gold standard for hypothermia decision and prognosis assessment by analyzing the results of 120 neonates admitted to hospital from January 2013 to February 2015 due to suspected HIE. Moreover, various time domain, frequency domain and joint time-frequency domain features have been extracted and then combined with various methods to grade HIE [Ambalavanan, Carlo, Shankaran et al. (2006); Korotchikova, Stevenson, Walsh et al. (2011); Ahmed, Temko, Marnane et al. (2016)]. Stevenson et al. [Stevenson, Korotchikova, Temko et al. (2013)] have proposed a method of automatically grading the degree of abnormality in an hour-long epoch of neonate. The automated HIE grading system is based on a multi-class linear classifier grading of short-term epoch of EEG which are converted into a long-term grading of EEG using majority vote operation. In a more recent study, Ahmed et al. [Ahmed, Temko, Marnane et al. (2016)] have proposed a cross-disciplinary method to classify the severity of HIE in neonates using EEG, which adopts the sequences of short-term features of EEG to grade an hour-long recording and improve the overall accuracy.

In this study, the paper presents an automated grading method for describing the severity of HIE infants. The aim of the paper is to classify grade of HIE (mild asphyxia, moderate asphyxia and severe asphyxia). In general, the classification accuracy of the model is closely related to the construction of features. Therefore, this paper focuses on feature

engineering and comprehensive analysis of long-term and short-term features from EEG signals. Long-term features which are extracted by decomposing the EEG data into 64 s epoch data and calculating the time domain and frequency domain features. Short-term features which are extracted by segmenting the 64 s epoch data into 8 s epoch data with 4 s overlap. Especially, the paper extracts two features that have significant discriminating capability in HIE infants classification, namely correlation coefficient and asymmetry. To the knowledge of the authors, although a large number of literatures have shown that the asymmetry of EEG signals play an important role in the diagnosis of encephalopathy (including HIE infants) [Low, Mathieson, Stevenson et al. (2014); Knott, Mahoney, Kennedy et al. (2001)], few studies [Schindler, Leung, Elger et al. (2006)] have been conducted on the classification of HIE infants using correlation coefficient. Moreover, since the support vector machine (SVM) algorithm has been successfully applied in the HIE infants classification model and behaves good performance [Ahmed, Temko, Marnane et al. (2016)]. On the one hand, the SVM algorithm behaves better performance than other algorithms on small dataset [Ahmed, Temko, Marnane et al. (2016)]. On the other hand, the SVM algorithm introduces the concept margin, which reduces the algorithm's requirements for data size and data distribution to enhance robustness and generalization capability. Consequently, the SVM algorithm is finally selected to construct the model of the paper. Finally, a multi-class classifier based on SVM algorithm is used to exploit this information and classify these epochs into three grades.

The rest of the paper is organized as follows: Section 2 presents the description of the experimental data used in this study, elucidates a brief description of the feature extraction and explicates the SVM classification in the proposed method. Results are presented in Section 3 and are discussed in Section 4. Finally, Section 5 presents the conclusions.

2 Materials and methods

The paper proposes an automated grading method for HIE infants based on neonatal EEG signals, following the step of data preprocessing, feature extraction, SVM training and SVM testing. Fig. 1 shows the procedure of the proposed method.

2.1 Dataset description of HIE infants

In general, the basic method of neonatal EEG recording is the same as conventional EEG, but considering the small neonatal head circumference, the number of recording electrodes can be appropriately reduced. Ideally, 8-channel recording electrodes are sufficient to cover the entire brain and describe the EEG information of infant. Moreover, according to the studies of Stevenson [Stevenson, Korotchikova, Temko et al. (2013)] and considering the age of subjects and the size of head circumference, this paper collects 8 channels of EEG signals. In fact, in order to improve the quality of the EEG signals, this paper uses alcohol to clean the infants' head to reduce the impedance between the electrode and the scalp. After this treatment, the impedance between electrodes and the scalp is distributed between 1000 Ω and 2000 Ω . Consequently, 8 channels (FP1, FP2, T3, T4, C3, C4, O1 and O2) EEG data with earlobe electrode as reference electrodes have been collected with the sampling rate of 128 Hz from 64 HIE infants (44 males and 20 females)

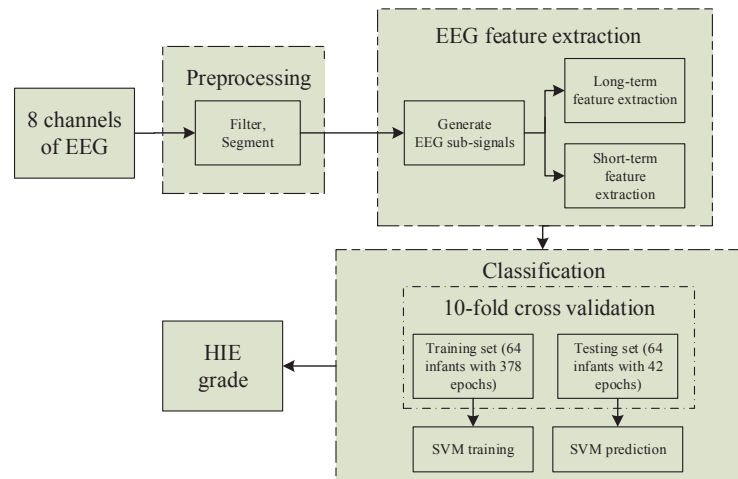


Figure 1: The automated grading procedure of neonatal EEG

in the Department of electrophysiology of Fujian Provincial Maternity and Children's Hospital and Affiliated Hospital of Fujian Medical University. In the data collection of HIE infants, all EEG recordings are tested within 24 hours of post-natal and the test time is similar across the cohort. Subsequently, the paper grades HIE according to the EEG recording. In the HIE infants dataset, 47 mothers are in healthy state during pregnancy, and the remaining 17 mothers have different health problems during pregnancy, such as gestational diabetes mellitus, gestational hypertension, α -thalassemia, etc., but all women have received treatment. Moreover, all of the mothers did not smoke, drink alcohol or take drugs during pregnancy.

2.2 Data preprocessing for EEG signals

In order to capture discriminative features of EEG data from 64 HIE infants, the data requires further preprocessing. In this study, the paper filters the EEG signal with a bandpass filter of 0.1-12 Hz, because there is very little power in the EEG signals in the frequency band higher than 12 Hz [Ahmed, Temko, Marnane et al. (2016)]. And then the paper segments the EEG signals of each patient into 64 s epochs which is similar to segmentation procedure in the study of Lofhede [Löfhede, Thordstein, Löfgren et al. (2010)]. Finally, this paper gets a total of 420 64 s epoch. Furthermore, according to the 1-min Apgar score, 5-min Apgar score, 10-min Apgar score, MRI information and Sarnat Score, etc. provided by the HIE infants dataset, two experienced neonatal EEG technicians are invited to perform visual analysis of each epoch independently and score the EEG epochs with 1, 2 or 3. Where 1 represents mild asphyxia (C1), 2 represents moderate asphyxia (C2), and 3 represents severe asphyxia (C3). In those cases, when two different grades were assigned to the same EEG epoch, the EEG epoch would be subsequently reviewed and discussed by the EEG technicians until consensus on the HIE grade was reached. Consequently, the paper gets C1 group including 128 epochs, C2 group including 169 epochs and C3 group including 123 epochs. The overview of this process is

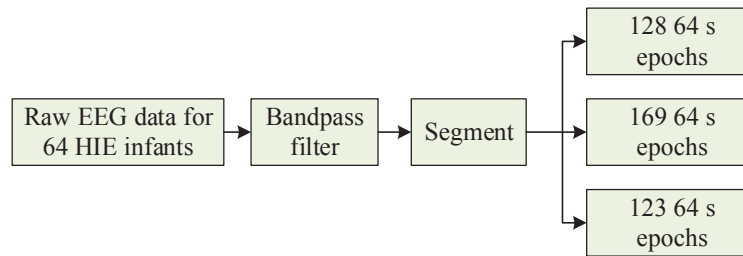


Figure 2: The preprocessing of HIE data

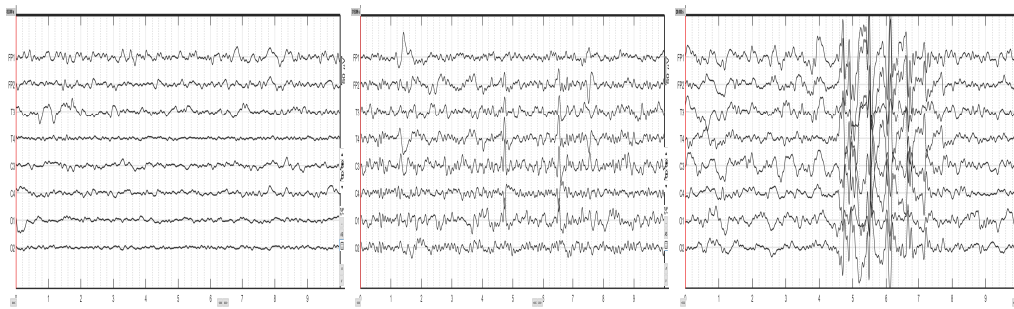
shown in Fig. 2. It is noteworthy that the state of HIE infants will change over time, so the single infant may exist in C1, C2 or C3 at the same time. Tab. 1 shows some basic statistics for HIE infants dataset classified by EEG technicians. Clearly, it is easy to find that the C3 group of mothers are more concentrated and older, whose age differs from 27 to 41. It suggests that the maternal age of subjects may be a factor affecting HIE infants classification. Additionally, in order to more intuitively display the EEG signals of different types of HIE infants, the paper shows the EEG signals of C1, C2 and C3 groups in Fig. 3. In Fig. 3(a), it can be seen that a continuous background pattern with mild asymmetric patterns and mild voltage depression, which is a typical feature of the C1 group EEG signal. In Fig. 3(b), it can be seen that discontinuous activity and clear asymmetry or asynchrony, which is a typical feature of the C2 group EEG signal. In Fig. 3(c), it can be seen that major discontinuous activity and continuous voltage depression, which is a typical feature of the C3 group EEG signal. Moreover, in the process of collecting EEG signals from infants with recording electrodes, some artifacts inevitably occur due to the young age of the subjects and the limitation of the professional level of the staff who collects the EEG signals. Clearly, the data collected in this study also contains artifacts, which affects the classification performance of HIE infants.

Table 1: Statistical information of 3 types of HIE infants dataset

Types	Features															
	Gender		Health state		Maternal age (year)				Gestational age (week)				Weight (kg)			
	Male	Female	With disease	Without disease	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std
C1	15	6	7	14	24.00	37.00	28.61	3.48	37.86	40.86	39.36	0.92	2.66	4.10	3.25	0.44
C2	20	9	5	24	21.00	39.00	30.84	5.47	35.00	41.43	39.08	1.40	2.40	5.90	3.24	0.70
C3	9	5	5	9	27.00	41.00	31.50	4.48	34.71	41.29	38.96	1.51	2.50	4.44	3.47	0.55
All groups	44	20	17	47	21.00	41.00	30.30	4.75	34.71	41.43	39.15	1.28	2.40	5.90	3.29	0.59

2.3 Feature extraction for EEG signals

In order to classify the HIE infants, the paper needs to extract feature information of EEG signals after preprocessing. The total features are outlined in Tab. 2 which mainly consists of two parts, one is time domain and frequency domain features of each 64 s EEG signals, and the other part contains the statistic features derived from 8 s epochs which are extracted from 64 s epoch data with 50% overlap.



(a) Mild asphyxia group. (b) Moderate asphyxia group. (c) Severe asphyxia group.

Figure 3: Types of EEG signals

In the first part, firstly, the paper extracts features from the perspective of the time domain based on previous research work [Glass, Nash, Bonifacio et al. (2011); Douglas-Escobar and Weiss (2015); Perlman and Shah (2011)], including Max, Min, Mean, Variance, Kurtosis [Hjorth (1970)] and Skewness [Mursalin, Zhang, Chen et al. (2017)]. Hence, a total of 48 statistical features are obtained from 8 channels. Then the line length [Esteller, Echauz, Tcheng et al. (2001)] and auto regressive (AR) model coefficients [Ge, Hou and Xiang (2007)] of each channel are also extracted to obtain 80 features. In order to improve the performance of the model, the correlation coefficients of the processed 8-channel EEG signals, δ (0.5 Hz-4 Hz), θ (4 Hz-8 Hz) and α (8 Hz-12 Hz) are extracted to measure the relationship between two channels and their related directions, and a total of 256 features are obtained. Furthermore, the paper extracts 8 features of Shannon entropy from each channel. Subsequently, the paper extracts features from the perspective of the frequency domain. The frequency domain features are limited up to 12 Hz as there is very little power in the frequency band higher than 13 Hz. As can be seen from Fig. 4 that when the frequency exceeds 12 Hz, the power spectral density (PSD) of all channels obtained is lower. Furthermore, Ahmed et al. [Ahmed, Temko, Marnane et al. (2016); Temko, Stevenson, Marnane et al. (2012)] have chosen power in the range of 0.1-12 Hz in the frequency domain. Hence, this paper also chooses the power of this interval. Moreover, Fig. 4 also shows significant differences between the left and right hemispheres, which indicates that there may be asymmetry between the left and right hemispheres. Therefore, extracting these features is beneficial to HIE classification. After passing a bandpass filter, the paper calculates the PSD to get the total power and the power of the 2 Hz width sub-bands. Meanwhile, in order to reduce the noise in spectral features, the paper calculates the energy of δ wave, θ wave and α wave as well as their power ratio: δ/θ , δ/α and θ/α respectively. Moreover, according to the definition of HIE grades in Murray et al. [Murray, Boylan, Ryan et al. (2009)] and the visualized EEG data in Fig. 3, it shows that asymmetry is an important feature of HIE infants classification. Therefore, the paper also calculates the asymmetry of 2 Hz width sub-bands power, which are calculated as $|\text{Left Channel}-\text{Right Channel}|/(\text{Left Channel}+\text{Right Channel})$. Left Channel means the power of electrodes in left hemisphere including FP1, T3, C3, O1 and Right Channel means the

power of electrodes in right hemisphere including FP2, T4, C4, O2 accordingly.

Table 2: Extracted features

Category	Domain	Feature parameters	Number
Type I (64 s epoch)	Time	Max, Min, Mean, Variance, Kurtosis, Skewness	8×6
		Line length	8
		AR model coefficient	8×9
		Correlation coefficient of original EEG signals, α , σ and θ waves	$8 \times 8 \times 4$
		Shannon entropy	8
		Total power (0.1 Hz-12 Hz)	8
		Power of 2 Hz width sub-bands (0.1-2 Hz, 1-3 Hz, ...10-12 Hz)	8×11
		Power of α , σ and θ waves	8×3
		Frequency Power ratio: δ/θ , δ/α and θ/α	8×3
		Asymmetry of 2 Hz width sub-bands power (0.1-2 Hz, 1-3 Hz, ..., 10-12 Hz), total power, δ , α and β power	8×15
Type II (8 s epoch)	Frequency	Max, Min, Mean, Variance of 2 Hz width sub-bands (0.1-2 Hz, 1-3Hz, ... , 10-12 Hz) power	$8 \times 4 \times 11$
Other	-	Gestational age (week)	1

In the second part, the 64 s epoch data will be divided into 8 s epoch with 4 s overlap, so as to obtain short-term features. In order to achieve this goal, the paper calculates Max, Min, Mean, Variance of each sub-band (0.1-2 Hz, 1-3 Hz, ... , 10-12 Hz) power and 352 statistical features are obtained. In addition, HIE abnormalities are often affected by gestational age, so this paper also considers gestational age as a feature.

2.4 SVM classifier

In proposed method, SVM is chosen to be used since it has shown good performance for HIE severity grading in neonatal EEG [Guler and Ubeyli (2007); Ahmed, Temko, Marnane et al. (2016); Temko, Doyle, Murray et al. (2015); Ansari, Cherian, Dereymaeker et al.

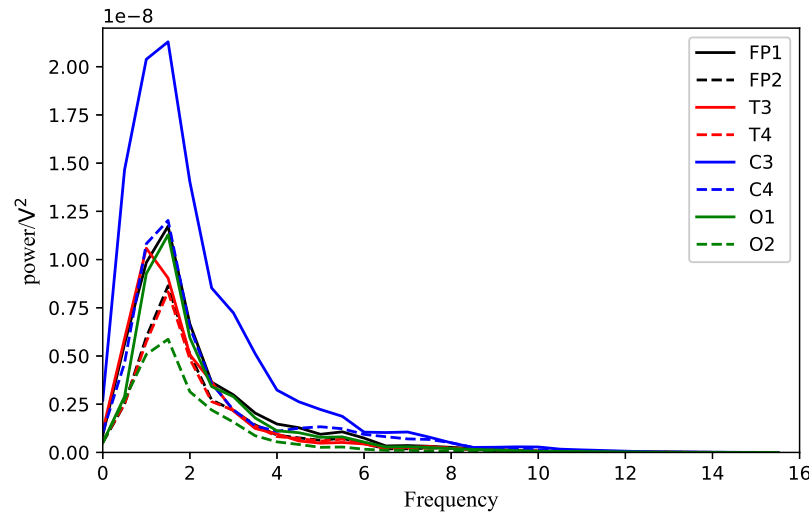


Figure 4: PSD distribution of the EEG signals in the frequency domain (64 s epoch)

(2016); Ahmed, Temko, Marnane et al. (2017)]. Clearly, SVM is a supervised learning method. Its main idea is to achieve the purpose of classification by finding a classification plane and separating the data from both sides of the plane. Since the problem is a three-category problem, the paper needs to build the SVM classifier into a suitable multi-class classifier. This paper uses the one-versus-one (OVO) classification method to classify HIE infants, which is to design a SVM classifier between any two types of samples, so k types of samples need to design $k(k-1)/2$ SVM classifiers. Fig. 5 shows the training process of the SVM. Considering that the SVM classifier needs to standardize the dataset, that is, the processed data conforms to a normal distribution with a mean of 0 and a variance of 1. According to Fig. 5, the classification process of this paper mainly includes two parts. The first part is to train the features of each channel to get the maximum probability of classification. The second part is to accumulate the maximum probability of the 8 channels, and then get the maximum probability of the overall classification of the 8 channels, and obtain the final classification results. Moreover, the paper uses the grid search method to set the penalty parameter C of SVM to be 20, kernel function of the SVM to be the radial basis function (RBF) and sigma component of the RBF to be 0.01.

3 Results

As stated above, this study treats the HIE grading problem as a multi-class classification problem in machine learning. And then, a multi-class classification model based on SVM is established.

3.1 Cross validation

In this study, a cross validation method is used to assess the performance of the method [Rodriguez, Perez and Lozano (2009)]. By using cross validation, not only can you get

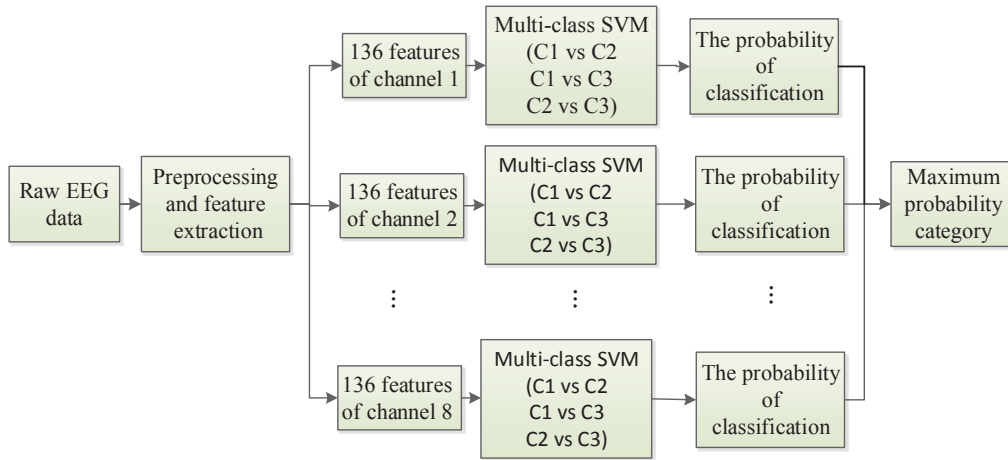


Figure 5: SVM training procedure

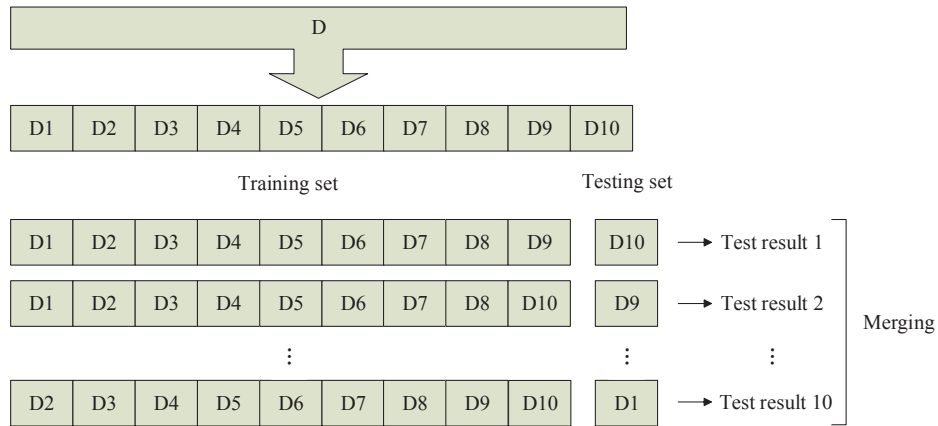


Figure 6: 10-fold cross validation

an estimate of the performance of the model, but you can also get the variance, which is important for judging the difference between the two models. This paper uses 10-fold cross validation method. Firstly, the original processed EEG data are randomly partitioned into 10 similar-sized subsets, and each subset is mutually exclusive. Of the 10 subsets, a single subset is retained as the validation data for testing the model, and the remaining 9 subsets are used as training data. The cross validation process is then repeated 10 times, with each of the 10 subsets used exactly once as the validation data. The 10 results are then merged to produce a single estimation. The advantage of this method over repeated random subsets is that all observations are used for both training and validation, and each observation is used for validation exactly once. The 10-fold cross training process is shown in Fig. 6.

In Fig. 6, D represents all EEG signal data (64 HIE infants with 420 epochs), and D_i is a subset of D which contains 42 epochs, they are similar in size and mutually exclusive. That is, $D = D_1 \cup D_2 \cup \dots \cup D_{10}$, $D_i \cap D_j = \emptyset$, ($i \neq j$).

3.2 Evaluation criteria

In order to evaluate the effectiveness of the proposed method, confusion matrix (CM) will be calculated to show the difference between the results given by the proposed method and marked by EEG technicians as shown in Eq. (1). S_{ij} represents the number of epochs that are marked to be class i and are classified to be class j . In the CM, C1 group, C2 group and C3 group are represented as 1, 2 and 3, respectively.

$$CM = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \quad (1)$$

In machine learning, the kappa is usually used to measure the accuracy of a model for multi-class classification problems, and its range is $[-1, 1]$. The higher the value, the higher the accuracy of the classification. And the calculation method of kappa is shown in Eq. (2).

$$k = \frac{p_o - p_e}{1 - p_e} \quad (2)$$

According to CM, p_o and p_e are calculated as follows:

$$P_o = \frac{\sum_{i=1}^3 S_{ii}}{\sum_{i=1}^3 \sum_{j=1}^3 S_{ij}} \quad (3)$$

$$P_e = \frac{\sum_{i=1}^3 (\sum_{j=1}^3 S_{ij} \sum_{j=1}^3 S_{ji})}{(\sum_{i=1}^3 \sum_{j=1}^3 S_{ij})^2} \quad (4)$$

Accuracy (AC), sensitivity (SE), specificity (SP) and precision (P) can also evaluate the performance of the proposed method as shown in Eq. (5) to Eq. (8), respectively.

$$AC = \frac{\sum_{i=1}^3 S_{ii}}{\sum_{i=1}^3 \sum_{j=1}^3 S_{ij}} \quad (5)$$

$$SE = \frac{S_{ii}}{\sum_{j=1}^3 S_{ij}} \quad (6)$$

$$SP = \frac{(\sum_{i=1}^3 S_{ii}) - S_{ii}}{\sum_{i=1}^3 \sum_{j=1}^3 S_{ij} - \sum_{i=1}^3 S_{ij}} \quad (7)$$

$$P = \frac{S_{ii}}{\sum_{i=1}^3 S_{ij}} \quad (8)$$

3.3 Result analysis

3.3.1 Performance of proposed method

In the 10-fold cross validation, all infants (64 infants with 420 epochs) are divided into 10 mutually exclusive subsets (approximately 6 infants per subset) by infant and the number of epochs for each subset is 42. According to Fig. 6, firstly, selecting D1 to D9 (378 epochs) with all features in Tab. 2 are input to train the SVM classifier and the remaining D10 (42 epochs) is used to test the performance of model. Secondly, D1 to D8 and D10 (378 epochs) are used to train the SVM classifier and D9 (42 epochs) is used to test the performance of model. Finally, this process will be repeated until each subset has been used as the testing data and an overall AC and kappa of the model will be reported. According to Eq. (5) and Eq. (2), the overall AC and kappa obtained by the SVM classifier are 79.5% and 0.69, respectively.

Tab. 3 presents the CM for the proposed method. It can be seen that 86 out of 420 epochs are misclassified. From Tab. 3, It can be clearly found that the EEG signals between adjacent groups can easily affect the AC of the proposed method. For example, when the actual data is C1 group, the proposed method misclassifies only 32 and 6 C1 group epochs to be C2 group epochs and C3 group epochs, respectively. This is also a direction that the paper can be improved in the future.

Table 3: CM of proposed model's output and actual assigned grade by EEG technicians

Actual Grade	Proposed Model's Output		
	C1	C2	C3
C1	90	32	6
C2	16	144	9
C3	9	14	100

Tab. 4 shows the classification performance for each HIE group. It can be found that the P of the C3 group is the best and the SE of the C2 group is the best, which shows that the proposed method can play a significant role in HIE infants classification. In medical diagnosis, doctors pay more attention to the SE of the model, which is also an advantage of the proposed method in this paper.

3.3.2 Comparison of different classifiers

In this part, 3 other classifiers including random forest (RF), logistic regression (LR) and naive bayes (NB) will be tested.

- RF classifier. RF proposed by Leo Breiman [Breiman (2001)] contains multiple decision trees and its output category is determined by the highest number of votes given by all trees. The essence of RF classifier is an improvement of the decision tree

Table 4: Performance of each group of proposed model

Types	P	SE	SP
C1	78.3%	70.3%	83.6%
C2	75.8%	85.2%	75.7%
C3	87.0%	81.3%	78.8%

algorithm. Multiple decision trees are merged together, and the establishment of each tree depends on an independently extracted sample. Each tree in the forest has the same distribution, and the classification error depends on each tree's classification ability and their correlation. In general, RF, an ensemble learning method, is highly competitive in both classification and regression tasks.

- LR classifier. LR is developed by statistician David Cox [Cox (1958)] in 1958 and can be considered as a special case of linear regression models. It usually uses a logistic function to model a binary dependent variable. Moreover, LR is also the most commonly used learning algorithm in machine learning and is often used as a benchmark classifier in classification tasks.
- NB classifier. NB is based on Bayes theory with strong (naive) independence assumptions between the features. With appropriate preprocessing, it is competitive in automatic medical diagnosis with more advanced methods [Pulmano and Estuar (2017)].

In order to compare the performance of the proposed method with the aforementioned classifier, all classifiers use the extracted features in Tab. 2 for experiments. Similarly, the paper uses the grid search method to obtain the number of the decision tree is 100 in RF. Likewise, LR selects L1 regularization and the performance of classifier is best when the regularization parameter C is set to be 1. Fig. 7 presents the comparison of kappa and AC between the proposed method and the RF classifier, LR classifier and NB classifier. It is obvious that the proposed method has obvious advantages in kappa and AC, especially in the kappa which can reach 0.69 in the proposed method, but can only reach 0.54, 0.50 and 0.18 in RF, LR and NB, respectively.

Moreover, since the AC and kappa of RF classifier are better than LR classifier and NB classifier, the paper compares the performance of the proposed method with RF classifier in the following analysis. Tab. 5 and Tab. 6 present the CM for obtaining the best AC and classification performance for each HIE group by RF classifier. It can be seen from Tab. 5 that the RF classifier is poorly identified in C3 group, but proposed method performs well in this group, which has great advantages in medical diagnosis. Similarly, comparing Tab. 4 with Tab. 6, it is easy to find that the proposed method has significant advantages in SE (C3 group) with 81.3% vs 62.6%. In clinical practice, it is important to identify serious patients, which can allow doctors to intervene early and reduce the risk of death.

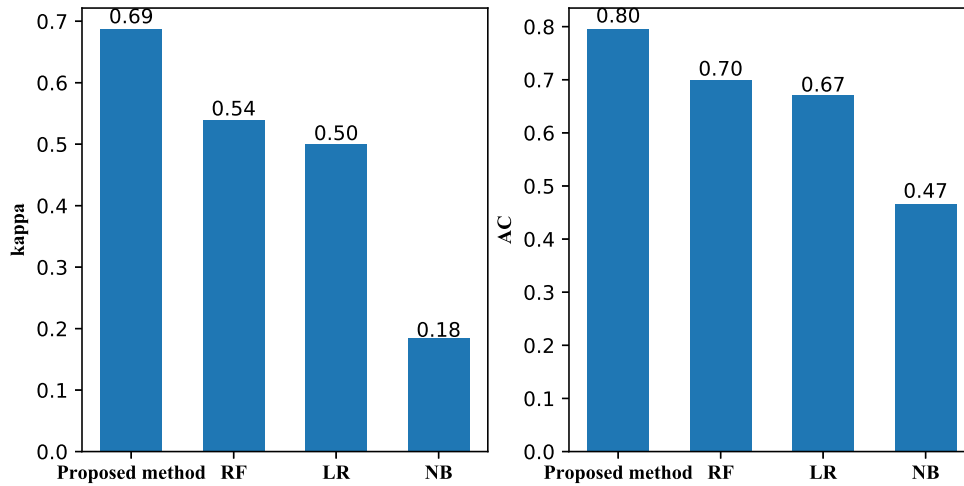


Figure 7: Comparison of different classifiers

Table 5: CM of RF model

Actual Grade	RF's Output		
	C1	C2	C3
C1	84	37	7
C2	30	133	6
C3	17	29	77

Table 6: Performance of each group of RF model

Types	P	SE	SP
C1	64.1%	65.6%	71.9%
C2	66.8%	78.7%	64.1%
C3	85.6%	62.6%	73.1%

3.3.3 Comparison of different features

Typically, different feature sets will have different effects on the proposed method. Therefore, this paper inputs different features into the proposed method for experiments, that is, all the proposed features in this paper, only the features of correlation coefficient and asymmetry, and the features of removing correlation coefficient and asymmetry. Fig. 8 shows that the proposed features have significant discriminating capability in HIE infants classification. In particular, the introduction of features (correlation coefficient and asymmetry) has significantly improves the performance of the proposed method in kappa and AC. If the paper removes these two features, the kappa and AC are reduced to be 0.46 and 65.0%, respectively, which is very bad and unacceptable. In fact, this also shows that these two features extracted in this paper have significant discriminating capability in HIE infants classification.

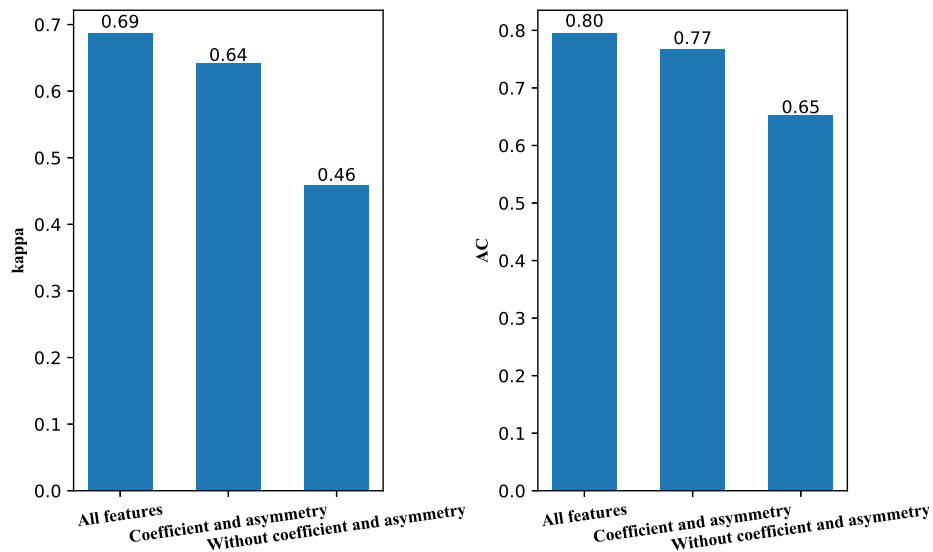


Figure 8: Comparison of different features in the proposed model

Tab. 7 and Tab. 8 present the CM for obtaining the best AC and classification performance for each HIE group by proposed method that removing correlation coefficient and asymmetry. It can be concluded from Tab. 7 and Tab. 8 that the classification of the proposed method in the C1 group and C3 group are seriously reduced due to the removal of the features correlation coefficient and asymmetry, which further indicates that the features proposed in this paper have important discriminating capability.

Furthermore, the paper uses recursive feature selection [Guyon, Weston, Barnhill et al. (2002)] method to obtain the most important 20 features which affect the performance of the proposed method, including Mean, Variance, Max of Type II-frequency, power of σ wave, Variance of Type I-time, power of α wave, total power of Type I-frequency, power

Table 7: CM of proposed model without correlation coefficient and asymmetry

Actual Grade	Proposed Model's Output		
	C1	C2	C3
C1	61	56	11
C2	19	139	11
C3	7	42	74

Table 8: Performance of each group of proposed model without correlation coefficient and asymmetry

Types	P	SE	SP
C1	70.1%	47.7%	72.9%
C2	58.6%	82.2%	53.8%
C3	77.1%	60.1%	67.3%

of 2 Hz width sub-bands of 1-3 Hz and 2-4 Hz, power of θ wave, power of 2 Hz width sub-bands of 5-7 Hz and 3-5 Hz, correlation coefficient of FP1 (δ wave) and T3, correlation coefficient of FP1 and T3, asymmetry of total power, asymmetry of 2 Hz width sub-bands power of 0.1-2 Hz and 3-5 Hz, power of 2 Hz width sub-band of 8-10 Hz, asymmetry of α power and asymmetry of 2 Hz width sub-band power of 9-11 Hz. It can be found that correlation coefficient and asymmetry play an important role in HIE infants classification.

Importantly, in order to evaluate the contribution of the extracted features to HIE infants classification, each type of feature in Tab. 2 (i.e., the feature of time domain in Type I, the feature of frequency domain in Type I and the feature of frequency domain in Type II) is used as input of the proposed method, and the evaluation performance is shown in Tab. 9. It can be seen from the Tab. 9 that the feature of Type I-time has obvious discrimination capability in HIE infants classification with AC and kappa reaching 74.8% and 0.61, respectively. The AC of Type I-time is 18.6% and 22.7% better than Type I-frequency and Type II-frequency, and the kappa of Type I-time is 0.3 and 0.38 better than Type I-frequency and Type II-frequency. Clearly, since Type I contains features correlation coefficient and asymmetry, it further indicates that features correlation coefficient and asymmetry have a significant contribution in the classification of HIE and that the contribution of correlation coefficient is slightly more than the contribution of asymmetry.

3.3.4 Comparison of different works

In fact, there are some other methods proposed for grading HIE. Tab. 10 presents a comparison on the results between the proposed method in this paper and other

Table 9: The contribution of the extracted features to HIE infants classification

Features	AC	Kappa
Type I-time	74.8%	0.61
Type I-frequency	56.2%	0.31
Type II-frequency	52.1%	0.23
Total	79.5%	0.69

Table 10: A comparison of the classification performance obtained by proposed method and others' work

Works	Methods	Features	Dataset	AC	Kappa
Korotchikova, Stevenson, Walsh et al. (2011)	Kruskal-Wallis testing with post hoc analysis and multiple linear regression	Skewness, kurtosis, discontinuity, amplitude, relative delta power, spectral edge frequency, fractal dimension, revised brain symmetry index, linear correlation coefficient	Continuous video-EEG data collected with NicoletOne EEG system (NOEEG)	87.5%	-
Stevenson, Korotchikova, Temko et al. (2013)	Multi-class linear discriminant classifier	Mean, standard deviation, skewness, kurtosis, covariance, relative delta power, IBI, symmetry, synchrony	Continuous video-EEG data collected with NicoletOne EEG system (NOEEG)	83%	0.76
Proposed method	Multi-class SVM classifier	Shown in Tab. 2	Children's Hospital and Affiliated Hospital of Fujian Medical University	79.5%	0.69

methods proposed in the previous literature. In Tab. 10, the studies of Korotchikova et al. [Korotchikova, Stevenson, Walsh et al. (2011)] and Stevenson et al. [Stevenson, Korotchikova, Temko et al. (2013)] on NOEEG dataset have a good performance on HIE infants classification. However, the description of the NOEEG dataset in korotchikova et al. [Korotchikova, Stevenson, Walsh et al. (2011); Stevenson, Korotchikova, Temko et al. (2013)] make it easy to find that the inclusion criteria of NOEEG dataset are more stringent, including 5-min Apgar score less than 5, initial capillary or arterial blood pH less than 7.1 mmol/l, and initial capillary or arterial blood lactate larger than 7 mmol/l, etc. However, the dataset in this paper does not have these requirements. In addition, the dataset of this paper is smaller than NOEEG dataset, which also limits the performance of the proposed method in the paper.

In addition, the kappa for each channel and the kappa obtained by accumulating the

maximum probability of each channel are shown in Fig. 9. The results in Fig. 9 show that if only a single channel is used for HIE infants classification, the T3 channel can achieve the best performance with a kappa of 0.57. However, if the HIE is classified by 8 channels, the kappa can reach 0.69, which significantly improves the classification performance of the proposed method.

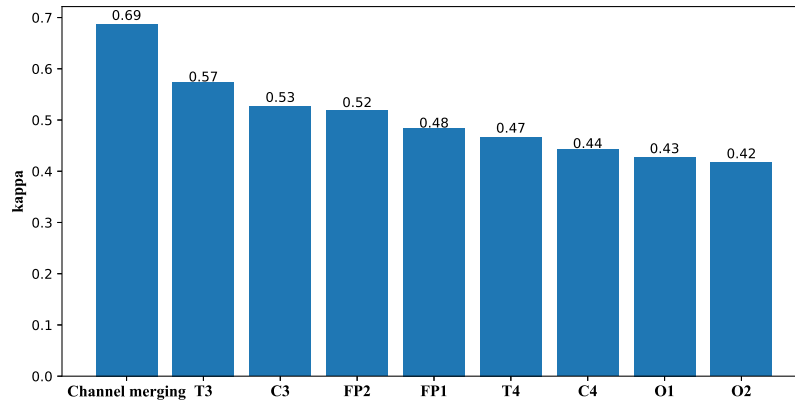


Figure 9: Kappa coefficient of different channel

However, there are certain areas that can be improved in the proposed automated grading method. Tab. 3 shows that the proposed method is more likely to misjudge predictions between adjacent classes, and that a future study could focus on HIE prediction between adjacent classes. In addition, since the EEG signals are image data, and deep learning has a very obvious advantage in the image field, a deep learning method can be introduced to classify HIE infants in the future [Litjens, Kooi, Bejnordi et al. (2017); Shen, Wu and Suk (2017); Suzuki (2017); Klang (2018)].

4 Discussion

The main contribution of this study is to propose a new method for HIE automatic classification, which combines the long-term and short-term features of EEG signals with SVM classifier. The proposed method has good performance in the real-world data provided by the Department of electrophysiology of Fujian Provincial Maternity and Children's Hospital and Affiliated Hospital of Fujian Medical University. Particularly, the feature of correlation coefficient and asymmetry have an obvious discriminating capability in the classification of HIE infants. Clearly, there are a large number of literatures and evidences showing that asymmetry plays an important role in the diagnosis of encephalopathy using EEG signals [Low, Mathieson, Stevenson et al. (2014); Knott, Mahoney, Kennedy et al. (2001)]. However, to the knowledge of the authors, there are few studies in the literature related to use correlation coefficient feature for HIE infants classification in EEG signals.

In Fig. 8, it shows that the features (correlation coefficient and asymmetry) extracted by the paper have a significant improvement on the classification accuracy and kappa of HIE infants classification. Even if the input of proposed method includes only the features of correlation coefficient and asymmetry proposed by the paper, there is still a larger improvement than using the remaining features. When using correlation coefficient and asymmetry, the AC and kappa of the model are 76.7% and 0.64, respectively. However, after removing these two features, the AC and kappa obtained by proposed method are only 65.2% and 0.46, which demonstrates the discriminating capability of these two features extracted in this paper. As expected, the proposed method and features have obvious advantages in AC, SE and SP.

However, the study is based on off-line analysis, which may be a limitation similar to other studies [Amsüss, Goebel, Jiang et al. (2013); Li, Wang, Yang et al. (2013)]. Furthermore, a relatively small number of infants with HIE are collected in this study, which also limit the performance of the proposed model. In the future, this study can collect more and more high-quality data for model training.

5 Conclusions

The treatment of HIE infants still has great challenges. In order to reduce the burden on doctors and improve the efficiency of diagnosis, this paper presents an automated method for HIE grading using EEG signals. Basing on the extraction of long-term features and short-term features of the EEG signals, SVM classifier is applied to classify HIE infants. Moreover, the features extracted in this paper, especially features correlation coefficient and asymmetry, have better discriminating capability in HIE infants classification. The proposed grading method has shown significant performance with an AC of 79.5% and the kappa of 0.69 by analyzing the EEG signals collected from 64 neonates. Therefore, the proposed method in the paper can provide great convenience for medical staff, and can also help non-experts to evaluate brain activity, which has great medical value.

Ethical standards: Study was approved by the Ethic committee of Fujian Provincial Maternity and Children's Hospital, Affiliated Hospital of Fujian Medical University and registered in Chinese Clinical Trail Registry (ChiCTR1800020438). All patients gave their informed consent prior to their inclusion in the study.

Acknowledgement: This work was supported in part by Natural Science Foundation of Zhejiang Province (grant numbers LGG19F030013 and LGF18F010007), Special Funds for Information Development in Shanghai (grant number 201801050) and Scientific research project of Zhejiang Provincial Department of Education (grant number Y201942165).

References

Ahmed, R.; Temko, A.; Marnane, W.; Lightbody, G.; Boylan, G. (2016): Grading hypoxic-ischemic encephalopathy severity in neonatal EEG using GMM supervectors and the support vector machine. *Clinical Neurophysiology*, vol. 127, no. 1, pp. 297-309.

- Ahmed, R.; Temko, A.; Marnane, W. P.; Boylan, G.; Lightbody, G.** (2017): Exploring temporal information in neonatal seizures using a dynamic time warping based SVM kernel. *Computers in Biology and Medicine*, vol. 82, pp. 100-110.
- Ambalavanan, N.; Carlo, W. A.; Shankaran, S.; Bann, C. M.; Emrich, S. L. et al.** (2006): Predicting outcomes of neonates diagnosed with hypoxemic-ischemic encephalopathy. *Pediatrics*, vol. 118, no. 5, pp. 2084-2093.
- Amsüss, S.; Goebel, P. M.; Jiang, N.; Graimann, B.; Paredes, L. et al.** (2013): Self-correcting pattern recognition system of surface emg signals for upper limb prosthesis control. *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1167-1176.
- Ansari, A. H.; Cherian, P.; Dereymaeker, A.; Matic, V.; Jansen, K. et al.** (2016): Improved multi-stage neonatal seizure detection using a heuristic classifier and a data-driven post-processor. *Clinical Neurophysiology*, vol. 127, no. 9, pp. 3014-3024.
- Awal, M. A.; Lai, M. M.; Azemi, G.; Boashash, B.; Colditz, P. B.** (2016): EEG background features that predict outcome in term neonates with hypoxic ischaemic encephalopathy: a structured review. *Clinical Neurophysiology*, vol. 127, no. 1, pp. 285-296.
- Breiman, L.** (2001): Random forests. *Machine Learning*, vol. 45, no. 1, pp. 5-32.
- Briatore, E.; Ferrari, F.; Pomero, G.; Boghi, A.; Gozzoli, L. et al.** (2013): Eeg findings in cooled asphyxiated newborns and correlation with site and severity of brain damage. *Brain and Development*, vol. 35, no. 5, pp. 420-426.
- Cox, D. R.** (1958): The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215-232.
- Douglas-Escobar, M.; Weiss, M. D.** (2015): Hypoxic-ischemic encephalopathy: a review for the clinician. *JAMA pediatrics*, vol. 169, no. 4, pp. 397-403.
- Dunne, J. M.; Wertheim, D.; Clarke, P.; Kapellou, O.; Chisholm, P. et al.** (2017): Automated electroencephalographic discontinuity in cooled newborns predicts cerebral MRI and neurodevelopmental outcome. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, vol. 102, no. 1, pp. F58-F64.
- Emilie, B. P.; Marie-Dominique, L.; Florence, F.** (2018): EEG in hypothermia in term neonates with HIE. *Neurophysiologie Clinique*, vol. 48, no. 4, pp. 236-237.
- Esteller, R.; Echauz, J.; Tcheng, T.; Litt, B.; Pless, B.** (2001): Line length: an efficient feature for seizure onset detection. *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, pp. 1707-1710.
- Ge, D. F.; Hou, B. P.; Xiang, X. J.** (2007): Study of feature extraction based on autoregressive modeling in egg automatic diagnosis. *Acta Automática Sinica*, vol. 33, no. 5, pp. 462-466.
- Glass, H. C.; Nash, K. B.; Bonifacio, S. L.; Barkovich, A. J.; Ferriero, D. M. et al.** (2011): Seizures and magnetic resonance imaging-detected brain injury in newborns cooled for hypoxic-ischemic encephalopathy. *Journal of Pediatrics*, vol. 159, no. 5, pp. 731-735.

- Guler, I.; Ubeyli, E. D.** (2007): Multiclass support vector machines for EEG-signals classification. *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 117-126.
- Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V.** (2002): Gene selection for cancer classification using support vector machines. *Machine Learning*, vol. 46, no. 1-3, pp. 389-422.
- Hjorth, B.** (1970): Eeg analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306-310.
- Klang, E.** (2018): Deep learning and medical imaging. *Journal of Thoracic Disease*, vol. 10, no. 3, pp. 1325-1328.
- Knott, V.; Mahoney, C.; Kennedy, S.; Evans, K.** (2001): Eeg power, frequency, asymmetry and coherence in male depression. *Psychiatry Research: Neuroimaging*, vol. 106, no. 2, pp. 123-140.
- Korotchikova, I.; Stevenson, N.; Walsh, B.; Murray, D.; Boylan, G.** (2011): Quantitative EEG analysis in neonatal hypoxic ischaemic encephalopathy. *Clinical Neurophysiology*, vol. 122, no. 8, pp. 1671-1678.
- Li, Z.; Wang, B.; Yang, C.; Xie, Q.; Su, C. Y.** (2013): Boosting-based emg patterns classification scheme for robustness enhancement. *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 545-552.
- Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F. et al.** (2017): A survey on deep learning in medical image analysis. *Medical Image Analysis*, vol. 42, pp. 60-88.
- Löfhede, J.; Thordstein, M.; Löfgren, N.; Flisberg, A.; Rosa-Zurera, M. et al.** (2010): Automatic classification of background EEG activity in healthy and sick neonates. *Journal of Neural Engineering*, vol. 7, no. 1, 016007.
- Low, E.; Mathieson, S. R.; Stevenson, N. J.; Livingstone, V.; Ryan, C. A. et al.** (2014): Early postnatal EEG features of perinatal arterial ischaemic stroke with seizures. *PLoS One*, vol. 9, no. 7, e100973.
- Mariani, E.; Scelsa, B.; Pogliani, L.; Introvini, P.; Lista, G.** (2008): Prognostic value of electroencephalograms in asphyxiated newborns treated with hypothermia. *Pediatric Neurology*, vol. 39, no. 5, pp. 317-324.
- Matić, V.; Cherian, P. J.; Jansen, K.; Koolen, N.; Naulaers, G. et al.** (2016): Improving reliability of monitoring background EEG dynamics in asphyxiated infants. *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 5, pp. 973-983.
- Murray, D. M.; Boylan, G. B.; Ryan, C. A.; Connolly, S.** (2009): Early EEG findings in hypoxic-ischemic encephalopathy predict outcomes at 2 years. *Pediatrics*, vol. 124, no. 3, pp. 459-467.
- Mursalin, M.; Zhang, Y.; Chen, Y.; Chawla, N. V.** (2017): Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. *Neurocomputing*, vol. 241, pp. 204-214.
- Perlman, M.; Shah, P. S.** (2011): Hypoxic-ischemic encephalopathy: challenges in outcome and prediction. *Journal of Pediatrics*, vol. 158, no. 2, pp. 51-54.

- Pulmano, C. E.; Estuar, M. R. J. E.** (2017): A multi-model approach in developing an intelligent assistant for diagnosis recommendation in clinical health systems. *Procedia Computer Science*, vol. 121, pp. 534-541.
- Rodriguez, J. D.; Perez, A.; Lozano, J. A.** (2009): Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569-575.
- Schindler, K.; Leung, H.; Elger, C. E.; Lehnertz, K.** (2006): Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial EEG. *Brain*, vol. 130, no. 1, pp. 65-77.
- Shah, P. S.** (2010): Hypothermia: a systematic review and meta-analysis of clinical trials. *Seminars in Fetal & Neonatal Medicine*, vol. 15, no. 5, pp. 238-246.
- Shen, D.; Wu, G.; Suk, H. I.** (2017): Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, vol. 19, pp. 221-248.
- Stevenson, N.; Korotchikova, I.; Temko, A.; Lightbody, G.; Marnane, W. et al.** (2013): An automated system for grading EEG abnormality in term neonates with hypoxic-ischaemic encephalopathy. *Annals of Biomedical Engineering*, vol. 41, no. 4, pp. 775-785.
- Suzuki, K.** (2017): Overview of deep learning in medical imaging. *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257-273.
- Temko, A.; Doyle, O.; Murray, D.; Lightbody, G.; Boylan, G. et al.** (2015): Multimodal predictor of neurodevelopmental outcome in newborns with hypoxic-ischaemic encephalopathy. *Computers in Biology and Medicine*, vol. 63, pp. 169-177.
- Temko, A.; Stevenson, N.; Marnane, W.; Boylan, G.; Lightbody, G.** (2012): Inclusion of temporal priors for automated neonatal EEG classification. *Journal of Neural Engineering*, vol. 9, no. 4, 046002.
- Tich, S. N. T.; Derambure, P.; Lamblin, M. D.; Betrouni, N.; Bourriez, J. L. et al.** (2018): Quantitative EEG analysis in early neonatal hypoxic ischaemic encephalopathy (HIE) to grade the severity and predict long-term prognosis. *Neurophysiologie Clinique*, vol. 48, no. 4, pp. 236-237.

