

A fair comparison of the EEG signal classification methods for alcoholic subject identification

Author

Awrangjeb, M, Rodrigues, JDC, Stantic, B, Estivill-Castro, V

Published

2020

Conference Title

2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)

Version

Accepted Manuscript (AM)

DOI

[10.1109/IVCNZ51579.2020.9290683](https://doi.org/10.1109/IVCNZ51579.2020.9290683)

Rights statement

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Downloaded from

<http://hdl.handle.net/10072/404377>

Griffith Research Online

<https://research-repository.griffith.edu.au>

A fair comparison of the EEG signal classification methods for alcoholic subject identification

Mohammad Awrangjeb¹, Jardel das C. Rodrigues², Bela Stantic¹ and Vladimir Estivill-Castro¹

¹*Institute for Integrated and Intelligent Systems, Griffith University, Nathan, QLD 4111 Australia*

²*Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Fortaleza, Ceará, Brazil*

Abstract—The electroencephalogram (EEG) signal, which records the electrical activity in the brain, is useful for assessing the mental state of the alcoholic subject. Since the public release of an EEG dataset by the University of California, Irvine, there have been many attempts to classify the EEG signals of ‘alcoholic’ and ‘healthy’ subjects. These classification methods are hard to compare as they use different subsets of the dataset and many of their algorithmic settings are unknown. The comparison of their published results using the inconsistent and unknown information is unfair. This paper attempts a fair comparison by presenting a level playing field where a public subset of the dataset is employed with known algorithmic settings. Two recently proposed high performing EEG signal classification methods are implemented with different classifiers and cross-validation techniques. While compared it is observed that the wavelet packet decomposition method with the Naïve Bayes classifier and the k -fold cross validation technique outperforms the other method.

Index Terms—EEG, alcoholic, alcoholism, alcohol-use disorder, signal decomposition, feature extraction, classification

I. INTRODUCTION

More than 3 million people died as a result of the harmful use of alcohol in 2016, according to a report released by the World Health Organisation (WHO) in September 2018, and overall, the harmful use of alcohol causes more than 5% of the global disease burden. An estimated 2.3 billion people were drinkers in 2018. Yet, far too many people, their families and communities suffered the consequences of the harmful use of alcohol through violence, injuries (accidents), mental health problems and diseases like cancer and stroke [1]. Excessive drinking of alcohol leads to behavioral, cognitive, and physiological changes in the subject. It affects all organs of the human body, mainly the central and peripheral nervous systems, affecting their functions, and causing aberrations in normal brain activities. These effects vary from person to person and are related to, among others, the amount of alcohol ingested, age and nutrition [2].

Therefore, the diagnosis of alcoholism, or alcohol-use disorder, is of great importance not only due to its effects on the individual and society but also on the cost to the national health system. Recently, the electroencephalography (EEG) signal (see Fig. 1) has been employed to evaluate the brain injuries related to alcohol abuse [3]. It is also used for the identification of epileptic [4], driving fatigue [5] and sleep spindle [6] symptoms while compared to the healthy subjects.

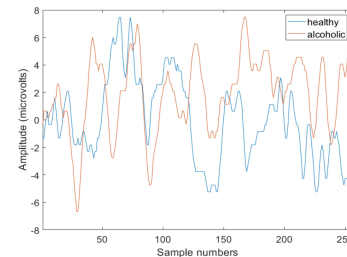


Fig. 1. Alcoholic and healthy EEG signals (1 second, 256 Hz).

For instance, Rodrigues et al. [3] apply the machine learning algorithms to separate alcoholic humans from the healthy ones and obtain high classification rates. Liu et al. [5] propose a prediction system to address the EEG regression problem in brain dynamics for driving fatigue. Goldstone et al. [6] analyse the EEG signals to conclude that the sleep spindle may serve as a marker of adolescent development.

The EEG signal, which can depict the state of the brain, is an ensemble of various neuronal activity recorded from different scalp regions having different characteristics and very low magnitude in microvolts. These factors make human interpretation difficult and time consuming to analyse this signal. Moreover, these highly varying EEG signals are susceptible to inter/intra variability errors [2]. Therefore, several Computer-Aided Diagnosis (CAD) methods have been proposed in recent years. They follow almost the same sequence of steps shown in Fig. 2. The EEG signals of the healthy and alcoholic subjects are first decomposed into different frequency bands (rhythms) such as alpha (α), beta (β), gamma (γ), delta (δ) and theta (θ). While some proposed methods select signals with specific EEG rhythms [7], others use all the relevant rhythms [3]. Many features such as max, mean, entropy, etc. are then estimated from each selected band signal. Sometimes a feature reduction technique such as the principal component analysis (PCA) is applied to obtain significant features [2]. The features are then divided into two sets: training and testing. One of the two cross-validation techniques, which are holdout [3] and k -fold [7], is applied to generate these sets. In the holdout validation, p percent of the input (observations, their features) data are randomly chosen and held out for testing. The rest of the data are used to generate a trained model. The data held for testing are then tested against the model. In the k -fold validation, the input data is randomly divided into k subsets and the training

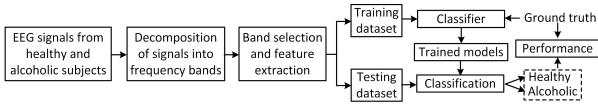


Fig. 2. General flow diagram of the EEG signal classification.

and testing phases are repeated for k times. In each time, one of the subsets is tested against the trained model generated using the remaining $k-1$ subsets of data. Obviously, the k -fold validation takes much longer time but is found more robust than the holdout validation. In the testing phase, each testing observation is classified into either ‘alcoholic’ or ‘healthy’ classes. The ground truth is used during the training and testing phases in both the validation techniques. The performance of an EEG classification method is estimated using a number of metrics such as accuracy, sensitivity, precision and specificity [2], [3], [7].

II. DATASETS

Since the collection of the EEG data requires the human subjects, it is time consuming. Therefore, there are not many datasets available publicly. The only publicly available dataset is from the University of California, Irvine, Knowledge Discovery in Databases (UCI KDD) archive [8]. It contains measurements of event-related potentials from 64 electrodes placed on the scalp sampled at 256 Hz (3.9 milliseconds epoch) for one second. There are 122 subjects and each subject completed 120 trials. Each subject was fitted with a 61-lead electrode cap (ECI, Electrocap International). Each subject was exposed to either a single stimulus (S_1) or two stimuli (S_1 and S_2) which were pictures of objects chosen from the 1980 Snodgrass and Vanderwart picture set. When two stimuli were shown, they were presented in either a matched condition where S_1 was identical to S_2 ($S_1 = S_2$) or in a non-matched condition where S_1 differed from S_2 ($S_1 \neq S_2$) [11].

Due to its availability, the majority of the researchers use the UCI KDD dataset, which has been released into three subsets: small, large and full. The small dataset contains data for 2 subjects only, one alcoholic and one healthy. For each of the 3 matching paradigms (S_1 , $S_1 = S_2$ and $S_1 \neq S_2$), there are 10 runs (trials) for each subject. The large dataset contains two sets (training and testing) of data. For the ‘training set’, there are 10 alcoholic and 10 control subjects, with 10 runs per subject per paradigm. The ‘testing set’ used the same 10 alcoholic and 10 control subjects as with the ‘training set’, but with 10 out-of-sample runs per subject per paradigm. The full dataset contains all data for 122 subjects. However, there are sometimes fewer trials, i.e., less than 120 trials, available for many subjects.

Among the other datasets, Mumtaz et al. [12] present a dataset from the University Malaya Medical Center but this dataset is not publicly available.

III. REVIEW

Table I shows the existing methods that mostly follow the general steps shown in Fig. 2 and use the UCI KDD dataset

for accuracy estimation.

Rodriguez et al. [3] apply the wavelet packet decomposition (WPD) to each EEG input signal using four different filters (e.g., Biorthogonal splines). All sub-bands (α , β , γ , δ and θ), including the noise (high-pass, 64-128 Hz), are reconstructed after the decomposition. Seven different features are estimated for each reconstructed signal and five different classifiers (including support vector machine - SVM and Naïve Bayes) are applied to classify the signals using the holdout cross-validation technique. After presenting an extensive experimental results the authors show that the combination of the Biorthogonal wavelet filter with two vanishing moments and the Naïve Bayes classifier with $p = 25\%$ percent testing signals offers the best result.

Taran and Bajaj [7] first apply the empirical mode decomposition (EMD) to convert the non-stationary EEG signal into stationery and orthogonal basis functions, which are known as the intrinsic mode functions (IMFs). Then, they apply the Hilbert-Huang transform to convert the real-valued IMFs into complex-valued IMFs, from where the instantaneous frequencies (IFs) that help separate the individual EEG sub-bands are estimated. They experimentally show that the low frequency such as δ and θ sub-bands offer better distinguishing characteristics between alcoholic and healthy subjects than the other sub-bands. Five statistical features are estimated from each chosen sub-band. Experimental results show that the extreme learning machine performs better than the least squares-SVM (LS-SVM) with the 10-fold cross-validation technique. It is reported that the EMD decomposition takes longer computational time than the wavelet decomposition [3].

Patidar et al. [2] decompose the EEG signal into low-pass and high-pass sub-band signals using the fourth level tunable-Q wavelet transform with the Daubechies filter featuring two vanishing moments. The correntropy (a nonlinear correlation) features are extracted from the detail sub-band and the PCA is applied to find 3 significant features, which are classified using the LS-SVM. Finally, they establish an alcoholism risk index using the three clinically significant PCA features to discriminate the given classes by means of a single number. Anuragi and Sisodia [10] propose a similar method.

Upadhyay et al. [9] choose the two specific channels (C_3 and C_4) only (out of available 64 channels). As a pre-processing they filter out the high-pass components using the Butterworth filter and keep only the low-pass components (δ , θ , α and β) from the raw EEG signals. They, then, decompose the filtered signals using the continuous wavelet transform with four mother wavelet functions. Four statistical parameters, i.e., mean, standard deviation, skewness and kurtosis are calculated from the selected wavelet coefficients of each mother wavelet as features of the input EEG signals. Out of the three classifiers they use, the SVM performs the worst in terms of accuracy but the best in terms of computational time, while the random forest performs the best in terms of accuracy.

Acharya et al. [11] estimate 7 features, including approximation entropy and sample entropy, from the original EEG signals without any signal decomposition step being followed.

TABLE I
SUMMARY OF THE METHODS USING THE UCI KDD DATASET [8].

Methods	Decomposition	Sub-bands	Features	F. reduction	Cross-validation	Classifiers	Subjects			Signals	S. length	Accuracy(%)
							Total	Alcoholic	Healthy			
Rodrigues [3]	Wavelet packet	All	7	-	Holdout	5	20	10	10	600	256	99.6
Taran [7]	Empirical mode	δ, θ	5	-	10-fold	2	-	-	-	240	2048	97.9
Patidar [2]	Tunable-Q wavelet	2	7	PCA	10-fold	1	-	-	-	240	2048	97.0
Upadhyay [9]	Continuous wavelet	$\delta, \theta, \alpha, \beta$	4	-	10-fold	3	-	-	-	140	256	100
Anuragi [10]	Tunable-Q wavelet	2	7	PCA	10-fold	1	-	-	-	-	256	97.1
Acharya [11]	-	-	7	t -test	3-fold	1	-	-	-	240	2048	91.7

They apply the statistical t -test to find highly discriminating features. A 3-fold cross-validation is used and the SVM classifier with a first order polynomial kernel is employed for classification.

IV. CONTRIBUTION

Although most of the methods for the alcoholic EEG signal classification show high accuracy on the KDD dataset [8], as summarised in Table I, many of them omit the number of subjects used, the number of alcoholic and healthy subjects and the number of signals evaluated. If the full KDD dataset is not used, for a fair comparison it is important to know which of the subjects and their trials are used. In addition, most of the works use complex techniques for the extraction of the attributes and classification [3]. Many detail algorithmic settings are mostly kept unknown to the audience. So, it is unfair to compare the methods using their published results.

For a fair comparison, this paper presents a level playing field and compares two recently proposed promising methods: Rodrigues et al. [3] and Taran and Bajaj [7]. After the step-by-step descriptions of these methods, we use the training subset of the large KDD dataset to evaluate and compare their performance. Extensive experimental results, running times and comparative discussions are presented.

V. PERFORMANCE STUDY

In this study, we use the ‘training set’ from the large UCI dataset [8] where there are 20 subjects, 10 in each of the alcoholic and healthy categories. All 30 trials for each subject are used. Therefore, we have a total of 600 EEG files. Each file contains 64 channels of EEG signals and each channel signal is recorded for 1 sec at 256 Hz. So, the signal length is $N = 256$ for each channel signal.

A. Methods

For each EEG channel signal, Rodrigues et al. [3] apply a five-level WPD using a filter (Biorthogonal splines, Coiflets, Daubechies or Symlets). At each level, the signal is decomposed (downsampled) into approximate and detail signals. The approximate signal in each level is decomposed again to find the approximate and detail signals in the next level. After the fifth level decomposition, the reconstructed detail signals from all five levels and the reconstructed approximate signal from the fifth level are considered as the designated EEG rhythms as described in Table II.

TABLE II
THE DECOMPOSED EEG RHYTHMS AND THEIR FREQUENCIES (Hz).

Rhythms	Rodrigues et al. [3]			Taran and Bajaj [7]	
	Levels	Signals	Frequency	Chosen	Frequency Chosen
Noise	1	Detail	64-128	✓	> 60
Gamma	2	Detail	32-64	✓	30-60
Beta	3	Detail	16-32	✓	13-30
Alpha	4	Detail	8-16	✓	8-13
Theta	5	Detail	4-8	✓	4-8
Delta	5	Approx	0-4	✓	0-4

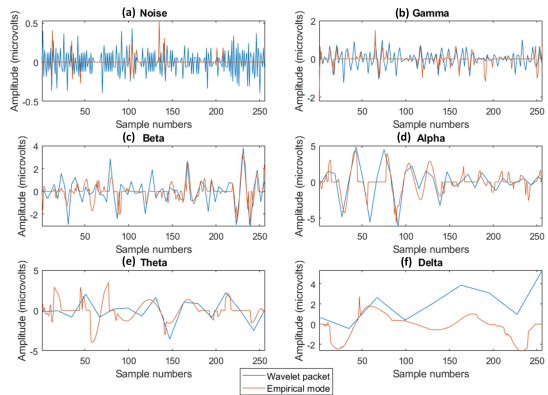


Fig. 3. Comparing the EEG rhythms generated using the wavelet packet [3] and empirical mode [7] decompositions from the alcoholic signal in Fig. 1.

In contrast, Taran and Bajaj [7] apply the EMD to decompose each EEG channel signal into n stationary IMFs $M = \{m_{i,j}\}$, where $1 \leq i \leq N$ and $1 \leq j \leq n$. The IMFs represent the well-defined frequency range in amplitude-frequency modulated components. The IFs $F = \{f_{i,j}\}$ are subsequently extracted from IMFs by the simple application of the Hilbert-Huang transform to M , thereby providing a multiscale analysis of the signal’s nonlinear phases. There are n columns in both M and F . Column j in M represents an IMF and the corresponding Column j in F represents its IFs. The i -th sample of a rhythm is finally estimated by summing the values in the i -th row of M if the corresponding frequencies in the i -th row of F fall into the given frequency range shown in Table II.

Fig. 3 compares the EEG rhythms generated after WPD [3] and EMD [7] decompositions. In many samples an EMD rhythm may have the zero value when there are no values in M within the given frequency range in F . In contrast, in most samples an WPD rhythm has a non-zero value as the signal

is reconstructed after the wavelet decomposition.

Thereafter, while Rodrigues et al. [3] select all rhythms including the noisy band, Taran and Bajaj [7] select only the two low pass rhythms namely θ and δ . From each selected band, the WPD method estimates seven features, i.e., min, max, power value, mean, standard deviation, absolute mean and entropy. In contrast, the EMD method estimates five features, i.e., mean absolute difference, interquartile range, covariance, entropy and negative entropy, from each selected band. This brings to the total number of features per subject in the WPD method is 42×64 , compared to 10×64 in the EMD method, where 64 is the number of EEG channels. For classification, while the WPD method adopts the holdout cross-validation, the EMD method adopts the 10-fold cross-validation. After using a number of classifiers, the Naive Bayes (NB) and the Extreme Learning Machine (ELM) are found to be the best performing classifiers by the WPD and EMD methods, respectively.

B. Experiments

The performance of the methods is estimated using four well-known metrics sensitivity (correctly classified alcoholic to the total classified alcoholic, also known as *true positive rate*), specificity (correctly classified healthy to the total classified healthy, also known as *true negative rate*), positive predictive value (PPV, also known as *precision*, correctly classified alcoholic to the total alcoholic subjects) and accuracy (correctly classified EEG signals to the total EEG signals) [3], [11]. The experimental results presented in this paper are the average of 100 runs over the chosen dataset. In each run, the training and testing samples are randomly chosen and cross-validated.

1) *The WPD Method:* The WPD method [3] follows the holdout cross-validation. Fig. 4 shows its performance under different training size using the NB classifier. The performance is averaged after two, three and four vanishing moments being used with each wavelet filter. In terms of sensitivity its performance does not change much, but there is a considerable fall in performance in terms of specificity, precision and accuracy as the training size decreases. Although it offers higher performance at a larger training size, it shows higher standard deviations which mean its performance is spread out from its average performance. Consequently, it is hard to decide a specific training size, which is different from Rodrigues et al. [3] where the authors find the best performance at 75% training size. In terms of the wavelet filter, although the Biorthogonal splines show the best performance in terms of specificity and precision, the Daubechies show the best in terms of sensitivity. Overall, at 75% training size the Biorthogonal splines offer the best performance, which is claimed in Rodrigues et al. [3] too.

Fig. 5 compares its performance under different vanishing moments of the wavelet filters. In each graph, the average performance line shows the mean of sensitivity, specificity, precision and accuracy for each filter. Although the Daubechies and Symlets filters perform the best with the two vanishing moments, the Biorthogonal filters show the best performance

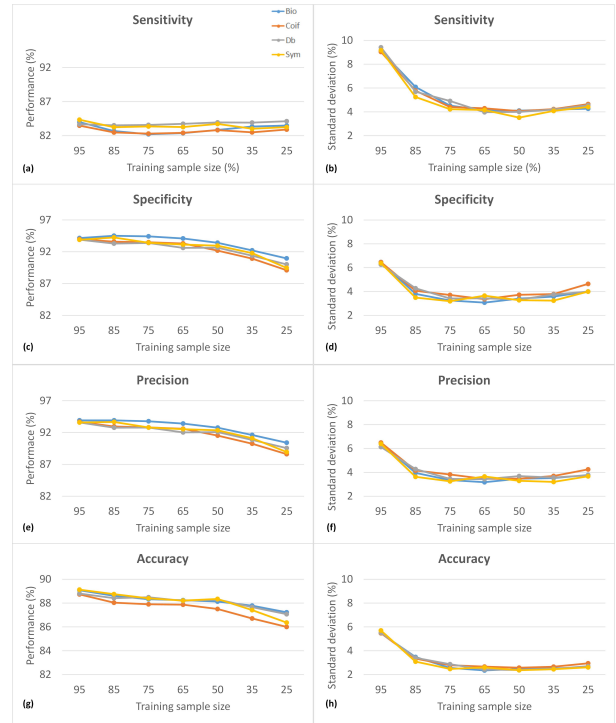


Fig. 4. Performance of Rodrigues et al. [3] under different training sample sizes using the Naive Bayes classifier.

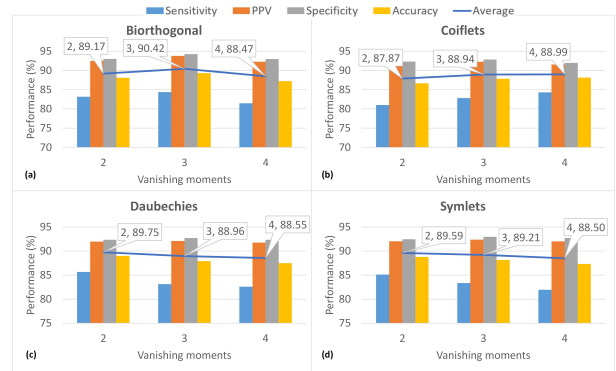


Fig. 5. Performance of Rodrigues et al. [3] under different vanishing moments of wavelet filters using the Naive Bayes classifier.

with the three vanishing moments and the Symlets filters work the best with the four vanishing moments. This contradicts the performance reported in Rodrigues et al. [3] where the Biorthogonal filters with the two vanishing moments work the best. However, since for each filter the performance difference is not significant under different vanishing moments, any vanishing moments (2, 3 or 4) can be used in general.

Fig. 6 compares its performance under different cross-validation techniques when the WPD method is set using the Biorthogonal filters with two vanishing moments. Its performance is more stable using the k -fold technique than using the holdout technique. With the holdout validation, its performance remains almost the same at a high training sample size (95 to 65%), however falls gradually with low training size

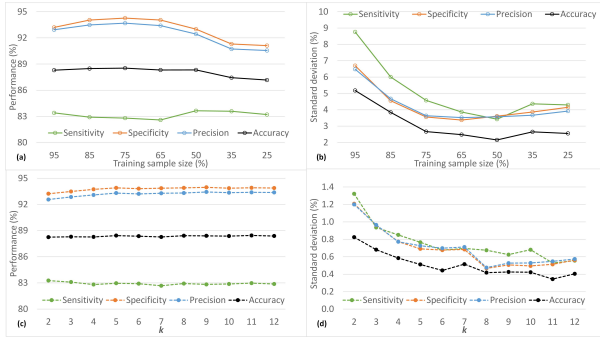


Fig. 6. Performance of Rodrigues et al. [3] under different cross-validation techniques using the Naïve Bayes classifier: (a-b) holdout and (c-d) k -fold.

(50 to 25%). Therefore, again the training size of 75% could be a good choice using the holdout validation. In contrast, with the k -fold validation, its performance remains almost the same for all k values. To make a trade-off between the computational time and performance $k = 5$ is a good choice. Moreover, the k -fold technique offers much less standard deviations than the holdout technique, which means using the k -fold technique most of the performance in terms of sensitivity, specificity, precision and accuracy are close to the average performance.

2) *The EMD Method*: The EMD method [7], instead of using all features, uses only the low-pass rhythm features. Fig. 7 compares its performance under different training size using the NB (a-b) and the ELM (c-d) classifiers. For both classifiers, the performance is shown when all and the selected features are employed. Using the ELM classifier, the ‘all features’ case outperforms the ‘selected features’ case in terms of sensitivity, specificity, precision and accuracy. This observation is opposite to the one reported in Taran and Bajaj [7]. Also, using the NB classifier, the ‘selected features’ case outperforms the ‘all features’ case only in terms of sensitivity. The same trend has been observed with the k -fold cross-validation technique (see Fig. 8). Therefore, the selection of specific rhythm features remain a mystery like many other earlier studies. It is also observed that using the ELM classifier the performance significantly drops when the training size is 50%. The standard deviations in Fig. 7 present a similar trend to those in Fig. 4, i.e., a higher performance is observed at a larger training size but at the cost of a higher standard deviation. Consequently, like the WMD method, the setting of the training size at 75% for the EMD method could be a good option to put a control over the performance uncertainty.

The EMD method [7] actually uses the k -fold cross-validation technique with $k = 10$. Fig. 8 shows its performance under different k values. Contrasting with the holdout cross-validation that offers a higher performance at a larger training size, the k -fold cross-validation may also offer a higher performance at a bigger k value. Because a high k value makes more partitions in the input observations resulting in a larger training size but a smaller testing size. However, a higher k value makes the training step slower as the number of trained models increases. Thus, there needs to be a trade-off between

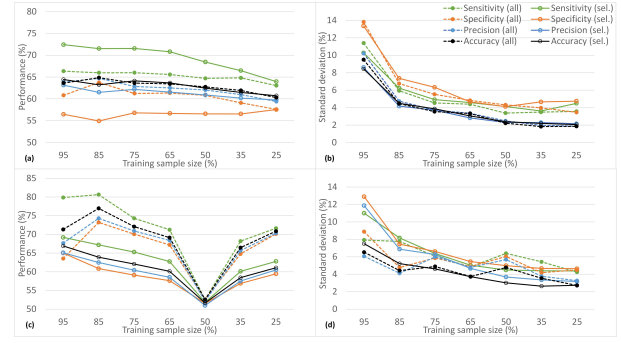


Fig. 7. Performance of Taran and Bajaj [7] under different training sample size using (a-b) Naïve Bayes classifier and (c-d) extreme learning machine.

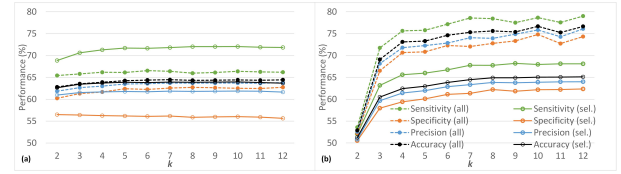


Fig. 8. Performance of Taran and Bajaj [7] under different k values in k -fold cross-validation using (a) Naïve Bayes classifier and (b) extreme learning machine.

the performance and the value of k .

When the NB classifier is used, there is a slight increase in performance for $k = 2$ to 4, but thereafter its performance remains almost the same although k increases. In contrast, when the ELM classifier is used, there is a significant increase in performance for $k = 2$ to 4 and after that its performance gradually increases till $k = 8$. For $k = 8$ to 12 its performance again remains almost the same. Therefore, $k = 5$ and 8 are good choices with the NB and ELM classifiers, respectively, to address the issue of performance and efficiency. The lower value of k for the NB than the ELM also shows that the NB is more robust than the ELM.

3) *Computational Time*: Table III compares the average computational times of the two methods. For signal decomposition and feature extraction the 64 EEG channel signals of a trial file of a subject are employed and the results are averaged after 100 runs. Since, all the EEG files (trials, signals) are of the same size this is a fair comparison. The training and testing times under the holdout and k -fold cross-validation techniques are averaged for the whole chosen dataset after running for different training sample sizes and k values.

In general, the decomposition and training times are higher than the feature extraction and testing times, respectively. For decomposition and feature extraction the EMD method is more than 6 times slower than the WPD method. However, the former is faster than the latter in training and testing steps: 6 times using the holdout and 1.2 times using the k -fold cross-validation techniques.

Fig. 9 compares the training and testing times under different combinations of the two methods, two classifiers and two cross-validation techniques. Overall, the training time is always higher than the testing time. Under the holdout cross-

TABLE III
AVERAGE COMPUTATIONAL TIMES (IN SECONDS) OF WAVELET PACKET DECOMPOSITION AND EMPIRICAL MODE DECOMPOSITION METHODS.

Methods	Processing		Holdout		k -fold	
	Decompose	Feature	Train	Test	Train	Test
WPD [3]	0.1114	0.0427	0.70160	0.1208	1.1234	0.0944
EMD [7]	0.6981	0.2650	0.07700	0.0133	0.7644	0.0735

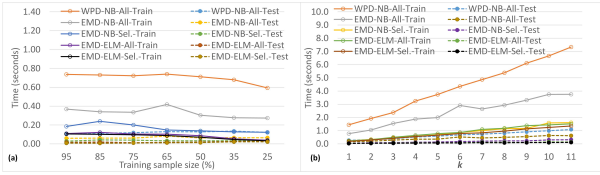


Fig. 9. Comparing computational times (in seconds) of wavelet packet decomposition [3] and empirical mode decomposition [7] methods using the Naïve Bayes (NB) and extreme learning machines (ELM) classifiers under (a) holdout and (b) k -fold cross-validation techniques.

validation while the training time slowly decreases with the decrease of the training sample size, the testing time increases insignificantly. The same trend is observed under the k -fold cross-validation where the training time considerably rises with the increase of k values. Again, the NB classifier is much slower than the ELM classifier and the holdout cross-validation is faster than the k -fold cross-validation.

4) *Comparisons and Discussions:* Comparing the results in Figs. 4 and 7, the WPD method offers much better performance than the EMD method. When the NB classifier is used with the holdout cross-validation technique, the performance ranges are (WPD vs EMD): [82,85] vs [63,73] for sensitivity; [89,94] vs [55,64] for specificity; [89,94] vs [59,63] for precision; and [86,89] vs [60,65] for accuracy (all in percentages).

Furthermore, the EMD method causes higher standard deviations than the WPD method. While the EMD method shows the highest standard deviations (9.5%) in sensitivity, the EMD method shows the highest standard deviations (11.5%) in specificity. Both methods show almost the same lowest standard deviations (2%) in accuracy measure.

For cross-validation, the k -fold technique should be chosen ahead of the holdout technique. Finally, in terms of classifiers, the NB performs much better than the ELM, although the former is a bit slower. Since new high performing computing machines are coming available every year, the emphasis should be placed on the performance in terms of sensitivity, specificity, precision and accuracy, rather than on the computation time. Therefore, the WPD method with k -fold cross-validation and the NB classifier should be given priority over any other combinations presented in this paper.

VI. CONCLUSION

This paper has addressed the issue of unfair comparisons among the methods that are used to classify ‘healthy’ and ‘alcoholic’ EEG signals. Although an EEG dataset is publicly available, the existing methods are compared when they use different subsets of the dataset and many of their algorithmic

parameters are not known. We have presented a level playing field for comparisons of any past and future methods. For instance, this paper presents a fair comparison of the two state-of-the-art methods, i.e., the WPD [3] and EMD [7] methods. The results and observations reported in this paper are sometimes quite different from those that the respective authors have reported in their papers. For example, for the WPD method the use of any number of vanishing moments (2, 3 or 4) is fine with the involved wavelet filter. Moreover, for the EMD method, the choice of the reduced set of the extracted features (from the low-pass sub-bands) is counterproductive.

The paper has presented some new observations. For instance, the k -fold cross-validation offers better performance than the holdout cross-validation at the expense of additional computational cost. The NB classifier performs better than the ELM classifier although the latter is faster in the training and testing steps. With the advent of high-performance computers frequently, it is advised to focus more on the classification effectiveness rather than on the efficiency. In this regard, the WPD method is preferred to the EMD method, the k -fold validation over the holdout validation and the NB classifier against the ELM classifier.

REFERENCES

- [1] WHO report. (2018) Harmful use of alcohol kills more than 3 million people each year, most of them men. [Online]. Available: <http://www.hydroquebec.com/electricite-et-vous/servitudes-droits-proprie/> (last accessed on 10 July 2020).
- [2] S. Patidar, R. B. Pachori, A. Upadhyay, and U. R. Acharya, “An integrated alcoholic index using tunable-Q wavelet transform based features extracted from EEG signals for diagnosis of alcoholism,” *Applied Soft Computing*, vol. 50, pp. 71–78, January 2017.
- [3] J. das C. Rodrigues, P. P. R. Filho, E. Peixoto, A. K. N., and V. H. C. de Albuquerque, “Classification of EEG signals to detect alcoholism using machine learning techniques,” *Pattern Recognition Letters*, vol. 125, pp. 140–149, July 2019.
- [4] S. Walsh, J. Donnan, Y. Fortin, L. Sikora, A. Morrissey, K. Collins, and D. MacDonald, “A systematic review of the risks factors associated with the onset and natural progression of epilepsy,” *NeuroToxicology*, vol. 61, pp. 64–77, July 2017.
- [5] Y. Liu, Y. Lin, S. Wu, C. Chuang, and C. Lin, “Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 2, pp. 347–360, November 2016.
- [6] “Sleep spindle characteristics in adolescents,” *Clinical Neurophysiology*, vol. 130, no. 6, pp. 893–902, June 2019.
- [7] S. Taran and V. Bajaj, “Rhythm-based identification of alcohol EEG signals,” *IET Science, Measurement Technology*, vol. 12, no. 3, pp. 343–349, April 2018.
- [8] UCI KDD database. (1999) Eeg database. [Online]. Available: <http://kdd.ics.uci.edu/databases/eeg/eeg.data.html> (last accessed on 15 July 2020).
- [9] R. Upadhyay, P. K. Padhy, and P. K. Kankar, “Alcoholism diagnosis from EEG signals using continuous wavelet transform,” in *2014 Annual IEEE India Conference (INDICON)*, February 2014, pp. 1–5.
- [10] A. Anuragi and D. S. Sisodia, “Alcoholism detection using support vector machines and centered coreentropy features of brain EEG signals,” in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, May 2017, pp. 1021–1026.
- [11] U. R. Acharya, S. V. Sree, S. Chattopadhyay, and J. S. Suri, “Automated diagnosis of normal and alcoholic EEG signals,” *International Journal of Neural Systems*, vol. 22, no. 3, pp. 1250011, 11 pages, June 2012.
- [12] W. Mumtaz, P. L. Vuong, L. Xia, A. S. Malik, and R. B. A. Rashid, “Automatic diagnosis of alcohol use disorder using EEG features,” *Knowledge-Based Systems*, vol. 105, pp. 48–59, August 2016.