2011 IEEE International Workshop on Machine Learning for Signal Processing
September 18-21, 2011, Beijing, China

DIMENSIONALITY REDUCTION FOR EEG CLASSIFICATION USING MUTUAL INFORMATION AND SVM

Carlos Guerrero-Mosquera, Michel Verleysen** and Angel Navia Vazquez**

*University Carlos III of Madrid, Signal Processing and Communications Department
Avda. Universidad, 30 28911 Leganes. Spain
{cguerrero,navia}@tsc.uc3m.es
**Machine Learning Group, Universite Catholique de Louvain
pl. du Levant 3, 1348 Louvain-la-Neuve, Belgium
michel.verleysen@uclouvain.be

ABSTRACT

Dimensionality reduction is a well known technique in signal processing oriented to improve both the computational cost and the performance of classifiers. We use an electroencephalogram (EEG) feature matrix based on three extraction methods: tracks extraction, wavelets coefficients and Fractional Fourier Transform. The dimension reduction is performed by Mutual Information (MI) and a forward-backward procedure. Our results show that feature extraction and dimension reduction could be considered as a new alternative for solving EEG classification problems.

1. INTRODUCTION

Since the integration between classical and modern biomedical signal processing with the engineering, new fields have been activated in a new area known called "neuroengineering." Clinical neuroengineering has active fields such as neural prosthesis, brain computer interface (BCI), new clinical imaging techniques and treatment tools with electroencephalogram (EEG), evoked potentials (EPs), magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI). Nowadays, there are several processing methods, tools and algorithms for helping in new treatments, obtaining new measurements of brain activity and detecting brain diseases.

EEG signals not only represent the brain function but also the status of the whole body, i.e., a simple action as blinking the eyes introduces oscillations in the EEG records. Then, the EEG is a direct way to measure neural activities and it is important in the area of biomedical research to understand and develop new processing techniques.

EEG signal pre-processing and processing could be considered as a "pattern recognition system" with focus on the classification algorithms. Pre-processing methods include EEG signal modelling, segmentation, filtering and denoising, and EEG processing methods consist of two tasks: feature extraction and classification.

Feature extraction consist in finding a set of measurements or a block of information with the objective of describing in a clear way the data or an event present in a signal. These measurements or *features* are the fundamental basis for detection, classification or regression tasks in biomedical signal processing and is one of the key steps in the data analysis process.

Features constitute a new form of expressing the data, and can be binary, categorical or continuous: they represent attributes or direct measurements of the signal. For example, features may be the age, health status of the patient, family history, electrode position or EEG signal descriptors (amplitude, voltage, phase, frequency, etc.). The aim of extracting features is to identify "patterns" of brain activity: features can be used as input to a classifier. The performance of a pattern recognition system depends on both the features and the classification algorithm employed.

This work at improving the performance of EEG signal classification. For this purpose, we construct a feature matrix using epileptic EEG signals and different feature extraction algorithms, and then we try to improve the classifier performance by dimension reduction. In this way, we could remove redundant features and improve the computational cost by simplification of the resulting models.

The paper is organized as follows: Section 2 explains the feature matrix construction, dataset used and also introduces the feature selection procedure. Section 3 shows results in EEG dimension reduction of the feature matrix. In Section 4 the main results are discussed and conclusions of the paper are given in Section 5.

This work has been funded by the Spanish Government under grant TEC2008-02473.

978-1-4577-1623-2/11/$26.00

2. METHODS

2.1. EEG feature extraction: related work

More formally, feature extraction assumes we have for N samples and D features, a $N \times D$ data matrix. It is also possible to obtain a feature vector at the sample n from the feature matrix, that is, x is a unidimensional vector $x = [x_1, x_2, \ldots, x_D]$ called as "pattern vector".
More specifically in EEG detection and classification sceneries, there are several features proposed in the literature for EEG signals such as power spectral density [1], wavelet transform [2, 3], Lyapunov exponents with wavelets [4], sampling techniques [5] and time frequency analysis [6].
In this paper, we evaluate three methods for feature extractions for EEG signals: tracks extraction (LFE) [7], wavelet transform (W) and Fractional Fourier transform (FrFT) [8]. Although there are more methods proposed for EEG extraction in the literature [9], we have chosen these algorithms because both wavelets and the FrFT show better results than others methods, for example: Fourier transform or time-frequency analysis (TFD). FrFT has the property that if we gradually increase its order, we can obtain more information in the form of coefficients than the Fourier transform. On the other hand, wavelets improve the time-frequency resolution through multi-resolution analysis [8]. Additionally, the tracks extraction has been selected because it works with only three features: energy (E), frequency (F) and track length (L), which together can solve complex classification problems [7]. Moreover, this method could be combined with other features and improve the results, as shown later.

2.2. Dimensionality reduction EEG

In practice, we need to know which features are sufficient and appropriate to specific problems but usually it is not easy to know a priori which features will be useful. It is here where feature selection methods play an important role. Feature selection techniques can help to achieve the following goals: (i) reducing the size of the feature matrix. This may involve removing some features and reducing unnecessary data storage. (ii) improving both the computational cost and the performance of the classifier, since we only retain the features useful to the classifier. This means that we do not choose the most potentially relevant, because the design could be suboptimal or conversely, the subset most useful, because it may exclude the most important relevant features [10].

2.3. Mutual information (MI)

Mutual information (MI) measures the relevance between a group of features X and the variable or output Y. This relationship is not necessarily linear. The mutual information between two variables is the amount of uncertainty (or entropy) that is lost on one variable when the other is known, and vice versa. The variables X and Y could be multidimensional, solving the drawback in correlation measurements that are based on individual variables [11].
The MI for discrete variables is calculated as

$$I(X,Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x) p_Y(y)} \qquad (1)$$

Note in Eq.1 that it is necessary to know the exact probability density functions for estimating the MI and this is the most sensitive part in the MI estimation. Different methods have been proposed in the literature to estimate such joint densities such as Parzen windows, the k-nearest neighbor algorithm (k-NN), the Kraskov method and other estimators derived from Kraskov method and oriented to classification problem [12].

2.4. Forward-backward algorithm

Several search strategies could be used for finding the most adequate subset of features, such as best-first, branch-and-bound, simulated annealing and genetic algorithms. Greedy search strategies such as forward selection, backward elimination or any combination of them are the most popular. The forward selection method starts from an empty set and progressively add features one by one according to some criterion. In a backward elimination procedure one starts with all the features and progressively eliminates the least useful ones [10].
With D input features, there are 2^{D-1} possible subsets that should be studied but this evaluation is unfeasible for large D due to its high computational cost. The combination of forward and backward procedures could alleviate the curse of dimensionality by avoiding the evaluation of features with dimension D. This algorithm works in the following steps:
(1) The first feature is selected by MI maximization between the original features $[X_1, ..., X_D]$ and the output variable Y:

$$X_1^{sel} = \arg\max_{X_j} \left\{ \hat{I}(X_j, Y) \right\}, \quad 1 \leq j \leq D$$

where $\hat{I}(X,Y)$ is the MI estimation of $I(X,Y)$ and X_{s1} is the first variable selected.
(2) Once X_1^{sel} is selected, the next components must be selected taking into account this first variable. The second variable X_2^{sel} is the one that maximizes the MI in conjunction with the first one and the output variable Y:

$$X_2^{sel} = \arg\max_{X_j} \left\{ \hat{I}(\{X_1^{sel}, X_j\}, Y) \right\}, 1 \leq j \leq D, X_j \neq X_1^{sel}$$

The next steps consist in selecting the variable X_{st}^{sel} in the t-th step given a subset of already selected features

$$S = [X_1^{sel}, X_2^{sel}, ..., X_{s(t-1)}^{sel}],$$

then X_{st}^{sel} is chosen according to

$$X_{st}^{sel} = \arg\max_{X_j} \left\{ \hat{I}(\{S, X_j\}, Y) \right\}, 1 \leq j \leq D, X_j \notin S$$

(3) Assuming that t variables have been selected after step t (i.e. the last variable selected is X_{st}^{sel}), the backward procedure consist in checking what happens with the MI when a variable is removed from the subset S. The variable chosen is the one (X_t^{rem}) that increases the estimation of the MI when eliminated. In other words, we apply the next maximization rule after forward step t:

$$X_t^{rem} = \arg\max_{X_j^{sel}} \left\{ \hat{I}(\{X_1^{sel}, ..., X_{j-1}^{sel}, X_{j+1}^{sel}, ..., X_{t-1}^{sel}\}, Y) \right\}, \quad 1 \leq j \leq t,$$

if

$$\hat{I}(\{X_1^{sel}, ..., X_{j-1}^{sel}, X_{j+1}^{sel}, ..., X_t^{sel}\}, Y) > \hat{I}(\{X_1^{sel}, ..., X_t^{sel}\}, Y)$$

Search methods must have a stopping criterion. The backward elimination is more intuitive than the forward selection. For the latter one could use a permutation test among the features chosen to evaluate if a new variable presents a significant increase of MI. Another way is using a ranking algorithm rather than a selection one. These methods are described in [12] respectively. However, studying stopping criteria for this problem exceeds the scope of this paper and is left as future work.

3. EXPERIMENTS AND RESULTS

The aim of this section is to evaluate the performance in dimension reduction of the feature matrix. Several experiments using different MI estimations have been conducted. A preliminary study on feature selection using Smooth pseudo Wigner-Ville distribution (SPWV), wavelets and FrFT is described in [8].

3.1. Data and experimental setup

This paper uses a database consisting of five sets (denoted as Z, O, N, F and S), each one containing 100 single-channel EEG segments each having 23.6 sec duration and sampling rate of 173.61 Hz. [1]. In our experiments we use three classifications problems to evaluate our features.

1) The first problem called N1, two classes are examined: normal (Z) and seizure (S).

[1]More details on the datasets and the classification problem are described in [13].

2) The second classification problem called N2, includes the classes normal, seizure-free and seizure (Z, F and S respectively).

3) In the third problem called N3, all the five classes are used.

Next, we create a feature matrix for each problem and individually test each feature or extraction method using a classifier. The feature matrix, with dimension (D) equal to 65, has been formed by three feature extraction methods: tracks extraction (3 features), wavelet coefficients (45 features) and fractional Fourier coefficients (17 features).
In this paper we will use SVMs with radial basis function (RBF) as classifiers. SVMs have proved to be one of the most appropriate alternative for solving classification problems and their solution is supported by Statistical Learning Theory [14, 15]. The SVM parameters are found by 10-fold cross-validation. Time-frequency distribution is the Reduced Interference Distribution (RID). EEG epochs was 5 secs., $\Delta = 0.5$ Hz and 30% of overlapping. K value for K-NN algorithm used in estimating the probability density function was 30.
The statistical relevance of the results has been verified by means of a Kruskall-Wallis test, which is a sort of nonparametric ANOVA test that does not assume Gaussianity in the data under study. In all cases (except between Fractional Fourier (FrF) and LFE+Wavelets (W) in the N1 case) a p-value smaller than 0.01 has been obtained, thereby rejecting the null hypothesis that data come from the same distribution. Note in these tables the difference in difficulty among N1 (easy), N2 and N3 (hard) problems.
Finally, the performance evaluations of the experiments was based on 1000 bootstrap runs [16] using two measures: "F_{score}" and accuracy, defined as:

$$F_{score} = 2*sensitivity*specifity/(sensitivity+specifity)$$

where sensitivity and specifity are defined as follows:

. $Sensitivity$: Percentage of EEG segments containing seizure activity correctly classified.

. $Specificity$: Percentage of EEG segments not containing seizure activity correctly classified.

and

$$Accuracy = \frac{\text{\# of correctly classified labels}}{\text{total \# of labels}}$$

3.2. Results

Table 2 shows individual performance (F_{score}) for each feature extraction method and their combinations on independent test sets. N2 problem achieves 99.74% in performance using features LFE+FrF (dimension D=20). N3 achieves

96.02% with all features (D=65). N1 problem is solved using just 2 features (LE) obtained from tracks extraction (100% with D=2).
The experiments that follow evaluate each feature (or feature subsets) using a feature selection algorithm based on forward-backward procedure and Mutual Information (MI) as relevant criteria. Although the N1 problem only needs two features (LE) to solve the classification problem (see Table 2), this problem was also included in feature selection analysis to see if there is possibility of finding other subset of features with the same performance (100%).
Table 1 shows the results on feature selection using reduced interference distribution (RID) and three methods for MI estimation: Kraskov, Parzen and K-NN. It should be noted that each method selects different features, but in most cases it is successful in F_{score} rate.

Table 1. F_{score} evaluations achieved by three different MI estimations and RID. F_{score} values average and variances correspond to over the 1000 bootstrap runs.

	Forward-backward selection		
	Kraskov	Parzen	Knn
N1	{E, 1 FrF} 99.77 ± 0.0013	{L,E} **100** ± 0	{F,E, 5 WC, 10 FrF } 99.16 ± 0.0058
N2	{L,F,E, 1 FrF} **100** ± 0.0023	{L,F,E} 83.50 ± 0.0552	{L,F,E, 2 WC, 5 FrF} 99.79 ± 0.0023
N3	{3 FrF} 85.45 ± 0.0116	{L,F,E} 83.34 ± 0.0721	{L,F,E, 2 WC, 6 FrF} **99.59** ± 0.0356

The following can be noted from these results:

- N1 is definitely solved with only two features obtained by tracks extraction method (LFE). Parzen method selected the appropriate features.
- Problem N2 has increased from 99.74% with dimension D=20 (see Table 2) to 100% with D=4 (see Table 1). Similarly, problem N3 from 96.02% with D=65 (see Table 2) to 99.59% with D=11 (see Table1). The selection method that showed better performance in this case was the K-NN.

Classification results clearly show the good performance of the LE detector in N1 problem (100%). They also show an increase of accuracy with less features for datasets N2 and N3. This validates the importance of choosing suitable features in each classification problem.
Finally, Table 3 shows a comparison between our features selected and other methods proposed in the literature. Results are presented in accuracy rate and the distribution used is the RID. Note that N1 and N2 has reached the maximum accuracy but with the difference in the dimension of the feature matrix (D=2 for N1, N2 D=4) than the best approach (D=4 both N1 and N2). Something similar occurs with N3 that has increased in accuracy from 99.28% to 99.59% but with a noticeable dimension reduction (D=11) compared with the best proposal in the literature, which has D=24.

4. DISCUSSION

Preliminary studies have shown the effectiveness of both wavelet coefficients and FrFT algorithm, as well as tracks extraction method for EEG classification [8]. There are three important remarks in this paper: (i) We made a preliminary selection of several time-frequency distributions; RID distribution was chosen for its good performance in classification tasks. (ii) We select features based on MI criteria and evaluate these features. (iii) We show a comparison between our features selected and other methods proposed in the literature.
Results presented in this work confirm that feature selection based on MI and forward-backward procedure is a good alternative for improving the classifier performance in EEG signals. In addition, these results also show the effectiveness of tracks extraction method for EEG signals with epilepsy and also show that these features could be combined with other features for classification tasks.

5. CONCLUSIONS

Dimensionality reduction of epileptic EEG feature matrix using mutual information and forward-backward procedure has shown to be a good alternative for improving classifier performance. On the other hand, features proposed based on tracks extraction, wavelet coefficients and fractional Fourier transform could be considered a good alternative to solve EEG classification problems with epilepsy.
Future work is oriented to detect other events on EEG signals such as event related potentials (ERPs), slow wave-sleep (SWS) and oxygen deprivation on fetal EEG, as well as the study of stopping criteria in the forward-backward algorithm and the comparison with non-greedy methods for MI estimation.

6. REFERENCES

[1] C. Lehmanna, T. Koenig, V. Jelic, L. Prichep, R.E. John, L. Wahlund, Y. Dodgee, and T. Dierks, "Application and comparison of classification algorithms for recognition of Alzheimers disease in electrical brain activity (EEG)," *Journal of Neuroscience Methods*, vol. 161, pp. 342–350, 2007.

[2] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, pp. 1084–1093, 2007.

[3] Ocak Hasan, "Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic al-

gorithm," *Signal processing*, vol. 88, pp. 1858–1867, 2008.

[4] I. Gler and E. Derya beyli, "Multiclass support vector machines for EEG-signals classification," *IEEE Trans. on Inf. Tech. in Biomed.*, vol. 11, pp. 117–126, 2007.

[5] Y.L. Siuly and P. Wen, "Classification of EEG signals using sampling techniques and least square support vector machine," *Lectures notes in Computer Science*, vol. 5589, pp. 375–382, 2009.

[6] A.T. Tzallas, M.G. Tsipouras, and D.I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis," *IEEE Trans. on Inf. Tech. in Biomed.*, vol. 13, pp. 703–710, 2009.

[7] C. Guerrero-Mosquera, A. Malanda Trigueros, J. Iriarte Franco, and A. Navia-Vazquez, "New feature extraction approach for epileptic EEG signal detection using time-frequency distributions," *Med. Biol. Eng. Comput.*, vol. 48, pp. 321–330, 2010.

[8] C. Guerrero-Mosquera, M. Verleysen, and A. Navia-Vazquez, "EEG feature selection using mutual information and support vector machine: A comparative analysis," *Proceedings of the 32nd Annual EMBS International Conference*, pp. 4946–4949, 2010.

[9] S. Sanei and J.A. Chambers, *EEG signal processing*, vol. 1, Wiley, 2007.

[10] I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh, *Feature Extraction, Foundations and Applications*, vol. 1, Springer, 2006.

[11] T.M. Cover and J. Thomas, *Elements of Information Theory*, vol. 2, Wiley, 1991.

[12] V. Gomez-Verdejo, M. Verleysen, and J. Fleury, "Information-theoretic feature selection for functional data classification," *Neurocomputing*, vol. 72, pp. 3580–3589, 2009.

[13] A.T. Tzallas, M.G. Tsipouras, and D.I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis," *IEEE Trans. on Inf. Tech. in Biomed.*, vol. 13, pp. 703–710, 2009.

[14] V. Vapnik, *The Nature of Statistical Learning Theory*, vol. 2, Springer-Verlag, 2000.

[15] B. Schlkopf and A. Smola, *Learning with kernels*, vol. 1, The MIT Press, 2002.

[16] F.E. Harrel, *Regression Modeling Strategies*, Springer, 2001.

[17] V. P. Nigam and D. Graupe, "A neural network based detection of epilepsy," *Neurol. Res.*, vol. 26, pp. 55–60, 2004.

[18] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-frequency domain features," *Journal of Medical Systems*, vol. 29, pp. 647–660, 2005.

[19] N. Kannathal, M.L. Choo, U.R. Acharya, and P.K. Sadasivan, "Entropies for detection of epilepsy in EEG," *Comput. Methods Prog. Biomed*, vol. 80, pp. 187–194, 2005.

[20] N. Kannathal, U.R. Acharya, C.M. Lim, and P.K. Sadasivan, "Characterization of EEG- A comparative study," *Comput. Methods Prog. Biomed*, vol. 80, pp. 17–23, 2005.

[21] K. Polat and S. Gnes, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast fourier transform," *App. Math. Comput.*, vol. 32, pp. 625–631, 2007.

[22] K. Chua, V. Chandran, U. Acharya, and C. Lim, "Application of higher order spectra to identify epileptic EEG," *Journal of Medical Systems*, pp. 1–9, 2010.

[23] S.F. Liang, H.C. Wang, and W. L. Chang, "Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection," *EURASIP Journal on Applied Signal Processing*, vol. 2010, pp. 853434, 2010.

[24] Yodchanan Wongsawat, "Epileptic seizure detection in EEG recordings using phase congruency," *Proceedings of the 30th IEEE Annual EMBS International Conference*, pp. 927 –930, 2008.

[25] N. F. Gler, E.D. beyli, and I. Gler, "Recurrent neural networks employing lyapunov exponents for EEG signals classification," *Exp. Syst. Appl.*, vol. 29, pp. 506–514, 2005.

[26] N. Sadati, H. R. Mohseni, and A. Magshoudi, "Epileptic seizure detection using neural fuzzy networks," *Proc. Of IEEE Int. Conf. On Fuzzy Syst.*, vol. 29, pp. 506–514, 2006.

[27] I. Gler and E.D. beyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *Journal of Neuroscience Methods*, vol. 148, pp. 113–121, 2005.

[28] E. D. beyli and I. Gler, "Features extracted by eingenvector methods for detecting variability of EEG signals," *Pattern Recognit. Lett.*, vol. 28, pp. 592–603, 2007.

Table 2. F_{score} values average and variances correspond to over the 1000 bootstrap runs using the TFD RID.

	L	F	E	LF	FE	LE	LFE	W	FrF	LFE+W	LFE+FrF	W+FrF	All
Dim	1	1	1	2	2	2	3	45	17	48	20	62	65
N1	67.24 ±0.0553	86.95 ±0.0357	99.38 ±0.0091	92.29 ±0.0261	99.82 ±0.0041	**100** ±0	99.57 ±0.0068	99.89 ±0.0032	98.70 ±0.0120	99.91 ±0.0031	99.06 ±0.0099	99.74 ±0.0048	99.36 ±0.0059
N2	36.84 ±0.1438	51.99 ±0.2895	75.81 ±0.0803	87.07 ±0.0322	87.41 ±0.0320	86.19 ±0.0359	83.51 ±0.0406	93.28 ±0.0250	99.18 ±0.0089	92.55 ±0.283	**99.74** ±0.0055	98.36 ±0.0145	98.06 ±0.0303
N3	20.14 ±0.1019	5.01 ±0.0139	9.80 ±0.0748	15.79 ±0.1204	67.34 ±0.1040	64.73 ±0.0841	82.13 ±0.0449	82.54 ±0.0459	83.59 ±0.0460	81.36 ±0.0434	89.82 ±0.0303	93.23 ±0.0270	**96.02** ±0.0301
Average	41.40	47.98	61.66	65.05	84.85	83.04	88.40	91.91	93.82	91.27	96.20	97.11	**97.81**

Table 3. Comparison of classification accuracy (in percent) obtained by LFE approach for epileptic seizure detection.

Authors	**Method**	**Problem**	**Dimension**	**Accuracy**
[17]	Non-linear pre-processing filter-Diagnostic neural network	N1	2	97.2
[18]	Time & frequency domain features-Recurrent neural network	N1	5	99.6
[19]	Entropy measures-Adaptive neuro-fuzzy inference system	N1	4	92.2
[20]	Chaotic measures-Surrogate data analysis	N1	4	$\approx$90
[21]	Fast Fourier transform-Decision tree	N1	129	98.72
[2]	Discrete wavelet transform-Mixture of expert model	N1	4	95
[22]	High order spectra (HOS)-Gaussian mixture model and Support vector machine	N1	6	93.11
[23]	Spectral and Entropy analysis-Linear and Non-linear classifiers	N1	16	98.51
[24]	Phase congruency-Linear discriminant analysis	N1	1	99
[6]	Time & frequency analysis-Artificial neural network	N1	4	100
This work	**LFE-Tracks extractions-Support vector machine**	**N1**	**2**	**100**
[25]	Lyapunov exponents-Recurrent neural network	N2	4	96.79
[26]	Discrete wavelets transform-Adaptive neural fuzzy network	N2	6	85.9
[23]	Spectral and Entropy analysis-Linear and Non-linear classifiers	N2	16	98.67
[24]	Phase congruency-Linear discriminant analysis	N2	1	96.5
[6]	Time & frequency analysis-Artificial neural network	N2	4	100
This work	**LFE Tracks extractions-Support vector machine and feature selection**	**N2**	**4**	**100**
[27]	Wavelet transform-Adaptive neuro-fuzzy inference system	N3	20	98.68
[4]	Wavelet transform-Support vector machine	N3	24	99.28
[28]	Eigenvector method-Modified of mixture of expert model	N3	12	98.60
[23]	Spectral and Entropy analysis-Linear and Non-linear classifiers	N3	16	85.9
[24]	Phase congruency-Linear discriminant analysis	N3	1	91
[6]	Time & frequency analysis-Artificial neural network	N3	4	89
This work	**LFE Tracks extractions-Support vector machine and feature selection**	**N3**	**11**	**99.59**