

Feature Space Reduction, using PCA in the Algorithm for Epilepsy Detection, using an Adaptive Neuro-Fuzzy Inference System and Comparative Analysis

Marjan Stoimchev, Vesna Ojleska Latkoska

Faculty of Electrical Engineering and Information Technologies, “Ss. Cyril and Methodius” University in Skopje, 1000 Skopje, Republic of Macedonia, ksjar3122013@feit.ukim.edu.mk, vojleska@feit.ukim.edu.mk

Abstract: This study presents an upgrade to our previously published algorithm for the detection of Epilepsy. The upgrade in the old algorithm, which was based on wavelet transform (WT) for feature extraction, and Adaptive Neuro-Fuzzy Inference System (ANFIS) for classification, is made by using Principal Component Analysis (PCA) in order to reduce the number of features used, for the training the ANFIS network. In order to make a comparison of the old and the upgraded algorithm with PCA, i.e. evaluating both algorithms, in terms of training performance and classification accuracies, comparative analysis of the both algorithms was made, when using different data splitting methods, and different input space partitioning methods. It was concluded that the upgraded algorithm exhibits a satisfactory performance, and in some cases, performs better than the old algorithm, even though the number of features is significantly reduced (from 20 to 7), which plays a crucial role in making the new algorithm more resistant to overfitting.

Keywords: Adaptive Neuro-Fuzzy Inference System (ANFIS); wavelet transform; fuzzy logic; Finite Impulse Response (FIR) filter; electroencephalogram (EEG); comparative analysis; normalization; input space partitioning; Principal Component Analysis (PCA)

1 Introduction

Epilepsy is chronic brain disorder, characterized by seizures, which can affect any person, at any age. It is characterized by recurrent convulsions over a time-period. Clinical diagnosis of epilepsy requires detailed history and neurological examinations [1]. There are many techniques to investigate the recurrent epileptic convulsions (namely, Computer Tomography-CT, Magnetic Resonance Imaging-MRI and Electroencephalogram-EEG). As it is stated in [1], the most common effective diagnostic method for the detection of epilepsy is the analysis of EEG signals, which can be based on different types of approaches [14] [26] [27].

Although it is possible for experienced neurophysiologist to detect the epilepsy by visually scanning of the EEG signals, for a more objective analysis and reproducible results, it is always advantageous to detect these activities from the EEG signals through some computer methods by extracting relevant features from the signals [1] [26] [27]. In order to solve this, there are many proposed methodologies. In general, all of the techniques consist of several steps, i.e. from data preprocessing, as the first step, feature extraction as a second step and to classification, as a third step.

In the first step, EEG signal de-noising is done, and it can be based using conventional filtering methods [8] [9], or filtering through wavelet analysis [11], [25]. For the second step (feature extraction), there are many, different methods (based on frequency domain analysis, or time domain analysis, or both), whereas the results of the studies in the literature have demonstrated that the WT is the most promising method to extract relevant features from the EEG signals [26], [27] [13]. For the final step there are also different ways for classifying the EEG signals [26], (feature extraction using genetic algorithms [28], the wavelet-based support vector machine (SVM) classifier [29], wavelet-based feed forward artificial neural network-FFANN [23] [24], fuzzy rule-based detection [30], Adaptive Neuro-Fuzzy Inference System (ANFIS) [13] [14] [7] [2], and many others).

In our previous study [2] and [3] an algorithm for classification of EEG signals was proposed, that combines Finite Impulse Response (FIR) filtering for artefact removal [8] [9] [7], WT for feature extraction [12] [13] [14] [26] [27], and ANFIS for classification [12] [13] [14]. ANFIS model learns how to classify the EEG signal, through the standard hybrid learning algorithm, whereas a special form of ANFIS model was used, which depending on the number of inputs, splits the model into appropriate number of substructures (sub-ANFIS models).

This study is a continuation of the study reported in [2] and [3], whereas in order to reduce the number of features in the old algorithm ([2] and [3]) an upgrade to the algorithm was made, introducing PCA. The researchers in [4] [5] [6] have also applied PCA to classify the epileptic EEG signals, using this time domain method to reduce the large number of data and select the most important components as feature vectors. The main difference of this papers with our study is that researchers in [4] used SVM based classification on EEG signals, whereas the researchers in [5] used EEG signal decomposition with the Wavelet Packet Decomposition (WPD) method, and the classification was obtained with the Gaussian Mixture Model (GMM) classifier and lastly, researchers in [6] have tested several classifiers, namely the K-Nearest Neighbor (KNN), SVM, the naive bayes (NB) classifier and the LDA method.

As the proposed algorithm in [2] [3] had a good performance in comparison to the several similar algorithms from different researchers [23] [24] [7] [14], in this study a comparison analysis of the old [2] [3]) and the upgraded algorithm was

made, i.e. evaluating the both algorithms in terms of training performance and classification accuracies. The comparative analysis of the both algorithms was made using different data splitting methods, and different input space partitioning methods. Firstly, an initial simulation comparison analysis of the both algorithms over the grid partitioning method [15] was made, by dividing the dataset into various ways (using 70%-30% and 50%-50% ratio of the training and testing dataset, and using K-Fold cross validation technique) [19]. The second comparative analysis was made by analyzing the old and the upgraded algorithm, through different input space partitioning methods (grid partitioning versus fuzzy c-means clustering, versus subtractive clustering) [15] [16] [20]. It was concluded that the new algorithm has satisfactory performance, and in some cases performs even better than the old algorithm. Nevertheless, due to the reduced number of features (from 20 to 7), used for training the ANFIS network, the upgraded algorithm is much more resistant to overfitting.

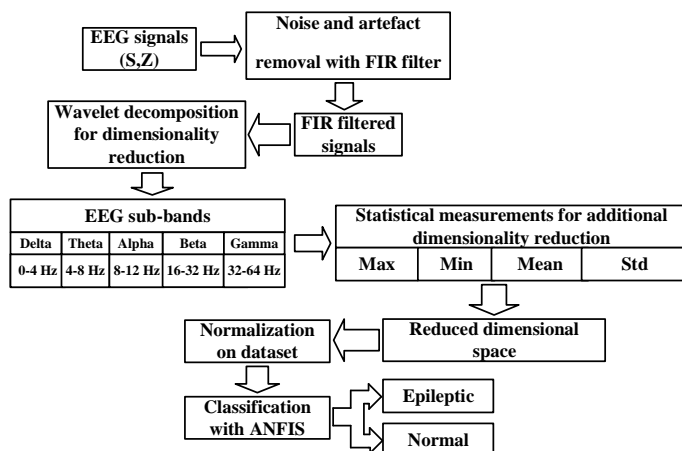
This paper is organized as follows. In Section 2 the upgraded algorithm for epilepsy detection using PCA is presented. In Section 3 a comparative analysis for the influence of the training and testing data was made, both for the old and the upgraded algorithm, whereas in Section 4 the comparative analysis of the both algorithms using different types of input space partitioning methods was made. Finally, the study is summarized giving the necessary conclusions.

2 Upgrade of the Algorithm for Epilepsy Detection with Fuzzy-Neural Network using PCA

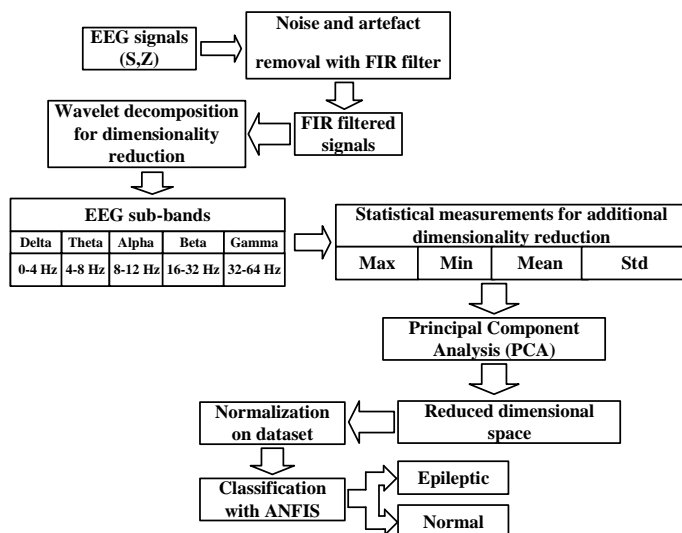
This work is a continuation of our previous work, published in [2] and [3], where the algorithm for detection of epilepsy with fuzzy-neural networks was presented. Here, except giving the short insight in the basic algorithm, presented in [2], and [3], an upgrade of the algorithm is introduced, using PCA method, in order to make further reduction of the used features. For the purpose of comparison on how different factors influence the system performance (both, when the old and the upgraded algorithm with PCA is used) bellow the old algorithm [2] [3], and the upgraded one are presented.

The algorithm for epilepsy detection with fuzzy-neural networks for classification of EEG signals [2] and [3], consists of three main steps (Figure 1-a)):

- 1) Filtering of the EEG signals, with FIR filter
- 2) Feature extraction and dimensionality reduction, with discrete wavelet transform (DWT)
- 3) Classification, using ANFIS



a)



b)

Figure 1

a) Data flow for the algorithm in [2], and [3], and b) Data flow for the upgraded algorithm, using PCA

whereas the upgraded algorithm in this study consists of one additional step, which implies the four following steps (Figure 1-b) and Figure 2):

- 1) Filtering of the EEG signals, with FIR filter
- 2) Feature extraction and dimensionality reduction, with DWT
- 3) Using PCA, for additional dimensionality reduction
- 4) Classification, using ANFIS

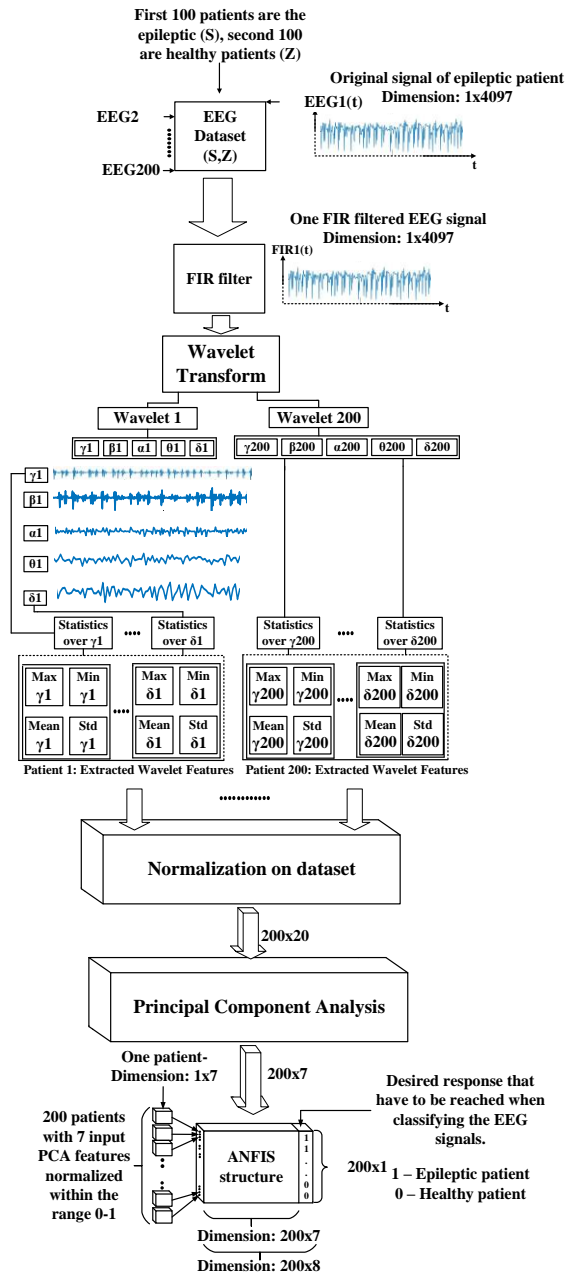


Figure 2

Detailed analysis for the overall upgraded algorithm from Figure 1-b)

Below we present a brief introduction for the used methodologies in the upgraded algorithm.

2.1 Input Data and De-noising of the EEG Signals

One of the major difficulties in analysis of EEG signals is the presence of artefacts [2] [7]. This disturbance represents serious obstructing factor that prohibits further processing to identify useful diagnostic features [2] [7].

In our case, as a first step in the upgraded algorithm (also in the old one, presented in [2] [3]), the band-pass FIR filter with the Hamming windowing method [8] [9] is used. The FIR filter is defined by two cutoff frequencies (in case of band-pass filtering), stopband attenuations and passband attenuation. The overall band of frequencies is defined by the Nyquist frequency, i.e. $F_s/2$ [8]. In our case, we used 1 Hz and 60 Hz, respectively. Below 1 Hz are the artefacts that are coming from the human body, and above 60 Hz is the power line noise [11].

The EEG data in this study was taken from the database of the university hospital in Bonn, Germany [10]. It consists of EEG signals that are recorded from three different events, namely, healthy subjects, epileptic subjects during seizure-free intervals (known as Interictal States) and epileptic subjects during a seizure (Ictal States).

The overall data consists of five subsets namely, O, Z, F, N and S. Set O and Z were obtained from healthy subjects with eyes open and closed respectively, sets F (epileptogenic zone) and N (hippocampal formation of the opposite hemisphere of the brain) were obtained during Interictal States in different zones of the brain and set S was taken from subjects during ictal state. Each subset contains 100 segments along with 4097 samples with sampling frequency of 173.61 Hz, each with duration of 23.6 seconds. We restrict ourselves to subsets S and Z, where the subset S denotes for epileptic subjects during epilepsy, whereas subset Z denotes for healthy subjects with eyes open. The dimension of our dataset is 200 segments by 4097 samples [10].

2.2 The use of Discrete Wavelet Transform for Feature Extraction

As a second step in the upgraded algorithm (also in the old one, presented in [2] [3]), the DWT is used, which analyses the signal at different frequency bands, with different resolutions, in terms of approximation and detail coefficients [12] [27].

The DWT is used for feature extraction [13] [14] where each EEG signal is decomposed into 4 levels, resulting in 4 detail coefficients and one final approximation coefficient. They are related to the EEG sub-bands, namely, α , β , γ , δ and θ [14]. After the DWT procedure, the dimension of the initial dataset was reduced. In [2], the wavelet coefficients were calculated using Daubechies wavelets of order 2 (db2) in MATLAB [17], whereas in [3] a comparative analysis of the influence of different wavelet families to the overall performance of the

system was made. As in [3], we concluded that Daubechies wavelets of order 2 (db2) gave the most satisfactory results, for all the analysis in this study db2 wavelets are used.

For further dimensionality reduction, statistics over the extracted wavelet coefficients is made, namely, maximum, minimum, mean value and standard deviation of the wavelet coefficients [14]. We present the initial dataset into a more compact representation, i.e. a dataset with dimensions of 200x20 (4 statistical measurements x number of extracted coefficients = 20 features for each EEG segment).

2.3 Further Dimensionality Reduction of the Feature Space by using the Principal Component Analysis (PCA) Method

As a third step in the upgraded algorithm (which was not present in the algorithms in [2] [3], PCA is used, which out of the given set of features selects the most important ones, in what way, further dimensionality reduction is performed. The researchers in [4] [5] [6] also applied the PCA to classify the epileptic EEG signals.

In general, PCA is a time domain method which is used to reduce the large number of data and select the most important components as feature vectors. Since patterns in data can be hard to find in data of high dimensions, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data. The objective is to represent data in a space that best expresses the variation in a sum-squared error sense.

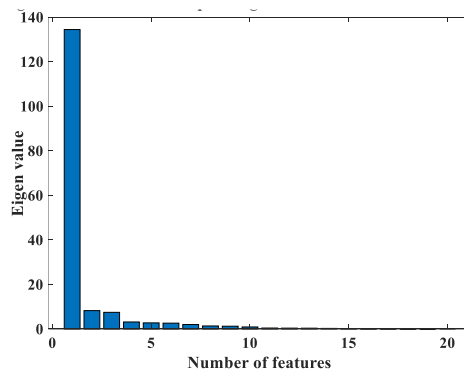


Figure 3

Eigen values for the corresponding features sorted in descending order

First, the d -dimensional mean vector μ and $d \times d$ covariance matrix is computed for the full data set. Next, the eigenvectors and eigenvalues are computed, and sorted according to decreasing eigenvalue. Subsequently, the largest k such eigenvectors are chosen [4].

In this extended version of the algorithm, the PCA method comes after the dimensionally reduced space from the DWT. The PCA replaces original features with new features, called principal components, which are orthogonal and have eigenvectors with adequate eigenvalues. Eigenvalues for the given 20 features are shown in decreasing order on Figure 3.

In order to decide which eigenvector(s) can be dropped without losing too much information from the construction of lower-dimensional subspace, inspection of the corresponding eigenvalues is made. The eigenvectors with the lowest eigenvalues bear the least information about the distribution of the data, and those are the ones that can be dropped. The common approach is to rank the eigenvalues from highest to lowest in order to choose the top k eigenvectors (as it is shown on Figure 3). In order to choose adequate number of principal components, a useful measure, a so-called explained variance, is used. The total variance is the sum of variances of all individual principal components, so the fraction of variance explained by a principal component is the ratio between the variance of the principal component and the total variance, as shown on Figure 4. The explained variance is a measure of how much information (variance) can be attributed to each of the principal components. It is clear that the first principal component explains the largest amount of variance (67.66%) compared to other principal components. Together, the first seven principal components contain 96.40% of the information. According to this, the first 7 principal components as features are used, implying reduced dataset of 200 instances by 7 features. Subsequently, those 7 feature vectors are used as an inputs to the ANFIS model [13] [14].

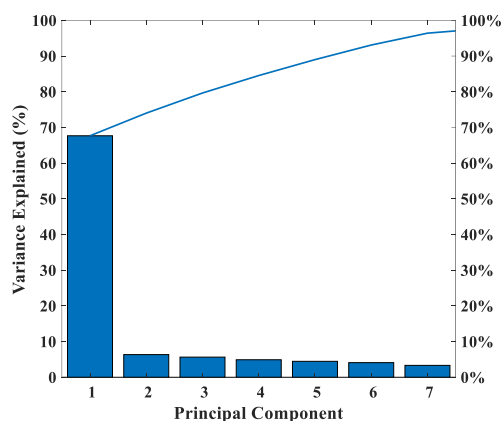


Figure 4

Variance explained for each principal component

It has to be emphasized that before the PCA, all the columns of the dataset have to be normalized within the range from 0 to 1 (Figure 2), in order to achieve stable convergence on the weighted factors of the lately used neural network training [18]. Min-max normalization is used, i.e. we are normalizing the feature column

vectors in the range from 0 to 1, [18] [22]. This method is also called “feature scaling”, and represents a preprocessing technique [22].

2.4 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is an Adaptive Neural Network (ANN) that is based on a fusion of ideas from fuzzy control and neural networks and possesses the advantages of both [15]. ANFIS is used as a fourth step in the upgraded algorithm (the third step in the old algorithm presented in [2] [3]), in order to make the final classification of the EEG patients.

The ANFIS classifier is trained with the hybrid learning algorithm [12] [14] [15]. In the old algorithm [2] [3], the 20 features were used as input patterns which represented the EEG signals, and output vector as the 21st column (epileptic patients were labelled with ones, and the healthy patients were labelled with zeros) which represented the desired response. In this paper only 7 features as an input in the ANFIS classifier are used, and the output vector as the 8th column (epileptic and healthy patients are also labelled with ones and zeros, respectively).

In the previous study [2], the ANFIS model used the grid partitioning method [15] for input space partitioning, where we presented a way of manipulating 20 inputs (with 3 MFs each) by partitioning the ANFIS model on sub-ANFIS models, as shown in Figure 6-a), surpassing the major obstacle of the “curse of dimensionality” [15]. In the upgraded algorithm simulations with grid partitioning method were also performed, but now only 7 inputs have to be manipulated (Figure 6-c)).

3 Comparative Analysis for the Influence of the Training and Testing Data, for the previous and the Upgraded Algorithm

In [3] different approaches on dividing the dataset were presented, i.e. how the size of the training and testing data influence the accuracy. In order to make a comparison on the performance of the old algorithm (that uses 20 features) and the upgraded algorithm with PCA (that uses 7 features) bellow 3 splitting methods were compared (divided dataset into 70%-30% and 50%-50% ratio, as well as, using the splitting method based on cross validation [19]). This initial comparative analysis is obtained using grid partitioning method, and use of Gauss MFs [17], as this type of MFs gave best results in our previous study [3].

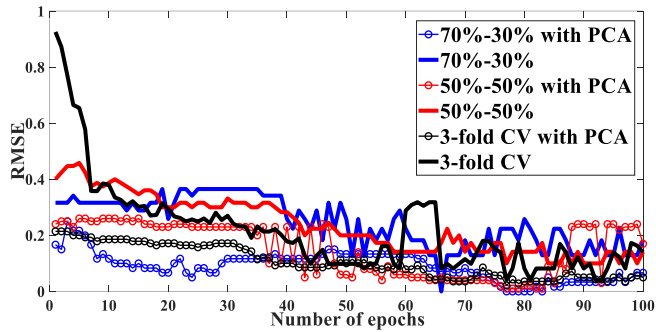


Figure 5

Comparison of the test set RMSEs over the three data split methods, using grid partitioning, for the old and the upgraded algorithm

When using the conventional splitting methods (70%-30%, or 50%-50% ratio of training and testing data, respectively) the data is simply divided into 2 appropriate sets. On the other hand, the K-Fold cross validation uses different approach.

With K-Fold cross validation, the available data is partitioned into K separate sets of approximately equal size [19]. The procedure involves K learning iterations, where for every iteration K-1 subsets are used for training, and the remaining set is used as the testing data. Every iteration leaves out a different subset, which means that each subset is used as test subset only once. In the end, all accuracies obtained from each iteration (testing fold) are averaged in order to obtain a reliable estimate of the model performance [19]. In our case 3-Fold cross validation is used.

Figure 5 presents the test set Root Mean Square Errors (RMSEs) when using the grid partitioning method during 100 epoch period, by applying the three different data split methods (70%-30%; 50%-50%; 3-Fold), for the old and the upgraded algorithm.

As can be seen from Figure 5, the lowest RMSEs are obtained in different number of epochs during the three partitioning methods. Nevertheless, it is evident that for the all three splitting methods, when the upgraded algorithm is used, RMSE values are lower than the RMSE values in the old algorithm. This implies that the use of PCA has a positive influence on the overall algorithm, when grid partitioning is used.

4 Influence of the Different Types of Input Space Partitioning

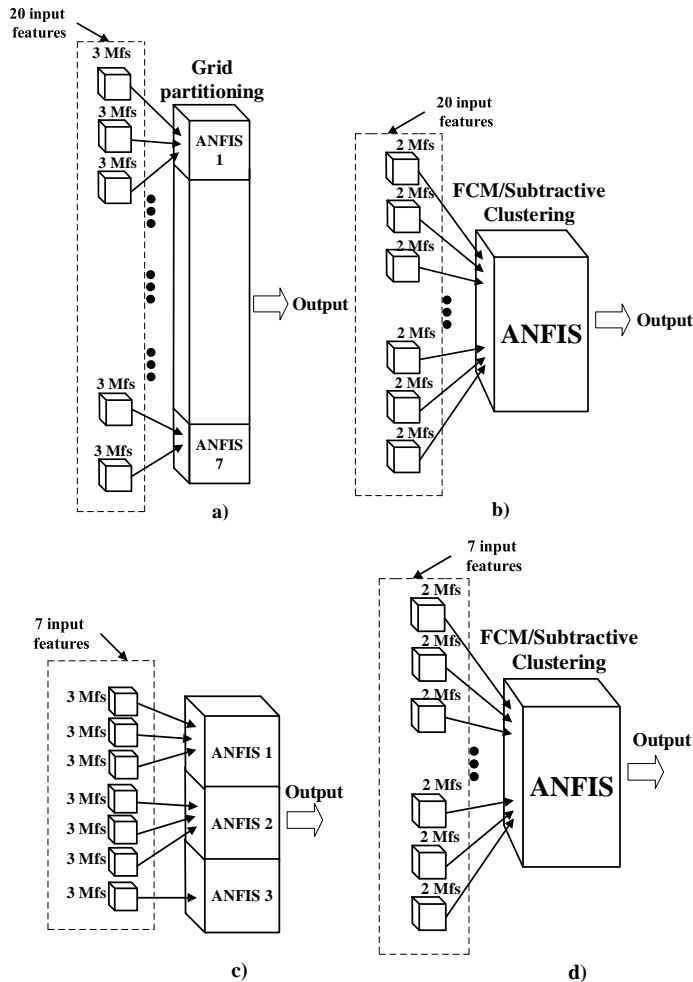


Figure 6

The structure of the model generated using different input space partitioning methods: a) Grid partitioning when the old algorithm with 20 input features is used ([2][3]), b) FCM/Subtractive clustering when the old algorithm with 20 input features is used ([2][3]), c) Grid partitioning when the upgraded algorithm with PCA and 7 input features is used; d) FCM/Subtractive clustering when the upgraded algorithm with PCA and 7 input features is used

This section presents different methods of input space partitioning in the ANFIS model and the overall overview is given in Figure 6. In the previous study, proposed in [2], only the grid partitioning method as shown in Figure 6-a) is used,

where in order to reduce the number of rules, sub-ANFIS models are formed. As it is clear from Figure 6-a), sub-ANFIS models from 1 to 6 accept 3 inputs and the last sub-ANFIS model accepts 2 inputs. In the study in [3], two new approaches for input space partitioning, namely, fuzzy c-means (FCM) clustering and subtractive clustering [15] were used. The resulting ANFIS structure in that case is shown on Figure 6-b), where all the 20 inputs are passed at once, i.e. the number of rules is equal to the number of clusters, thus we do not face the problem called “curse of dimensionality” as presented in [2]. In this paper further reduction of the number of features (from 20 down to 7 features using PCA) was made, whereas the ANFIS input/output structure, when grid partitioning and FCM/Subtractive clustering are used is given on Figure 6-c) and Figure 6-d), respectively.

FCM is a data clustering algorithm in which each data point belongs to a cluster to a degree specified by a membership grade (i.e. given data point can belong to several groups with the degree of membership between 0 and 1). The cluster centers are manually specified, where the performance depends on the initial cluster centers [15] [20].

Subtractive clustering, on the other hand, considers each data point as a potential cluster center, where the measure of potential is based on the distance of the data point from other data points (a data point located in a mound of different data points has a greater chance of being a cluster center) [15] [16].

Adequately to Figure 5, Figure 7 and Figure 8, present the test set RMSEs for FCM clustering and Subtractive clustering, respectively, using different data split methods, when both the old and the upgraded algorithm are used. 2 clusters for FCM clustering are used, and 0.8 radius of coverage (influence) for the subtractive clustering (for more information of the parameters in these clustering algorithms see [15] [17]). In both cases Gauss MFs [17] are used, and db2 wavelets, as this parameters gave best results in our previous study [3]. As we can see from Figure 7 and Figure 8, same as for the grid partitioning method (Figure 5), satisfying results are obtained when trained between 40 and 60 epochs approximately (old algorithm performs better at about 40 epochs, but the upgraded one needs more epochs for training, which is expected as in the upgraded algorithm we use less data compared to the old algorithm).

In the case of FCM clustering (Figure 7), different results are obtained, when the old and the upgraded algorithm are used, for the three splitting methods. At the beginning, for all the three splitting methods the upgraded algorithm shows worse behavior, but as the number of epochs increases this algorithm seems to stabilize after 60 epochs and does not show overfitting afterwards [21], as is the case with the old algorithm (for 50%-50% splitting method).

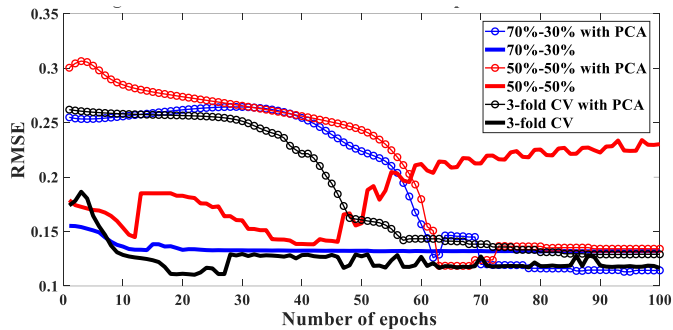


Figure 7

RMSE values for FCM clustering for different data split methods

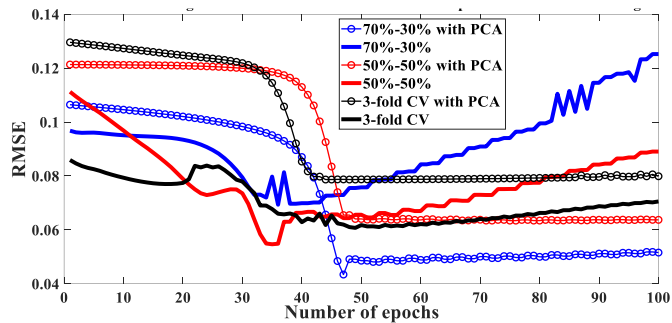


Figure 8

RMSE values for Subtractive clustering for different data split methods

In the case of subtractive clustering (Figure 8), we also get different results when the old and the upgraded algorithm are used, for the three splitting methods. Nevertheless, it is evident that for this type of input space partitioning, the upgraded algorithm with PCA shows better results, as after about 45 epochs it is stable for all the three types of splitting, and it does not overfit, as is the case with the old algorithm.

In order to conclude the comparisons, in Table 1 and Table 2 test set accuracies of the old and the upgraded algorithm are presented for the three different data splitting methods, using the three different input space partitioning methods.

Firstly, we choose to make comparative analysis using 40 training epochs (Table 1), as for some of the cases, the first algorithm overfits after 40 epochs (Figure 5, Figure 7, and Figure 8). On the other hand, as the upgraded algorithm in some of the cases needs more than 40 epochs for training, we choose to make comparative analysis of the both algorithms using 60 epochs for training (Table 2).

Table 1

Accuracies Obtained for Grid Partitioning, FCM and Subtractive clustering, for the old and the upgraded algorithm, when 40 epochs for training are used

Grid partitioning				70-30%	50-50%	3-Fold
Filter	Wavelet	MFs type	Old versus upgraded algorithm			
<i>FIR</i>	db2	Gauss MFs	Old algorithm from [2], and [3]	95	87	95.51
			Upgraded algorithm with PCA	97.33	89	96.03
Fuzzy c-means clustering (number of clusters = 2)				70-30%	50-50%	3-Fold
Filter	Wavelet	MFs type	Old versus upgraded algorithm			
<i>FIR</i>	db2	Gauss MFs	Old algorithm from [2], and [3]	98.25	95.41	96.39
			Upgraded algorithm with PCA	83.43	82.21	85.33
Subtractive clustering (radius of coverage – influence = 0.8)				70-30%	50-50%	3-Fold
Filter	Wavelet	MFs type	Old versus upgraded algorithm			
<i>FIR</i>	db2	Gauss MFs	Old algorithm from [2], and [3]	98.07	99.45	99.59
			Upgraded algorithm with PCA	97.21	93.33	98.22

Table 2

Accuracies Obtained for Grid Partitioning, FCM and Subtractive clustering, for the old and the upgraded algorithm, when 60 epochs for training are used

Grid partitioning				70-30%	50-50%	3-Fold
Filter	Wavelet	MFs type	Old versus upgraded algorithm			
<i>FIR</i>	db2	Gauss MFs	Old algorithm from [2], and [3]	98.33	97	92.96
			Upgraded algorithm with PCA	96.67	98	96.53
Fuzzy c-means clustering (number of clusters = 2)				70-30%	50-50%	3-Fold
Filter	Wavelet	MFs type	Old versus upgraded algorithm			
<i>FIR</i>	db2	Gauss MFs	Old algorithm from [2], and [3]	99.74	90	98.98
			Upgraded algorithm with PCA	98.85	97.93	95.33

<i>Subtractive clustering (radius of coverage – influence = 0.8)</i>				70-30%	50-50%	3-Fold
<i>Filter</i>	<i>Wavelet</i>	<i>MFs type</i>	<i>Old versus upgraded algorithm</i>			
FIR	db2	Gauss MFs	Old algorithm from [2], and [3]	92.21	93.33	97.43
			Upgraded algorithm with PCA	99.5	98.63	97.94

From Table 1 it is evident that the upgraded algorithm in most of the cases have worse performance than the old algorithm, but this is expected, as according to Figure 5, Figure 7, and Figure 8, the upgraded algorithm needs more than 40 epochs for training. On the other hand, the test set accuracies for both algorithms, trained with 60 epochs, show better performance in most of the cases for the upgraded algorithm.

In general, it is evident that the upgraded algorithm has satisfactory performance, and in some cases performs even better than the old algorithm, although the number of features is significantly reduced (from 20 to 7), which also plays a crucial role in making the new algorithm more resistant to overfitting.

As in our previous study [2] [3], the papers are concluded giving comparison of different relevant works, that have also used the Bonn database [10], here we expand that comparison. The results are shown in Table 2.

We have to note that the authors in [14], also use WT and ANFIS, but our approach differs as we use FIR filtering, as well as normalization. Our approach gives similar, or even better results, they have a 98.63% test accuracy of the test set Z (healthy), and a 98.25% test accuracy on the test set S (epileptic) patients. However, they make 5 class classification, which differs from our 2 class classification, for we did not summarize their results in Table 2.

Table 3

Comparison between accuracy in this study and other related studies

Related studies	Test set Accuracy (%)
Old algorithm-Grid Partitioning [2]	98.33
Old algorithm-FCM [3]	99.74
Old algorithm-Subtractive clustering [3]	99.59
Upgraded algorithm with PCA - Grid Partitioning	98
Upgraded algorithm with PCA – FCM	95.33
Upgraded algorithm with PCA – Subtractive clustering	99.5
E.Juarez-FFANN [23] (WT and NN, using six features. Several filters and wavelets were used, namely,	93.23

Haar, Db2 and Db4, getting 93.23% as the highest accuracy)	
I.Overhodzic-Wavelet+NN [24] (Wavelet and NN. DWT with Multiresolution analysis (MRA), based on db4 was used)	94

Conclusions

This paper represents a continuation of our previous work, representing, an extension of the previously presented algorithm, reducing the number of used features, using PCA. In order to show that the upgraded algorithm performs in a similar, or even better method, as related to the old algorithm, several comparative analyses were made, comparing the old and the upgraded algorithm for different data splitting methods (70%-30%, 50%-50%, and 3-Fold cross validation), and the different types of input space partitioning methods (grid partitioning, versus FCM clustering, versus subtractive clustering). As in our previous study, it was concluded that the combination of db2 wavelets, and the Gauss MFs gave satisfactory results, here the comparisons were made using only db2 wavelets and Gauss MFs. The comparisons of the accuracies of the old and the upgraded algorithm for the different data split methods and the different input space partitioning were made both for 40 and 60 epochs, as with the old algorithm, in some of the cases, overfits after 40 epochs and 60 epochs seemed a reasonable sublimata according the RMSE simulation results. It is well known that size reduction methods are preferred, in applications where the number of input parameters in the data set is too large [31]. In this sense, the PCA in the upgraded algorithm plays a key role, i.e. although it sufficiently reduces the data dimension features from 20 to 7, according to the results, the essence of the original data is preserved, but the algorithm has a potential of greatly reducing the computational costs.

References

- [1] U. Rajendra Acharya, S. Vinitha Sree, G. Swapna, Roshan Joy Martis, Jasjit S. Suri, "Automated EEG analysis of epilepsy: A review," in Knowledge-based Systems, Vol. 45, pp. 147-165, June 2013
- [2] Marjan Stoimchev, Vesna Ojleska Latkoska, "Detection of Epilepsy Using Adaptive Neuro-Fuzzy Inference System," Journal of Electrical Engineering and Information Technologies, in press
- [3] Marjan Stoimchev, Vesna Ojleska Latkoska, "Comparative Analysis for the Influence of the Tuning Parameters in the Algorithm for Detection of Epilepsy Based on Fuzzy Neural Networks," in Proceedings of the 14th International Conference - ETAI 2018, Struga, R.Macedonia, September 20-22, 2018

-
- [4] Subasi, A. and Gursoy, M. I.: 'EEG Signal Classification Using PCA, Ica, LDA and Support Vector Machines', *Expert Systems with Applications*, 2010, 37, (12), pp. 8659-8666
- [5] Acharya, U. R., Sree, S. V., Alvin, A. P. C., *et al.*: 'Use of Principal Component Analysis for Automatic Classification of Epileptic EEG Activities in Wavelet Framework', *Expert Systems with Applications*, 2012, 39, (10), pp. 9072-9078
- [6] Siuly, S. and Li, Y.: 'Designing a Robust Feature Extraction Method Based on Optimum Allocation and Principal Component Analysis for Epileptic EEG Signal Classification', *Computer methods and programs in biomedicine*, 2015, 119, (1), pp. 29-42
- [7] D. Najumnissa, T. R Rangaswamy, "Detection and Classification of Epilepsy Seizures using Wavelet feature extraction and Adaptive Neuro-Fuzzy Inverence System," *International Journal of Computational Engineering Research*, Vol. 2, pp. 755-761, May-June 2013
- [8] P. Sinha, *Speech Processing in Embedded Systems*, Springer Science+Business Media, LLC 2010, pp. 25-32
- [9] S. H. Mneney, *An Introduction to Digital Signal Processing: A Focus on Implementation*, River Publishers, 2008, pp. 153-158
- [10] R. G Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David, C. E. Elger, "Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3
- [11] A. Nakate, P. D Bahirgonde, "Feature Extraction of EEG Signals using Wavelet Transform," *International Journal of Computer Applications*, Vol. 124, No. 2, August 2015
- [12] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Computers in Biology and Medicine*, Vol. 37, Iss. 2, pp. 227-244, December 2005
- [13] İ. Güller, E. D. Übeyli, "Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction," *Expert Systems with Applications*, Vol. 27, Iss. 3, pp. 323-330, October 2004
- [14] İnan Güller, Elif Derya Übeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *Journal of Neuroscience Methods*, Vol. 148, Iss. 2, pp. 113-121, 30 October 2005
- [15] J. S. R Jang, C. T Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing-A Computational Approach to Learning and Machine Intelligence*, Prentice Hall Upper Saddle River, 1997

- [16] N. Pandey, N. Tiwari, "Predictive Accuracy of Modified Subtractive Clustering Algorithm on Large Dataset," *International Journal of Research and Development in Applied Science and Engineering (IJRDASE)*, Vol. 8, Iss. 2, December 2015
- [17] <https://www.mathworks.com/products/matlab.html>
- [18] Z. Mustaffa, Y. Yusof, "A Comparison of Normalization Techniques in Predictiong Dengue Outbreak," *Internation Conference on Business and Economics Research*, Vol. 1, 2011
- [19] Z. Omary, F. Mtenzi, "Machine Learning Approach to Identifying the Dataset Treshold for the Performance Estimators in Supervised Learning," *International Journal for Informatics (IJI)*, Vol. 3, Iss. 3, September 2010
- [20] S. S. Ghuman, "Clustering Techniques-A Review," *International Journal of Computer Science and Mobile Computing*, Vol. 5, Iss. 5, May 2016
- [21] P. Domingos, "A Few Useful Things to Know about Machine Learning," *Communications ACM*, Vol. 55, Iss. 10, pp. 78-87, October 2012
- [22] B. K. Singh, K. Verma, A. S. Thoke, "Investigations of Impact of Feature Normalization Techniques on Classifier's Performance in Breast Tumor Classification," *International Journal of Computer Applications*. Vol. 116, No. 19, April 2015
- [23] E. J. Guerra, V. A. Aquino, P. G. Gil, "Epilepsy Seizure Detection in EEG Signals Using Wavelet Transforms and Neural Networks," *Computer, Information and Systems Sciences and Engineering (CISSE)*, December 12-14, 2013
- [24] I. Omerhodzic, S. Avdakovic, A. Nuhanovic, K. Dizdarevic, *Energy Distrubution of EEG Signals: EEG Signal Wavelet-Neural Network Classifier*, <https://arxiv.org/abs/1307.7897v1>, July, 2013
- [25] L. Wang, W. Xue, Y. L., M. Luo, L. Huang, W. Cui, C. Huang, "Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis," *Entropy*, Vol. 19, Iss. 6, 27 May 2017
- [26] H. Adeli, Z. Zhou, N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *Journal of Neuroscience Methods*, Vol. 123, Iss. 1, pp. 69-87, 15 February 2003
- [27] H. Adeli, S. G. Dastidar, *Automated EEG-Based Diagnosis of Neurological Disorders: Inventing the Future of Neurology*, Taylor and Francis Group, 2010
- [28] T. Wen, Z. Zhang, *Effective and Extensible Feature Extraction Method Using Genetic Algorithm-Based Frequency-Domain Feature Search for Epileptic EEG Multi-classification*, <https://arxiv.org/abs/1701.06120v1>, January, 2017

- [29] P. K. Bhatia, A. Sharma, "Epilepsy Seizure Detection Using Wavelet Support Vector Machine Classifier," *International Journal of Bio-Science and Bio-Technology*, Vol. 8, No. 2, pp. 11-22, 2016
- [30] A. F. Rabbi, R. F. Rezai, "A Fuzzy Logic System for Seizure Onset Detection in Intracranial EEG," *Computational Intelligence and Neuroscience*, Vol. 2012, 4 November 2011
- [31] Ahmet Yüksek, Halil Arslan, Oguz Kaynar, Emre Delibaş, "Comparison of the Effects of Different Dimensional Reduction Algorithms on the Training Performance of Anfis (Adaptive Neuro-Fuzzy Inference System) Model," *Cumhuriyet Science Journal*, Vol. 38-4, pp. 716-730, 2017