ICACSIS 2017

2017 International Conference on Advanced Computer Science and Information Systems

CERTIFICATE OF PRESENTER

Is Awarded To

Elly Matul Imah

Dean of Faculty of Computer Science
Universitas Indonesia,

Mirna Adriani, Ph.D.

ICACSIS 2017 General Chair,

Prof. Dr. Eng. Wisnu Jatmiko

ICACSIS

FACULTY OF COMPUTER SCIENCE | KANTOR PENGELOLAAN PRODUK RISET & INOVASI | IEEE INDONESIA SECTION | FILKOM | UNESA | Departemen Ilmu Komputer

A Comparative Study of Machine Learning Algorithms for Epileptic Seizure Classification on EEG Signals

Elly Matul Imah
Mathematics Department
Universitas Negeri Surabaya
Surabaya, Indonesia
ellymatul@unesa.ac.id

Arif Widodo
Electrical Engineering Department
Universitas Negeri Surabaya
Surabaya, Indonesia

***Abstract*—Electroencephalography (EEG) is a tool for monitoring brain activity which is important for identifying epilepsy seizure. Automatic epileptic seizure identification in EEG is a challenging task and useful for helping neurophysiologists. This study compares some algorithms in machine learning algorithm that combine features extraction and classification algorithm for epilepsy seizure identification based on EEG data. The classification algorithms compared in this study are Generalized Relevance Learning Vector Quantization (GRLVQ), Backpropagation, SVM, and Random Forest, combined with Wavelet and PCA feature extraction. The EEG signals used in this study were obtained from EEG dataset which was developed by University of Bonn. EEG epilepsy seizure dataset has five classes. Class A and B are from five healthy subjects in open and closed eyes. Class C, D, and E from five elliptic subjects, where C and D are no-seizure signals, and E contains only seizure signal. The tasks that are used to compare the performance of feature extraction and classification algorithm is classifying 5 classes of EEG epilepsy seizure on EEG dataset. The measurements for evaluating methods are: accuracy, recall, precision training and testing times. The best feature extraction method at our experiment is PCA. The best performance in recognizing the five classes in EEG epileptic seizure dataset is GRLVQ, with the accuracy, precision and recall is 0.9866 and testing time is less than 0.1 seconds.**

Keywords*—EEG, epilepsy seizure,GRLVQ, SVM, Random Forest, Backpropagation, PCA, Wavelet

I. INTRODUCTION

Epilepsy is a neurological disorder that causes deterioration in consciousness and also leads to random and frequent body convulsions. The diagnosis and analysis of epilepsy widely used Electroencephalogram (EEG). EEG is signal that records electrical activities of the cerebral cortex by measuring electrical potential of neuron. Manual analysis and identification epilepsy seizure from EEG data by neurologist is very time consuming, and difficult to reliable visual inspection[1]. Automatic epilepsy seizure detection from EEG signal has attracted much interest for research because it can save neurologist from searching for seizure in a sheer amount of EEG data.

Automatic EEG signal classification is not easy and still challenging. There is no apparent difference in EEG signal between non-epileptic seizures in people that close eyes to epileptic seizure patient. It is also difficult to recognize EEG signal between epileptic seizures to non-epileptic seizures in patient with the region of the tumor. Many researchers have published their study on automatic epileptic seizure, but most of which only used two classes, epileptic seizure EEG signal and non-epileptic seizure EEG signal. In machine learning domain, it is more difficult to classify five classes than two classes. If we can classify correctly the EEG data in five classes that was annotated in the dataset, it is very helpful for neurologist and medical practice to diagnose their patients.

The EEG dataset used to identification epileptic seizure may contain some redundant information and noise signal. Some researchers have published their study that focus on feature extraction. Sharma et all, proposed a new method for EEG signal feature extraction to identify epileptic seizure[2]. Sharma used analytic time-frequency flexible wavelet transform and fractal dimension to approach characteristics of epileptic seizures. Wang et al published their study on an automatic epileptic seizure detection that focused on feature extraction, in two classes, epileptic seizure and non-seizure in EEG signal[3]. Pippa et al publish their study that also focus on feature extraction for identification epileptic seizure[4]. Chen et al studied feature extraction for epileptic seizure classification in EEG using Wavelet-Fourier Feature[5]. Birjandtalab et al studied non-linear dimension reduction for feature extraction to classify epileptic seizure, they reported that non-linear PCA had good performance[6]. Sharma et al and Wang et al suggest that wavelet has good performance for the identification of the epileptic seizure in EEG signal. Pippa et al studied suggest multi-array decomposition has good performance. Based on their study, we compare wavelet and multi-array decomposition as feature extraction for identification epileptic seizure in EEG signal.

978-1-5386-3172-0/17/$31.00© 2017 IEEE

The machine learning research in epileptic seizure identification is interesting to study, not only in feature extraction domain, but also in classification-algorithm domain. Karlik et all studied that k-NN, Backpropataion, SVM have a better performance than Naïve Bayes as algorithm for Epileptic Seizures in EEG[7]. Subadi et al used hybrid SVM for epileptic seizure detection in EEG[8], the result shown good performance but it is complex algorithm that need many memory resources. Ebrahimpour et al used ANN ensemble method and wavelet transform, but the result showed an average performance[9]. Based on some literatures in this study we select Generalized Relevance Learning Vector Quantization (GRLVQ) as classification algorithm because this algorithm is rarely used for epileptic seizure identification in EEG, although this algorithm is very robust algorithm especially for multiclass classification problem. The result of GRLVQ will be compared to the other algorithm that has been used by other researchers and has reported has good performance.

Compared to the other algorithm, GRLVQ is rarely used in epileptic seizure identification in EEG signal. GRLVQ is competitive based learning that uses prototypes of each class to classify. Prototype is determined in the training process from training dataset and captures the essential features of the data in the same space[10]. GRLVQ is a modification of Relevance Learning Vector Quantization (RLVQ) by using adaptive metric. GRLVQ is proposed by Hammer et al and used stochastic gradient descent on an energy function[11]. Based on that, we choose GRLVQ as classification algorithm for classify five classes from EEG dataset which is developed by University of Bonn.

This paper is organized as follow, section II describes Epileptic seizure in EEG dataset and preprocessing methods used in this study. Section III describes the basic concept of machine learning classification algorithm used. Section IV presents the experimental setup, result, and discussion of the study. Section V presents the conclusion of the study.

II. Epileptic Seizure in EEG Preprocessing and Feature Extraction

A. *Eliptical Seizure in EEG Preprocessing*

EEG signal proved significant information of neurological condition and other neuro disorders. There are five frequency bands which are usually used for EEG signal analysis, Delta (up to 4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-26 Hz), and Gamma (26-100 Hz). Dataset was used in this study taken from EEG dataset which is developed by University of Bonn. The dataset is recording brain activity of five healthy subjects and five epileptic subjects for 23.6 seconds. The dataset contains 100 single-channel EEG segment, and then is classified into five classes. Class A is EEG signal from a healthy subject in an awake state with opened eyes. Class B consists of EEG signal from a healthy subject with closed eyes. Class C, D, and E, are taken from five difference patient EEG records that are archives of pre-surgical diagnosis. Class C and D, contain seizure-free activity, while set E only contains seizure activity. EEG signals record at 173,61Hz using 12 bit A/D converter. This study classifies five classes, A, B, C, D, and E.

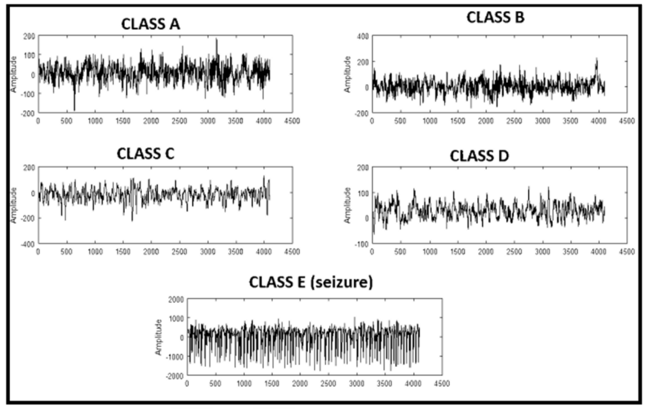

Fig. 1. Samples of EEG signal in each classes

The visualization of EEG signal epileptic seizure and free epileptic seizure in differences condition was presented in Fig.1. A single data consists of 4097 data points. In this study a single EEG signal will be partitioned into 17 segments, such that every segment consists of 241 data points. In UCI ML, epileptic seizure dataset that has been uploaded in May, 2017 partitioned from EEG dataset which is developed by University of Bonn, a single signal to be some segment, such that every segment has 170 data points. We did not use those datasets, because in partition some data points from original data signal has been removed. Some data points that have been removed from original data might dangerous, so we partition dataset by cutting every 241 points, so we get 17 segments. We divide a single EEG dataset into 17 segments because 17 is the biggest first prime factor of 4097, where 4097 is the number of original data points. The result of partitions in a single EEG signal can be seen on Fig. 2.

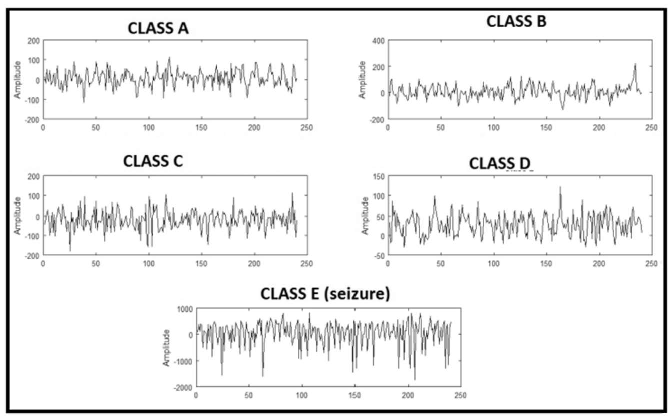

Fig. 2. Sample of EEG signal each classes after partition

B. *EEG Feature extraction*

This study used two kind of feature extraction, Principal Component Analysis (PCA) and Wavelet Transform.

- Principal Component Analysis (PCA)

PCA is a feature extraction method that simplifies the complexity in high-dimensional data while retaining as much as number of interrelated variable and patterns. PCA works by

978-1-5386-3172-0/17/$31.00© 2017 IEEE

transforming geometrically projecting of dataset into lower dimension, it is called principal components (PCs). Nowadays PCA is still the one of the most effective algorithm for feature extraction[12] and used as baseline algorithm in feature extraction algorithm. For given X, which is X is p-dimension dataset, the$m$ principal axes $U_1, U_2, U_3, \dots, U_m$, where $1 \leq m \leq p$, are orthonormal axes onto which the retained variance is maximum in the projection space. Set of U are given from m leading eigenvectors of sample covariance matrix X. Visualization of PCA feature extraction result of epileptic seizure in EEG signal, show in Fig.3.

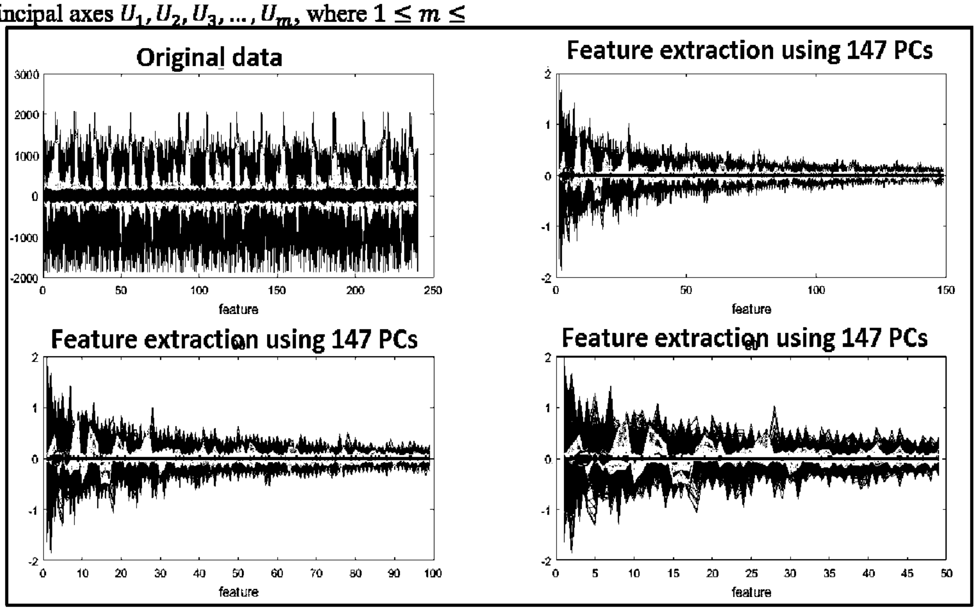

Fig. 3. Visualization of feature extraction result using PCA

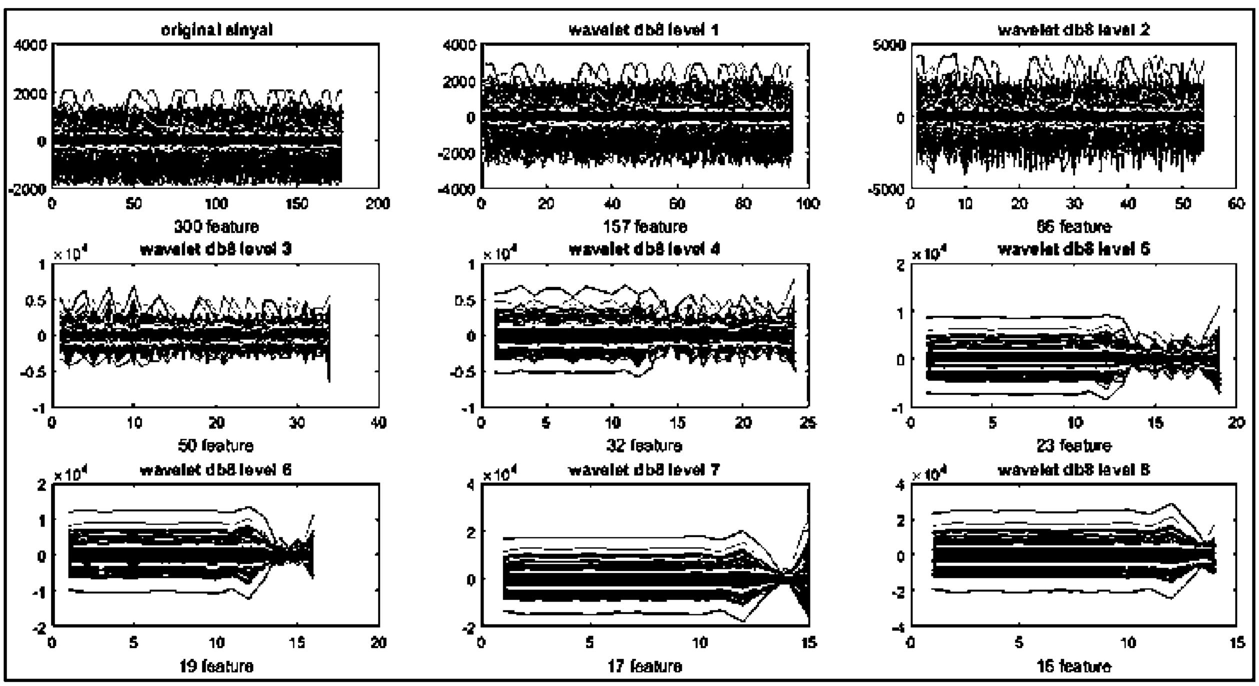

Fig. 4. Visualization of feature extraction result using WT

- Wavelet Transform

  Wavelet feature extraction has reported good performance in identifying epileptic seizure in EEG datasets[2], [7]. Wavelet Transform (WT) feature extraction works in time domain, different from PCA that works in special domain. WT of a signal $f(x)$ is defined as:

$$W_s f(x) = f(x) * \psi_s(x) = \frac{1}{s}\int_{-\infty}^{+\infty} f(t)\psi(\frac{x-t}{s})\,dt \quad (1)$$

Which s is scale factor, $\Psi_s(x) = \frac{1}{s}\Psi(\frac{x}{s})$ is the dilation of a basic wavelet $\Psi(x)$ by the scaling factor s. Let $s = 2j$ ($j \in Z$, and Z is the integer), then the WT is called dyadic WT[13]. In wavelet theory, selecting the appropriate mother wavelet and the number of decomposition level is an important part. The proper selection aims to retain the important part of information and still remain in the wavelet coefficients. The Mother wavelet that we used in this study is Daubechies. Result visualization of epileptic seizure in EEG dataset can be seen on Fig.4.

This study used PCA and WT as feature extraction algorithm that are combined with classification algorithms. The results will be compared each other.

III. Classification Algorithm

A. Random Forest

Random forest algorithm is classification algorithm that is basically based on random tree. In random forest, every input feature vector is compared to the one stored in the train dataset in order to find the best match. Growing an ensemble of random trees for recognition using a probabilistic scheme is called random forest of trees. Recognition accuracy is as high as the trees vote for the most popular class. Trees drawn at random from a set of possible trees is called random tree. Random tree is a decision tree that considers k randomly chosen attributes at each node. The class probabilities on each node are based on back fitting with no pruning [14]. The steps involved in growing a random tree are as follow:

1. The training set for growing the tree is obtained by selecting N cases at random but with replacement from original dataset.
2. A random number of attributes m are chosen for each tree. The attributes from the nodes and leaves using standard tree building algorithms. The best split on m is used to split the nodes and m is held constant.
3. Each tree is growing to the fullest extent possible without pruning.

A new object is classified using its input vector down each of the trees in the forest. The forest chooses the class with the most vote, the new object input vector is classified.

B. Support Vector Machine (SVM)

The SVM was proposed using basic concept of Maximal Margin Classifier, which is applicable for linearly separable data. It is simple to understand the basic ideas behind more sophisticated SVMs. Consider a linearly separable dataset $\{(X_i, d_i)\}$, where X_i is the input pattern for the i:th example and d_i is the corresponding desired output {-1, 1}. The assumption, *the dataset is linearly separable*, means that there exists a hyper plane working as the decision surface. We can write:

$$\begin{aligned} W^T X_i + b \geq 0, then d_i = +1 \\ W^T X_i + b \leq 0, then d_i = -1 \end{aligned} \quad (2)$$

where $W^T X_i + b$, is the output function. The distance from the hyper plane to the closest point is called the geometric margin. The idea is, to have a good machine, so the geometric margin needs to be maximized. First, we introduce the marginal function $W^T X_i + b$ because the dataset is linearly separable we can rewrite as (3), as follow:

$$\begin{aligned} W^T X_i + b = +1 \\ W^T X_i + b = -1 \end{aligned} \quad (3)$$

where $X^+(X^-)$ is the closest data point on the positive (negative) side of the hyperplane. Now it is straight forward to compute the geometric margin.

$$\begin{aligned} \gamma &= \frac{1}{2}\left(\frac{W^T X^+ + b}{|w|} - \frac{W^T X^- + b}{|w|}\right) \\ &= \frac{1}{2|w|})W^T X^+ + b - W^T X^- - b) \\ &= \frac{1}{2|w|}(1-(-1)) = \frac{1}{|w|} \end{aligned} \quad (4)$$

Hence, equivalent to maximize the geometric margin is fixing the functional margin to one and minimizing the norm of the weight vector |w|. This can be formulated as a quadratic problem with inequality constraints

$d(w^T x_i + b) \geq 1.$

$$min: \frac{1}{2} W^T W \ (quadratic - problem) \quad (5)$$

subject to: $d(w^T x_i + b) \geq 1$

By the use of Lagrange multipliers$\alpha_i \geq 0$ the original problem is transformed into the dual problem. From the Kuhan–Tuker theory we have the following condition:

$$\alpha_i[d_i(W^T x_i + b) - 1] = 0 \quad (6)$$

It means that only the points with functional margin unity contribute to the output function. These points are called the Support Vectors, which support the separating hyper plane. Non-linear SVM has been added Mercer Theorem for handling feature transformation into high dimension space.

C. Backpropagation

Backpropagation is one of Artificial Neural Network (ANN) algorithms that develop from Multi Layers Perceptron (MLP) by adding delta rule as backward phases. The basic idea of Backpropagation is to efficiently computing partial derivatives of an approximating function realized by the network and toward the entire processing element (neuron). It is an adjustable weight vector for a given value of input vector. Using nonlinear activation function can be easy to classify non-linearly separable data. Backpropagation and many ANN algorithms have special advantaged for solving multiclass classification task. Architecture Network of Backpropagation in this study has 3 layers, input, hidden, and output layers. Output layers consist of 5 neurons.

D. Generelaized Relevance Learning Vector Quantization (GRLVQ)

GLRVQ is a competitive based learning classification algorithm modified of GLVQ. GLVQ has been proposed by A. Sato Yamada, using a steepest descent method which minimizes a cost function to define the codebook or prototype vectors update [15]. Relative distance difference is defined as (7) and cost function as (8).

$$\mu(x) = \frac{d_j - d_k}{d_j + d_k} \quad (7)$$

$$S = \sum_{i=1}^{N} f(\mu(x_i)) \quad (8)$$

where N is number of input vector and f is a monotonically increasing function.

In GRLVQ the distance was modified using weighted distance between input vector x_i and a codebook vector w_j[16]:

$$D_{ij} = \sqrt{\sum_{k=1}^{N} \lambda_k (x_{ik} - w_{jk})^2} \quad (9)$$

where $\sum_{k=1}^{N} \lambda_k = 1$. Modification of distance formula, Eq. (7) must be reformulated to minimized and objective function based on this modified distance as in (10).

$$\mu_\lambda(x_i) = \frac{D_{ij} - D_{ik}}{D_{ij} + D_{ik}} \quad (10)$$

Obtained a modified rule of GLVQ, which is the GRLVQ rule:

$$\Delta w_j = \pm \eta \lambda \mathrm{I} \frac{\partial f}{\partial \mu} \frac{D_{ij}}{(D_{ij} + D_{ijk})^2} (x_i - w_j) \quad (11)$$

If x_i and w_k are different classes, the sign of Δw_j is (+), and if different classes is (-).

The relevance is updating using Eq. 12.

$$\lambda^{(t+1)} = \lambda^{(t)} - \alpha \frac{1}{4\sigma^2} G(y_1 - y_2, 2\sigma^2 \mathrm{I}).(y_2 - y_1)\mathrm{I}.(x_1 - w_{j(1)} - x_2 + w_{j(2)}) \quad (12)$$

Update on-line both the relevance and feature ranks algorithm as follow:

1. Initialize , α, and relevance vector $\lambda_k = \frac{1}{n}, k = 1, \dots, n$.
2. Initialize codebook vector.
3. Update codebook vector using Eq. (11).
4. Update the relevance vector using Eq. (12).
5. Normalize the relevance vector.
6. Compute the weight of each feature as an average of its before ordering position index in the input vector, for all previous steps.
7. Repeat step 3-6 for each training pattern.

IV. Result and Discuss

A. Experimental Setup

The task used to evaluate the performances of machine learning that we study to compare is to classify five classes of epileptic seizure in EEG database which is developed by University of Bonn. Different from the UCI-ML epileptic seizure dataset set up, in this study a single EEG signal we partition in 17 segments, such that every segment consists of 241 data point. In UCI ML, a single segment only has 170 data points, this partition is not suitable because some data points from original data signal has been removed. We divide a single EEG dataset in 17 segments because 17 is a biggest first prime factor of 4097, therefore we have five classes so every class has 1700 instances. The detail of distribution of data can be seen in Table 1.

TABLE I. EPILEPTIC SEUZURE DATASET

Class	Original Dataset		After partition	
	# attributes	***#instances***	***# attributes***	***#instances***
A	4097	100	241	1700
B	4097	100	241	1700
C	4097	100	241	1700
D	4097	100	241	1700
E	4097	100	241	1700

The original Epileptic seizure datasets consist of five different classes, but many classified the dataset in tow classes, namely epileptic seizure against non-seizure or free-seizure. Therefore the task to classify five classes is still challenging area. This study classified the data sets into five classes based on the original class provided by the University of Bonn. In this experiment, we used full feature attributes and extracted feature using PCA and WT, then we classified them by using Random Forest, SVM, Backpropagation, and GRLVQ. The training and testing data ratio in this study are 2:1.

B. Experiment Result

The experiment in this study to do tasks for classifying EEG datasets into five classes, using Random Forest, SVM, Backpropagation, and GRLVQ in full feature condition. The detail of the accuracy, precision, and recall can be seen in Table 2.

TABLE II. EXKPERIMEN RESULT

Methods	Evaluation Measure		
	Accuracy	***Recall***	***Precision***
Randome Forest	0.9788	0.9792	0.9744
SVM	0.833	0.8449	0.8367
Backpropagation	0.9529	0.9537	0.9529
GRLVQ	**0.9866**	**0.9866**	**0.9866**
WT+Randome Forest	0.9805	0.9764	0.9992
WT+SVM	0.8222	0.828	0.822

WT+Backpropagation	0.9546	0.9551	0.9546
WT+GRLVQ	0.9761	0.9762	0.9761
PCA+Randome Forest	0.9822	0.983	0.982
PCA+SVM	0.8444	0.8511	0.8458
PCA+Backpropagation	0.955	0.9556	0.955
PCA+GRLVQ	0.9845	0.9846	0.9845

Not only in full feature, this study also does the task using feature extraction combined as WT+Random Forest, WT+SVM, WT+Backpropagation, WT+GRLVA, PCA+Random Forest, PCA+SVM, PCA+Backpropagation, and PCA+GRLVQ. Table 2 shows that GRLVQ and Random Forest have a performance better than the others. The accuracy, precision and recall of GRLVQ in the same value 0.9866, it means that GRLVQ has a good performance in all classes. Random forest accuracy is 0.9822, precision is 0.983, and recall is 0.982, this performance obtained after extract the feature using PCA. Comparing the result of PCA and WT as feature extraction, from Table 2 can be seen that PCA has better performance than WT. GRLVQ has the best performance is in full feature condition than extracted features. GRLVQ used weighted distance that generally is proposed for selected feature, therefore if feature of the input data have been extracted or reduced some information are lose for selection by relevance factor in GRLVQ.

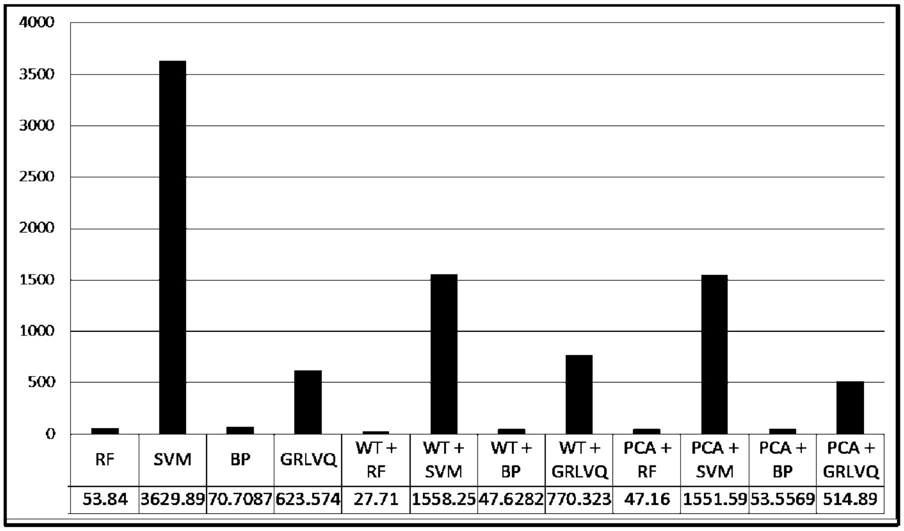

Fig. 5. CPU time of Training for build the models

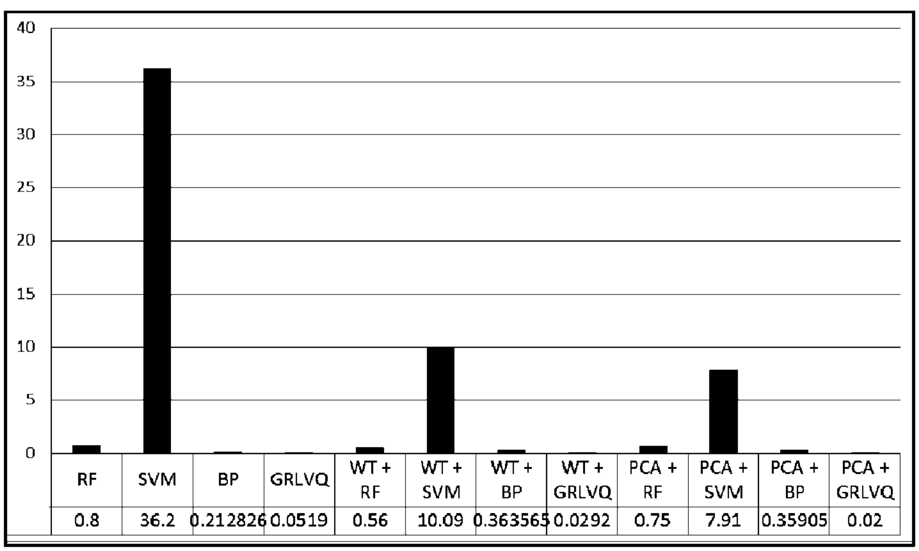

Fig. 6. CPU time of Testing for classify the data

Fig.5 shows that Random Forest is able to build the fastest models, it is better than the other classification methods that were compared in this study. SVM is the longest in building the models. The fastest algorithm in recognition is GRLVQ. The testing time or recognition time of GRLVQ is less than 0.1 seconds; it is faster than SVM, Random Forest, or Backpropagation because GRLVQ is the simplest algorithm that only used minimum distance of data to prototype. The details can be seen in Fig.6.

V. Conclussion

This study examined various feature extraction and classification algorithms for automatically classifying five classes annotation EEG signals provided by University of Bonn. The dataset records brain activities of five healthy subjects and five epileptic subjects for 23.6 seconds. The dataset contains 100 single-channel EEG segments, and was then classified into five classes. Class A is an EEG signal from a healthy subject in an awake state with opened eyes. Class B consists of an EEG signal from a healthy subject with closed eyes. Class C, D, and E, are taken from five different patients' EEG records that were the archives of pre-surgical diagnosis. Class C and D contain seizure-free activity, while set E only contains seizure activity. EEG signals record at 173,61Hz using 12 bit A/D converter.

The experiments used full feature attributes and extracted feature using PCA and WT, combined with classification algorithms: Random Forest, SVM, Backpropagation, and GRLVQ. The training and testing data ratio in this study are 2:1. The best feature extraction method in our experiment is PCA. The best performance in recognizing the five classes in EEG epileptic seizure dataset is GRLVQ, with the accuracy, precision and recall is 0.9866 and testing time is less than 0.1 seconds.

Acknowledgment

Thank you for Universitas Negeri Surabaya, especially for Vice Rector of Academic Deputy, Universitas Negeri Surabaya for supporting this research publication.

References

[1] H. Stefan *et al.*, "Objective quantification of seizure frequency and treatment success via long-term outpatient video-EEG monitoring: A feasibility study," *Seizure*, vol. 20, no. 2, pp. 97–100, 2011.

[2] M. Sharma, R. B. Pachori, and U. R. Acharya, "A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension," *Pattern Recognit. Lett.*, 2017.

[3] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, and W. Cui, "Automatic Epileptic Seizure Detection in EEG Signals Using Multi-Domain Feature Extraction and Nonlinear Analysis," *Entropy*, vol. 9, pp. 1–17, 2017.

[4] E. Pippa *et al.*, "EEG-based Classification of Epileptic and Non-epileptic Events using Multi-array Decomposition To cite this version : EEG-based Classification of Epileptic and Non-epileptic Events," 2017.

[5] G. Chen, W. Xie, and T. D. Bui, "Automatic Epileptic Seizure Detection in EEG Using Nonsubsampled Wavelet – Fourier Features," 2017.

[6] J. Birjandtalab, M. B. Pouyan, and M. Nourani, "Nonlinear Dimension Reduction for EEG-Based Epileptic Seizure Detection," *IEEE Trans. Biomed. Heal. Informatics*, pp. 595–598, 2016.

[7] B. Karlık and Ş. B. Hayta, "Comparison Machine Learning Algorithms for Recognition of Epileptic Seizures in EEG," pp. 1–12, 2014.

[8] A. Subasi, "Epileptic seizure detection using hybrid machine learning methods," 2017.

[9] R. Ebrahimpour, "EPILEPTIC SEIZURE DETECTION USING A NEURAL NETWORK ENSEMBLE METHOD," pp. 291–310, 2012.

[10] E. M. Imah, W. Jatmiko, and T. Basaruddin, "Electrocardiogram for biometrics by using adaptive multilayer generalized learning vector quantization (AMGLVQ): Integrating feature extraction and classification," *Int. J. Smart Sens. Intell. Syst.*, vol. 6, no. 5, 2013.

[11] M. Kästner, B. Hammer, M. Biehl, and T. Villmann, "Functional relevance learning in generalized learning vector quantization," in *Neurocomputing*, 2012, pp. 1–22.

[12] J. Lever, M. Krzywinski, and N. Altman, "Principal component analysis," *Nat. Publ. Gr.*, vol. 14, no. 7, pp. 641–642, 2017.

[13] R. Schneider and F. Kr, "Daubechies Wavelets and Interpolating Scaling Functions and Application on PDEs," pp. 1–44, 2007.

[14] N. Belgacem, "ECG Based Human Authentication using Wavelets and Random Forests," *Int. J. Cryptogr. Inf. Secur.*, vol. 2, no. 2, pp. 1–11, Jun. 2012.

[15] A. Sato and K. Yamada, "Generalized Learning Vector Quantization," in *Advances in Neural Information Processing Systems 8 Proceedings of the 1995 Conference*, 1996, vol. 7, pp. 423–429.

[16] A. Caţaron and R. Andonie, "Energy generalized LVQ with relevance factors," *IEEE Int. Conf. Neural Networks - Conf. Proc.*, vol. 2, pp. 1421–1426, 2004.

978-1-5386-3172-0/17/$31.00© 2017 IEEE