FUSION OF BAGGING BASED ENSEMBLE FRAMEWORK FOR EPILEPTIC SEIZURE CLASSIFICATION

Farrikh Alzami[1*], Aries Jehan Tamamy[2], Ricardus Anggi Pramunendar[1] dan Zaenal Arifin[2]

[1]Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia
[2]Faculty of Technic, Universitas Dian Nuswantoro, Semarang, Indonesia

*)E-mail: alzami@dsn.dinus.ac.id

Abstract

The ensemble learning approach, especially in classification, has been widely carried out and is successful in many scopes, but unfortunately not many ensemble approaches are used for the detection and classification of epilepsy in biomedical terms. Compared to using a simple bagging ensemble framework, we propose a fusion bagging-based ensemble framework (FBEF) that uses 3 weak learners in each oracle, using fusion rules, a weak learner will give results as predictors of the oracle. All oracle predictors will be included in the trust factor to get a better prediction and classification. Compared to traditional Ensemble bagging and single learner type Ensemble bagging, our framework outperforms similar research in relation to the epileptic seizure classification as 98.11±0.68 and several real-world datasets

Keywords: Epileptic seizure detection, wavelet analysis, feature selection, ensemble, fusion, bagging

Abstrak

Pendekatan *ensemble learning* terutama dalam klasifikasi sudah banyak dilakukan dan berhasil di banyak ruang lingkup, tetapi sayangnya tidak banyak pendekatan *ensemble* yang digunakan untuk deteksi dan klasifikasi epilepsi dalam hal biomedis. Dibanding menggunakan kerangka *ensemble* bagging sederhana, kami mengusulkan fusi kerangka *ensemble* berbasis *bagging* yang menggunakan 3 *weak learner* di setiap *oracle*. Dengan menggunakan aturan *fusion*, *weak learner* akan memberikan hasil sebagai prediktor *oracle*. Semua prediktor *oracle* akan dimasukkan ke dalam faktor kepercayaan untuk mendapatkan prediksi dan klasifikasi yang lebih baik. Dibandingkan dengan *Ensemble bagging* tradisional dan *single learner type Ensemble bagging*, kerangka kerja kami mengungguli penelitian sejenis pada kaitan dengan klasifikasi penyakit epilepsi sebesar 98.11±0.68 dan pada beberapa dataset dunia nyata.

Kata kunci: Deteksi kejang epilepsi, analisis wavelet, pemilihan fitur, ensemble, fusi, bagging

1. Introduction

Fisher et.al [1] explain that epilepsy is a group of neurological disorder that can be characterized by epileptic seizures which result by recurrent, spontaneous and abnormal cortical nerve cell activity in the brain. Shoeb [2] also mentions that Electroencephalogram (EEG) is widely used in epileptic seizure detection because epilepsy diagnosis can be performed by identifying the abnormalities from EEG. Since manual detection by expert neurologist is time consuming, expensive and concerns about accuracy caused by fatigue, using computer aided diagnosis for epilepsy diagnosis using EEG time series data analysis can be useful and efficient solution. Epileptic seizure EEG dataset usually contains several groups of subjects such as 1) epileptic subject during seizure (ictal), 2) epileptic subject during seizure free interval (inter-ictal) and 3) healthy subject [2].

Due to EEG signals have non-stationary nature, wavelet analysis is used by many researchers because wavelet transform can give precise time information at high frequencies and frequency information at low frequencies[3], [4] and [5]. Wavelet analysis using DWT can be used to preprocess the EEG signals by decomposing it into sub-bands level signals.

By using feature selection, Alzami et.al [6] explain that there are many features that can be extracted from EEG signal decomposition, but not all features contribute to the improvement of epileptic seizure classification performance. Alotaiby et.al [7] also provide comprehensive survey that most epileptic seizure detection and classification research focus on EEG signal preprocessing such as artifact removal, signal decomposition techniques and using single classifier rather than multi classifier such as ensemble learning.

Ensemble approach especially for classification is primary used for real world dataset problem rather than epileptic seizure detection and classification, some example is Ho [8] that using random subspace method for constructing decision forest (RSCE), Yu et.al [9] by using Hybrid Adaptive Classifier Ensemble (HAEL) that compare RSCE method with their research such as HAEL and singly adaptive ensemble learning (SAEL).

In order to improve the accuracy of epileptic seizure classification and to validate that ensemble approach is more robust and suitable for epileptic seizure classification, we design a fusion of bagging-based ensemble framework for epileptic seizure classification. Our contribution is two-fold: First, rather than using traditional ensemble bagging, we use fusion rules that calculate three different classifiers (weak learner) in every oracle bootstrap which can be used to improve the ensemble accuracy. Second, propose the most useful classifier fusion that can optimize and improve the bagging classifier performance for epileptic seizure classification.

The remainder of the paper is organized as follows. Section 2 describes the fusion of bagging-based ensemble framework (FBEF) methods. Section 3 experimentally investigates the performance of the FBEF methods and section 4 presents the conclusion and future works

2. Methods

Figure 1 provides the overview of the epileptic seizure classification approach. First, we set a band-pass FIR filter from 0.5 Hz to 60 Hz to preprocess the EEG raw data and extract 5 sub-band signals such as delta, theta, alpha, beta, gamma and band limited (0-60 Hz) using DWT db4 level 4. Then, features are extracted from each sub-band signal such as: embedding-dimension, fractal-dimension, correlation dimension, kurtosis, approximate-entropy, sample-entropy, standard-deviation, max, min, median, mean, skewness, time- lag, Largest-Lyapunov-Exponent and energy. These features then implemented into several feature selection such as MDEFS [6], mRMR [10] and genetic algorithm. Finally, all those features that been selected will be used with fusion of bagging-based ensemble framework.

Figure 1. Epileptic seizure classification overview

Figure 2 illustrates the framework of FBEF approach, specifically, after clean dataset is obtained from Figure 1, we split dataset into two parts, training set data (D) and test set data. Then, FBEF generates a set of datasets $D = \{D_1, D_2, \dots, D_B\}$ (where B is the total number of datasets) based on original training set using bagging technique.

Then, every dataset is feed into 3 different classifiers to obtains a set of classifiers $\Omega\{C\alpha_1, C\beta_1, C\gamma_1, \dots, C\alpha B, C\beta B, C\gamma B\}$. Next the set of classifiers Ω are evaluated through test set using decision profile (fusion) ϕ in same oracle respectively. Prediction results in every oracle is denoted as $\{L_1, L_2, \dots, L_B\}$. Finally, FBEF combines those prediction results based on a set of confidence factors Φ to obtain the final result. In here, classifier is denoted as weakLearn.
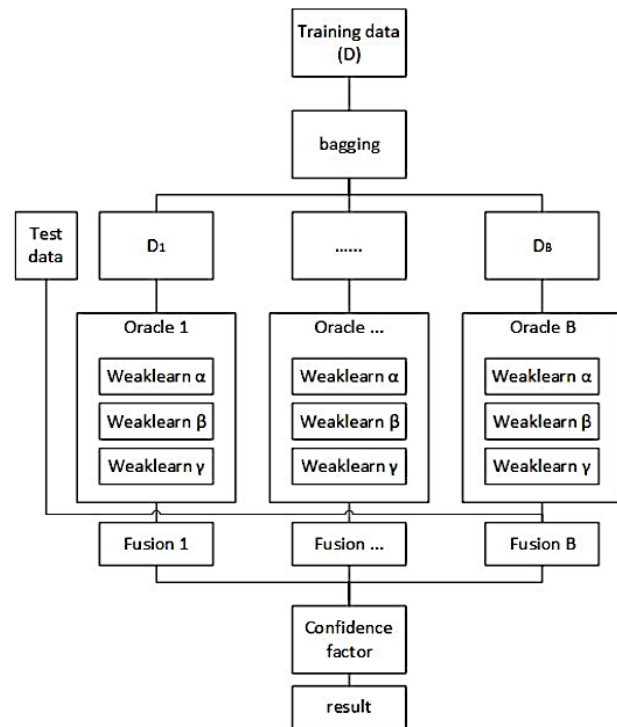

Figure 2. Fusion of Bagging-Based Ensemble Framework

We use min-max as scaling options to a fixed range because it can make standard deviations become smaller which the result is suppress the effect of outliers.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Using explanation from Xu.et.al [11], in every member of oracle Ω, every weakLearn α, β, γ will produce several Type of Information (TOI) [12], hypothesis that can be group into three levels:
- Abstract level (predicted class): each weakLearn outputs the class label for each input pattern.
- Measurement level (decision profile): each weakLearn outputs a posterior probability, score or confidence level for each input pattern
- Rank level: each weakLearn outputs a ranking list of possible classes for each input pattern (simply said, we sort the measurement level by descend sorting).

Fusion rule ϕ based on Kittler et.al [13] is used when all weakLearn in respective oracle already produce the TOI. Let's assume we have ϑ as possible classes $(\psi_1, \dots, \psi_\vartheta)$, η as weak learner, μ_i as measurement vector by ith η weak learner, class ψ_k in the measurement space can be modeled by probability density function $P(\mu_i|\psi_k)$, ψ_j as predicted class and $P(\psi_k|\mu_i)$ as measurement level output. After we have all those notations, we can put into fusion rule which is:

- Maximum: finds the maximum score of each class between the weakLearn using measurement level and assigns the input pattern to the class with the maximum score among the maximum scores.

$$max_{i=1}^{\eta}P(\psi_j|\mu_i) = max_{k=1}^{v}max_{i=1}^{\eta}P(\psi_k|\mu_i) \quad (2)$$

- Minimum: finds the minimum of each class between the weakLearn using measurement level and assign the input pattern to the class with the maximum score among the maximum scores

$$min_{i=1}^{\eta}P(\psi_j|\mu_i) = max_{k=1}^{v}min_{i=1}^{\eta}P(\psi_k|\mu_i) \quad (3)$$

- Average(sum): finds the mean of the scores of each class between the weakLearn using measurement level and assigns the input pattern to the class with the maximum score among the means.

$$avg_{i=1}^{\eta}P(\psi_j|\mu_i) = max_{k=1}^{v}avg_{i=1}^{\eta}P(\psi_k|\mu_i) \quad (4)$$

- Product: multiplies the score provided by each base weakLearn using measurement level and assigns the class label with the maximum score to given input pattern

$$\Pi_{i=1}^{\eta}(\psi_j|\mu_i) = max_{k=1}^{v}\Pi_{i=1}^{\eta}P(\psi_k|\mu_i) \quad (5)$$

- Majority vote: using abstract level from TOI, voting method finds what is the class output of each weakLearn and count it output as a vote for a class and assigns the input pattern to the class with the majority vote

$$\Upsilon_{ki} = \begin{cases} 1, if\ P(\psi_k|\mu_i) = \ max_{j=1}^{\mu}P(\psi_j|\mu_i) \\ 0, otherwise \end{cases} \quad (6)$$

$$\sum_{I=1}^{\eta}\Upsilon_{ji} = max_{k=1}^{v}\sum_{i=1}^{\eta}\Upsilon_{ki} \quad (7)$$

After all oracle is produced the prediction result, confidence factor Φ is used to get the final result. This confidence factor can also use fusion rule or just simply use majority vote.

3. Experiment and Results

We used Bonn dataset from Andrzejak [14] which is referred as SZN. This SZN means we use three classes of epileptic seizure subjects: S denoted as subject during epileptic seizure (ictal), Z denoted as normal data and N denoted as during seizure free interval (interictal). We chose three classes due to it represent real world situation, which is: healthy, patient when seizure and when the patient did not get seizure. These Dataset contains sub-band signals as follows: delta, theta, alpha and beta. Those four sub-band signals then extracted to obtain features such as: embedding-dimension, fractal-dimension, correlation dimension, kurtosis, approximate-entropy, sample-entropy, standard-deviation, max, min, median, mean, skewness, time- lag, Largest-Lyapunov-Exponent and energy. Those 15 features are fed into MDFS feature selection, mRMR feature selection and genetic algorithm feature selection. The number of obtained features is determined from the inner operation of respective feature selection mechanisms; thus, we did not set how many features that we will use.

The proposed fusion of bagging-based ensemble framework is measured by the average accuracy on the datasets. five-fold crossover validation is used to reduce the randomness effect. The number iteration of bagging T, number bagging Ω, scaling dataset σ is set, respectively. We also consider to not use the scaling into dataset to understand the effect of fusion classifier

The classifier combination in the experiment include k nearest neighbor classifier (kNN), the Naive Bayes classifier (NB), the decision tree classifier (DT), Discriminant analysis which use linear (LDA) and pseudoQuadratic (QDA), the support vector machine (libSVM), the levenberg-Marquardt back propagation neural network (LMBPNN).

In the following experiments, we first explore the effect of FBEF using fusion rules and 3 classifiers. Then, FBEF is compared with another epileptic seizure single classification. FBEF also compared with some ensemble framework which used the real datasets.

3.1. The Effect of FBEF Using Fusion Rules and Classifiers

In order to explore the effect of fusion rules, we compare 5 fusion rules into SZN dataset using 3 feature selection. Table 1 shows that (1) majority vote have tendency get better result because it will predict the label if at least half the number weaklearner plus 1 weaklearner give same answer. (2) average fusion rules can also be considered because once the individual weaklearn are trained, without any further training, their output can be fused to produce ensemble decision.

Table 1. Fusion Comparison using SVM-LMBPNN-LDA

Fusion Names	MDEFS	mRMR	GA
Voting	**98.11±0.68**	97.78±0.16	92.89±0.42
Max	98.11±0.16	**98.44±0.31**	93.33±0.27
Avg	98±0.54	98±0.27	**94.33±0.72**
Min	97.56±0.68	96.89±0.16	94.44±0.83
Product	97.88±0.68	97.44±0.31	94.67±0.54

3.2. The Effect of FBEF Using 3 Fusion Classifiers

Table 2 shows our method in comparison to other ensemble learning with different learning model (such as ensemble LMBPNN, ensemble SVM, etc.) can give us some insight, when dataset is scaled and using voting fusion rule, SVM can gain more better result but KNN will have slightly bad result. Good combination fusion 3 classifier is LMBPNN-SVM-LDA with scaling dataset. In here, we only used voting, max, avg, min, product. Another fusion rule should be considered and tested.

Table 2. FBEF and other ensemble comparison with scale on

Dataset	FBEF	LMBPNN	SVM
MDEFS	**98.71±0.08**	98.11±0.16	95.67±0.98
mRMR	**97.78±0.68**	97.56±0.16	95.44±0.16
GA	**95.22±0.42**	93.76±0.27	82.33±0.54
Dataset	**KNN**	**DT**	**LDA**
MDEFS	77.11±0.16	96.44±0.54	97.33±0.54
mRMR	96.67±0.54	96.56±0.31	97.11±0.16
GA	84.43±0.27	95.11±0.57	92.89±0.42

3.3. Comparison with Other Epileptic Seizure Classification

From Table 3, it is observed that our proposed approach achieved 98.11±0.68. Using standard deviation, our method is comparable with the Alzami method. By using 5 features, FBEF with 3 sub-classifiers for bagging can improve the accuracy rather than alzami method by using single classifier. The processing time of FBEF is faster than alzami method because alzami is using pattern recognition neural network and we using levenberg-Marquardt BackPropagation Feed Forward Neural network

Table 3. comparison with other seizure classification using Bonn Dataset where m' is number of features

Methods	m'	Classifier	Accuracy
Samanwoy [15]	9	BPNN	96.7
Guler [16]	24	SVM	75.6
Hsu [17]	13	SVM	87.6
Alzami [6]	5	Bagging+MDEFS	98.67
Our proposed	5	SVM-BPNN-LDA	**98.71**

3.4. Comparison with other Method for Real Dataset

From Table 4, it is observed that our proposed approach, is can also be used into another dataset. For the remainder, some real-world dataset cannot use QDA but Linear Discriminant Analysis, this can be happened because we didn't put any feature selection or dimension reduction into real world dataset. Thus, the features characteristic is not changed in real datasets. RSCE and SAEL is focused in random subspace. RSCE have tendency ignore the importance degree of different subspaces and classifier. RSCE also generate different random subspace set which contains set of random subspaces that will produce different combinations of classifier which will affect the final prediction result. SAEL is improved RSCE by using adaptively adjust the weight of base classifier and using weighted voting to combine predicted labels from different base classifier and obtain final result. The reason FBEF is outperformed RSCE and SAEL is 1) FBEF treated the base classifier equally, 2) FBEF do not need tuning anything and weighting but SAEL need tuning the parameters that prespecified by the user to get better accuracy and need iterations to obtain precise weight.

Table 4. FBEF and other approaches using real world dataset in matter of accuracy

Dataset	FBEF	SAEL[9]	RSCE [8]
Australia	**86.76**	84.1	85.7
Bands	66.85	68.5	**73.6**
Bupa	**71.5**	65.1	67.2
contraceptive	**54.63**	49.5	50.5
Dermatology	96.84	95.5	**97**
Haberman	**73.59**	70.6	73.1
Hayes-roth	73.54	**76.4**	65.7
Heart	82.35	76.4	**82.5**
Housevotes	**96.82**	96	94.4
Led7digits	**73.07**	59.8	54.5
Saheart	**71.21**	65.4	68.3
Segment	95.41	**97.3**	**97.3**
Wine	**97.94**	95.1	97.4
wisconsin	93.99	96.1	**97.1**

4. Conclusion and future works

This paper has proposed fusion of bagging-based ensemble framework for epileptic seizure classification. The main contribution is using multi weak learner in oracle to improve the accuracy, robustness and suitable for epileptic seizure classification. The Experiment show that epileptic seizure classification based on FBEF with weak learner LMBPNN, SVM and LDA using voting and MDEFS as feature selection get 98.11%±0.68. FBEF also suitable to be used in real world dataset because it outperforms SAEL and RSCE. From the performance on epileptic seizure dataset and real-world dataset, it can be concluded that FBEF is better than using traditional Ensemble Bagging Framework. Moreover, the combination of weak learner (classifier) take major effect in accuracy, such as worst weaklearner have tendency to lower the accuracy and cannot put the weaklearner that sensitive into scaling data together such as SVM with KNN. Furthermore, LMBPNN-SVM-LDA is the best fusion weaklearner using scale data.

Beside voting, average can be considered as fusion rule when using FBEF.

Future work would inclusion of more fusion rule like Behavior Knowledge Space, Dempster-Shafer, etc. or combination of bagging and adaboost inside the weakleaner α, β, γ.

Reference

[1]. R. S. Fisher *et al.*, "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE).," *Epilepsia*, vol. 46, no. 4, pp. 470–2, Apr. 2005.

[2]. A. Shoeb and J. Guttag, "Application of Machine Learning to Epileptic Seizure Detection," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010, pp. 975–982.

[3]. F. Alzami *et al.*, "Adaptive Hybrid Feature Selection-Based Classifier Ensemble for Epileptic Seizure Classification," *IEEE Access*, vol. 6, pp. 29132–29145, 2018.

[4]. M. Li, W. Chen, and T. Zhang, "Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble," *Biomed. Signal Process. Control*, vol. 31, pp. 357–365, Jan. 2017.

[5]. N. S. Tawfik, S. M. Youssef, and M. Kholief, "A hybrid automated detection of epileptic seizures in EEG records," *Comput. Electr. Eng.*, vol. 53, pp. 177–190, Jul. 2016.

[6]. F. Alzami, D. Wang, Z. Yu, J. You, H. S. Wong, and G. Han, "Robust epileptic seizure classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9772, 2016, pp. 363–373.

[7]. T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. Abd El-Samie, "EEG seizure detection and prediction algorithms: a survey," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, p. 183, Dec. 2014.

[8]. Tin Kam Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, 1998.

[9]. Zhiwen Yu, Le Li, Jiming Liu, and Guoqiang Han, "Hybrid Adaptive Classifier Ensemble," *IEEE Trans. Cybern.*, vol. 45, no. 2, pp. 177–190, Feb. 2015.

[10]. Hanchuan Peng, Fuhui Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.

[11]. L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Syst. Man. Cybern.*, vol. 22, no. 3, pp. 418–435, 1992.

[12]. L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms: Second Edition*, 2nd Editio., vol. 9781118315. New Jersey: John Wiley & Sons, 2014.

[13]. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.

[14]. R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E*, vol. 64, no. 6, p. 061907, Nov. 2001.

[15]. S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-Band Wavelet-Chaos-Neural Network Methodology for Epilepsy and Epileptic Seizure Detection," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1545–1551, Sep. 2007.

[16]. I. Guler and E. D. Ubeyli, "Multiclass Support Vector Machines for EEG-Signals Classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 2, pp. 117–126, Mar. 2007.

[17]. K.-C. Hsu and S.-N. Yu, "Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm," *Comput. Biol. Med.*, vol. 40, no. 10, pp. 823–830, Oct. 2010.