EEG Signal Classification for Alcoholic and Non-Alcoholic Person using Multilevel Wavelet Packet Entropy and Support Vector Machine

Cahyantari Ekaputri[1]
School of Electrical Engineering,
Telkom University
Bandung, Indonesia
[1]cahyantarie@telkomuniversity.ac.id

Rahmat Widadi[2]
Faculty of Telecommunication and
Electrical Engineering,
Institut Teknologi Telkom Purwokerto
Purwokerto, Indonesia
[2]rahmat@ittelkom-pwt.ac.id

Achmad Rizal[3]
School of Electrical Engineering,
Telkom University
Bandung, Indonesia
[3]achmadrizal@telkomuniversity.ac.id

***Abstract*— EEG signal provides information about brain conditions such as brain activity or consciousness level of a person. The consciousness level of a person can also be determined by alcohol. The use of alcohol for a long time can raise specific patterns in EEG signals. Several studies have shown a pattern of differences in EEG signals between alcoholic and non-alcoholic subjects. In this study, EEG signal for alcoholic and non-alcoholic was classified using Multilevel Wavelet Packet Entropy (MWPE) method in the feature extraction stage. MWPE was used to measure the signal complexity at different wavelet decomposition levels. These features are used as Support Vector Machine (SVM) input. The results of the test showed the highest accuracy of 77.8% with quadratic SVM. These results indicated that signal complexity could be used as a differentiator of EEG signals for alcoholic and non-alcoholic persons.**

***Keywords*— EEG, alcoholic, wavelet analysis, entropy**

I. INTRODUCTION

Heavy alcohol consumption can cause various diseases, such as liver cirrhosis, diabetes, and cancer [1]. It can also cause the loss of body control, in addition to lead to various criminal acts and driving accidents. A treatment is then deemed necessary to reduce the adverse effects of heavy alcohol consumption. The classification of alcoholic (addicted to alcohol) and non-alcoholic people in the community is vital before conducting the treatment. This classification process can use the signal, which is obtained from the human brain, called an EEG signal.

In previous studies, EEG signal classification on alcoholic and non-alcoholic subjects was carried out on Gamma signals [2]. These signals were extracted from the EEG signal, and several parameters were calculated as the characteristic of the EEG signal. Another study in [3] used Wavelet Packet Decomposition (WPD), Principal Component Analysis (PCA), and Back-Propagation Neural-Network (BP-NN) optimized by genetic algorithms for the classification of alcoholic EEG. Here, the highest accuracy of 94% was achieved using a decomposition level of 3.

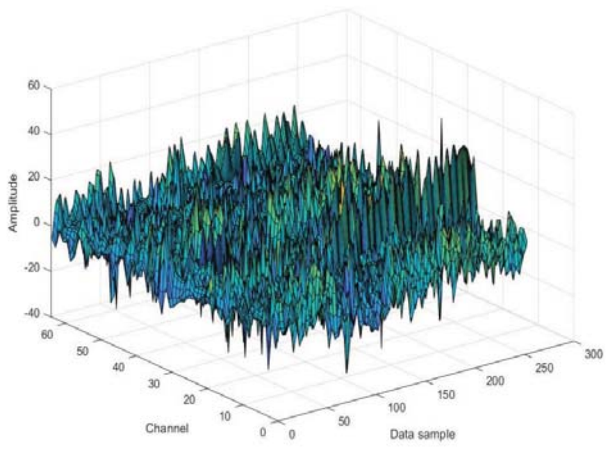
Previous research showed that wavelet analysis could produce a promising performance. One method of wavelet analysis is Multilevel Wavelet Packet Entropy (MWPE) - an entropy-based feature extraction method on wavelet sub-band [4]. With multilevel wavelet packet entropy, the information of the wavelet entropy signal's wavelet entropy can be obtained at the desired decomposition level range. A signal that is concentrated in one sub-band will produce a lower packet entropy wavelet (WPE) compared to the signal spread over all frequencies [5]. MWPE provides a quite good performance in lung sound analysis [6], EEG signal analysis, and speech analysis [7].

In this study, we use MWPE to obtain alcoholic and non-alcoholic EEG signal features. The EEG signal was decomposed using WPD, and then the entropy was calculated as a feature. Support Vector Machine (SVM) was used to classify the signal. MWPE was considered suitable as a feature extraction method considering that WPD decomposed the EEG signal into 2N sub-band with the equal bandwidth. Thus, the distribution of signal energy in the frequency band could be directly identified. This proposed method can be an alternative method of the analysis with low computation and a small number of features.

II. MATERIALS AND METHOD

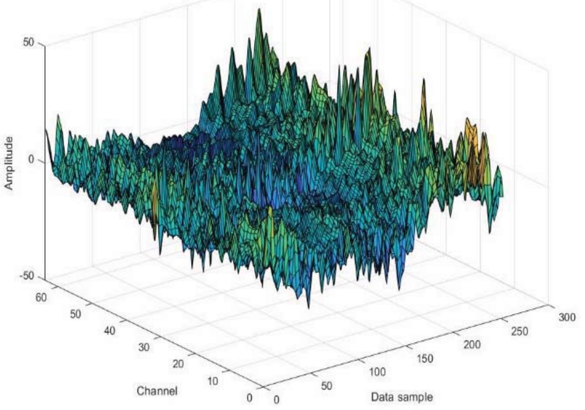
A. EEG dataset

EEG signal data were obtained from the UCI Machine Learning Repository with detailed process descriptions reported in [8]. The frequency sampling for the signal was set at 256 Hz for 1 second on 64 channels. EEG signals were taken when the subject was given a single visual stimulus S1) or two stimuli (S1 and S2). There were two conditions, S1 matched S2, and S1 differed from S2. There were 120 subjects, each of which comprised 120 trials. In this study, only 600 data were taken for each alcoholic and normal control EEG signal. The examples of alcoholic and non-alcoholic EEG signals show in Fig 1.



(a)

978-1-7281-6142-6/2020/$31.00 ©2020



(b)

Fig. 1. EEG Signal for (a) Alcoholic Subject and (b) Normal Subject

B. Multilevel wavelet packet entropy

Entropy is a measure of signal complexity [9]. One method of measuring entropy is Wavelet Entropy (WE), which uses sub-band of DWT [10]. Another variation of WE is Wavelet Packet Entropy (WPE), utilising the sub-band resulted from WPD [5]. WPE is frequently often used for biology signal analysis, such as heart sounds and lung sounds [4][5]. The calculation process of WPE is described below.

Wavelet packet decomposition (WPD) on signal $S(t)$ can be described as follows:

$$d_{j,n}(k) = 2^{\frac{j}{2}} \int_{-\infty}^{+\infty} S(t)\psi_n(2^{-j}t - k)dt,$$
$$0 \leq n \leq 2^N - 1 \quad (1)$$

where $S(t)$ refers to the original signal. j , n and k refer to scale, band, and surge parameter respectively. From Eq. 1, computation of the energy from each sub-band described as in Eq. 2.

$$E_{j,n} = \sum_k \left| d_{j,n}(k) \right|^2 \quad (2)$$

where *j, n,* and *k* refer to the scale, band, and surge parameter, respectively. The total energy of WPD sub-band described as in Eq. 3.

$$E_{tot} = \sum_n E_{j,n} \quad (3)$$

The energy for each sub-band in scale j relatively described as in Eq. 4.

$$p_{j,n} = \frac{E_{j,n}}{E_{tot}} \quad (4)$$

Wavelet packet entropy (WPE) described as in Eq. 5.

$$WPE_N = -\sum p_{j,n} \ln p_{j,n} \quad (5)$$

The notation N on WPE_N refers to the level of decomposition used in WPD.

In previous studies, single WPE value was frequently used as a feature for signal analysis. In this paper, N (multiple) WPE was used to improve the extraction of EEG signal features as in [4]. The characteristics used in this study are as shown in equation (6) with N = 7; thus, it became possible to have multilevel wavelet packet entropy as expressed in Eq. (6).

$$MWPE = [WPE_1, WPE_2, \dots, WPE_N] \quad (6)$$

Decomposition level N = 7 would produce 128 sub-bands with a width of 1.95Hz at level 7 decomposition. Here, we used Haar, Db2, Db8, Bior1.5, and Bior2.8 as mother wavelets as in [11].

C. Support Vector Machine

Support Vector Machine (SVM) is one method of machine learning that can solve the problem of the classification of two groups. Conceptually, the learning process is done by mapping non-linear input vectors to a feature space with a very high dimension. Then, in this feature space, the surface is built by making several linear decisions. The special nature of the surface is related to its ability to ascertain the high generalization ability of the machine learning process. Thus, it is possible to implement it for limited cases where training data (training data) can be separated without any misclassification. A high generalization ability is shown in the process of transforming input polynomial [12].

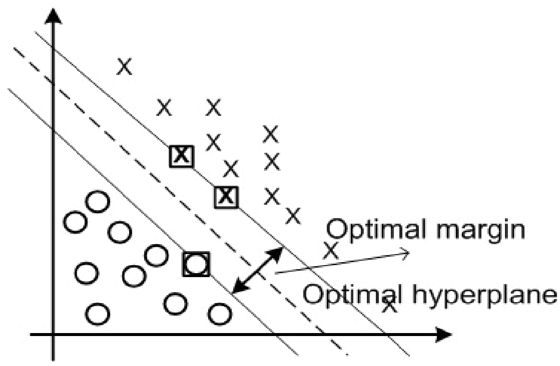


Fig. 2. Example of SVM classification in 2D space

Figure 2 shows an example of the process of classifying/separating two classes into a single 2-dimensional space. Support Vector, marked by grey box, calculates distance/margin between two classes. The basic idea of SVM resembles a neural network that finds an optimal Hyperplane to separate the classes linearly. Then, SVM will expand the linearly inseparable classes by transforming data to be mapped into new spaces using the Kernel function.

Training data A is described as in Eq. 7.

$$A = \{(x_i, y_i), i = 1, \dots, n | x_i \in R^d, y_i \in \{+1, -1\}\} \quad (7)$$

where x_i refers to the input with dimensions d and y_i , and A refers to the output in the form of 2 classes, which are -1 and 1. To separate classes linearly, a line is defined that separates classes, often called a hyperplane. The hyperplane is described as in Eq. 8.

$$w.x_i + b = 0 \quad (8)$$

where w refers to the weight variable matrix, x refers to the input vector, and b refers to the bias.

An optimal hyperplane will be found that meets in Eq. 9

$$\begin{cases} \text{if} \quad w \cdot x_i + b \geq 1, \quad \text{then} \quad y_i = 1 \\ \text{if} \quad w \cdot x_i + b \leq -1, \quad \text{then} \quad y_i = -1 \end{cases} \quad (9)$$

SVM will maximize a distance/margin between two classes, i.e. $2/\|w\|$. Therefore, to find the optimal hyperplane can be calculated by resolving the optimization problem limited as in Eq. 10 and Eq. 11.

$$\text{minimize} \quad \|w\|^2/2 \tag{10}$$

$$\text{subject to} \quad y_i(w \cdot x_i + b) \geq 1 \quad i = 1,2,\ldots,n \tag{11}$$

This limited optimization problem can be solved by using Quadratic Programming (QP). However, to separate the linearly inseparable classes, the kernel function is used to transform data into several new spaces by adding new variables/dimensions.

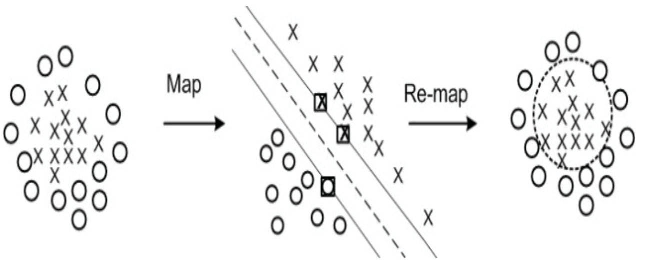


Fig. 3. SVM for the nonlinear problem

The use of kernel functions in SVM is often categorized as nonlinear SVM. One of the kernel functions included in nonlinear SVM is the polynomial kernel function described as in Eq. 12.

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^p \tag{12}$$

where p refers to the parameters that can be changed. Some polynomial kernel functions used in this study besides linear SVM included quadratic SVM and Cubic SVM.

Considering SVM as a supervised learning, N-fold cross-validation (NFCV) was used to divide training and testing data. With NFCV, the signal data were divided into N data sets. An N-1 data set was used as training data and one dataset as test data. The process was repeated up to N times until each data set has been used as a test data. In this study, N = 5 and 10 were used to enable each dataset to contain 120 and 60 data for each data class.

III. Result and Discussion

Figure 4 displays the results of the features extraction result of EEG signal in alcoholic and non-alcoholic subjects using MWPE with Db8 as mother wavelet. On average, WPE EEG signals in alcoholic subjects showed higher values than that of non-alcoholic subjects. The deviation values of WPE at each decomposition level overlapped between alcoholic and non-alcoholic. Such condition was the potential to cause some errors in the classification process.

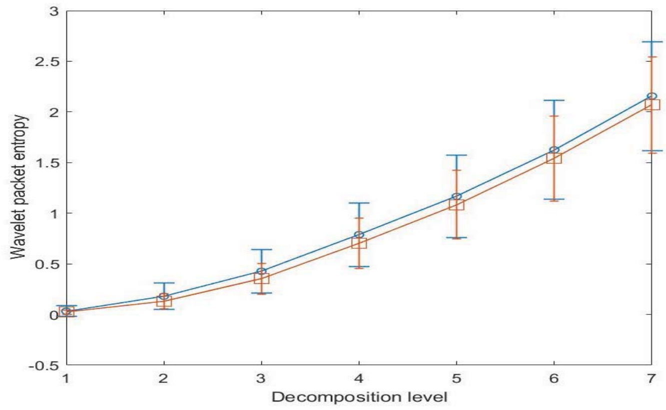


Fig. 4. WPE for level decomposition N = 1-7 using Db8 as mother wavelet

Figure 5 displays the classification accuracy using Db8 with several levels of decomposition in 5-fold cross-validation. The highest accuracy yield of 77.8% was obtained using quadratic SVM at level 6. Reducing the decomposition level would decrease the accuracy. Here, a higher level of decomposition did not guarantee higher accuracy because of the possibility of an inappropriate sub-band division.

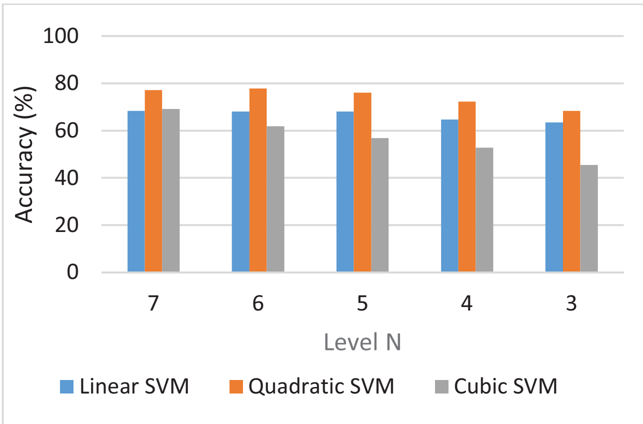


Fig. 5. Accuracy (%) MWPE using Db8 as mother wavelet and 5-fold cross-validation

Figure 6 displays the accuracy using DB8 and 10-fold cross-validation. The highest accuracy achieved was 77.8% using quadratic SVM. This result had no differences from the one using a 5-fold CV. This means that the accuracy achieved is not dependent upon the validation process.

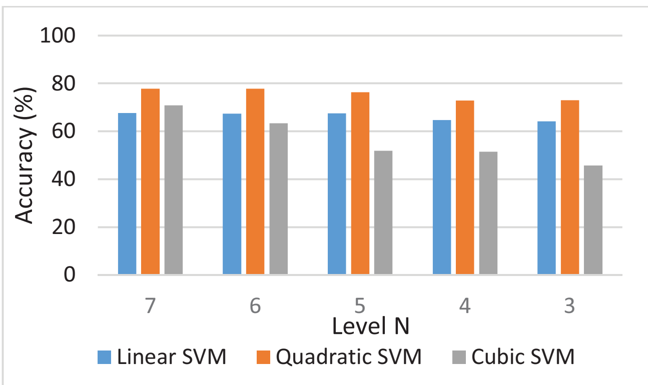


Fig. 6. Accuracy (%) MWPE using Db8 as mother wavelet and 10-fold cross-validation

In this study, MWPE was carried out to level 7 that was considered adequate as, with the decomposition of level 7, it would produce 128 sub-band. With a sampling frequency of 256 Hz, a sub-band of 1 Hz width could be produced. Decomposition with a higher level will provide a narrower sub-band. As shown in Figure 2 and Figure 3, reducing the level of decomposition does not result in any higher accuracy. The weakness of this study was that the EEG signal from 64 channels was put together, thus making it to be a signal along 64 x 256 samples. With the proposed method, seven characteristics could be obtained. The results obtained would be different if MWPE was calculated on each EEG signal channel. As a consequence, the number of features to be produced was much higher, namely 7x64 = 448 features.

Compared to previous studies using the same dataset, the accuracy obtained was lower [2-3]. The proposed method, however, had some advantages such as not requiring any filtering and its flexibility to choose the width of the sub-band. MWPE has the potential to be applied to other biological signals. Choosing the right level of decomposition and mother wavelet will improve the accuracy of the classification performed. The use of different classifiers will also affect accuracy.

IV. Conclusion

This research proposed the MWPE method to classify the alcoholic EEG signals. The highest accuracy was produced by Db8 with quadratic SVM as classifier. Even though the best accuracy was not as good as previous studies, the MWPE method had an opportunity to be improved in various ways. Some techniques that can be used to improve accuracy include the calculation of MWPE on each EEG signal channel, the selection of the right channel that can be used to classify the EEG signals, the selection of mother wavelets and the appropriate decomposition level. A combination with other classifiers will be an interesting research topic in the future.

References

[1] J. Rehm, “The risks associated with alcohol use and alcoholism,” *Alcohol Res. Health*, vol. 34, no. 2, pp. 135–143, 2011.

[2] R. Widadi, I. Soesanti, and O. Wahyunggoro, “EEG Classification Using Elliptic Filter and Multilayer Perceptron Based on Gamma Activity Features,” in *Proceedings - 2018 4th International Conference on Science and Technology, ICST 2018*, 2018, vol. 1, pp. 1–5.

[3] M. Saddam, H. Tjandrasa and D. A. Navastara, "Classification of alcoholic EEG using wavelet packet decomposition, principal component analysis, and combination of genetic algorithm and neural network," *2017 11th International Conference on Information & Communication Technology and System (ICTS)*, Surabaya, 2017, pp. 19-24.

[4] A. Rizal, R. Hidayat, and H. A. Nugroho, “Multilevel wavelet packet entropy: A new strategy for lung sound feature extraction based on wavelet entropy,” in *2017 International Conference on Robotics, Automation and Sciences (ICORAS)*, 2017, pp. 1–5.

[5] F. Safara, S. Doraisamy, A. Azman, A. Jantan, and S. Ranga, “Wavelet packet entropy for heart murmurs classification,” *Adv. Bioinformatics*, vol. 2012, 2012.

[6] A. Rizal, R. Hidayat, and H. A. Nugroho, “Multilevel wavelet packet entropy: A new strategy for lung sound feature extraction based on wavelet entropy,” in *Proceeding of 2017 International Conference on Robotics, Automation and Sciences, ICORAS 2017*, 2018, vol. 2018-March, pp. 1–5.

[7] R. Hidayat, A. Rizal, A. Bejo, and S. Sumaryono, “Vowel Sound Analysis in the Indonesian Language using Multilevel Wavelet Entropy and Multilevel Wavelet Packet Entropy,” in *International Conference on Science in Information Technology (ICSITech)*, 2019.

[8] X. L. Zhang, H. Begleiter, B. Porjesz, W. Wang, and A. Litke, “Event related potentials during object recognition tasks,” *Brain Res. Bull.*, vol. 38, no. 6, pp. 531–538, 1995.

[9] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K. H. Ng, and J. S. Suri, “Automated diagnosis of epileptic EEG using entropies,” *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 401–408, 2012.

[10] O. A. Rosso *et al.*, “Wavelet entropy : a new tool for analysis of short duration brain electrical signals,” *J. Neurosci. Methods*, vol. 105, pp. 65–75, 2001.

[11] I. Wijayanto, A. Rizal, and S. Hadiyoso, “Multilevel Wavelet Packet Entropy and Support Vector Machine for Epileptic EEG Classification,” in *2018 4th International Conference on Science and Technology (ICST)*, 2018, pp. 9–14.

[12] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.