ISSN: 04532198
Volume 62, Issue 03, April, 2020

# Epileptic EEG Signal Classification using Wavelet Time Entropy

Sugondo Hadiyoso[1], Achmad Rizal[2], Suci Aulia[3], Mohamad Ramdhani[4]

[1,3] Applied Science School,[2,4] School of Electrical Engineering, Telkom University

**Abstract-** Electroencephalogram (EEG) signal interpretation has been developed for various purposes such as brain health examination, brain detection, brain trauma, emotional condition, and even predict the response that will occur. The complex form of EEG signals will complicate one's interpretation visually so that it requires neurologists to deduce it. One of the brain disorders that are of concern and can be detected through EEG is epilepsy. EEG signal patterns can be identified through excessive brain cell activity before or after a person experiences seizures without cause. In this study, we proposed an EEG epilepsy signal recognition using Wavelet Time Entropy (WTE) as the main modality to obtain signal features. 300 EEG signal consisting of 3 classes (normal, interictal, seizure) has been tested with the highest accuracy result of 86.3% generated by Db 2 with decomposition level 2 or 3 using cubic Support Vector Machine (SVM).

**Keywords-** EEG, epilepsy, Wavelet Time Entropy, SVM

## 1. Introduction

One of the most common disorders of the human brain nervous system is epilepsy. Epilepsy occurs due to the activity of a group of excess neuron cells which causes various reactions in the human body. The response can be in the form of a momentary daze, tingling, disturbance of consciousness, convulsions and or muscle contractions. A new understanding of epilepsy was put forward by neurologist Hughlings Jackson in 1873, stating that seizures are the result of a brief and sudden electrochemical release in the brain. In the 1930s, a psychiatrist named Hans Berger experimented using EEG to observe epilepsy. Until now, EEG is the main modality for observing brain wave discharge associated with different types of seizures. EEG can also be used to find the location of seizure sources that can help neurosurgery in epilepsy patients.

One of the tools used to analyze EEG signals is wavelet entropy, in previous research [1], Osvaldo and his colleagues proved that the advantages of entropy wavelet could analyze EEG signals even to changes in light conditions based on their level of contrast and interference with EEG signals. Then in 2002, Osvaldo continued his research [2] to evaluate EEG signals in unimodal and bimodal ways from ERP theta components (4-8 Hz) in other bands.

Based on the description above, in this research, we proposed the classification of EEG epileptic using entropy wavelet and packet entropy wavelet as a comparison. In this study, the classification of EEG epileptic was carried out as a continuation of previous studies [3], epileptic EEG classification using sample entropy on multidistance signal level difference (MSLD) and verification using SVM method showed the highest accuracy of 97.7 %. Besides, the relationship between measurement results of wavelet time entropy and wavelet packet entropy has also been carried out by Saminu [4] with test data using SVM for classification of EEG signals with satisfactory results reaching 98%.

Wavelet packet entropy (WPE) was chosen as a comparison of WE which was reviewed in the journal [5]. WPE is recommended as one of the EEG signal extraction features and signal disturbance power quality classification [6]. In other research [7] WPE is used by decomposition to eliminate the interference displayed in the signal. The difference between WE and WPE is its decomposition. Where the WPE can present a signal decomposition that cannot be presented only with WE alone [8], so the testing of the WPE results will be much better than WE.

## 2. Material and Methods

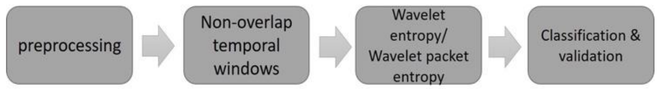



Fig.1: Block diagram of proposed system

Figure 1 depict the detailed process of the proposed system. In preprocessing, the Raw EEG signal that observed will be normalized in its amplitude in the range -1 to +1. Windowing without overlap is applied to get a small subset of the signal so as to simplify the analysis. WE and WPE are used to get signal features that become training datasets for classification and validation.

### 2.1 EEG Data

In this study, we used EEG data in the database that provided open access at the University of Bonn [9]. EEG data consists of three classes, namely seizure, interictal and normal. Data was recorded using a sampling frequency of 173.61 Hz and filtered using a 40 Hz LPF. Each data has a length of 4098 samples. Each data class consists of 100 data. Sample data for each class can be seen in Figure 2.

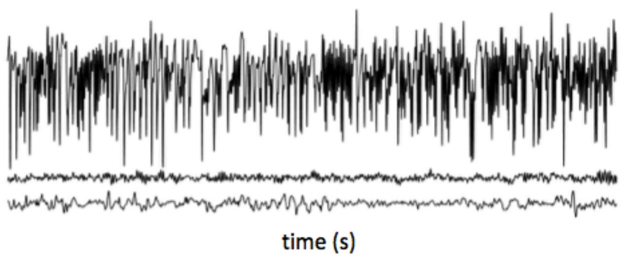



Fig. 2: EEG signal. Top; Seizure EEG, Middle; Normal EEG, Below; interictal EEG

### 2.2 Wavelet Entropy and Wavelet Packet entropy

Wavelet entropy is an entropy calculation in a wavelet subband resulting from discrete wavelet transform. Discrete wavelet transform (DWT) for any signal S (t) can be expressed by (1):

$$(W_\psi S)(j,k) = \int_{-\infty}^{+\infty} S(t)\psi_{j,k}(t)dt \quad (1)$$

With $\psi_{j,k}(t)$ is a discrete mother wavelet function, j and k are the scale and translation parameter respectively, with $j \neq 0$.

If given the wavelet coefficient as follows $C_j(k) = \langle S, \psi_{j,k} \rangle$ which is the result of DWT, then the energy of the signal is on a scale $j=1,2,\ldots N$ formulated as follows:

$$E_j = \sum_k |C_j(k)|^2 \quad (2)$$

The total energy signal results from DWT can be expressedby

$$E_{tot} = \|S\|^2 = \sum_j \sum_k |C_j(k)|^2 = \sum_j E_j \quad (3)$$

Relative wavelet energy for scale j is considered as

$$P_j = \frac{E_j}{E_{tot}} \quad (4)$$

Then wavelet entropy (WE) can be define as (2)

$$WE = -\sum p_i \ln p_i \quad (5)$$

If *WE* is obtained from the DWT process, then the Wavelet packet entropy (WPE) is an entropy that is calculated from the subband resulting from wavelet packet decomposition (WPD). WPD on the S (t) signal can be defined as follows:

$$d_{j,n}(k) = 2^{j/2}\int_{-\infty}^{+\infty} S(t)\psi_n(2^{-j}t-k)dt, \qquad 0 \le n \le 2^N-1 \quad (6)$$

With S (t) is the original signal, j is the scale, n and k are band and surge parameter respectively. From equation (6) energy can be calculated for each subband as follows:

$$E_{j,n} = \sum_k |d_{j,n}(k)|^2 \quad (7)$$

Where j, n, k represent the scale, band, and surge parameter, respectively. Total energy from WPD results is:

$$E_{tot} = \sum_n E_{j,n} \quad (8)$$

In the same way as equation (4), relative energy for each subband in scale j can be expressed as:

$$p_{j,n} = \frac{E_{j,n}}{E_{tot}} \quad (9)$$

Thus WE from the WPD process can be called wavelet packet entropy (WPE) which is expressed as follows:

$$WPE_N = -\sum p_{j,n} \ln p_{j,n} \quad (10)$$

In this research, WE and WPE were calculated on non-overlap windows along 256 samples. The EEG signal is expected to be fairly stable in that range so that it represents the overall condition of the signal.

## 3. Classification and Validation

### 3.1 Support vector machine (SVM)

Support vector machine (SVM) is a supervised learning method that was first introduced in 1992 by Cortes and Vapnik [11]. SVM is a pattern recognition method for mapping functions from a set of labeled training data. SVM has a good performance in pattern recognition so it is widely used in data classification which has many attractive features. As one of the learning machine methods, SVM works based on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane to make it easier to separate two classes. SVM was developed by combining computational theories that have existed earlier decades, such as hyperplane margins introduced in 1965 and 1973 and the kernel method was introduced by Aronszajn in 1950. SVMs have good performance in numerous real applications [12-13], such as bioinformatics [14-16], text mining, face recognition, and object detections [17].

Essentially SVM works by finding the best hyperplane where this condition does not exist on the neural network. Basically SVM works on a linear classifier, and is further developed to work on nonlinear problems. The best separator hyperplane between the two classes is obtained by measuring the hyperplane margins and finding the maximum point. In other words hyperplane will divide the vector space into two different parts for each class. By calculating the margins will be obtained the distance between the hyperplane and the closest pattern in each class [18]. The closest positions of this pattern are called support vectors [19]. The greatest margin can be found by maximizing the distance value between the hyperplane and its nearest point.

Practically, training datasets can be defined as [20]:

$X_i \in Rp$ (11)

With,

$Y_i \in \{-1,+1\}, \quad i=1,...,n$ (12)

The function of hyperplane is defined by the equation:

$f(x) = hx + b = 0, \quad h = (h_1, h_2, h_3, ..., h_p)$ (13)

h is the weight of the hyperplane vector and b is defines the bias

In order to maximize margin then used SVM optimization equation for case of linear classification in primal space, as equation bellow:

$$min\frac{1}{2}\|h\|^2 \quad (14)$$

With,

$y_i\,(b + hx_i) \geq 1,$ i= (15)

1,...,n

In the implementation, many cases two classes on the input space cannot be separated completely. In this non-linear case it can be separated by the soft margin concept introduced in SVM [2] [7]. With the concept of SVM soft margin optimization made the following equation:

$$(w) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\varepsilon_i \quad (16)$$

Where

$$\boldsymbol{\varepsilon_i \geq 0} \quad (17)$$

$$\boldsymbol{y_i\,(b + hx_i) \geq 1 - \varepsilon_i, i = 1, \dots, n} \quad (18)$$

### 3.2 N-fold Cross Validation (CV)

Cross validation (CV) is used to evaluate the model that implemented in this study. The CV will divide the test dataset into K section and K-1 other sections as training dataset, this is repeated as many times as K. Then evaluating the ability to generalize the model by averaging the performance of each iteration. In this work, we use 5-fold cross-validation so that every test dataset is consist of 5 dataset that will repeat 60 times so that each data set will be test data. This cross validation process can be seen in Figure 3.

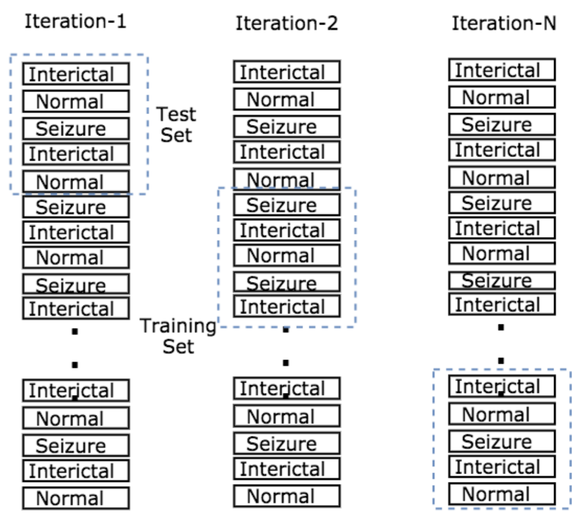
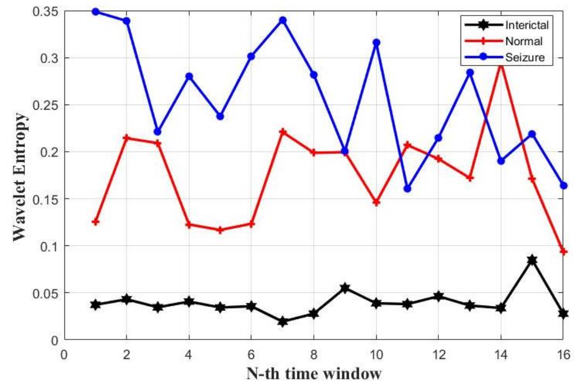


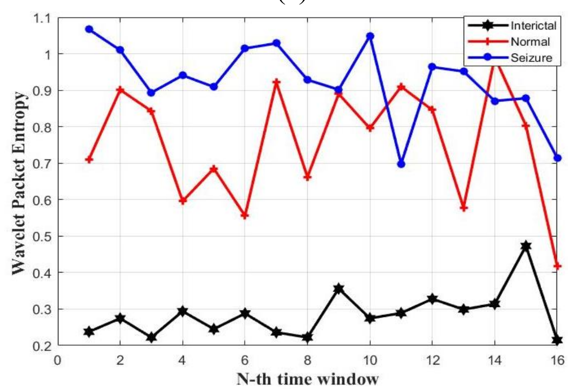

Fig. 3: 5-fold Cross-validation illustration in this research

## 4. Result and Discussion

Figure 4 shows the characteristics of each data class using WE and WPE with DB2 level 2 for feature extraction. While Figure 5 displays the same data as Bior2.8 level 7 for decomposition of EEG signals. In Figure 4 it can be seen that the WE and WPE values of the seizure signal tend to be higher than the normal signal. Meanwhile the interictal signal has a very low entropy value. in Figure 5 the pattern of the three data classes tends to be irregular so that no different patterns appear between data classes. Intuitively it can be said that DB2 is better at differentiating data between classes.

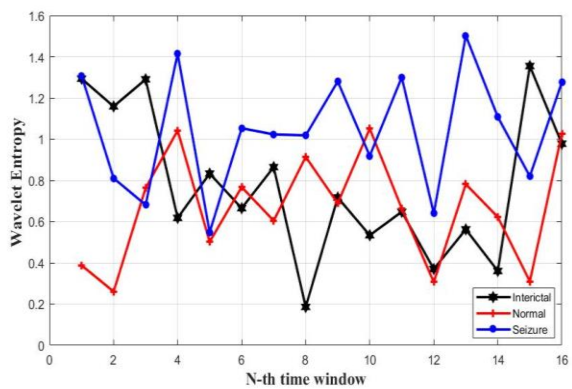



(a)

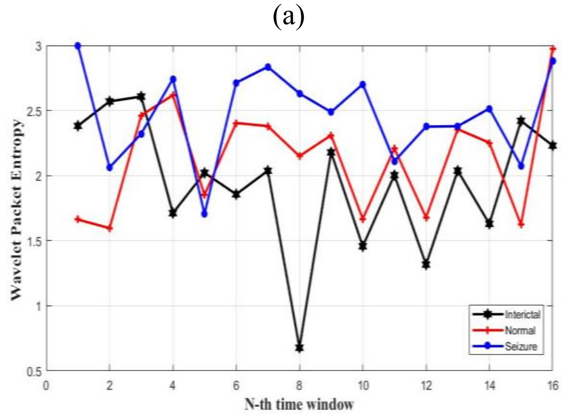



(b)

Fig. 4: (a) Wavelet entropy using Db2 level 2 (b) Wavelet packet entropy using Db2 level 2




(a)

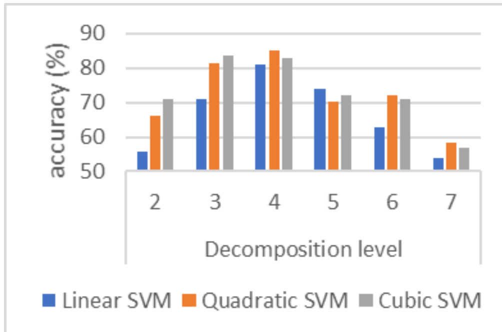



(b)

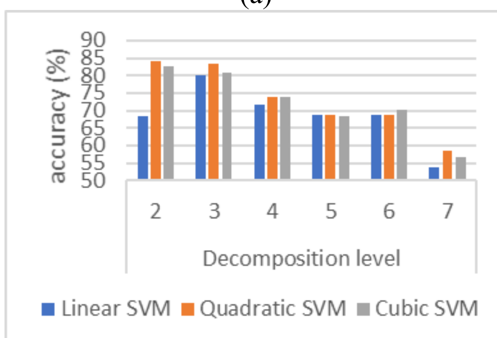
ISSN: 04532198
Volume 62, Issue 03, April, 2020

Fig.5: (a) Wavelet entropy using Bior2.8 level 7 (b) Wavelet packet entropy using Bior2.8 level 7

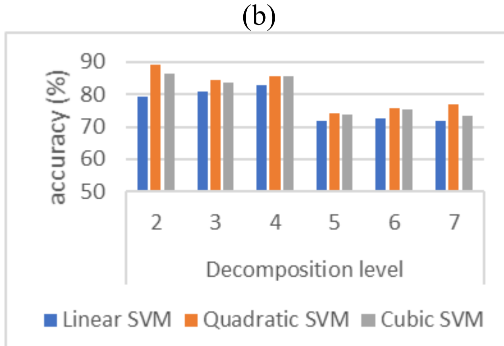
Figures 6, 7, and 8 show classification accuracy using DB2, DB8 and Bior2.8 as mother wavelets with decomposition levels 2 - 7. Entropy that used is WE, WPE and a combination of WE and WPE. The highest accuracy of 86.3% is achieved by WPE Db 2 with decomposition level 2 or 3 using cubic SVM or quadratic discriminant.




(a)

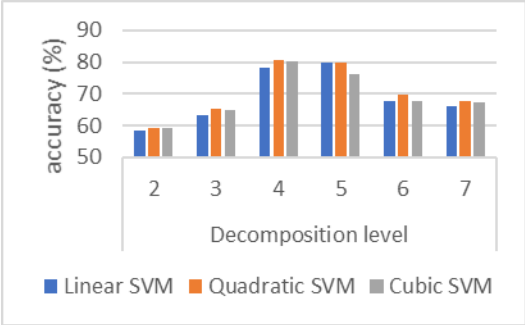



(b)

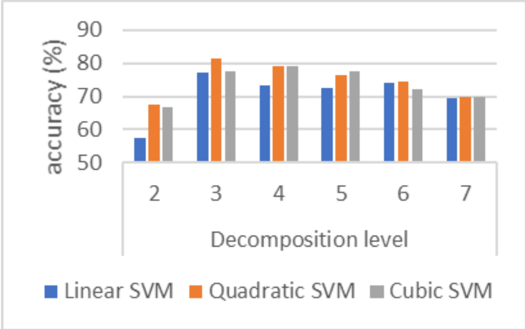



(c)

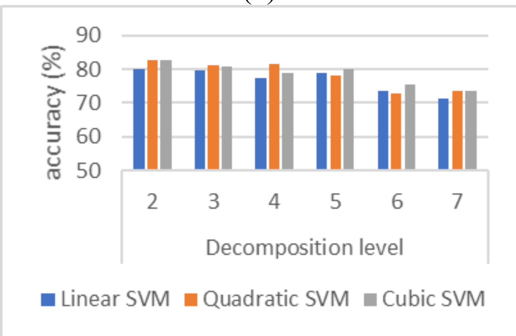
Fig. 6: Accuracy using DB2 (a) WTE (b) WPTE (c) Composite




(a)

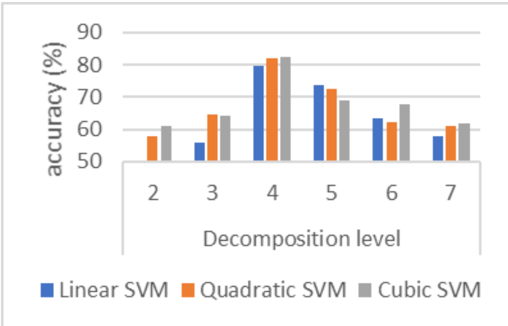



(b)

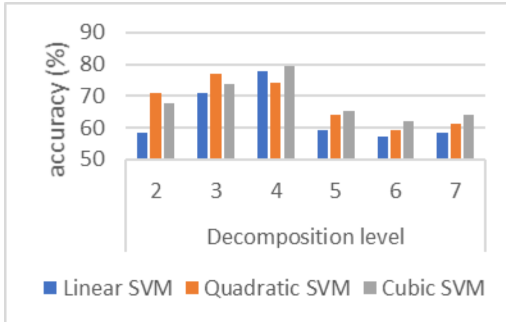



(c)

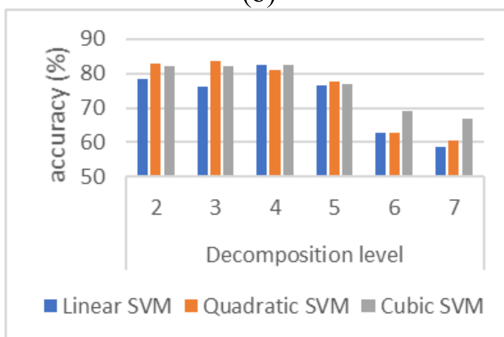
Fig.7: Accuracy using DB8 (a) WTE (b) WPTE (c) Composite




(a)




(b)




(c)

Fig.8: Accuracy using Bior2.8 (a) WTE (b) WPTE

(c) Composite

It is seen that accuracy is greatly influenced by the selection of mother wavelet, decomposition level, and entropy selection. In general, WPE produces better accuracy compared to WE. This is caused by the sub-band used has a uniform sub-band width so that the distribution of energy from the signal can be known more fully.

The proposed method has several weaknesses. Because it uses windowing, cutting data will greatly affect the characteristics generated. The use of windows requires the same length of data used. In addition, the selection of window length will greatly affect accuracy. Compared to previous studies using multilevel wavelet entropy (MWE) and multilevel wavelet packet entropy (MWPE) [12], the proposed method produces lower accuracy. However, the proposed method has advantages in terms of representing signal dynamics over time.

## 5. Conclusion

In this research, EEG signal classification has been successfully simulated for normal cases, seizure and interictal in normal subjects and epilepsy patients. 300 EEG datasets for all three conditions have been analyzed. The WE and WPE methods are used to obtain signal features which will then be classified. From the analysis carried out, we obtained the WE and WPE values in the seizure signal tend to be higher than the normal signal. Meanwhile the interictal signal has a very low entropy value. The highest accuracy achieved 86.3% is generated by WPE Db 2 with decomposition level 2 or 3 using cubic SVM. Selection of base mother wavelet, decomposition level, and entropy selection are criteria that affect accuracy. In this study, WPE produces better accuracy compared to WE. This is caused by the sub-band used has a uniform sub-band width so that the distribution of energy from the signal can be known more fully. The proposed method still leaves the potential to be explored. Some parameters such as window length, sample overlap length, signal initial determination, and determination of sub-band used for entropy calculations can still be explored to improve accuracy. The trial of the use of the proposed method for analysis of other biology signals can also be a research topic in the future.

**6. Reference**

[1] 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.

[2] J. Yordanova *et al.*, “Wavelet entropy analysis of event-related potentials indicates modality-independent theta dominance,” *J. Neurosci. Methods*, vol. 117, no. 1, pp. 99–109, 2002.

[3] A. Rizal and S. Hadiyoso, “Sample Entropy on Multidistance Signal Level Difference for Epileptic EEG Classification,” *Sci. World J.*, vol. 2018, pp. 1–6, Sep. 2018.

[4] S. Saminu and N. Özkurt, “Stationary Wavelet Transform and Entropy-Based Features for ECG Beat Classification,” *Int. J. Res. Stud. Sci. Eng. Technol.*, vol. 2, no. 7, pp. 23–32, 2015.

[5] T. Li and M. Zhou, “ECG Classification Using Wavelet Packet Entropy and Random Forests,” *Entropy*, vol. 18, no. 8, p. 285, Aug. 2016.

[6] M. Zhang, K. Li, and Y. Hu, “Classification of Power Quality Disturbances Using Wavelet Packet Energy Entropy and LS-SVM,” *Energy Power Eng.*, vol. 2, no. August, pp. 154–160, 2010.

[7] M. Varanis and R. Pederiva, “Wavelet Packet Energy-Entropy Feature Extraction and Principal Component Analysis for Signal Classification,” in *Proceeding Series of the Brazilian Society of Applied and Computational Mathematics*, 2015, no. 3.

[8] G. Ghodrati Amiri and A. Asadi, “Comparison of different methods of wavelet and wavelet packet transform in processing ground motion records,” *Int. J. Civ. Eng.*, vol. 7, no. 4, pp. 248–257, 2009.

[9] 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.

[10] A. Xu, Q. Ye, J. Qiang, S. Shan, and D. Zhang, “A wavelet time entropy algorithm for ultrasonic signal detection,” *Comput. Model. New Technol.*, vol. 18, no. 12, pp. 69–73, 2014.

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

[12] W. Astuti and Adiwijaya, “Support vector machine and principal component analysis for microarray data classification,” *J. Phys. Conf. Ser.*, vol. 971, no. 1, p. 012003, Mar. 2018.

[13] M. Turhan, D. Şengür, S. Karabatak, Y. Guo, and F. Smarandache, “Neutrosophic Weighted Support Vector Machines for the Determination of School Administrators Who Attended an Action Learning Course Based on Their Conflict-Handling Styles,” *Symmetry (Basel).*, vol. 10, no. 5, p. 176, May 2018.

[14] A. Statnikov, C. F. Aliferis, I. Tsamardinos, D. Hardin, and S. Levy, “A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis,” *Bioinformatics*, vol. 21, no. 5, pp. 631–643, Mar. 2005.

[15] B. Al Kindhi, T. A. Sardjono, and M. H. Purnomo, “Optimasi Support Vector Machine ( SVM ) untuk memprediksi adanya mutasi pada DNA Hepatitis C Virus ( HCV ),” *JNTETI7*, vol. 7, no. September, 2018.

[16] C. Orphanidou and I. Drobnjak, “Quality Assessment of Ambulatory ECG Using Wavelet Entropy of the HRV Signal,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 5, pp. 1216–1223, Sep. 2017.

[17] N. Dalal, B. Triggs, and D. Europe, “Histograms of Oriented Gradients for Human Detection,” 2005.

[18] S. Aulia, S. Hadiyoso, and D. N. Ramadan, “Analisis Perbandingan KNN dengan SVM untuk Klasifikasi Penyakit Diabetes Retinopati berdasarkan Citra Eksudat dan Mikroaneurisma,” *J. ELKOMIKA -Teknik Elektro Itenas - ISSN 2338-8323*, vol. 3, no. 1, pp. 75–90, 2015.

[19] 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.

[20] N. M. R. Kashif Mahmood Rajpoot, “Wavelets and support vector machines for texture classification,” in *8th International Multitopic Conference, 2004. Proceedings of INMIC 2004.*, 2004, pp. 328–333.