Bipolar Disorder Classification Based On Electrocardiogram Signal Using Support Vector Machine

Izzatunnisa Ainunhusna, Achmad Rizal, Sony Sumaryo

Abstract: Bipolar Disorder (BD) is one of kinds of mental disease that is quite common found in Indonesia. Those suffering from this disease will drastically experience a shift in mood in a certain period of time. This shift in mood, in turn, can cause many undesired things. The detection of Bipolar Disorder can be done through various diagnosing methods, one of which is by using EEG (Electroencephalogram) signal or ECG (electrocardiogram). One of the methods to detect BD using the ECG signal is by assessing the heart-rate variability (HRV) in which HRV in the patients of Bipolar Disorder tends to be lower than that of normal persons. In this research, an analysis method of HRV was developed to detect Bipolar Disorder using the ECG signal. The method proposed consisted of notch filter, wavelet decomposition, R-R detection, and HRV analysis using Mean Heart Rate (MHR), Standard Deviation of Normal to Normal (SDNN) and Root Mean Square of successive RR interval differences (RMSSD), and SVM for classification. From the experiment, it was found the highest accuracy of 93.8% using three features and quadratic SVM. This showed that the ability of method designed to differentiate HRV with the high accuracy. Here, the verification using the larger dataset was required to test the consistency of the proposed method.

Index Terms: bipolar disorder, classification, electrocardiogram, heart-rate variability, pattern recognition, support vector machine

1 INTRODUCTION

Bipolar Disorder (BD) is one of kinds of mental disease that is quite common found in Indonesia. The patients of this disease will drastically experience a shift in mood in a certain period of time [1]. In bipolar disorder there are two very significant changes of the patients’ condition, those are mania condition or the increasing phase and the depression condition or the decreasing phase [2]. Those two phases will bring a negative effect on both the daily life and the day-to-day activities of the patient. Hence, there is a need for an intensive healthcare and treatment either from the patients themselves or from the people surrounding. One way to detect BD is by analyzing the heart-rate variability (HRV) of the patient with BD [3]. A number of researches have used various signals to calculate the HRV to patients with BD. Quintana et al. calculated the HRV of the BD patients using the signal from pulse-oximeter [4]. In other research, HRV was analyzed from ECG signal [5]. Here, the ECG signal was not only analyzed in regard to its form to see any disorder in heart but also used in analyzing any diseases that were directly related to the heart [6]. From a previous various research, it is found a fact that the patients of BD had the lower HRV compared to the healthy people [7]. Any methods have been used to calculate HRV in the patients of BD either in the domain of time or in the domain of frequency [8]. Parameters in the domain of time include mean RR, mean HR, Standard Deviation of Normal to Normal (SDNN) and Root mean square of successive RR interval differences (RMSSD) [7]. Meanwhile, the parameters in the domain of frequency include high-frequency (HF) power band, low-frequency (LF) power band and LF/HF ratio [9]. In this research, the classification of BD and normal control based upon the ECG signal was conducted using a number of HRV parameters.

To obtain an accurate R-R interval, the filtering process was conducted in the ECG signal. It was continued with the calculation of some HRV parameters as the feature of both BD and normal control. SVM with some configurations was used as the classifier. In this research, some tests on any feature combination resulted from R-R interval were also carried out. It is expected that the results of this research can be a recommendation for a method in processing the ECG signal to detect Bipolar Disorder.

2 MATERIAL AND METHOD

The signal processing process in this paper is illustrated in Figure 1. Notch filter was used to remove the power-line noise 50 Hz. The wavelet decomposition was done to obtain the components of low frequency from the signal of notch filter output. The results of wavelet decomposition were used to remove the baseline wander in signal. In the following phase, the detection of R-R was carried out using certain algorithm. The results of R-R detection were used to calculate HRV and its parameter. The detail of each process is explained in the sub-parts as follows.

2.1 Dataset

The dataset used consisted of 18 signals of ECG from the patients with BD and 14 signals of normal ECG. The ECG signal took the recording duration for 10 seconds and the...
sampling frequency at 250 Hz. The number of R waves in each data was in the range of 13 to 16. In the signal there were some noises such as 50Hz power-line noise and baseline noise as shown in Figure 2. For this, in the following process, filtering was conducted to remove any noises and to guarantee that each R was detectable.

2.2 Filtering Process

Notch Filter
Notch filter was used to remove the component of 50 Hz on the EEG signal. Even though most of ECG signals were at the frequency of < 50 Hz, the existence of Notch filter could result in the smoother signals. This then could make the process of RR detection easier in the following phase. North filter used was IIR filter with the measurement in which the value of numerator was $b = [0.579, -0.357, 0.579]$ and denominator was $a = [1, -0.358, 0.158]$. The response of the frequency from the designed filter notch is illustrated in Figure 3.

Wavelet Transform
In this research, the wavelet transformation was used to obtain the approximation components of the signal. Afterwards, the approximation components from the signal would be used to remove the baseline wander from ECG signal. The wavelet transformation in this research used ‘Db6’ level 10. Furthermore, the approximation components at Level 10 (A10) were reconstructed to obtain the components of the signal from the low frequency. Subsequently, the ECG signals were reduced with A10 to remove the fluctuation in the baseline of ECG signal. The samples of ECG signal and A10 are shown in Figure 4.

2.3 RR detection and Heart-Rate variability

R-R detection was conducted by seeking the peak of the ECG waves. Initially, it was to detect R using the logics as shown in Figure 5.

In the HRV analysis, some variables in the domain of time were used such as Mean Heart Rate (MHR), Standard Deviation of Normal to Normal (SDNN) and Root Mean Square of successive RR interval differences (RMSSD) [10]. MHR, SDNN, and RMSSD were stated in Equation (3)–(4).

\[ THB = N \]
\[ MHR = \bar{f} = \frac{1}{N-1} \sum_{n=2}^{N} I_n \]
\[ SDNN = \sqrt{\frac{1}{N-2} \sum_{n=2}^{N} [I(n) - \bar{f}]^2} \]
\[ RMSSD = \sqrt{\frac{1}{N-3} \sum_{n=3}^{N} [I(n) - I(n-1)]^2} \]

where $THB = N$ is the total of heartbeat, and $MRR$ is the mean of RR interval [11]. The simplest way to look HRV is by plotting the results of time interval of each beat into the graph. If the results of the plotting were far from each other, then HRV was stated high; whereas if the results of plotting were closer to each other, then the HRV was stated low [10].

2.4 Classification
Support vector machine (SVM) is a linear classifier that has
been developed to cope with any classification of non-linear problems. The basic concept of SVM is to seek the best hyperplane to separate two data classes [12]. Hyperplane is a straight line or field separating the inter-classes data. The best hyperplane is obtained by maximizing a margin between two objects from the different classes. Margin refers to a distance between the hyperplane and the closest patterns in each data class. The closest position among the patterns in each class is named as support vector. To cope with the non-linear problem, it can use a method called as Kernel Trick. This method uses the non-linear field to separate the data among the classes that might be inseparable by the straight line. In this research, Quadratic SVM and Cubic SVM were used as the comparison of SVM linear. As SVM is a method requiring a supervised learning, then the data in this research were divided into four training data and testing data. As the number of data were limited, then the division of training data and testing data was by 50%-50% randomly.

3 RESULTS AND DISCUSSION

Figure 6 shows the ECG signals in the patients with BD. It was found that the noise frequency was high and the noise was also found in the base line of the signal. To remove the noise, the filtering process was used as explained previously. Figure 7 shows the signal of the filtering process output. Visually, it was found that the noise frequency reduced and the base line from the signal had started to be even. Due to this process, the ECG signal seemed to rather change from its original signal form. As the purpose of the signal processing was to guarantee the R wave, then the shift in the signal form no longer becomes the serious problem.

Figure 8 presents the results of R wave detection process. It can be seen that all R waves were detected rapidly. The accuracy in determining R wave can guarantee the accuracy in the HRV calculation. Visually, the difference in R-R interval was found in the patients with BD and those normal as shown in Figure 9. It can be seen that the values of RR interval in patients with BD tended to be closer; while in normal one they were scattered. This showed that HRV in BD patients tended to be lower compared to normal people [3] [13]. In previous studies it was stated that HRV alone was inadequate to determine BD or other abnormalities. For that, in HRV it was calculated several other parameters such as MHR, SDNN, and RMSSD calculated to distinguish BD and normal control.

Table 1 and Table 2 show the accuracy in the classification of BD and normal control using HRV parameter. As shown in Table 1, one HRV parameter was used as the feature and three types of SVM were used as the classifier. Using MHR or SDNN with Cubic SVM or Quadratic SVM as the classifier, the accuracy obtained was 62.5% at maximum. Table 2 shows the accuracy by means of a number of combination of HRV
parameters in which the highest accuracy of 93.8% was resulted from three HRV parameters using quadratic SVM as the classifier. This showed that with only one HRV parameter, it could differentiate BD and normal control. Overall, to the patients with BD, HRV tended to be lower compared to normal control. This was caused by the distortions of autonomic function in the patients with BD. The decrease of HRV was not dependent upon the medication, age or body mass index (BMI) [4]. More value of HRV could not be directly used as the feature for being relying on the diet or physical activities from the subject [4]. More HRV parameters might be needed to differentiate HRV to the patients with BD and normal control as in ECG-RR intervals (inter-beat intervals or RRIs), total power (TP) or the feature in the domain of frequency such as high-frequency (HF) power band, low-frequency (LF) power band and LF/HF ratio [9].

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Classification result (%) using individual feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classifier</td>
<td>MHR</td>
</tr>
<tr>
<td>Linear SVM</td>
<td>56.3</td>
</tr>
<tr>
<td>Cubic SVM</td>
<td>56.3</td>
</tr>
<tr>
<td>Quadratic SVM</td>
<td>62.5</td>
</tr>
</tbody>
</table>

TABLE 2 CLASSIFICATION RESULT (%) USING COMBINATION FEATURES

<table>
<thead>
<tr>
<th>Classifier</th>
<th>MHR + SDNN</th>
<th>MHR + RMS</th>
<th>SDNN + RMS</th>
<th>MHR + SDNN + RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear SVM</td>
<td>65.6</td>
<td>56.3</td>
<td>68.8</td>
<td>68.8</td>
</tr>
<tr>
<td>Cubic SVM</td>
<td>81.3</td>
<td>87.5</td>
<td>75</td>
<td>81.3</td>
</tr>
<tr>
<td>Quadratic SVM</td>
<td>78.1</td>
<td>87.5</td>
<td>90.6</td>
<td>93.8</td>
</tr>
</tbody>
</table>

The weakness of the proposed system was related to the very short time of ECG recording making the trend unable to be seen in a quite long time. In other research, the ECG recording took 5 minutes [14]. Even though HRV in the patients of BD is not dependent on the healthcare, age and BMI, the data should be taken from the condition of the same patient. In other research, HRGC in the patient of BD was found different for the mania condition and euthymic condition [7]. The research with the data in any states in the patients with BD becomes an interesting topic for the next research.

4 CONCLUSION

In this research, HRV-based BD classification was conducted using SVM. The HRV parameters used included MHR, SDNN, and RMSSD. The test resulted in the highest accuracy of 93.8% using three features and quadratic SVM as the classifier. In this research, it only used the feature in the domain of time – not using the one in the domain of frequency. The method in this research needs to be tested for the larger dataset with the longer duration of recording to test the consistency of the performance achieved.

5 REFERENCES


