

Support Vector Method with Radial Basis Function and Multi-Segment of Electrocardiogram for Paroxysmal Atrial Fibrillation Recognition

Nuryani Nuryani¹

¹Physics Department, Universitas Sebelas Maret, Surakarta, Indonesia
nuryani@mipa.uns.ac.id

ABSTRACT

Paroxysmal Atrial fibrillation (PAF) is a heart problem relating to irregular and rapid beating of the heart atria. It has risk of stroke and is independently associated with risk of mortality. Early information of PAF episode is important for a patient to have appropriate treatment to reduce atrial fibrillation complications. This article presents a new strategy to detect PAF with base of statistical electrocardiographic features and a support vector machine (SVM). R-peak series of electrocardiogram were segmented and were extracted to find the statistics of RR intervals. Different approaches in relation with the segmentation were investigated. Two-class SVM with radial basis function (RBF) and the statistics of RR intervals were examined for PAF detection. Using clinical data of patients with PAF, the proposed strategy showed excellent performance of 99.17% in terms of accuracy. The experimental result indicated that the appropriate statistics of RR intervals and SVM-RBF with its suitable parameters can perform well for PAF detection.

Keywords: heart problem, irregular beat, atrial fibrillation, support vector machine, statistics, multi-segment

INTRODUCTION

Paroxysmal Atrial fibrillation (PAF) is a heart problem with an irregular and often rapid heartbeat. It leads to the number of complications and death. The origin of atrial fibrillation is suggested from the atrial. When atrial fibrillation is happening the atrial beats chaotically and irregularly. PAF was suggested as independent cause of mortality^[1]. It was estimated that by 2050 atrial fibrillation in the USA would be 6–12 million and in Europe would be 17.9 million by 2060^[2, 3].

PAF is able to be diagnosed by using electrocardiogram^[4]. In terms of electrocardiogram parameter, the essential criterion of PAF is irregular RR interval and the absence of P-wave^[5]. It is caused by degeneration of the electrical impulses in atria yielding a rapid, chaotic heart rhythm. It is often rapid and irregular with no clear pattern or is known as irregularly irregular. Episodes of PAF in patients may be symptomatic or asymptomatic^[6]. A screening of older participants with mean age of 64 years results in atrial fibrillation of 2.3% in 122.571 participants^[7].

A detection system for PAF is meaningful to provide early information of a PAF episode. This system is very useful for asymptomatic PAF. Early information of PAF is essential for the base of the appropriate treatment to prevent further complications or permanent PAF. In addition, the system can help improve patient outcomes, as its long duration is associated with recurrence^[8].

An automatic PAF detection using single-lead electrocardiogram with RR intervals was initiated in 1983^[9]. Around a decade later an artificial neural network was introduced for atrial fibrillation detection^[10, 11]. The detection utilized RR intervals. Possibility of RR intervals for PAF analysis was also stated by other study^[11]. In 2011, the relation of atrial fibrillation and RR intervals for atrial fibrillation detection was studied again using different approach.

Around the past decade, various machine learning techniques were investigated to be applied for atrial fibrillation detection using RR intervals^[12-17]. Furthermore, in the past three years, deep learning-based classifications techniques were evaluated for atrial fibrillation detection. Studies investigated a deep convolutional neural networks (DCNN) for atrial fibrillation in 2017-2018.^{23,28} A recent study presented a method for atrial fibrillation detection using Convolutional- and Recurrent-Neural Networks (CNN and RNN)^[18].

Atrial fibrillation detection techniques with base of support vector machine (SVM) have been studied. A study implemented SVM and a stationary wavelet transform for atrial fibrillation detection^[19]. In the study, ECG was transformed using a stationary wavelet to obtain electrocardiographic features. These features were used for the input of SVM classification. Different study^[17] presented atrial fibrillation detection using SVM and heart rate variability (HRV). HRV was obtained from ECG and then it is for input for classification using SVM.

Regarding the electrocardiographic features for atrial fibrillation detection, different lengths of ECG for the base of detection have been studied. The ECG record with a certain length was commonly said as one segment. A study has examined three different segment lengths: 10, 15 and 30 seconds^[17]. A different study investigated three lengths: 32, 64, and 128 RR intervals.^[20] It utilized terms of the number of RR intervals, rather than time, to express the different length of segments.

Until recently, in atrial fibrillation detection, the ECG was used in the form of one segment. The length of a segment could be different, as explained above. This article introduces an PAF detection with more than one segment. ECG of one segment and two segments can be the same length. For example, one segment of 60 seconds has the same length as the two consecutive segments of 30 seconds. These different segment numbers, although with the same length, might affect the performance of PAF detection. This issue is novel and is presented in this article.

This article also presents the contribution of RR-interval in terms of the difference between one RR-interval and the others in a segment, rather than only the difference between two consecutive RR intervals. The RR-interval in terms of the difference between two consecutive RR intervals and in terms of RR interval for atrial fibrillation detection has been studied for atrial fibrillation detection^[15, 21].

Furthermore, this article presents an AF detection utilizing SVM dealing with the contribution of different segment numbers and RR interval in different terms. SVMs with different parameters are studied.

The rest of this article is presented as follows. The second section presents the proposed method which provides the features extraction, to find the statistical features of RR intervals, and SVM classification. The experimental results and discussion are presented in the third and fourth sections, respectively. Finally, the conclusion is presented in the fifth section.

METHODS

The proposed detection strategy is generally presented in Figure 1. ECG was electrocardiogram obtained from the patients. R-peaks series was the sequential peaks position in time. This series was obtained from the Physionet. [9, 22]. The R-peaks series was then segmented in certain length. RR intervals were obtained from the segment. The statistics of the RR intervals were calculated and were used for the input of SVM. The single output of SVM was a number which indicated PAF or normal.

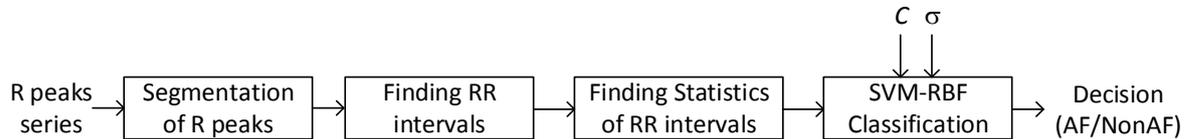


Figure 1. Strategy for PAF detection using statistical features of RR intervals and SVM-RBF.

Feature Extraction

Feature extraction was conducted to obtain electrocardiographic features. These features were for the input of SVM. We collected electrocardiographic R-peaks position in time and the attributes, PAF or normal, from the MIT-BIH atrial fibrillation database.

The R-peaks was arranged in a time series. This series was then segmented, as presented in Figure 2. Each segment was t seconds in length. An RR-interval was defined as the time interval between two consecutive R-peaks. From each segment we calculated features $F_1 - F_6$ as presented in Table 1.

Table 1. Electrocardiographic features

Feature Names	Feature Definition
F_1	The time interval between two consecutive R peaks, or RR interval, in a segment
F_2	The difference between two consecutive RR intervals, or ΔRR interval, in a segment.
F_3	The average difference of an RR interval and all the other RR interval in a segment
F_4	The RR intervals of the previous segment.
F_5	The ΔRR of the previous segment
F_6	The average difference of an RR interval and all the other RR intervals of the previous segment.

F_1 to F_6 are presented in terms of eleven their statistical parameters, S_1 - S_{11} , listed in Table 2.

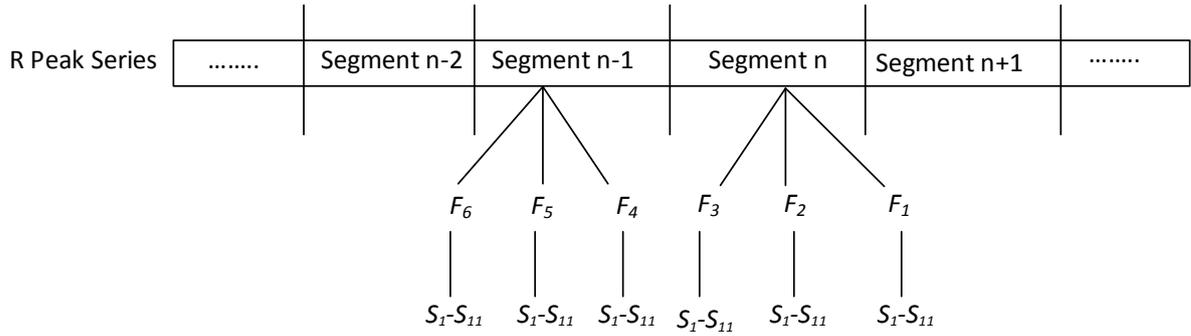


Figure 2. Finding features $F_1 - F_6$ and S_1-S_{11} from R peaks series

F_1 to F_6 were presented in terms of their statistical parameters, listed in Table 2. Eleven statistical parameters were utilized. Each F_1 to F_6 related to those eleven parameters. Thus, the total statistical parameters were 66. These parameters were studied for the input of SVM.

Table 2. Statistical parameters of the features F_1-F_6 as presented in Table 1.

Feature Names	Statistical Features
S_1	Arithmetic Mean
S_2	Geometric Mean
S_3	Harmonic Mean
S_4	Trimmed Mean
S_5	Range
S_6	Interquartile Range
S_7	Standard Deviation
S_8	Minimum
S_9	Maximum
S_{10}	Kurtosis
S_{11}	Skewness

SVM classification

SVM classification was to classify the electrocardiographic features presented above. The detail of SVM classification was presented by Burgers (1998). Essentially, an SVM technique created a hyperplane optimally separating two class data. Let (\mathbf{x}_i, y_i) was n two-class data. $\mathbf{x}_i \in R^k$ was data with k dimensional space and $y_i \in [-1, 1]$ was the associated class label. The optimal hyperplane was defined as $\mathbf{w} \cdot \mathbf{x} + b = 0$ maximally separating the data. \mathbf{w} was the hyperplane perpendicular vector and $|b| / \|\mathbf{w}\|$ was the distance of the hyperplane to the origin.

For common world problems such separating hyperplane did not exist. Hence it was introduced a positive slack variable ξ_i and then

$$y_i(\mathbf{w} \cdot \mathbf{x} + b) \geq 1 - \xi_i \quad (1)$$

The optimal hyperplane was determined by minimizing

$$C \sum_{i=1}^n \xi_i + 0.5 \|\mathbf{w}\|^2 \quad (2)$$

where C was called soft margin parameter controlling the trade-off between margin size and classification error. Obtaining the optimal hyperplane was conducted utilizing Lagrange multiplier approach by maximizing

$$L(\alpha) = \sum_{i=1}^n \alpha - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (3)$$

Subject to $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^n y_i \alpha_i$; α_i was Lagrange multiplier.

Input data was mapped to higher dimensional space by replacing the inner-product in the above equation by a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$. This approach was to find that nonlinearly separable data could be linearly classified. In this article, radial basis function (RBF) kernel function was used. RBF function was defined as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\sigma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (4)$$

σ was a constant defined by user. In this article σ was varied from 1 to 10.

Finally, the class prediction for any test vector $\mathbf{x} \in R^N$ was given by

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b) \quad (5)$$

where sign was a signum function. The $f(x)$ which was greater than 0 was associated with +1 class and the less than 0 was associated with -1 class.

Metrics for Experimental Result Evaluation

This article uses Atrial Fibrillation Database provided by Physionet.^{[22],[9]} The AFDB consists of 25 ECG records. The length of each record is about 10 hours. The sampling rate is 250 Hz with resolution of 12 bits. The most atrial fibrillation type in this database is paroxysmal. Physionet provides information which ECGs are atrial fibrillation and which one is not atrial fibrillation. For simplicity, in this article term of normal is used to stated the one which is not atrial fibrillation.

The proposed detection strategies were evaluated using 5-fold cross validation.^[23] The performance of the detection was measured in terms of sensitivity, specificity and accuracy. Sensitivity was the ability to determine PAF class correctly. Specificity was the ability to determine normal class correctly. Accuracy was the ability to differentiate PAF and normal classes. Mathematically, the three terms were defined as follows:

$$\text{Sensitivity} = tp / (tp + fp)$$

$$\text{Specificity} = tn / (tn + fp)$$

$$\text{Accuracy} = (tp + tn) / (tp + tn + fp + fn)$$

where tp = true positive, tn = true negative, fp = false positive, and fn = false negative.

RESULTS

This article presents strategies to detect atrial fibrillation. In general the strategies utilized an SVM and the statistical features of RR intervals. Eleven statistical features were measured from RR intervals of electrocardiogram. The features which belong to PAF and normal were presented in Table 3. The number was normalized to be between 0 and 1.

In general the statistics of PAF and non-PAF were statistically significantly-different, with p -value < 0.0001 . The means in PAF were lower than which in normal for all four terms of means: arithmetic, geometric, harmonic and trimmed means. The ranges in PAF were higher than which in normal for both terms: range and Interquartile range. In terms of measures of shapes, kurtosis and skewness, PAF was lower than normal. The minimum in PAF was lower than

which in normal. Conversely, the maximum in PAF was higher than which in normal. Finally, standard deviation in PAF was higher than which in normal.

Table 3. Statistics of electrocardiographic RR intervals of PAF and normal

Statistical Parameters	normal (Mean±SD)	PAF (Mean±SD)	<i>p</i> -value
Arithmetic Mean	0.4717±0.1483	0.3052±0.1378	<i>p</i> <0.0001
Geometric Mean	0.5094±0.1564	0.3251±0.1409	<i>p</i> <0.0001
Harmonic Mean	0.5403±0.1632	0.3397±0.1423	<i>p</i> <0.0001
Trimmed Mean	0.4902±0.1515	0.3177±0.1414	<i>p</i> <0.0001
Range	0.0385±0.0492	0.1396±0.0524	<i>p</i> <0.0001
Interquartile Range	0.0496±0.0840	0.2266±0.1076	<i>p</i> <0.0001
Standard Deviation	0.0540±0.0783	0.2089±0.0853	<i>p</i> <0.0001
Minimum	0.6077±0.1731	0.3166±0.0906	<i>p</i> <0.0001
Maximum	0.1594±0.0551	0.1643±0.0699	<i>p</i> <0.0001
Kurtosis	0.0561±0.0717	0.0371±0.0268	<i>p</i> <0.0001
Skewness	0.1920±0.1026	0.1707±0.0530	<i>p</i> <0.0001

We evaluated the PAF detection using SVM-RBF with C and σ of 100 and 10, respectively, and varied electrocardiographic features, F₁-F₆. This approach provided performances presented in Table 4. The segment length used was 30 seconds. Among those performances, the detection utilizing F₄ provided the highest accuracy, 98.27%, followed by which utilized F₁, 96.97%. F₁ and F₄ were the RR intervals of a segment and of the segment before it, respectively. The third and fourth highest accuracies were which utilized F₅ and F₂, respectively. F₂ and F₅ were Δ RR interval of a segment and of the previous one. The fifth and the lowest were F₆ and F₅, respectively. F₅ and F₆ were the average difference of an RR interval and all the other RR intervals of a segment and of the previous segment.

Table 4. The performances of PAF detection using SVM-RBF and single group of feature.

Features	Sensitivity (%)	Specificity (%)	Accuracy (%)
F ₁	96.51	97.29	96.97
F ₂	93.01	93.77	93.46
F ₃	88.59	92.43	90.83
F ₄	98.07	98.41	98.27
F ₅	94.64	94.78	94.72
F ₆	90.15	93.30	91.99

The performances of the detection systems which utilized combinations of feature group are presented in Table 5. The combinations were arranged by adding the feature group one by one, with the sequence from the feature group which provided the highest performance to the lowest, considering Table 4. Addition F₁ to F₄ provided higher accuracy from 97.95 to 98.81%. The next, by including F₂ the accuracy increased to 99.05%. Adding each F₃, F₅ and F₆ did not change the performance significantly. Thus, by employing only the three features (F₄,F₁,F₅) the highest performance was achieved. As mentioned above, F₁ and F₄ were the RR intervals of a segment and of the previous segment, respectively. F₅ was the difference between two consecutive RR intervals.

Table 5. The performances of PAF detection using SVM-RBF and combinations of feature groups.

Features	Sensitivity (%)	Specificity (%)	Accuracy (%)
F ₄	98.07	98.41	98.27
F ₄ ,F ₁	98.72	99.20	99.00
F ₄ ,F ₁ ,F ₅	98.79	99.38	99.13
F ₄ ,F ₁ ,F ₅ ,F ₂	98.61	99.42	99.08
F ₄ ,F ₁ ,F ₅ ,F ₂ ,F ₅	98.64	99.45	99.11
F ₄ ,F ₁ ,F ₂ ,F ₃ ,F ₅ ,F ₃	98.58	99.46	99.09

The above approach employed features from the electrocardiographic segment with length of 30 seconds. F₁ and F₄ were two consecutive segments where each of them was 30 seconds. Thus, F₁ and F₄ were totally 60 seconds. While the two segments F₁ and F₄ were joined, it became one segment with 60 seconds in length. We called this feature as F₁₄. Using similar explanation, the other two features were F₂₅ and F₃₆. Table 5 compares the performances of which used the segment of 30 seconds and of which utilize the 60 seconds. It can be seen that employing the 30 seconds provided higher performance in all the three performance terms used. It provided sensitivity, specificity and accuracy of 98.58%, 99.46%, and 99.09%, respectively.

Table 5. The performances of PAF detection using SVM-RBF and combination of features.

Features	Sensitivity (%)	Specificity (%)	Accuracy (%)
F ₄ ,F ₁ ,F ₂ ,F ₃ ,F ₅ ,F ₆	98.58	99.46	99.09
F ₁₄ ,F ₂₅ ,F ₃₆	96.83	98.19	97.61

The performances of the detections which utilized each of the statistical parameters of features F₄,F₁,F₅ are presented in Table 6. The highest accuracy was obtained by which utilized interquartile range, 94.84% in terms of accuracy. The second and third were which used arithmetic mean and trimmed mean, respectively. The two least were skewness and kurtosis.

Table 6. The performances of the detections which utilized each of the statistical parameters of the features constituting F₄,F₁,F₅

Feature Names	Statistical Features	Sensitivity (%)	Specificity (%)	Accuracy (%)
S ₁	Arithmetic Mean	98.03	91.17	94.02
S ₂	Geometric Mean	77.15	91.56	85.57
S ₃	Harmonic Mean	78.45	91.03	85.81
S ₄	Trimmed Mean	98.09	91.46	94.22
S ₅	Range	93.38	87.96	90.21
S ₆	Interquartile Range	94.48	95.10	94.84
S ₇	Standard Deviation	93.47	93.50	93.49
S ₈	Minimum	93.61	91.22	92.22
S ₉	Maximum	97.84	85.88	90.85
S ₁₀	Kurtosis	34.48	74.57	57.92
S ₁₁	Skewness	32.41	78.59	59.41

The performances of the detection which used the combinations of the statistical parameters by adding the statistical type one by one are presented in Table 7. The sequence of the addition considered the table 6, from the highest performance to the lowest. Each statistical type provided contribution to increase the accuracy, except harmonic range (S_5) and harmonic mean (S_3). It indicated that including all the eleven statistical parameters could provide the higher accuracy than the others.

Table 7. The performances of the detection which used the combinations of statistical features by adding the statistical types one by one.

Features	Accuracy (%)
S_6	94.84
S_6S_4	97.69
$S_6S_4S_1$	98.24
$S_6S_4S_1S_7$	98.71
$S_6S_4S_1S_7S_8$	98.77
$S_6S_4S_1S_7S_8S_9$	99.00
$S_6S_4S_1S_7S_8S_9S_5$	98.97
$S_6S_4S_1S_7S_8S_9S_5S_3$	98.94
$S_6S_4S_1S_7S_8S_9S_5S_3S_2$	98.98
$S_6S_4S_1S_7S_8S_9S_5S_3S_2S_{11}$	99.06
$S_6S_4S_1S_7S_8S_9S_5S_3S_2S_{11}S_{10}$	99.13

According to Table 4, the highest performance of the detection was which utilized F_4, F_1, F_5 , 99.13% in terms of accuracy. As mentioned before, this performance was obtained while the SVM utilized σ and C of 10 and 100, respectively. Utilizing different σ and C , the performances were presented in Figure 2. C was varied from 1 to 500 and σ was varied from 1 to 10. Figure 2 shows that the highest accuracy was 99.17%. This was obtained while $\sigma = 5$ and $C = 200$. The Figure also indicated that lower C and σ tended to lower accuracy. After an optimum value was achieved, increasing C resulted in almost constant accuracy.

In more detail, Figure 3 shows that while C was more than 100 the accuracy remained stable. Compared to which utilizes $\sigma = 5$, by using $\sigma = 10$ the accuracy was higher while C was less than 100. But, utilizing $\sigma = 5$ provided higher accuracy while C was more than 100.

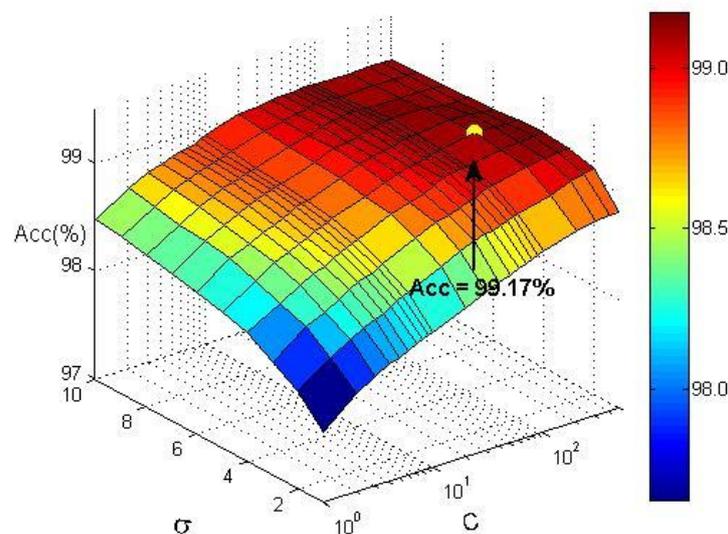
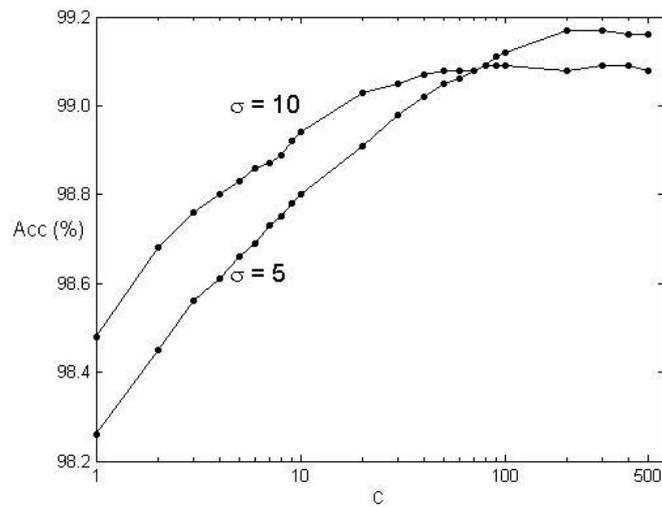
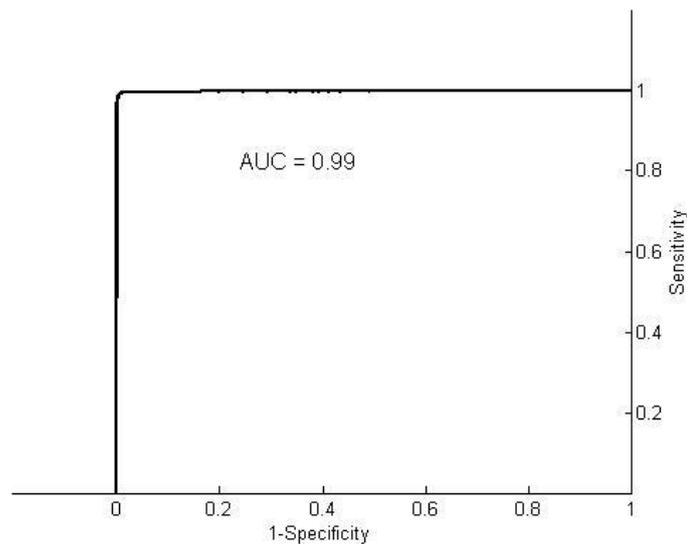


Figure 3. The accuracy the detection utilizing varied σ and C .**Figure 4.** The accuracy the detection utilizing $\sigma=5$, $\sigma=10$ and varied C .**Figure 5.** The Area under curve (AUC) of the detection.

The performance of the detection in term of receiver operating characteristics (ROC) curve is presented in Figure 5. The area under curve (AUC) of the ROC was 0.99. As the AUC was close to 1, the detection approach showed an excellent model. In other word, it had good measure of separability. It had an ability to differentiate electrocardiogram belonging to PAF.

The detection strategy which utilized different segment lengths is presented in Table 7. Three different lengths were 20, 30, and 40 seconds. Using the different lengths, the accuracies were nearly the same. But, the highest one was which utilized 30 seconds where the accuracy was 99.09%. For this comparison, utilizing all features were applied.

Table 7. The performances of the detection which used the combinations of statistical features by adding statistical types one by one.

Length (Seconds)	Sensitivity (%)	Specificity (%)	Accuracy (%)
20	98.54	99.38	99.03
30	98.58	99.46	99.09
40	98.39	99.54	99.06

Performance comparison with other studies

The performance of the proposed method compared with other methods studied in the last lustrum is presented in Table 8. Study by Yong Xia et al. ^[24] presented a technique for atrial fibrillation using a deep convolutional neural networks (DCNN). Using this approach, a segment of ECG was analysed using two types of signal processing, a short-term Fourier Transform (STFT) and stationary wavelet transform (SWT). The output of either STFT or SWT was used as input of DCNN. STFT-DCNN provided sensitivity and specificity of 98.34% and 98.24%, respectively. With nearly the same performance, SWT-DCNN provided of 98.79% and 97.87%, in terms of sensitivity and specificity, respectively.

Andersen R. S. et al. ^[18] introduced an atrial fibrillation detection using Convolutional- and Recurrent-Neural Networks (CNN and RNN). This approach utilized RR interval of electrocardiogram, rather than STFT or SWT, as presented by ^[19]. The performance found by this approach was 98.98% and 96.95%, in terms of sensitivity and specificity, respectively.

Pourbabaee, B. et al. ^[25] proposed atrial fibrillation detection technique using deep convolutional neural network. The electrocardiographic feature utilized by this method was electrocardiogram in time series. Thus, the feature was raw data of electrocardiogram. The performance obtained by this approach was 90.20% and 90.48%, in terms of sensitivity and specificity, respectively.

Asgari, S. et al. ^[19] utilized a support vector machine (SVM) for atrial fibrillation detection. This approach used a stationary wavelet transform to find electrocardiographic feature. The performance found by this approach was 97.0% and 97.1%, in terms of sensitivity and specificity, respectively.

To straightforwardly compared the techniques presented by the other studies presented above with the proposed method we used terms of geometric mean (*gm*). ^[26, 27] Mathematically, *gm* was defined as $gm = \sqrt{\text{Sensitivity} \times \text{Specificity}}$. In general the proposed method outperformed the others, as it had higher *gm*. In terms of specificity, the proposed technique was higher than the others. In terms of sensitivity, the introduced method was a bit lower than the method proposed by Andersen R. S. et al. ^[18] and higher than the others.

Table 8. Comparison of the proposed method with other methods presented in the last lustrum for PAF detection.

Authors	Methods	Sensitivity (%)	Specificity (%)	<i>gm</i> (%)
Asgari et al. ^[19]	SVM with wavelet transform	97.0	97.1	97.05
Pourbabaee, B. et al. ^[25]	CNN with ECG in time series	90.20	90.48	90.34
Yong Xia ^[24]	STFT-DCNN	98.34	98.24	98.29
	SWT-DCNN	98.79	97.87	98.33
Andersen R. S. et al. ^[18]	CNN with RR intervals	98.98	96.95	97.96
Porposed Method	SVM with RR intervals statistics	98.79	99.38	99.08

$$gm = \sqrt{\text{Sensitivit} \times \text{Specificity}}$$

DISCUSSION

This article presents a new strategy to detect an atrial fibrillation happening in human body. In general, the strategy utilized electrocardiogram and SVM. Essentially, SVM was a computer algorithm which creates an optimal separating-hyperplane. The creation was based on labelled training data. In other words, using given labelled training dataset, the algorithm could provide optimal hyperplane, which can do classification for test dataset.

This article used the atrial fibrillation database (AFDB) provided by Physionet ^[28]. This database has been used by other articles for creating and testing models of atrial fibrillation detection. ^[19, 24] The benefit of using this database is that the performance of a proposed algorithm is reasonable to be compared with the available algorithms which also use this database.

Statistical parameters of electrocardiographic RR intervals in relation to PAF and normal are presented in this article. Statistical features of electrocardiogram was reasonable for PAF detection as PAF rhythm was not regular, where a normal rhythm was regular. Four terms of mean were included. PAF was lower than normal in all four terms of mean. The lower means in PAF than in normal indicated that the RR intervals in PAF tended to be shorter than in normal. In other words, the heart rate in PAF was faster than which in normal ^[29]. Statistical mean refered to the average which was used to derive the central tendency of data. Arithmetic mean was mostly used and was often preferred over simply "mean". Geometric, harmonic and trimmed means were also included in this article. The effect of outliers was greatly dampened in these three types of mean.

The range of RR intervals in PAF was higher than which in normal. Range was simply measure of dispersion and was the difference between the minimum and maximum in a dataset. It was a measure of dispersion. The higher dispersion of RR interval in PAF was also indicated by the higher interquartile range and standard deviation. Interquartile range was a measure of variability. It indicated the extent to which the central 50% of values within the dataset were dispersed, and it was based upon the median. Standard deviation was also a measure of spread out a data set. Standard deviation summarised the amount by which every value within a dataset varied from the mean. The higher variation in PAF indicated that the heart rate in PAF was

more unstable than which in normal, as presented in other studies ^[4]. During PAF, the QRS complexes occurred in an irregular pattern. This was presented as varying RR intervals.

The detection utilizing RR interval outperformed which utilizes the other two terms of RR interval difference: Δ RR interval and the difference between an RR interval and the others. In other words, the contribution of RR interval was higher than the other two terms.

The interesting result of the proposed strategy was that the inclusion of the previous segment increased the detection performance. The available methods, such as ^[19, 24], utilized a segment without the previous segment for one detection. Two consecutive segments, where each segment for example had 30 seconds in length, might have different statistical parameters with one segment of 60 seconds in length. The consecutive segments were a segment and its previous segment. Thus, by using two consecutive segments, there were two values which might inform more complete than which of one value. It implied that using two values, representing two consecutive state, was suitable for PAF detection, as RR intervals of PAF fluctuated rapidly. However, this article did not yet present electrocardiographic features using more than two consecutive segments. This needed further studies.

Detections utilizing electrocardiographic features with different length of segments did not result in significantly different performances. The length were 20, 30, and 40 seconds with the accuracies of 99.03%, 99.09% and 99.06%, respectively. The length of more than 40 seconds and less than 20 seconds were not yet tested for the proposed method. The 30 second length was chosen as this length was used for the standard definition of PAF ^[5].

The electrocardiographic statistical features in terms of mean and variability provided significant contribution for the performance. It was showed by that the approaches utilizing those features individually provided accuracies of more than 85% (Table 5).

By using only the six statistical features (Table 7), rather than all 11 features, the method provided accuracy of 99%. By addition of the other five features, the accuracy became 99.13%. The six features above were the feature which resulted in accuracy of more than 90% while the features were used individually.

The detections using kurtosis and skewness individually performed the least with the accuracies of less than 60%. Skewness measures symmetry. The symmetric data had the same situation to the left and right of the centre. Kurtosis measured the tail relative to a normal distribution. Data with large kurtosis had heavy tails, and conversely, data with low kurtosis had light tails.

Variations of SVM parameters were commonly conducted to find the optimal parameters ^[30, 31]. The SVM parameters, C -SVM and also σ -RBF, were user defined and their optimal values might differ for different applications. C -SVM was the maximum value of α , the Lagrangian. α was possible from zero to C . It was tuned to find the best state between the complexity of SVM model and empirical error. If too large, the model became very complex which could fall into the overfitting situation ^[32]. Overfitting of a model was an error occurring when the function of the model was too closely fit to training data. The error might happen where real-world data often consist of some degree of error or random noise within it. Conversely, when C was too small, the model was poor to separate the different class data during training. Intuitively, the model did not fit the data well enough. It often resulted in an excessively simple model. Another parameter, α -RBF, controlled the width for Gaussian distribution.

CONCLUSION

A proposed strategy to detect atrial fibrillation (AF) has been presented. Statistical parameters of electrocardiogram were used as the main predictor for PAF detection. The statistical parameters in terms of mean and variability provided significant contribution for the performance of PAF detection. In relation to the segmentation, the approach of the PAF detection using two segments resulted in higher performance than which used one segment. The detection using different segment lengths did not show significant different performance. Regarding the SVM-RBF algorithm, different C -SVM and σ -RBF provided different detection performances. The SVM-RBF using the optimum C -SVM and σ -RBF and the selected statistical parameters provided great performance of 99.17% in terms of accuracy.

REFERENCES

- 1 Andersson, T., A. Magnuson, I.-L. Bryngelsson, O. Frøbert, K.M. Henriksson, N. Edvardsson, and D. Poçi. 2013. All-cause mortality in 272 186 patients hospitalized with incident atrial fibrillation 1995–2008: a Swedish nationwide long-term case–control study. *European heart journal*. Vol. 34. No. 14. page. 1061-1067.
- 2 Krijthe, B.P., A. Kunst, E.J. Benjamin, G.Y. Lip, O.H. Franco, A. Hofman, J.C. Witteman, B.H. Stricker, and J. Heeringa. 2013. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *European heart journal*. Vol. 34. No. 35. page. 2746-2751.
- 3 Miyasaka. 2006. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence (vol 114, pg 119, 2006). *Circulation*. Vol. 114. No. 11. page. E498-E498.
- 4 Carrara, M., L. Carozzi, T.J. Moss, M. de Pasquale, S. Cerutti, M. Ferrario, D.E. Lake, and J.R. Moorman. 2015. Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy. *Physiological Measurement*. Vol. 36. No. 9. page. 1873-1888.
- 5 Kirchhof, P., S. Benussi, D. Kotecha, A. Ahlsson, D. Atar, B. Casadei, M. Castella, H.-C. Diener, H. Heidbuchel, and J. Hendriks. 2016. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *European journal of cardiothoracic surgery*. Vol. 50. No. 5. page. e1-e88.
- 6 Savelieva, I. and A.J. Camm. 2000. Clinical relevance of silent atrial fibrillation: prevalence, prognosis, quality of life, and management. *Journal of Interventional Cardiac Electrophysiology*. Vol. 4. No. 2. page. 369-382.
- 7 Lowres, N., L. Neubeck, J. Redfern, and S.B. Freedman. 2013. Screening to identify unknown atrial fibrillation. *Thrombosis and haemostasis*. Vol. 110. No. 08. page. 213-222.
- 8 Scherr, D., P. Khairy, S. Miyazaki, V. Aurillac-Lavignolle, P. Pascale, S.B. Wilton, K. Ramoul, Y. Komatsu, L. Roten, and A. Jadidi. 2015. Five-year outcome of catheter ablation of persistent atrial fibrillation using termination of atrial fibrillation as a procedural endpoint. *Circulation: Arrhythmia and Electrophysiology*. Vol. 8. No. 1. page. 18-24.
- 9 Moody, G., B and R. George, G. 1983. A new method for detecting atrial fibrillation using RR intervals. *Computers in Cardiology*. . page. 227-230.
- 10 Artis, S.G., R. Mark, and G. Moody. 1991. *Detection of atrial fibrillation using artificial neural networks*. IEEE.
- 11 Cubanski, D., D. Cyganski, E.M. Antman, and C.L. Feldman. 1994. A neural network system for detection of atrial fibrillation in ambulatory electrocardiograms. *Journal of cardiovascular electrophysiology*. Vol. 5. No. 7. page. 602-608.

- 12 Larburu, N., T. Lopetegi, and I. Romero. 2011. Comparative study of algorithms for atrial fibrillation detection. in *Computing in Cardiology, 2011*.
- 13 Nuryani, N., B. Harjito, I. Yahya, A. Lestari, E. Anzihory, and K. Farosi. 2015. Atrial Fibrillation Detection Using Swarm Fuzzy Inference System and Electrocardiographic P-Wave Features. *Procedia Computer Science*. Vol. 72. . page. 154-161.
- 14 Kurzweil, R.C., L. Gibson, P. Albrecht, and P. Grimshaw, 2016, *Atrial fibrillation detection and associated methods*. Google Patents.
- 15 Nuryani, N., B. Harjito, I. Yahya, M. Solikhah, R. Chai, and A. Lestari. 2017. Atrial fibrillation detection using support vector machine and electrocardiographic descriptive statistics. *International Journal of Biomedical Engineering and Technology*. Vol. 24. . page. 225-236.
- 16 Solari, D., E. Bertero, R. Miceli, C. Brunelli, P. Ameri, and M. Canepa. 2017. Methods, accuracy and clinical implications of atrial fibrillation detection by cardiac implantable electronic devices. *International Journal of Cardiology*. Vol. 236. . page. 262-269.
- 17 Andersen, R.S., E.S. Poulsen, and S. Puthusserypady. 2017. A novel approach for automatic detection of atrial fibrillation based on inter beat intervals and support vector machine. in *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*.
- 18 Andersen, R.S., A. Peimankar, and S. Puthusserypady. 2019. A deep learning approach for real-time detection of atrial fibrillation. *Expert Systems with Applications*. Vol. 115. . page. 465-473.
- 19 Asgari, S., A. Mehrnia, and M. Moussavi. 2015. Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. *Computers in biology and medicine*. Vol. 60. . page. 132-142.
- 20 Lian, J., L. Wang, and D. Muessig. 2011. A simple method to detect atrial fibrillation using RR intervals. *The American journal of cardiology*. Vol. 107. No. 10. page. 1494-1497.
- 21 Huang, C., S. Ye, H. Chen, D. Li, F. He, and Y. Tu. 2011. A Novel Method for Detection of the Transition Between Atrial Fibrillation and Sinus Rhythm. *IEEE Transactions on Biomedical Engineering*. Vol. 58. No. 4. page. 1113-1119.
- 22 Goldberger, A., L. Amaral, L. Glass, J. Hausdorff, P.C. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, 2000, *PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals*. *Circulation [Online]*. 101 (23), pp. e215–e220.
- 23 Wong, T.-T. 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*. Vol. 48. No. 9. page. 2839-2846.
- 24 Xia, Y., N. Wulan, K. Wang, and H. Zhang. 2018. Detecting atrial fibrillation by deep convolutional neural networks. *Computers in biology and medicine*. Vol. 93. . page. 84-92.
- 25 Pourbabaee, B., M.J. Roshtkhari, and K. Khorasani. 2017. Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. No. 99. page. 1-10.
- 26 Nuryani, N., S.S. Ling, and H. Nguyen. 2012. Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection. *Annals of biomedical engineering*. Vol. 40. No. 4. page. 934-945.
- 27 Georgoulas, G., D. Stylios, and P. Groumpos. 2006. Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using support vector machines. *IEEE Transactions on Biomedical Engineering*. Vol. 53. No. 5. page. 875-884.
- 28 PhysioBank, P. 2000. PhysioNet: Components of a New Research Resource for Complex Physiologic Signals [Circulation Electronic Pages]. *Circulation*. Vol. 101. No. 23. page. e215-e220.

- 29 Van Gelder, I.C., M. Rienstra, H.J.G.M. Crijns, and B. Olshansky. 2016. Rate control in atrial fibrillation. *The Lancet*. Vol. 388. 10046. page. 818-828.
- 30 Ghafoori, Z., S.M. Erfani, S. Rajasegarar, J.C. Bezdek, S. Karunasekera, and C. Leckie. 2018. Efficient unsupervised parameter estimation for one-class support vector machines. *IEEE transactions on neural networks and learning systems*. Vol. 29. No. 10. page. 5057-5070.
- 31 Zadeh, A.E. and A. Khazaei. 2011. High efficient system for automatic classification of the electrocardiogram beats. *Annals of biomedical engineering*. Vol. 39. No. 3. page. 996-1011.
- 32 Kao, I.-H., Y.-W. Hsu, Y.H. Lai, and J.-W. Perng. 2019. Laser Cladding Quality Monitoring Using Coaxial Image Based on Machine Learning. *IEEE Transactions on Instrumentation and Measurement*. Vol. 69. No. 6. page. 2868-2880.