Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 1, February 2024, pp. 97-104
DOI: https://doi.org/10.52436/1.jutif.2024.5.1.1267
p-ISSN: 2723-3863
e-ISSN: 2723-3871

ANALYSIS FEATURE EXTRACTION FOR OPTIMIZING ARRHYTHMIA CLASSIFICATION FROM ELECTROCARDIOGRAM SIGNALS

Yusril Ramadhan[1], Satria Mandala[*2]

[1]School of Computing, Telkom University, Indonesia
[2]Human Centric (HUMIC) Engineering, Telkom University, Bandung, Indonesia
Email: [1]yusrilramadhan@student.telkomuniversity.ac.id, [2]satriamandala@telkomuniversity.ac.id

(Article received: August 1, 2023; Revision: August 11, 2023; Published: February 31, 2024)

Abstract

Heart disease is the primary cause of death globally, with arrhythmias, such as Premature Atrial Contraction (PAC), Atrial Fibrillation (AF), and Premature Ventricular Contraction (PVC), being critical heart rhythm abnormalities. Although numerous studies have utilized feature extraction from electrocardiogram (ECG) signals to detect these conditions, optimal accuracy has not been achieved. Therefore, this research aims to identify relevant features and achieve better results by using dynamic feature extraction methods. The extracted features used are RR Interval, PR Interval, and QRS Complex. By combining 2 feature extractions - RR Interval & PR Interval, RR Interval & QRS Complex, and PR Interval & QRS Complex - this study achieves a high level of accuracy on the RR Interval & QRS Complex feature extraction, reaching 97.60%, with a specificity of 98.30% and sensitivity of 96.58%.

Keywords: *Arrhythmia, Dynamic Feature, Electrocardiogram, Feature Extraction.*

1. INTRODUCTION

In modern times, individuals face mounting stress from demanding work settings and the fast-paced routines of everyday life [1], [2]. The heightened stress levels and the rapid tempo of work have contributed to an elevated likelihood of experiencing arrhythmias or heart attacks. Cardiovascular diseases (CVDs) have emerged as one of the most significant health concerns in contemporary society. Annually, approximately 17.9 million people lose their lives to CVDs, accounting for 31% of all global fatalities [3]. Arrhythmia, a cardiovascular disease (CVD), is distinguished by irregular alterations in the regular heart rhythm. There are various forms of arrhythmias that can manifest, including Premature Atrial Contraction (PAC), Premature Ventricular Contraction (PVC) and Atrial Fibrillation (AF). The effective treatment of arrhythmias relies on their early detection [2]. AF is characterized by an irregular heart rate and carries significant health consequences, including a heightened risk of stroke [4, 5]. While a healthy individual typically maintains a normal heart rate between 60, and 100 beats per minute, AF disrupts this balance by causing the heart rate to exceed 100 beats per minute [6].

The presence of an underlying condition, such as hypertension, is frequently linked to Atrial Fibrillation (AF). While it can affect individuals of all age groups, it is more prevalent in the elderly population and less common in younger generations. On the other hand, Premature Atrial Contraction (PAC) is another frequently encountered type of heart arrhythmia, characterized by premature heartbeats originating from the atria. This particular arrhythmia is commonly observed in both healthy young individuals and the elderly, though its precise cause remains uncertain [7]. Premature Ventricular Contraction (PVC) are extra abnormal heartbeats originating from the ventricles. They typically happen when the ventricles contract earlier than the next normal beat and are followed by a forceful heartbeat, disrupting the proper pumping of blood. Frequent occurrence of Premature Ventricular Contraction (PVC) can pose a risk for arrhythmia-induced cardiomyopathy, a condition where the heart muscle becomes less efficient, potentially leading to heart failure [8]. Difficulties in diagnosing heart disease are mainly related to their paroxysmal and complex nature. Clinically, doctors usually diagnose it using an ECG. An electrocardiogram (ECG) is a method used to graphically depict the heart's activity over a period of time. It provides significant understanding heart's regular functioning. and the physiological condition of various body parts. Consequently, an ECG serves as a crucial tool in diagnosing heart diseases [9].

This study is currently examining efforts to optimize arrhythmia detection using the dynamic feature extraction method from electrocardiogram signals. Several related studies are being conducted to improve arrhythmia detection through the utilization of dynamic feature extraction from electrocardiogram signals, such as [10]-[16], in a related study conducted by Solikhah, Mar'Atus et al [10], The main aim of this study was to utilize the Multilayer Perceptron-Backpropagation (MLP-BP) neural network method to detect arrhythmias in an electrocardiogram, specifically targeting three types: Premature Ventricular Contraction (PVC), Left Bundle Branch Block(LBB), and Premature Atrial Contractions (PAC). The features considered for detection include the QRS Interval, RR Interval, and R Wave Gradient. This study explores various combinations of input features for neural networks, finally finding that the best results are obtained by using two specific features: the QRS interval and R wave gradient. The selected features exhibited satisfactory performance with regards to sensitivity (86.18%), specificity (75.07%), and accuracy (84.15%).

Chen, Xianjie et, al [11], the research focuses on developing AF algorithm detection that utilizes a combination of another feature extraction and a CNN or Convolutional Neural Network applied to electrocardiograph. The algorithm achieves impressive performance with accuracy (98.92%), specificity (97.04%), and sensitivity (97.19%). When comparing this algorithm to others, the average accuracy of those alternatives is found to be 80.26%. Remarkably, the proposed algorithm outperforms these alternatives with an accuracy rate that is 23.25% higher, showcasing its superiority in detecting atrial fibrillation from electrocardiograph signals.

Karri, Meghana and Sekhara, Chandra [12], The research aims to create an integrated system tailored to individual patients' ECG data for the detection of the QRS complex and arrhythmia classification. The proposed model combines DSM or Delta Sigma Modulation, DWT or Discrete Wavelet Transform, and a local max / min point algorithm to accurately identify QRS complex occurrences. Additionally, extracted features, including onset, peak, offset, and RR Interval, contribute to enhancing the classification accuracy. Arrhythmia classification is accomplished through the implementation of a LSTM or Long Short-Term Memory. The algorithm achieves exceptional performance, boasting accuracy rates of 99.64%, 99.15%, 99.87%, and 98.18% for all metrics (sensitivity, positive predictivity, accuracy, and F1 score, respectively).

Rohan, Banerjee et, al [13], In their study, the authors propose an innovative RNN, comprising two LSTM, designed to analyze the temporal patterns of PR Intervals and RR intervals in ECG recordings. The outputs of these Long-Short Term Memory are combined in the dense layer, along with manually crafted statistical features relating to heart rate variability (HRV) measurements. The results exhibit promising performance, with sensitivity (93%), specificity (98%), and F1-score (89%) in effectively classifying atrial fibrillation (AF).

Hui, Yang and Zhiqiang, Wei [14], This research introduces a novel feature to examine visual patterns of changes in QRS complex, along with the introduction of a new feature extraction algorithm based on clustering. The feature vectors acquired through this approach are then applied to three widely recognized classifiers (KNN, neural network, and SVM) and purpose automated diagnosis. The method underwent evaluation using complete set of fifteen heartbeat types recommended by the Association. The method achieved its highest overall accuracy (97.70%) when utilizing KNN, by incorporating the combined visual pattern features and parametric of Electrocardiogram morphology.

Chen, Chen et al [15], In this research, a deep learning approach was introduced, combining CNN or Convolutional Neural Networks and LSTM or Long Short-Term Memory automatically know six types Electrocardiogram signals: sinus bradycardia (SBR), atrial fibrillation (AFIB), atrial flutter (AFL), normal (N) sinus rhythm segments, pacing rhythm (P), and ventricular bigeminy (B). The proposed utilized a multi-input approach to examine 10-second segments of ECG signals along with their RR Intervals, both obtained from the MITDB. Using cross-validation with five-fold, this network get an impressive accuracy rate of 99.32%.

In the study [10], research was conducted for arrhythmia detection in electrocardiograms using the MLP-BP or Multi-layer Perceptron-Backpropagation ANN or Artificial Neural Network with multi-class classification. The study employed feature extraction from the RR interval, QRS interval, and R-wave gradient. However, the results obtained did not yield an optimal accuracy level.

Therefore, the objective of this research is to achieve more optimal accuracy by employing dynamic feature extraction methods on the RR interval, PR interval, and QRS complex features. Afterward, a performance analysis is conducted on the extracted features of RR Interval, PR Interval, and QRS Complex to determine the best feature extraction algorithm in terms of sensitivity, specificity, and accuracy for optimizing the detection of arrhythmia related to PVC, AF, and PAC.

2. RESEARCH METHODS

The research methodology employed in conducting this study is illustrated in Figure 1 below:

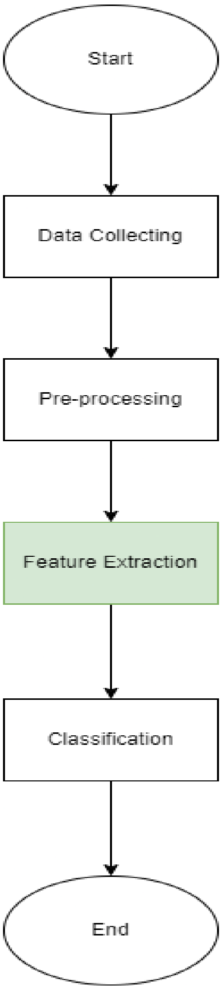

Figure 1. Reseach Methods

2.1. Data Collecting

This study utilized data acquired from two sources: the MIT-BIH Atrial Fibrillation Database (AFDB), the AFDB comprises 23 records, and the MIT-BIH Arrhythmia Database (MITDB), MITDB consists of 48 records. Both of which were made available by PhysioNet [17]. Later the data will be stored in the record patient, then details for the abnormal category data are made as shown in table 1.

table 1. Abnormal Category

Arrhythmia	Symbol	Total
PAC	A	2546
PVC	V	7130
AF	f	982
Normal	N	75052

After getting the data and categorized into an abnormal category then make the function make dataset is used to create a dataset from ECG (Electrocardiogram) data based on pre-defined categories of heart abnormalities. The steps begin with initializing a NumPy array and empty lists to store data from all patients.

Next, it loops through each patient in the record patients list, where for each patient, it loads the ECG signal and the corresponding heart abnormality symbols from the data files. The first ECG signal from each patient is selected for further analysis.

The obtained data is then organized into a Data Frame to manage information about heart abnormalities based on the abnormality symbols and their sample indices. The filtered data only includes heart abnormality symbols that fall within the pre-defined categories.

Following that, sections of the ECG signal are selected, spanning 'num_sec' seconds before and after each heart abnormality symbol. In case the length in the ECG segment corresponds to the anticipated duration, it is stored in a NumPy array labeled as X, along with the corresponding heart abnormality symbol recorded in array Y.

The above steps are performed for all patients in record patients, and the results from all patients are combined into ‘X_all’, ‘Y_all’, and ‘sym_all’. The outcome of this function is an ECG dataset (X_all) containing ECG signal segments, their corresponding heart abnormality category labels (Y_all), and the heart abnormality symbols (sym_all) from all patients that fall within the pre-defined categories of heart abnormalities. By using this function, ECG data can be prepared and further processed for the analysis and modelling of existing heart abnormalities in the patients.

2.2. Pre-processing

In this study, the data preprocessing process involves a series of sequential steps aimed at preparing the signal data for further analysis. First, it iterates through the records in record_patients, loading the signal data and corresponding annotations using the wfdb.rdrecord() and wfdb.rdann() functions, respectively. The ECG signal is then extracted from each loaded record, and any potential NaN values are replaced with zeros.

Next, the ECG signal sig is normalized to have zero mean and unit standard deviation. This normalization step ensures that the signal is scaled appropriately and becomes more suitable for further processing.

Subsequently, the signal is resampled to a new sampling rate defined as new_sampling_rate, which in this case is set to 250 samples per second (new_sampling_rate = 250). Resampling is employed to adjust the sampling frequency while retaining the signal's essential characteristics. Following that, a bandpass filter is applied to the resampled signal to isolate specific frequency components of interest. The bandpass_filter() function, implementing the Butterworth filter design, carries out the filtering operation with cutoff frequencies set between 5.0 Hz and 15.0 Hz, as defined by the lowcut and highcut parameters.

Subsequent to bandpass filtering, the signal sig_filtered undergoes wavelet denoising to eliminate undesired noise while preserving crucial features of the ECG signal. This is achieved by employing a discrete wavelet transform with the Daubechies 6 (db6) wavelet and soft thresholding to remove noise from the wavelet coefficients. The result, a denoised signal, is obtained by reconstructing the modified wavelet coefficients. Finally, the code generates a synthetic ECG signal sig with a frequency of 500 Hz using the integrate.cumulative_trapezoid() function from the SciPy library. This synthetic signal spans a duration of 10 seconds and is sampled at a rate of 1000 samples per second.

Overall, the code demonstrates a robust signal processing pipeline for ECG data, encompassing data loading, normalization, resampling, filtering, denoising, and even the generation of synthetic ECG signals for further analysis or experimentation.

2.3. Feature Extraction

According to Marinho, Leandro et al [18], The study introduced feature extraction methods that encompassed different features of the RR intervals and the morphology ECG signal, resulting total that get 155 attributes. Among these attributes, seven were specifically related to the events occurring in the heartbeats, such as post-RR interval, pre-RR interval, local RR interval, mean RR interval, T wave duration, QRS duration, and the P wave. Additionally, there were 148 attributes linked to the morphology of the ECG signal, extracted from each channel recorded during the ECG, resulting in 74 attributes from each channel.

These morphological attributes captured aspects such as PF-QRS complex, QRS complex, normalized QRS complex, normalized, T wave, PF-T wave, normalized T wave, and normalized PF-T wave. The combination of these RR Interval and ECG or electrocardiogram signal morphology attributes provided a comprehensive set of features for their analysis and classification tasks.

Therefore, in this research, the utilized feature extraction includes RR Interval, PR Interval, and QRS Complex.

In the search for RR Intervals in figure 2, the electrocardiogram (ECG) recording data is processed to generate a dataset containing RR interval segments from multiple patients. First, the parameters are set, namely the number of seconds taken as the length of RR interval segments (‘num_sec’) and the sampling frequency (**fs**), which indicates the number of samples taken per second from the ECG signal.

Next, the ‘make_dataset’ function is used to process the patients' recording data (‘record_patients’). This function takes the first ‘num_sec’ seconds from each ECG recording and labels whether the recording belongs to the "normal" or "abnormal" category based on the ‘abnormal_category’ variable.

After obtaining the RR interval segments from the dataset, stored in the ‘rr_intervals’ variable, the next step is to calculate features from each RR interval segment. This is done by iterating through each RR interval segment using a loop, and in each iteration, the ‘extract_features(rr)’ function is called to calculate the features of that RR interval segment. The results of these feature calculations are stored in the ‘features’ variable.

Finally, the previously computed features are organized into a Data Frame using ‘pd.DataFrame’, and the result is stored in the ‘df_features’ variable. This Data Frame will contain information about the features of the RR interval segments, which will be used for further analysis. All these steps aim to provide further insights into analysing the heart condition of patients based on their ECG recording data.

In the search for PR Intervals in figure 3, the process of analysing electrocardiogram (ECG) recording data to extract PR interval segments and compute their corresponding features is performed. First, the parameters ‘num_sec’ and **fs** are set, representing the number of seconds for PR interval segments and the sampling frequency, respectively.

Next, call the ‘make_dataset()’ function, this function extracts the initial ‘num_sec’ seconds from each ECG recording and categorizes them as "normal" or "abnormal" based on the ‘abnormal_category’ variable.

Subsequently, the PR interval segments are obtained and stored in the variable ‘pr_intervals’. These PR interval segments correspond to the time duration between the QRS complex, and P-wave on the ECG signal, which signifies the electrical conduction time from the atria to the ventricles of the heart.

The next step involves calculating features from each PR interval segment. This is accomplished by iterating through each PR interval segment using a loop. In each iteration, the function ‘extract_features(pr)’ is called to compute the features specific to that PR interval segment. The resulting features are then stored in the ‘features’ variable. Finally, the computed features are organized into a Data Frame using ‘pd.DataFrame()’, and the Data Frame is saved in the variable ‘df_features’. This Data Frame contains information about the features of the PR interval segments, which will be useful for further analysis.

In the search for QRS Complex in figure 4, the processing of electrocardiogram (ECG) data is performed to extract QRS complex segments and calculate their corresponding features. The first step involves setting the parameters, namely ‘num_sec’, which represents the number of seconds used as the length of each QRS complex segment, and **fs**, denoting the sampling frequency, indicating the number of samples taken per second from the ECG signal.

Next, the ‘make_dataset’ function is utilized to create a dataset from the ECG recording data (‘record_patients’). This function extracts ‘num_sec’ seconds of ECG data from each recording and labels them as "normal" or "abnormal" based on the ‘abnormal_category’ variable. Following this, the QRS complex segments are obtained and stored in the variable ‘qrs_complexes’. The subsequent step involves calculating features from each QRS complex segment. This is achieved by iterating through each QRS complex segment using a loop. During each iteration, the function ‘extract_features(qrs)’ is called to compute the features specific to that QRS complex segment. The resulting features are then stored in the ‘features’ variable.

Finally, the computed features are organized into a Data Frame using ‘pd.DataFrame’, and the Data Frame is saved in the variable ‘df_features’. This Data Frame contains information about the features of the QRS complex segments, which will be useful for further analysis. The entire process aims to gain insights into the heart condition of patients based on their ECG recording data, particularly focusing on the QRS complexes and their derived features.

As seen in Figure 2 illustrates the RR Interval detection process, which involves identifying all R-wave peaks and then performing calculations between each pair of R-wave peaks.

As seen in Figure 3 illustrates the PR Interval detection process, it is necessary to identify all R-wave peaks, P-wave peaks, and the starting point of the QRS Complex. Only after obtaining all the required points, the calculations can be carried out.

As seen in Figure 4, for QRS Complex detection, it is necessary to identify all R-wave peaks, Q waves, and S waves. Only after obtaining all the required points, the calculations can be carried out.

2.4. Classification

At the beginning of the process classification, the model utilizes K-Fold Cross Validation to partition the data into folds and applies standardization and reshaping to the input features to enhance model performance. The CNN model is subsequently constructed using the Sequential API, which involves implementing Conv1D layers for feature extraction, MaxPooling1D for dimensionality reduction, and Dense layers for classification. During the training process, the model is set up with the 'categorical_crossentropy' loss function, 'adam' optimizer, and 'accuracy' as the evaluation metric.

Once the training is completed, the model's performance is assessed on the test data using the evaluate method, which provides the loss and accuracy metrics. After performing cross-validation, the average accuracy and confusion matrix are calculated. A classification report is generated which includes sensitivity and specificity measures for each abnormal class. The selected test metric is the multi-class confusion matrix, which gives a comprehensive view of the accuracy of predictions for each heart rhythm category. The following test matrix is used:

$$Sensitivity = \frac{TP}{TN+FN} \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2)$$

$$accuracy = \frac{TP + TN}{TP + TN+FP+FN} \quad (3)$$

In this context, TP (True Positive) refers to the number of correctly detected Atrial Fibrillation beats, while TN (True Negative) represents the count of accurately identified normal beats. FP (False Positive) indicates normal beats mistakenly and detected as Atrial Fibrillation (AF), and FN (False Negative) signifies the count of AF beats erroneously identified as normal.

3. RESULTS AND DISCUSSION

The results presented here are based on a comprehensive feature extraction analysis involving the RR Interval, QRS Complex, and PR Interval, followed by a classification process using a Convolutional Neural Network (CNN).

The main objective of this analysis is to achieve the most optimal results, with special emphasis on evaluating the accuracy, sensitivity and specificity values. By focusing on these performance metrics, this study aims to ensure the effectiveness of the model in identifying accurately in order to obtain more optimal results compared to previous studies.

3.1. Variation 1 Feature

Table 2. Variation 1 Feature

Feature	Sensi-tivity (%)	Specificity (%)	Accua-cy (%)
RR Interval	95.28%	97.78%	96.99%
PR Interval	94.18%	97.22%	96.46%
QRS Complex	94.61%	97.47%	96.66%
RR Interval	95.28%	97.78%	96.99%
PR Interval	94.18%	97.22%	96.46%

In Table 2, it can be observed that there are three variations of features evaluated in this research, namely RR Interval, QRS Complex, and PR Interval. The performance evaluation results for each feature variation are also recorded in the table. For RR Interval, it was found to have a sensitivity (95.28%), specificity (97.78%), and accuracy (96.99%). As for the PR Interval, they achieved sensitivity (94.18%), specificity (97.22%), and accuracy (96.46%). Furthermore, for the QRS Complex, the evaluation results showed sensitivity (94.61%), specificity (97.47%), and accuracy (96.66%).

Interestingly, among the three evaluated feature variations, RR Interval showed the best performance. RR Interval in figure 5 achieved the highest values in terms of sensitivity (95.28%), specificity (97.78%), and accuracy (96.99%). This indicates that RR Interval is the most reliable feature in the analysis and detection conducted in this research.

These findings contribute significantly to the field of analysis related to the research subject. With high sensitivity, RR Interval accurately detected most positive cases. On the other hand, with high specificity, this feature also identified negative cases effectively. Moreover, the high accuracy of RR Interval demonstrates its capability to provide results that closely approximate the truth overall.

```
Mean Accuracy: 0.9698993921279907
Mean Confusion Matrix:
[[ 902    5   75]
 [   3 2432  101]
 [  72   64 6977]]
Classification Report:
              precision    recall  f1-score   support

           f       0.95      0.89      0.92       190
           A       0.96      0.96      0.96       506
           V       0.97      0.98      0.98      1430

    accuracy                           0.97      2126
   macro avg       0.96      0.95      0.95      2126
weighted avg       0.97      0.97      0.97      2126

Mean Sensitivity: 0.9528501929204044
Mean Specificity: 0.9778894845218706
```

Figure 5. RR Interval

3.2. Variation 2 Features

Table 3. Variation 2 Features

Feature	Sensi-tivity (%)	Specificity (%)	Accua-cy (%)
RR Interval & PR Interval	96.43%	98.19%	97.45%
PR Interval & QRS Complex	95.68%	98.03%	97.44%
RR Interval & QRS Complex	96.58%	98.30%	97.60%
RR Interval & PR Interval	96.43%	98.19%	97.45%
PR Interval & QRS Complex	95.68%	98.03%	97.44%

In table 3 reveals that RR Interval & PR Interval obtained sensitivity (96.43%), specificity (98.19%), and accuracy (97.45%). For PR Interval & QRS Complex, the values were sensitivity (95.68%), specificity (98.03%), and accuracy (97.44%). As for RR Interval & QRS Complex, the corresponding values were sensitivity (96.58%), specificity (98.30%), and accuracy (97.60%). The highest performance among the combinations of two features was achieved by RR Interval & QRS Complex in figure 6, with sensitivity (96.58%), specificity (98.30%), and accuracy (97.60%).

```
Mean Accuracy: 0.9759195685386658
Mean Confusion Matrix:
[[ 931    6   45]
 [   1 2451   84]
 [  54   66 6993]]
Classification Report:
              precision    recall  f1-score   support

           f       0.95      0.98      0.97       194
           A       0.95      0.97      0.96       478
           V       0.99      0.98      0.98      1454

    accuracy                           0.98      2126
   macro avg       0.97      0.98      0.97      2126
weighted avg       0.98      0.98      0.98      2126

Mean Sensitivity: 0.9658588298094017
Mean Specificity: 0.9829710823522274
```

Figure 6. RR Interval & QRS Complex

These findings highlight the strong performance of RR Interval when paired with either PR Interval or QRS Complex. The combination of RR Interval & PR Interval showed excellent sensitivity and specificity, with an overall accuracy (97.45%). Similarly, when RR Interval was combined with QRS Complex, the results remained consistently high, with sensitivity (96.58%), specificity (98.30%), and accuracy (97.60%).

The study's emphasis on combining different features to enhance the accuracy of analysis and detection is crucial in medical research. The results indicate that leveraging multiple features can lead to improved diagnostic capabilities and a deeper comprehension underlying patterns in the data.

3.3. Best Variety of Features

Table 4. Best Variety of Features

Feature	Sensi-tivity (%)	Specificity (%)	Accua-cy (%)
RR Interval	95.28%	97.78%	96.99%
PR Interval	94.18%	97.22%	96.46%
QRS Complex	94.61%	97.47%	96.66%
RR Interval & PR Interval	96.43%	98.19%	97.45%
PR Interval & QRS Complex	95.68%	98.03%	97.44%

After conducting experiments with single feature extraction and two feature extraction variations, it can be observed from Table 4 that the two-feature combination of RR Interval and QRS Complex outperforms the single-feature approach. The fusion of QRS Complex and RR Interval features achieved higher values for Sensitivity (96.58%), Specificity (98.30%), and Accuracy (97.60%).

The results presented in Table 4 demonstrate the advantage of utilizing multiple features in the analysis. By combining RR Interval and QRS Complex, the diagnostic performance improved significantly, as indicated by the higher sensitivity, specificity, and accuracy values compared to using only a single feature.

4. DISCUSSION

The evaluation results highlight the significance of the RR Interval feature in detecting and analyzing the studied condition. With sensitivity (95.28%), specificity (97.78%), and accuracy (96.99%) respectively, RR Interval consistently outperformed the other two features, PR Interval and QRS Complex. This suggests that RR Interval is a reliable and robust indicator for the detection and assessment of the condition under study.

Moreover, the investigation into combinations of features provided valuable insights. When RR Interval was combined with either PR Interval or QRS Complex, the outcomes demonstrated outstanding performance in regard to sensitivity, specificity, and accuracy. This indicates that incorporating RR Interval into feature combinations significantly improves the overall effectiveness of the detection process. Specifically, the combination of RR Interval & PR Interval achieved an impressive accuracy (97.45%), while RR Interval & QRS Complex exhibited the highest performance among all combinations with accuracy (97.60%). These findings highlight the synergistic effect of RR Interval when combined with other ECG features, which enhances the diagnostic power and potential applications in clinical settings.

In previous research [10], variations on 2 features, namely the QRS Interval and R Wave Gradient, obtained values of sensitivity (86.18%), specificity (75.07%), and accuracy (84.15%). This proves that the combination of QRS Complex and RR Interval gets more optimal values, namely Sensitivity (96.58%), Specificity (98.30%), and Accuracy (97.60%).

ACKNOWLEDGMENTS

This research was supported by the School of Computing and Human Centric Engineering (HUMIC), Telkom University, Bandung, Indonesia.

5. CONCLUSSION

The conclusion of the above discussion is that the RR Interval is a crucial feature in detecting and analyzing the investigated conditions. The RR Interval consistently outperforms two other features, namely the PR Interval and QRS Complex, as evidenced by its high sensitivity (95.28%), specificity (97.78%), and accuracy (96.99%). The research results indicate that the RR Interval is a reliable and robust indicator for detecting and assessing the investigated conditions.

Furthermore, investigating feature combinations provides valuable insights. When the RR Interval is effectively combined with either the PR Interval or the QRS Complex, the results show outstanding performance in terms of sensitivity, specificity, and accuracy. This indicates that incorporating the RR Interval with other EKG features significantly enhances the effectiveness of the detection process. The combination of RR Interval & PR Interval achieves an impressive accuracy (97.45%), while the combination of RR Interval & QRS Complex shows the highest performance among all combinations with an accuracy of (97.60%). These findings highlight the synergistic effect of the RR Interval when combined with other EKG features, improving diagnostic power and potential applications in clinical settings.

Previous research [10] showed suboptimal results in terms of sensitivity, specificity, and accuracy. Therefore, the aim of this study is to demonstrate that the use of the combination of QRS Complex and RR Interval features results in more optimal values, with value sensitivity (96.58%), specificity (98.30%), and accuracy (97.60%). For future research, efforts can be made to explore different variations of feature extraction and further optimize the methods for detecting arrhythmia. This endeavor aims to identify more informative and relevant features in the analysis of electrocardiogram (ECG) signals to enhance the ability to detect and classify arrhythmias.

REFERENCES

[1] A. Nigolian, N. Dayal, H. Nigolian, C. Stettler, H. Burri, Diagnostic accuracy of multi-lead ECGs obtained using a pocket-sized bipolar handheld event recorder, J. Electrocardiol. 51 (2018) 278–281.

[2] H. Fujita, D. Cimr, Computer aided detection for fibrillations and flutters using deep convolutional neural network, Inform. Sci. 486 (2019) 231–239.

[3] WHO 2016. https://www.who.int/cardiovascular_diseases/en/. (Accessed 28 July 2023).

[4] M. Abdar, M. Zomorodi-Moghadam, X. Zhou, R. Gururajan, X. Tao, P.D. Barua, R. Gururajan, A new nested ensemble technique for automated diagnosis of breast cancer, Pattern Recognit. Lett. (2018).

[5] M. Abdar, N.Y. Yen, J.C.-S. Hung, Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees, J. Med. Biol. Eng. 38 (2018) 953–965.

[6] H. Fujita, D. Cimr, Decision support system

for arrhythmia prediction using convolutional neural network structure without preprocessing, Appl. Intell. (2019) 1–9.

[7] B. M. Mathunjwa, Y.-T. Lin, C.-H. Lin, M. F. Abbod, and J.-S. Shieh, "Ecg arrhythmia classification by using a recurrence plot and convo- lutional neural network," Biomedical Signal Processing and Control, vol. 64, p. 102262, 2021.

[8] B. Akdemir, H. Yarmohammadi, M.C. Alraies, W.O. Adkisson, Premature ventricular contractions: Reassure or refer? Cleve. Clin. J. Med. 83 (7) (2016) 524–530.

[9] C. Che, P. Zhang, M. Zhu, Y. Qu, and B. Jin, "Constrained transformer network for ecg signal processing and arrhythmia classification," BMC Medical Informatics and Decision Making, vol. 21, 06 2021.

[10] SOLIKHAH, M., NURYANI, N., AND DARMANTO, D. Deteksi aritmia pada elektrokardiogram dengan metode jaringan syaraf tiruan kelas jamak menggunakan fitur interval rr, lebar qrs, dan gradien gelombang r. Jurnal Fisika dan Aplikasinya 11 (01 2015), 36.

[11] X. Chen, Z. Cheng, S. Wang, G. Lu, G. Xv, Q. Liu, and X. Zhu, "Atrial fibrillation detection based on multi-feature extraction and convolutional neural network for processing ecg signals," Computer Methods and Programs in Biomedicine, vol. 202, p. 106009, 2021.

[12] M. Karri and C. S. R. Annavarapu, "A real-time embedded system to detect qrs-complex and arrhythmia classification using lstm through hybridized features," Expert Systems with Applications, vol. 214, p. 119221, 2023.

[13] R. Banerjee, A. Ghose, and S. Khandelwal, "A novel recurrent neural network architecture for classification of atrial fibrillation using single- lead ecg," in 2019 27th European Signal Processing Conference (EU- SIPCO), pp. 1–5, 2019.

[14] H. Yang and Z. Wei, "Arrhythmia recognition and classification using combined parametric and visual pattern features of ecg morphology," IEEE Access, vol. 8, pp. 47103–47117, 2020.

[15] CHEN, C., HUA, Z., ZHANG, R., LIU, G., AND WEN, W. Automated arrhythmia classification based on a combination network of cnn and lstm. Biomedical Signal Processing and Control 57 (2020). Cited by : 88.

[16] DUAN, J., WANG, Q., ZHANG, B., LIU, C., LI, C., AND WANG, L. Accurate detection of atrial fibrillation events with r-r intervals from ecg signals. PLoS ONE 17, 8 August (2022). Cited by: 0; All Open Access, Gold Open Access, Green Open Access.

[17] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220. M.

[18] L. B. Marinho, N. de M.M. Nascimento, J. W. M. Souza, M. V. Gurgel, P. P. Rebouc ̧as Filho, and V. H. C. de Albuquerque, "A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification," Future Generation Computer Systems, vol. 97, pp. 564–577, 2019.