Heart Diseases Diagnosis Based on ECG Harmonic Analysis and Pattern Classification

Alejandro Vidales-Esquivel[0009-0009-7213-9435], Fernando Ornelas-Tellez[0000-0002-6428-8184], and José Ortiz-Bejar[0000-0001-8023-8008]

Faculty of Electrical Engineering, Universidad Michoacana de San Nicolas de Hidalgo, F.J. Mujica S/N, Cd. Universitaria, Morelia, 58030, Mexico

Abstract. Heart diseases have been a critical issue to deal with to improve people's health. Medical research and technology are being developed to obtain accurate diagnoses and treatments. This paper contributes to the design of an automated diagnosis system to classify electrocardiogram (ECG) signals to detect cardiac diseases. The proposed diagnosis system is based on the Fourier series analysis, which uses a dynamical state observer to instantaneously obtain salient features and patterns from the ECG harmonic content in real-time, whose information is classified through a K-Nearest neighbor algorithm (KNN), named as classifier, which determines the possible disease. The ECG signals used in this paper are obtained from the free online available PhysioNet databases, which contain information that can be used for the diagnosis and classification of healthy patients, arrhythmia cases, myocardial infarction, and heart failure. The proposed automated procedure is 93% effective in disease detection for the explored databases, highlighting its potential as a classification tool for ECG-based diagnosis.

Keywords: Heart diseases diagnosis · ECG · Fourier series · State observer · Harmonic content · Computational classifiers.

1 Introduction

Currently, derived from technological developments and electronic devices, different disciplines are improving their information analysis to make better decisions. It is even desirable to automate processes when possible. For instance, mechanisms in medicine are progressing to automate diagnosis, such as identifying cardiac pathologies from biomedical signals through electrocardiogram datasets, among other means. In 2021, around 220,000 people in Mexico died due to cardiovascular diseases, making it the leading cause of death in people over 55 years old [18]. The myocardial infarction was the primary cardiovascular disease and caused 177,000 deaths in 2021 [17]. Meanwhile, in the same year, 20.5 million deaths were recorded worldwide, according to reports from the World Heart Association [23].

Electrocardiograms are essential and non-invasive methods for diagnosing heart diseases. However, manual analysis based on human visualization/interpretation

https://doi.org/10.61728/AE20255084

of these ECG signals requires highly specialized knowledge, as they can be misinterpreted or misdiagnosed. Currently, an ECG is the result of electrocardiography, as the process of recording the electrical activity of the heart [22]. The obtained data can be displayed in an electrocardiograph for medical analysis, and additionally, the data can be digitally stored for computerized analysis purposes. In this sense, the importance of our proposed research is the faster and more automated analysis of ECG signals as a tool for detecting heart pathologies for adequate treatment by a specialist.

Techniques to automatically classify ECG signals and for detecting heart diseases have been developed. In the case of arrhythmias, for example, authors in [10] use Deep Separable Convolutional Neural Networks with Focal Loss (DSC-FL-CNN) to analyze and classify arrhythmias in ECG signals; in [24], authors use a Mixed-Kernel-based Extreme Learning Machine-based Random Forest binary classifier (MKELM-RF) with 98.1% accuracy in the classification. Also, in [3] employ a deep arrhythmia diagnosis method named Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BLSTM), with an accuracy of 96.59%. In [14], they employ a method to detect arrhythmias based on a combination of convolutional neural network and short-term memory. Another way to analyze the ECG signals is by preprocessing the information in the frequency domain. For instance, in [12], the data is preprocessed using the reduced-time Fourier transform and then image-focused neural networks are applied to classify the ECGs into two categories: healthy or presenting a heart anomaly. However, the method does not give a specific diagnosis the disease, that is, it only gives the result of whether the patient is healthy or has a heart disease. In this study, little real data is used, and therefore, artificial signals are used to simulate the ECGs. Another alternative is the usage of the Wavelet transform as done in [16], in which a model is developed, and use a support vector machine (SVM) classifier for the arrhythmias detection. In [9], a K-Nearest Neighbors algorithm is used to classify the ECG signals in the time domain, obtaining an accuracy of 80.9%. Additionally, the approaches above are focused only on binary classification, distinguishing only between healthy and arrhythmias, while our proposal tackles a wider problem by identifying between four specific heart conditions (i.e., heart failure, myocardial infarction, arrhythmia, and healthy). Apart from the fact that most studies focus on arrhythmias, there are also investigations dedicated to other diseases, such as in [11], where a system is proposed for detecting myocardial infarction. This model combines a convolutional neural network (CNN) with a long short-term memory model (LSTM), achieving an accuracy of 88.89%. However, in [25], a study is presented on a deep learning system that employs a classification CNN and a LSTM network to extract the spatial and temporal features of ECG signals.

This paper contributes by proposing a method for the automated diagnosis of heart diseases. The method uses the Fourier series and an optimal state observer (Kalman filter) to quickly obtain the harmonic content of ECG signals as a means of determining salient features of different heart pathologies, that a posteriori by employing a K-Nearest-Neighbors algorithm, the obtained harmonic content features are classified to determine the possible pathologies. Unlike other studies, our method not only determines whether a patient is healthy but also classifies three different heart diseases: arrhythmia, myocardial infarction, and heart failure. The accuracy of the proposed diagnosis methodology is greater than 93%.

The organization of the paper is as follows. Section 2 describes the modeling of the periodic signals using the Fourier series and the design of the optimal observer. Section 3 presents the optimal estimator as a means of determining the ECG harmonic content quickly, which will be used for classification purposes. Section 4 describes the classification method used for heart disease detection and presents the metrics used to evaluate the performance of the proposed system. Finally, Section 5 presents the conclusions and future work.

2 ECG Modeling through a Fourier Harmonic Description

The heart is an organ that depends on coordinated electrical activity for its proper functioning [13]. The electrocardiogram is a fundamental tool for recording heart activity and reflects the electrical events that occur during each cardiac cycle [6]. The heart is an organ that works as a pump and is responsible for pumping blood through the body. The heart muscles carry out this pumping, which can work correctly thanks to electrical impulses [13]. One of the main tasks of the heart is to contract to generate enough pressure in its chambers to properly distribute oxygenated blood throughout the body [20]. Its ability to function as a pump depends on the pumping movements of the atria, ventricles, atrioventricular, and semilunar valves, which direct blood flow within the heart. As mentioned above, heart movements occur every certain period, known as the cardiac cycle. In more concrete words, the cardiac cycle refers to the events that occur between the start of one heartbeat and the start of the next.

2.1 The Electrocardiogram

One way to know the health state of the heart is through the ECG, which is a record of its electrical activity. The ECG represents the impulses generated during each cardiac cycle, which are graphically recorded and possibly digitally stored. ECG signals are composed of several wave-shaped features that represent the different events of the cardiac cycle, as depicted in Fig. 1, and described as [6,8]: *P wave:* Represents the depolarization of the atria; *QRS complex:* It is the depolarization of the ventricles, which occurs just before ventricular repolarization; *T wave:* Corresponds to the repolarization of the ventricles, or in other words, the return to a resting state.

2.2 Fourier Harmonic Modeling

The Fourier transform and the Fourier series are widely used tools for analyzing biomedical signals. This proposal employs the latter. Fourier series allows

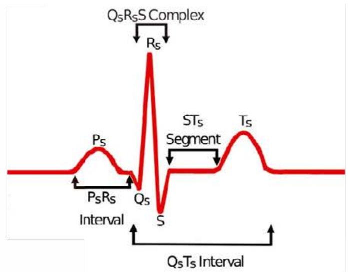

Fig. 1: ECG with their characteristics.

periodic functions/signals to be decomposed into an infinite sum of cosines and sines, which is mathematically described as

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{N} [a_n \cos(w_n\, t) + b_n \sin(w_n\, t)] \tag{1}$$

where $s(t)$ is a periodic signal, $\dfrac{a_0}{2}$ is the constant value (usually named DC component) of the signal, a_n, b_n are the Fourier coefficients, $N \to \infty$ is the number of harmonics (for finite values of N, the series only approximates the signal), the angular frequency $w_n = 2\, n\, \pi\, f$, with $f = \dfrac{1}{T}$ as the signal frequency of period T. For the purposes of modeling in this paper, it is convenient to present (1) as

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{N} A_n \sin(w_n\, t + \theta_n) \tag{2}$$

where $A_n = \sqrt{a_n^2 + b_n^2}$ is the amplitude of the sinusoidal function, and $\theta_n = \arctan\left(\dfrac{b_n}{a_n}\right)$ is the phase angle [15]. Hence, the harmonic content estimation of a signal consists in the determination of a_0, A_n, and θ_n in (2).

2.3 Periodic Signal Modeling in a State Space

Signal (2) can be presented in state space, such that a posterior a state observer can be designed with the aim of determining the supposed unknown values of the parameters a_0, A_n and θ_n.

To this end, consider the basic case by taking $a_0 = 0$, $N = 1$, the signal amplitude as Am, and employing the following trigonometric identity $\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$ for (2), one obtains[1]:

$$s(t) = Am\, \cos\theta\, \sin(w\, t) + Am\, \sin\, \theta \cos(w\, t). \tag{3}$$

[1] The n-subscript notation is omitted for easy of notation.

By defining the following state space variables as

$$x_1 \triangleq \sin(w\,t); \qquad x_2 \triangleq \cos(w\,t) \qquad \text{and} \qquad y = s(t)$$

the corresponding dynamical description in state space for (3) becomes

$$\begin{aligned} \dot{x}_1 &= w\,x_2 \\ \dot{x}_2 &= -w\,x_1 \\ y &= Am\,\cos\theta\,x_1 + Am\sin\theta\,x_2. \end{aligned} \tag{4}$$

Additionally, for the case when the signal $s(t)$ contains a constant component, given in (2) as $\frac{a_0}{2}$, then the following state variable can be added to (4) as

$$\dot{x}_0 = 0 \tag{5}$$

with initial condition $x(0) = a_0/2$. Then, the complete signal model described by (4)–(5) can be presented in a matrix form as

$$\dot{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & w \\ 0 & -w & 0 \end{bmatrix} x \tag{6}$$

with $x = [x_0\ x_1\ x_2]^T$ and corresponding output as

$$y = \begin{bmatrix} 1 & Am\,\cos\theta & Am\,\sin\theta \end{bmatrix} x. \tag{7}$$

Notice the previous analysis is for one harmonic, i.e. $N = 1$, where the definition of two state variables is needed. Now, the previous modeling procedure can be easily extended to finite values of N, and additionally, from practical considerations, signal $s(t)$ is assumed to be affected by noise, resulting in a general system as

$$\dot{x} = A\,x + v_1 \tag{8}$$

and output

$$y = C\,x + v_2 \tag{9}$$

where the space state $x \in \mathbb{R}^{2N}$ is defined as

$$x = \begin{bmatrix} x_0\ x_1\ x_2 \cdots x_{2N-1}\ x_{2N} \end{bmatrix}^T$$

and matrices

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & w & \cdots & 0 & 0 \\ 0 & -w & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & w_N \\ 0 & 0 & 0 & \cdots & -w_N & 0 \end{bmatrix}$$

and

$$C = \begin{bmatrix} 1 & A_1 \cos\theta_1 & A_1 \sin\theta_1 & \cdots & A_N \cos\theta_N & A_N \sin\theta_N \end{bmatrix}$$

while $v_1 \in \mathbb{R}^{2N}$ and $v_2 \in \mathbb{R}$ are the process noise and measurement noise, respectively. Noises v_1 and v_2 are considered Gaussian and zero-mean, with covariance given by $E\left\{v_1(t)\, v_1^T(\tau)\right\} = Q_{v_1}\,\delta(t-\tau)$ and $E\left\{v_2(t)\, v_2^T(\tau)\right\} = R_{v_2}\,\delta(t-\tau)$, respectively, with matrices Q_{v_1} and R_{v_2} nonnegative. Model (8)–(9) will serve for the design of an optimal state estimator (known as Kalman filter), as follows.

2.4 Optimal Observer Design

State observers are algorithms used in dynamical systems to estimate variables (when possible) that are unmeasured or difficult to measure, which is achieved through the available measurements. Hence, considering system (8)–(9), the optimal state observer, based on the Kalman filter [2], is described as

$$\begin{aligned} \dot{\hat{x}} &= A\,\hat{x} + L\,(y - \hat{y}) \\ \hat{y} &= \overline{C}\,\hat{x} \end{aligned} \tag{10}$$

where the space state $\hat{x} \in \mathbb{R}^{2N}$ is defined as

$$\hat{x} = \begin{bmatrix} \hat{x}_0 & \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_{2N-1} & \hat{x}_{2N} \end{bmatrix}^T$$

and $L = P\,\overline{C}^T R_{v_2}^{-1}$ represents the optimal gain of the observer, and P is the solution to the Riccati equation, given as

$$\dot{P} = Q_{v_1} - P\,\overline{C}^T R_{v_2}^{-1}\,\overline{C}\,P + A\,P + P\,A^T. \tag{11}$$

and

$$\overline{C} = \begin{bmatrix} 1 & k_1 & k_2 & \cdots & k_{2N-1} & k_{2N} \end{bmatrix}^T$$

with $k_1, k_2, \ldots, k_{2N}$ as design parameters, which are defined by the user.

Remark 1 *For the observer design, usually are used the matrices given in* (8)–(9)*; nonetheless, matrix C in* (9) *is not available since the values for A_n and θ_n are unknown; indeed, those are the parameters to be determined. Hence, as a part of the observer design, it is required to specify the values of the parameters in $\overline{C}$, which will be used to calculate the values for matrix C.*

Based on the analysis of (3), the harmonic estimation for Am and θ is given by considering that $\hat{A}m\,\cos\hat{\theta}\,\sin(w\,t) + \hat{A}m\,\sin\hat{\theta}\cos(w\,t) = k_1\,\hat{x}_1 + k_2\,\hat{x}_2$ and by taking its time derivative, then a set of two algebraic equations and two variables is formed, whose solution becomes

$$\hat{A}m = \sqrt{(k_1^2 + k_2^2)(\hat{x}_1^2 + \hat{x}_2^2)}$$

and

$$\hat{\theta} = -w\,t + \arctan\left(\frac{k_1\,\hat{x}_1 + k_2\,\hat{x}_2}{k_1\,\hat{x}_2 - k_2\,\hat{x}_1}\right).$$

For this research, the parameter of interest is only the amplitude of each n-th harmonic, i.e. A_n. Hence, in a general way, the amplitude estimation of each harmonic in a signal is calculated as

$$\hat{A}_N = \sqrt{\left(k_{2N-1}^2 + k_{2N}^2\right)\left(\hat{x}_{2N-1}^2 + \hat{x}_{2N}^2\right)}. \tag{12}$$

As an example, numerical values are used in a digital simulation for signal (2), with $N = 1$, $a_0/2 = 0.5$, $A_1 = 0.2$, $w = 0.9$, $\theta_1 = 30\,deg$, $k_1 = 6$ and $k_2 = 0.01$. Then, the estimation of the signal $s(t)$ through $\hat{y}$, and the harmonic content amplitude by means of (12), are depicted in Fig. 2 and Fig. 3, respectively. As can be noted, the variables of the observer converge after a few seconds to the actual values, demonstrating the capability in the harmonic content estimation of the optimal observer. Once the convergence is achieved, the signal can be used to extract harmonic remarkable information over time, which is a difference with respect to the basic Fourier series, where a complete period of the signal is required to determine the harmonic content.

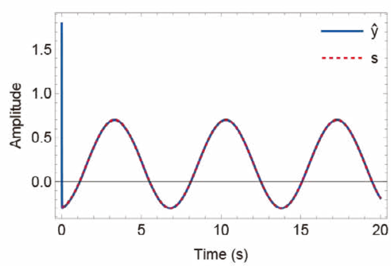

Fig. 2: $\hat{y}$ vs s

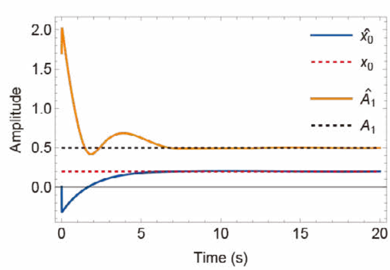

Fig. 3: Estimated components vs real components

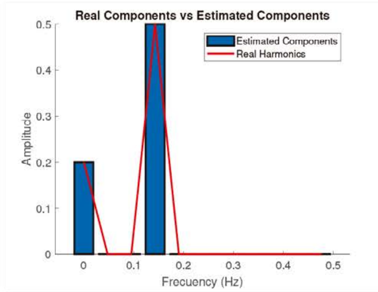

Fig. 4: Estimated components vs real components (Fourier Spectrum)

3 Harmonic Content Determination of ECG

This section presents the usage of the optimal observer to estimate the harmonic content of the ECG signals, and based on the content, its salient characteristics will be used as patterns to classify heart pathologies. Notice that this procedure is performed over time; that is, the harmonic content is fastly determined and, in the same way, immediately classified to obtain the heart diagnosis. Additionally, and only for comparison purposes, the fast Fourier transform (FFT) is used to validate the estimation of the harmonic content. It is worth remarking that the coefficient in the Fourier series requires at least a period of time to be calculated; for its part, using a state estimator allows the Fourier coefficients to be instantaneously adapted along the signal acquired over time, which is known as an online process.

3.1 Diagnosis General Scheme

Fig. 5 shows the general scheme of the process that is carried out, starting with the acquisition of the ECG signal, which is preprocessed by normalizing the data to an appropriate range and separating the signal into individual cycles to determine its current frequency. Then, the harmonic content is obtained through an optimal observer. Finally, the harmonic content is classified by a KNN algorithm, which determines the heart diagnosis.

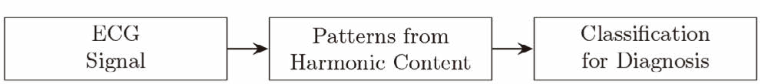

Fig. 5: Process flow diagram

3.2 Simulation results

Fig. 6 displays the estimated harmonics (blue bars), which are compared with the FFT (red line), showing the correct harmonic estimation. Fig. 7 presents the estimation error, which is small, giving confidence in the proposed fast procedure to be used by the classifier. On the other hand, Fig. 8 and Fig. 9 show the convergence over time of the optimal state observer toward the references provided by the FFT. In Fig. 8, the reference is the DC component, and in Fig. 9, the reference corresponds to the first harmonic of the Fourier spectrum. The rest of the harmonics present a similar behavior. Finally, the harmonic content is processed by the classifier as described as follows.

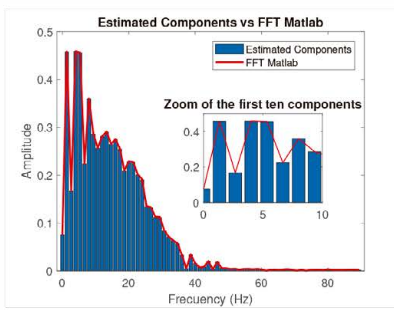

Fig. 6: FFT Matlab vs Estimated Harmonics

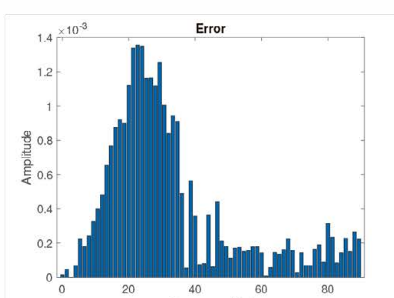

Fig. 7: Error

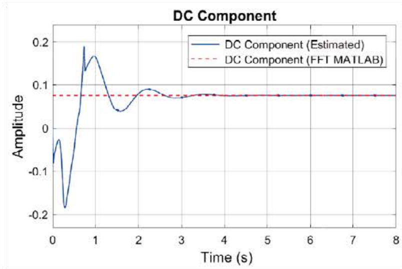

Fig. 8: DC Component

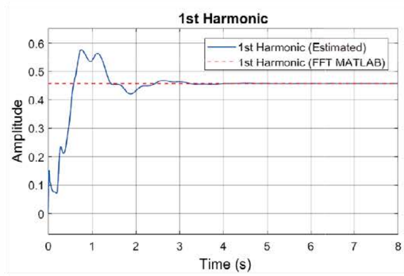

Fig. 9: 1st Component

4 ECG Classification and Diseases Detection

ECG classification is necessary to determine different heart diseases. The number of classes to classify depends on the number of available classes on the training dataset. In this research, including the healthy condition, three cardiac pathologies are also considered: arrhythmia, heart failure, and myocardial infarction. The pathology classification will be performed by identifying patterns in the harmonic content of the ECG by using a machine learning algorithm. The patient databases are obtained from the PhysioNet website [1]. For this research, a total of 105 ECG signals were used, including 24 from healthy patients, 29 with arrhythmia, 42 with myocardial infarction, and 10 with heart failure. Each signal contains 10 cardiac cycles, resulting in a total of 1050 cardiac cycles analyzed and used to train the KNN classifier.

There are various techniques for ECG classification, such as Traditional ECG interpretation by experts, where doctors visually analyze an ECG to diagnose heart problems, mainly when visible issues can be identified in the ECG, such as changes in the duration of the P_s and T_s waves or changes in the amplitude of the QRS complex. Another ECG classification method is the Classification by Machine Learning. Currently, there are many types of classifiers. Examples that have been used for ECG analysis through different approaches are artificial neural networks, such as those exposed in [3], [10], and [12]. In [24], authors use the MKELM-based binary random forest classifier (MKELM-RF), while [16] uses a Support Vector Machine. Those classifiers are trained using data sets such that when exposed to new data, this new information can be classified.

4.1 K-Nearest Neighbors

This research uses the KNN algorithm to classify ECG signals according to their harmonic characteristics. Although KNN has been applied in other studies, it is implemented differently since multiple classes are used in this work. KNN is a supervised machine learning classifier, meaning that it is trained using a labeled dataset [21]. In our case, the database did not originally contain labeled classes, so the data are manually labeled to make it suitable for a classification task. The algorithm follows the following steps to classify a new data point whose class is unknown: *(a)* distance computation: the distance between the new data point and all training data points is reckoned using a distance function. In MATLAB, Euclidean, Chebychev, Correlation, Cosine, and Mahalanobis are available. Generally, Euclidean distance is used, as done in this research; *(b)* identify the nearest neighbors: the K closest data points to the new data point are selected; *(c)* classify: the class of new data is computed as a majority vote decision, selecting the most common class (i.e., the mode) between the K nearest neighbors.

Fig. 10 illustrates the classification process, in which the KNN algorithm is observed using a new data point and $K = 3$; then, in this case, the 3 closest data are considered, and the majority vote results in class A because there are 2 data of class A and 1 data of class B. Now, for $K = 12$, the majority vote considers class B the winner because there are 7 data from class B and 5 from class A, and the new data will be classified as class B. An alternative to classifying is by a weighted vote using the distance between the data query and their neighbors. In this case, the closest points will have greater weight in the decision, while the farthest ones have less influence.

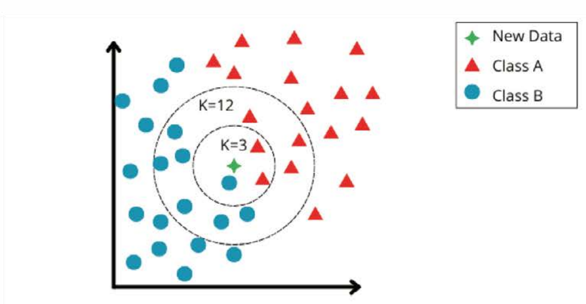

Fig. 10: K-Nearest Neighbor (KNN).

4.2 Metrics

The tools (metrics) used to evaluate the classifier performance and their values are shown in Table 1. *The Accuracy* expresses how well the classifier works in terms of the ratio of correct predictions out to the total number of predictions. [19]. *The Precision* measures how many of the predicted classes for a specify class are actually correct. [19]. *Recall* calculates the number of actual instances of a specific class that were correctly predicted [19]. *The F1 score* is a metric that combines precision and recall in a single value. This metric is particularly useful when working with datasets with imbalanced classes [7]. The area under the curve (AUC) is a metric used to evaluate the performance of a classification model, specifically regarding its ability to discriminate between classes. The AUC varies between 0.5 and 1, where a value closer to 1 means a better classifier, a value of 1 is a perfect classifier, and a value closer to 0.5 is equivalent to a very bad classifier; this means that the classifier is working randomly. The AUC of a classifier represents the probability that the classifier ranks a randomly chosen positive instance higher than a randomly chosen negative instance [4]. In other words, a higher AUC value means that the classifier has a greater probability of correctly predicting that a data item belongs to the specified class.

Table 1: Metrics

Metrics	Value
Accuracy	91.8831%
Precision	90.7487%
F1 Score	92.0694%
F1 Score (Micro)	91.8831%
AUC	0.9336

In this research, the KNN classifier was implemented in MATLAB, and it was configured with two closest neighbors ($K = 2$).

4.3 Confidence

Confidence in classification models evaluates how certain the model is about its predictions. In this research, the receiver operating characteristic (ROC) curve will be used to evaluate the performance of the KNN classifier. The ROC curve is a visual representation of the model performance, which is the relationship between the false positive rate and the true positive rate [5]. The false positive rate is the data that is false but classified as true, and the true positive rate is the true data that is classified as true [4]. For the KNN classifier, the ROC curves shown in Fig. 11 are generated including four classes: Class 0 (Healthy person), class 1 (Arrhythmia), class 2 (Myocardial Infarction) and class 3 (Heart failure). These curves are interpreted by the AUC, whose results are presented in Table 1.

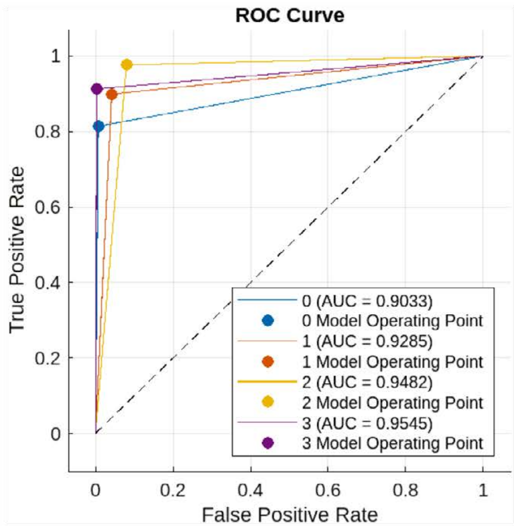

Fig. 11: Curve ROC.

4.4 Results Discussion

The proposal of this research requires high computational effort; however, modern computer equipment is able to perform the required algorithms. Although the system works correctly with the available data, work is being done to have a larger database to improve the results. These limitations do not affect the performance of the entire system because it has a high performance that can be seen in the metrics presented in Table 1. Also, from Fig. 11 it can be seen that the classifier used in this research has a better performance for class 3 (Heart failure) with a 95.45% probability of being correct for that pathology and the lowest but not bad performance for class 0 (Healthy person) with 90.33%.
Finally, the system was tested for the four categories: heart failure, myocardial infarction, arrhythmia and healthy people, using new ECG test signals that were not included in the training of the classifier. Each ECG consisted of 10 cardiac cycles and predictions were made cycle by cycle to finally assign a final pathology to each patient based on the category with the highest number of predictions. The effectiveness of the research proposal is summarized as follows. The results presented in Table 2 show a performance greater than 93%, highlighting the effectiveness in correctly classifying ECG signals in the four categories: heart failure, myocardial infarction, arrhythmia, and healthy individuals. The metrics obtained emphasize the ability of the system to provide accurate and reliable classifications. Furthermore, an important characteristic of this work is that each cardiac cycle of the ECG is analyzed individually, ensuring more precise classifications and improving the overall diagnostic accuracy.

Table 2: Results

Condition	Patients	Effectiveness (%)
Heart Failure	3	93.3
Myocardial Infarction	11	93.7
Arrhythmia	9	93.6
Healthy	5	93.1

5 Conclusions and Future Work

This research develops a methodology based on ECG signals capable of identifying heart diseases. The obtained results show high confidence levels and elevated effectiveness in classifying three pathologies: arrhythmia, myocardial infarction, and heart failure, as well as the condition of a healthy person. In this sense, this contribution demonstrated its effectiveness and to be a functional method for the automatic detection of cardiac pathologies. For future studies, there is the possibility of implementing other algorithms for signal classification, such as support vector machines, artificial neural networks, etc. In addition, there is the possibility of expanding the database with new data on the pathologies studied in this research by adding new pathologies. Also, it is expected to develop an experimental prototype to perform tests on real patients.

References

1. Physionet, https://physionet.org/
2. Brian D. O. Anderson, John B. Moore: Optimal Control: Linear Quadratic Methods. Prentice-Hall, Englewood Cliffs, NJ, USA (1990)
3. Dang, H., Sun, M., Zhang, G., Qi, X., Zhou, X., Chang, Q.: Deep arrhythmia-diagnosis network for atrial fibrillation classification using electrocardiogram signals. IEEE Access (2019)
4. Fawcett, T.: Introduction to roc analysis. Pattern Recognition Letters **27**, 861–874 (06 2006). https://doi.org/10.1016/j.patrec.2005.10.010
5. G., S., Vardhan, V.: Multi-class classification using mixtures of univariate and multivariate roc curves. Journal of Biostatistics and Epidemiology **8** (04 2022). https://doi.org/10.18502/jbe.v8i2.10418
6. Guyton, Hall: Textbook of Medical Physiology. Saunders/Elsevier, twelfth edn. (2011)
7. Han, J., Kamber, M., Pei, J.: Data mining concepts and techniques. Morgan Kaufmann Publishers, Waltham, Mass., 3rd edn. (2012)
8. Ismail, S.N.M.S., Aziz, N.A.A., Ibrahim, S.Z., Nawawi, S.W., Alelyani, S., Mohana, M., Chun, L.C.: Evaluation of electrocardiogram: numerical vs. image data for emotion recognition system. F1000Research **10**(1114), 1114 (2021)
9. Losada, D., Gómez, J., Vela, J.: Classification of ecg signals using machine learning techniques. Academic Journal of Interdisciplinary Studies **13**, 92 (05 2024). https://doi.org/10.36941/ajis-2024-0067
10. Lu, Y., Jiang, M., Wei, L., Zhang, J., Wang, Z., Wei, B., Xia, L.: Automated arrhythmia classification using depthwise separable convolutional neural network

with focal loss. ScienceDirect - Biomedical Signal Processing and Control 69, 2021 (2021)

11. Manimekalai, K., Kavitha, D.: Deep learning methods in classification of myocardial infarction by employing ecg signals. Indian Journal of Science and Technology **13**(28), 2823–2832 (2020). https://doi.org/10.17485/IJST/v13i28.445
12. Martínez, J.Q.: Análisis De Electrocardiogramas Mediante Técnicas De Ingeniería Para La Detección De Enfermedades Cardiacas. Master's thesis, Universidad Pontificia De Comillas (2019)
13. Newman, T., article reviewed by Dr. Payal Kohli M.D. FACC: El corazón: Anatomía, cómo funciona y más. MedicalNewsToday (2021)
14. Oh, S.L., Ng, E., Tan, R.S., Acharya, U.: Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats. Computers in Biology and Medicine **102** (06 2018). https://doi.org/10.1016/j.compbiomed.2018.06.002
15. Ramos-Paz, S., Ornelas-Tellez, F., Rico-Melgoza, J.J.: Dynamic harmonics–interharmonics identification and compensation through optimal control of a power conditioning application. Electrical Engineering (2022)
16. Saini, S.K., Gupta, R.: Artificial intelligence methods for analysis of electrocardiogram signals for cardiac abnormalities: state-of-the-art and future challenges. Springer (2021)
17. Secretaría de Salud: Cada año, 220 mil personas fallecen debido a enfermedades del corazón, https://www.gob.mx/salud/prensa/490-cada-ano-220-mil-personas-fallecen-debido-a-enfermedades-del-corazon
18. Secretaría de Salud: Hasta 80% de decesos por enfermedades cardiovasculares son prevenibles: Hospital general de méxico, https://www.gob.mx/salud/prensa/319-hasta-80-de-decesos-por-enfermedades-cardiovasculares-son-prevenibles-hospital-general-de-mexico
19. Sokolova, M., Japkowicz, N., Szpakowicz, S.: Beyond accuracy, f-score and roc: A family of discriminant measures for performance evaluation. vol. Vol. 4304, pp. 1015–1021 (01 2006). https://doi.org/10.1007/11941439_114
20. Tamargo, J., Delpón, E.: La función de bomba del corazón. McGraw-Hill Education, New York, NY (2016), accessmedicina.mhmedical.com/content.aspx?aid=1132161556
21. Uddin, S., Ibtisham Haque, Haohui Lu, M.A.M., Gide, E.: Comparative performance analysis of k-nearest neighbour (knn) algorithm and its different variants for disease prediction. Scientific Reports (2022)
22. Uribe, W., Duque, M., Arango, E.: Electrocardiografía y arritmias. Revista Iberoamericana de Arritmología (01 2010). https://doi.org/10.5031/v1i2.RIA1012
23. World Heart Federation: World heart report 2023: Confronting the world's number one killer (2023), https://world-heart-federation.org/wp-content/uploads/World-Heart-Report-2023.pdf
24. Yang, P., Wang, D., Zhao, W.B., Fu, L.H., Du, J.L., Su, H.: Ensemble of kernel extreme learning machine based random forest classifiers for automatic heartbeat classification. ScienceDirect - Biomedical Signal Processing and Control 63, 2021 (2021)
25. Zhang, C.J., Yuan-Lu, et al., F.T.: Heart failure classification using deep learning to extract spatiotemporal features from ecg. BMC Medical Informatics and Decision Making **24**, 17 (2024). https://doi.org/10.1186/s12911-024-02415-4