European Journal of Engineering Science and Technology
ISSN 2538-9181

Classification of Cardiac Arrhythmias Based on Electrocardiogram Data Using a Convolutional Neural Network Model

Rodrigo Alexandre Dos Santos

Department of Software Development, CPQD Foundation, Campinas, SP, Brazil

ARTICLE INFO

Keywords:
Cardiac Arrhythmias,
Electrocardiogram,
Machine Learning,
Deep Learning

ABSTRACT

Cardiac arrhythmias are a disease with considerable incidence and prevalence worldwide, and their diagnosis can be complex due to the existence of different types of arrhythmias that share similar characteristics and make an accurate diagnosis difficult. Making a correct diagnosis of the kind of arrhythmia that affects an individual is important to define the most appropriate type of treatment for the case. Machine Learning and Deep Learning techniques have been proposed to automate the diagnosis of arrhythmias to assist healthcare professionals in decision-making. This study proposes a Convolutional Neural Network model for classifying cardiac arrhythmias using electrocardiogram data. The objective is to present a model that achieves high accuracy rates in identifying types of arrhythmias and presents an adequate balance between performance and computational costs. The model was trained with a dataset composed of electrocardiogram exams with 32 types of arrhythmias. In the pre-processing phase, the dataset was restructured to allow the data to be treated as a time series to explore the potential of Convolutional Neural Networks in dealing with data organized in this way. Training was carried out using a state-of-the-art Deep Learning model and the model achieved an accuracy rate of 98.37% in its predictions. This excellent performance confirms the ability of Convolutional Neural Networks to efficiently deal with pattern learning in time series. The results obtained demonstrate the potential of Deep Learning techniques as aiding tools to provide improvements in medical processes.

1. Introduction

Cardiac arrhythmias are a disease that generates considerable concern and requires careful diagnosis because they are associated with several risk factors. Some of the associated risk factors are high blood pressure, high cholesterol, diabetes, obesity, hyperthyroidism, hypothyroidism, electrolyte disorders, structural heart disease, smoking, alcoholism, anemia, atherosclerosis, emotional stress, and genetic predisposition. Arrhythmias are classified according to their level of criticality as benign or malignant. Arrhythmias are considered benign when they generate unpleasant symptoms but do not pose a risk of death, and they are

* Corresponding Author's E-Mail Address: rodrigoasantos1981@gmail.com

Cite this article as:
Dos Santos, R. A. (2024). Classification of Cardiac Arrhythmias Based on Electrocardiogram Data Using a Convolutional Neural Network Model. *European Journal of Engineering Science and Technology, 7*(2): 14-30. https://doi.org/10.33422/ejest.v7i2.1354

considered malignant when they compromise cardiac function and can even cause sudden death due to a fulminant heart attack.

Arrhythmias are a disease with considerable incidence and prevalence worldwide. According to Li et al. (2022), a study that analyzed and calculated epidemiological data related to cardiac arrhythmias in 204 countries reported an incidence of around 4.72 million cases and a prevalence of around 59.7 million cases in 2019. The results showed in the study reinforce the importance of investments in the development of strategies and technologies that can assist doctors and healthcare professionals in preventing and treating this disease.

There are different types of arrhythmias and there is usually a treatment modality best suited for each type. According to Kingma et al. (2023), there are treatments based on antiarrhythmic medications (beta blockers and calcium channel blockers, for example), implantable electrical devices (pacemakers and defibrillators, for example), medical procedures (cardioversion and ablation, for example) and surgeries. According to Cruickshank (2008), an accurate diagnosis of the type of arrhythmia that affects the individual is essential to assess the severity of the disorder and the associated risks and define the type of treatment most suitable for the case.

This study aims to propose a Deep Learning model for classifying cardiac arrhythmias through the analysis of data from electrocardiograms. The model must be able to classify electrocardiogram exams accurately considering different types of arrhythmias. The results obtained in the study are expected to demonstrate the potential of using Deep Learning techniques to assist healthcare professionals in diagnosing cardiac arrhythmias. A model that can accurately classify different types of arrhythmias provides healthcare professionals with a very useful tool to improve the accuracy of their diagnoses, mitigate the occurrence of errors, and provide cost reduction by eliminating the need for additional tests to be carried out to define the specific treatment to be given to each patient. This study also seeks to contribute to improving the clinical processes involved and the exchange of information between different professionals by encouraging the use of standardized ways of using Artificial Intelligence techniques to make diagnoses and conduct treatments in clinical practice.

2. Literature Review

The cardiac cycle is the set of events that occur in the valves and chambers of the heart between the beginning of a heartbeat and the beginning of the next beat. These events are responsible for pumping blood to the lungs, where the blood oxygenation process occurs, and pumping oxygenated blood to the aorta artery, which has the function of ensuring that blood is transported to the different parts of the organism with the ultimate objective of nourishing cells and to provide the energy necessary for the proper functioning of all organs. The cardiac cycle is carried out through movements of contraction (systole) and relaxation (diastole) of the myocardium. These movements are performed regularly and synchronized to ensure the correct flow of blood throughout the cardiovascular system.

Cardiac arrhythmias are changes in the rhythm of the heartbeat caused by disorders or diseases that affect the functioning of the heart. The events of the cardiac cycle occur irregularly or inappropriately when the body suffers an arrhythmia. These changes create an imbalance in the body that can range from impaired pumping of oxygenated blood to other organs in the body to sudden death in more extreme cases. According to Kingma et al. (2023), most arrhythmias occur as a result of structural abnormalities in the myocardium, but arrhythmias can also occur due to risk factors derived from genetic or environmental conditions.

Figure 1 shows a representation of the basic anatomy and electrical conduction system of the heart. The heart is divided into 4 chambers: 2 atria (right and left) and 2 ventricles (right and left). Due to their location, the atria are called upper chambers, and the ventricles are called lower chambers. Oxygen-poor blood reaches the right atrium through the vena cava and is drained into the right ventricle to be transported to the lungs through the pulmonary valve and pulmonary artery. Oxygenated blood reaches the left atrium through the pulmonary veins and is drained into the left ventricle to be transported throughout the body through the aortic valve and the aorta artery. The contraction and relaxation movements of the chambers, as well as the opening and closing events of the valves, are carried out in an orderly and synchronized manner, and this process ensures that the blood has the correct destination at the correct time.

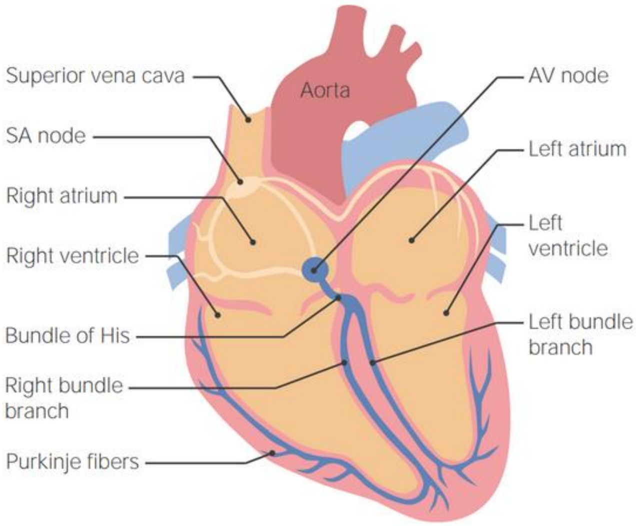

Figure 1. Basic anatomy and conduction system of the heart

The conduction system has a set of specialized cells responsible for generating and conducting electrical impulses that coordinate the systole and diastole movements of the heart chambers. Electrical impulses propagate across the membrane of each cell due to action potentials that allow ions to flow through membrane channels. According to Williams (2005), the generation of action potentials occurs through a set of complex interactions that occur inside and outside cells involving sodium, potassium, and calcium ions. Action potentials are divided into 2 main processes: depolarization and repolarization. According to Chakrabarti and Stuart (2005), initially, the cell is in a state called polarization, also called resting phase, which comprises the period between two subsequent action potentials and in which the membrane potential (difference between the intracellular environment and the extracellular environment) has a negative charge of about -90 mV. The depolarization process occurs when a polarized cell receives a stimulus and undergoes a change in its membrane potential to around +30mV caused by the entry of positively charged sodium ions into the cell. The depolarization process causes muscle contraction. The repolarization process occurs when a depolarized cell undergoes a change in its membrane potential to around -90 mV caused by the exit of potassium ions to the outside of the cell. The repolarization process causes muscle relaxation and prepares the cell for the onset of the next action potential.

The conduction system is formed by 4 main structures: sinoatrial node (SA node), atrioventricular node (AV node), bundle of His and the branches of the bundle (right and left), and Purkinje fibers. The sinoatrial node is located at the junction of the superior vena cava and the right atrium. According to Williams (2005), the sinoatrial node is composed of a set of cells specialized in generating electrical impulses automatically and rhythmically, and acts as the system's main pacemaker, controlling the frequency of the heartbeat. The electrical impulses generated by the sinoatrial node travel through the atrium until they reach the atrioventricular node. The atrioventricular node is located at the base of the right atrium. According to Williams (2005), the atrioventricular node retransmits the electrical current received to the bundle of His with a small delay that is necessary to ensure that ventricular systole occurs after atrial systole so that the transport of blood to the ventricles is carried out at the correct time. According to Chakrabarti and Stuart (2005), the atrioventricular node can also act as a secondary pacemaker of the system in situations in which the sinoatrial node presents some dysfunction that prevents it from performing its function as expected. The bundle of His and the branches of the bundle are located in the ventricular walls. This structure is formed by cells that conduct rapid electrical impulses and have the function of transmitting action potentials from the atria to the ventricles. Purkinje fibers are the endings of the branches of the bundle of His and spread throughout the entire length of the ventricles to enable synchronized contraction of the ventricular myocardium.

Any abnormality in the functioning of any of the structures of the conduction system can cause arrhythmias and compromise the functioning of the system as a whole. According to Chakrabarti and Stuart (2005), cardiac arrhythmias can arise as a result of abnormalities in the generation or conduction of electrical impulses and can be classified according to the abnormal speed of the heart rate (bradycardia or tachycardia) and according to the location of origin of the disorder (atrium, atrioventricular junction or ventricle).

The usual heart rate of a person at rest ranges from 60 to 100 beats per minute (bpm). Bradycardia occurs when the heart rate is less than 60 bpm. According to Chakrabarti and Stuart (2005), bradycardia can occur when the sinoatrial node cannot generate an action potential fast enough to meet the demands of the conduction system, or when there is some blockage in the atrioventricular node or in the structure of the His-Purkinje system that compromises the propagation of the action potential. According to Williams (2005), bradycardias can also be caused by excessive activation of the vagal nerve, which causes the expansion of blood vessels and a reduction in the return of blood to the heart.

Tachycardia occurs when the resting heart rate is greater than 100 bpm. According to Chakrabarti and Stuart (2005), tachycardia can be caused by 3 mechanisms: reentry, enhanced automaticity, or triggered activity. Reentry is a disorder in which an electrical impulse propagates through a closed circuit and follows a retrograde path, causing the re-excitation of cells that are already in another phase of the cardiac cycle. Enhanced automaticity is a disorder in which electrical impulses are generated in an accelerated manner by the pacemaker as a result of physiological or pathological causes. According to Chakrabarti and Stuart (2005), the activated activity combines characteristics of both reentry and enhanced automaticity and can cause the generation of extrasystoles by allowing depolarization events to occur during or immediately after repolarization events.

The classification of arrhythmias considering both the speed of the heart rhythms and their place of origin can generate several possible combinations and, according to Liu et al. (2022), there may be imprecise or redundant terms in the classifications found in the literature or clinical practice. The American Heart Association (AHA) has organized a statement that presents a list of terms for diagnosing arrhythmias. The statement was published by Mason et al. (2007) and has the main objective of providing a concise list of standardized terms to improve diagnosis accuracy. The recommendation proposed by the AHA contains 117 main classifications and is widely accepted worldwide.

The electrocardiogram (ECG) is a widely used tool for diagnosing cardiac arrhythmias. It is a device that uses electrodes attached to the surface of the skin to detect and record the electrical activity generated by the heart. ECG voltage is measured in microvolts (µV) and represents the systole and diastole movements of the myocardium. According to iMotions (2024), the fact that it is a non-invasive, low-cost, and high-resolution tool favors the use of ECG in physiological examinations. According to Rafie et al. (2021), the ECG enables the diagnosis of various cardiovascular diseases and the cost-benefit ratio of its use is evident when compared to other advanced modalities of imaging exams or invasive procedures.

The 12-lead system is the electrode placement configuration considered the universal standard. This system uses 10 electrodes, 6 of which are attached to the chest and 4 are attached to the limbs. Figures 2 and 3 show a representation of the positioning of the electrodes and the configuration of the system. The system is called 12-Lead because the positioning of the 10 electrodes allows 12 leads to be generated that correspond to 12 points of view of the myocardial rhythm. Leads are divided into 2 groups: precordial leads and limb leads. Precordial leads (or chest leads) are obtained through electrodes attached to the chest and are designated by the acronyms V1, V2, V3, V4, V5, and V6. Limb leads are obtained through the combination of 2 or 3 electrodes attached to the limbs and are designated by the acronyms I, II, III, aVR, aVL, and aVF.

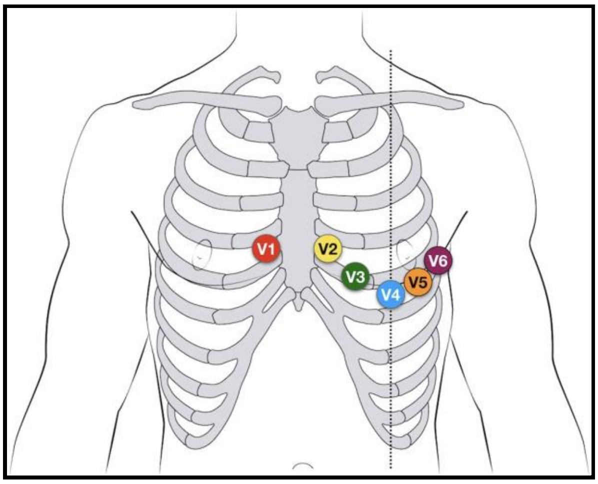

Figure 2. Precordial leads

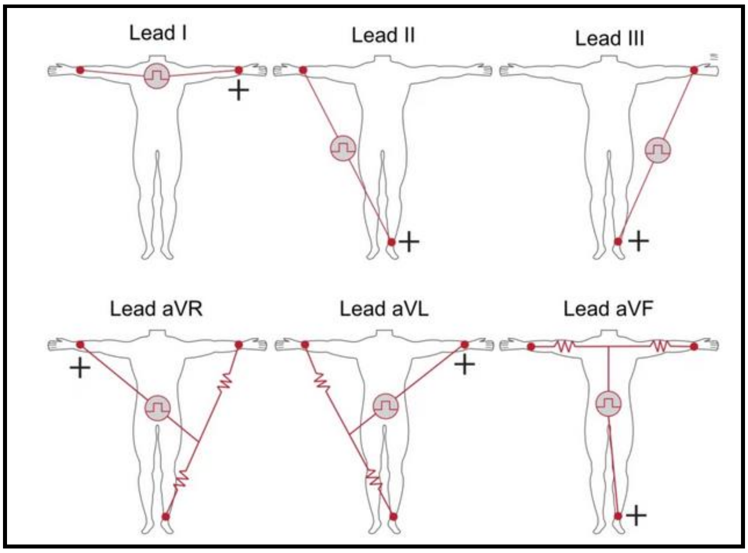

Figure 3. Limb leads

The ECG records the phases of the cardiac cycle and presents the result as a trace in which the vertical axis shows the recorded voltage and the horizontal axis shows the temporal sequence of occurrence of action potentials. Figure 4 shows an example of an ECG in which the trace of the 12 recorded leads can be viewed. According to Becker (2006), the ECG can record events generated by muscle cells (atrium and ventricle), but cannot record events generated by specialized cells (sinoatrial node, for example) because these cells generate very low voltages. Although the activity of specialized cells is not recorded, it is possible to deduce these events through the interpretation of trace information. The phases are represented by the ECG through waves, as shown in the diagram in Figure 5. The phases recorded are P wave (depolarization of the atrial muscle), QRS complex (depolarization of the ventricular muscle), and T wave (repolarization of the ventricular muscle and return to baseline). According to Becker (2006), the ECG of healthy hearts shows a sequence of waves that follows a certain order and regularity, and the ECG of arrhythmic hearts may present an absence of waves, inconsistent intervals between waves, extra waves, or waves with altered morphology.

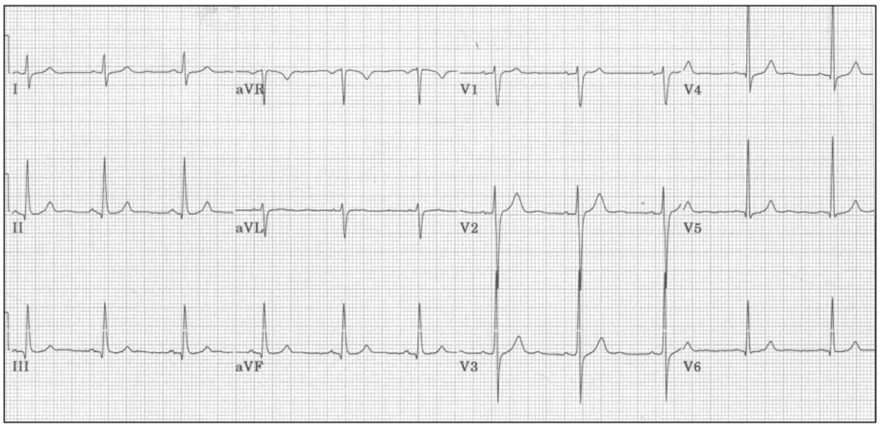

Figure 4. Example of an ECG exam

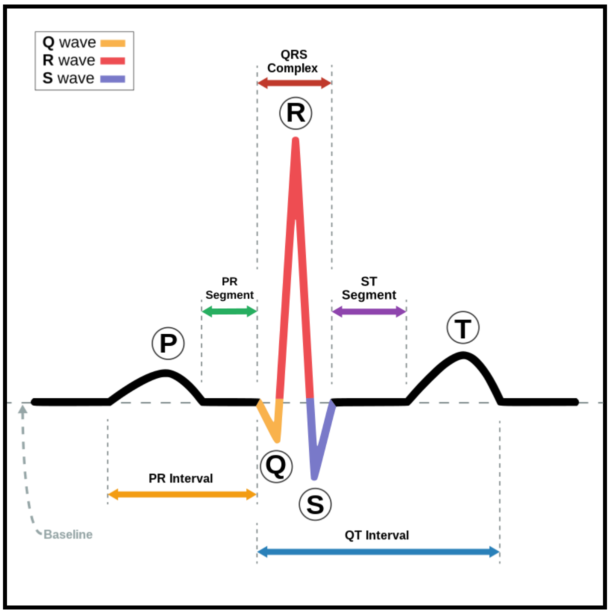

Figure 5. The ECG waves

According to Rafie et al. (2021), the use of Machine Learning techniques for analyzing ECG exams has proven to be a promising alternative for improving clinical processes by enabling the reduction of time and costs required to carry out the analyses and providing healthcare professionals with an auxiliary tool to increase the accuracy of diagnoses. Several studies have shown that health professionals can benefit from tools that help improve the accuracy of diagnoses to provide greater confidence in making decisions about the treatments to be adopted.

According to Liu et al. (2022) and Zheng et al. (2020), the development and dissemination of wearable devices in recent years is another factor responsible for the growth of interest in the application of Machine Learning techniques for automatic ECG interpretation. Mobile devices such as smartwatches or smart vests can be used as cardiac monitoring tools to enable healthcare services to be extended beyond clinics and laboratories.

Several studies have proposed Machine Learning and Deep Learning models for classifying cardiac arrhythmias. Studies cited by Singh et al. (2023) and Nagarajan et al. (2021) presented models that proved to be efficient tools for diagnosing arrhythmias by achieving high accuracy rates in their predictions. According to Harmon et al. (2023), despite the advances achieved in this field of research, there is still a need for improvements in technologies and processes to attenuate limitations that may hinder the adoption of technologies based on Artificial Intelligence in medical practice. An issue that has been mentioned as an important factor in facilitating the adoption of Machine Learning and Deep Learning models in medical practice is the improvement of the interpretability of model predictions, that is, the ability to explain to healthcare professionals what rules or information were considered most important by the model to reach the conclusions presented.

3. Methods

3.1. Dataset Description

The dataset used in this study was derived from the work developed by Liu et al. (2022). The dataset contains 25,770 ECG records obtained from 24,666 individuals who were examined at Shandong Provincial Hospital (Jinan, China) between 2019 and 2020. According to Liu et al. (2022), the main motivators for creating the dataset were the lack of public large-scale ECG datasets and the lack of standardization in the diagnoses used in existing datasets. The authors made the dataset publicly available for use in research that addresses the development and evaluation of arrhythmia classification methods.

The recordings were made using equipment configured according to the 12-lead system with a sampling frequency of 500 Hz, and the duration of each recording ranged from 10 to 60 seconds. The exams were diagnosed by a cardiologist following the standard recommended by the AHA. It was considered 44 classifications among the 117 suggested by the AHA, with certain tests receiving more than one diagnosis. Table 1 presents the 44 classifications considered in the diagnoses.

Table 1.
Classification of arrhythmias according to AHA

Code	Description	Code	Description
1	Normal ECG	102	Left posterior fascicular block
21	Sinus tachycardia	104	Left bundle-branch block
22	Sinus bradycardia	105	Incomplete right bundle-branch block
23	Sinus arrhythmia	106	Right bundle-branch block
30	Atrial premature complex(es)	108	Ventricular preexcitation
31	Atrial premature complexes, nonconduct	120	Right-axis deviation
36	Junctional premature complex(es)	121	Left-axis deviation
37	Junctional escape complex(es)	125	Low voltage
50	Atrial fibrillation	140	Left atrial enlargement
51	Atrial flutter	142	Left ventricular
54	Junctional tachycardia	143	Right ventricular hypertrophy
60	Ventricular premature complex(es)	145	ST deviation
80	Short PR interval	146	ST deviation with T-wave change

Code	Description	Code	Description
81	AV conduction ratio N:D	147	T-wave abnormality
82	Prolonged PR interval	148	Prolonged QT interval
83	Second-degree AV block, Mobitz type I	152	TU fusion
84	Second-degree AV block, Mobitz type II	153	ST-T change due to ventricular hypertrophy
85	2:1 AV block	155	Early repolarization
86	AV block, varying conduction	160	Anterior MI
87	AV block, advanced (high-grade)	161	Inferior MI
88	AV block, complete (third-degree)	165	Anteroseptal MI
101	Left anterior fascicular block	166	Extensive anterior MI

In this study, for standardization reasons, it was decided to use only the first 10 seconds of each ECG recording. Thus, the dataset considered for the study has 128,850,000 samples with 12 features, in which each feature represents the voltage in microvolts obtained in each of the 12 recorded leads. Each ECG exam consists of 5,000 samples since a sampling frequency of 500 Hz was used. It was decided to consider only the first diagnosis of each ECG in cases in which more than one diagnosis was assigned. As a result, 40 classifications were used since 4 classifications that were considered only as secondary diagnoses were eliminated from the dataset.

3.2. Preprocessing

It was decided to use information from a single lead in this study, that is, only one of the 12 features in the dataset was used. The justification for this choice is the fact that, as it is a large-scale dataset, the use of all leads could consume a large amount of computational resources and make model training unfeasible. Lead II was selected in this case because, according to Meek and Morris (2002), this lead is the most used for detecting arrhythmias because it is located close to the cardiac axis and provides the best view of the P waves. When compared to the other leads, the positioning of lead II is better aligned with the direction of propagation of electrical impulses from the sinoatrial node towards the Purkinje fibers. Another reason for choosing to use data from a single lead is that using all leads could result in a more complex model due to having to deal with a greater number of features, which could result in worse performance.

It was verified through exploratory data analysis that there was no problem with missing or inconsistent data in the dataset, eliminating the need to carry out specific treatments for these types of problems. However, the presence of outliers and a moderate negative asymmetry in the data distribution were observed, which required the use of robust normalization techniques to adjust the data so that model training was less affected by outliers.

The 500 Hz sampling frequency used by the ECG made it possible to obtain high-resolution recordings. High sampling rates like this make it possible to analyze events at specific moments in time but can bring the trade-off of requiring a greater amount of computational resources to process large amounts of data. For this reason, it was decided to resample the records to a sampling frequency of 125 Hz. Thus, each ECG record had 1,250 samples from now on. This decision was based on the results obtained from the work presented by Habib et al. (2020), in which the authors compared the training of Convolutional Neural Network (CNN) models using ECG recording datasets with different sampling frequencies and concluded that CNNs demonstrate good generalization capacity even in datasets with lower sampling rates such as 100 Hz and 250 Hz for example.

The dataset was restructured to present the data in the form of a "time window". The 1,250 samples corresponding to each ECG exam were converted into a single sample made up of 1,250 features, in which each feature represents the recording of lead II voltage at a given instant of time. This data structuring is aligned with Cardiology concepts as doctors analyze changes in recording patterns over time intervals. The number of samples in the dataset became 25,770 (number of ECG exams registered).

Table 2 shows the number of samples belonging to each arrhythmia classification. The dataset was considerably unbalanced, as AHA code 1 (without arrhythmia) had 53% of the samples while certain AHA codes had less than 1% of the samples. Upsampling and downsampling techniques were used to balance the distribution of classes so as not to harm the model's performance during training. The technique used for upsampling was the Synthetic Minority Over-sampling Technique (SMOTE). The SMOTE technique requires that each class has at least 6 instances as it uses a K-Nearest Neighbors algorithm considering 5 neighbors to generate each synthetic instance. For this reason, the 8 categories that had less than 6 instances (AHA codes 31, 37, 84, 87, 102, 143, 148, and 152) were removed from the dataset. The dataset now contains 200,000 instances belonging to 32 classes after applying the balancing techniques and was considered suitable for training.

Table 2.
Distribution of categories in the dataset

AHA Code	Quantity	%	AHA Code	Quantity	%
1	13905	53.96	165	64	0.25
22	2659	10.32	104	62	0.24
147	1334	5.18	36	44	0.17
23	1123	4.36	160	35	0.14
145	1045	4.06	155	28	0.11
105	917	3.56	108	22	0.09
60	786	3.05	88	20	0.08
21	723	2.81	54	12	0.05
50	663	2.57	80	9	0.03
146	540	2.10	83	8	0.03
106	473	1.84	140	7	0.03
30	384	1.49	166	7	0.03
125	201	0.78	102	5	0.02
120	122	0.47	31	4	0.02
121	111	0.43	148	4	0.02
82	98	0.38	87	3	0.01
142	96	0.37	152	3	0.01
51	94	0.36	37	2	0.01
101	77	0.30	84	2	0.01
161	77	0.30	143	1	0.00

Table 3 shows the class identifier assigned to each AHA code for "internal" use in training and evaluating the Deep Learning model. The use of numerical and sequential values to identify classes is a requirement of the library used to train the model.

Table 3.
Class identification for each AHA code

Class ID	AHA Code	Class ID	AHA Code
0	1	16	105
1	21	17	106
2	22	18	108
3	23	19	120
4	30	20	121
5	36	21	125
6	50	22	140
7	51	23	142

Class ID	AHA Code	Class ID	AHA Code
8	54	24	145
9	60	25	146
10	80	26	147
11	82	27	155
12	83	28	160
13	88	29	161
14	101	30	165
15	104	31	166

3.3. Model Training

A Convolutional Neural Network (CNN) model was developed for the task of predicting cardiac arrhythmias. The choice to use a CNN model is because, according to Wibawa et al. (2022), CNN models have been explored in recent years for time series analysis tasks and have shown promising results, including surpassing the performance of other types of Deep Learning models by achieving better accuracy in predictions with less consumption of computational resources. Figure 6 shows the architecture of the proposed CNN model. The architecture is composed of 4 convolutional layers, 2 Max Pooling layers, 1 Flatten layer, 1 Fully Connected layer, and 1 output layer. Additionally, there are Batch Normalization layers after the convolutional and Fully Connected layers. The proposed architecture provides a balance between complexity and performance as it uses more convolutional layers to extract features while also using Max Pooling layers to reduce dimensionality and computational cost. This architecture provided better training results when compared to other traditional architectures with 3 convolutional layers that were also tested. Table 4 shows the hyperparameters setup. The definition of hyperparameters was based on some recommendations suggested in the literature and on previous works by the author himself.

The model was implemented using the TensorFlow framework with the Keras library. The training was carried out with a limit of 300 epochs using the Adam optimizer, a learning rate of 0.001, and a batch size of 32. The dataset samples were split in a stratified manner into 80% of samples for training and 20% of samples for testing.

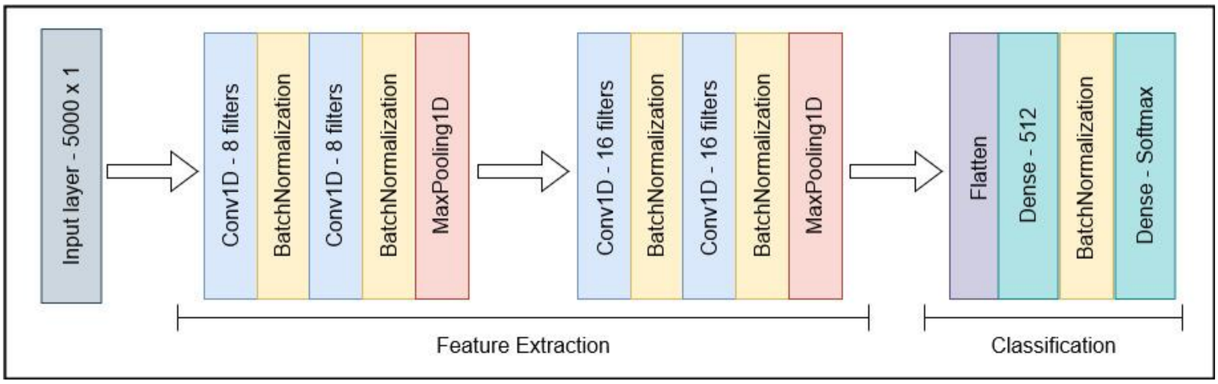

Figure 6. CNN Model architecture

Table 4.
Hyperparameters setup

Layer	Layer type	Filters	Kernel size	Activation	Kernel initializer	Units
1	Conv1D	8	3	ReLU	he_uniform	-
2	BachNormalization	-	-	-	-	-
3	Conv1D	8	3	ReLU	he_uniform	-
4	BachNormalization	-	-	-	-	-
5	MaxPooling1D	-	-	-	-	-
6	Conv1D	16	5	ReLU	he_uniform	-
7	BachNormalization	-	-	-	-	-
8	Conv1D	16	5	ReLU	he_uniform	-
9	BachNormalization	-	-	-	-	-
10	MaxPooling1D	-	-	-	-	-
11	Flatten	-	-	-	-	-
12	Dense	-	-	ReLU	he_uniform	512
13	BachNormalization	-	-	-	-	-
14	Dense	-	-	Softmax	-	32

4. Results and Discussion

The CNN model achieved an accuracy rate of 98.37% on the test set. This performance is considered satisfactory, as it indicates that the model is capable of achieving a high level of accuracy in its predictions. Table 5 shows the detailed training metrics (precision, recall, and F1-Score) for each of the classes. Rates greater than 78% were obtained for these metrics in predicting class 0 (without arrhythmia) and rates greater than 93% for these metrics in predicting all other classes. These results show that the training execution was consistent and the model was able to identify all 32 classes in the dataset with a high level of accuracy.

Table 5.
Detailed metrics

Class	Precision	Recall	F1-Score	Class	Precision	Recall	F1-Score
0	0.84	0.78	0.81	16	0.97	0.95	0.96
1	0.99	1.00	0.99	17	0.99	1.00	0.99
2	0.95	0.98	0.97	18	1.00	1.00	1.00
3	0.96	0.94	0.95	19	1.00	1.00	1.00
4	0.99	0.99	0.99	20	1.00	1.00	1.00
5	1.00	1.00	1.00	21	0.99	1.00	1.00
6	0.98	0.98	0.98	22	1.00	1.00	1.00
7	1.00	1.00	1.00	23	1.00	1.00	1.00
8	1.00	1.00	1.00	24	0.94	0.96	0.95
9	0.99	0.98	0.98	25	0.95	1.00	0.97
10	1.00	1.00	1.00	26	0.95	0.93	0.94
11	1.00	1.00	1.00	27	1.00	1.00	1.00
12	1.00	1.00	1.00	28	1.00	1.00	1.00
13	1.00	1.00	1.00	29	1.00	1.00	1.00
14	1.00	1.00	1.00	30	1.00	1.00	1.00
15	1.00	1.00	1.00	31	1.00	1.00	1.00

The fact that the rates obtained for class 0 are slightly lower than those obtained for the other classes is a detail that deserves attention, as it is an indication that the model had a little difficulty identifying samples of this class. Figure 7 shows the confusion matrix of the model predictions. An analysis of the confusion matrix makes it possible to verify the results of the predictions for samples of class 0. The model obtained a recall rate of 78% for this class. This recall indicates that for every 100 samples that belong to class 0, the model confused 22 samples with other classes. It can be seen that most of the samples misclassified in this case were confused with classes 2, 3, 16, 24, and 26. The model obtained a precision rate of 84% for class 0. This precision indicates that for every 100 samples that were classified by the model as belonging to this class, 16 samples belonged to other classes. It can be seen that most of the samples misclassified in this case were confused with classes 3, 16, 24, and 26.

Classes 2, 3, 16, 24, and 26 correspond respectively to the following AHA codes: 22 (sinus bradycardia), 23 (sinus arrhythmia), 105 (incomplete right bundle-branch block), 145 (ST deviation), and 147 (T-wave abnormality). A possible reason for the confusion of class 0 with these classes is that these types of arrhythmia are considered difficult to diagnose in the medical literature because they may present similar characteristics in ECG data patterns. Sinus bradycardia, for example, normally presents ECG waves with the same morphology as those presented in a normal ECG despite presenting a smaller number of waves per minute. Therefore, the confusion of a sample belonging to class 0 with a sample belonging to class 3 may be an indication that the model is considering wave morphology as the main pattern to differentiate the classes and, therefore, has some difficulty in classifying a sinus bradycardia sample as a type of arrhythmia.

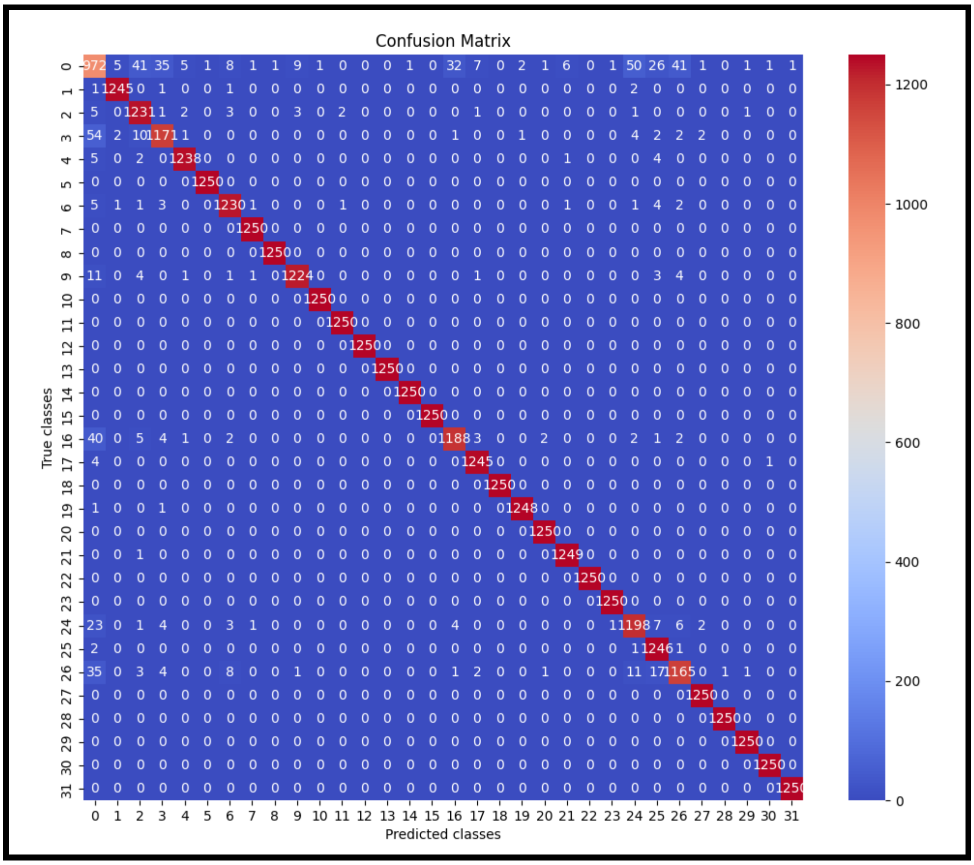

Figure 7. Confusion Matrix

The Local Interpretable Model-agnostic Explanations (LIME) method was used to identify which features contributed most (or were most important) to the prediction of each class. LIME was applied to one sample of each class to identify among the 1,250 available features which were the 50 features that most contributed to the classification of that sample. Figures 8 to 12 show as an example the result of applying LIME to samples of classes 0, 2, 3, 7, and 28. The red dots in the graphs represent the 50 features identified by LIME as the biggest contributors to the identification of each class.

The results presented by LIME can be useful in helping healthcare professionals interpret how the CNN model reached its conclusions, and evaluate whether the results presented are reliable and whether the model's decisions adhere to requirements related to responsibility, ethics, compliance, and regulation. Taking Figure 10 as an example, the features that most contributed to the classification of a sinus arrhythmia sample (AHA code 23) are mostly located in the P wave region and, according to cardiology concepts, sinus arrhythmia is characterized by variations in the P-P interval. Taking Figure 12 as an example, the features that most contributed to the classification of a sample of anterior myocardial infarction (AHA code 160) are mostly located in the ST segment region and, according to cardiology concepts, myocardial infarction is characterized by unevenness in the ST segment.

Efforts to improve the interpretability of model decisions tend to contribute to increasing the acceptance of AI techniques by healthcare professionals as the "black box" nature of AI models has represented a challenge for the adoption of these techniques in the medical field. AI tools must offer the user an explanation of why a given patient has a given disease and the use of interpretability techniques such as LIME can help increase the comprehensibility of the model's decisions and contribute positively to the healthcare professionals' decision-making process.

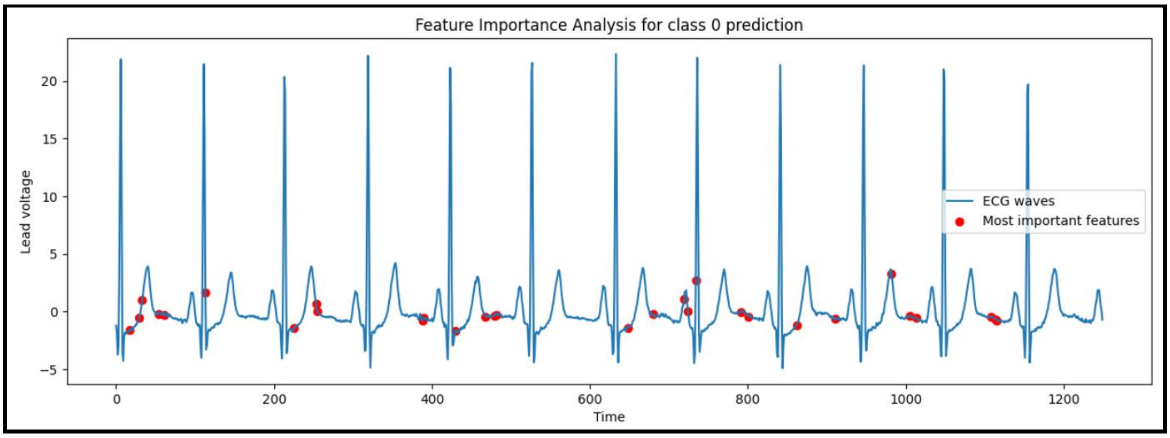

Figure 8. Feature importance for class 0 (Normal ECG)

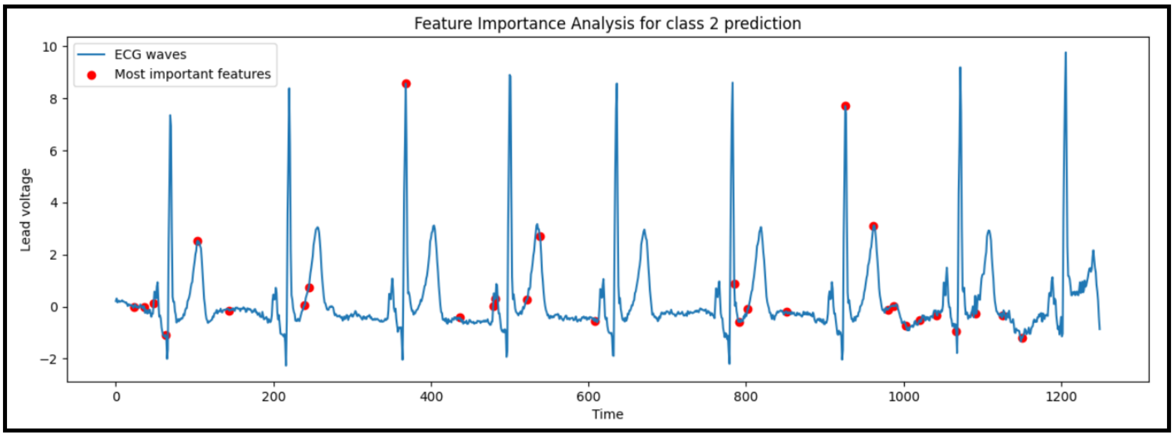

Figure 9. Feature importance for class 2 (Sinus bradycardia)

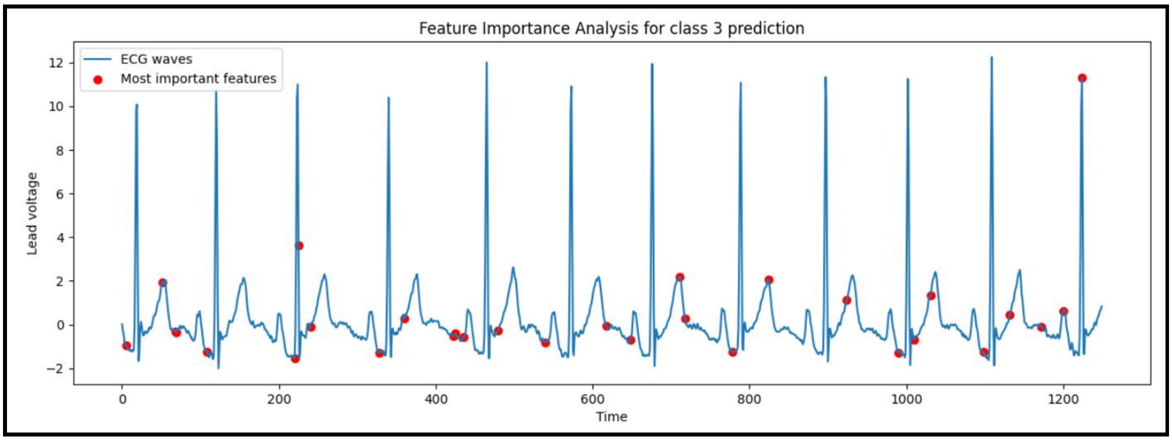

Figure 10. Feature importance for class 3 (Sinus arrhythmia)

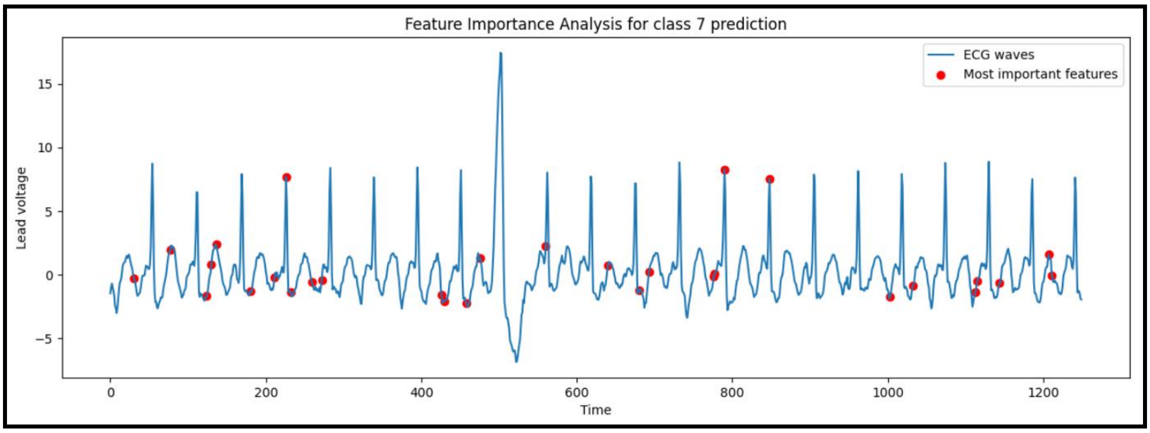

Figure 11. Feature importance for class 7 (Atrial flutter)

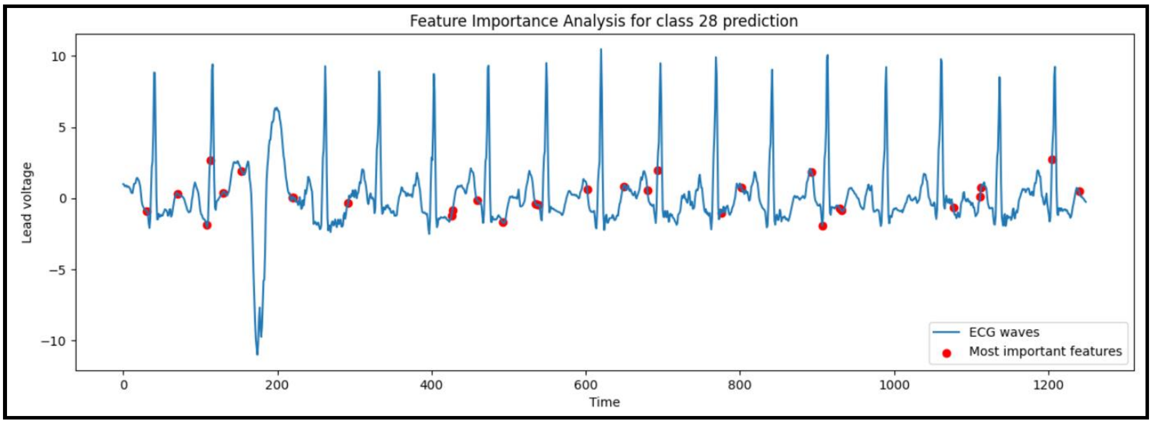

Figure 12. Feature importance for class 28 (Anterior MI)

Figures 13 and 14 show the history of accuracy and loss during training. It can be seen that the CNN model achieved high accuracy rates in the first iterations and converged with less than 40 training epochs. It was observed that the use of Batch Normalization layers resulted in a more stable evolution of validation loss values and better test set accuracy when compared to training the same architecture without Batch Normalization layers.

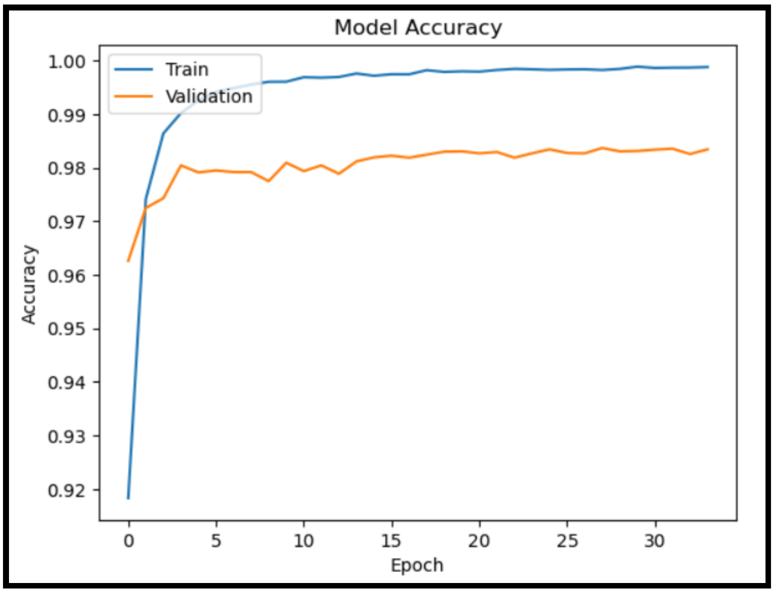

Figure 13. Accuracy history

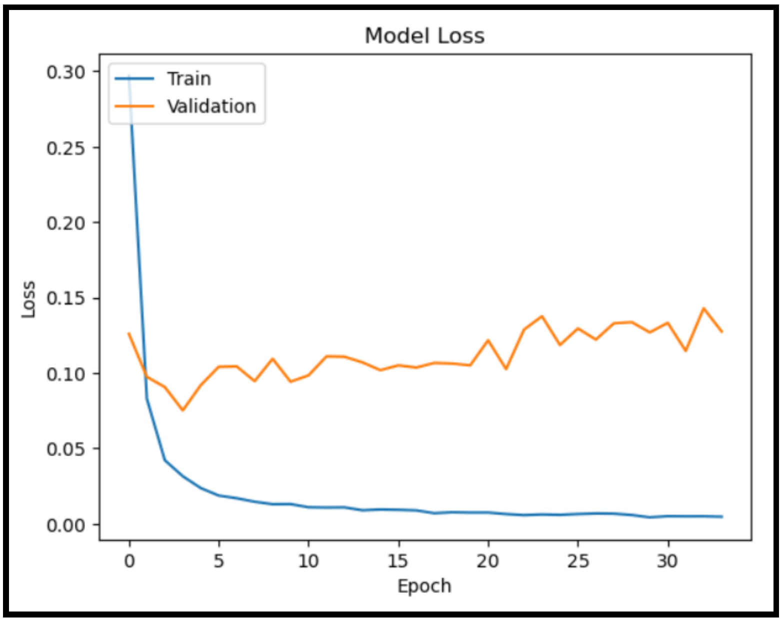

Figure 14. Loss history

The CNN model uses 2,535,176 trainable parameters. A model with this number of parameters is normally considered to be of moderate complexity. This characteristic meets the objective of presenting a Deep Learning model that achieves a satisfactory accuracy rate in its predictions with the use of fewer computational resources for training. The decision to perform a resampling of the records to a frequency of 125 Hz proved to be correct as it made it possible to reduce the computational cost of training without compromising the model's performance.

It was possible to obtain excellent performance using only data from lead II in the training and this result is interesting because wearable devices such as smartwatches, for example, make use of a single lead and the results obtained in this study demonstrate the feasibility of performing accurate cardiac monitoring using such devices. However, the decision to disregard the other leads in the study meant losing the possibility of carrying out comparative analyses to validate the assumption that lead II provides better results than other leads or perform analyses that can provide insights into the data of other leads.

5. Conclusion

Deep Learning techniques have great potential for helping healthcare professionals identify and treat diseases. The possibility of analyzing exams quickly and automatically allows large amounts of information to be analyzed in less time, which can speed up processes in the medical field. Furthermore, professionals can count on a tool that helps reduce diagnostic errors and costs associated with the process.

The development of low-cost Deep Learning models generates an alternative to enable the availability of cardiac monitoring technologies that can be used in wearable devices in environments outside laboratories and clinics. This possibility can be useful in cardiology practice because certain arrhythmias occur intermittently at different times of the day and it can be difficult to diagnose the occurrence of the arrhythmia during examinations in laboratories and clinics.

Aspects related to ethics, transparency, responsibility, and legality represent a challenge for the adoption of AI-based technologies in medical practice because it is a highly regulated sector and AI-based systems need to meet a series of requirements to ensure that their results are considered valid and accepted. Concern about data quality, security, and bias is also an issue that poses challenges for the adoption of AI-based systems because healthcare professionals may become resistant to using the technology if they understand that there is not good data governance in conducting the processes. The need for healthcare professionals to be trained to use this type of technology also creates challenges for its adoption because it creates the need for the training of new healthcare professionals to be adapted to develop skills in the use of digital technologies in their daily practice. To overcome these challenges, investments are needed in adequate training of healthcare professionals to deal with AI-based technologies, the establishment of data governance frameworks that ensure data quality and security, the evolution of regulatory mechanisms associated with the use of AI in healthcare, and collaboration between healthcare and technology professionals to facilitate knowledge sharing.

The proposed CNN model achieved results that satisfactorily meet the performance expectations intended for this study and proved to be an interesting option to deal with the problem in question. As a recommendation for future work, a comparative study between the leads is suggested to verify whether the performance of the model using the other leads could outperform the result obtained with lead II. A comparative study of the differences between leads could reveal useful insights into the data and contribute to research in the area.

References

Becker, D. E. (2006). Fundamentals of electrocardiography interpretation. *Anesthesia Progress*, *53*(2), 53–64. https://doi.org/10.2344/0003-3006(2006)53[53:FOEI]2.0.CO;2

Chakrabarti, S., & Stuart, A. G. (2005). Understanding cardiac arrhythmias. *Archives of Disease in Childhood*, *90*(10), 1086-1090. https://doi.org/10.1136/adc.2005.076984

Cruickshank, J. (2008). Initial management of cardiac arrhythmias. *Australian Family Physician*, *37*(7), 516-520.

Habib A., Karmakar, C., & Yearwood, J. (2020). Choosing a sampling frequency for ECG QRS detection using convolutional networks. *arXiv*. https://doi.org/10.48550/arXiv.2007.02052

Harmon, D. M., Sehrawat, O., Maanja, M., Wight, J., & Noseworth, P. A. (2023). Artificial Intelligence for the detection and treatment of atrial fibrillation. *Arrhythmia & Electrophysiology Review*, *12*(1), e12. https://doi.org/10.15420/aer.2022.31

iMotions. (2024). Electrocardiography (ECG): the complete pocket guide. *iMotions*. Retrieved February 5, 2024, from https://imotions.com/support/document-library/

Kingma, J., Simard, C., & Drolet, B. (2023). Overview of cardiac arrhythmias and treatment strategies. *Pharmaceuticals*, *16*(6), 844. https://doi.org/10.3390/ph16060844

Li, H., Song, X., Liang, Y., Bai, X., Liu-Huo, W., Tang, C., Chen, W., & Zhao, L. (2022). Global, regional, and national burden of disease study of atrial fibrillation/flutter, 1990–2019: results from a global burden of disease study. *BMC Public Health*, *22*(1), 2015. https://doi.org/10.1186/s12889-022-14403-2

Liu, H., Chen, D., Zhang, X., Li, H., Bian, L., Shu, M., & Wang, Y. (2022). A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements. *Scientific Data*, *9*(1), 272. https://doi.org/10.1038/s41597-022-01403-5

Mason, J. W., Hancock, E. W., & Gettes, L. S. (2007). Recommendations for the standardization and interpretation of the electrocardiogram. *Circulation*, *115*(10), 1325–1332. https://doi.org/10.1161/CIRCULATIONAHA.106.180201

Meek, S., & Morris, F. (2002). ABC of clinical electrocardiography. Introduction. I-Leads, rate, rhythm, and cardiac axis. *BMJ*, *324*(7334), 415-418. https://doi.org/10.1136/bmj.324.7334.415

Nagarajan, V., Lee, S., Robertus, J., Nienaber, C. A., Trayanova, N. A., & Ernst, S. (2021). Artificial intelligence in the diagnosis and management of arrhythmias. *European Heart Journal*, *42*(38), 3904–3914. https://doi.org/10.1093/eurheartj/ehab544

Rafie, N., Kashou, A. H., & Noseworthy, P. A. (2021). ECG Interpretation: clinical relevance, challenges, and advances. *Hearts 2021*, *2*(4), 505–513. https://doi.org/10.3390/hearts2040039

Singh, G., Agarwal, C., Kaur, I., & Gupta, P. (2023). Machine Learning for cardiac arrhythmia detection: a systematic survey. *Journal of Physics: Conference Series*, *2570*(1), 012028. https://doi.org/10.1088/1742-6596/2570/1/012028

Wibawa, A. P., Utama, A. B. P., Elmunsyah, H., Pujianto, U., Dwiyanto, F. A., & Hernandez, L. (2022). Time-series analysis with smoothed Convolutional Neural Network. *Journal of Big Data*, *9*(44), 1-18. https://doi.org/10.1186/s40537-022-00599-y

Williams, H. (2005). Arrhythmias: overview of the condition. *The Pharmaceutical Journal*. https://doi.org/10.1211/PJ.2021.1.66303

Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., & Rakovski, C. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific Data*, *7*(1), 48. https://doi.org/10.1038/s41597-020-0386-x