Classification of heart arrhythmia by digital image processing and machine learning

Gustavo Henrique de Oliveira[1]**, Franklin César Flores**[2]

[1]Departamento de Informática – Universidade Estadual de Maringá (UEM).
Maringá – PR – Brazil

`pg402912@uem.br, fcflores@uem.br`

Abstract. The electrocardiogram (ECG) exam can be used reliably as a measure to monitor the functionality of the cardiovascular system. Although there are many similarities between different ECG conditions, the focus of most studies has been to classify a set of database signals known as PhysionNet MIT-BIH and PTB Diagnostics data sets, rather than classifying problems in real images. In this article, we propose methods to extract features from the exam image and then algorithms such as CNN, decision tree, extra trees and random forest are used for the classification of exams, which is able to accurately classify according to the AAMI EC57 standard . According to the results, the suggested method is capable of making predictions with an average accuracy of 97.4 %.

Resumo. O exame de eletrocardiograma (ECG) pode ser utilizado de forma confiável como medida para monitorar a funcionalidade do sistema cardiovascular. Embora existam muitas semelhanças entre diferentes condições de ECG, o foco da maioria dos estudos tem sido classificar um conjunto de sinais de banco de dados conhecido como conjuntos de dados PhysionNet MIT-BIH e PTB Diagnostics, em vez de classificar problemas em imagens reais. Neste artigo, propomos métodos para extrair recursos da imagem do exame e, em seguida, algoritmos como CNN, árvore de decisão, árvores extras e floresta aleatória são usados para a classificação dos exames, que é capaz de classificar com precisão de acordo com o padrão AAMI EC57. De acordo com os resultados, o método sugerido é capaz de fazer previsões com uma precisão média de 97,4 %.

1. Introduction

According to the World Health Organization ([WHO 2017]), cardiovascular diseases (CVD) are the leading causes of death today. More than 17.7 million people died from CVD in 2019, representing about 31% of all deaths and more than 75% of these deaths occur in low- and middle-income countries. Arrhythmias are a representative type of CVD that refers to any irregular change in the normal heart rate. There are several types of arrhythmias, including atrial fibrillation, premature contraction, ventricular fibrillation, among others. Although single arrhythmia heartbeats may not have a serious impact on life, they can still result in fatal outcomes. Thus, it is understood that it is important to screen patients with risk factors as a way to prevent CVD.

According to [Fairley 2017] the ECG corresponds to the recording of electrical phenomena, which originate during cardiac activity, through a device called an electrocardiograph, being one of the most common complementary tests to assess heart rhythm.

Regarding its functionality, an example given [Ramos and Sousa 2007],it serves as a valuable aid in the diagnosis of a large number of heart diseases and other conditions such as electrolyte disturbances, for example .

The ECG needs to be interpreted by a professional to obtain a correct diagnosis. A study done by [Benner et al. 2007] demonstrated that the analysis and interpretation of ECGs changed patient management in 32% of cases. Unlike many laboratory tests, there are few well-defined criteria to give an accurate result when interpreting an ECG. It is common for cardiologists, electrophysiologists and emergency physicians to have low agreement rates between raters (Kappa agreement test) analyzing the same ECG. [Holmvang et al. 1998] and [Massel et al. 2000] point out that even the same evaluator reading the same ECG at different times will arrive at inconsistent results. The American College of Cardiology Foundation estimates that you need to read 3500 supervised ECGs to become a specialist in [Myerburg et al. 2008]. As a result, it is not surprising that errors are common in interpreting ECGs.

Similarly, studies done by [Salerno et al. 2003] show that non-cardiologists have an accuracy of 36% to 96% detecting ECG abnormalities, and from 87% to 100% detecting acute myocardial infarction . As these are diseases with a high mortality and morbidity, the consequences of these errors in interpretation can be significant. Therefore, the use of machine learning algorithms in order to improve accuracy is justifiable.

Most of the studies performed use databases known as the MIT-BIH created by [Moody and Mark 2001], which are databases that contain only two derivations and were created in the 90s. The diagnosis of a more realistic way of classifying the exam and the use of current exam images, which contain 12 leads, improves the amount of characteristics to be extracted to perform the classification.

For a practical way to get the diagnosis, the ECG image can be captured in an uncontrolled environment, where the ECG exam and the camera do not have a fixed place and placement in the environment. With this, the search and segmentation of the region of interest (ROI) of the examination in the image is done, using thresholding [Otsu 1979] and contour detection by [Suzuki and be 1985] and [Xavier and Reshmi 2016] and then, corrections are applied in the perspective of the image using a transformation matrix as proposed by [Reis et al. 2008] to correct deformations on the electrocardiogram.

It is important to highlight that the electrocardiogram data basically consists of twelve recording derivations, six from the frontal electrical plane and six from the horizontal electrical plane ([Feldman and Goldwasser 2004]). Each derivation is formed by waves that are different curvatures of the ECG tracing, up and down. Waves are products of action potentials that occur during cardiac stimulation, and are repeated from beat to beat, except for changes ([Ribeiro 1991]). The objective is to obtain the extraction of features in each derivations and create a model for classification, hoping to achieve valid results with an optimal value of accuracy with just an image of the electrocardiogram exam. In addition, we will seek to build a model with these data from all types of possible diagnoses and, with the use of classification algorithms, classifying the exam as [Acharya et al. 2017, Ji et al. 2019, Jin and Dong 2018], getting good classification results in the diagnosis.

2. Methodology

In this section, the steps of the proposed method will be detailed, which are: data acquisition and pre-processing, ECG detection using contour detection [Suzuki and be 1985, Xavier and Reshmi 2016], segmentation of the useful region of the exam, perspective correction using transformation matrix like [Reis et al. 2008], segmentation of derivations using color range and contour detection algorithm [Suzuki and be 1985, Xavier and Reshmi 2016] through thresholding [Otsu 1979], extraction of characteristics of each derivation using local maximum points using [Schwartzman et al. 2011] and baselines of [Al-Ani 2018]. Finally, to classify the exams, the Decision Tree [carvalho 2004], Random Forest [Ho 1995] and Extremely Randomized Trees (extra tree) [Zhou et al. 2020] methods were used.

The algorithms used in this work for all steps of the proposed method were implemented using the python language with the assistance of image processing libraries and tools for predictive data analysis like Scikit-learn [Pedregosa et al. 2011].

2.1. Data Colletion and Data Pre-Processing

The database was formed from the medical records of patients. Images of electrocardiogram exams were captured from the patients' medical records and information such as age, height, weight, sex and comorbidities of each patient were also taken. The images were captured in the horizontal position, in 4K resolution and with good lighting/illumination, preventing the shadows from generating noise in the image. All this information and the characteristics found from the next steps will be used for the classification of cardiac alterations.

Information was taken from more than 1200 different patients and 2000 ECG images were captured with medical reports of seven cardiovascular alterations in a balanced way. The alterations are: 1st degree atrioventricular block (1stAVB), 2nd degree atrioventricular block (2stAVB), Complete atrioventricular block (CAVB), Flutter and atrial fibrillation (AF), Flutter and ventricular fibrillation (AV), Supraventricular tachycardia (ST) and ventricular tachycardia (VT).

For the extraction of features found in the exams, the next steps are performed.

2.2. Detection of Region of Interest

To execute the detection, it is necessary to locate the region of interest in the image. On some electrocardiograms, in addition to information on electrical pulses, there is also information about the patient and the exam, such as the patient's name, date and time when the ECG was performed. As this information will not be used, the region of interest for detection is only the squared part of the exam, shown in Figure 1 (a) outlined in blue, which contains the leads formed by the electrical pulses.

Thus, as the region of interest is the gridded area is Figure 1, the Otsu threshold algorithm [Otsu 1979] is applied to obtain a binarized image as shown in Figure 2.

Furthermore, the algorithm of [SUZUKI AND BE 1985] was used to contour detection in order to extract the gridded region, as well as to eliminate the elements that will not be used. Consequently, the region of interest in the image is the one with the largest contour found.
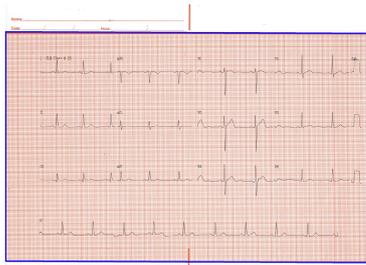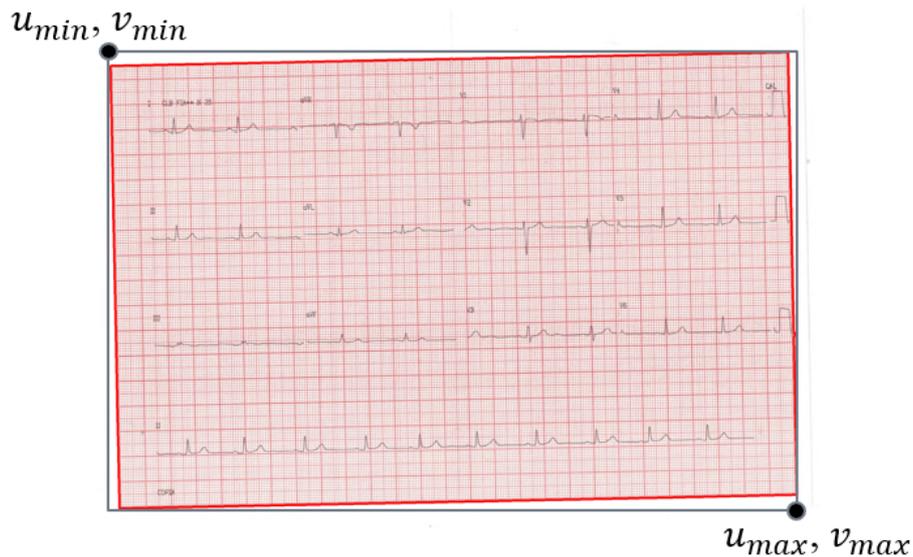
Figura 1. Region if interest in the exam

Figura 2. Binarized exam

2.3. Region of Interest Segmentation

After defining the contours of the exam, they are segmented. The smallest and largest values of u and v of the board outline are used, that is, *umin*, *vmin*, *umax* and *vmax* as shown in Figure 3. In such manner, the regions of interest for the exam are obtained.


Figura 3. Segmented Exam

2.4. Perspective Correction

Each exam is usually represented in the image with a viewpoint that may have deformations caused at the time of capture. These deformations affect the attaining of the derivations, as they can interfere with the final classification if the lead sign is modified.

In order to correct these deformations, the ECG perspective correction was performed using a transformation matrix by [Reis et al. 2008]. The matrix holds the values of the four coordinate points (x_0, y_0), (x_1, y_1), (x_2, y_2) and (x_3, y_3) of the outline of the exam in the image and also the four points of the region of interest (u_0, v_0), (u_1, v_1), (u_2, v_2) and (u_3, v_3).

The coordinates (u, v) are obtained from the values $(u_{min}, v_{min}, u_{max}, v_{max})$ of the region of interest, which contains the segmented exam, and the coordinates (x, y) were determined from this region in the lines that compose the contours of the ECG in the

image. In Figure 4 are shown the coordinates (u, v) of the region of interest of the image and the coordinates (x, y) that represent the contour of the board in red color.
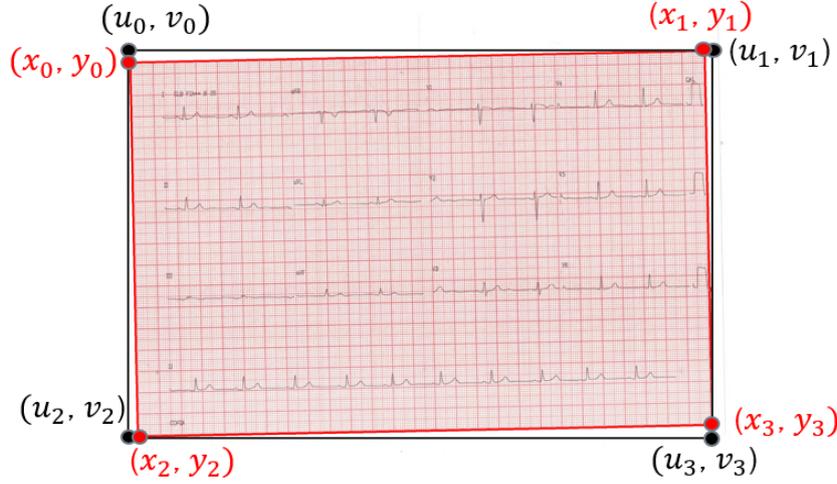
Figura 4. Coordinates (x,y) and (u,v) of the electrocardiogram.

To perform image distortion and transform coordinates (x, y) (A) to coordinates (u, v) (B) is used Eq. 1:

$$A.X = B \tag{1}$$

where X represents the unknown coefficients, which are needed to complete the transformation matrix Eq. 2 ([REIS ET AL. 2008]):

$$
\overset{A}{\begin{bmatrix}
x_0 & y_0 & 1 & 0 & 0 & 0 & -x_0u_0 & -y_0u_0 \\
x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -y_1u_1 \\
x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2u_2 & -y_2u_2 \\
x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3u_3 & -y_3u_3 \\
0 & 0 & 0 & x_0 & y_0 & 1 & -x_0v_0 & -y_0v_0 \\
0 & 0 & 0 & x_1 & y_1 & 1 & -x_1v_1 & -y_1v_1 \\
0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_2 & -y_2v_2 \\
0 & 0 & 0 & x_3 & y_3 & 1 & -x_3v_3 & -y_3v_3
\end{bmatrix}}
\cdot
\overset{X}{\begin{bmatrix}
a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \\ c_0 \\ c_1
\end{bmatrix}}
=
\overset{B}{\begin{bmatrix}
u_0 \\ u_1 \\ u_2 \\ u_3 \\ v_0 \\ v_1 \\ v_2 \\ v_3
\end{bmatrix}}
\tag{2}
$$

As it is necessary to find the values of X, a linear system is formed, which is solved by the LU decomposition (Lower and Upper) such as in [BUNCH AND HOPCROFT 1974]. The LU decomposition separates the matrix in two:

$$A = L.U \tag{3}$$

where L is the lower triangular matrix and U is the upper triangular matrix. To obtain the matrices L and U, the Gauss elimination method [Kaw et al. 2018] was used. Gauss elimination aims to zero out all elements below the main diagonal of the matrix, thus obtaining the matrix U. After this process, the matrix U is formed by the matrix with

the elements below the main diagonal zeroed, and the matrix *L* is formed with the value of the multipliers used to reset the respective elements of the matrix *U*, and with the values above the main diagonal zeroed out.

After getting both *Lower* and *Upper* matrices, the matrix *A* can be replaced by the matrices *L* and *U*, thus forming *(LU).X=B* or *L.(UX)=B* . With this it is possible to obtain:

$$L.Y = B \tag{4}$$

$$U.X = Y \tag{5}$$

After solving Eqs. 4 and 5, the values of the unknown coefficients are found. With all the coefficients already found, the following equations are used to find the values of *U*, *V*:
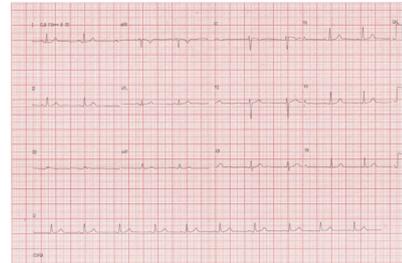
$$u_i = \frac{(a_0 * x_i + a_1 * y_i + a_2)}{c_0 * x_i + c_1 * y_i + 1} \tag{6}$$

$$v_i = \frac{(b_0 * x_i + b_1 * y_i + b_2)}{c_0 * x_i + c_1 * y_i + 1} \tag{7}$$

Then, for each pixel coordinate, Eqs. 6 and 7 are applied, obtaining the new coordinate. After calculating the new coordinates of all the *pixels*, the image has its perspective corrected as shown in Figure 5 and Figure 6.
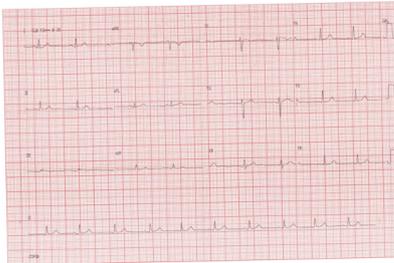
Figura 5. Exam with deformations

Figura 6. Exam with corrected perspective

2.5. Derivation Segmentation

The electrocardiogram contains twelve lead traces and three calibrators which are the regions of interest in this section [FELDMAN AND GOLDWASSER 2004].

By default, the traces of leads and calibrators contained in the electrocardiogram are all black, which facilitates the visualization by contrast with the rest of the exam. The black color in the HSV channel (*Hue*, *Saturation*, *Value*) is represented with the value of *V* as zero, then, using the threshold [Otsu 1979] of the variable *V* it is possible to perform the separation of derivation and background. The result is a binarized image as shown in Figure 7.

After this process, the contour detection [XAVIER AND RESHMI 2016] was applied to identify all regions of interest in the exam.

Figura 7. Binarized electrocardiogram.

From the contours obtained from the derivations, the segmentation of each of them was performed, using the same technique as the segmentation step of the gridded area, with the minimum and maximum values of height and the width of each derivation: *xmin*, *ymin*, *xmax* and *ymax*. An example of the result of this character segmentation step can be seen in Figure 8.

[V5 lead.]

[Binarized V5 lead.]
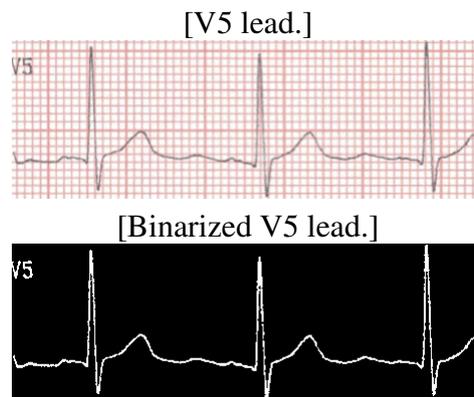
Figura 8. Example of segmented V5 lead

2.6. Feature Extraction

To carry out the classification, it is necessary to gather features that are relevant for the diagnosis. When the health professional evaluates an electrocardiogram exam, he/she looks for certain characteristics and, if there is any alteration or anomaly, the exam is classified as non-normal.

Characteristics extracted from each derivation:

- Heart Rate;
- Horizontal and Vertical size of each wave found in the lead (P, Q, R, S)

Before extracting the information, it is necessary to understand a little about the waves found in the leads. Although there may be small differences, electrocardiogram waves follow patterns in duration, amplitude, and the order in which they appear.

2.6.1. Heart Rate Extraction

To calculate the heart rate, it is first required to spot the intervals between the R waves. For this, it is important to find the R waves present in the lead. The location of the waves is performed by comparing neighboring values to find local maximums and then only the largest peaks are filtered out.

After obtaining the location of the R waves, the difference between the beginning and the end of one wave to another is removed. The formula to find the beats per minute (bpm) is obtained by dividing 1500 by the number of millimeters between waves. This number is used because, as the standard ECG paper speed is 25mm/s, in 60 seconds, 1.5000mm would be traversed. In order for the standard formula to be applied, the extracted derivation is resized. Each derivation has a fixed size of 60 horizontal squares. Then, the formula for calculating the heart rate is done by the horizontal size of the lead divided by 60.

$$bpm = len_{lead} \div 60 \times 1500 \times len_{squares} \qquad (8)$$

Whether there is more than one range, all are analyzed. If all intervals are normal it is averaged and sent, otherwise the single rate is sent (less than 60 bpm or greater than 100 bpm).

To obtain the size of the waves, it is first necessary to detect each one. Considering that the order in which they appear is standardized, it is used to detect the R wave already found for heart rate calculation. As what makes the waves different is that they alternate between positive and negative, it is possible to detect them using the isoelectric line that separates positive from negative voltages:
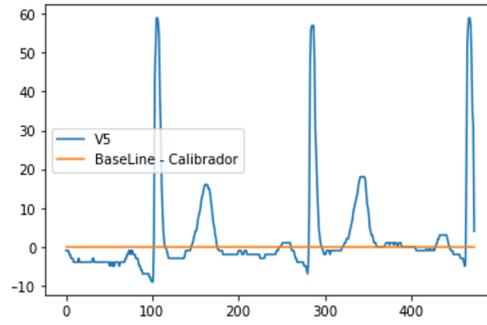
- calibrators;
- baseline [Al-Ani 2018]

The calibrators are present on the electrocardiogram to make comparisons with waves. They start as isoelectric at position 0 and there is an electrical pulse of 1 mV in amplitude that is equivalent to 0.2 seconds in duration, printing a rectangular image 10 mm in height and 5 mm in duration on the exam. As we have both the positions of the leads and the calibrators in the image, only the difference between them is made, as if the start of the calibrator is zero. Every wave above this position is positive and every wave below is negative as seen in the Figure 9 .

The baseline used by Al-ani is based on the idea that waveforms are isoelectric lines [AL-ANI 2018]. So most of it would be registered as zero. Thus, all black *pixels* present in each line are counted and the line that achieves the highest amount (mode) of black *pixels* is chosen as the base line. If there is more than one mode, the average is obtained . Figure 10 depicts an example of the formed line.

2.6.2. Wave segmentation

After obtaining the two lines that separate the positive waves from the negative waves, each wave is segmented.

Figura 9. Baseline using the calibrator.
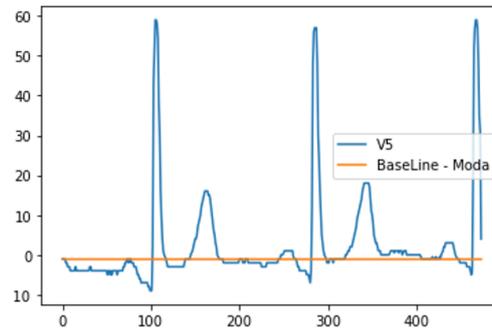
Figura 10. Baseline using the Al-Ani method

As we already get the R wave from the calculation of the local maximum, the extraction of the P, Q and R waves is obtained through the order in which they meet: The P wave is the first positive wave before the R wave; The Q wave is the first negative wave before the R wave; The S wave is the first negative wave after the R wave.

After the waves and their intervals are found, the width and height of each wave are acquired. The width means the time in seconds and the height means the voltage of the wave. Eq. 9 is used to convert the width in seconds, based on the checkered grid described in Figure 11.

$$Time_{wave} = len_{wave} \times 60 \div len_{lead} \times 0.04 \tag{9}$$

2.7. Classification

To perform the classification, three methods were used: Decision Tree, Random Forest and Extra Tree. These methods were chosen because they simulate the decision in stages (devide and conquer) in the same way as cardiologists according to [Podgorelec et al. 2002] and also because they obtain better performances than artificial neural networks.

3. Results

The experiment was performed using a single CPU at 2.00 GHz and 16 GB of RAM. To validate the performance of the classification algorithm, the whole data set was arbitrarily split into 80% and 20% for training and testing, respectivel. Good validation accuracy of 97.4% was obtained.
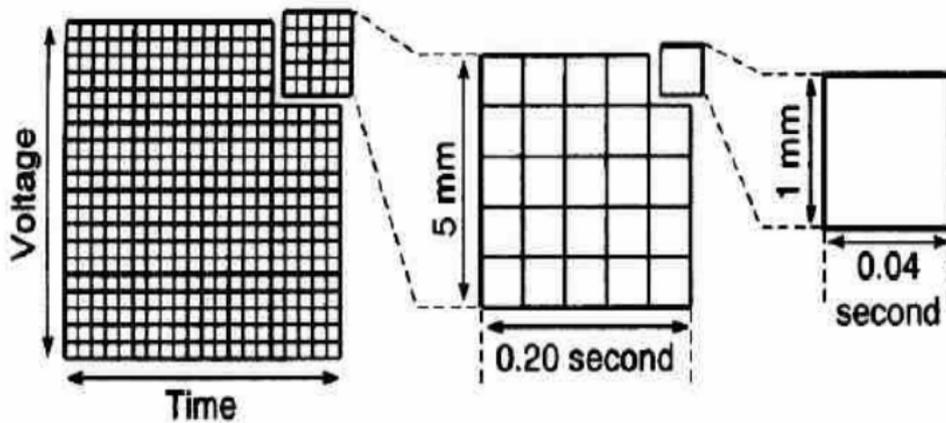
Figura 11. Checkered grid values

To investigate the performance of the models, validation accuracy, precision, recall, and F-1 Score are determined. Precision, recall and F-1 Score are acquired using the confusion matrix. Figure 12 shows the seven-class confusion matrix obtained by achieved accuracy of the Extra Tree algorithm.

Figura 12. Extra Tree confusion matrix.

The table 1 shows the values of various performance metrics obtained. The values show that classification algorithms can successfully classify cardiovascular changes on ECG images.

Tabela 1. Performance Metrics

	Decision Tree	Random Forest	Extra Tree
Accuracy	90.9%	94.2%	97.4%
Precision	90.9%	94.3%	97.4%
Recall	90.9%	94.0%	97.3%
F1-Score	91.0 %	93.9%	97.3%

4. Conclusion

This article proposes an effective method to extract features from images from electrocardiogram exams and diagnose them. As the inputs of the classification algorithms are a list of numerical features, training and inference is performed quickly. But on the other hand, the entire image processing part and extraction of features from the image of the exam takes one to two seconds per exam due to limited computational resources. Our experimental results in the database formed by the Cardiology records show a classification accuracy of 97.4% for the dataset of 7 different cardiac alterations. Our expectation is that this system can be developed and adjusted for practical application.

Referências

Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., and Adam, M. (2017). Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. *Information Sciences*, 405:81–90.

Al-Ani, M. (2018). Ecg waveform classification based on p-qrs-t wave recognition.

Benner, J. P., Borloz, M. P., Adams, M., and Brady, W. J. (2007). Impact of the 12-lead electrocardiogram on ED evaluation and management. *Am J Emerg Med*, 25(8):942–948.

Bunch, J. R. and Hopcroft, J. E. (1974). Triangular factorization and inversion by fast matrix multiplication. *Mathematics of Computation*, 28(125):231–236.

carvalho, I. C. (2004). Métodos de mineração de dados (data mining) como suporte à tomada de decisão.

Fairley, R. (2017). Ecg analysis: Where do we go wrong and how can we improve?

Feldman, J. and Goldwasser, G. P. (2004). Eletrocardiograma: recomendações para a sua interpretação. *Revista da SOCERJ*, 17(4):251–256.

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282 vol.1.

Holmvang, L., Hasbak, P., Clemmensen, P., Wagner, G., and Grande, P. (1998). Differences between local investigator and core laboratory interpretation of the admission electrocardiogram in patients with unstable angina pectoris or non-q-wave myocardial infarction (a thrombin inhibition in myocardial ischemia [TRIM] substudy). *Am J Cardiol*, 82(1):54–60.

Ji, Y., Zhang, S., and Xiao, W. (2019). Electrocardiogram classification based on faster regions with convolutional neural network. *Sensors (Basel)*, 19(11).

Jin, L. and Dong, J. (2018). Normal versus abnormal ecg classification by the aid of deep learning.

Kaw, A., Kalu, E., and Nguyen, D. (2018). Numerical methods with applications: chapter 04.06 gaussian elimination.

Massel, D., Dawdy, J. A., and Melendez, L. J. (2000). Strict reliance on a computer algorithm or measurable st segment criteria may lead to errors in thrombolytic therapy eligibility. *American Heart Journal*, 140(2):221–226.

Moody, G. and Mark, R. (2001). The impact of the mit-bih arrhythmia database. *IEEE engineering in medicine and biology magazine : the quarterly magazine of the Engineering in Medicine e Biology Society*, 20:45–50.

Myerburg, R. J., Chaitman, B. R., Ewy, G. A., and Lauer, M. S. (2008). Task force 2: Training in electrocardiography, ambulatory electrocardiography, and exercise testing. *Journal of the American College of Cardiology*, 51(3):348–354.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Podgorelec, V., Kokol, P., Stiglic, B., and Rozman, I. (2002). Decision trees: An overview and their use in medicine. *Journal of Medical Systems*, 26(5):445–463.

Ramos, A. and Sousa, B. S. (2007). Eletrocardiograma: princípios, conceitos e aplicações. centro de estudos de fisiologia do exercício.

Reis, B. F., Teixeira, J. M., Teichrieb, V., and Kelner, J. (2008). Perspective correction implementation for embedded (maker-based) augmented reality. universidade federal de pernambuco, centro de informática. *Workshop de Realidade Virtual e Aumentada, Bauru*.

Ribeiro, A. (1991). Sistema para simulação de ecg/arritmias cardíacas para ensino e treinamento de pessoal.

Salerno, S. M., Alguire, P. C., and Waxman, H. S. (2003). Competency in interpretation of 12-lead electrocardiograms: a summary and appraisal of published evidence. *Ann Intern Med*, 138(9):751–760.

Schwartzman, A., Gavrilov, Y., and Adler, R. J. (2011). MULTIPLE TESTING OF LOCAL MAXIMA FOR DETECTION OF PEAKS IN 1D. *Ann Stat*, 39(6):3290–3319.

Suzuki, S. and be, K. (1985). Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46.

WHO (2017). Cardiovascular diseases (cvds).

Xavier, S. and Reshmi, R. (2016). Automatic detection and recognition of tet in traffic sign boards based on word recognizer.

Zhou, Q., Lan, W., Zhou, Y., and Mo, G. (2020). Effectiveness evaluation of anti-bird devices based on random forest algorithm. In *2020 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, pages 743–748.