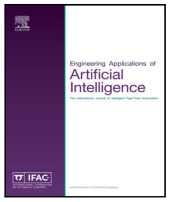
Engineering Applications of Artificial Intelligence 107 (2022) 104536

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



A novel ECG diagnostic system for the detection of 13 different diseases

Iñigo Monedero

Escuela Politécnica Superior, Departamento de Tecnología Electrónica, Universidad de Sevilla, C/Virgen de África 7, 41011, Seville, Spain

ARTICLE INFO

Keywords:
Electrocardiogram
Expert system
Wavelet transform
Decision tree

ABSTRACT

Manual analysis of electrocardiogram (ECG) signals is a laborious and prone-to-error task, even for a specialist with many hours of experience. For this reason, research on automatic ECG diagnosis is widespread in the literature and continues to grow each year. The present paper describes a novel and fully functional expert system for automatic diagnosis of 13 different diseases using standard 12-lead ECGs. This system makes three significant contributions to the state of the art: (a) the large number of different diseases diagnosed; (b) the use of 5 leads for a more precise identification and measurement of the ECG waves; and (c) a novel noise indicator that measures the quality of the acquired ECG signal. The kernel of the system consists of a set of rules that replicate a specialist's diagnostic process but with the speed of an automatic system. The rules use a set of parameters generated after a noise-filtering process of the ECG signal and subsequent identification of its different waves (P, QRS complex, T, and Delta). The design of the rules was carried out with the collaboration of a specialist with more than 20 years of experience in ECG diagnosis and using a database of 284,000 ECGs as support. The system was validated by the specialist, obtaining a reliability of 80.8%. Given the complexity of the problem and the number of diagnoses covered, the results are considered satisfactory and make the system a useful support tool for diagnosis.

1. Introduction

The electrocardiogram (ECG) is the most widely used noninvasive technique in heart disease diagnoses. Since it reflects the electrical activity within the heart during contraction, the time it occurs and its pattern provides much information about the state of the heart.

Each heartbeat constitutes a cardiac cycle, which is reflected in the ECG graph: a P-wave, a QRS complex and a T-wave. In addition, a small U-wave may be visible in certain ECGs. The line voltage ECG base is also known as the isoelectric line or baseline. Normally, the baseline is the portion of the stroke that follows the T-wave and precedes the next P-wave. Fig. 1 shows a schematic record of a normal cardiac cycle.

An ECG can have different configuration of leads that reflect the set of cardiac cycles of the heart from different perspectives or angles. Thus, the standard 12-lead ECG is one of the most commonly used medical studies in the assessment of cardiovascular disease. It provides a complete picture of the heart's electrical activity by recording information through 12 different perspectives calculated using 10 electrodes (Anderson and DiCarlo, 2000). Each lead of the 12 (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, and V6) represents a part of the heart and specialists/cardiologists use one or another lead depending on the disease to be detected. Fig. 2 shows the different ECG leads and the view/perspective of the heart that each one provides. It is possible to perform ECGs with a single lead (for example, this is done by some of the smart watches currently on the market) (Samol et al., 2019), but this implies limited studies and the possibility of detecting a small number of diseases and with a lower degree of reliability.

The work presented in this paper arises in the framework of a collaboration by the Department of Electronic Technology at the University of Seville with the Preving Investment Company, an important labor risk company that covers more than 70,000 companies and 500,000 workers. The work consists of the design of a rule-based expert system, developed as a software application, for the detection of 13 different diseases using standard 12-lead ECGs. Looking at the current horizon of the state of the art, analyzed in Section 2, the presented work is a relevant contribution in terms of both the number of diseases covered and the number of ECG leads used.

A specialist with more than 10 years of experience in ECG diagnostics was available during the development of the work. This specialist resolved the doubts that arose throughout the process and collaborated in the system design. In addition, he was in contact with his peers to discuss those aspects that required consensus.

In order to carry out the design of the system, a large ECG database provided by the Preving Investment Company was used as support. This database is described in Section 3.

The developed system works on the Microsoft Windows operating system and was programmed in R language. The design of the system is described in Sections 4 and 5. The system was validated with the collaboration of the specialist. This validation process is described in Section 6.

E-mail address: imonedero@us.es.

https://doi.org/10.1016/j.engappai.2021.104536
Received 21 January 2020; Received in revised form 13 September 2021; Accepted 31 October 2021
Available online 22 November 2021

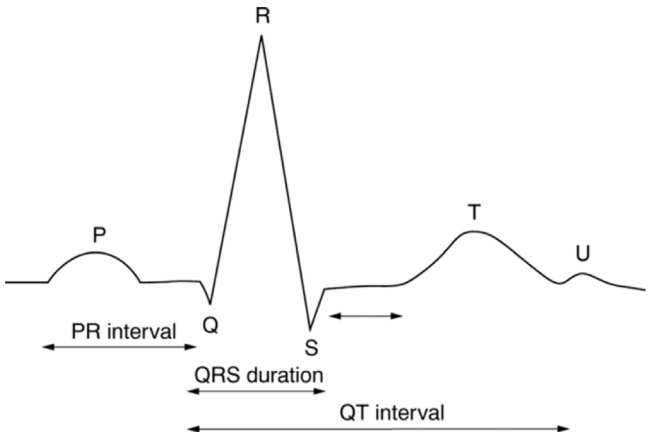


Fig. 1. Waves reflected in an ECG corresponding to a cardiac cycle.

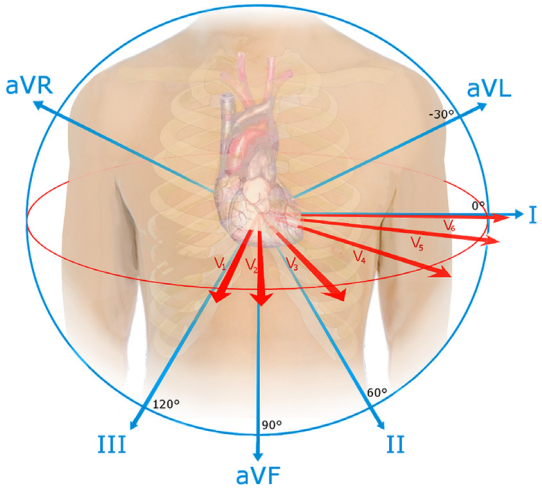


Fig. 2. ECG views of the heart. Image created by Nicholas Patche, Boston Medical Center (CC BY-SA 4.0).

2. State of the art

Automatic analysis and diagnosis of ECGs is an unresolved and much-debated topic in the literature (Berkaya et al., 2018; Martis et al., 2014). In this section the most relevant studies of recent years related to the objective of the manuscript will be referenced and discussed.

2.1. Noise filtering

In the studies on ECG classification referenced in this paper (Adam et al., 2018; Alarsan and Younes, 2019; Altay et al., 2019; Chun-Cheng et al., 2019; Golrizkhatami and Acan, 2018; Jun et al., 2018; Kanungo and Sabut, 2015; Kasar and Joshi, 2016; Li et al., 2016; Li and Zhou, 2016; Haddadi et al., 2014; Nayak et al., 2019; Pingale and Daimiwal, 2014; Sahoo et al., 2017; Shao et al., 2018; Singh and Sunkaria, 2017; Singh et al., 2018; Tripathy and Dandapat, 2016), the first problem that researchers face is the filtering of noise embedded within the ECG signal. This arises because in practice, the acquired electrocardiographic signal is a weak signal which is often heavily contaminated by interference and high-frequency noise, such as power-line interference, electromyography noise, and instrumentation noise (Limaye and Deshmukh, 2016). These noise signals cannot be completely removed by a differential instrumentation amplifier. In addition, the signal is mixed with various low-frequency artifacts generated as a result of a deep breath or the movement of the patient. Finally, there is also a certain amount of white noise generated by the amplifier system itself. The types of noise and the different noise removal techniques are described in the literature (Haritha et al., 2016; Limaye and Deshmukh, 2016).

The wavelet transform (Rajini, 2016) is presented as an extended tool when filtering the ECG signal. It divides the analyzed data into different frequency components and then studies each component in resolutions that fit its scale. Compared to other types of frequency analysis, such as Fourier transform, it has the advantage of being localized both in time and in the frequency domain, and enable the researcher to observe and analyze data at different scales.

In many of the published studies on the topic (Chun-Cheng et al., 2019; Haddadi et al., 2014; Kanungo and Sabut, 2015; Li et al., 2016; Sharma et al., 2019; Singh and Sunkaria, 2017), wavelet transform is used as the main tool for filtering the signal and/or extracting its features. The aforementioned works obtain good results in ECG filtering for small databases (mainly composed of arrhythmias).

There are other ECG signal filtering tools, used in various published studies (Altay et al., 2019; Kasar and Joshi, 2016; Shao et al., 2018; Singh et al., 2018), based on the classic filtering techniques, such as the use of low-pass filters, the filters of moving average, the use of the average of beats, and the approximation through functions. These studies (again applied to a small number of diseases) show that classic techniques can be a good filtering tool or a useful complement to other more modern techniques.

The main problem that exists is that certain noise components overlap in frequency with components of the different ECG waves. It makes total noise filtering in an ECG a subject of high complexity and implies additional complexity in tackling the problem of disease detection.

2.2. Wave identification

In the works published on automatic ECG processing, after handling the noise problem, the generation of a set of indicators (feature extraction process) that allow the detection of possible diseases in the ECG is carried out. For this, a first ECG processing is usually necessary to identify the different waves, time intervals, amplitudes, morphologies, etc. This processing is also a complex problem (Wijaya et al., 2018). The reasons for this complexity lie in the great difficulty involved in programming accurately the calculation of the necessary temporal parameters, considering the different possible morphologies of ECGs and, above all, in ensuring that the noise found in the signal does not alter the obtaining of the measurements with a certain degree of accuracy.

Within the state of the art, wavelet transform is again seen as an extended tool for this objective (Chun-Cheng et al., 2019; Golrizkhatami and Acan, 2018; Haddadi et al., 2014; Li and Zhou, 2016; Kanungo and Sabut, 2015; Nayak et al., 2019; Sharma et al., 2019; Sahoo et al., 2017; Singh et al., 2018). There are also studies that use other types of filters to separate the waves (Altay et al., 2019; Pingale and Daimiwal, 2014; Shao et al., 2018). Finally, it is possible to find some studies based on supervised models, which directly perform a feature extraction of the ECG signal without prior detection of the waves (Li et al., 2016; Li and Zhou, 2016; Mathews et al., 2018; Tripathy and Dandapat, 2016; Yildirim et al., 2019). The feature extraction and correct measurement of waves and intervals are topics currently under research.

2.3. Modeling

Once the necessary waves, intervals and morphologies have been measured, it is necessary to program a model or sets of these to perform the classification. There are different published solutions, based on different algorithms such as clustering (Fuzzy C-Means clustering in Golrizkhatami and Acan, 2018; Haddadi et al., 2014, artificial neural networks (Jun et al., 2018; Mathews et al., 2018; Singh et al., 2018;

Yildirim et al., 2019)), decision trees (Alarsan and Younes, 2019; Kasar and Joshi, 2016; Li and Zhou, 2016; Shao et al., 2018) or the generation of mathematical rules and other techniques (Golrizkhatami and Acan, 2018; Mathews et al., 2018; Nayak et al., 2019; Pingale and Daimiwal, 2014; Singh et al., 2018; Tripathy and Dandapat, 2016). The selection of the appropriate algorithm depends on the data sample and the disease or diseases to be detected. Thus, if there is no solid data set on which to use supervised algorithms (such as decision trees or artificial neural networks), it is necessary to go to other tools (such as the expert systems).

Regarding the number of different diseases classified, it is currently not possible to find published works regarding the diagnosis of a significant number. Thus, most of the studies found in the literature are focused on detecting arrhythmias (Alarsan and Younes, 2019; Golrizkhatami and Acan, 2018; Jun et al., 2018; Kasar and Joshi, 2016; Nayak et al., 2019; Pingale and Daimiwal, 2014; Shao et al., 2018; Sharma et al., 2019; Singh et al., 2018; Yildirim et al., 2019). Only three works targeting other diseases have been found: (Tripathy and Dandapat, 2016) focused on bundle branch block, heart muscle disease and myocardial infarction, (Adam et al., 2018) on dilated cardiomyopathy, hypertrophic cardiomyopathy and myocardial infarction, and (Sahoo et al., 2017) on left bundle branch block, right bundle branch block and paced beats. On the other hand, few studies have been seen in which several ECG leads are used for the analysis and detection of the diseases (Kasar and Joshi, 2016; Tripathy and Dandapat, 2016). The reason why published studies focus on arrhythmias, use a single-lead ECG, and supervised models for classification is clarified in Section 3.

3. Database

In relation to the databases used for modeling, the vast majority of the published studies use long-term ECGs, such as the MIT-BIH (Massachusetts Institute of Technology – Beth Israel Hospital) and the AHA (American Heart Association). These databases are freely accessible and each of the records belonging to them is more than 30 min long, containing about 2000 cardiac cycles. MIT-BIH database (MIT-BIH Database and Software Catalog, 2019) is in turn composed of several databases with characteristics that encompass different diseases (although in a limited number). Specifically, these databases are: MIT-BIH Arrhythmia Database, Creighton University Ventricular Tachyarrhythmia Database, MIT-BIH Noise Stress Test Database, MIT-BIH ST Change Database, MIT-BIH Malignant Ventricular Arrhythmia Database, MIT-BIH Atrial Fibrillation/Flutter Database, MIT-BIH ECH Compression Test Database, MIT-BIH Supraventricular Arrhythmia Database, MIT-BIH Long-Term Database and MIT-BIH Normal Sinus Rhythm Database.

The main problems of the above-mentioned databases are the number of records, which is very limited (for example, the MIT-BIH arrhythmia database has 48 records) and mainly focused on arrhythmias, the number of leads (the MIT-BIH arrhythmia database contains two-lead ECG signals), and the length of the ECG (30 min each, which logically exceeds the normal duration of an ECG performed in a medical center). The advantage of these public databases is that all the registered diagnoses are validated. This allows for supervised modeling in novel classification techniques. Furthermore, it is not easy for researchers to access other more extensive ECG databases.

For the development of the presented system, an ECG database provided by the Preving Investment Company was used. This database contained around 284,000 anonymous ECGs (specifically 283,939 ECGs), stored in electronic files in SCP format (European standard EN1064 for communication and storage of information related to ECGs). In addition, the corresponding diagnosis associated with each of the ECGs (performed by the doctors of the company), the date of completion, and the age and gender of the patient were provided. The sample covered the time interval between 2006 and 2017. Each of the ECGs had a duration of 10 s with a sampling frequency of 500 Hz (sampling period = 2 msec), which involved a total of 5000 samples per ECG.

Table 1
Set of 15 diagnoses and number of cases in the database.

Diagnosis	Number of cases
Normal	212069
Artifacted or bad performance	1138
Incomplete right branch block	26375
Complete right branch block	2407
Incomplete right branch block with narrow QRS	Not registered
1st degree atrioventricular block	565
Wolff–Parkinson–White preexcitation	163
Complete arrhythmia due to atrial fibrillation	181
Long QT	Not registered
Short QT	Not registered
Sinus tachycardia	2248
Sinus bradycardia	21439
Nodal/ectopic atrial rhythm	Not registered
Sinus arrhythmia	Not registered
Cardiac arrhythmia	2164

Fig. 3 shows the distribution of the ECGs in the database with respect to ECGs registered with some type of disease versus nonpathological ECGs. As can be seen in the figure, 25% of the diagnoses were registered with some type of disease.

The main problem of this database was the existence of a large number of incorrect diagnoses for certain diseases. This was discovered quickly with ECG diagnoses that included simple rules or parameters as checks. For example, in sinus bradycardia, in which the heart rate must be below 60 beats per minute (bpm), a large number of ECGs diagnosed as such but with a value greater than 60 bpm could be found in the database. Subsequently it was verified with the help of the specialist that this was extrapolated to other diagnoses. A considerable number of diagnostic errors arose because the professional who made the diagnosis was often not a specialist but a general practitioner. In addition, there was the problem of the lack of unification of criteria among the different doctors involved. It was not possible for the specialist to filter incorrect diagnoses since the high ECG number did not allow one-to-one validation of the entire database. As a second problem, there were also diagnoses (such as long QT or nodal rhythm) that had not been registered in the database to date and therefore there were no ECG samples with these diagnoses.

The consequence of the above-mentioned problems was that the database could not be used for the supervised training of models (for example through deep learning with convolutional neural networks). This database, thanks to the large number of ECGs registered in it, was used as support for the development of the diagnostic system (for example to verify the noise filtering and identification of waves or to validate the outputs of the generated diagnostic rules).

Table 1 shows the set of 15 diagnoses (13 of them pathological) included in the system, together with the number of ECGs registered in the database for each case.

As can be seen, there are five diagnoses (incomplete right branch block with narrow QRS, long QT, short QT, nodal/ectopic atrial rhythm, and sinus arrhythmia) for which there were no ECGs with that diagnosis in the database. This is because these diagnoses had not been taken into account by the doctors who performed the ECGs.

4. Noise filtering and wave identification

For automatic signal processing corresponding to ECGs, it is necessary, as a starting point, to identify the different waves by means of algorithms that process the signals. But as was introduced in Section 2, a captured ECG signal is so weak that it is heavily contaminated with noise coupled in the form of interference from the power grid. In addition, the acquired signal is mixed with various low-frequency artifacts generated as a result of the patient's breathing and other causes. Hence, filtering of the noise before the identification of the different waves becomes highly complex, but totally necessary.

For the design of the noise filtering and wave identification in the diagnostic system, the set of stages shown in Fig. 4 was carried out.

Value	Proportion	%	Count
F		74.69	212069
T		25.31	71870

Fig. 3. Distribution of pathological ECGs versus nonpathological ECGs.

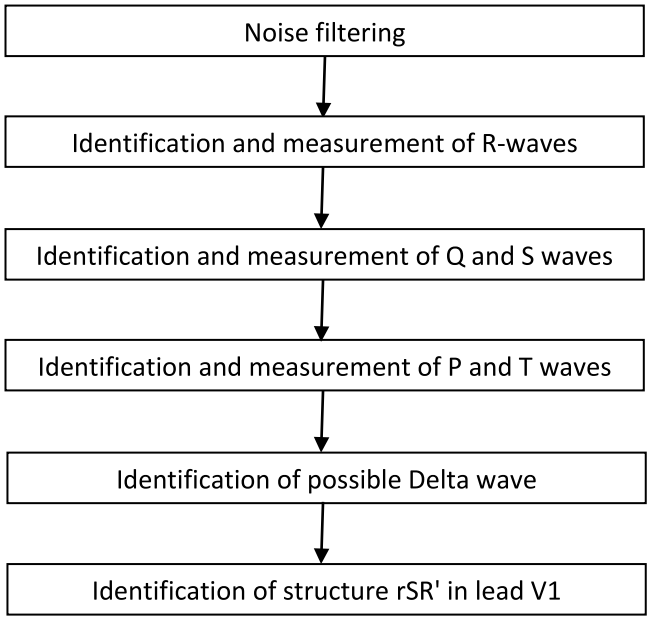


Fig. 4. Stages for the identification and measurement of waves.

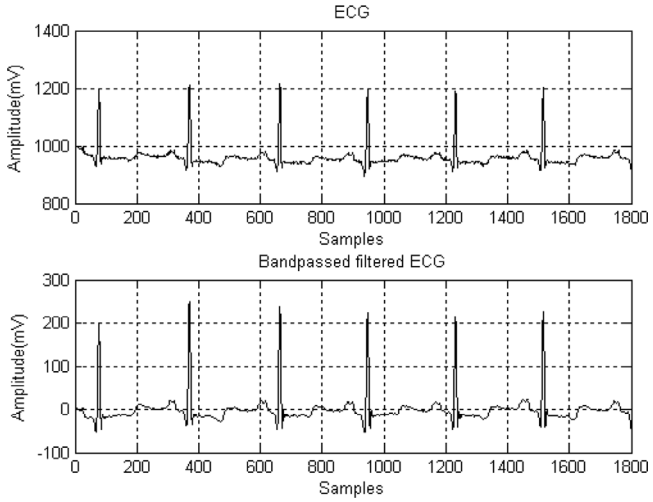


Fig. 5. Second-order Butterworth bandpass filter in ECG (Fazel-Rezai et al., 2011).

4.1. Noise filtering

Regarding noise filtering, it was essential in the first place to identify the different frequencies that can be found in the signal of an electrocardiogram (Haritha et al., 2016; Limaye and Deshmukh, 2016). The noise in ECGs contains the following frequency components:

Relating to the components of the ECG:

- Heart rate: 0.67–5 Hz (i.e. 40–300 bpm)
- P-wave: 0.67–5 Hz
- QRS component: 10–50 Hz
- T-wave: 1–7 Hz.

Relating to the artifact and noise on the ECG:

- Muscle contractions: 5–50 Hz
- Respiratory: 0.12–0.5 Hz (e.g. 8–30 bpm)
- External electrical: 50 or 60 Hz
- Electrode contact: 60 Hz.

Others: Typically >10 Hz (muscle stimulators, strong magnetic fields, pacemakers with impedance monitoring)

Taking into consideration the above frequencies, an algorithm was designed in order to filter the signal. This algorithm performs two successive filtrations:

In the first filtering stage, a second-order Butterworth (Fig. 5) bandpass filter is applied. This filter has been used and analyzed in previous work related to ECGs (Jagtap, 2012; Liu et al., 2018; Salsekar and Wadhwani, 2012). It is applied to remove frequencies below 0.5 Hz (corresponding to a baseline displacement and a modulation of ECG amplitude with respiration) and above 150 Hz (corresponding to the interference of the power line and noise by electrode contact).

The second signal filtering is based on a wavelet transform. The purpose of this filtering is to soften the original ECG signal (without removing valid frequencies of the different ECG components). This filtering is performed in two steps: First, a discrete wavelet transform (DWT) according to Mallat's pyramid algorithm (Mallat, 1989) is applied in order to obtain the wavelet coefficients. Subsequently, these coefficients are modified by applying signal smoothing (Nason and Silverman, 1995). Finally, the signal is reconstructed from the modified coefficients.

The filtering process is applied to the different leads of the ECG required for the identification of the diseases (which were specifically leads *I*, *II*, *V1*, *V5*, and *V6*). Since the noise of the ECG signals could not be filtered in its entirety (because the frequencies of the ECG overlap with ECG information), it was decided to generate a noise level indicator (called *NC50*). This indicator would allow the professional who performs the ECG to know if the level of noise has affected the proper processing by the ECG. Thus, the system would indicate, by means of the 'artifacted or bad performance' diagnosis, when the noise level exceeds what is advisable to perform a reliable diagnosis. It would enable the doctor to repeat the test. The design of the noise indicator is described in Section 4.2 since it uses the signal processing described in that section.

4.2. Identification and measurement of r-waves

For detecting the different R-waves in the ECG, wavelet transform (Rajini, 2016) was used. This transform allows frequency intervals to be detected and placed in time within the analysis window. It breaks down the signal into different levels, each of which covers a frequency range, and is composed of approximation coefficients (covering the lowest frequencies of that level) and detail (covering the highest frequencies). Fig. 6 shows these levels along with their frequency range for the signal sampled in the ECG. The original signal contains frequencies of up to 250 Hz since each ECG is sampled at 500 Hz (according to Nyquist theorem, the ECG can contain frequencies of up to half of the sampling).

It is shown that the R-wave is located at frequencies between 31.2 and 62.5 Hz (Haritha et al., 2016; Limaye and Deshmukh, 2016). Thus, in order to detect the R-wave, an algorithm was designed that performed a wavelet decomposition using the Haar transform function. The algorithm detects the peaks of the level 3 detail coefficients (D3) from a threshold value. The ECG lead used for the algorithm is *I*, which

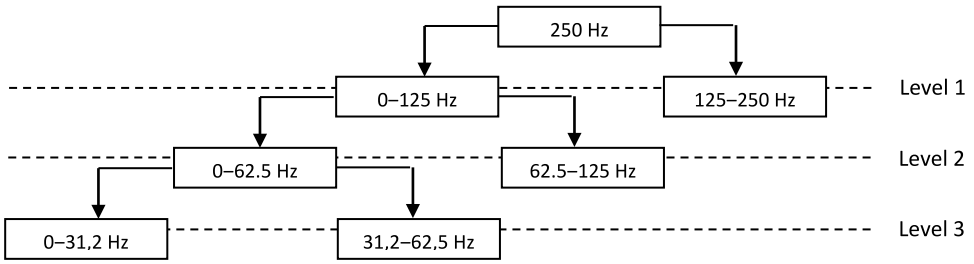


Fig. 6. Frequency range for each level of wavelet decomposition.

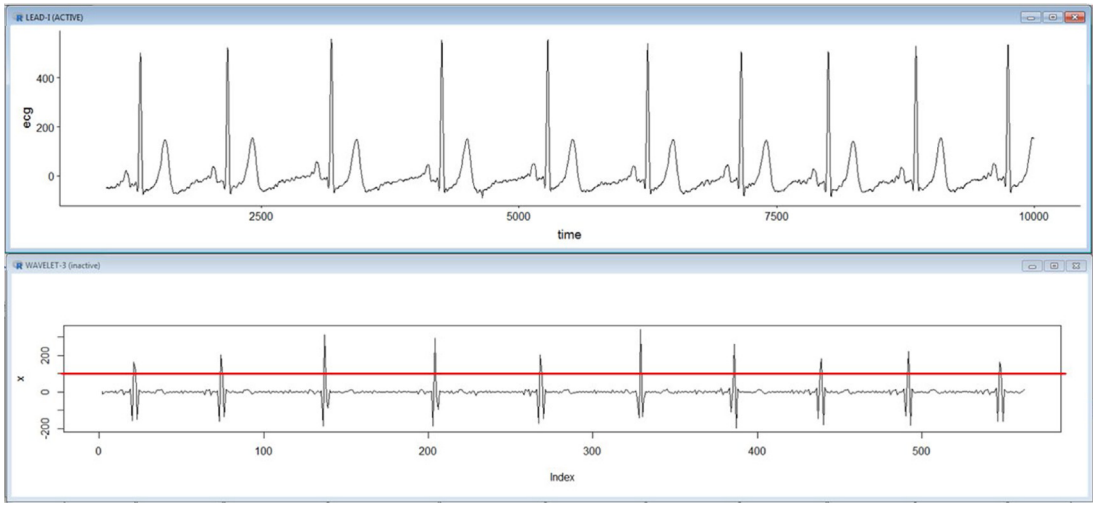


Fig. 7. Identification of R-waves.

was suggested by the specialist as the best lead for identifying R-waves. Fig. 7 shows an example of this detection process.

The threshold must be dynamic since the amplitude of a peak in the wavelet detail coefficients depends on the amplitude of the R-wave in the original signal. This amplitude may vary depending on the patient's heartbeat strength, his body fat, electrode placement, etc.

For the calculation of the threshold, a large battery of tests was carried out looking for the correspondence between the amplitude levels of the R-waves of the original signal and the amplitude obtained in the wavelet coefficients. After this analysis, the following calculation for the threshold was reached (with *maxR* being the maximum value of the signal):

```
threshold <- 0.25 * maxR - 25
if (maxR < 150)
    threshold <- 12.5
if (maxR > 900)
    threshold <- 200
```

This dynamic threshold allows the R-wave detection algorithm to work well at acceptable noise levels. The problem at this point is that an artifact ECG with a severe noise in the frequency spectrum of the R-wave produces additional peaks at this level of detail coefficients (Fig. 8). This means that the R-wave identification algorithm will not work correctly in those cases since the R-waves overlaps in frequency with the noise signal.

In order to control the above problem, the noise indicator *NC50* was created. This parameter evaluates the level of overlapping noise at frequencies that make it difficult to detect the R-waves. It is calculated as the number of times that the values of the detail coefficients cross, on the Y axis, 50% of the threshold value calculated for that signal and subtracting from this value the number of identified R-waves. Thus, *NC50* identifies the overlapping noise in the frequency range in which the R-waves are located in the ECG.

Fig. 9 shows the resulting histogram with normal curve of parameter *NC50* calculated for all ECGs in the database. A threshold of 40 was configured in order to identify the ECG as 'artifacted or bad performance.' Above this threshold there were 17,792 ECGs, which implies a filtration rate of 6.3% of the ECGs.

Additionally, as a noise indicator that also considers other leads, the number of R-waves in leads *II* and *V5* was calculated. This subsequently allows the system to filter based on the mismatch of the number of R-waves identified in lead *I* with the number of R-waves identified in leads *II* and *V5* (in both leads the R-waves also have to be perfectly identifiable). Thus, if they do not match, the ECG is filtered and identified as 'artifacted or bad performance.'

After applying the above filters to all the ECGs in the database, the total number of records filtered was around 10%.

4.3. Identification and measurement of P, Q, S, and t waves

Once the R-wave set for an ECG was identified by the system, the next step was the detection of the rest of the ECG waves. An algorithm was designed to detect, for each of the R-waves, its attached P, Q, S and T waves. In relation to the Q and S waves, the algorithm searches for the relative minimums attached to the R-wave that follows any of these possible QRS patterns (Fig. 10). Thus, it searches for a peak (this is done in lead *I* and *V5*) to the left of the R-wave (Q-wave) and a peak to the right (S-wave). It first finds the lower peak and then the end of the wave.

The end of a Q or S wave is not indicated by the crossing of the signal through the baseline, since it does not normally coincide, but a change in the slope of the signal. Thus, as can be seen in Fig. 11, the measurement of the S-wave covers up to the vertical line (which coincides with the change in signal slope). The threshold angle from

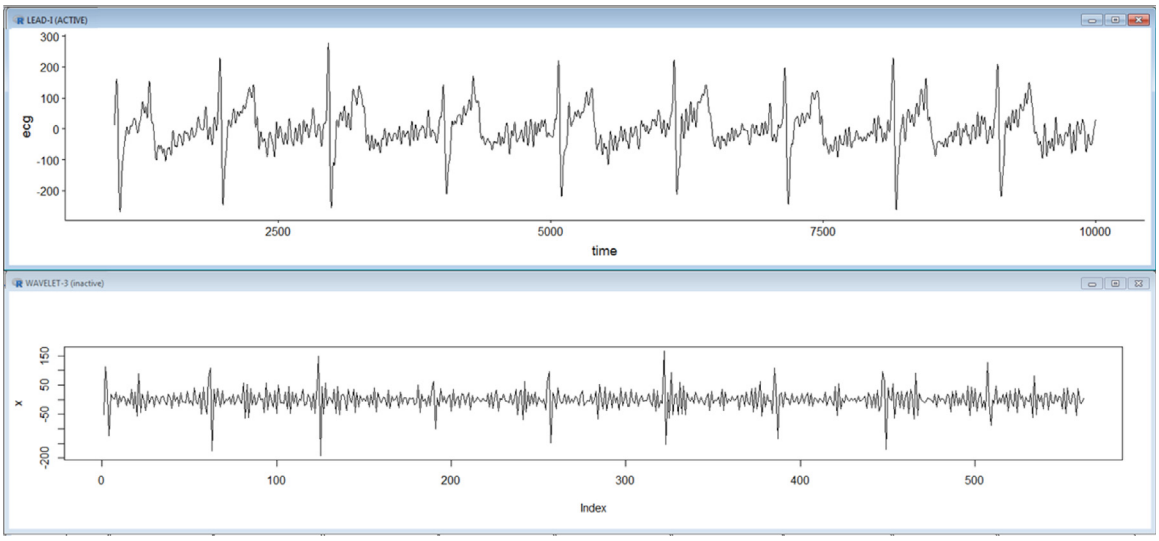


Fig. 8. Severe noise in ECG signal.

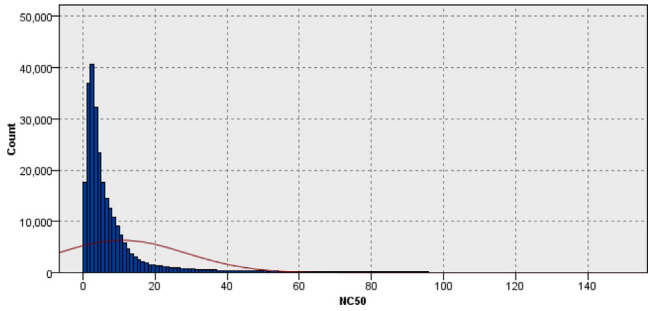


Fig. 9. Histogram of parameter *NC50* for all ECGs in the database.

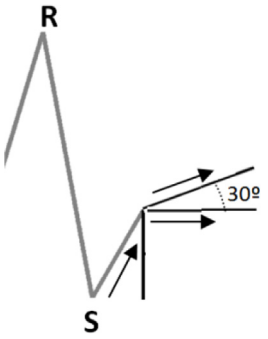


Fig. 11. Detection of slope change in S-wave.

the baseline that marks the end of the wave was configured at 30° for S and −30° for Q. This value was set after an extensive battery of tests to find the optimal value.

According to the specialist's recommendation, the values related to the Q-wave are calculated by the algorithm for lead *I* but also for lead *V5*, since in certain ECGs that wave is not observed in lead *I*.

The next step that the algorithm carries out is the detection of the peaks relative to the P and T waves. This detection is performed by searching for relative maximums in two time intervals of the filtered signal (Fig. 12). Specifically for the P-wave, the 40% of the interval from the start of the S-wave of the previous cycle to the Q-wave of the analyzed pulse is searched for the relative maximum of the signal. On the other hand, for the T-wave, the remaining 60% interval marked in the figure (which starts at the end point of the S-wave) is searched for the relative maximum of the signal.

The designed algorithm cannot detect negative P and T waves (which may appear in some ECGs). However, it is not necessary for the 13 diseases diagnosed by the system.

Once the peaks of the P and T waves have been found, the algorithm find the beginning and end of the P and T waves respectively in order to calculate their widths. This algorithm searches for a slope change in the decrease of the wave value from the calculated maximums (Fig. 13). The threshold angle from the baseline for the start and end of the P and T waves was configured at 70°.

The calculation corresponding to the P-wave is also performed in lead *II*, since in some ECGs this wave can be seen in lead *II* but not in lead *I*. Moreover, the processing of the S-wave, in addition to lead *I*, is also performed in lead *V6* as a condition for the diagnosis of certain diseases (specifically both right branch blocks).

Once the P, Q, S, and T waves of the ECG have been identified, the following parameters were calculated by the algorithm: QRS width (msec), QT interval (msec) and PR interval (msec).

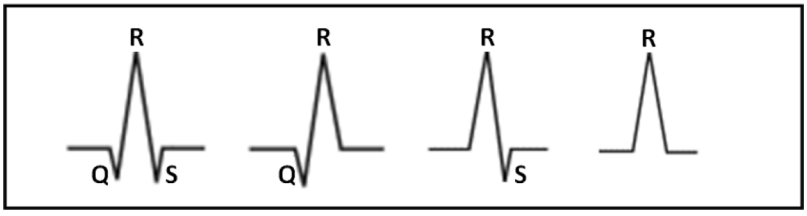


Fig. 10. Possible QRS patterns to be detected.

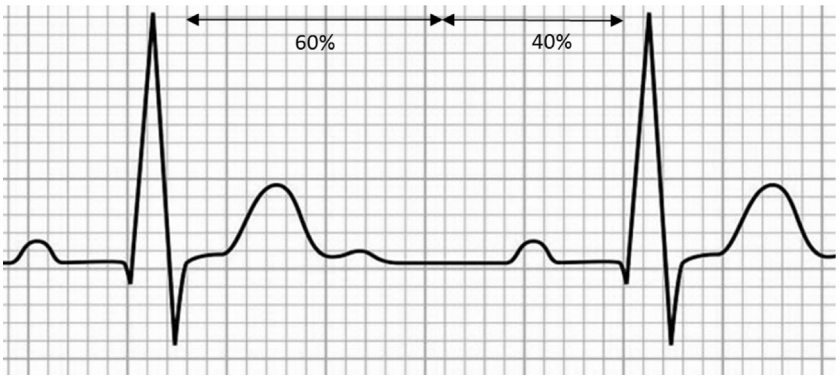


Fig. 12. Intervals for the detection of P and T waves.

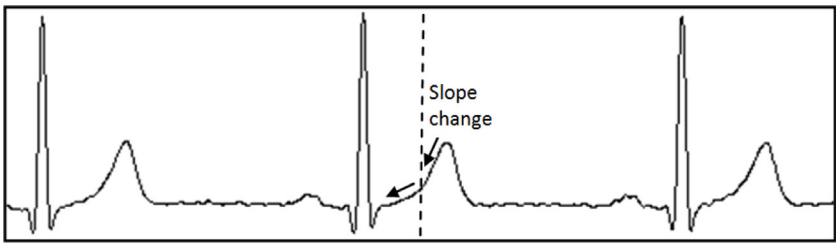


Fig. 13. Detection of slope change in T-wave.

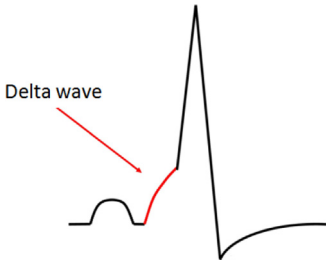


Fig. 14. Delta wave in ECG.

4.4. Identification of possible Delta wave

The Delta wave is an additional wave that may appear on the ECG due to the appearance of Wolff–Parkinson–White preexcitation. This disease is a preexcitation syndrome of the ventricles of the heart due to an accessory pathway known as the Kent beam. Thus, the realization of this diagnosis implies the identification of this wave, which can appear as an extension of the R-wave, on the left side of it (Fig. 14).

For its detection, an algorithm was designed to identify the appearance or not of this Delta wave, as well as to calculate its width. For the detection of said wave, the wavelet coefficients of level 3 are used in lead *I*. The Delta wave does not appear in the detail coefficients at level 3 because it is not in the range of frequencies that they detect. Hence, to perform the detection, an additional descent, with respect to that marked by the wavelet detail coefficients at its crossing of the baseline, is identified by the algorithm (Fig. 15). Thus, the Delta wave is identified by measuring that width of this part of the signal.

4.5. Identification of structure rSR' in lead V1

For diagnosis of some of the most common heart diseases (including the right branch blocks), the detection of the appearance of an rSR' structure in the lead *V1* is necessary. The structure rSR' (or, which is the same, the stocks of double wave R) follows the pattern shown in Fig. 16. However, in most cases the highs of both R-waves are different.

Thus, for identifying the rSR' structure, an algorithm was designed to count, within the QRS complex found in lead *I*, the number of peaks (or R-waves) existing in lead *V1*, identifying the intermediate S-wave (whose peak is below the baseline). Fig. 17 shows a representation of the window analysis performed for the rSR' detection.

The number of R-waves in lead *V1* is obtained by calculating the number of relative maximums with intermediate S-waves. An intermediate S-wave is identified as a signal drop of at least 50% of the previous R-wave or below the baseline.

5. Rule-based system design

Once it had been validated that the different ECG waves are detected by the system (Fig. 18) and the necessary intervals and parameters had been measured, the design of the kernel of the diagnostic system was carried out.

The design of the system is based on rules (as cardiologists and ECG specialists work when making a diagnosis). It was not considered a kernel designed from supervised training (except for a specific diagnosis that will be described below) due to the low reliability for the diagnoses registered in the ECG database. In addition, there were diagnoses (such as long QT and short QT) that had not been considered in the past by doctors who performed the ECGs and, therefore, were not registered in the database.

For the modeling process, the work and knowledge of the ECG diagnosis specialist were indispensable. Thus, a document describing the diagnoses was provided by the specialist. This document contains the following information for each disease:

- Necessary ECG parameters to be measured (intervals, waves, morphologies...).
- Conditions and rules for its diagnosis (required ECG leads, key observations...).

By doing this, and with the permanent support of the specialist, the different diagnostic rules were modeled. The parameters involved in the rule conditions are described in Table 2. These parameters are calculated by the system from the noise filtering and wave identification process (see Section 4) of the analyzed ECG. For 12 of the

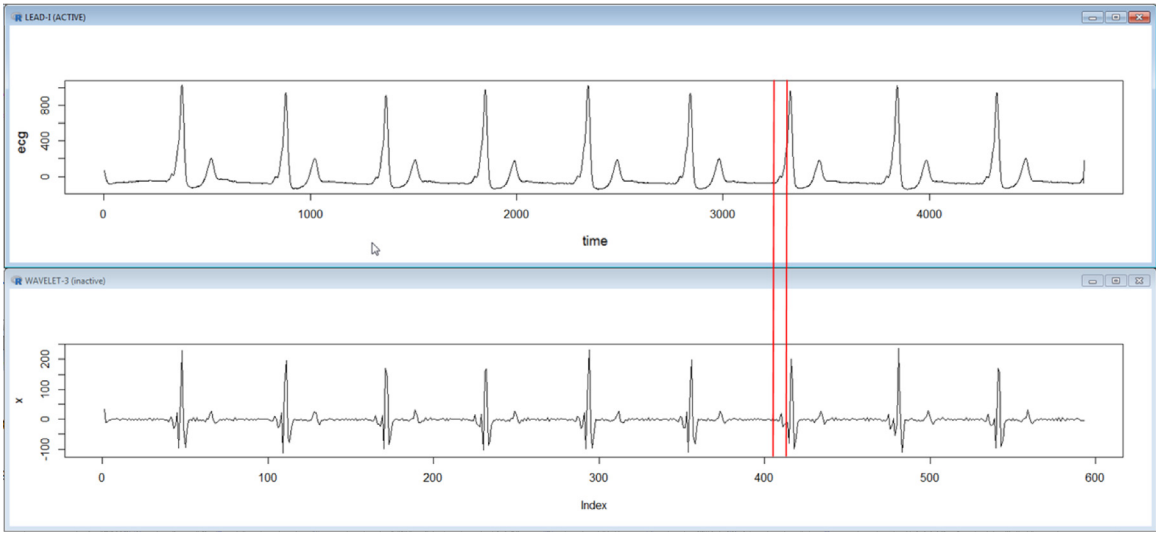


Fig. 15. Descent in Delta wave from the crossing of the wavelet coefficients on the X axis.

13 pathological diagnoses (all but 'complete arrhythmia due to atrial fibrillation'), the rules were designed from the information collected in the specialist's document and subsequently validated with the approval of the specialist. The ECGs in the database in their entirety were used as support for the design process of the expert system (first, verify correct noise filtering; second, check the correct identification of waves by the system; and third, test the output of the generated rule).

For the remaining diagnosis ('complete arrhythmia due to atrial fibrillation'), the diagnostic rules for this disease are not too clear. Thus, for example, it is known that the intervals between R waves must be irregular, but this irregularity is not defined. It was decided to ask the specialist to generate, from the database, a validated sample of ECGs with said disease. The objective for this diagnosis in particular was to perform a supervised detection algorithm. To generate the sample, the specialist took the 181 ECGs diagnosed with this disease (Table 1), filtering the incorrect diagnoses, generating a sample of 144 ECGs. From this sample and the rest of the ECGs in the database, a decision tree was designed for the diagnosis of this disease (Fig. 19). The chosen algorithm was the CHAID (Chi-squares Automatic Interaction Detection), which provided the best results. The model was generated with IBM Modeler 18 which used 30% of data for testing. From the decision tree, a rule was extracted that covered 91.6% of the ECGs with this disease and that was supported by 843 ECGs. It is important to emphasize that the diagnostic errors of the system in this disease imply some other disease related to some type of arrhythmia and they are not diagnosed as nonpathological. This implies that these false-positive errors are not critical.

The diagnostic rules for all of the 13 diseases covered are detailed in Table 3. Those ECGs not identified as 'artifacted or bad performance' and not detected by some of the 13 rules were diagnosed as 'normal' (although they included, in addition to nonpathological ones, those ECGs with any of the diseases not detected by the system).

The adjustment of the parameters of the rules was performed sequentially by means of an extensive battery of tests and with the support of the specialist. In many rules in which percentages are used (for example: *P_exv2_porc <25*), certain thresholds were set (in the previous case 25). It avoids errors in the detection of a disease due to the incorrect identification in any cardiac cycle of any of the waves.

At each step of the design, all of the ECGs (specifically 283,939 ECGs) were processed in order to verify the correct functioning of the system and to validate those parameter settings. For the generation and adjustment of the rules, the IBM Modeler 18 software was used. This tool is designed for data mining processes and allows, in a fast and complete way, the processing of large databases. Once the diagnostic rules had been designed, the expert system was programmed entirely in R language.

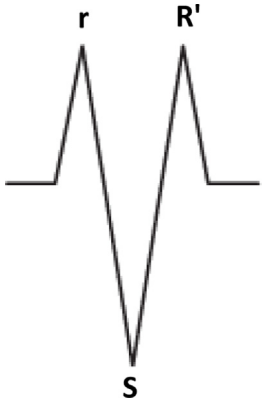


Fig. 16. rSR' structure in ECG.

With regard to temporal parameters, the total processing time for the entire ECG database was 9 h and 53 min. This time involves about 130 msec per processed ECG. For testing it was used with an Intel Core i5 9400 desktop PC with 16 GB of RAM.

6. Results

Once the system had been designed and implemented, it was validated with the help of the specialist. For this, for each of the diseases, a set of diagnoses made by the system on ECGs randomly selected from the database was sent to the specialist. Later the specialist returned a report with the correct and incorrect diagnoses made by the system, and with that data the reliability calculation was made.

The diagnoses used for validation were classified into three types:

- Type 1: Cases registered in the database with the disease that the system confirms with that disease.
- Type 2: Cases diagnosed in the database as 'normal' (nonpathological) but that the system detects with that disease.
- Type 3: Cases diagnosed in the database with the disease and that the system detects that the diagnosis is not correct.

This classification in each of the diagnoses was adequate for the objective of validation of the present work. In addition, it allows

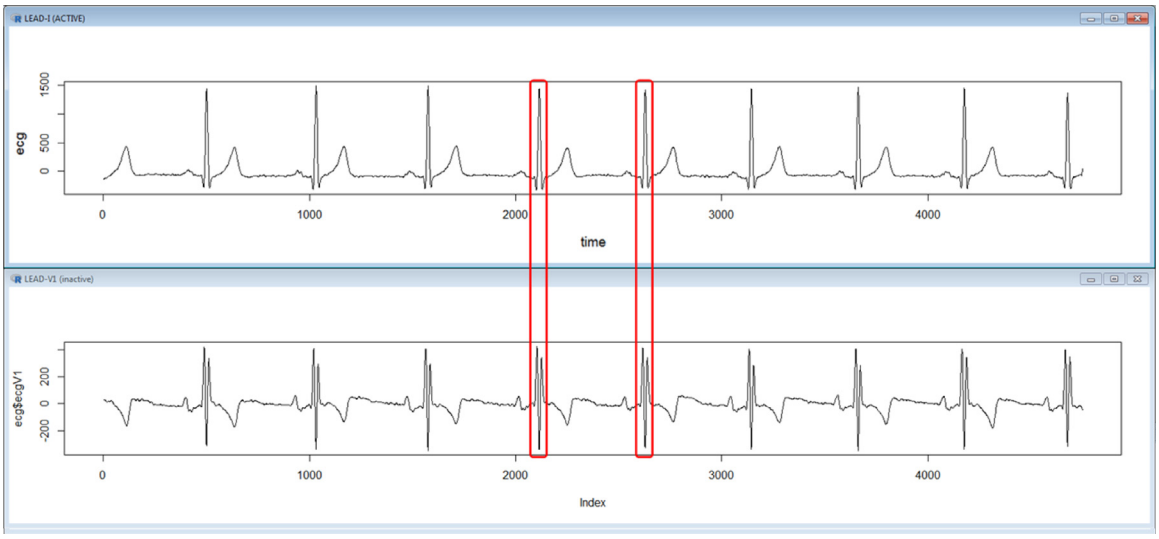


Fig. 17. Detection of rSR' in lead V1.

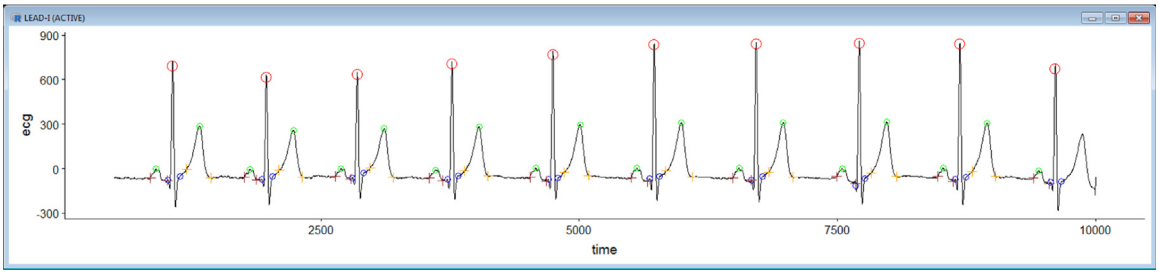


Fig. 18. ECG with the different waves identified.

Table 2
Parameters used for diagnostic rules.

Parameter	Description
HR	Heart rate (calculated on the average time between the different R-waves).
P_exv2_porc	Percentage of P-waves in lead I over total cardiac cycles.
P_rv_mean	Mean value (mV) of the amplitudes of the P-wave in lead *I*.
PD2_exv2_porc	Percentage of P-waves in lead *II* over total cardiac cycles.
PP_msec_max	Maximum value (msec) of the intervals between P-waves.
PP_msec_min	Minimum value (msec) of the intervals between P-waves.
PR_exc_porc	Percentage of PR intervals in lead I that are below 120 msec.
PR_exl_porc	Percentage of PR intervals in lead I that are above 200 msec.
PR_msec_median	Median value (msec) of the PR intervals in lead *I*.
QRS110_ex_porc	Percentage of QRS components whose width is below 110 msec.
QRS1120_ex_porc	Percentage of QRS components whose width is between 100 and 120 msec.
QRS120_ex_porc	Percentage of QRS components whose width is above 120 msec.
QT_corr_long_porc	Percentage of QT intervals that are above normal values for a heart rate. The intervals were calculated using the normal value tables of Lepeschkin/Ashman.
QT_corr_short_porc	Percentage of QT intervals that are below normal values for a heart rate. The intervals were calculated using the normal value tables of Lepeschkin/Ashman.
RnV1_porc	Percentage of rSR' complexes over the total cardiac cycles in lead *V1*.
RR_msec_max	Maximum value (msec) of the intervals between R-waves.
RR_msec_min	Minimum value (msec) of the intervals between R-waves.
RR10_mean	Number of R-waves whose distances deviate at least 10% from the average of the RR intervals.
RR10_mm	Number of R-waves whose distances deviate at least 10% from the maximum/minimum of the RR intervals.
RR20_mean	Number of R-waves whose distances deviate at least 20% from the average of the RR intervals.
RR20_median	Number of R-waves whose distances deviate at least 20% from the median of the RR intervals.
RR20_mm	Number of R-waves whose distances deviate at least 20% from the maximum/minimum of the RR intervals.
S_e_porc	Percentage of occurrences of S waves in lead *I* over total cardiac cycles.
SeV6_porc	Percentage of occurrences of S waves in lead *V6* over total cardiac cycles.
T_ex_porc	Percentage of T-waves in lead I over total cardiac cycles.
W_delta_max	Maximum width (number of samples) of the different Delta waves in lead *I*.

(against, for example, a confusion matrix) information to be obtained regarding the quality of the diagnoses registered in the database.

The number of cases sent to the specialist for each disease depended on the diagnosis in question as well as its frequency of appearance

Table 3
Rules for the 13 pathological diagnoses of the system.

Disease	Rule
Complete arrhythmia due to atrial fibrillation	*(RR20_mean>0) and (P_exv2_porc<=50) and (RR10_mm>2) and (P_rv_mean<=1.32)*
1st degree atrioventricular block	*(PR_exl_porc>25) and (PR_msec_median>170) and (QRS120_ex_porc<=25) and ((PD2_exv2_porc>25) or (P_exv2_porc>25))*
Wolff–Parkinson–White preexcitation	*(QRS110_ex_porc<40) and (PR_exc_porc>60) and (W_delta_max>5)*
Long QT	*(QT_corr_long_porc>75) and (T_ex_porc>75)*
Short QT	*(QT_corr_short_porc>75) and (T_ex_porc>75)*
Nodal/ectopic atrial rhythm	*(RR10_mean==0) and (P_exv2_porc<25) and (PD2_exv2_porc<25) and (QRS120_ex_porc<75) and (HR>=40) and (HR<=70)*
Incomplete right branch block	*(RnV1_porc>50) and (SeV6_porc>50) and (S_e_porc>50) and (QRS1120_ex_porc>25) and (QRS1120_ex_porc>=QRS120_ex_porc)*
Complete right branch block	*(RnV1_porc>50) and (SeV6_porc>50) and (S_e_porc>50) and (QRS120_ex_porc>25)*
Incomplete right branch block with narrow QRS	*(RnV1_porc>50) and (SeV6_porc>50) and (S_e_porc>50) and (QRS1120_ex_porc<=25)*
Sinus tachycardia	*(HR>100) and ((P_exv2_porc>25) or (PD2_exv2_porc>25)) and (RR20_mean==0) and (QRS120_ex_porc<=75) and (PR_exl_porc<75)*
Sinus bradycardia	*(HR<60) and ((P_exv2_porc>25) or (PD2_exv2_porc>25)) and (RR20_mean==0) and (QRS120_ex_porc<=75) and (PR_exl_porc<75)*
Sinus arrhythmia	*((PP_msec_max-PP_msec_min>=120) or (RR_msec_max-RR_msec_min>=120)) and ((P_exv2_porc>25) or (PD2_exv2_porc>25)) and (QRS120_ex_porc<=40) and (PR_exl_porc<75)*
Other cardiac arrhythmias	*(RR20_mm>0) or (RR20_mean>0) or (RR20_median>0)*

in the database. The number of cases had to be limited due to the significant time spent by the specialist validating an ECG and carrying out a diagnosis.

For the selection of the ECG sample for validation, random sampling for each disease was carried out using the subset of ECGs corresponding to the last two years of the sample (2016 and 2017). The results of the validation for each of the diagnoses in the final version of the system are shown in Table 4. These results include both the 12 diagnoses of the expert system and the complete arrhythmia due to atrial fibrillation diagnosis generated from the decision tree. For complete arrhythmia due to atrial fibrillation, the result is marked with *. It refers to the confidence of the rule on the total of the sample selected for that diagnosis (since for that disease the total number of cases registered in the database was used). The percentage obtained for these conditions and number of cases is considered excellent. On the other hand, in those diseases whose results are linked to the three types in the table, there were no cases previously identified in the database.

The results obtained by the system in its validation were 97.51% on average in the percentage of correct answers in type 1 diagnoses, 60% on average in type 2 and 78.8% in type 3. It is necessary to bear in mind that the most significant results are focused on type 1, since a good percentage implies achieving at least the reliability of the diagnoses of the doctors who performed the ECGs. On the other hand, in relation to diagnoses without previous registration in the database (and therefore not validated in the three types), the average percentage of success was 82.5%.

As a means of reliability, taking into account all cases sent to the specialist for validation and excluding those related to type 1 for 'complete arrhythmia due to atrial fibrillation' (which were really previously validated data), the system correctly diagnosed a total of 135 cases of 167 pathological ECGs sent to be checked by the specialist. This implies 80.8% of average reliability in the diagnosis of diseases. Thus, with the previous numbers, the validation results of the system can be considered good enough in the context of a high-complexity problem.

7. Conclusions and discussion

The paper describes the development of an expert diagnostic system for standard 12-lead ECGs. The system includes a total of 15 diagnoses (13 of them pathological). A database of 283,939 anonymous ECGs compiled by the Preving Investment Company was used as support for the design and validation process. In addition, a specialist with extensive experience in ECG diagnosis was available during the development of the work.

A total of 26 parameters were generated for the feature extraction of the ECG. The kernel of the system consists of a set of 13 rules (one per pathological diagnosis) that were generated with the collaboration of the specialist and using as support the data recorded in the database.

Furthermore, an algorithm was designed to filter the noise from the signals of the ECG leads used (*I*, *II*, *V1*, *V5*, and *V6*). Since total noise filtering is not possible when the frequency overlaps with the ECG information, a noise indicator was generated. This indicator allows the professional performing the ECG to know its noise level.

For the validation process, a set of diagnoses made by the system for each modeled disease were sent to the specialist. In terms of the reliability of the system, it correctly diagnosed a total of 135 cases out of a total of 167 pathological ECGs. This is 80.8% effective in diagnosing diseases. With this success rate and given the extreme complexity of the problem, the results are considered adequate to make the system a useful support tool for diagnosis.

Research on automatic ECG diagnosis is currently widespread, with many supervised techniques (for example with convolutional neural networks). These techniques are powerful but require a large amount of reliable data for training. The main problem is that the diagnoses registered in an extensive database contain errors or depend on the judgment of the professional who made the diagnosis. The present work adds an alternative approach to the majority of the studies found in the literature. The contributions of this work are the following:

- A fully functional expert system that performs diagnoses following the approach of a specialist. The set of rules in the system is concise and based on electrocardiography parameters.
- A large number of diseases are covered by the system with a high reliability rate. Most of the studies found in the literature are focused on detecting arrhythmias. There is no study that includes 13 different diseases.
- The use of 5 leads of a standard 12-lead ECG to perform the diagnosis. Most published studies use a single-lead ECG. This implies that the diagnostic possibilities are limited since certain diseases require several leads for their diagnosis.
- A novel noise indicator that measures the quality of the acquired ECG signal. This allows the user to repeat the ECG if its noise level is excessive. In this way, the system provides a solution for those ECGs whose noise cannot be filtered.
- The unification of criteria for the diagnosis of ECGs. This provides a better organization of the Preving Investment Company's database. In addition, the system allows the company to detect previously undiagnosed diseases.

Currently, the Preving Investment Company's medical staff is testing the system as a support for diagnoses. The company's intention for the future is to expand the project to include all heart diseases.

Table 4
Validation results for the system.

Disease	Type 1	Type 2	Type 3
Complete arrhythmia due to atrial fibrillation	132/144 (91.6%)*	2/10 (20%)	N/A
1st degree atrioventricular block	7/7 (100%)	6/7 (85.7%)	6/6 (100%)
Wolff–Parkinson–White preexcitation	4/4 (100%)	0/4 (0%)	2/4 (50%)
Long QT		7/10 (70%)	
Short QT		4/5 (80%)	
Nodal/ectopic atrial rhythm		5/10 (50%)	
Incomplete right branch block	10/10 (100%)	10/10 (100%)	9/10 (90%)
Complete right branch block	6/7 (85.7%)	3/7 (42.8%)	5/6 (83.3%)
Incomplete right branch block with narrow QRS	10/10 (100%)	10/10 (100%)	9/10 (90%)
Sinus tachycardia	7/7 (100%)	7/7 (100%)	4/6 (66.6%)
Sinus bradycardia	7/7 (100%)	6/7 (85.7%)	4/6 (66.6%)
Sinus arrhythmia		10/10 (100%)	

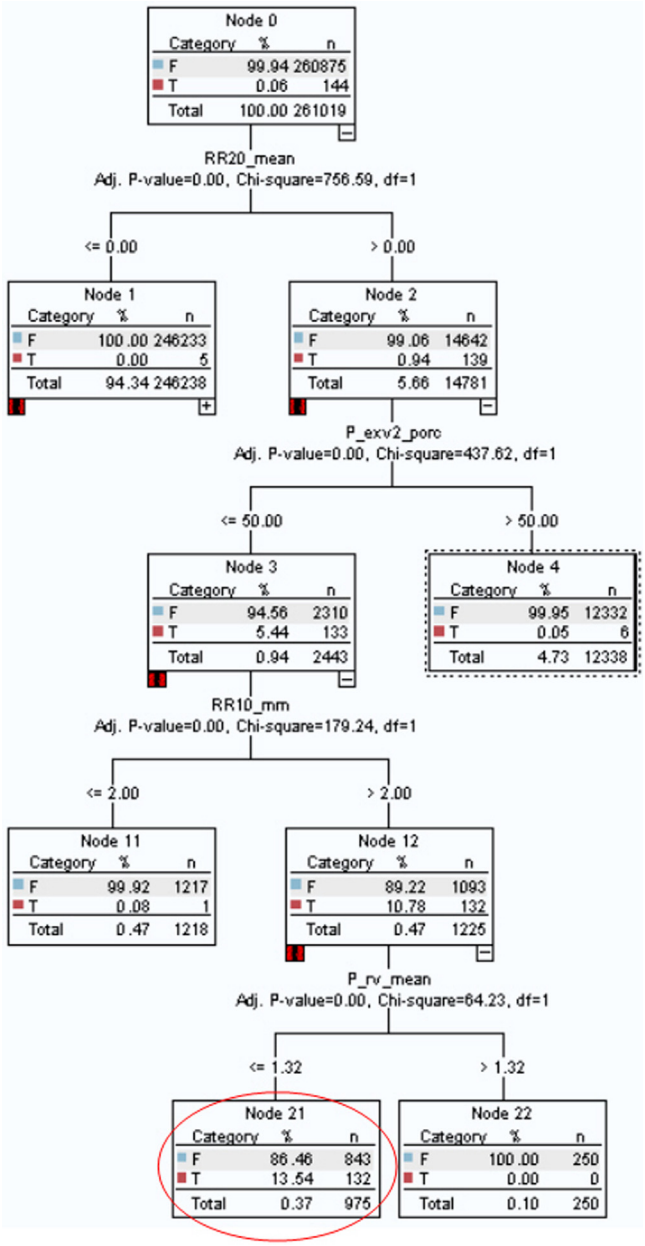


Fig. 19. Decision tree for 'complete arrhythmia due to atrial fibrillation'.

CRediT authorship contribution statement

Iñigo Monedero: Conception and design, or analysis and interpretation of the data, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was carried out in the framework of the project ECARDIO, funded by the Preving Investment Company. The author thanks Jesús Nieto and the Preving Investment staff for the collaboration provided, as well as for the facilities at the time of publishing this work. Furthermore, the author would like to thank Fernando Martínez for his invaluable work as a specialist in electrocardiogram diagnosis and for his collaboration throughout the research process.

References

Adam, M., Oh, S.L., Sudarshan, V.K., 2018. Automated characterization of cardiovascular diseases using relative wavelet nonlinear features extracted from ECG signals. Comput. Methods Programs Biomed. 161, 133–143.

Alarsan, F.I., Younes, M., 2019. Analysis and classification of heart diseases using heartbeat features and machine learning algorithms. J. Big Data 6.

Altay, Y.A., Kremlev, A.S., Margun, A.A., 2019. ECG signal filtering approach for detection of P, QRS, T waves and complexes in short single-lead recording. IEEE Conf. Russ. Young Res. Electr. Electron. Eng. 1135–1140.

Anderson, J., DiCarlo, S.E., 2000. Virtual experiment for understanding the electrocardiogram and the mean electrical axis advances in physiology education. 23 (1), 1–17.

Berkaya, S.K., Uysal, A.K., Gunal, E.S., Ergin, S., Gunal, S., Gulmezoglu, M.B., 2018. A survey on ECG analysis. Biomed. Signal Process. Control 43, 216–235.

Chun-Cheng, L., Hung-Yu, C., Yan-Hua, H., Cheng-Yu, Y., 2019. A novel wavelet-based algorithm for detection of QRS complex. Appl. Sci. 9 (10), 1–19.

Fazel-Rezai, Reza, Root, Noah, Rabbi, Ahmed, Lee, DuckHee, Ahmad, Waqas, ECG signal processing: A practical approach. Biomedical Engineering - From Theory to Applications. IntechOpen, https://www.intechopen.com/books/2241.

Golrizkhatami, Z., Acan, A., 2018. ECG classification using three-level fusion of different feature descriptors. Expert Syst. Appl. 114, 54–64.

Haddadi, R., Abdelmounim, E., Belaguid, A., 2014. Discrete Wavelet Transform based algorithm for recognition of QRS complexes. In: International Conference on Multimedia Computing and Systems. ICMCS, pp. 375–379.

Haritha, C., Ganesan, M., Sumesh, E.P., 2016. A survey on modern trends in ECG noise removal techniques. In: International Conference on Circuit, Power and Computing Technologies. ICCPCT, pp. 1–7.

Jagtap, S.K., 2012. The impact of digital filtering to ECG analysis: Butterworth filter application. In: International Conference on Communication, Information & Computing Technology. ICCICT, pp.1–6.

Jun, T.J., Nguyen, H.M., Kang, D.Y., 2018. ECG arrhythmia classification using a 2-D convolutional neural network. In: Proceedings of the Computer Vision and Pattern Recognition. pp. 1–22.

Kanungo, B., Sabut, S.K., 2015. Feature extraction of ECG signal based on wavelet transform for arrhythmia detection. In: IEEE International Conference on Electrical, Electronics, Signals, Communication and Optimization. EESCO, pp. 1–5.

Kasar, S.L., Joshi, M.S., 2016. Analysis of multi-lead ECG signals using decision tree algorithms. Int. J. Comput. Appl. 134 (16), 27–30.

Li, H., Yuan, D., Wang, Y., Cui, D., Cao, L., 2016. Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system. Sensors 16 (10), 1–16.

Li, T., Zhou, M., 2016. ECG classification using wavelet packet entropy and random forests. Entropy 18 (8), 1–16.

Limaye, H., Deshmukh, V.V., 2016. ECG noise sources and various noise removal techniques: A survey. Int. J. Appl. Innov. Eng. Manag. 5 (2), 86–92.

Liu, M., Hao, H., Xiong, P., Lin, F., Hou, Z., Liu, X., 2018. Constructing a guided filter by exploiting the butterworth filter for ECG signal enhancement. J. Med. Biol. Eng. 38, 980–992.

Mallat, S., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. 11, 674–693.

Martis, R.J., Acharya, U.R., Adeli, H., 2014. Current methods in electrocardiogram characterization. Comput. Biol. Med. 48, 133–149.

Mathews, S., Kambhamettu, C., Barner, K.E., 2018. A novel application of deep learning for single-lead ECG classification. Comput. Biol. Med. 99, 53–62.

MIT-BIH Database and Software Catalog, 2019. http://ecg.mit.edu/dbinfo.html.

Nason, G.P., Silverman, B.W., 1995. The stationary wavelet transform and some statistical applications. In: Lecture Notes in Statistics, vol. 103, pp. 281–299.

Nayak, S.K., Banerjee, I., Pal, K., 2019. Electrocardiogram signal processing-based diagnostics: applications of wavelet transform. In: Bioelectronics and Medical Devices: From Materials To Devices-Fabrication, Applications and Reliability, Vol. 24. pp. 591–614.

Pingale, S.L., Daimiwal, N., 2014. Detection of various diseases using ECG signal in matlab. Int. J. Recent Technol. Eng. 3 (1), 1–5.

Rajini, G.K., 2016. A comprehensive review on wavelet transform and its applications. ARPN J. Eng. Appl. Sci. 11 (19), 713–723.

Sahoo, S., Kanungo, B., Behera, S., Sabut, S., 2017. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. Measurement 108, 55–66.

Salsekar, B., Wadhwani, A., 2012. Filtering of ecg signal using butterworth filter and its feature extraction. Int. J. Eng. Sci. Technol. 4, 1292–1298.

Samol, A., Bischof, K., Luani, B., Pascut, D., Wiemer, M., Kaese, S., 2019. Single-lead ECG recordings including einthoven and wilson leads by a smartwatch: A new era of patient directed early ECG differential diagnosis of cardiac diseases. Sensors 19 (20), 1–11.

Shao, M., Bin, G., Wu, S., Bin, G., Huang, J., Zhou, Z., 2018. Detection of atrial fibrillation from ECG recordings using decision tree ensemble with multi-level features. Physiol. Meas. 39 (9), 1–16.

Sharma, M., Swapnil, S., Kumar, A., 2019. Automated detection of shockable and non-shockable arrhythmia using novel wavelet-based ECG features. Comput. Biol. Med. 115, 1–10.

Singh, R., Mehta, R., Rajpal, N., 2018. efficient wavelet families for ECG classification using neural classifiers. In: Proceedings of the International Conference on Computational Intelligence and Data Science. Vol. 132, pp. 11–21.

Singh, O., Sunkaria, R.K., 2017. Ecg signal denoising via empirical wavelet transform. Australas. Phys. Eng. Sci. Med. 40, 219–229.

Tripathy, R.K., Dandapat, S., 2016. Detection of cardiac abnormalities from multilead ECG using multiscale phase alternation features. J. Med. Syst. 40 (6), 1–9.

Wijaya, C., Andrian, Harahap, M., Christnatalis, Turnip, M., Turnip, A., 2018. Abnormalities state detection from P-wave, QRS complex, and T-wave in noisy ECG. J. Phy.: Conf. Ser.2nd Inter. Conf. Mech. Electron. Comput. Indus. Technol. 1230, 1–9.

Yildirim, O., Pławiak, P., Tan, R.S., U.R., Acharya, 2019. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput. Biol. Med. 102, 411–420.