Automated Atrial Fibrillation Detection by ECG Signal Processing: A Review*

Jader Giraldo-Guzmán,a, Sonia H. Contreras-Ortiz,a Marian Kotas,b Francisco Castells,c & Tomasz Morońb

aFaculty of Engineering, Universidad Tecnológica de Bolívar, Cartagena de Indias, Colombia
bDepartment of Cybernetics, Nanotechnology and Data Processing, Silesian University of Technology, Gliwice, Poland
cInstituto ITACA, Universitat Politècnica de València, Spain


ABSTRACT: Cardiovascular diseases are the main cause of death in the world, according to the World Health Organization. Among them, ischemic heart disease is at the top, followed by a stroke. Several studies have revealed that atrial fibrillation (AF), which is the most common cardiac arrhythmia, increases up to five fold the overall risk of stroke. As AF can be asymptomatic, approximately 20% of the AF cases remain undiagnosed. AF can be detected by analyzing electrocardiography records. Many studies have been conducted to develop automatic methods for AF detection. This paper reviews some of the most relevant methods, classified into three groups: analysis of heart rate variability, analysis of the atrial activity, and hybrid methods. Their benefits and limitations are analyzed and compared, and our beliefs about where AF automatic detection research could be addressed are presented to improve its effectiveness and performance.

KEY WORDS: atrial fibrillation detection, ECG signal processing

I. INTRODUCTION

For more than 15 years, cardiovascular diseases have been the leading cause of death in the world. At least 75% of these deaths happen in developing countries. Cardiovascular diseases represent an economic burden that can reduce the gross domestic product (GDP) up to 6.77%. Atrial fibrillation (AF) is the most common cardiac arrhythmia affecting between 1% and 2% of world population. It increases the risk of serious diseases, such as stroke, heart failure, renal disease, or sudden death. A study by Wolf et al. revealed that AF increases five times the incidence of stroke, which is the third cause of death in the world according to the World Health Organization (WHO).

AF is caused by problems in the synchronization of atrial activity due to abnormal electrical impulses that cause irregular contractions. As a result, blood remains in the heart, which increases the risk of clotting or stroke. Studies reveal that 20% of the AF cases remain undiagnosed. This condition is more common in the elderly. During AF episodes, the P wave cannot be detected or is replaced by irregular fibrillatory waves (F waves) and heart rate tends to vary.

AF can be classified in five categories: first diagnosed AF, paroxysmal AF, persistent AF, long-standing persistent AF, and permanent AF. In first diagnosed AF, it appears for the first time without cardiopulmonary disease. After diagnosis, if the events are random and self-terminating without the use of medications, it is classified as a paroxysmal AF. When events last longer than seven days, including the use of cardioversion either by drugs or by direct current, it is classified as persistent AF. When the arrhythmia lasts more than one year, a rhythm control strategy is required and it becomes a long-standing persistent AF. Finally, when the episodes do not stop, even when applying cardioversion, the treatment focuses on controlling the heart rhythm, and it is classified as permanent AF. When none of the aforementioned methods seem to work, an alternative is the ablation procedure. It is performed by destroying or scarring tissues, where the irregular electrical impulses are coming from, to
restore the normal rhythm. For most of the cases, ablation works; however, there is a risk of recurrent AF. Several studies have shown that there is a strong relationship between the difference in left and right atrial electrical activity and the recurrence of AF. Early detection of AF is important to ensure timely management of the condition and avoid the recurrence of the arrhythmia as much as possible. However, early detection may require continuous electrocardiography (ECG) monitoring for up to 30 months. Therefore, the most effective way to detect AF is using long-term portable cardiac monitoring devices. With these devices, people continue with their normal activities while ECG signals are recorded. The resultant signals are affected by noise and artifacts, such as baseline wandering, motion artifacts, and interference from other biopotentials. Most of the works on the analysis of ECG records for AF detection are based on heart rate variability. However, the most relevant information in atrial arrhythmias is contained in the P wave, which registers the depolarization of the right and left atria. P-wave analysis is more challenging because of the P wave’s low amplitude compared to the QRS complex. Therefore, the detection of AF is a challenging task. The AF episodes can be short (6–7 min), appear randomly, and the implementation of algorithms for real-time signal analysis is still challenging.

The development of algorithms for AF detection has been fostered by PhysioNet and the Computers in Cardiology conference through challenges. In 2021, the Predicting Paroxysmal Atrial Fibrillation/Flutter challenge was launched aimed to the development of a fully automated method to predict the onset of paroxysmal atrial fibrillation/flutter, based on the ECG prior to the event. Later, in 2014 the spontaneous termination of the AF challenge was launched aimed to predict if (or when) an episode of AF will end spontaneously. In 2017, they proposed a challenge for AF classification from a short single-lead ECG recording from 30 to 60 s in length. One of the most recent is the classification of the 12-lead ECGs challenge launched in 2020 to develop new machine-learning approaches for creating accurate and automatic detection of cardiac arrhythmias. These challenges are incentives to perform research in this area with hundreds of algorithms publicly available and published papers as outcomes.

The purpose of this paper is to offer an overview of the methods for automatic AF detection. AF detection methods have been classified in three groups according to the features considered: heart rate variability, atrial activity, and hybrid methods. Heart rate variability is measured using features like standard deviation of RR intervals, root square of the mean of the difference between successive RR, successive RR pairs with greater difference than 50 ms (NN50), statistical measures based on the dispersion of the variability of the intervals, slope analysis, density histogram, Shannon entropy, among others. Typically, machine-learning techniques are used with these features for detection and classification. On the other hand, atrial activity approaches aim to note the absence of P waves or the presence of fibrillation waves. As P waves have different frequency components than fibrillation waves, spectral decomposition can be used. Some common techniques for atrial activity analysis include spectral decomposition, correlation with a P-wave model, and analysis of measures such as duration, amplitude, shape, and statistical features. Hybrid methods perform both RR interval analysis and atrial activity analysis, which make them more robust. Usually, a QRS detector is used to measure RR intervals variation, and then, R peaks are used as a fiducial point to detect P waves.

The paper is organized as follows: Section 2 discusses the acquisition and analysis of ECG signals; Section 3 reviews some of the most common AF detection methods in each group, heart rate variability analysis, atrial activity analysis, and hybrid methods; and Section 4 concludes the paper.

II. ECG SIGNAL ACQUISITION AND ANALYSIS

Electrocardiography is a diagnostic test that records the electrical activity of the heart using electrodes placed on the skin. It is a noninvasive and painless technique. There are several types of devices that allow ECG signal recording in clinical environments
at rest, during exercise, and long-term monitoring in daily life.

Standard ECG records the biopotentials from 12 leads, including precordial and limb leads. Signals are detected using electrodes on the surface of the forearms, thorax, and on the left ankle for reference. It can take between 1 and 5 min to obtain the results.

When irregularities occur sporadically, standard ECG may not detect the event during the evaluation; there are other alternatives for those cases. The Holter monitor is a portable device capable of recording continuous ECG signals from 24 to 48 h; although currently, there are devices that can record up to seven days. Generally, three precordial leads are used to perform the test. Holter monitors use a battery and can be carried in a pocket, belt, or on a strap fastened to the shoulder. The patient can perform normal activities while the device is recording, but electrodes should not be wet, or their conductivity may be affected. It is possible to take notes of activities to establish relations when the symptoms appear.

When the events appear sporadically, the specialist can suggest an event monitor. It can be used for more time than a Holter, up to one or two months, but it only records electrical activity for a few minutes when the symptoms appear. Some event monitors can automatically detect anomalies in the heart rate that cause them to turn on and start recording. In the simplest, the patient can push a button to start recording when feeling the symptoms. The records can be sent by a mobile device to the doctor to be evaluated.13

The implantable loop recorder is a single-lead device designed to record continuous ECG signals for more time than a standard ECG or Holter element. It is required when the patient is suffering unexplained and recurrent events of palpitation or syncope. The device is implanted by surgery in a subcutaneous pocket in the chest. It is a simple surgery, and the device can last for a maximum of three years.14

In 2016, Apple Inc. launched the first ECG monitor free of wires or patches. It is a chest strap that allows continuous monitoring of ECG and measures heart rate variability, stress level, and breathing rate. It can record more than 20 million of data per day and send it to a mobile device. Currently, the Apple watch includes two electrodes that acquire a signal similar to a Lead I ECG. It can detect irregular heartbeats that may be associated with AF.15,16

Under the same framework of technological innovations, Alivecor Inc. introduced Kardia mobile. This device allows obtaining ECG monitoring using two electrodes to measure bipolar leads from the right and left forearms. It works with either iOS or Android devices for visualization and other functions. It is also possible to find a watch version of this device named the Kardia band. It has been shown that the algorithms that work with the Kardia band can accurately differentiate AF from normal sinus rhythm.17

III. ATRIAL FIBRILLATION DETECTION BASED ON HEART RATE VARIABILITY

RR intervals exhibit an irregularly irregular behavior during AF episodes, because the heart’s activity is completely arrhythmic. RR variability can be quantified using mathematical, statistical, and digital signal-processing approaches. Examples of these methods include turning point ratio, Shannon entropy, density histogram, Poincare map, hidden Markov models, and waveshape analysis.18 Additionally, statistical measures are frequently used; these include standard deviation, root mean square, and the mean of the RR intervals.19–24

The purpose of the turning point ratio is to test the time series’ randomness, comparing the value of each RR with the neighbors. The Shannon entropy offers a measure of uncertainty of the source information. It is high when data show high dispersion. A normal sinus ECG is expected to have a lower value for Shannon entropy than an AF ECG. A density histogram is used to show every class’ distribution for a normal sinus ECG and AF ECG. A Poincare map expresses every RR interval, depending on the previous one to quantify RR variability. Some works propose modifications on the existing techniques to improve the feature extraction,12,25,26 while others only measure the duration of RR and RT intervals and focus on the application of artificial intelligence techniques.27,27–32
A. Turning Point Ratio

The turning point test allows one to know if a series of random variables is independent. It is also used to test for randomness, such as in the case of AF. For instance, a data set with periodic behavior has less turning points than a random data set.

The number of turning points expected in a series of duration $n$ can be expressed as follows:

$$\frac{2n-4}{3}$$

and its respective variance as

$$\frac{16n-29}{90}$$

As $n$ grows, the number of turning points should be between $n(2n/3, 8n/45)$, with a level of significance of 5%. If the null hypothesis states that the series is random, it can be rejected if the number of turning points lies outside of the range $[2n/3 \pm 1.96 8n/45]$.

In 2016, Dash et al. used the turning points ratio with the root mean square of successive RR differences and Shannon entropy to characterize AF, achieving results in specificity and sensitivity of 94.4 and 95.1%, respectively, for the MIT-BIH AF database. In 2019, Bashar et al. used the same approach and a support vector machine for classification.

B. Shannon Entropy

The Shannon entropy measures the uncertainty of a data set. It can be considered as a measure of the amount of information contained in a signal. The most likely events are those that provide the least information; therefore, Shannon entropy will tend to 0, and on the contrary, for the less probable events, Shannon’s entropy will be around the maximum value of 1.

Shannon entropy can be estimated as follows:

$$E(x) = - \sum_{i=1}^{k} p(i) \log_2(p(i))$$

where $p(i)$ represents the probability distribution of each data point in the set. In AF, the ECG signal must be segmented in several bins, and the probability distribution can be estimated as follows:

$$p(i) = \frac{N_{\text{bin}}(i)}{T_{\text{bin}}}$$

where $N_{\text{bin}}(i)$ represents the number of beats at $i$th bin and $T_{\text{bin}}$ is the total number of beats in the signal.

The Shannon entropy is one of the most used techniques to quantify RR variability. It has also been widely used in other fields, such as thermodynamics, communications, and information theory. However, there are some limitations that should be considered. Shannon entropy depends on the probabilistic model to characterize the information. It may vary depending of the length of bins, for instance, for a sequence “1212121212,” if the probabilistic model considers individual numbers as independent, Shannon entropy will be 1, but if the sequence is taken every two characters as “12 12 12 12,” the entropy will be 0; so if the bins are very large, the estimation of entropy may become low.

C. Density Histogram

This technique has been used to distinguish between normal sinus rhythm and AF, using either a variation coefficient or a threshold from the density histogram of RR intervals and $RR$, which is the difference between two successive RR intervals. The ECG time series is segmented into blocks with a specific number of beats, which must be part of a specific class to build the histogram. Some of the authors who have used this method to detect AF are: Huang et al. and Yaghoubi et al., among others. Figures 1 and 2 show a typical density histogram for normal sinus rhythm and AF, respectively. The signals are from the MIT-BIH atrial fibrillation database from the Physionet record 06453m, and the MIT-BIH normal sinus rhythm record 17052m. It is possible to see that the histogram of AF is wider than in normal sinus rhythm, where the maximum frequency is related to the heart beat.

D. Poincare Map and Scatter Plots

Poincare map has been used to quantify the variability of RR intervals for several purposes.
FIG. 1: A density histogram for normal sinus rhythm taken from the MIT-BIH normal sinus rhythm database, record 17052m

FIG. 2: A AF density histogram taken from the MIT-BIH atrial fibrillation database from Physionet, record 06453m

is built by expressing every RR interval in the function of the previous one. The Poincare map offers a visual inspection of the dispersion between the intervals. Typical statistical measures, such as standard deviation, mean, or root-mean square can characterize the variation of RR intervals. In the case of AF, it is expected that the Poincare plot takes an irregularly irregular shape; whereas in a normal sinus rhythm, it shows a pattern.

Lian et al. 43 and Sarkat et al. 44 used scatter plots of RR intervals versus changes in RR intervals. AF is characterized by a high dispersion in the plot.
E. Hidden Markov Models

A hidden Markov model is a statistical model with known output but unobservable states. The goal of the method is to determine the unknown parameters based on the known ones. Every state has a likelihood distribution influencing the output sequence. In this way, the output sequence generated by a hidden Markov model yields information about the states. Figure 3 presents the classical architecture of a hidden Markov model, where $x_n$ represents the hidden state, $Y_n$ the observable output for every state, respectively, $w_n$ the probability of transition between the states, and $z_n$ the probability of the output. The probability to observe the sequence $Y = Y_0, Y_1, ..., Y_I$ is given by

$$P(Y) = \prod_{n=0}^{I} P(Y_n|X_n)P(X_n)$$  (5)

A hidden Markov model can be used to characterize RR intervals, modeling the RR sequence as a three-state Markov process (short, regular, and long). Based on the transitions between states, the heart rate is characterized as regular by the likelihood of transition from an estate R to itself. It is expected to occur when the RR intervals present the same length. Previous works have used this method.45,46

F. Spectral Analysis

Spectral analysis allows the detection of components related to the QRS complex, T wave, P wave, and noise separately even in cases when they are overlapped or shifted in time. Heart rate variability is associated with frequency components of the ECG signal; for this reason, spectral analysis provides useful information when there are changes in the cardiac rhythm generated by AF.47 Most these works are based on the Fourier transform and the power spectral density. Mei et al.48 extracted features related to heart rate variability, such as

FIG. 3: Hidden Markov model classical architecture
standard deviation, mean, Poincaré plot, among others. These descriptors were combined with the Welch power spectrum density calculated with the fast Fourier transform (FFT) around every beat detected. These features were used as inputs for the classifiers: support vector machines and bagged trees. Zhao et al. introduced a new spectro-temporal analysis method by assuming the time-varying Fourier coefficients of the signal have Gaussian process priors. After computing a two-dimensional (2D) matrix with the spectro-temporal information of the signals, a densely connected convolutional network was used for the classification into four groups: AF, non-AF normal rhythms, non-AF abnormal rhythms, and noisy segments.

G. Wavelet Transform

AF is characterized by the variability of the heart rate. In normal conditions, the QRS complex tends to be regular with a dominant peak in the range of 0–20 Hz. During AF episodes, this behavior can vary, introducing new components in the spectrum related to the variability of the RR intervals. This variability in the ECG spectrum has been quantified mainly by applying the wavelet transform.

The wavelet transform provides a time-frequency representation using coefficients for displacement and scaling of a base function (mother wavelet). For a space in $\mathbb{R}^2$, the set of basic functions is defined as follows:

$$\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n)$$  

where $m$, $n$ represent scaling and displacement, respectively. The discrete wavelet transform can be obtained by applying the pyramid algorithm, which uses filtering and decimation to obtain two output signals with half of the original signal’s length. A high pass filter allows obtaining the detail coefficients, and a low pass filter generates the approximation coefficients.

Duverney et al. and Yuan et al. used the wavelet transform to find the beginning and the end of episodes with high heart rate variability that can be related to AF. The fractal analysis was used to classify these high variability periods into normal or pathological (AF). Andersen et al. implemented wavelet decomposition to segment the signal and extract features such as entropy and peak-to-average power ratio of each of the coefficients (details and approximation); the classification was done by a support vector machine.

H. Other Methods

1. Pattern Similarities

Marsili et al. implemented a methodology for AF detection suitable for wearable tele-Holter devices. After applying the QRS complex detector proposed by Gutierrez et al., RR interval irregularity is quantified applying the method proposed by Petrenas et al., which measures the similarity between RR intervals in a window of length $l$, as follows:

$$M[n] = \frac{2}{l(l-1)} \sum_{j=0}^{l-2} H(|r_m(n-j) - r_m(n-k)| - \gamma)$$  

where $H$ represents the heaviside step function, $r_m$ is the output of the median filter, and $\gamma$ is the similarity threshold. The pattern similarity is obtained by averaging the forward-backward filtered version of $M$ and $r$.

2. Symbolic Dynamics

It was first proposed by Zhou et al., and has been used in other works. The method builds a set of symbols to encode the heart rate and then the irregularity of the symbols is quantified using the Shannon entropy. The mapping function to encode the heart rate series into the symbolic series used by Marsili et al. is as follows:

$$sy[n] = \begin{cases} 63 & \text{if } hr[n] \geq 315 \\ 1\lfloor hr[n]/5 \rfloor & \text{other cases} \end{cases}$$  

where $\lfloor \cdot \rfloor$ represents the floor operator. The complexity of the symbolic series is quantified by the combination of a series of words. The words can be composed either by eight symbols or by two successive symbols to reduce memory space, which allows the implementation in an embedded device.
3. **Multiparametric RRI Feature**

Islam et al. introduced a new multiparametric feature based on the heaviside step function to characterize RR interval irregularity. This is to avoid the use of the Shannon entropy due to the required length in the data set. The irregularity of a parameter $\gamma$ can be estimated as follows:

$$h(\gamma) = \frac{2}{n(n-1)} \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} H(|r(j)-r(k)| - \gamma) \quad (9)$$

There is an optimal value of $\gamma$ in the literature. The multiparametric method is proposed using several values for $\gamma = [\gamma_1, \gamma_2, ..., \gamma_m]$ to obtain multiple values for $h$ forming a RR irregularities features vector $h = [h_1, h_2, ..., h_m]$. The final choice of $\gamma$ is obtained by a support vector machine.

4. **AFD Algorithm**

Linker proposed a five-step algorithm for AF detection based only on the RR intervals variability, when the algorithm detects a QRS complex, calculates the RR duration, and converts it to heart rate. Several beats are stored to form a segment. Then, the mean and linear trend are subtracted from the values of the segment. The median of the absolute value of the previous step is calculated to measure the variability. Then it is compared to the previous, current, and following segments, and finally, the median value is returned. It was found that in a normal sinus rhythm, the mean heart rate has a linear relationship with the AFD algorithm’s output in a normal sinus rhythm.

5. **Time-Varying Coherence Function**

Lee et al. introduced a time-varying coherence function that combined with Shannon entropy, provides high accuracy for AF detection. The time-varying coherence function is calculated by the multiplication of two time-varying transfer functions. The first one is obtained using two adjacent data segments with one as the input signal and the other as the output, and for the second one, the input and output signals are shifted. The time-varying coherence function by nonparametric time-frequency spectra is defined as follows:

$$|\gamma(t,f)|^4 = \frac{|S_{yx}(t,f)|^2}{|S_{xx}(t,f)||S_{yy}(t,f)|} \times \frac{|S_{yx}(t,f)|^2}{|S_{xy}(t,f)||S_{yy}(t,f)|} \quad (10)$$

where $S_{xy}(t,f)$ and $S_{yx}(t,f)$ represent the time-frequency cross spectrum, and $S_{xx}(t,f)$ and $S_{yy}(t,f)$ represent the auto spectra of the signals $x$ and $y$, respectively. The time-varying transfer function in terms of time-frequency spectra for a linear time-varying system is defined as follows:

$$H_{x\rightarrow y}(t,f) = \frac{S_{xy}(t,f)}{S_{xx}(t,f)} \quad (11)$$

where $H_{x\rightarrow y}(t,f)$ is the time-varying transfer function from the input $x$ to the output $y$, in the same way the time-varying transfer function from $y$ as input and $x$ as output is as follows:

$$H_{y\rightarrow x}(t,f) = \frac{S_{yx}(t,f)}{S_{yy}(t,f)} \quad (12)$$

Therefore, by multiplying the two transfer functions, the time-varying magnitude is obtained, as follows:

$$|\gamma(t,f)|^2 = |H_{x\rightarrow y}(t,f)H_{y\rightarrow x}(t,f)| \quad (13)$$

Since it is expected to find a high coherence value (close to 1) for a normal sinus rhythm, and lower values for AF segments, this parameter can be used for classification.

6. **Bimodal Classification**

Kruger et al. developed a method for automatic detection of AF for a single-lead ECG device. The algorithm picks 6 s from a 20 s ECG record. Time-frequency domain analysis is performed to characterize the variability of RR intervals. Classification is performed by a bimodal scattering clusterization, where the classes are perfectly distinguishable using a hyperbolic function.
7. Orderings of Conjugate Symmetric Complex Hadamard Transform

Annavarapu and Kora\textsuperscript{57} segmented the signal around each R peak using different ordering of the conjugate symmetric complex Hadamard transform. Several orderings were used: natural order, Paley order, sequence order, and CalSal order. The classification was done using a Levenberg Marquardt Neural Network.\textsuperscript{57}

IV. Atrial Fibrillation Detection Based on Atrial Activity Characterization

It is well known that AF is mainly characterized by heart rate variability and loss of coordination between atrial and ventricular activity. All cardiac arrhythmias exhibit RR variability, either with regular or irregular behavior. However, a remarkable fact in atrial arrhythmias is the change in atrial activity, more specifically, the lack of P waves or the presence of F waves. F waves can have different shape, amplitude, and duration in every patient, and the P wave can be superimposed by the QRS complex. As a result, atrial activity characterization for AF detection can be challenging.\textsuperscript{58} The methods for AF detection based on the analysis of atrial activity can focus on either the detection of the absence of P waves, or the detection of F waves.\textsuperscript{59}

A. P-Wave Detection Methods

These methods aim to detect and characterize P waves. Ladavich and Ghoraani\textsuperscript{60} extracted morphological and statistical features of ECG signals from a training set. These features were used as the input for an expectation-maximization algorithm to create a Gaussian mixture model. The model was used to detect the lack of P waves, and hence, AF.\textsuperscript{50} Other approaches for P-wave detection are correlation with a model P wave, blind source separation, statistical measurements, among others. Correlation is based on the detection of the P wave using a known template of the wave in normal conditions. The template is chosen either manually from a local P wave in the signal or using a universal standard of the P wave. Statistical measurements, such as the root mean square of the P wave voltages, give a measure of the variability in P-wave shape, showing lower values for patients with AF. Blind source separation can be used to extract the P wave using multiscale analysis. As the P wave can be retained over several resolution scales, other techniques are usually required to improve the detection and characterization. Previous works have shown that the best performance in P-wave detection is achieved by combining time and frequency domain techniques.

1. Correlation

Correlation allows one to measure the degree of similarity between two signals. It can be used to match a template with a signal. The mathematical procedure consists of shifting one signal over the other and calculating the summation of the product of the two signals, as follows:

\[
y(x) = \sum_{m=0}^{M-1} f(m)g(x + m)
\]

where \(M\) represents the length of one of the signals. It is expected to have a maximum when the signal matches with the model. The correlation coefficient can be used to know if there is a relationship between the signals and can be estimated as follows:

\[
\rho = \frac{S_{xy}}{S_x S_y}
\]

where \(S_x\) and \(S_y\) are the standard deviation of the signals \(x\) and \(y\), respectively, and \(S_{xy}\) is the covariance between them. The correlation has been widely used in ECG processing to detect P waves using a wave template and establishing a threshold. Last et al.\textsuperscript{61} extracted a template from the ECG records and then applied correlation to match P, QRS, and T waves. Dotsinsky et al.\textsuperscript{62} generated a synthesized sinusoidal P-wave template and applied a modified convolution for P-wave detection.

2. P-Wave Duration

P-wave duration is an important criteria in AF diagnosis, since it shows if there are irregularities in
atrial conduction. Despite the fact that atrial depolarization presents the first detectable change in ECG records, it is not easy to detect the onset and offset of the P wave, and it may change based on the lead. The challenge lies on establishing a criteria for the onset and offset of the wave, which allows one to classify the signal as a normal sinus rhythm or AF.

The New York Heart Association established 110 ms as the duration of a normal P wave. Longer values may be due to an atrial arrhythmia. However, this value is not reliable to detect patients at risk of AF, since only 38% of the patients with an intratricular conduction defect have longer P waves. Andrikopoulos et al. use a simple criteria that searches the first point that crosses the baseline with positive slope, and the next one with a negative slope for the onset and offset, respectively. 64

In 2000, Mehta et al. 65 introduced a new concept to measure the P-wave prolongation named the P terminal force. It is calculated as the product between the duration of the terminal part of the P wave (negative part) in lead V1 and its depth in millimeters, as shown in Fig. 4. If the terminal part of the P wave is positive, then the whole window must be considered. 65 P-wave terminal force is frequently used to detect left atrial abnormality. Similarly, Passman et al. showed that a prolongation in the PR interval duration from lead V1 is a risk factor of AF after coronary artery bypass grafting. 66

To estimate the P-wave duration, onset and offset should be calculated as accurately as possible. Several algorithms have been developed to perform this search. Sasikala and WahidaBanu 67 used the wavelet transform to implement a search algorithm to detect the onset and offset of the P wave. After the detection of the QRS complex, the P peak is detected as the highest peak within a window of 200 ms before the onset of the QRS complex, to the onset of the QRS complex. P waves are presented as a modulus maxima pair with opposite signs, as is shown in Fig. 5. The onset is chosen as the point at the left of the zero-crossing, where the amplitude decreases 5% from the maximum module. In the same way, the offset is chosen as the point at the right of the zero-crossing, where the amplitude decreases 5% of the modulus minimum. 67

3. Blind Source Separation

Blind source separation is a well-known method to separate sources that are mixed into a signal. The method is based on the assumption that the sources are independent and there are at least as many observations as independent sources. Being $x(t)$, the vec-

![FIG. 4: Negative part, and depth of the P wave to calculate the terminal force](image-url)
The problem lies in how to obtain $A$ and $s(t)$ from the observations $x(t)$. Two approaches have been proposed to solve this: principal component analysis (PCA) and independent component analysis (ICA). ICA is preferred because the PCA success depends on the orthogonality of the mixing matrix; whereas, ICA does not have that condition. Several works have shown the usefulness of blind source separation in ECG signal processing, taking into account that ventricular and atrial activity come from independent sources and do not have a Gaussian distribution. 68,69

In 2004, Rietal et al. 70 used blind source separation and ICA to extract atrial activity from real ECG records to analyze AF. After the application of blind source separation, kurtosis-based reordering and spectral analysis were performed on the separated sources to detect the source containing the atrial activity and classify it as AF or a normal sinus rhythm. 70 One year later, Castells et al. 69,70 introduced a modification of the blind source separation by the application of the spatio-temporal operation, which allows the exploitation of the temporal information found out in previous works. 69,70 Signals are classified as AF or noise based on the location of the main peak in the spectrum of the separated source. 71

B. F Wave Detection Methods

During AF episodes, atrial activity is uncoordinated and faster than usual. As a result, P waves can be replaced in the ECG records by multiple waves that vary in shape and duration, called F waves. Methods based on the detection of F waves exploit this information using spectral analysis to detect high-frequency components related to abnormal atrial activity.

Atrial flutter (AFL) is another atrial arrhythmia that can have a similar behavior in the ECG signal. However, in atrial flutter, F waves present a particular sawtooth shape with a higher frequency than in AF. 72 Therefore, spectral analysis can be useful for detecting F waves and distinguishing AF from AFL, even in the presence of noise.

In the early 1990s, methods based on atrial activity’s characterization were focused on delimiting the P wave in the time domain. However, in the case of F waves, it is challenging because of their irregular behavior. Therefore, spectral analyses have been

FIG. 5: Maximum, minimum, and zero-crossings from a window of the ECG signal
useful for this task. In 1995 Hiraki et al. concluded that frequency domain analysis is as useful as time domain analysis to detect paroxysmal AF. They analyzed the ratio of the power spectrum area (Ar20) in the interval 0–20 Hz to 20–100 Hz, and the magnitude ratio at 30 Hz (%Mag30). They found that Ar20 was significantly higher and %Mag30 was significantly lower in patients with paroxysmal AF, reaching a predictive accuracy of 83 and 73% with Ar20 and %Mag30, respectively.

After applying a template-matching algorithm to suppress QRS-T components, Bollman et al. applied Fourier transform over the remaining signal to analyze the fibrillatory baseline. They found a dominant component at 5.1 ± 0.7 Hz. This peak was strongly correlated with the duration of the AF episode. Shorter AF with a duration of < 15 min had a lower frequency dominant component at 4.8 ± 0.6 Hz; whereas, longer episodes had a dominant component at 5.3 ± 0.7 Hz. Longer episodes showed higher spectral components for AF episodes of > 5 min. The same relationship was found by Kanuru et al. They observed lower spectral components for AF episodes of < 5 min and higher spectral components for AF episodes of > 5 min.

In 2015, Ródenas et al. computed the wavelet entropy over the TQ segments to suppress F waves during AF episodes using a single lead. Signals were classified as normal sinus rhythm and atrial fibrillation using a threshold. One year later, García et al. used a similar approach. They extracted the relative wavelet energy over the stationary wavelet transform of the TQ segments.

V. AF DETECTION USING HYBRID METHODS

Hybrid methods offer a robust solution for AF detection since they exploit information from both atrial and ventricular activity. If only ventricular activity (i.e., heart rate variability) is considered, then it is possible to get false negatives since this feature is common among all the arrhythmias. On the other hand, atrial activity analysis can be challenging due to the P wave’s low amplitude. Therefore, methods that combine information from both atrial and ventricular activity can offer a more accurate detection because they can overcome the limitations of using only one feature.

Hybrid approaches can be divided into two groups: (i) classical methods that use signal processing for feature extraction and statistical or machine-learning techniques for classification and (ii) deep learning approaches.

A. Methods That Use Signal Processing for Feature Extraction

Two of the most relevant works in hybrid approaches are those by Couceiro and Babaeizadeh. Couceiro calculated features related to the P-wave presence, heart rate variability, and atrial activity. The P wave was detected with the correlation of the signal with a P wave model. Heart rate variability was computed using a three-state Markov process, where every RR interval was characterized as small, regular, or long based on the probability of transition from state R to itself, since this transition is more likely to occur when the RR intervals have the same length. Finally, atrial activity was characterized using wavelet analysis. One year later, Babaeizadeh et al. developed a real-time methodology based on a hidden Markov model to characterize heart rate variability. The Markov model used a transition probability matrix, which represents the probability of R-R interval transition from one state to another. A fixed threshold was settled to detect AF. To improve the method, they also analyzed PR interval duration and the similarity between the morphology of two consecutive P waves. PR interval duration was characterized by the deviation from the average. For AF episodes, the PR interval duration can be nonmeasurable if there is no P wave, or very large, in the presence of F waves. P wave similarities were measured by the coincidence of two consecutive P waves, and it is expected to be low during AF. A decision tree was used for classification.

In 2012, Stridh and Rosenqvist proposed a methodology for AF detection using one lead ECG from a home monitor. The method was developed in five steps. First, beat detection and classification is performed using cross correlation to separate beats with different morphology in several classes. In the
second step, an analysis of RR intervals is performed using the median and the standard deviation. In the third step, the P wave is detected by waveform characterization. Spectral analysis is developed to find a model of the P wave and verify the similarity with the rest of the P waves. For the fourth step, quality control is achieved by analyzing the number of beats with similar morphology. Finally, the rhythms are sorted as a low priority for normal sinus rhythm, high priority for signals with abnormal beats, and the third group contains unreliable rhythms.

Recently, Pürerfellner et al. showed that the use of P-wave information together with heart rate variability substantially reduces inaccurate detection of AF episodes. The proposed algorithm uses the Lorenz plot of the difference between RR intervals. Based on this, an AF evidence score is computed every 2 min and compared to a threshold. In the improved version, a P-wave evidence score is calculated and subtracted from the AF evidence score of the difference between RR intervals before comparing the threshold. The P-wave evidence score is computed by averaging 600 ms of baseline ECG before R waves for four consecutive beats. Morphological features are extracted and used to calculate the score, and finally, the P-wave evidence score is used as evidence against AF’s presence.

In 2017, Bruun et al. computed the root mean square of successive differences of the RR intervals, Shannon entropy, and sample entropy to quantify the heart rate variability. The atrial activity was evaluated by extracting the log-energy entropy and peak-to-average band power from the detail coefficients of the discrete wavelet transform. Classification is performed by a bagged tree.

A different approach was developed by Asgari et al. Instead of detecting R and P waves, they used the stationary wavelet transform to characterize the ECG signal. They measured power spectral density and log-energy entropy and used a support vector machine for classification.

**B. Deep Learning Methods**

Deep learning has seen a rise during the last years. The availability of faster hardware and highly optimized open source libraries allow creating deeper neural networks to solve a wide range of problems. AF detection has not been the exception, and convolutional neural networks (CNN) have been used on ECG records to classify normal sinus rhythm and other arrhythmias. The main advantage of using deep learning methods in AF detection lies in the volume of data that can be processed in a short time with high accuracy. Deep learning methods can be used to perform both feature extraction and classification stages, due to their ability to model complex systems when properly designed and trained.

The approaches that use deep learning methods for AF detection typically perform some preprocessing and basic feature extraction on the ECG signal and then use deep learning networks to complete the characterization and perform the classification. Yuan et al. use autoencoder neural networks to detect AF. The signal was previously filtered, the R waves were detected for signal segmentation, and some features related to the P waves were extracted. In 2018, He et al. calculated the continuous wavelet transform of the P-QRS-T segment to characterize atrial and ventricular components. A 2D CNN was then used to extract features and perform the classification. Similarly, Xia et al. defined a bidimensional matrix using the short-term Fourier transform and the stationary wavelet transform of the ECG signal. This matrix was used as the input for a 2D CNN, which extracts features and classifies the signals.

Among the most recent approaches, Shi et al. proposed a loop-locked framework integrating AF diagnosis, label query, and model fine-tuning. The model is pretrained on a multiple-input deep neural network using labeled samples from an original training set. The model can be used in practical application for AF detection when new data are collected to form a candidate set. The learning stage is increased continuously using the most informative samples from the candidate set. Yildirim et al. used deep neural networks to detect cardiac arrhythmia on more than 10,000 patients. The model works on each lead of the 12 leads in ECG records. The model included learning and sequence learning.
After the representation of learning layers, the sequence learning stage involved a long short-term memory unit. 86

VI. DISCUSSION

Table 1 summarizes the works considered in this review, and Table 2 contains the acronyms of the databases for each of them. For each work, specificity (SP) and sensitivity (SE) are shown to evaluate the methods’ performance.

In general, methods that use only atrial activity show lower sensitivity and specificity than methods based on heart rate variability and hybrid approaches. The reason may be that the detection and characterization of P waves is challenging due to their low energy. Some authors reported the inability of algorithms to detect the P wave because the wave shape varies from a patient to another. However, as the P wave contains information on atrial contraction, it is important to differentiate AF from other arrhythmias. Atrial activity characterization methods can be useful for AF detection. However, their performance is highly dependant on the signal-to-noise ratio.
ratio of the ECG records. Therefore, signal quality metrics should be considered in the detection algorithms as medical long-term ECG recordings are often characterized by a substantial variation in the level of noise. 59

Regarding the methods based only on heart rate variability, they can generate many false alarms, because this feature is presented in most of the arrhythmias. For instance, AF can be misclassified as atrial flutter, ventricular premature beats, or as bigeminy and trigeminy. 59 Heart rate variability in AF is known to be not wholly unpredictable, as there is a nonzero correlation between the observed and the predicted R-R interval at different correlation lags. 59,95 This predictability is weak and its physiological causes are not clear. 96 Therefore, heart rate variability cannot be used by itself to detect AF. However, when properly characterized, it can improve the performance of an AF detection system.

Previous studies have shown that including information about atrial activity in an algorithm based on heart rate variability improves the performance of AF detection. 79 Hybrid approaches that combine information from atrial activity and heart rate variability offer improved prediction performance of machine-learning models compared to algorithms that use only one of these features. 97

As AF can be confound with atrial flutter, it is important to explore new features that allow one to distinguish these two conditions. Table 3 shows a set of features of AF, atrial flutter, and normal sinus rhythm. Some of them have been used previously. 71,98–100 Atrial activity in AF is characterized by a narrowband spectrum with fundamental frequency between 3.5–9 Hz with higher spectral concentration around the main frequency peak and not harmonics and bandwidth of 2 Hz. In the case of atrial flutter, the spectrum contains harmonics with decreasing amplitude due to the reentry rhythm. One interesting feature to exploit is the atrio-ventricular coupling, since it is evident in normal sinus rhythm and in atrial flutter, and it is not present in AF. Another feature that is present in AF and atrial flutter is overlapping between atrial and ventricular activity, and it can be characterized with spectral analysis.

Finally, it is well known that the big challenge in AF detection relies on asymptomatic AF. It is estimated that there is at least one-third of the population undiagnosed. Therefore, it is essential to improve the sensitivity of the current methods. One alternative is using implantable cardiac monitors. However, these devices are recommended only for patients with frequent symptoms considered at risk. Paroxysmal AF is characterized by random and

| TABLE 3: Main differences and resemblances among AF, AFL, and a normal sinus rhythm |
|--------------------------------|--------------------------------|--------------------------------|
| **Normal Rhythm**              | **Atrial Fibrillation**        | **Atrial Flutter**             |
| RR Interval                    | Regular Same as P-wave         | Irregular                      | Regular/Irregular Integer multiple of the atrial cycle length |
| Atrio-ventricular coupling     | Coupled                        | Uncoupled                      | Coupled                      |
| Overlapping of atrial and ventricular | Not overlapped                | Overlapped                     | Overlapped                   |
| Atrial signal                  | pulse                          | Continuous                     | Continuous                   |
| Atrial period                  | 500 ms–1 s                    | 120 ms–200 ms                 | 220 ms–300 ms               |
| Regularity of atrial waveform  | Regular                        | Irregular                      | Regular                      |
| Fundamental freq. of the atrial component | Typ. 1.2 Hz                   | Typ. 6 Hz                      | Typ. 4 Hz                    |
| Harmonics                      | Yes, firstly increasing, then decreasing | No                            | Yes, with decreasing amplitude |
| Bandwidth                      | 20 Hz                          | 10 Hz                          | 20 Hz                        |
| Bandwidth of fundamental peak  | Delta                          | 1–3 Hz                         | Delta                        |
self-termination events. Thus, the most suitable alternative is the use of portable devices. When the ECG records are acquired during normal daily activities outside a clinical environment, the low signal-to-noise ratio imposes more requirements for signal preprocessing and the performance of the methods may be reduced. Additionally, portable devices can record up to three leads per patient, and robust and accurate P-wave detection often requires a higher number of leads (up to 12 leads). Therefore, it is important to develop new methods or improve current methods for AF detection with the following characteristics:

- high temporal resolution to detect short episodes of AF
- use of up to three leads
- long-term monitoring
- robust algorithms to deal with a low signal-to-noise ratio.

AF has a high incidence in stroke, even short episodes can be related to it, and reduces the quality of life (QoL) of the patients. 101 Therefore, the search for an optimal solution for detecting patients at risk of AF is fundamental to start treatment opportune and improve their QoL.

ACKNOWLEDGMENTS

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper. This research is partially supported by Universidad Tecnológica de Bolívar, Cartagena-Colombia, (Grant No. C2018P022) and by Erasmus+ KA107 (ICM). This research did not receive any funding from NIH. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

REFERENCES

15. Barr CS. Comparison of accuracy and diagnostic valid-
ity of a novel non-invasive electrocardiographic monitoring device with a standard 12-lead ECG recording device; 2013.


100. Husser D, Stridh M, Sommo L, Olsson SB, Bollmann A.