UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACULTY OF ENGINEERING
MASTER'S DEGREE IN BIOMEDICAL ENGINEERING

A novel deep-learning method for fibrillatory waves extraction

Candidate:
Luca Goffi

Advisor:
Dr. Agnese Sbrollini

Coadvisor:
Prof. Laura Burattini

Academic Year 2022-2023

UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACULTY OF ENGINEERING
MASTER'S DEGREE IN BIOMEDICAL ENGINEERING

A novel deep-learning method for fibrillatory waves extraction

Candidate:
Luca Goffi

Advisor:
Dr. Agnese Sbrollini

Coadvisor:
Prof. Laura Burattini

Academic Year 2022-2023

Università Politecnica delle Marche
FACULTY OF ENGINEERING
MASTER'S DEGREE IN BIOMEDICAL ENGINEERING
Via Brecce Bianche – 60131 Ancona (AN), Italy

To you who have always been there,
to your teachings and your love.

To my grandmother Tina.

Acknowledgments

I would like to express my deep gratitude to all the people who contributed to the completion of this thesis. This journey was as beautiful as tough and would not have been possible without the support and precious contribution of all of you.
First of all, I would like to thank Dr. Agnese Sbrollini for her constant support and patience in guiding me through the complex stages of this project. His expertise and enthusiasm have made this research journey an extraordinary educational experience.
At the same time, I would like to thank my coadvisor, Prof. Laura Burattini and all the professors and researchers of the Cardiovascular Bioengineering Lab. of the University, who welcomed me and made me feel at home during my internship with them. A special thanks go to MHD Jafar Mortada, the shoulder I relied on during my first steps with Python.
A heartfelt thanks goes to my parents, Andrea and Cristina and my brother Davide for their unceasing encouragement, their love and unconditional support. Thanks also to all my aunts, Antonella, Stefania and Sabrina, your trust in me has been the necessary push to face the challenges and persevere in reaching this goal.
I can't help but thank you, Marina. You were the closest to me during this period. You always supported me, celebrated with me in the happy moments, and supported me in the most difficult ones, when I saw only black and you tried to colour my days.
I would also like to thank my fellow students and friends who have shared this academic journey with me. Your carelessness between lessons and your support before and after each exam made the journey more meaningful and enjoyable.
Thank you to all those who made this achievement possible. Your presence in my life is a precious gift.

Ancona, February 2024

Luca Goffi

Abstract

Atrial fibrillation (AF) is the most common sustained supraventricular arrhythmia caused by a dysfunction of the sinus atrial node, which is no longer able to guide atrial depolarization and, consequently, ventricular depolarization. As a result of this, atrial contractility is lost causing an inability to completely empty blood from atrial appendage leading to the risk of clot formation and subsequent thromboembolic events.
The ECG analysis represents the most well-established noninvasive technique used to detect atrial fibrillation. Most of the works on the analysis of ECG records for AF detection are based on heart rate variability, i.e., on the R-R intervals, even if, the most relevant information in atrial arrhythmias is contained in the fibrillatory waves (F-waves) which replace the ordinary P-waves, related to the depolarization of the atria. Due to their stochastic shapes and little amplitude, the process of detection, extraction and visual inspection by clinicians of the F-waves represents a really challenging task. Classical techniques used to carry out these tasks involve the use of different signal processing principles such as principal component analysis; however, recently, deep neural networks such as 1D convolutional neural networks (CNNs), achieved high results in feature extraction and filtering of biomedical signals. For this reason, in this thesis, a two-stage deep learning method based on 1D CNNs and multipath modules is proposed to extract F-waves signals from 1-second length windows of recorded ECG of patients affected by atrial fibrillation. The system was trained and tested on a reference database, available online, for validation of methods of extraction of atrial fibrillatory waves in the ECG, which consists of records of simulated AF 12-lead ECG signals that are different combinations of real F-waves and QRST complexes.
According to the results related to the testing dataset, in terms of evaluation metrics analyzed, the performances of the implemented method are really promising, with a mean correlation between the output and the target signal of 0,82 and mean values of the sum of square distances and maximum absolute distance of 0,05 au and 0,03 au, respectively. Moreover, also the mean absolute errors of the dominant frequencies and the amplitudes computed between the output and the target signals are low, with values of 0,05 Hz and 0,01 μV, respectively. To our knowledge, no other work exists in the literature which employs deep learning algorithms to extract F-waves from ECG signals of AF patients. For this reason, also according to the promising results obtained, this work can be considered a forerunner for this branch of research.

Contents

List of Figures

List of Tables

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia affecting roughly 33.5 million people worldwide and its prevalence is expected to increase significantly in the coming years.
AF is caused by a dysfunction of the sinus atrial node, which is no longer able to guide atrial depolarization and, consequently, ventricular depolarization. As a result of this, atrial contractility is lost causing an inability to empty blood from atrial appendage leading to the risk of clot formation and subsequent thromboembolic events.
Since the pathophysiological mechanisms causing and maintaining AF are still not completely understood, various types of noninvasive techniques have been developed to better understand the mechanisms. In addition, the lack of a comprehensive understanding of its pathological mechanism, made is timely diagnosis becoming a problem . People often miss the optimal treatment time because the early stages of atrial fibrillation are usually paroxysmal and asymptomatic. Therefore, the development of an automatic atrial fibrillation detection system, to provide accurate and reliable diagnostic information as early as possible, is of great significance for improving the quality of treatment and reducing the further deterioration of the patient's health.
The ECG analysis represents the most well-established noninvasive technique used to detect atrial fibrillation. Even if most of the works on the analysis of ECG records for AF detection are based on heart rate variability, i.e., on the R-R intervals, recently the attention has been directed towards atrial fibrillatory waves (F -waves) and their characterization. In episodes of AF, the F- waves replace the ordinary P-waves related to the depolarization of the atria. F-waves can have different shapes, amplitude, and duration in every patient and, additionally, their amplitude is low and generally superimposed by the QRS complex. For these reasons, for clinicians, detecting F-waves by visual inspection is a challenging task.
Recently, machine learning techniques, and in particular deep neural networks, achieved high results in the evaluation and classification of biomedical images and signals and have started to be used to aid in the process of diagnosis and treatment of patients and augmenting physicians' capabilities. Especially, 2D convolutional neural networks (CNN) have achieved great success in the field of computer vision research and are currently widely used in image processing tasks due to their unique ability to capture position and translation invariant patterns and thus extract features from complex data and classify them. This may not be a viable option in numerous

applications over 1D signals, such as ECG, especially when the training data is scarce or application-specific. To address this issue, 1D CNNs have recently been proposed and immediately achieved state-of-the-art performance levels in several applications such as personalized biomedical data classification and early diagnosis, structural health monitoring, etc...
In this thesis, a two-stage deep learning method based on 1D CNNs and multipath modules is proposed to extract F-waves signals from recorded ECG of patients affected by atrial fibrillation. The performances of the method were evaluated in terms of the most used signal evaluation metrics.

Chapter 1

Clinical and Technical Background

1.1 Electrical event of the heart

1.1.1 Introduction

Even if the cardiac muscle shares a few characteristics with both skeletal muscle and smooth muscle, it has a unique property known as autorhythmicity. This is related to its ability to initiate an electrical potential at a fixed rate that spreads rapidly from cell to cell to trigger the contractile mechanism. Although cardiac muscle has autorhythmicity, heart rate is modulated by the endocrine and nervous systems.
There are two major types of cardiac muscle cells:

- myocardial contractile cells: 99 % of the cells in the atria and ventricles. They conduct impulses and are responsible for contractions that pump blood through the body.

- myocardial conducting cells: 1 % of the cells in the atria and ventricles that form the conduction system of the heart. Except for Purkinje cells, they are generally much smaller than the contractile cells and have few of the myofibrils or filaments needed for contraction. Myocardial conduction cells initiate and propagate the action potential (the electrical impulse) that travels throughout the heart and triggers the contractions that propel the blood.

Each embryonic heart cell can generate its own electrical impulse followed by contraction. When two independently beating embryonic cardiac muscle cells are placed together, the pace is set by the cell with the higher inherent rate, and the impulse spreads from the faster to the slower cell to trigger a contraction. As more cells are joined together, the fastest cell continues to assume control of the rate. A fully developed adult heart maintains the capability of generating its own electrical impulse, triggered by the fastest cells, as part of the cardiac conduction system.
The left and the right sides of the heart share the same electrical system for stimulation and control. The components of the cardiac conduction system (Figure 1.1) include the sinoatrial node, the atrioventricular node, the atrioventricular bundle, the atrioventricular bundle branches, and the Purkinje cells.

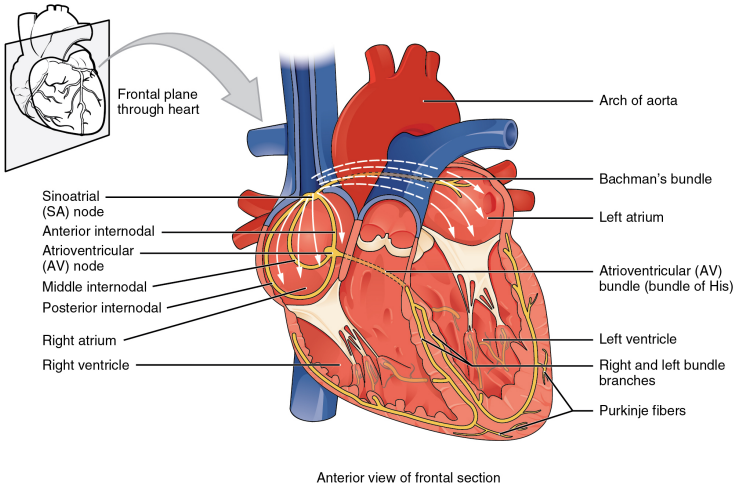

Figure 1.1: Conduction System of the Heart.

1.1.2 Sinoatrial Node

Normal cardiac rhythm is established by the sinoatrial (SA) node, a specialized clump of myocardial conducting cells (modified cardiocytes) located in the superior and posterior walls of the right atrium close to the orifice of the superior vena cava. The SA node is known as the pacemaker of the heart since generates the rhythmic pulse (sinus rhythm, or normal electrical pattern) followed by contraction of the heart through an action potential, i.e., an electrochemical signal that propagates as a travelling wave along the neurons. This impulse spreads from the SA node throughout the atria through specialized internodal pathways, to the atrial myocardial contractile cells and the atrioventricular node.
The internodal pathways are made of three bands (anterior, middle, and posterior) that directly link the SA node to the next node in the conduction system, the atrioventricular node. The impulse takes approximately 50 ms (milliseconds) to travel between these two nodes. The relative importance of this pathway has been debated since the impulse would reach the atrioventricular node simply following the cell-by-cell pathway through the contractile cells of the myocardium in the atria. Moreover, there is also a specialized pathway called Bachmann's bundle or the interatrial band that conducts the impulse directly from the right atrium to the left atrium. Regardless of the chosen pathway, when the impulse reaches the atrioventricular septum, the connective tissue of the cardiac skeleton prevents its dispersion into the ventricular myocardial cells, except for the atrioventricular node. This electrical event, that generates the wave of depolarization, is the trigger for muscular contraction of the atria. The wave of depolarization begins in the right atrium, and the impulse spreads from the superior portions of both atria and then down through the contractile cells, efficiently pumping blood into the ventricles.

1.1.3 Atrioventricular Node

The atrioventricular (AV) node is a second clump of specialized myocardial conductive cells, located in the inferior portion of the right atrium within the atrioventricular septum. The septum prevents the impulse from spreading directly to the ventricles without passing through the AV node. There is a critical pause before the AV node depolarizes and transmits the impulse to the atrioventricular bundle. This delay in transmission is partially related to the small diameter of the cells of the node, which slows the impulse, and also based on the fact that the conduction between nodal cells is less efficient than between conducting cells.
For these reasons, the impulse takes approximately 100 ms to pass through the node. This allows the atrial cardiomyocytes to complete their contraction that pumps blood into the ventricles before the impulse is transmitted to the cells of the ventricle itself. With extreme stimulation by the SA node, the AV node can transmit impulses maximally at 220 per minute. This establishes the typical maximum heart rate in a healthy young individual. Damaged hearts or those stimulated by drugs can contract at higher rates, but at these rates, the heart can no longer effectively pump blood.

1.1.4 Atrioventricular Bundle, Bundle Branches and Purkinje Fibers

Arising from the AV node, the atrioventricular bundle, or bundle of His, proceeds through the interventricular septum and then divides into two atrioventricular bundle branches, commonly called the left and right bundle branches. The left bundle branch has two fascicles and supplies the left ventricle, while the right bundle branch supplies the right ventricle. Since the left ventricle is much larger than the right, the left bundle branch is also considerably larger than the right. Both bundle branches descend and reach the apex of the heart where they connect with the Purkinje fibers. The impulse takes approximately 25 ms to reach them.
The Purkinje fibres are additional myocardial conductive fibres that spread the impulse to the myocardial contractile cells in the ventricles. They constitute a more elaborate network in the left ventricle than in the right. They extend throughout the myocardium from the apex of the heart toward the atrioventricular septum and the base of the heart. The Purkinje fibres have a fast inherent conduction rate, and the electrical impulse reaches all the ventricular muscle cells in about 75 ms. The contraction, as the electrical stimulus, begins at the apex and travels toward the base of the heart. This allows the blood to be pumped out of the ventricles and into the aorta and pulmonary trunk.
The total time from the initiation of the impulse in the SA node until the depolarization of the ventricles is approximately 225 ms [1]. Figure 1.2 shows the complete cardiac conduction cycle, from the initiation of the pulse in the SA to its diffusion throughout the atria and ventricles.

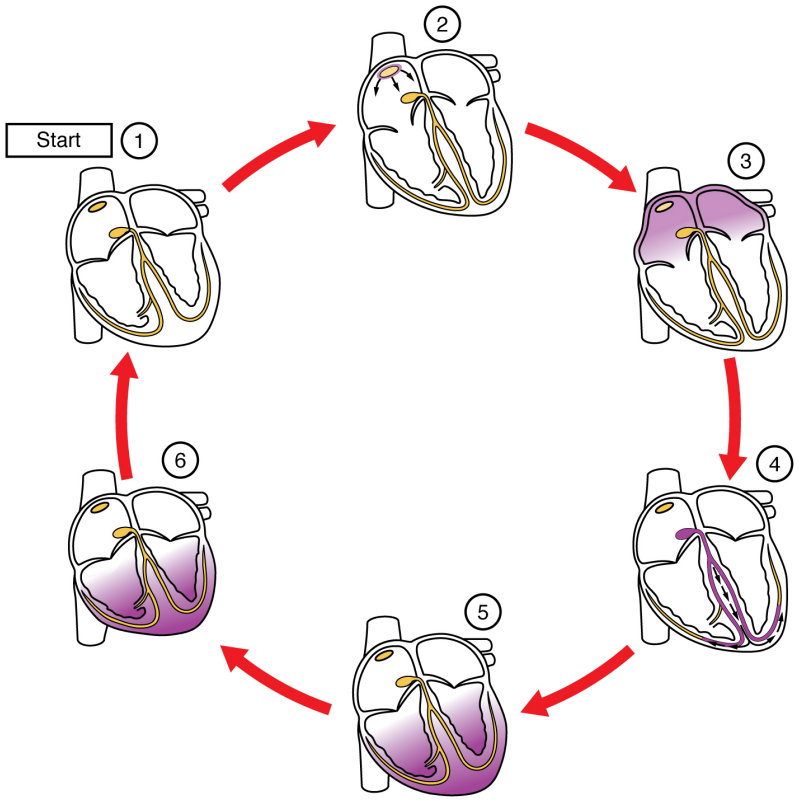

Figure 1.2: Cardiac conduction cycle.

1.1.5 Electrical event in pathological conditions

The normal electrical conduction system of the heart allows electrical impulses generated by the heart's pacemaker (the sinoatrial node) to spread to and stimulate the muscular layer of the heart (myocardium) in both the atria and the ventricles. Stimulation of the myocardium allows its contraction and, if this occurs in an orderly manner, this permits the blood to be pumped to the body. A normal heartbeat consists of a sequential contraction of atria followed by ventricles in a series of cardiac cycle events. The succession of 3 such regular heartbeats displaying identical waveforms leads to a steady rhythm.
Abnormal heart rate or rhythm, which is not physiologically justified, is known as arrhythmia. Arrhythmias are almost always pathological except for sinus arrhythmia, which is physiological. All pathological arrhythmias can be further classified based on heart rate into tachyarrhythmia (fast), bradyarrhythmia (slow), or tachy-brady (fast-slow) arrhythmia. All tachyarrhythmia originating above the ventricles, including atria and atrioventricular node (AV node), are grouped under supraventricular tachycardia (SVT). [2] Examples of SVT include atrial flutter, atrial fibrillation (AF), atrioventricular nodal reentrant tachycardia (AVNRT), also known as paroxysmal supraventricular tachycardia (PSVT), atrioventricular reentrant tachycardia (AVRT), and multifocal atrial tachycardia (MAT).[3]

1.2 Electrocardiography

1.2.1 Introduction

By means of the electrocardiogram (ECG), a measure of the changes in the electrical activity of the heart over time, it is possible to obtain a vision of the structure and functions of both healthy and diseased hearts. Over the years, the use of ECG has become a standard of care in cardiology, particularly for the detection of arrhythmias and acute myocardial infarction.
During the cardiac cycle, the heart contracts in response to a travelling electrical impulse in the form of action potentials that move through the cells of the atria and ventricles. According to the latter, at each stage of the cardiac cycle, there will be one part of the heart tissue that is depolarized and another part, at rest, that is polarized. This results in a charge separation, or dipole, which causes a fluctuating electric field throughout the body that can be detected via electrodes attached to the skin.
The use of electrodes on the surface of the skin to detect the voltage of this electrical field is what the electrocardiogram provides. The intensity of the voltage detected depends on the orientation of the electrodes with respect to that of the dipole ends.[4]
An ECG track can be obtained using a series of different positions or configurations of the electrodes (unipolar, bipolar, modified bipolar) that, however, have been standardized by universal application of certain conventions.
There are mainly three types of ECG recordings:

- Rest ECG: patients are simply required to lie down or sit up for the duration of the test which takes about 5 to 10 minutes. This is the most common type of ECG and one of the easiest to complete. The results recorded are typically reflective of the heart at rest.

- Exercise ECG, also called stress test: this test is done in controlled environments with the patient that walks on a treadmill or pedals on a stationary bike for about 10 to 20 minutes gradually increasing the intensity of the exercise. This ECG monitors the heart's capabilities and activity under physically demanding conditions, such as exercise.

- Holter monitor: this is a portable ECG used when there is the need to be monitored for an extended period. The Holter monitor may need to be worn for 24-72 hours, and sometimes up to 14 days to record any irregularities that may not be picked up during shorter ECG tests. To accurately record the heart's activity during this period, these devices can continuously record ECG and disperse the accumulated heat.

1.2.2 Device

Functionally, the electrocardiograph can be divided into 6 blocks: ECG acquisition, ECG signal processing, real-time ECG display, automated ECG interpretation, storage, and transmission of the ECG reports. Figure 1.3 shows a typical ECG machine.

Figure 1.3: ECG acquisition device.

The ECG signal is acquired through the patient module, made of a microprocessor, an analog application-specific integrated circuit, 10-lead wires connected to the patient using adhesive electrodes, and a cable back to the main part of the cardiograph.
The surface ECG is a low amplitude signal recorded in the presence of significant interference, so great care is necessary to obtain a quality result. The analog signal is filtered, amplified by a gain of 1000, and converted to a digital signal to allow computer processing. Note that careful attention to skin preparation and electrode adhesion has a large impact on reducing artifacts and interference.
The interference is actively controlled by the right leg drive system. This removes the interference by using a negative feedback loop in which a small current is driven into the right leg opposite to the common-mode signal until the measured one has been reduced as much as possible. Generally, the common-mode interference is dominated by the power line signal, with a frequency of 50 or 60 Hz. If the common-mode interference were a pure sine wave (the form of the signal in AC power), the right leg

drive system would output the opposite signal to delete the original common-mode sinusoid.
Another important part of the ECG device is the leads-off indication system. This warns the clinician when a lead wire or lead wires are either not connected or poorly connected to the patient, situations that will result in a poor-quality signal being recorded. The quality of the connection to the patient can be expressed as electrical impedance (i.e., the resistance to the flow of electrical current). When a lead is off, the resistance is infinite, so no current can flow. Special circuitry in the cardiograph seen in Figure 1.4 is devoted to measuring the resistance between the lead wires to assess the quality of the patient connection.[5]

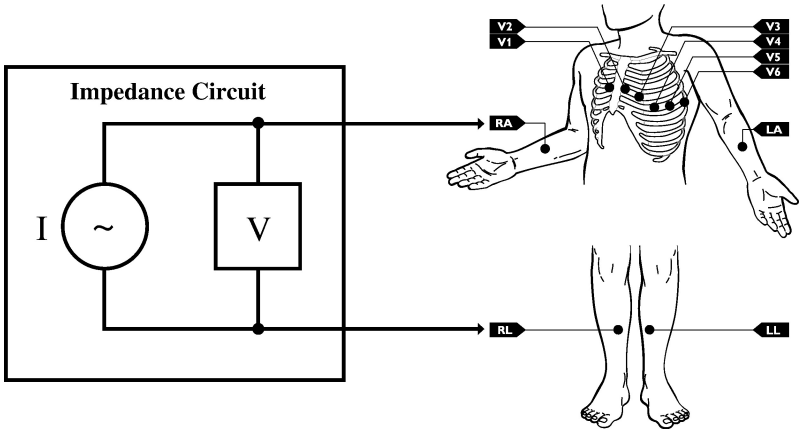

Figure 1.4: Schematic diagram of the impedance measurement used for leads-off indication.

Together with the patient module, there is the part devoted to the real signal pre-processing and registration. A traditional hospital electrocardiograph presents a main processor, real-time printer, real-time display, and a part to transmit the ECG for long-term storage.
Typical real-time ECG processing includes filtering, buffering, and heart rate detection. Many ECG applications use CPU time and computer resources in real-time so that the ECG signal processing can only use a fraction of what is available. That constraint does not exist for the traditional diagnostic ECG report because real-time processing is not required.

1.2.3 12-lead ECG

Commonly, 10 electrodes attached to the body are used to form the 12 ECG leads, with each lead measuring a specific electrical potential difference.
The three most used lead positions are referred to as lead I, II and III and form the so-called Einthoven triangle. These are the bipolar limb leads since they are placed in the two upper limbs and the left lower limb (at each vertex of the triangle). A

single ECG record (lead I, II, or III) is measured along the corresponding side of the triangle using the electrodes at both ends (Figure 1.5). Considering LA the voltage

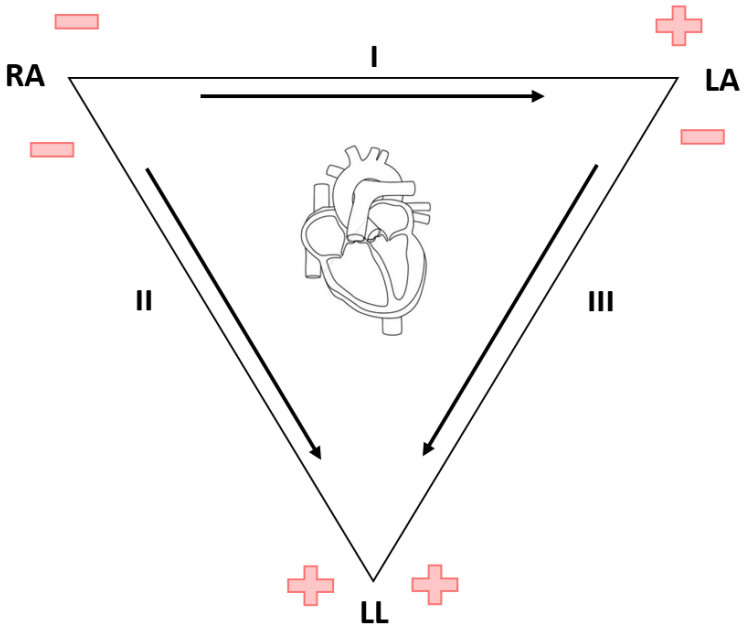

Figure 1.5: Einthoven's Triangle.

of the electrode on the left arm, RA the one on the right arm and LL the one on the left leg, the voltage of the described lead can be computed using the following equations:

$$I = LA - RA \tag{1.1}$$

$$II = LL - RA \tag{1.2}$$

$$III = LL - LA \tag{1.3}$$

The net dipole occurring in the heart at any step of the cardiac cycle is detected by each lead (I, II, and III) in a different way because of the different orientations of each lead set relative to the dipole in the heart. Each of these lead placements can be thought of as viewing the electrical dipole from three different directions: lead I from the top, lead II from the lower right side of the body, and lead III from the lower left side, all looking at the heart in the frontal plane.
Three other leads use the limb electrodes in a way that each of these uses a pair made by one limb electrode and a "neutral reference lead", known as Wilson's central terminal (V_W). These are referred to as the augmented unipolar limb leads. The voltage at V_W is produced by averaging the measurements from the electrodes RA, LA, and LL as follows:

$$V_W = \frac{1}{3}\,(RA + LA + LL) \tag{1.4}$$

The voltage recorded between the left arm limb lead and the neutral reference lead is called lead aVL; similarly, the right arm limb lead is aVR, and the left leg lead is aVF.

$$aVR = RA - \frac{1}{2}\,(LA + LL) = \frac{3}{2}\,(RA - V_W) \tag{1.5}$$

$$aVL = LA - \frac{1}{2}\,(RA + LL) = \frac{3}{2}\,(LA - V_W) \tag{1.6}$$

$$aVF = LL - \frac{1}{2}\,(RA + LA) = \frac{3}{2}\,(LL - V_W) \tag{1.7}$$

Together with leads I, II, and III, augmented limb leads aVR, aVL, and aVF form the basis of the hex axial reference system, which is used to calculate the heart's electrical axis in the frontal plane.
The last 6 are the chest leads, also known as the precordial leads. These leads are unipolar, and they measure the electrical activity of the heart in the traverse plane, instead of the frontal plane. For them, the neutral reference lead is "created," using all 3 limb leads connected to the negative ECG lead, which puts it in the center of the chest (Wilson's central terminal). The 6 positives, or "exploring", electrodes are labelled from V1 to V6.(Figure 1.6).

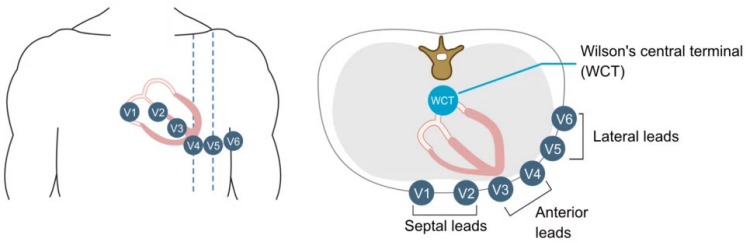

Figure 1.6: Precordial leads and Wilson's central terminal.

The 3 bipolar limb leads, 3 unipolar limb leads, and 6 precordial leads make up the 12-lead ECG.[4]

1.2.4 ECG signal

A recorded ECG represents the detected change in voltage of the electrical activity of the heart, which is normally displayed as millivolts (mV), with respect to time, expressed as seconds. A typical lead II ECG waveform shows a series of peaks and valleys that correspond to ventricular or atrial depolarization and repolarization.
The cardiac cycle begins with the firing of the sinoatrial node in the right atrium. This firing is not detected by the surface ECG since this signal does not have enough amplitude to be recorded with distal electrodes and it dissipates through the conductive medium. The atria then depolarize giving rise to the P-wave of around 80-100 ms in duration. As the P-wave ends, the atria are completely depolarized and begin their contraction. The signal then returns to baseline, and action potentials (not large enough to be detected) spread to the atrioventricular node and bundle of His.
Then, roughly 160 ms after the beginning of the P-wave, the right and left ventricles begin to depolarize, resulting in the QRS complex, which represents the beginning of ventricular contraction, which is around 80 (60- 100) ms in duration. Simultaneous with the QRS complex, the atria are repolarizing. The effect of this global atrial repolarization is sufficiently masked by the much larger amount of tissue involved in ventricular depolarization and, thus, is not normally detected in the ECG. During ventricular contraction, the ECG signal returns to baseline.
Then the ventricles, after contraction, repolarize, giving rise to the T wave. Note that the T-wave is normally the last-detected potential in the cardiac cycle and so followed by the P-wave of the next cycle. Figure 1.7 shows the recorded ECG of a healthy subject.
Sometimes also the so-called U-wave is detected after the T-wave, with the same polarity as this latter (positive deflection). This one has a much shorter amplitude and usually ascends more rapidly than it descends (which is the opposite of the T-wave). Even if its presence is not fully understood, it is considered by some to be caused by the late repolarization of the Purkinje system.[4] The frequency of an ECG signal is generally between 0.5 Hz and 100 Hz.[6]

1.2.5 Noise and Artifacts

ECG signal conveys a large amount of information about the structure of the heart and the function of its electrical conduction system. However, this information often interfered with noise generated during measurement which may lead to wrong interpretation. The noise can be divided into continuous and transient noise.
The continuous noise is associated with signals coming from all the leads with a similar temporal distribution but different intensity levels.
These noises dominate different frequency bands. The low-frequency range signifies baseline wander (BW), the medium frequency signifies the power line interference

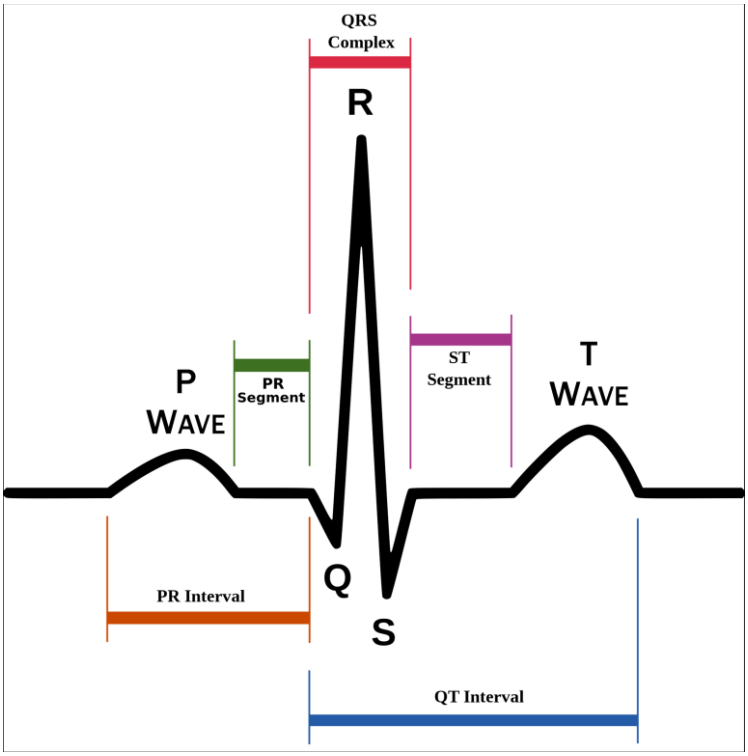

Figure 1.7: Electrocardiogram for a healthy subject during one cardiac cycle.

(PLI), and the high-frequency components signify the electromyography (EMG) noise.[7]

- Low frequency: Baseline Wander (BW) noise
  Baseline wander is a low-frequency noise component present in the ECG signal which is especially present in the exercise ECG and during ambulatory/Holter monitoring. This noise is mainly due to respiration, and body movement,[8] which affect the base x-axis of the ECG signal making it appear to 'wander' or move up and down rather than be straight. The drift of the baseline with respiration can be modelled as a non-stationary sinusoidal signal of time-varying amplitude and frequency of respiration. That is, the effect of baseline drift can be considered as an amplitude modulation to the ECG signal.[9] In other words, BW noise causes a shift of the entire signal from its normal base, which normally is a zero-mean signal (Figure 1.8). The frequency range of baseline wander is usually less than 1.0 Hz, but this range could be enlarged in real situations.

- Medium frequency: Power-line interference (PLI) noise
  The power-line interference represents a common noise source in the ECG as well as in any other bioelectrical signal recorded from the body surface. This noise is related to the electromagnetic susceptibility of the cables transporting signals and is characterized by a 50 or 60 Hz sinusoidal interference within

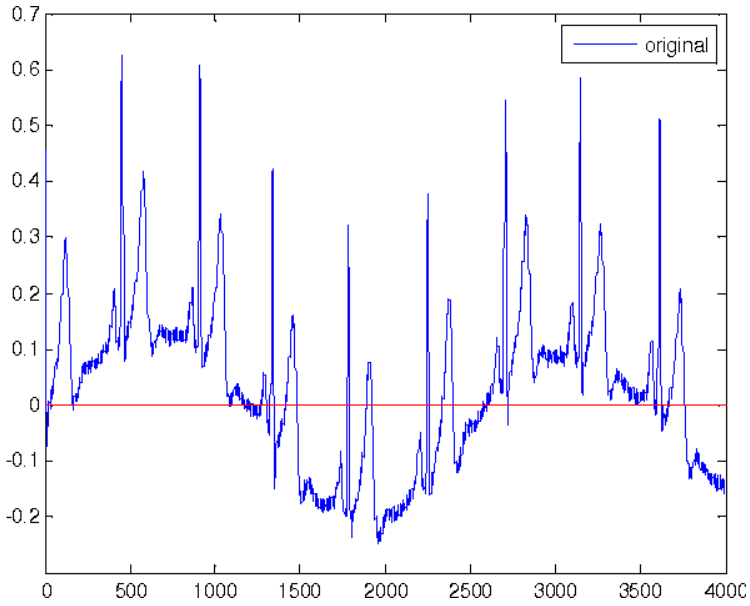

Figure 1.8: ECG signal with baseline wander artifact.

the ECG signals' frequency range, sometimes accompanied by harmonics. PLI noise is classified as a narrow band and can prevent the clinician from analyzing the ECG signal exactly, mainly due to the low-amplitude waveform making P-waves and T-waves unidentifiable boundary regions (Figure 1.9).[7]

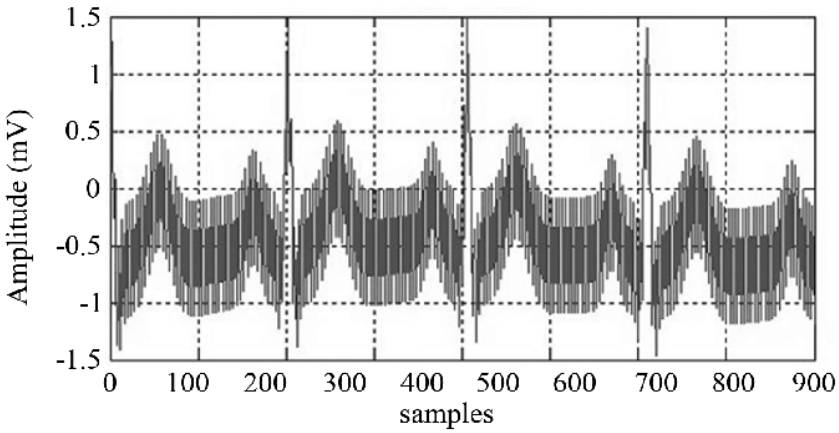

Figure 1.9: ECG signal with power-line interference.

- High frequency: Electromyography (EMG) noise
  The electromyographic (EMG) noise is caused by the detection in the ECG signal of the contraction of muscles different from the heart. When people move their bodies, muscles around electrodes will contract, generating depolarization and repolarization waves that will be detected by the electrodes and then appear in the ECG signals. The EMG signal is completely stochastic, and its amplitude depends on the intensity and frequency of the muscles' contraction

(Figure 1.10).[7] EMG noise consists of a maximum frequency of 10 KHz.

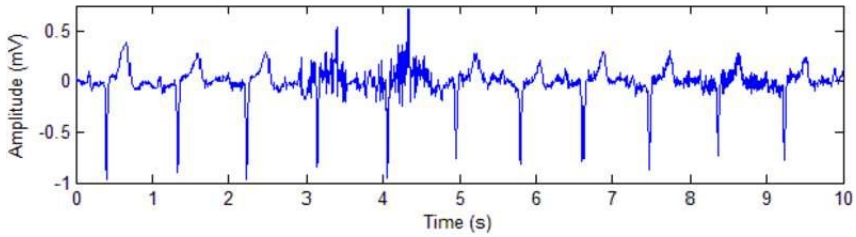

Figure 1.10: ECG signal with EMG noise.

On the other hand, the transient noise typically lasts for a short time and is categorized into white Gaussian noise (WGN). Since its instantaneous value indicates Gaussian distribution and power, its spectral density is distributed uniformly.[7] Examples of this type of noise are patient electrode motion artifact and instrumentation noise. Different from continuous noise, this cannot be defined in terms of frequency.

- Patient electrode motion artifact
  Motion artifact is the noise that results from the motion of the electrode in relation to the patient's skin. This artifact is usually the most difficult type of noise to detect because its spectrum completely overlaps that of the ECG, and its morphology often resembles that of P, QRS, and T-waves (Figure 1.11).[10] Therefore, it is hardly detected when only using the ECG signal.

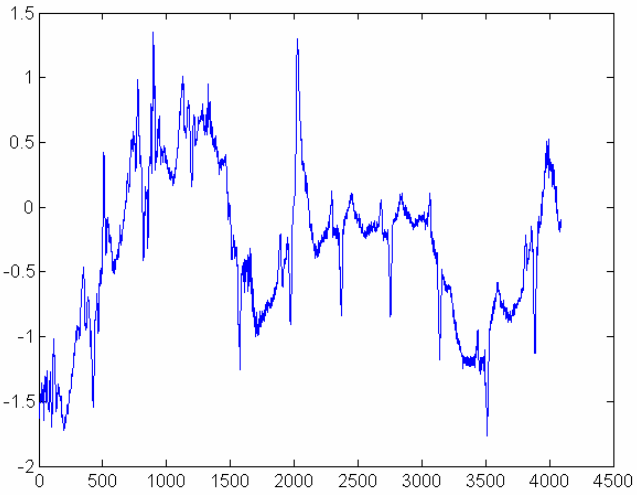

Figure 1.11: ECG signal with electrode motion artifact.

- Instrumentation noise
  The electrical device related to the ECG signal also contributes to noise. All the

components in equipment like the electrodes, cables, amplifiers, and converters are the major sources of instrumentation noise. This noise cannot be eliminated and can only be reduced by applying high-quality equipment and a well-designed circuit.[7]

1.2.6 Filtering techniques

The usual pre-processing stage of the ECG signal consists of using filters to attenuate noises and artifacts components while preserving the morphological characteristics important in the diagnostic interpretation of the signal. Different filters can be employed, according to the type of noise that needs to be removed:

- Finite impulse response (FIR) filters:
  In FIR design, the output of the filter is the weighted sum of past input values which is finite and can be represented by Eq. 1.8:

$$Y[n] = \sum_{k=0}^{M} b_k x(n-k) \tag{1.8}$$

  where $x[n]$ denotes the input signal, b_k are the filter coefficients and $Y[n]$ is the output response.[10]
  FIR filters are simple and stable filters. The window method is the simplest FIR filter design method. Here all frequencies below the cut-off frequency are passed with unity amplitude and others are blocked. The different windows used are the Rectangular Window, Hanning window, Hamming window, and Blackman window. Using these windows, high pass filters and low pass filters are designed with cut-off frequencies of 3 Hz and 100 Hz to remove baseline wander (BW) noise and electromyography (EMG) noise respectively.[8] Band-stop FIR filters can be used also to remove PLI noise. In Figure 1.12 are shown examples of frequency response of ideal FIR filters.

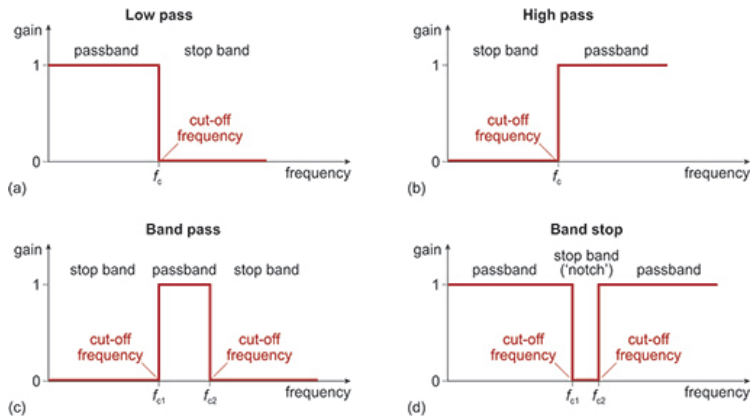

Figure 1.12: Frequency response of ideal FIR filters.

- Infinite impulse response (IIR) filters:
  IIR filter has infinite impulse response and acts like a feedback loop which never terminates when a single impulse is applied to it. It has both zeros and poles in the system. IIR filters may not be stable because of the infinite response. IIR filter can be mathematically expressed by Eq. 1.9:

$$Y[n] = \sum_{i=0}^{N} a_i x[n-i] + \sum_{j=1}^{N} b_j Y[n-j] \tag{1.9}$$

  where N is the filter's order, a_i and b_j are the filter coefficients and the output depends on past inputs and past outputs.[11]
  The stationary power line interference can be removed using a notch IIR filter. If a notch filter has a higher attenuation level, it will be able to remove PLI noise to a greater extent from the ECG signal, but practically it eliminates power line interference at 50 Hz or 60 Hz frequency.[8]

- Adaptive filters:
  An adaptive filter can adapt to the change in the signal over time. An adaptive filter has two input signals: one is the base input signal and the other is the reference signal. The filter compares them and calculates the error. The error is then minimized iteratively based on some objective function. Some popular algorithms for adaptive filters are Least Mean Square (LMS), Normalized Least Mean Square (NLMS) and Recursive Least Squares (RLS).[11] Figure 1.13 reports the block diagram of an adaptive filter, where $x[n]$ is the input signal, $d[n]$ is the reference signal and $e[n]$ is the computed error.

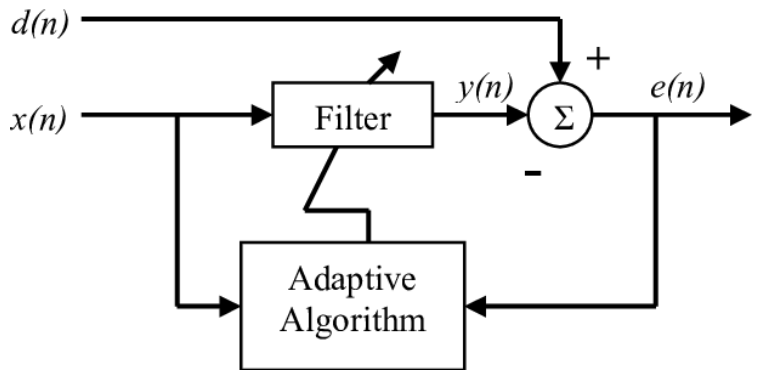

Figure 1.13: Block diagram of an adaptive filter.

The advantages of the adaptive filter method are that the filtering response is fast, and the residual errors are small, while its main drawback is that this method requires reference signal (either signal or noise characteristics) information for the effective filtering process.[8] Adaptive filters can be employed to remove motion artifacts and power line noise.

1.3 Atrial Fibrillation

1.3.1 Introduction

Atrial fibrillation (AF) is a pathology that generates an irregularly irregular heart rhythm. AF is the most common sustained supraventricular arrhythmia [12] caused by a dysfunction of the sinus atrial node, which is no longer able to guide atrial depolarization and, consequently, ventricular depolarization. Specifically, the normal regular electrical impulses generated by the sinoatrial node are overwhelmed by disorganized electrical waves, usually originating from the roots of the pulmonary veins. These disorganized waves conduct intermittently through the atrioventricular node, leading to irregular activation of the ventricles that generate the heartbeat.
As a result of the above effects, atrial contractility is lost causing an inability to completely empty blood from atrial appendage leading to the risk of clot formation and subsequent thromboembolic events. Typically, the heart rate varies from 120 to 160 beats per minute; however, a heart rate as fast as 200 beats per minute can be seen.[3]

1.3.2 Classification types

AF may be described in terms of the duration of episodes. Episodes often increase in frequency and duration over time.[12]

- Paroxysmal AF comes and goes, i.e. it begins and ends spontaneously. The irregular heartbeat may last anywhere from several seconds to a week and episodes may recur with variable frequency. However, most episodes of paroxysmal AF resolve themselves within 24 hours. Paroxysmal AF may be asymptomatic, which means without any apparent symptoms. The first line of treatment for asymptomatic paroxysmal AF may be lifestyle changes, such as reducing daily caffeine consumption and reducing stress, in addition to medications as preventative measures.[13]

- Persistent AF also begins spontaneously. It lasts at least seven days and may or may not end on its own. Medical intervention such as cardioversion, in which the clinician shocks the heart into rhythm, may be needed to stop an acute, persistent AF episode. Lifestyle changes and medications may be used as preventive measures.

- Long-standing persistent AF lasts at least a year without interruption. It's often associated with structural heart damage. This type of AF can be the most challenging to treat since usually the medications to maintain a normal heart rate or rhythm are often ineffective. More invasive treatments, such as electrical cardioversion, catheter ablation or pacemaker implantation, may be needed.

- The term "permanent AF" is used when the patient and clinician make a joint decision to stop further attempts to restore and/or maintain sinus rhythm. Acceptance of AF represents a therapeutic attitude on the part of the patient and clinician rather than an inherent pathophysiological attribute of AF. According to [14] this type of AF may result in more severe symptoms, lower quality of life, and an increased risk of a major cardiac event.

Further classification is done in terms of:

- Nonvalvular AF, in the absence of rheumatic mitral stenosis, a mechanical or bioprosthetic heart valve, or mitral valve repair.
- Valvular AF, when results from a problem with a heart valve, such as mitral valve stenosis, in which not enough blood can pass from the left atrium into the left ventricle.

Clinicians must determine whether a valvular problem is responsible for AF before they recommend a treatment plan. Newer medications that aim to prevent blood clots are available, but the Food and Drug Administration (FDA) has not approved them as safe and effective treatments for nonvalvular AF.[12]
As described above, the characterization of patients with AF by the duration of their AF episodes has clinical relevance in those outcomes of therapy such as catheter ablation are better for paroxysmal AF than for persistent AF. [15]
Furthermore, both paroxysmal and persistent AF may occur in a single individual. "Lone AF" is a historical descriptor that has been variably applied to younger persons without clinical or echocardiographic evidence of cardiopulmonary disease, hypertension, or diabetes mellitus.[16] Because the definitions are variable, the term lone AF is potentially confusing and should not be used to guide therapeutic decisions.

1.3.3 Epidemiology

The incidence and prevalence of AF are increasing worldwide. According to data from the FHS (Framingham Heart Study), the prevalence of AF increased 3-fold over the last 50 years. The Global Burden of Disease project estimated the global prevalence of AF around 46.3 million individuals in 2016. Europe has a higher prevalence of AF compared to the United States.
Similar to the prevalence, the incidence of AF increases with age. In all age groups, males are more commonly affected than females. Despite a high prevalence of risk factors, African Americans tend to have lower AF incidence compared to Caucasians.[17]
The lifetime risk of AF was estimated at 1 in 4 among white men and women older than 40 years in 2004; a decade later, lifetime risk estimates reached about 1 in 3 in white individuals and 1 in 5 for black individuals.
In the United States alone, at least 3 to 6 million people have AF, and this number is expected to rise from 6 to 16 million by 2050. Figure 1.14 shows the projected

numbers of patients with AF in the United States by 2050 according to different studies. In Europe, prevalent AF in 2010 was around 9 million among individuals older than 55 years and is expected to reach 14 million by 2060. It was estimated that by 2050 AF will be diagnosed at least in 72 million individuals in Asia, of which 3 million with AF-related strokes.
Awareness and enhanced detection of AF have improved over the past decade, which is important since about one-third of the total AF population is asymptomatic. Therefore, the global AF burden is certainly underestimated.[18]

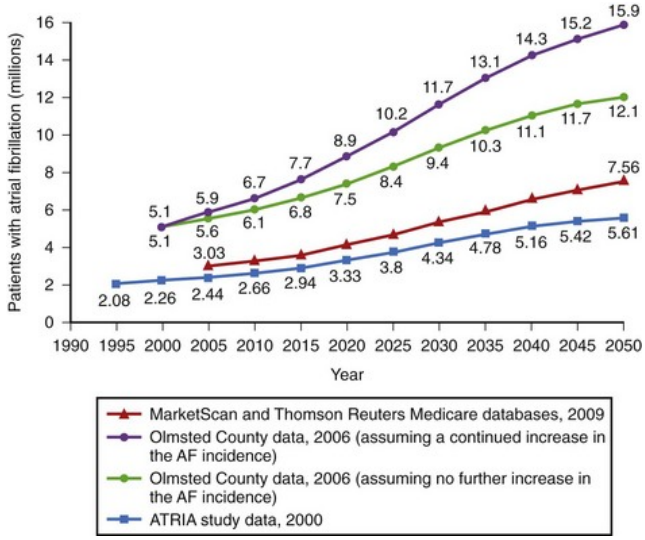

Figure 1.14: Estimates of the number of individuals with atrial fibrillation in the United States by 2050.

1.3.4 Etiology

Atrial fibrillation is commonly associated with conditions that alter the structure of the heart. Important causes and risk factors for AF are as follows [3]:

Cardiac Causes:

- Hypertensive heart disease
- Coronary artery disease
- Valvular heart disease
- Heart failure
- Congenital heart disease
- Cardiomyopathy

- Infiltrative cardiac disease
- Sick sinus syndrome
- Pre-excitation syndrome

Non-Cardiac Causes:

- Chronic lung disease
- Pulmonary embolism
- Electrolyte abnormalities
- Acute infections
- Thyroid disorders
- Pheochromocytoma
- Hypothermia
- Post-surgical (seen in 35% to 50% of patients post coronary artery bypass graft)

Risk Factors:

- Age-related fibrosis
- Diabetes
- Obesity
- Metabolic syndrome
- Obstructive sleep apnea
- Chronic kidney disease
- High-intensity exercise
- Genetic factors

1.3.5 Pathophysiological and electrophysiological mechanism

AF occurs when structural and/or electrophysiological abnormalities alter atrial tissue to promote abnormal impulse formation and/or propagation. These abnormalities can be related to several pathophysiological mechanisms, such that AF represents a common phenotype for multiple disease pathways and mechanisms that are incompletely understood.[19]

The progressive fibrosis of the atria is the primary pathological change seen in AF. This fibrosis is primarily due to atrial dilation, but also to genetic causes and inflammation. Dilation of the atria is a subsequence of a rise in the pressure within the heart related to several diseases, such as valvular heart disease (mitral stenosis, mitral regurgitation, and tricuspid regurgitation), hypertension, and congestive heart failure. Any inflammatory state that affects the heart can cause fibrosis of the atria. Mutation of the lamin AC gene is also associated with fibrosis of the atria which can lead to atrial fibrillation.
Fibrosis is not limited to the muscle mass of the atria and may occur in the sinus node (SA node) and atrioventricular node (AV node), correlating with sick sinus syndrome. Prolonged episodes of atrial fibrillation have been shown to correlate with prolongation of the sinus node recovery time [19], this suggests that dysfunction of the SA node is progressive with prolonged episodes of atrial fibrillation.
Along with fibrosis, also alterations in the electrical properties of the atria can be related to atrial fibrillation, such as their ability to be stimulated by the autonomic nervous system.
Several hypotheses have been proposed to explain the electrophysiological mechanisms that initiate and maintain AF.[15]
An important theory is that the regular impulses produced by the sinus node for a normal heartbeat are overwhelmed by disorganized electrical waves, usually originating from the roots of the pulmonary veins.
Sources of these disturbances are either automatic foci, often localized at one of the pulmonary veins or a small number of localized sources in the form of either a re-entrant leading circle or electrical spiral waves (rotors). Figure 1.15 shows the different mechanisms of conduction of the electric impulse in the heart of healthy and affected by atrial fibrillation individuals. Three fundamental components favour the establishment of a leading circle or a rotor: slow conduction velocity of the cardiac action potential, a short refractory period, and a small wavelength. If the action potential has fast conduction, with a long refractory period and/or conduction pathway shorter than the wavelength, an AF focus would not be established. In multiple wavelet theory, a wavefront will break into smaller daughter wavelets when encountering an obstacle, through a process called vortex shedding. But, under the proper conditions, such wavelets can reform and spin around a center, forming an AF focus.[20]
Although the electrical impulses of AF occur at a high rate, most of them do not result in a heartbeat. A heartbeat results when an electrical impulse from the atria passes through the atrioventricular (AV) node to the ventricles and causes them to contract. During AF, if all of the impulses from the atria passed through the AV node, there would be severe ventricular tachycardia, resulting in a severe reduction of cardiac output. This dangerous situation is prevented by the AV node since its limited conduction velocity reduces the rate at which impulses reach the ventricles during AF.[21]

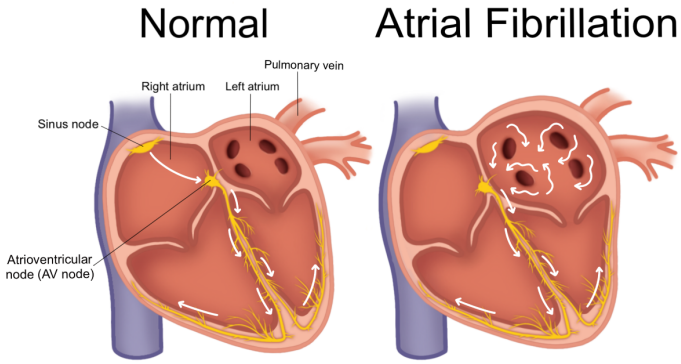

Figure 1.15: Conduction of the electrical impulse in a healthy and affected by AF heart.

1.3.6 Diagnosis by ECG

Even if early detection of AF is important to ensure timely management of the condition and avoid the recurrence of the arrhythmia as much as possible, several studies reveal that ∼20% of AF cases remain undiagnosed. This condition is more common in the elderly.[22] The diagnosis of AF requires rhythm documentation acquired by a single-lead ECG tracing of $\geq$ 30 s or 12-lead ECG showing AF analyzed by a physician with expertise in ECG rhythm interpretation. When AF detection is not based on an ECG recording (e.g. with devices using photoplethysmography) or in case of uncertainty in the interpretation of device-provided ECG tracing, a confirmatory ECG diagnosis has to be obtained using additional ECG recording, e.g., Holter monitoring.[23] According to [24], in patients presenting with ischemic stroke and with no prior history of AF, 72-hour Holter monitoring improves the detection rate of silent paroxysmal AF.
Most of the works on the analysis of ECG records for AF detection are based on heart rate variability, i.e., on the R-R intervals. RR intervals exhibit an irregularly irregular behaviour during AF episodes because the heart's activity is completely arrhythmic.
However, the most relevant information in atrial arrhythmias is contained in the P-waves, which consist of the depolarization of the atria. During AF episodes, the P-wave cannot be detected or is replaced by irregular fibrillatory waves: F-waves. F-waves (Figure 1.16) can have different shapes, amplitude, and duration in every patient. Moreover, the P/F-waves have lower amplitude and can be also superimposed by the QRS complex. For these reasons P-wave analysis for AF detection is much more challenging.[22]

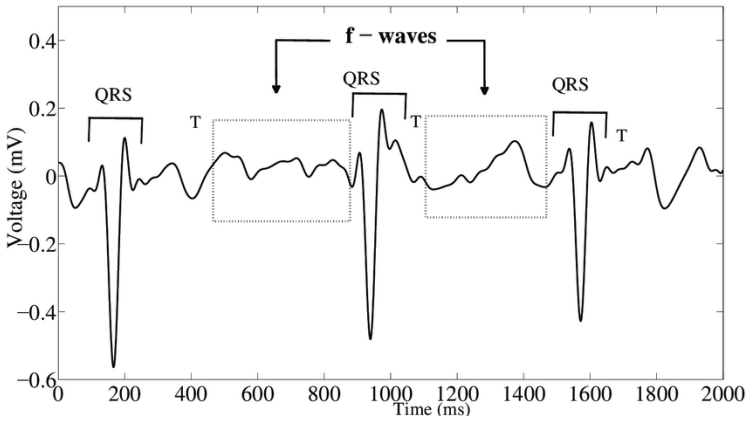

Figure 1.16: Characteristic F-waves in the ECG of AF patients.

1.4 Artificial Intelligence and Deep Learning

1.4.1 Introduction

The term Artificial Intelligence (AI) was coined by John McCarthy in 1956 and refers to the possibility of machines being able to simulate human behaviour and think.
Over the years, the computational power has grown to the point of instant calculations and the ability to evaluate new data, according to previously assessed data, in real-time. For this reason, today, AI is integrated into our daily lives in many forms and has also begun to be incorporated into medicine to improve patient care by speeding up processes and achieving greater accuracy. For instance, machine learning algorithms can be used to evaluate medical images (X-ray, MRI, etc..) or signals (ECG, EMG, etc...) to aid in the process of diagnosis and treatment of patients and augmenting physicians' capabilities.[25]
Anyway, conventional machine-learning techniques are limited in their ability to process natural data in their raw form. For decades, constructing a pattern recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image or samples of a signal) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input.
Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification.
Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned.[26]
Deep learning methods can be used to solve problems of classification, clustering

or regression and so they can be "supervised", "semi-supervised" or "unsupervised" algorithms.

1.4.2 Artificial Neural Network

Artificial Neural networks (ANN) are the basis of deep learning techniques with representation learning that attempt to imitate the way a human brain works.
In its simplest form, a biological brain is a huge collection of neurons. Each neuron takes electrical and chemical signals as inputs through its many dendrites and transmits the output signals through its axon. Axons contact other neurons at specialized junctions called synapses where they pass on their output signals to other neurons to repeat the same process over and over millions and millions of times.
Taking inspiration from the brain, an ANN is a collection of connected units, called artificial neurons.[27]
The first artificial neuron was modelled in 1943 by McCulloch and Pitts as a switch that receives input data and, depending on the total weighted input and the bias, is either activated or remains inactive. The decision of the state is devoted to the activation function of the neuron. Making a parallelism with the human brain, the weight, by which an input is multiplied, corresponds to the strength of a synapse. These weights can be both positive (excitatory) and negative (inhibitory).[26] The bias instead, allows to shift the activation function by adding a constant to the inputs (it can be thought of as analogous to the role of a constant in a linear function, whereby the line is effectively transposed by the constant value).
Figure 1.17 shows the structure of an artificial neuron.

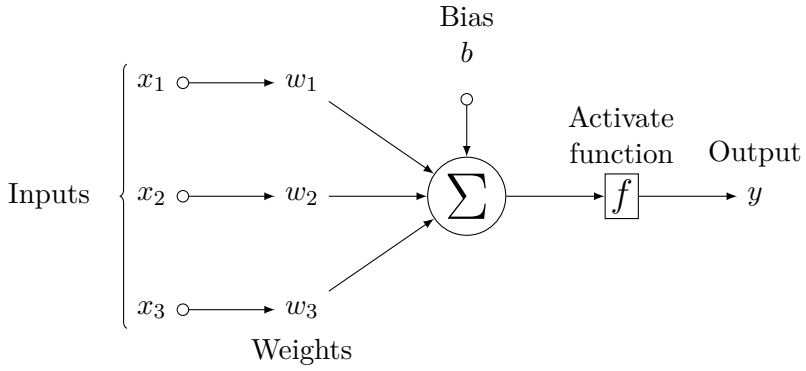

Figure 1.17: Artificial neuron model.

According to the previous definition, Eq. 1.10 relates the input with the output of an artificial neuron.

$$Y_k = \Phi \left(\sum_n x_n \cdot W_{kn} + b_k \right) \tag{1.10}$$

Where Y_k is the output, x_n are the n input, W_{kn} are the weights associated with each input, b_k is the bias and Φ represents the activation function of the neuron.

A typical neural network has from a few dozen to hundreds, thousands, or even millions of artificial neurons, also called units, arranged in a series of layers, each of which connects to the layers on either side. Some of them, known as input units, are designed to receive the input data from the outside world, other units, sit on the opposite side of the network, represent the response of the network and are known as output units. In between the input units and output units are one or more layers of hidden units. Most neural networks are fully connected, which means each hidden unit and each output unit is connected to every unit in the layers on either side.[27] The connections between one unit and another are represented by the weights W_{kn}. The higher the weight, the more influence one unit has on another. A simple fully connected Neural Network can be designed as shown in Figure 1.18.

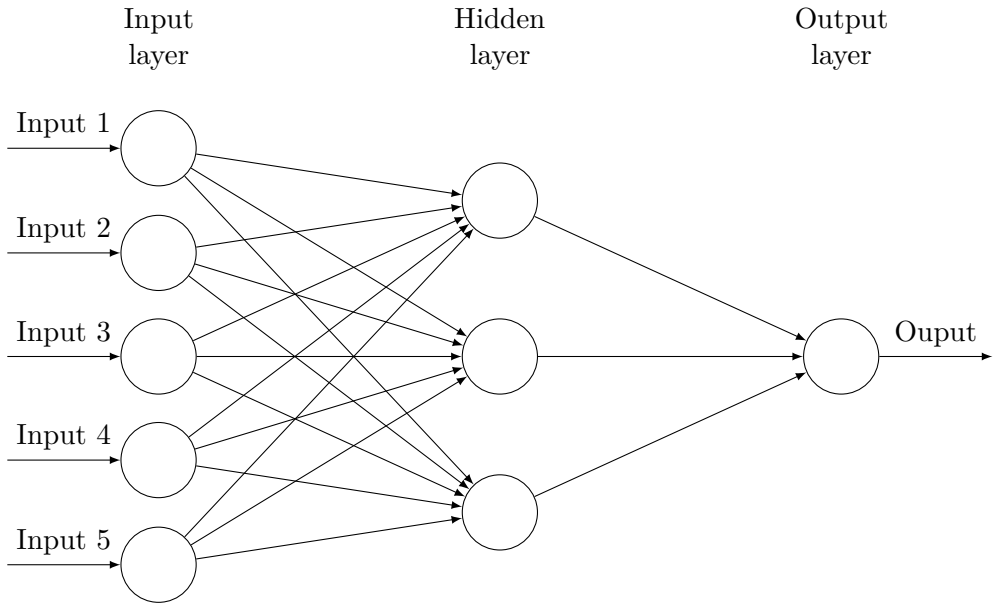

Figure 1.18: Architecture of a fully connected neural network.

1.4.3 Activation functions

If an activation function is not used in a neural network, then the output would simply be a simple linear function (which is a polynomial of degree one). Although linear equations are simple and easy to solve, their limited complexity gives no ability to the networks to learn and recognize complex mappings from data.
In literature, different kinds of activation functions exist, but there is no thumb rule for one of them. The choice of activation function is context-dependent, i.e., it depends on the task that is to be accomplished. The most used activation functions are [28]:

- Binary Step Function
  The Binary Step Function is the simplest activation function that exists, and it simply implements an if-else condition for the input. It is generally used in binary classifiers, but it cannot be used in the case of multiclass classification. The main problem of this activation function is that its gradient is zero, which may cause a problem during the backpropagation step in the training phase of the network. Mathematically the binary step function can be defined by Eq. 1.11.

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 1 \end{cases} \tag{1.11}$$

  Figure 1.19 shows a plot of this function.

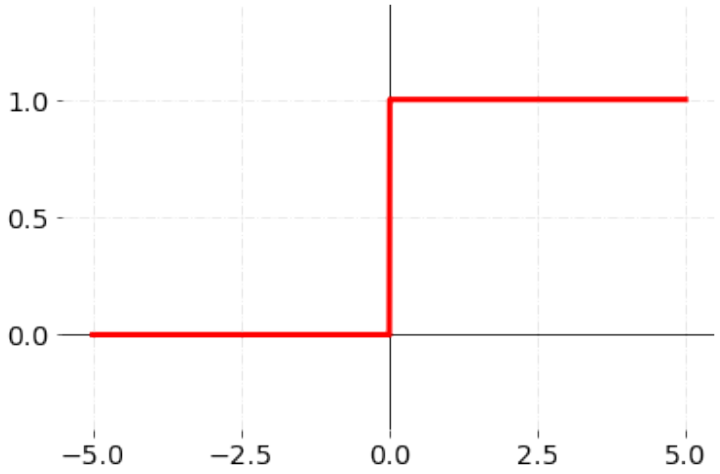

Figure 1.19: Binary step activation function.

- Linear Function
  The linear activation function is directly proportional to the input. To remove the problem of the zero gradient of the binary step function, the linear function can be used. It can be defined by Eq. 1.12.

$$F(x) = a \cdot x \tag{1.12}$$

  The value of variable a can be any constant value chosen by the user, and it will correspond to the value of the gradient. There isn't much benefit of using linear function, since the neural network would not improve the error due to the same value of gradient for every iteration and so it will not be able to identify complex patterns from the data. Therefore, linear functions are ideal where interpretability is required and for simple tasks.
  Figure 1.20 shows a plot of this function.

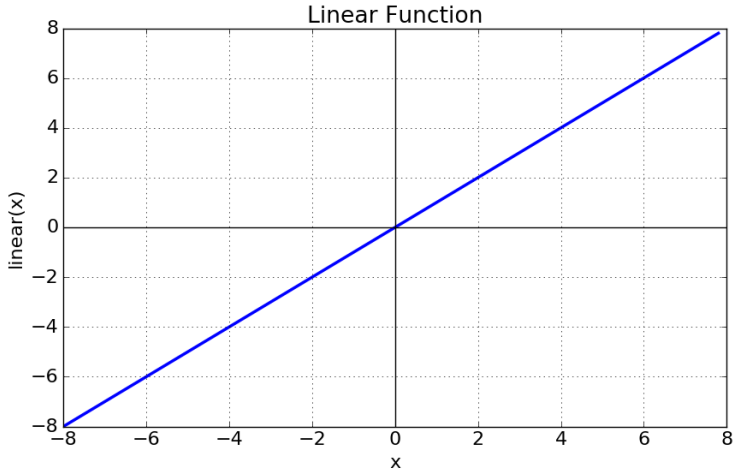

Figure 1.20: Linear activation function.

- Sigmoid Function
  It is the most widely used activation function as it is non-linear. The sigmoid function transforms the values in the range 0 to 1. It can be defined by Eq. 1.13.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1.13}$$

The sigmoid function is continuously differentiable and defined for all real input values. The derivative of the function is non-negative at each point. Moreover, the sigmoid function is not symmetric about zero, which means that the signs of all output values of neurons will be the same. This issue can be improved by scaling it. Figure 1.21 shows a plot of this function.

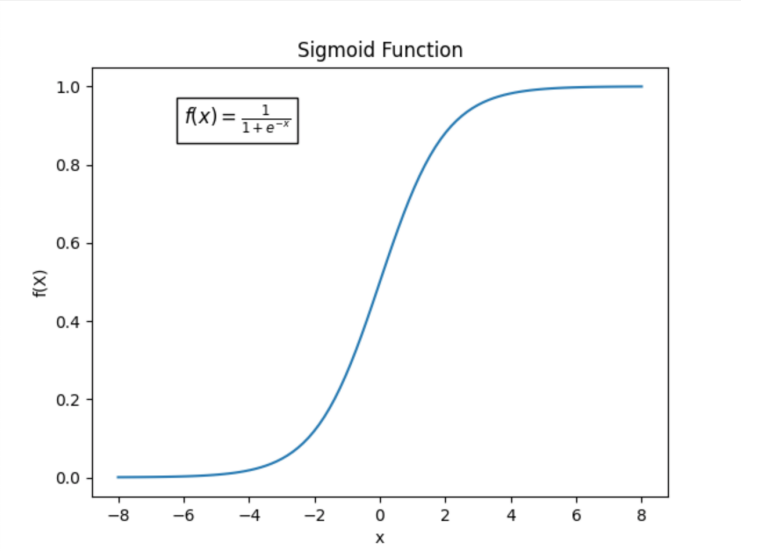

Figure 1.21: Sigmoid activation function.

- Hyperbolic Tangent function
  The hyperbolic Tangent function (Tanh) function is similar to the sigmoid function, but it is symmetric around the origin. This results in different signs of outputs. It can be defined by Eq. 1.14.

$$f(x) = 2 \cdot sigmoid(2x) - 1 \tag{1.14}$$

  Tanh function is continuous and differentiable, the values lie in the range -1 to 1. As compared to the sigmoid function the gradient of the tanh function is steeper. Tanh is preferred over sigmoid function as it has gradients which are not restricted to vary in a certain direction and, it is zero-centered. Figure 1.22 shows a plot of this function.

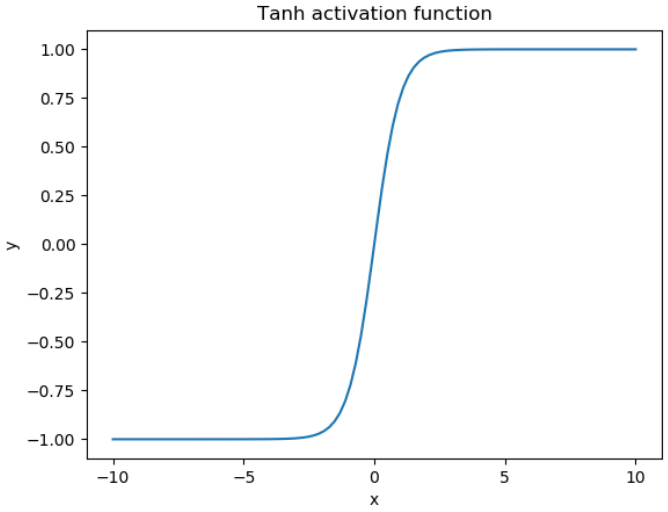

Figure 1.22: Hyperbolic Tangent activation function.

- ReLU function
  ReLU stands for rectified linear unit and is a non-linear activation function which is widely used in neural networks.
  ReLU is more efficient than other functions because all the neurons are not activated at the same time, rather a certain number of neurons are activated at a time. In some cases, the value of the gradient is zero, due to which the weights and biases are not updated during the back-propagation step in neural network training. It can be defined mathematically by Eq. 1.15.

$$f(x) = \max(0, x) \tag{1.15}$$

  Figure 1.23 shows a plot of this function.

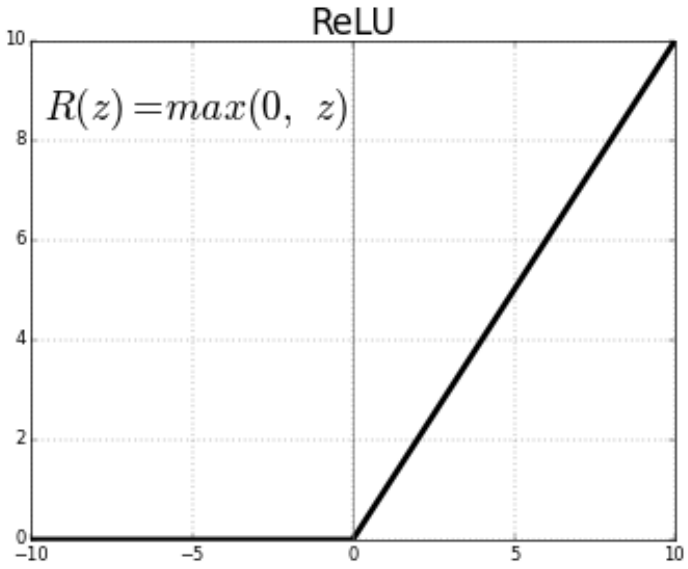

Figure 1.23: ReLU activation function.

- Exponential Linear Unit function
  Exponential Linear Unit or ELU is a variant of the Rectified Linear Unit. ELU introduces a parameter slope for the negative values of x. It uses a log curve for defining the negative values. It can be defined mathematically by Eq. 1.16.

$$\begin{cases} f(x) = x, & x \geq 0 \\ f(x) = a \cdot \left(e^{x-1}\right), & x < 0 \end{cases} \tag{1.16}$$

  Figure 1.24 shows a plot of this function.

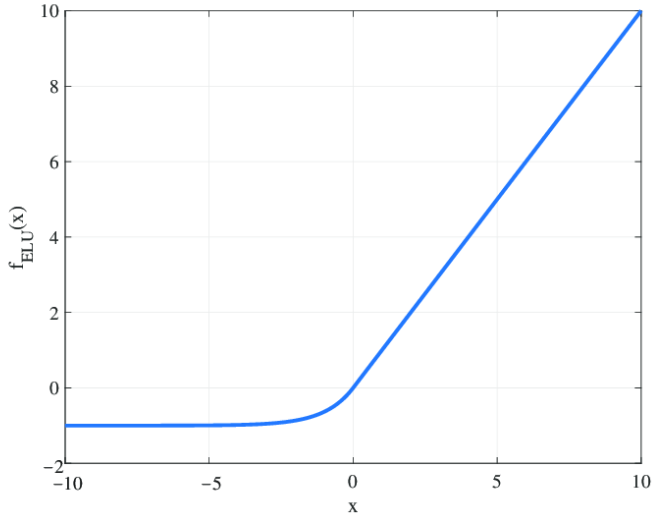

Figure 1.24: Exponential Linear Unit activation function.

- SoftMax function
  Softmax function is a combination of multiple sigmoid functions. Since a sigmoid function returns values in the range 0 to 1, these can be treated as probabilities of particular class data points. Softmax functions unlike sigmoid functions, which are used for binary classification, can be used for multiclass classification problems. The function, for every data point of all the individual classes, returns the probability. It can be defined mathematically by Eq. 1.17.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad for\ j = 1, \ldots, K \tag{1.17}$$

  Figure 1.25 shows a plot of this function.

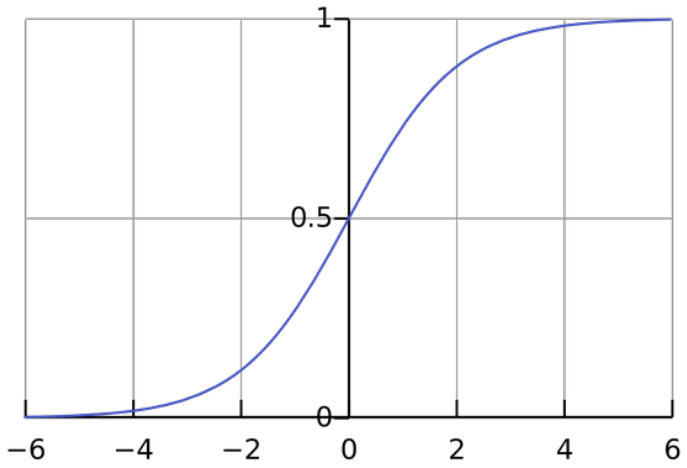

Figure 1.25: SoftMax activation function.

1.4.4 1D Convolutional Neural Network

During the last decade, Convolutional Neural Networks (CNNs) have become a standard for various Computer Vision and Machine Learning operations. CNNs are feed-forward Artificial Neural Networks (ANNs) with alternating convolutional and subsampling layers, also called pooling layers. Deep 2D CNNs with many hidden layers and millions of parameters can learn complex objects and patterns providing that they can be trained on a massive size visual database with ground-truth labels. With proper training, this unique ability makes them the primary tool for various engineering applications for 2D signals such as images and video frames. Thus, each neuron of this network contains 2-D planes for weights, which is known as the kernel, and input and output which is known as the feature map.
Yet, this may not be a viable option in numerous applications over 1D signals especially when the training data is scarce or application-specific. To address this issue, 1D CNNs have recently been proposed and immediately achieved state-of-the-art performance levels in several applications such as personalized biomedical

data classification and early diagnosis, structural health monitoring, etc... Another major advantage is that a real-time and low-cost hardware implementation is feasible due to the simple and compact configuration of 1D CNNs that perform only 1D convolutions (scalar multiplications and additions).
Different studies show that compact 1D CNNs have superior performance on those applications with limited labelled data and high signal variations acquired from different sources (i.e. patient ECG). [29]
The basic structure of a 1D CNN tries to imitate that of the older 2D CNN, and, as illustrated in Figure 1.26, it presents two distinct layer types:

1. the so-called "CNN-layers" where both 1D convolutions, activation function and sub-sampling (pooling) occur.
2. Fully connected (dense) layers that are identical to the layers of a typical Multi-layer Perceptron (MLP) and therefore called "MLP-layers". The scalar outputs of the last convolutional layer are forward propagated through the MLP-layers to produce the final output that represents the classification or regression of the input signal.

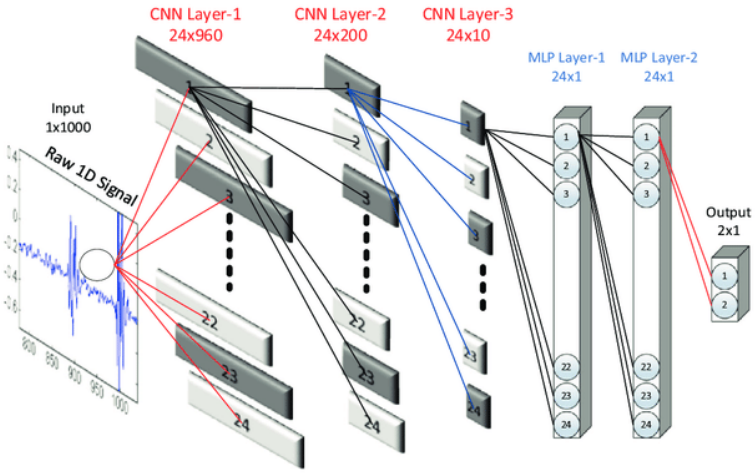

Figure 1.26: 1D CNN configuration with 3 CNN and 2 MLP layers.

The configuration of a 1D-CNN is formed by the following hyperparameters:

- Number of hidden CNN and MLP layers/neurons (in the sample 1D CNN shown in Figure 1.26, there are 3 and 2 hidden CNN and MLP layers respectively).
- Filter (kernel) size in each CNN layer (in the sample 1D CNN shown in Figure 1.26, the filter size is 41 in all hidden CNN layers).
- Subsampling factor in each CNN layer (in the sample 1D CNN shown in Figure 1.26, the subsampling factor is 4).
- The choice of pooling and activation functions.

As in the conventional 2D CNNs, the input layer is a passive layer that receives the raw 1D signal and the output layer is a MLP layer with the number of neurons equal to the number of classes.
Three consecutive CNN layers of a 1D CNN are presented in Figure 1.27. As shown in this figure, the 1D filter kernels have size 3 and the sub-sampling factor is 2 where the k^{th} neuron in the hidden CNN layer, l, first performs a sequence of convolutions, the sum of which is passed through the activation function, f, followed by the sub-sampling operation. This is indeed the main difference between 1D and 2D CNNs, where 1D arrays replace 2D matrices for both kernels and feature maps. As a next step, the CNN layers process the raw 1D data and "learn to extract" such features which are used in the classification task performed by the MLP-layers. Therefore, both feature extraction and classification operations are fused into one process that can be optimized to maximize the classification performance.
This is the major advantage of 1D CNNs which can also result in a low computational complexity since the only operation with a significant cost is a sequence of 1D convolutions which are simply linear weighted sums of two 1D arrays.[29]

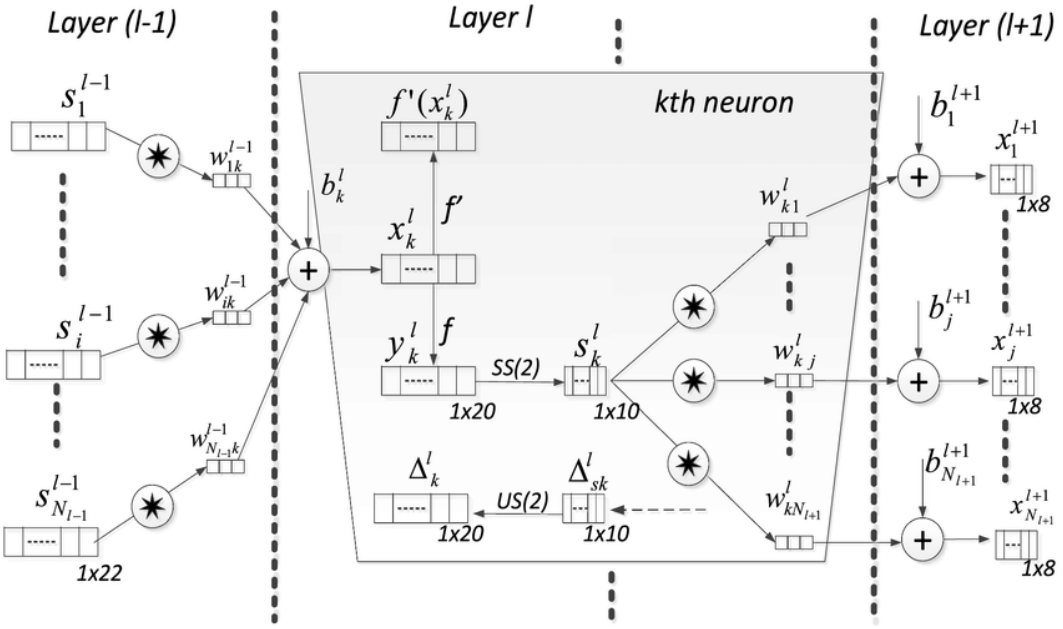

Figure 1.27: Three consecutive hidden CNN layers of a 1D CNN.

Thus, the types of layers generally found in a 1D CNN are:

1. Convolutional layers
   These are the essential blocks of the CNN. In each CNN layer, 1D forward propagation (1D-FP) is expressed by Eq. 1.18.

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv1D(w_{ik}^{l-1}, s_i^{l-1}) \tag{1.18}$$

where x_k^l is defined as the input, b_k^l is defined as the bias of the k^{th} neuron at layer l, s_i^{l-1} is the output of the i^{th} neuron at layer $l-1$, w_{ik}^{l-1} is the kernel from the i^{th} neuron at layer $l-1$ to the k^{th} neuron at layer l. $conv1D(.,.)$ is used to perform "in-valid" 1D convolution without zero-padding.

2. Pooling layers
   These are generally applied after each convolutional layer. Different functions can be used in this layer, but the max is the most used. A max 1-D pooling layer performs downsampling by dividing the input into 1-D pooling regions and then computing the maximum of each region. Thus, after the convolution is performed, the intermediate output of each layer of the network, y_k^l, can be expressed by passing the input x_k^l through the activation function of the neuron $f(.)$, and then applying the downsampling. These operations can be expressed by Eq. 1.19.

$$y_k^l = f\left(x_k^l\right) \; and \; s_k^l = y_k^l \quad \downarrow ss \tag{1.19}$$

   where s_k^l stands for the output of the k^{th} neuron of the layer, l, and "$\downarrow ss$" represents the downsampling operation with a scalar factor, ss. Pooling layer parameters are not trainable.

In most of the networks, Batch normalization layers are also present to standardize the inputs to a layer for each mini-batch during the training of the network, stabilizing the learning process and dramatically reducing the number of training epochs required.

1.4.5 Training artificial neural networks

Deep learning discovers intricate structures in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer.[26]
Backpropagation is an abbreviation for "the backward propagation of errors", since an error is computed at the output of the network and distributed backwards throughout the network's layers. It is commonly used to train deep neural networks. In the context of learning, backpropagation is commonly used by the gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function.[27]
In other words, the learning (training) process of a neural network (supervised algorithm) is an iterative process in which the calculations are carried out forward and backwards through each layer in the network until a loss function is minimized (Figure 1.28).
The entire learning process starts initializing the weights of the network. A typical strategy in neural networks is to initialize the weights randomly, and then start

optimizing. At each iteration, also called epoch, the predicted output of the network is used, together with the expected output, to compute the loss. Several loss functions can be found in the literature. The definition and choice of the best loss function for a network depends on the type of the problem and the nature of the input data.

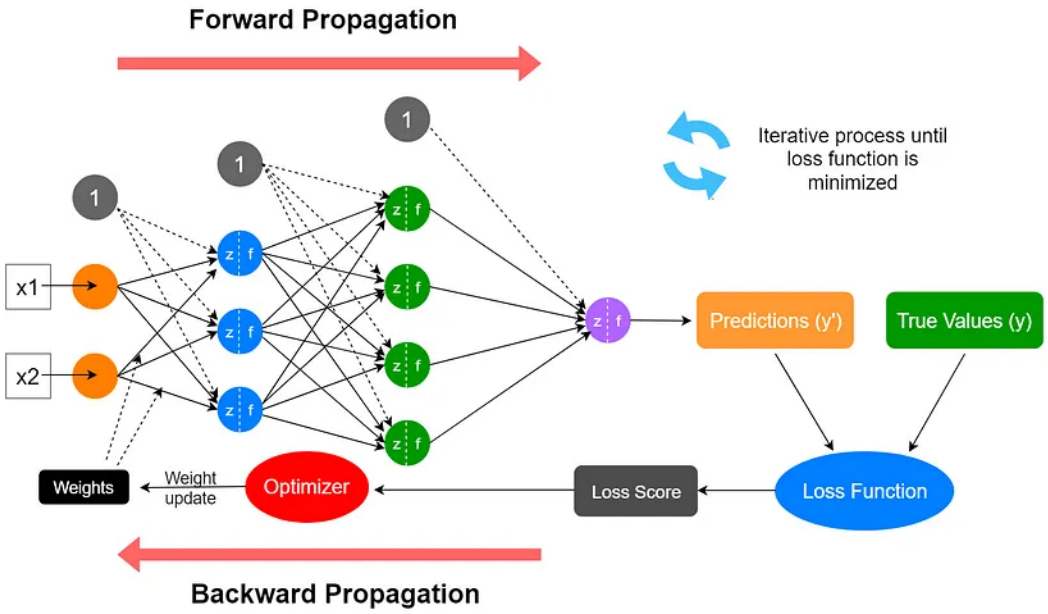

Figure 1.28: Training process of Artificial Neural Network.

1.4.6 Overfitting, Underfitting and Generalization problems

The generalization error of a machine learning model is the difference between the empirical loss of the training set and the expected loss of a test set. This measure represents the ability of the trained model (algorithm) to generalize well from the learning data to new unseen data. It is typically understood that good generalization is obtained when a machine learning model does not memorize the training data, but rather learns some underlying rule associated with the data generation process, thereby being able to extrapolate that rule from the training data to new unseen data and generalize well.[30]
On the opposite side, when the generalization error is too high, it could be related to two different situations:

1. Overfitting: this happens when the network works too hard to find the very best fit to the training data and the model learns their detail and noise. In other words, the noise or random fluctuations in the training data are picked up and learned as concepts by the model. The problem is that these concepts do not apply to new data and negatively impact the model's ability to generalize.

2. Underfitting: this refers to a model that can neither model the training data nor generalize to new data. An underfit machine learning model is not a suitable

model and will be obvious as it will have poor performance on the training data. This is an easy-to-detect situation and it could be solved by designing a different machine learning algorithm.

To reduce the effect of overfitting, multiple solutions based on different strategies are proposed to inhibit the different triggers.
One method, called "Early-stopping", has been widely used in iterative algorithms, especially in neural networks starting from the 1990s.[31] This strategy divides the initial set of training data into two subsets, one used to train the model (training data), and the other used to validate it (validation data). After each epoch of training, the network is validated using the data of the validation set and the relative training and validation errors (losses) are computed. Trends of training and validation error are reported in Figure 1.29.

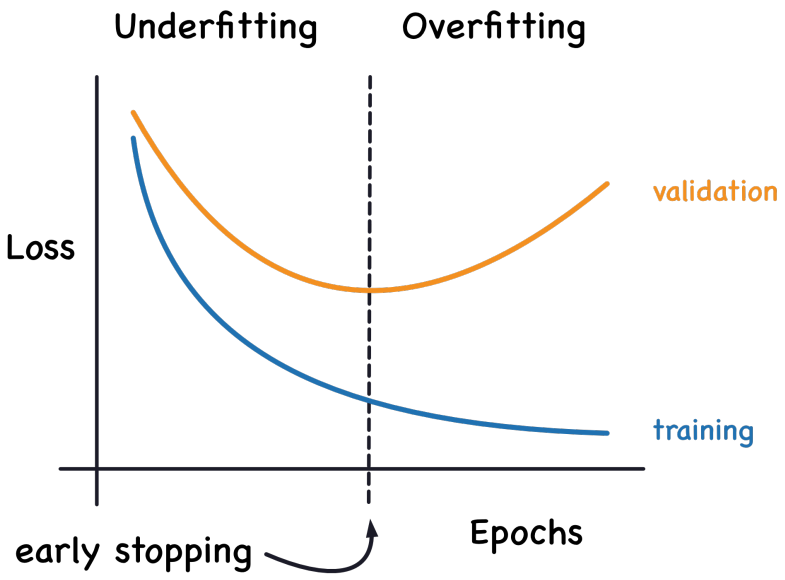

Figure 1.29: Overfitting and early stop point.

If the model continues learning, after a certain point known as "early-stopping", the validation error will start to increase, while the training error will continue decreasing. If we stop learning before that point, it's underfitting, after is overfitting. So the aim is to find the exact point to stop training. In other words, the validation set is used to figure out a perfect set of values for the hyperparameters, weights and biases of the network, while, later, the test set will be used to do the final evaluation of accuracy. Other techniques to solve the problem of overfitting try to reduce noises in the training set, as in the "reduce the size of network" strategy, or try to distinguish noises, meaning and meaningless features and assign different weights to them, as in the "Regularization" strategy.[31]

Chapter 2

Literature review

2.1 Introduction

Most of the recent works on atrial fibrillation and neural networks rely on the latter for the extraction of features useful to classify ECG signals as normal or affected by the various types of AF. No research has been found in the literature using neural networks to extract F-waves from the ECG signals of patients with atrial fibrillation. The identification and extraction of F-waves from the ECG signal is an arduous task, as the F-waves overlap the QRS complex and T-waves in both the time and frequency domain.
For this reason, it cannot be achieved by standard signal-filtering techniques. Since this study aims to create a filter using neural networks for the extraction of F-waves, the literature research has concentrated on papers that use deep learning algorithms to filter the ECG signal.

2.2 Method

The systematic literature search was conducted in two electronic bibliographic databases: PubMed and Scopus.
The root 'ecg' was used to search for studies about the methodology of signal acquisition. The roots 'denoising*', 'filter*' and 'elimination' were used to search for studies concerning the filtering techniques of the signal, and the roots 'network*', and 'deep learning' to search for studies referring to machine learning techniques. Moreover, studies that contain the keywords 'classification' and 'extraction' were excluded from the results.
Terms within each concept were combined with the Boolean operator 'OR' and then combined with the Boolean operator 'AND', except for the terms 'classification' and 'extraction' that were combined with the Boolean operator 'AND NOT'. As a limit for the field of search, 'Title' was used for all the concepts, expect for the exclusion criteria where 'Title/Abstract' was used. Moreover, the English language was used as a limit to filter the documents and only papers of the last 10 years were considered in the results.
Documents (research papers and conference articles) were imported into the Zotero

reference management system for duplicate removal. Titles and abstracts were analyzed to include only documents about the topic of interest.

2.3 Results

Overall, 43 studies were identified in the bibliographic and organizational databases; of these, 7 were duplicated. After title and abstract analysis, 14 studies were selected, from 2013 to 2023.

2.3.1 S. Poungponsri et al. (2013)

In this paper [9], a novel adaptive filtering approach based on wavelet transform and artificial neural networks is investigated for ECG signal noise reduction.
In this work, Daubechies 4-tap wavelet (D4) is employed. Once the wavelet transform coefficients are obtained, sub-band thresholding is then performed on these coefficients. This thresholding discards high-frequency noise and also performs feature extraction of the ECG signal to provide the inputs to the neural network. A neural network is then employed as the final filtering process to further remove the remaining noise that is "embedded" in DWT coefficients. In the meantime, the neural network effectively performs an "inverse discrete wavelet transform (IDWT)" at the output. That is, the inputs of the neural network are DWT coefficients while the output of the neural network is the filtered ECG signal in the time domain.
The overall system block diagram for neural network training is shown in Figure 2.1.

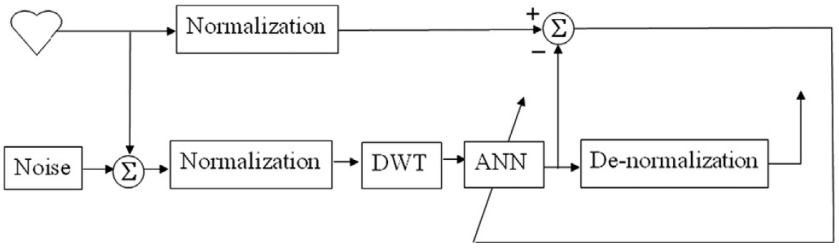

Figure 2.1: The neural network training scheme proposed by S. Poungponsri et al. (2013).

The proposed approach is tested on various noises and artifacts, including power-line interference, baseline wander noise, electrode motion artifact, muscle contraction artifact, and white noise. Computer simulation results show this approach can successfully remove them with significant SNR improvement in the ECG signal, over other algorithms which are typically limited to removing only one or two types of noise.

2.3.2 S. O. Rajankar et al. (2015)

In this paper [32] ECG denoising is achieved using a wavelet neural network by approximating the signal to the maximum possible accuracy.
The feed-forward back propagation neural network is designed to consist of four layers namely input, two hidden layers having 10 neurons each and an output layer. The first hidden layer has having bias vector and sigmoidal activation function. The input provides weights to the first layer. Each subsequent layer has weights coming from its previous layers. For the second hidden layer instead of using the sigmoidal activation function the various mother wavelets from library wavelets available in MATLAB toolbox such as daubachies, meyer, coiflet, etc. are used. The network architecture is shown in Figure 2.2.

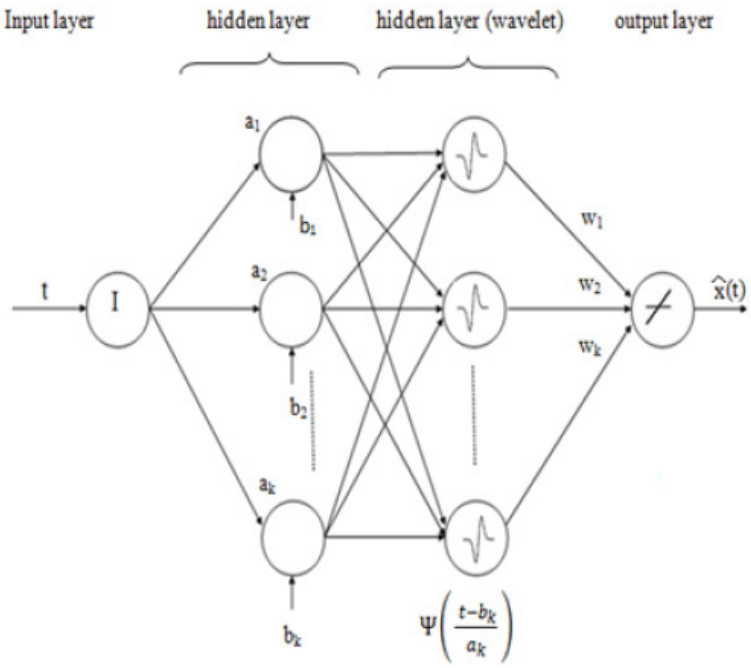

Figure 2.2: The neural network architecture proposed by S. O. Rajankar et al. (2015).

The input to the neural network are the samples of ECG signal with white Gaussian noise of various noise levels, while the ECG signal to be denoised is considered as the target. Performance is measured according to the mean square error (MSE) and the one achieved with db6 wavelet is found to be superior. Comparing the results with other standard methods is shown that the neural network approach gives a strong alternative to discrete wavelet transform with soft thresholding for denoising or compression of an ECG signal.

2.3.3 Yue Qiu et al. (2017)

This study [33] proposes a method for the elimination of 50-Hz power line interference (PLI) in ECG signals, using recurrent neural networks (RNN). It uses a deep Long Short-Term Memory (LSTM) RNN model with 3 hidden layers and each layer has 16 LSTM neurons, which could guarantee enough model capability to acquire the signal pattern.
A dense connection with linear activation is put on the top of the third hidden layer for the regression function.
Figure 2.3 shows the architecture of the network.

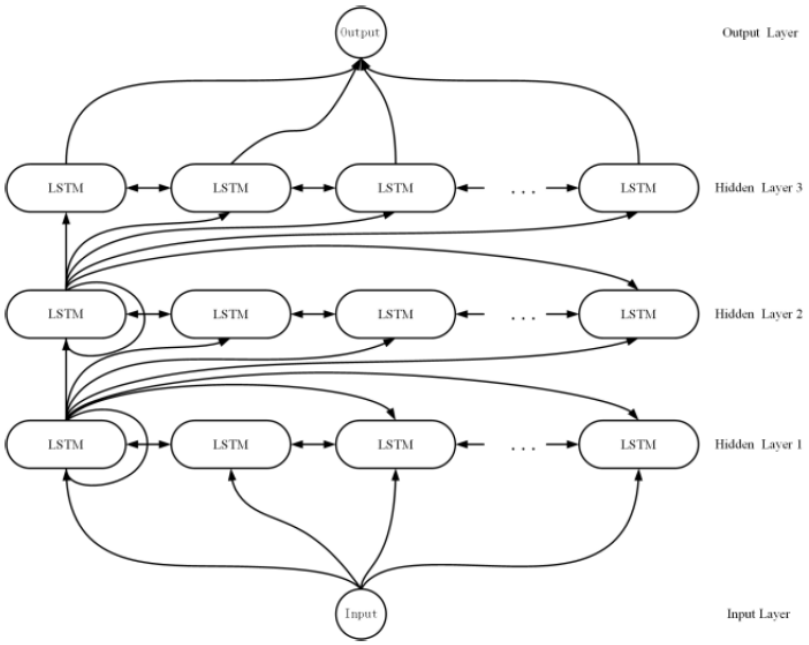

Figure 2.3: The neural network architecture proposed by Yue Qiu et al. (2017).

After the training procedure, the trained model was used to extract the PLI signals from contaminated ECG signals. Then the ECG signals are filtered by subtracting the extracted PLI signals.
To compare the effectiveness of the proposed method with the traditional linear filtering approach, a 50-Hz Butterworth notch filter (10th order with a center frequency equal to 50 Hz) was also implemented on the same signal. It could be noticed that the distortions of the QRS complex are negligible when filtered by the proposed method while the applied notch filter causes observable distortions at each QRS complex.

2.3.4 C. T. C. Arsene et al. (2019)

This paper [34] presents two DL models for denoising ECG signals.
First, a Convolutional Neural Network (CNN) is depicted and applied to noisy ECG signals. The CNN model was obtained by experiment, and it consists of six 2-dimensional convolutional layers, each having 36 filters with a kernel size of 19x1 per filter. Each convolutional layer is followed by a batch normalization layer with 36 channels, a rectified linear unit (ReLU) layer and an average pooling layer with a stride of 4 and pooling size of 2. Before the final regression output layer, the signal goes through a fully connected layer for regression.
The second DL model is a Long Short-Term Memory (LSTM) model, consisting of two LSTM layers with 140 hidden nodes per layer. Before the final regression output layer, the signal goes through a fully connected layer.
Figure 2.4(a,b) depicts respectively the structure of the CNN and LSTM models.

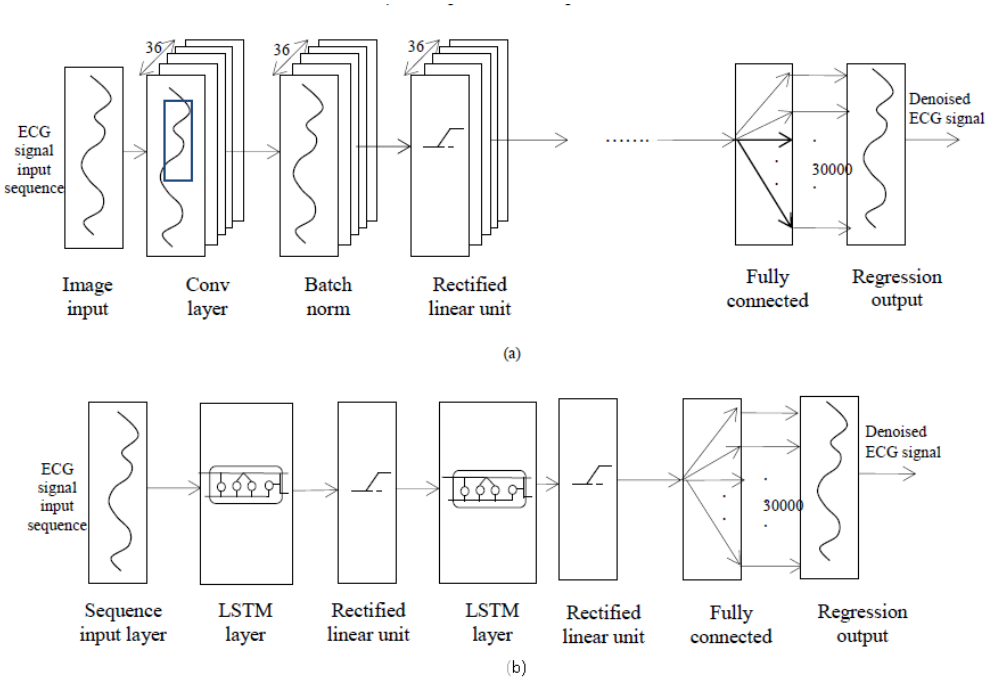

Figure 2.4: The CNN (a) and LSTM (b) models proposed by C. T. C. Arsene et al. (2019).

To train and test the models three datasets were used. Two comprise synthetic data while a third dataset is a real dataset. In each dataset, different levels of noise were considered, which included baseline wander, muscle artifact, and electrode motion artifact. Each dataset was divided into a training (3/4) and a testing dataset (1/4).
The results demonstrate that while both DL models could deal with heavy and drifting noise, the CNN model was markedly superior to the LSTM model in terms of the Root Mean Squared (RMS) error.

2.3.5 F. P. Romero et al. (2021)

This paper [35] proposes a novel algorithm for BLW noise filtering using deep learning techniques. The basic hypothesis of this work is that by learning "smart filters" it is possible to discriminate between the desired ECG signal and the undesired noise. Multiple deep filters can learn how to properly filter small sections of the input ECG signal while conserving its morphology using a similarity loss function.
The proposed model is a fully convolutional architecture based on multipath modules that receive an ECG signal contaminated with BLW as input and give the clean ECG signal as output.
The approach of this paper uses multipath modules, which place different convolutional layers at the same level and let the backpropagation algorithm choose not only the weights but also the best path for the signal to pass through. Figure 2.5 shows the proposed Multi-Kernel Linear And Non-Linear (MKLANL) filter module.

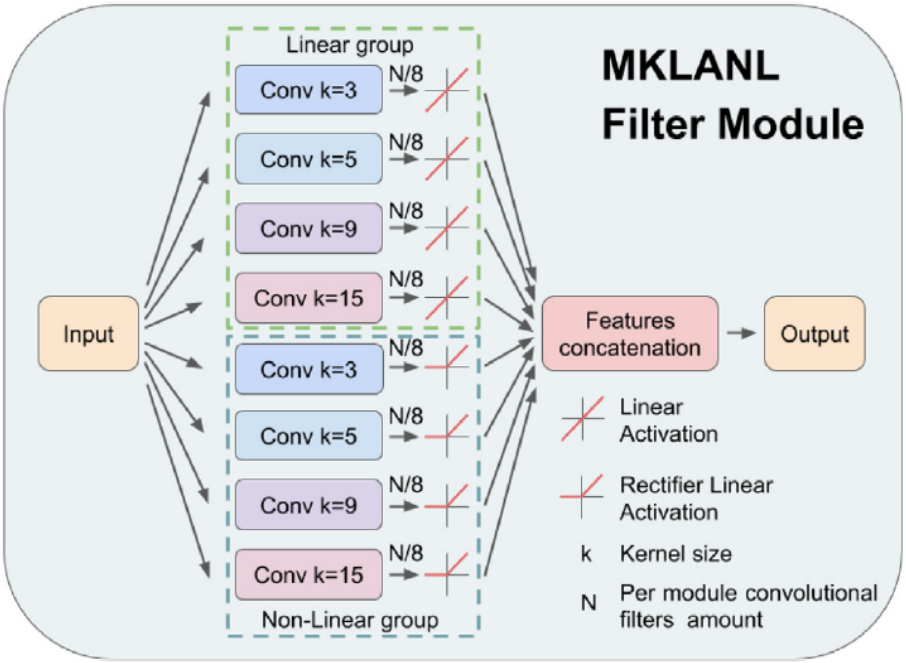

Figure 2.5: Multi-Kernel Linear And Non-Linear (MKLANL) filter module proposed by F. P. Romero et al. (2021).

The proposed MKLANL filter module is composed of two internal groups: the linear group and the non-linear group. Each internal group contains four types of convolutional layers with kernels equal to (3, 5, 9, and 15) followed by a linear activation or a rectified linear unit (ReLU), depending on the group. The rationale behind having convolution with linear and non-linear activations is the same as having different convolutional kernels: let the model choose during training which path is better, and how much each of them will contribute to the output. The

proposed deep learning architecture is composed of six MKLANL filter modules arranged sequentially (Figure 2.6).

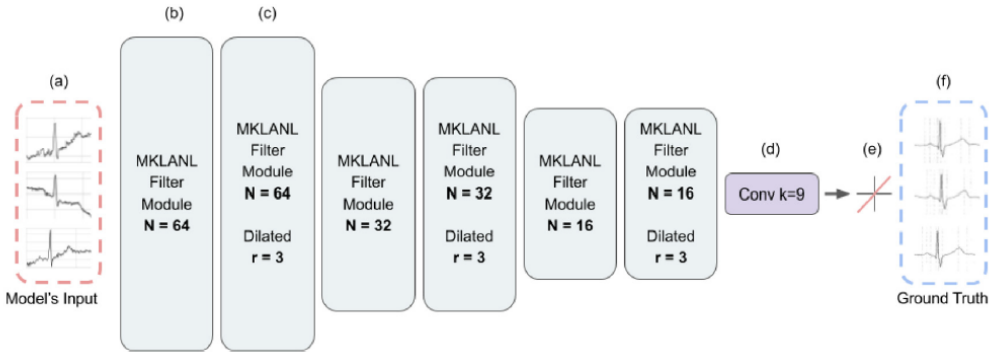

Figure 2.6: Deep learning network architecture proposed by F. P. Romero et al. (2021).

The proposed method was compared in terms of performance with other state-of-the-art approaches. FIR and IIR filters were used as a representation of non-learnable approaches together with two deep learning for ECG denoising. The first one is based on deep recurrent neural networks (DRNN) and the second one is based on fully convolutional denoising autoencoder (FCN-DAE).
Comparing the performances of all the methods, it can be observed that the proposed one obtained the best results on all the metrics used.

2.3.6 P. Singh et al. (2021)

This paper presents [36] an ECG denoising approach based on the generative adversarial network (GAN). GAN method trains two different models/networks simultaneously: a generator (G) model that learns to capture the training data distribution, and a discriminator (D) model that computes the probability that a sample is not from the G model but rather comes from the training data distribution.
In the proposed technique, the G network is made fully convolutional, and it does not contain any dense layer. The architecture of G is similar to an auto-encoder (Figure 2.7) and this network aims to learn the clean data distribution and produce the denoised output.

The overall training process of the ECG-GAN denoising model is shown in Figure 2.8.

The proposed framework is tested for multiple noisy conditions including white Gaussian noise, baseline wander, muscle artifact and electrode motion artifact. The qualitative and quantitative evaluations used in this work show a significant improvement over existing state-of-the-art ECG denoising methods, and in particular, the noted advantage of such a framework is the generalization to several noise conditions using a single generative model.

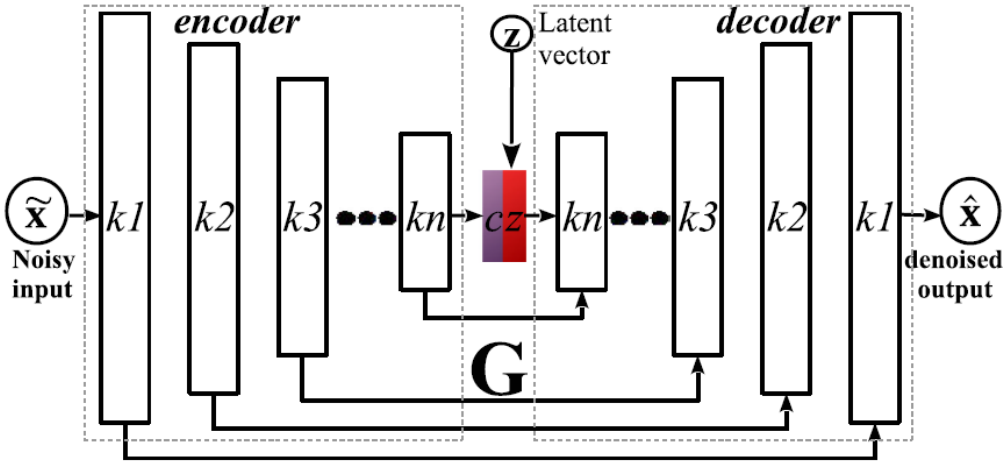

Figure 2.7: Generator network (G) proposed P. Singh et al. (2021).

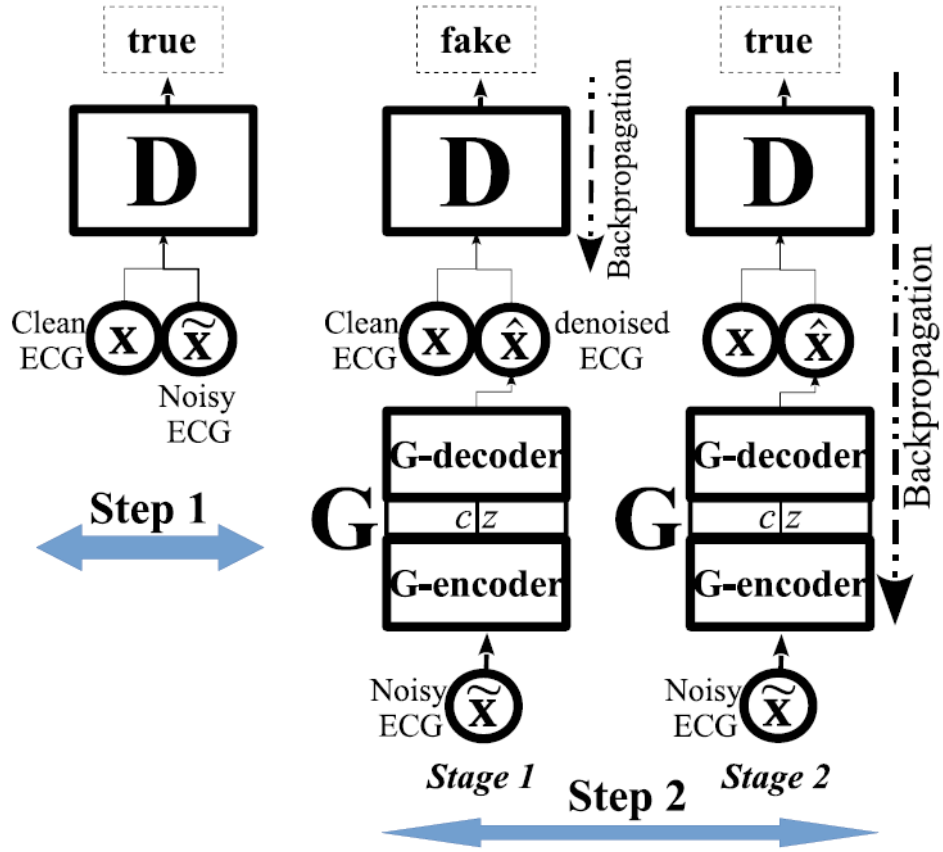

Figure 2.8: ECG-GAN training process proposed P. Singh et al. (2021).

2.3.7 B. Xu et al. (2021)

This paper [37] proposes a new method of ECG denoising based on the combination of the Generative Adversarial Network (GAN) and Residual Network. This work can be seen as a step further than the one done in [36]. The method adopted is based on the GAN structure, and it restructures the generator and discriminator. In the generator network, residual blocks and Skip-Connecting are used to deepen the network structure and better capture the in-depth information in the ECG signal. In the discriminator network, the ResNet framework is used.
Figure 2.9 shows the overall structure of the method used, while Figure 2.10 and Figure 2.11 are shown respectively the structure of the generator and discriminator networks.

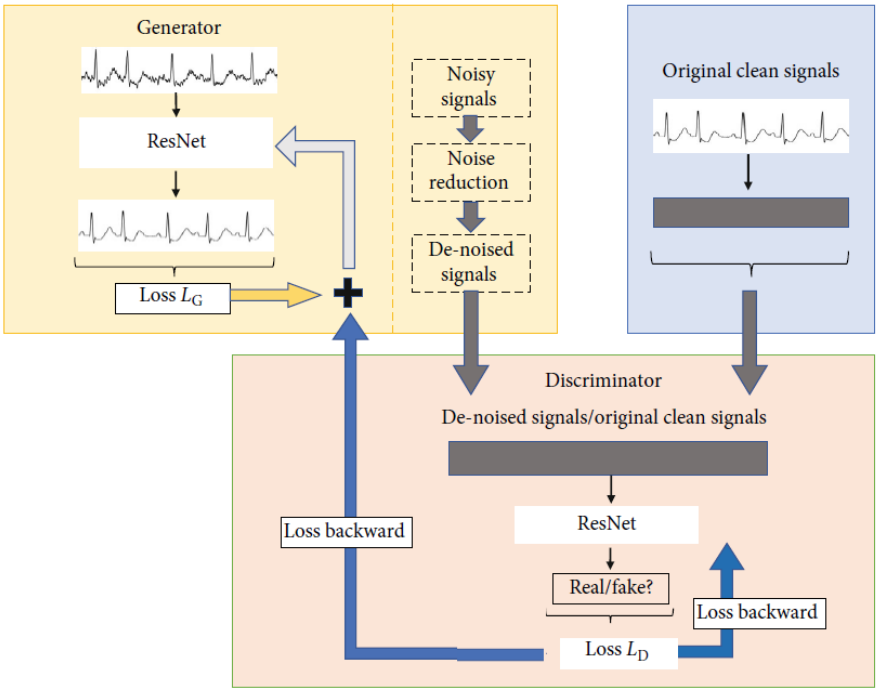

Figure 2.9: The overall structure of the network proposed by B. Xu et al. (2021).

In this work, to optimize the noise reduction process and solve the lack of local relevance considering the global ECG problem, the differential function and overall function of the maximum local difference are added to the loss function.
The experimental results prove that the method used in this article has better performance than the S-Transform (S-T) algorithm, Wavelet Transform (WT) algorithm, Stacked Denoising Autoencoder (S-DAE) algorithm, and Improved Denoising Autoencoder (I-DAE) algorithm. It is demonstrated that this method can effectively retain the important information conveyed by the original signal performing an optimal denoising.

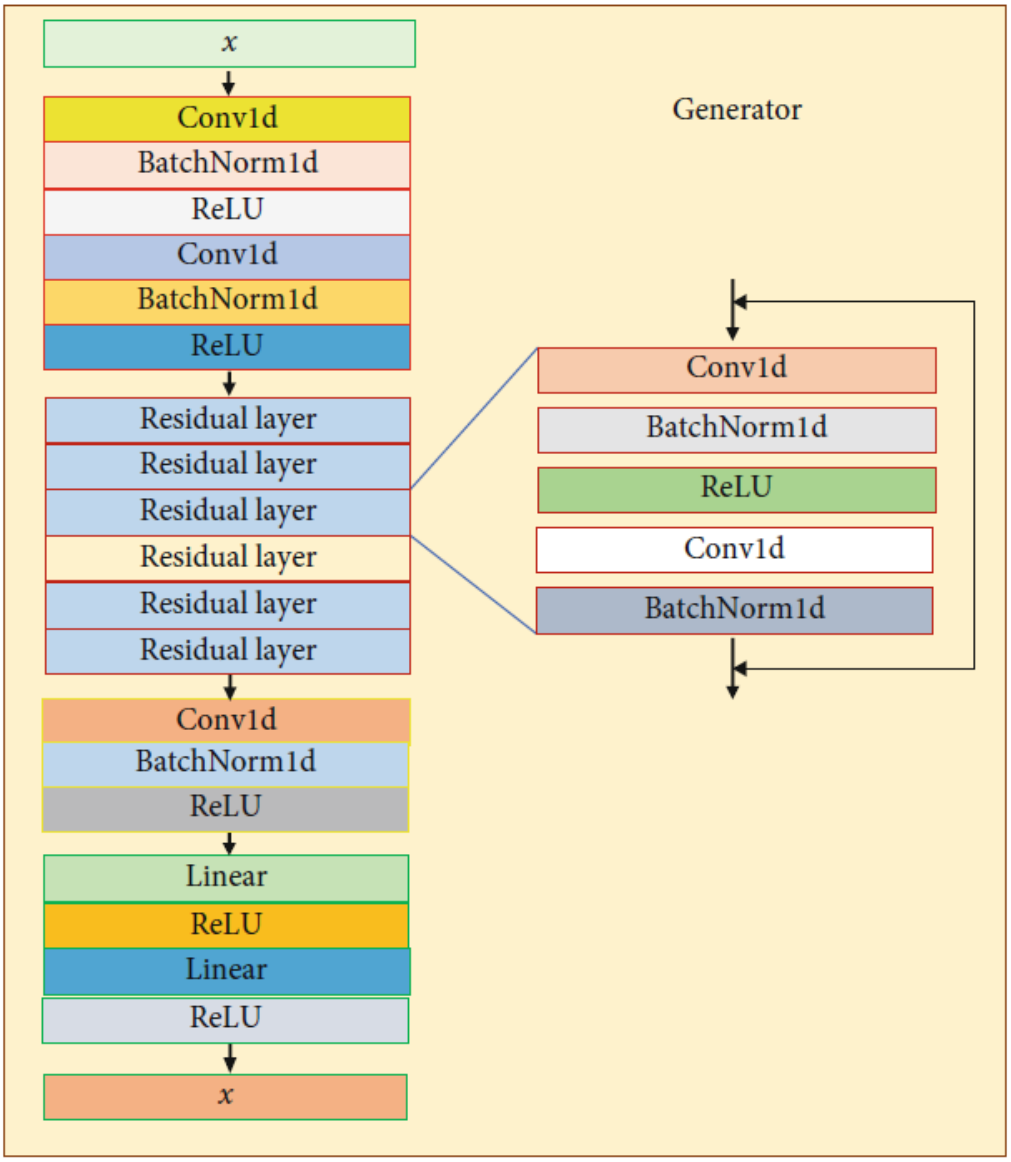

Figure 2.10: The structure of the generator network proposed by B. Xu et al. (2021).

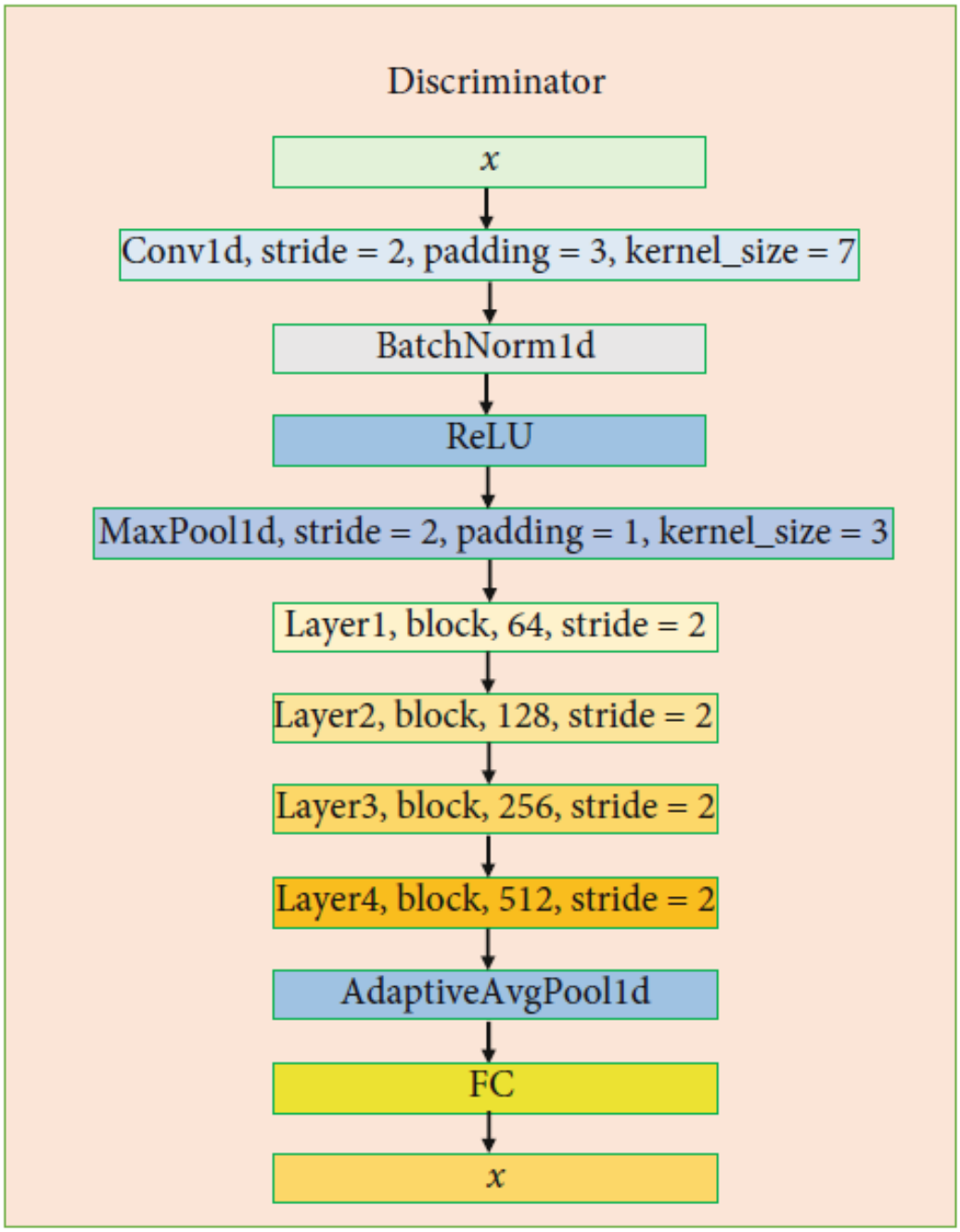

Figure 2.11: The structure of the discriminator network proposed by B. Xu et al. (2021).

2.3.8 L. Qiu et al. (2021)

The method proposed in this paper [38] is divided into two stages. In Figure 2.12 is depicted the general flowchart of the proposed ECG denoising method.

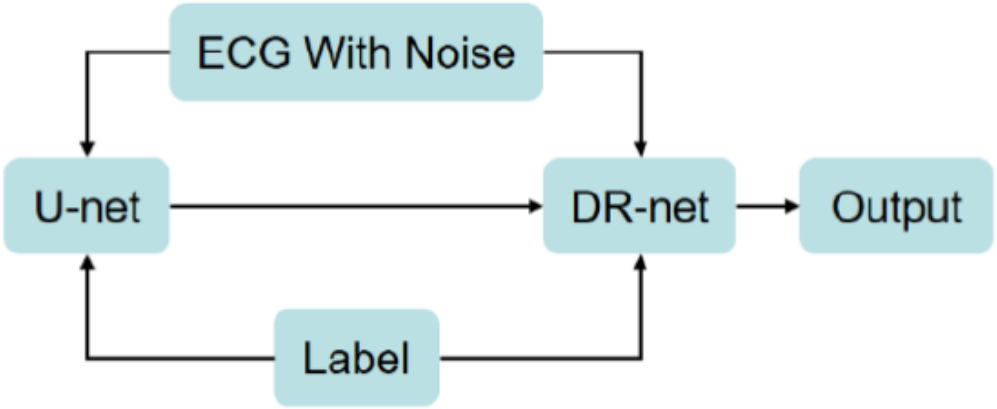

Figure 2.12: General flowchart of the ECG denoising method proposed by L. Qiu et al. (2021).

In the first stage, a Ude-net model is designed for ECG signal denoising to eliminate noise. After the first stage, even if the noise in the ECG signal is eliminated, the waveform of the ECG signal is inevitably distorted.
Figure 2.13 shows the structure of the proposed Ude-net model.

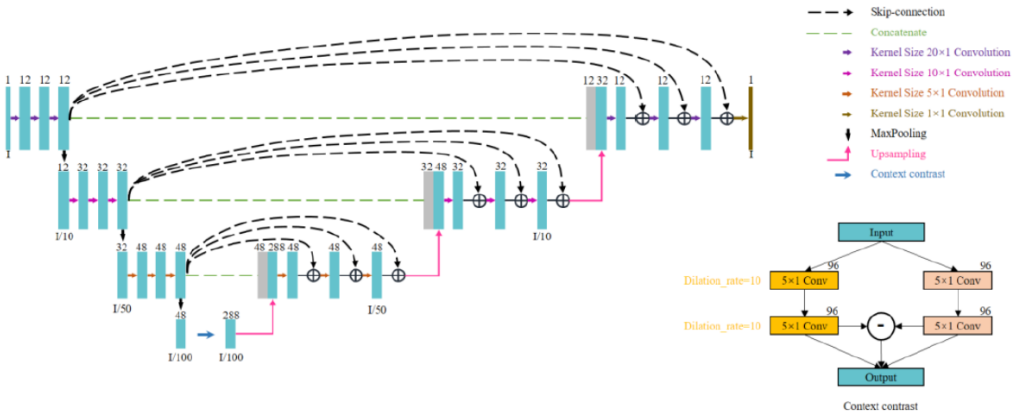

Figure 2.13: Improved one-dimensional U-net model structure proposed by L. Qiu et al. (2021).

The DR-net model in the second stage is used to reconstruct the ECG signal and to correct the distortion caused by noise removal in the first stage. This model has two inputs: one input is the output of the U-net after the first stage, and the other is the original noisy signal. The output instead is the corresponding ground-truth ECG. Thus, this second model does not learn how to remove noise but rather learns

to recover the effective part based on the first stage of U-net denoising.
Figure 2.14 shows the structure of the proposed DR-net model.

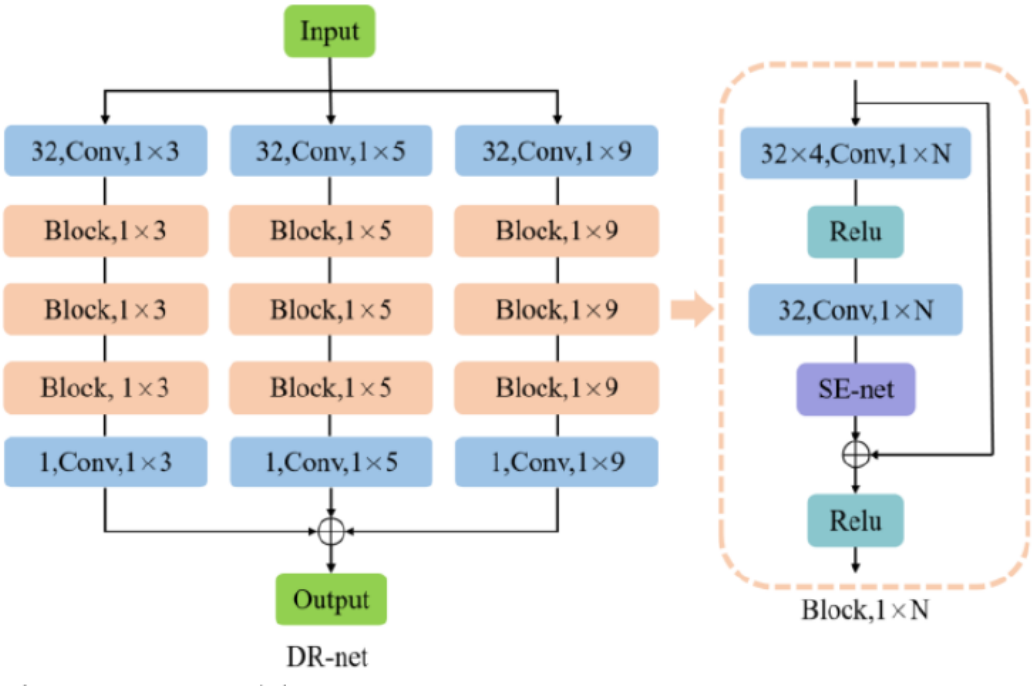

Figure 2.14: DR-net model structure proposed by L. Qiu et al. (2021).

This paper demonstrates how the proposed two-stage method can achieve both the elimination of noise and the preservation of effective details to a large extent of the signals.

2.3.9 Z. He al. (2021)

This paper [39] proposes an effective and simple model of encoder-decoder structure with a skip connection embedded for denoising ECG signals (APR-CNN). Specifically, Adaptive Parametric ReLU (APReLU) and Dual Attention Module (DAM) are introduced in the model. The Rectified Linear Unit (ReLU) is replaced with the APReLU for better negative information retainment. The DAM is an attention-based module consisting of a channel attention module and a spatial attention module, through which the inter-spatial and inter-channel relationship of the input data are exploited. DAM is introduced at the back of each layer. The overall structure of this model is presented in Figure 2.15.

The results of this work show that the APR-CNN can handle ECG signals with a different signal-to-noise ratio (SNR).

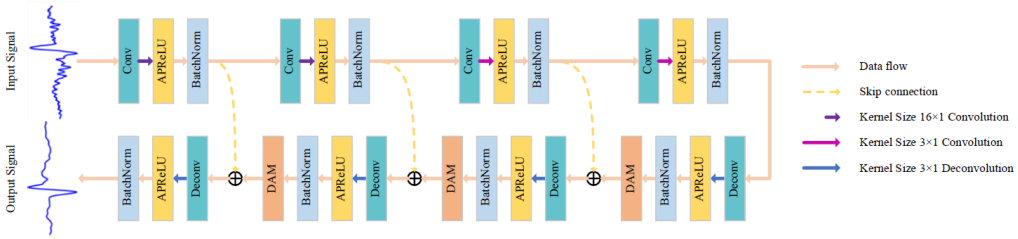

Figure 2.15: Structure of the model proposed by Z. He al. (2021).

2.3.10 A. Mohammadisrab et al. (2022)

This paper [40] analyzes the performance of a deep adaptive denoising auto-encoder network (DeepADAENet) for ECG signal noise cancellation in the time-frequency domain for practical use cases.
To achieve a higher resolution in distinguishing the noise from valuable data, the fractional Stockwell transform (FrST) is exploited to convert the ECG to the time-frequency image. The magnitude of the time-frequency version of the ECG is noise-canceled using DeepADAENet. Then, inverse FrST is utilized to return the denoised time-frequency ECG into the time domain. The architecture of DeepADAENet used is shown in Figure 2.16.

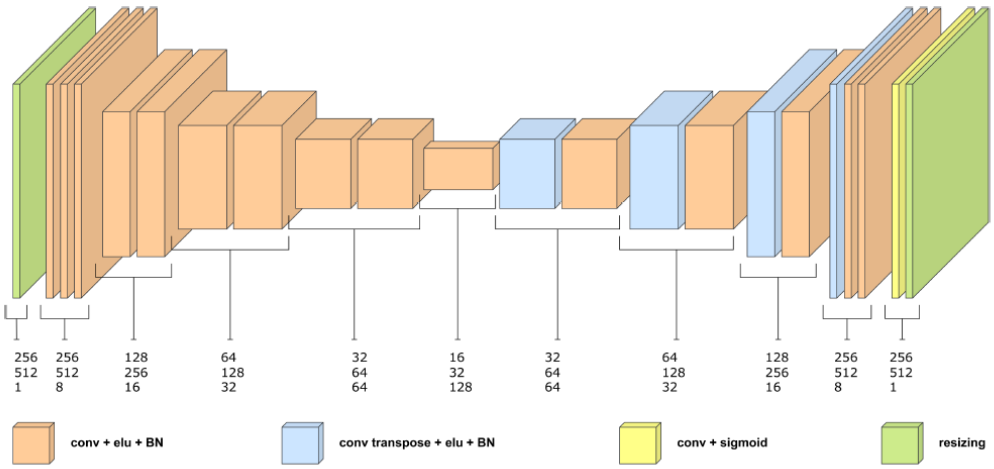

Figure 2.16: The architecture of DeepADAENet proposed by A. Mohammadisrab et al. (2022).

Muscle artifacts (MA), baseline wander (BW), and electrode motion (EM) from the MIT-BIH Noise Stress Test Database (NSTDB) are utilized to make noisy the clean ECG signals dataset according to Eq. 2.1:

$$Noise = \alpha BW + \beta EM + \lambda MA + \omega Rand \tag{2.1}$$

where α, β, λ, and ω represent the coefficients of each noise element and specify the influence of the corresponding noise element in the simulated noisy signal. Moreover, *Rand* stands for additive random noise in addition to BW, MA, and EM.
Experimental results reveal that DeepADAENet is remarkably adaptive to noises resembling ECG waves (P, QRS, and T) and different heart physiologies. A comparative evaluation also proved that DeepADAENet had a greater output signal-to-noise ratio and achieved lower root mean square errors and percentage root mean square differences compared to a fully convolutional network-based denoising auto-encoder network (FCN-based DAE).

2.3.11 E. Brophy et al. (2022)

This paper [41] proposes a custom loss function capable of denoising electrode motion artifact in ECG data to a higher standard than other, more common loss functions. This work implements a personalized custom loss function with a convolutional neural network (CNN) to return high-quality ECG. The CNN is composed of four-layer 1-D with batch normalization and ReLU (Rectified Linear Units) followed by a fully connected layer. The model architecture is shown in Figure 2.17.

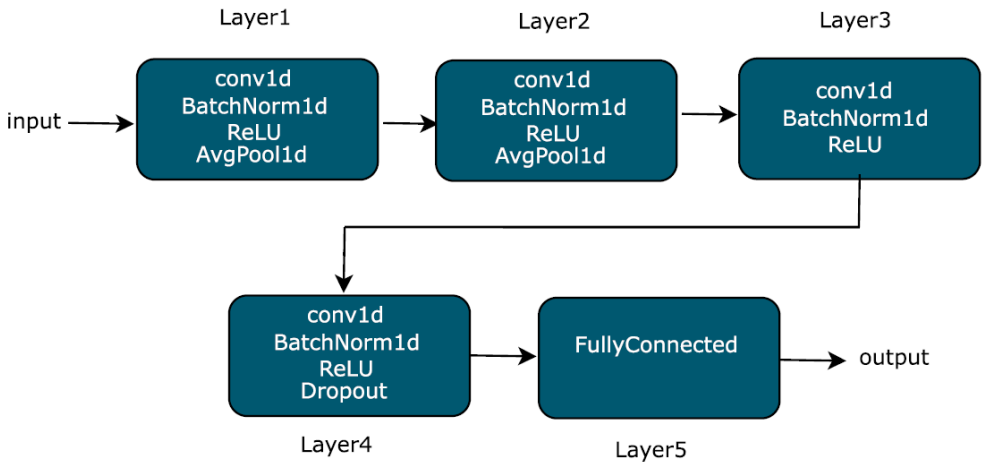

Figure 2.17: The network architecture proposed by E. Brophy et al. (2022).

The custom loss function used is composed of two parts: the first is the global mean squared error (MSE), and the second is the MSE pertaining only to regions where the QRS wave features exist. This can be represented by Eq. 2.2:

$$\mathcal{L} = MSE\left(X_n, X\right) + \alpha \cdot \sum_{i=1}^{j} MSE(R_{X_n i}X, R_{Xi}) \tag{2.2}$$

where j is the total number of QRS complexes in the 3-second signal segment, R_{Xi} is the i^{th} QRS complex in signal X and $R_{X_n i}$ is the i^{th} QRS complex in signal X_n. The value α is a hyperparameter that is used to determine the level of importance placed on the QRS complexes by the loss function relative to the ECG signal as a whole.
The proposed model and custom loss function compute a weighted combination of global and local mean square errors and improve the denoising performance of the ECG in terms of the SNR and heart rate. This demonstrates the capability of the algorithm to balance between denoising the signal and preserving the peaks effectively.

2.3.12 R. Badiger et al. (2023)

This paper [42] presents a deep learning-based scheme for ECG signal filtering, which is based on the deep autoencoder module. According to this scheme, the data is processed through the encoder and decoder layer to reconstruct by eliminating noises. Figure 2.18 shows the proposed DAE-ASCNet architecture.

The proposed deep learning architecture is based on the same concepts used in the work by He et al. [39].
It uses a modified ReLU function to improve the learning of attributes since standard ReLU cannot adapt to huge variations. Further, a skip connection is incorporated to retain the key feature of the encoder layer while mapping these features to the decoder layer. Finally, an attention model is also included, which performs channel and spatial attention, which generates the robust map by using channel and average pooling operations, resulting in improving the learning performance.
The proposed approach is tested on a publicly available MIT-BIH dataset where different types of noise, such as electrode motion, baseline wander and motion artifacts, are added to the original signal at varied SNR levels.
The comparative analysis with existing techniques shows a significant improvement in the performance of the proposed approach in terms of RMSE and SNR.

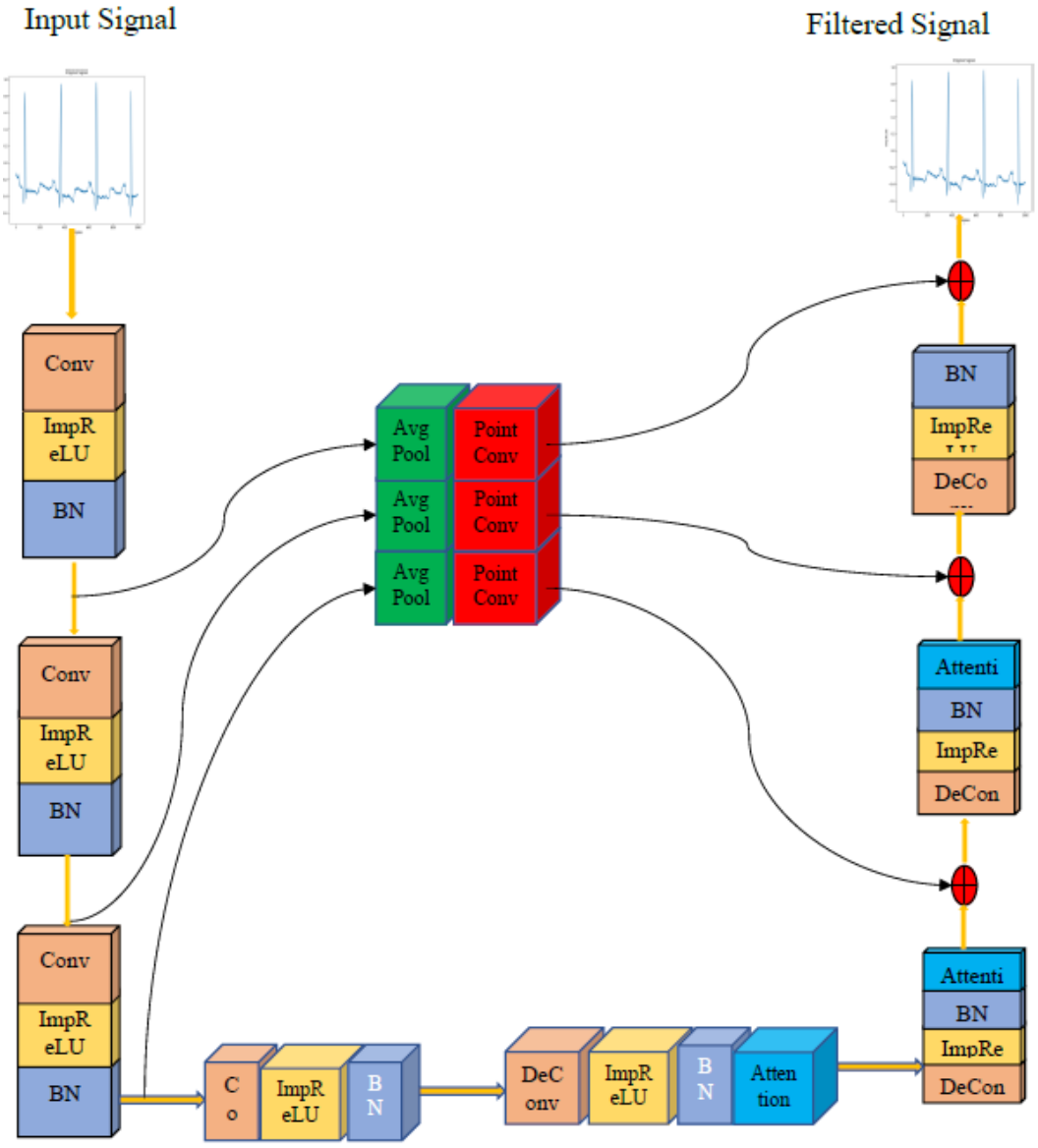

Figure 2.18: DAE-ASCNet for ECG filtering proposed by R. Badiger et al. (2023).

2.3.13 H. Wang et al. (2023)

In this paper [43], it is proposed an ECG denoising method referred to as LSTM-DCGAN which is based on an improved generative adversarial network (GAN). This is a modified version of [37], where LSTM structures are included in the architecture of the generator and discriminator instead of using only CNN layers.
The overall network structure is shown in Figure 2.19 and it is composed of multiple layers of convolutional networks. Furthermore, the convolutional features can be connected to their time series order dependence by adding LSTM layers after each convolutional layer. The LSTM is used to preserve the global time domain information of ECG.

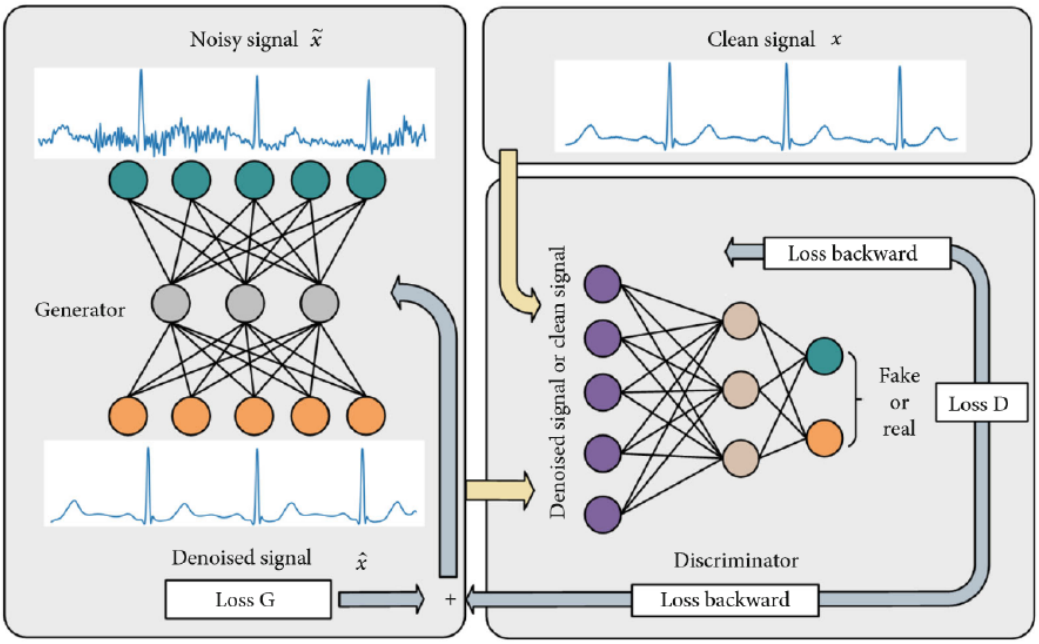

Figure 2.19: Denoising model based on GAN proposed by H. Wang et al. (2023).

Figure 2.20 and Figure 2.21 show respectively the architecture of the generator and discriminator.

Experimental results show that this method can remove the single noise and the mixed noise while retaining the complete ECG information.
Compared with the state-of-the-art methods, this method obtains higher SNR improvement and lower RMSE and PRD scores.

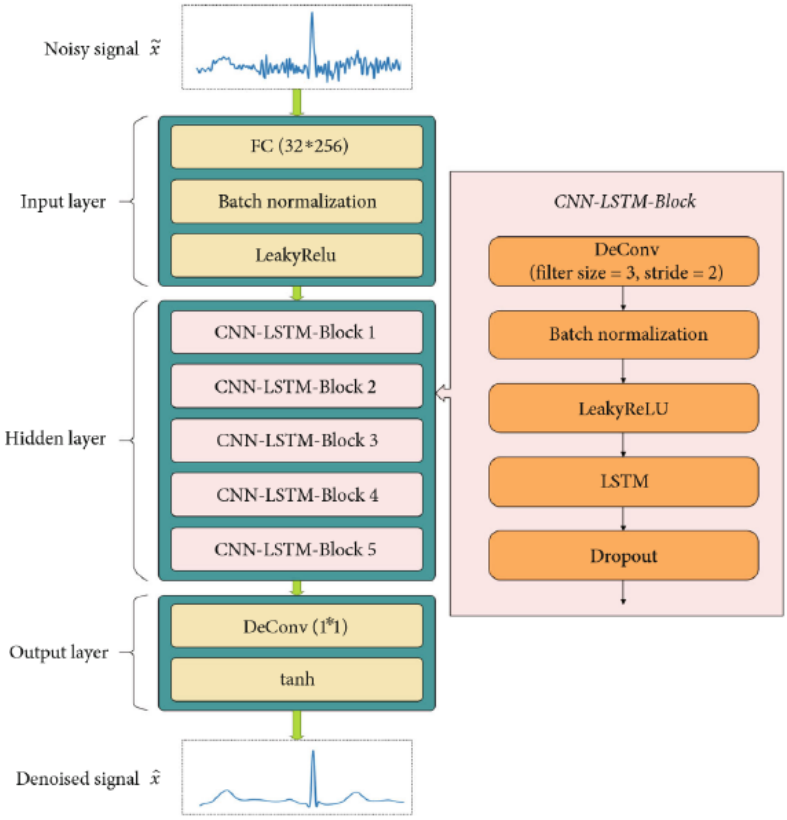

Figure 2.20: Structure of the generator of the model proposed by H. Wang et al. (2023).

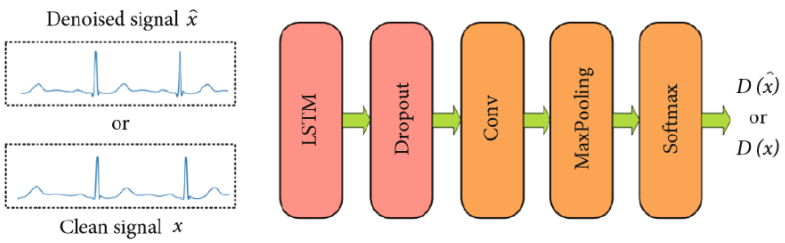

Figure 2.21: Structure of the discriminator of the model proposed by H. Wang et al. (2023).

2.3.14 Y. Jin et al. (2024)

This paper [44] proposes a novel signal-denoising method based on a deep wavelet convolutional neural network.
The architecture of the network (DW-CNN), shown in Figure 2.22, is inspired by the structure of the denoising self-encoder with the convolution layers that replace the simple full-connected layers. Moreover, based on automatic feature extraction in the convolution layer, the discrete wavelet transform (DWT) is used to convert the signal into high-frequency and low-frequency components for replacing pooling layers to compress the input data and fully preserve the effective information.

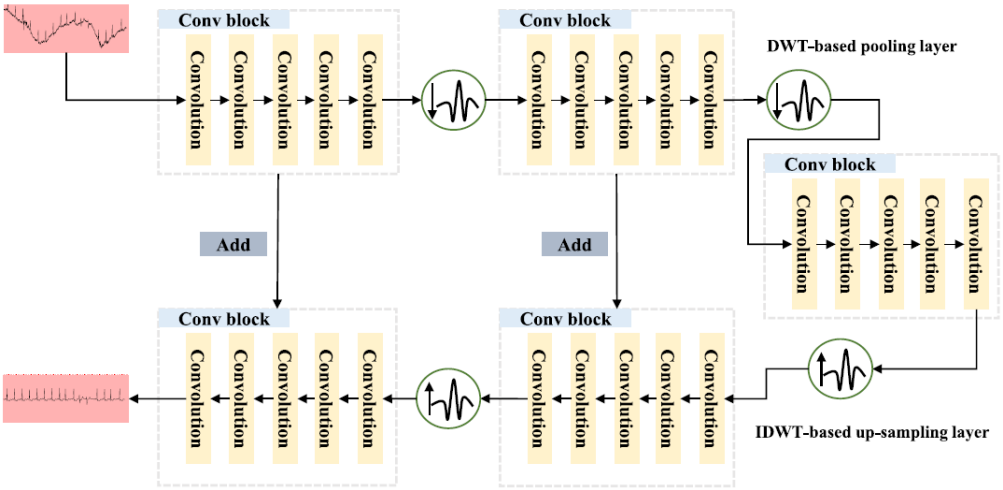

Figure 2.22: The network architecture proposed by Y. Jin et al. (2024).

Compared with the existing methods, DW-CNN has better denoising performance under different noise intensities and noise types.

2.4 Comparison tables and discussion

Table 2.1,Table 2.2 and Table 2.3 provide a comparison of the main features of the studies analyzed in this literature review.
As it is possible to notice, each paper bases the evaluation of the performances of the proposed deep learning structures on the use of different metrics. In addition, the input ECG signals to the various networks have different characteristics, including sampling frequency, length and amplitude of the signal. Moreover, the noise removed by the various signal-denoising techniques proposed is also different. The main ECG artifacts considered for the denoising techniques are white Gaussian noise (WGN), baseline wander (BW), muscle artifact (MA), and electrode motion artifact (EM).
Table 2.4 shows a comparison of the results obtained from the proposed studies in terms of the relative metrics used. For the above-reported reasons and differences between the works, the metrics reported are only the ones most widely used. In particular, are reported the values of the signal-to-noise ratio of the output signal (SNR output), the signal-to-noise ratio improvement between the input and the output signal (SNRimp), the percent root mean square difference (PRD), the mean square error (MSE) and the root mean square error (RMSE).
In the case in which in a paper the evaluation metric results are reported divided by the type of noise removed in the ECG signal and the input ECG signal amplitude, an average value has been calculated. For the paper [40] was not possible to report the numerical results as the performances of the analyzed deep learning method are shown only in the form of graphs or denoised signal plots.

Table 2.1: Comparison among different studies [1/3].

Ref.	Title	Author	Year	Dataset	Noise Removed	Method Used	Loss Function	Evaluation Metrics
[9]	An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks	Suranai Poungponsri; Xiao-HuaYu	2013	MIT–BIH (Massachusetts Institute of Technology – Beth Israel Hospital) Database.	Power-line Interference; Baseline Wander; Electrode Motion Artifact; Muscle Contraction Artifact; White Gaussian Noise	Discrete Wavelet Transform + ANN	Weighted Sum of Squared Errors	Signal to Noise Ratio Improvement (SNRimp)
[32]	An optimum ECG Denoising with Wavelet Neural Network	S. O. Rajankar; S. N. Talbar	2015	MIT-BIH arrhythmia datasets	Gaussian noise of various noise levels	Wavelet Neural Network Architecture	Not Mentioned	Signal to Noise Ratio (SNR); Mean Square Error (MSE)
[33]	Elimination of Power Line Interference from ECG Signals Using Recurrent Neural Networks	Yue Qiu; Feng Xiao; Haibin Shen	2017	3000 synthetic samples of ECG signal contaminated by Power Line Noise randomly generated.	Power Line Interference	Deep LSTM RNN model	Mean Squared Error	Percentage Root Mean Square Difference (PRD)
[34]	Deep Learning Models for Denoising ECG Signals	C. T. C. Arsene; R. Hankins; H. Yin	2019	Three datasets were used. Two comprise synthetic data while a third dataset is a real dataset (MITBIH Arrhythmia Database).	Baseline Wander; Muscle Artifact; Electrode Motion Artifact	Two DL Models: CNN and Long Short-Term Memory (LSTM)	Not Mentioned	Root Mean Square Error (RMSE)
[35]	DeepFilter: An ECG baseline wander removal filter using deep learning techniques	Francisco P. Romero; David C. Piñol; Carlos R. Vázquez-Seisdedos	2021	QT Database from Physionet	Baseline Wander	Deep Filter: CNN composed by MKLANL filter modules	Sum of Squared Distance (SSD) + Maximum Absolute Distance (MAD)	Sum of the Square of the Distances (SSD); Percentage Root Mean Square Difference (PRD); Maximum Absolute Distance (MAD); Cosine Similarity

Table 2.2: Comparison among different studies [2/3].

Ref.	Title	Author	Year	Dataset	Noise Removed	Method Used	Loss Function	Evaluation Metrics
[36]	A New ECG Denoising Framework Using Generative Adversarial Network	Pratik Singh; Gayadhar Pradhan	2021	MIT-BIH Arrhythmia datasets	White Gaussian Noise; Baseline Wander; Muscle Artifact; Electrode Motion Artifact	Generative Adversarial Network (GAN)	The modified value function of G network is used to minimize the distance between denoised and clean samples	Signal to Noise Ratio (SNR); Signal to Noise Ratio Improvement (SNRimp); Percentage Root Mean Square Difference (PRD); Mean Square Error (MSE); Root Mean Square Error (RMSE)
[37]	An ECG Denoising Method Based on the Generative Adversarial Residual Network	Bingxin Xu; Ruixia Liu; Minglei Shu; Xiaoyi Shang; Yinglong Wang	2021	MIT-BIH arrhythmia database by the Massachusetts Institute of Technology and Beth Israel Hospital and the MIT-BIH noise stress test database of PhysioBank	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	Combination of the Generative Adversarial Network and Residual Network	Differential function and overall function of the maximum local difference	Signal to Noise Ratio (SNR); Root Mean Square Error (RMSE)
[38]	A two-stage ECG signal denoising method based on deep convolutional network	Qiu Lishen; Cai Wenqiang; Zhang Miao; Zhu Wenliang; Wang Lirong	2021	The ECG data used are from CPSC2018, and the noise signal is from MIT-BIH Noise Stress Test Database	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	Ude-net model + DR-net model	Not Mentioned	Signal to Noise Ratio Improvement (SNRimp); Root Mean Square Error Decrease (RMSEde)
[39]	Dual Attention Convolutional Neural Network Based on Adaptive Parametric ReLU for Denoising ECG Signals with Strong Noise	Zixiao He; Xinwen Liu Hao He; Huan Wang	2021	Real ECG signals from MIT-BIH Arrhythmia Database and Real ECG noise from MIT-BIH Stress Test Database	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	Encoder-Decoder structure with Adaptive Parametric ReLU (APReLU) and Dual Attention Module (DAM)	Not Mentioned	Signal to Noise Ratio (SNR); Mean Square Error (MSE)
[40]	Deep Adaptive Denoising Auto-Encoder Networks for ECG Noise Cancellation via Time-Frequency Domain	A. Moham-madisrab et al	2022	MIT-BIH Apnea-ECG database (APNEA-ECG) and MIT-BIH Noise Stress Test Database (NSTDB)	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	Fractional Stockwell transform (FrST) + Deep ADAENet + Inverse FrST	Not Mentioned	Signal to Noise Ratio (SNR); Root Mean Square Error (RMSE); Percent Root Mean Square Difference (PRD)

Table 2.3: Comparison among different studies [3/3].

Ref.	Title	Author	Year	Dataset	Noise Removed	Method Used	Loss Function	Evaluation Metrics
[41]	Improved Electrode Motion Artefact Denoising in ECG Using Convolutional Neural Networks and a Custom Loss Function	E. Brophy; B. Hennelly; M. De Vos; G. Boylan; T. Ward	2022	MIT-BIH Arrhythmia Database and MIT-BIH Noise Stress Test Database	Electrode Motion Artifact	CNN with four 1-D layers	Two MSE components: Global + Local MSE	Signal to Noise Ratio Improvement (SNRimp); Heart Rate Error Prediction; Interbeat Interval (IBI) + Heart Rate Variability (HRV) of the denoised vs noisy ECG signals
[42]	ASCNet-ECG: Deep Autoencoder based Attention aware Skip Connection network for ECG filtering	Raghavendra Badiger; M. Prabhakar	2023	MIT-BIH dataset	White Gaussian Noise; Baseline Wander; Muscle Artifact; Electrode Motion Artifact	Deep autoencoder module with a modified ReLU function, skip connection and attention model	Not Mentioned	Signal to Noise Ratio (SNR); Root Mean Square Error (RMSE)
[43]	Deep Convolutional Generative Adversarial Network with LSTM for ECG Denoising	H. Wang; Y. Ma; A. Zhang; D. Lin; Y. Qi; J. Li	2023	MIT-BIH Arrhythmia Database with different levels of noise from the MIT-BIH Noise Stress Test Database	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	Deep Convolutional Generative Adversarial Network (DCGAN) + LSTM	Not Mentioned	Signal to Noise Ratio Improvement (SNRimp); Percent Root Mean Square Difference (PRD); Root Mean Square Error (RMSE)
[44]	A novel deep wavelet convolutional neural network for actual ECG signal denoising	Yanrui Jin; Chengjin Qin; Jinlei Liu; Yunqing Liu; Zhiyuan Li; Chengliang Liu	2024	MIT-BIH Arrhythmia Database and MIT-BIH Noise Stress Test Database	Electrode Motion Artifact; Baseline Wander; Muscle Artifact	DW-CNN: Deep Wavelet Convolutional Neural Network (with Haar wavelet).	Mean Squared Error	Signal to Noise Ratio (SNR); Root Mean Square Error (RMSE)

Table 2.4: Comparison of evaluation metric results among studies.

Ref.	SNR OUTPUT	SNRimp	PRD	MSE	RMSE
[9]		12,11			
[32]	27,72			4,010	
[33]			2,42		
[34]					CNN: 0,029 LSTM: 0,232
[35]			50.45		
[36]		14,58	6,48	0,002	
[37]	40,88				0,010
[38]		9,96			0,004
[39]	7,21				0,045
[40]					
[41]	29,74				
[42]	46,22				0,024
[43]		19,25	12,91		0,035
[44]		5,99			0,194

Chapter 3

F-waves extraction by deep learning

3.1 Introduction

Early detection of AF is important to ensure timely and correct management of the condition and avoid the recurrence of the arrhythmia as much as possible.
According to the European guidelines for the diagnosis and management of patients with atrial fibrillation [23], the diagnosis process requires rhythm documentation acquired by a single-lead or 12-lead ECG tracing.
The most relevant information searched by clinicians in the evaluation of atrial arrhythmias is contained in the irregular fibrillatory waves: F-waves, which replace the standard atrial depolarization waves (P-waves). F-waves can have different shapes, amplitude, and duration in every patient and, additionally, their amplitude is low and generally superimposed by the QRS complex. For these reasons, for clinicians, detecting F-waves by visual inspection is a challenging task.
Nowadays different signal processing principles have been employed to perform F-waves extraction from ECG signals as average beat subtraction (ABS) and variants, principal component analysis (PCA), interpolation and singular spectral analysis, independent component analysis (ICA), adaptive filtering using an echo state network and diffusion geometry.
Moreover, also deep learning algorithms are widely used to analyze and extract features from biomedical signals such as ECG. Many recent works use machine learning algorithms to classify different types of atrial fibrillation, but no work has been found in the literature using deep learning methods to extract F-waves from ECG signals of patients with AF.

3.2 Materials and methods

3.2.1 Dataset

The data used in this research come from a reference database for validation of methods of extraction of atrial fibrillatory waves in the ECG, available online and described in [45]. This dataset contains 240 records of simulated AF 12-lead ECG signals, which are different combinations of real or synthetic F-waves and QRST complexes. The synthetic F-waves and QRST complexes are generated by an extended

version of the sawtooth model [46] and an extended version of the single-dipole model [47], respectively. The real F-waves without QRS-related residuals are carefully selected from a proprietary clinical AF database [48], and the real QRST complexes are selected from the PTB database [49].
The database is composed of eight signal sets together accounting for a wide range of characteristics known to represent major challenges in F-waves extraction, including high heart rates, QRST with high morphological variability (HMV), and the presence of ventricular premature beats (VPBs). Each set contains 30 5-minute signals with six different values of amplitude (5, 10, 20, 30, 40, 50 μV). Each record is sampled at a frequency of 1000 Hz. Moreover, this dataset is noise-free, except for the noise which is present in real signals.
In Table 3.1 are reported the composition of sets of signals available in this database.

Table 3.1: Composition of the dataset.

SET	Real F-waves	Synthetic F-waves	Real QRST complexes	Synthetic QRST complexes	Real QRST complexes with HMV	Real VPBs
S1	×		×			
S2		×	×			
S3	×			×		
S4		×		×		
S5	×				×	
S6		×			×	
S7	×				×	×
S8		×			×	×

3.2.2 Proposed model

This work proposes a two-stage deep learning method to extract F-waves from ECG signals of patients affected by atrial fibrillation.
The architecture of the models employed in both two stages is the same, while the two loss functions are different. The first model takes as input data the real ECG signals with F-waves and the corresponding real F-waves as ground truth, while the second model takes the output of the first model as input data and always the real F-waves as ground truth. The basic block scheme of the designed architecture is shown in Figure 3.1.

Both the proposed models are designed with 2 fully convolutional architectures based on multipath modules. This approach places different convolutional layers at the same level and lets the backpropagation algorithm choose not only the weights but also the best path for the signal to pass through.
In deep learning models, often the kernel size is one hyperparameter challenging to set. Therefore, typically a grid search or multiple empirical choices are performed to

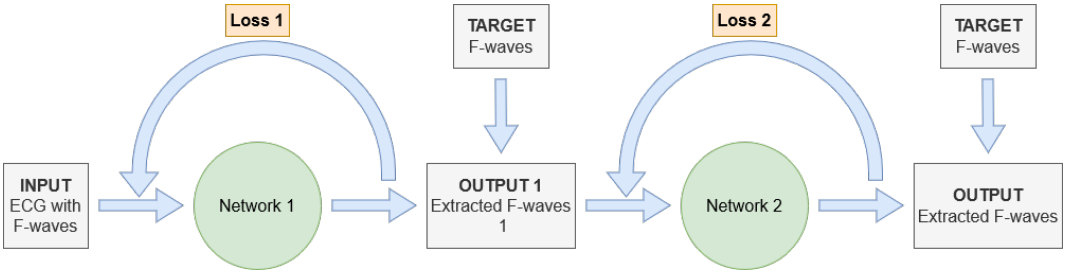

Figure 3.1: Block scheme of the proposed two-stage deep learning method.

find an optimal value, which can be time-consuming. Employing multipath modules, an optimal kernel will be learned at each level. The proposed Multi-Kernel Linear And Non-Linear (MKLANL) filter module is inspired by the Inception module, originally introduced by [50] and adopted also in the DeepFilter designed in [35]. The used MKLANL filter module (Figure 3.2) is composed of two internal groups: the linear group with linear activation functions and the non-linear group with rectified linear unit (ReLU) activation functions. Each internal group contains four types of convolutional layers with 1D kernels equal to (3, 5, 9, and 15). The number of convolutional filters for each type is $N/8$, where N is a hyperparameter to control the total amount of filters per multipath module.

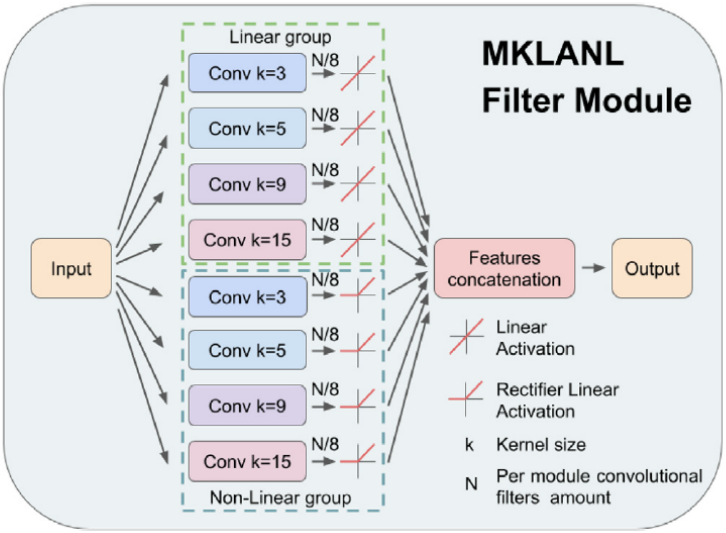

Figure 3.2: Structure of the MKLANL filter module.

Moreover, also dilated convolutions using non-consecutive kernels (Figure 3.3) were introduced in the filter module in order to increase the kernel receptive fields without increasing the computational load. For a one-dimensional signal $x[i]$, the output $y[i]$ of a dilated convolution operation with a dilation rate r and a filter mask $w[s]$ is defined by Eq. 3.1.

$$y[i] = \sum_{s=1}^{S} x[i + r \cdot s]w[s] \tag{3.1}$$

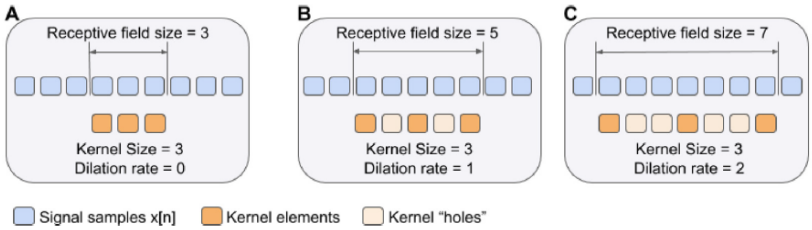

Figure 3.3: Dilated convolutions.

Each of the two designed networks is composed of six MKLANL filter modules arranged sequentially. Figure 3.4 illustrates one of the two proposed deep learning network architectures.

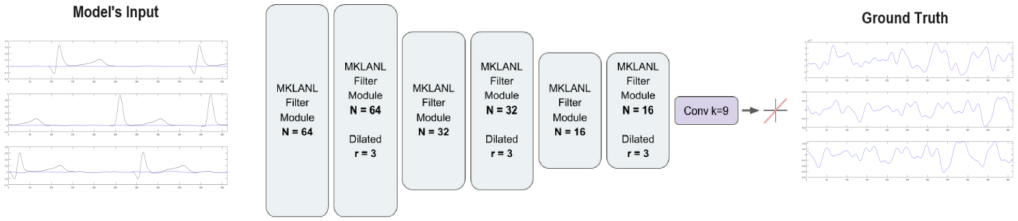

Figure 3.4: Architecture of each proposed network.

The first two layers have a total of 64 convolutional filters and are internally structured as shown in Figure 3.2. In the second filter module, convolutional operations were configured with a dilation rate $r = 3$, enabling dilated convolutions. This same dilation rate was also set in the fourth and sixth layers. The number of extracted features decreases along the network, starting with 64 in modules one and two, then 32 features in modules three and four. Lastly, the modules five and six contain 16 features each. The final step has one convolutional filter with $kernel = 9$, which conforms to the output signal. Since F-waves are bipolar signals, linear activation was used for this final step, thereby allowing the output to have either positive or negative values.

3.2.3 Training strategy

This thesis has been implemented partly in the MATLAB environment and partly running Python code in the Google Colab environment.

Preprocessing

From the above-reported dataset composed of eight signal sets, only the data of the first set, i.e. the one which contains signals with both real F-waves and QRST complexes, were used.

Data were provided with the ".mat" extension, thus the first steps of preprocessing of the data were implemented in the MATLAB environment. The first set contains 30 5-minute ECG and F-wave signals (12-lead acquisition) with six different values of amplitude (5, 10, 20, 30, 40, 50 μV) sampled at a frequency of 1000 Hz. To reduce the computational effort of the training of the next step, signals were down-sampled at a frequency of 512 Hz.

All the next preprocessing steps described have been performed similarly for ECG and F-wave signals.

For every 5 minutes of the signal acquired through the 12 leads, 1-second windows were extracted. To increase the randomization of the input data, a window shifting has been imposed based on the amplitude of the signal considered. For the signal with an amplitude of 5 μV, no shifting of the generated window has been considered, while for the signal of 10 μV a shift of 0.1 seconds of the window has been imposed with respect to the previous one, for the signal of 20 μV 0.2 seconds, for that of 30 μV of 0.3, 0.4 for that of 40 μV and 0.5 for that with amplitude equal to 50 μV. This procedure was necessary because, despite the ECG signals of the original dataset being all different, the F-wave component contained in them was repeated for some amplitude pairs of signals. Specifically, the signals with amplitude of 10 and 20 μV contained the same component of F-waves and similarly the pair of signals with amplitude of 30 and 50 μV.

For each 5-minute signal, windows containing the first and the last second were excluded, resulting in 298 windows of one second for each recording and therefore a total of 107.280 windows of ECG and F-waves signal respectively. Finally, the windows corresponding to the signals acquired through the leads aVR, aVF, aVL and III were also excluded as they did not contain any additional information compared to those acquired through the other 8 leads.

According to all the steps described, the final dataset obtained consists of 71.520s of ECG signal recordings and the respective 71.520s of F-waves.

The dataset was imported via Google Drive into the Colab environment, which provides a Jupyter notebook to execute Python codes. The lists of data corresponding to the input signals of the implemented two-stage deep learning method (ECG with F-waves) were shuffled to increase the randomization while maintaining the correlation between the ECG recording and the corresponding component of F-wave.

Finally, the dataset was divided into three sets: the training set which contained 80% of the data (57.120 seconds) that were used to train the implemented deep learning method, the validation set which contained 10% of the data (7.200 seconds) to determine when to stop training to avoid the overfitting problem, and the test set containing the last 10% of the data (7.200 seconds); this data were kept away from the system during training and were used only to evaluate it.
The described data split is shown in Figure 3.5.

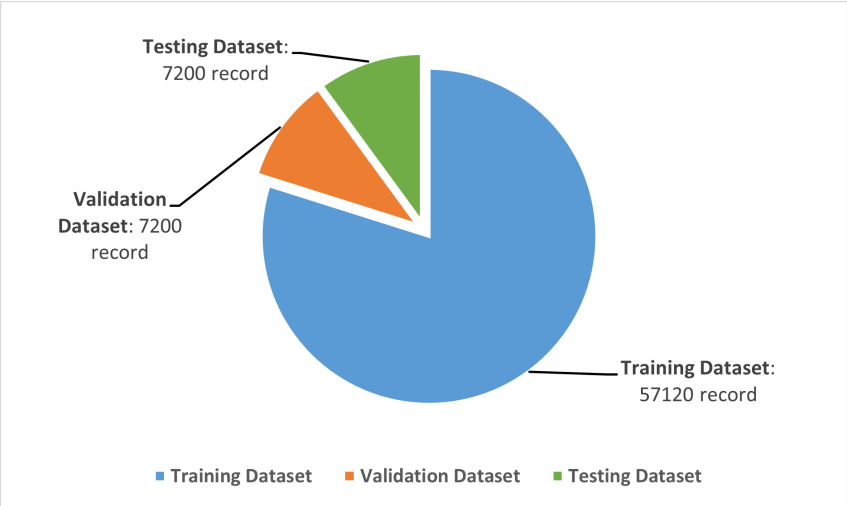

Figure 3.5: Data subdivision.

Loss functions

As previously mentioned, the architecture of the two models, employed in the two-stages deep learning method implemented, is the same, while the loss functions change.
Both loss functions used during the network optimization process combine the Cosine Similarity (CosSim) and the Mean Squared Error (MSE) or Mean Squared Deviation (MSD).
The Cosine similarity is a measure of similarity between two vectors. It is a normalized bounded inner product by L2 norms. The cosine similarity comes because the normalized dot product by the Euclidean L2 normalization is the cosine of the angle between the points denoted by the vectors in the unit sphere. It is defined by Eq. 3.2.

$$CosSim(s_1, s_2) = \frac{\sum_{n=1}^{N} s_1(n)s_2(n)}{\sqrt{\sum_{n=1}^{N} {s_1}^2(n)}\sqrt{\sum_{n=1}^{N} {s_2}^2(n)}} = \frac{\langle s_1, s_2 \rangle}{\|s_1\| \, \|s_2\|} \quad 1 \leq n \leq N \quad (3.2)$$

where s_1 and s_2 are the extracted signal (F-waves) and the ground truth one to be compared, n is the index of the current sample and N is the length of the signals

(512 samples). The more similar the two vectors (or the two signals), the more the cosine similarity gets closer to 1. Thus, $CosSim$ looks more at the shape of the extracted signals with respect to their amplitude.
The MSE instead takes the difference between each respective sample of the ground truth signal and the signal predicted by the model, squares it, and averages it out across the whole sample of the signal given in input. In other words, it provides an idea of how similar the two signals are along their entire duration.
The MSE will never be negative, since it always squares the errors, and it could be formally defined by Eq. 3.3.

$$MSE(s_{1,}s_2) \;=\; \frac{1}{N}\sum_{n=1}^{N}\left(s_2(n) \;-\; s_1(n)\right)^2 \quad 1 \leq n \leq N \tag{3.3}$$

The MSE is great for ensuring that the trained model has no outlier predictions with huge errors since it puts a larger weight on these errors due to the squaring part of the function.
Since both losses have different value ranges, have been balanced by different factors in the two models. The loss function adopted in the first model is simply the sum of the two above-described functions (Eq. 3.4), while in the second loss function, the MSE term was balanced by a λ term (Eq. 3.5). It was empirically found that $\lambda = 500$ works well in the designed deep learning setup.

$$Loss_1 = CosSim\left(s_1, s_2\right) + \; MSE(s_{1,}s_2) \tag{3.4}$$

$$Loss_2 = CosSim\left(s_1, s_2\right) + \; \lambda \cdot MSE(s_{1,}s_2) \tag{3.5}$$

3.2.4 Evaluation Metrics

The metrics used to evaluate the performances of the designed two-stage deep learning model are the sum of the square of the distances (SSD), the maximum absolute distance (MAD) and the normalized sample correlation coefficient between the extracted and the ground truth signal. Moreover, for each of the two signals also the amplitude (A) and dominant frequency (DF) were computed.

Sum of the square of the distances (SSD)
This metric measures the sum of squared distances between the ground truth and the extracted F-wave signals. It provides an idea of how similar the signals are along their entire duration. It can be computed by Eq. 3.6.

$$SSD(s_{1,}s_2) \;=\; \sum_{n=1}^{N}\left(s_2(n) \;-\; s_1(n)\right)^2 \;\; 1 \leq n \leq N \tag{3.6}$$

where s_1 and s_2 are the extracted signal (F-wave) and the ground truth one to be compared, n is the index of the current sample and N is the length of the signals (512 samples).

Maximum absolute distance (MAD)
This metric measures the maximum absolute distance between the two signals. The formula is given by Eq. 3.7.

$$MAD(s_{1,}s_2) \;=\; max\,|s_1(n) - s_2(n)| \quad 1 \leq n \leq N \tag{3.7}$$

Normalized sample correlation coefficient (ρ)
This index measures the similarity between two signals. The normalized sample correlation coefficient varies between -1 and +1 and can be expressed by Eq. 3.8.

$$\rho(s_1, s_2) \;=\; \frac{C(s_{1,}s_2)}{\sqrt{E(s_1)E(s_2)}} \;=\; \frac{1}{\sqrt{E(s_1)E(s_2)}} \sum_{n=1}^{N} s_1(n)s_2(n) \quad 1 \leq n \leq N \tag{3.8}$$

where:

$$E(s_1) \;= C(s_1, s_1) \;=\; \sum_{n=1}^{N} s_1(n)s_1(n) \tag{3.9}$$

$$E(s_2) \;= C(s_2, s_2) \;=\; \sum_{n=1}^{N} s_2(n)s_2(n) \tag{3.10}$$

Thus, signals that are as positively correlated as possible have a normalized correlation of 1 and signals that are as negatively correlated as possible have a normalized correlation of -1.

Amplitude (A)
Since F-waves are stochastic signals, the amplitude was computed as four times the standard deviation (σ_n) of all the sample's amplitude in the 1-second length window (Eq. 3.11).

$$A(s_1) = 4 \cdot \sigma_n \tag{3.11}$$

Dominant frequency (DF)
Dominant Frequency (DF) is a term used to identify the frequency related with the greatest power/amplitude in a signal. The dominant frequency is determined from a spectral analysis of the time series, i.e., Fast Fourier Transform (FFT). In this work, the DF of the signal was computed as the maximum value of the FFT of the signal in the 1-second length window (Eq. 3.12).

$$DF(s_1) = max\,|FFT(s_1)| \tag{3.12}$$

3.3 Results

Both the networks were trained for 100.000 epochs, with an early stop of the first network at epoch number 35 and an early stop of the second network at epoch number 39. The output of the two-stage deep learning method implemented corresponds to the output of the second network. For this reason, all the evaluation metrics of the results reported involve the F-wave signals extracted as the output of the second network and the original F-wave signals considered as a target.
Table 3.2, Table 3.3 and Table 3.4 report the results in the training, validation and testing dataset, in terms of evaluation metrics considered for the comparison of the two signals subdivided for the original amplitude of the input signal of the two-stage deep learning network (ECG with F-waves).

Table 3.2: Evaluation metrics results for the training dataset in terms of input signal amplitudes. Values are reported in terms of mean ± std.

TRAINING							
A INPUT signal (μV)	**A TARGET (μV)**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
05	0,008 ± 0,002	0,019 ± 0,004	6,093 ± 2,084	6,218 ± 1,774	0,754 ± 0,115	0,008 ± 0,004	0,010 ± 0,003
10	0,021 ± 0,009	0,033 ± 0,011	6,006 ± 2,134	6,227 ± 1,806	0,812 ± 0,099	0,014 ± 0,007	0,014 ± 0,005
20	0,039 ± 0,010	0,052 ± 0,011	6,949 ± 2,532	6,772 ± 2,167	0,854 ± 0,061	0,026 ± 0,012	0,021 ± 0,007
30	0,063 ± 0,027	0,066 ± 0,019	6,563 ± 2,324	6,577 ± 2,039	0,845 ± 0,070	0,059 ± 0,055	0,029 ± 0,015
40	0,091 ± 0,034	0,085 ± 0,023	5,995 ± 2,000	6,146 ± 1,664	0,841 ± 0,073	0,111 ± 0,102	0,042 ± 0,020
50	0,079 ± 0,025	0,077 ± 0,020	6,069 ± 2,159	6,167 ± 1,763	0,826 ± 0,082	0,093 ± 0,060	0,037 ± 0,015

Table 3.3: Evaluation metrics results for the validation dataset in terms of input signal amplitudes. Values are reported in terms of mean ± std.

VALIDATION							
A INPUT signal (μV)	**A TARGET (μV)**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
05	0,008 ± 0,002	0,019 ± 0,004	6,002 ± 1,956	6,186 ± 1,809	0,757 ± 0,116	0,008 ± 0,004	0,010 ± 0,003
10	0,021 ± 0,009	0,033 ± 0,011	5,990 ± 2,104	6,255 ± 1,881	0,813 ± 0,100	0,014 ± 0,007	0,014 ± 0,005
20	0,040 ± 0,010	0,051 ± 0,011	6,962 ± 2,621	6,988 ± 2,366	0,855 ± 0,062	0,026 ± 0,010	0,021 ± 0,006
30	0,062 ± 0,027	0,066 ± 0,019	6,588 ± 2,296	6,519 ± 1,890	0,848 ± 0,066	0,056 ± 0,049	0,029 ± 0,015
40	0,091 ± 0,035	0,084 ± 0,023	5,928 ± 2,092	6,086 ± 1,645	0,841 ± 0,072	0,113 ± 0,104	0,041 ± 0,019
50	0,079 ± 0,026	0,077 ± 0,020	6,078 ± 2,035	6,092 ± 1,665	0,825 ± 0,083	0,095 ± 0,069	0,037 ± 0,016

Table 3.4: Evaluation metrics results for the testing dataset in terms of input signal amplitudes. Values are reported in terms of mean ± std.

TESTING							
A INPUT signal (μV)	**A TARGET (μV)**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
05	0,008 ± 0,002	0,020 ± 0,004	6,131 ± 2,089	6,136 ± 1,725	0,749 ± 0,119	0,008 ± 0,004	0,010 ± 0,003
10	0,021 ± 0,009	0,034 ± 0,012	6,056 ± 2,139	6,245 ± 1,889	0,814 ± 0,099	0,014 ± 0,008	0,014 ± 0,005
20	0,040 ± 0,010	0,052 ± 0,011	6,818 ± 2,542	6,700 ± 2,039	0,854 ± 0,060	0,027 ± 0,012	0,021 ± 0,007
30	0,063 ± 0,027	0,066 ± 0,020	6,568 ± 2,340	6,626 ± 2,016	0,843 ± 0,071	0,060 ± 0,052	0,030 ± 0,015
40	0,091 ± 0,034	0,085 ± 0,022	5,985 ± 2,101	6,187 ± 1,805	0,840 ± 0,073	0,112 ± 0,097	0,042 ± 0,019
50	0,080 ± 0,025	0,078 ± 0,021	6,090 ± 2,114	6,104 ± 1,655	0,824 ± 0,083	0,093 ± 0,054	0,037 ± 0,015

Table 3.5, Table 3.6 and Table 3.7 instead, report the results in the training, validation and testing dataset, subdivided for the corresponding lead of the original input signals of the two-stage deep learning network.

Table 3.5: Evaluation metrics results for the training dataset in terms of input signal leads. Values are reported in terms of mean ± std.

TRAINING							
LEAD	**A TARGET (μV**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
I	0,037 ± 0,017	0,046 ± 0,014	6,286 ± 2,058	6,374 ± 1,837	0,839 ± 0,074	0,029 ± 0,025	0,019 ± 0,009
II	0,045 ± 0,021	0,053 ± 0,017	6,168 ± 2,121	6,182 ± 1,804	0,853 ± 0,065	0,037 ± 0,033	0,023 ± 0,010
V1	0,101 ± 0,021	0,097 ± 0,019	6,295 ± 1,197	6,270 ± 1,139	0,859 ± 0,069	0,125 ± 0,069	0,046 ± 0,017
V2	0,082 ± 0,026	0,079 ± 0,019	6,073 ± 1,712	6,151 ± 1,431	0,840 ± 0,076	0,101 ± 0,087	0,037 ± 0,017
V3	0,060 ± 0,025	0,063 ± 0,018	5,900 ± 1,831	6,127 ± 1,559	0,825 ± 0,082	0,064 ± 0,065	0,030 ± 0,014
V4	0,039 ± 0,015	0,047 ± 0,013	5,914 ± 2,170	6,173 ± 1,787	0,809 ± 0,089	0,032 ± 0,023	0,022 ± 0,009
V5	0,020 ± 0,010	0,030 ± 0,009	6,399 ± 3,081	6,598 ± 2,574	0,782 ± 0,100	0,014 ± 0,009	0,014 ± 0,006
V6	0,016 ± 0,008	0,027 ± 0,008	7,123 ± 3,470	6,910 ± 2,864	0,769 ± 0,111	0,012 ± 0,009	0,013 ± 0,005

Finally, Figure 3.6, Figure 3.7 and Figure 3.8 show some examples of F-wave signals extraction performed by the deep learning method implemented. In each of them is reported the original ECG signal with the F-wave component used as input to the network, the F-waves signal used as target and finally, the F-waves signal extracted by the implemented method.

Table 3.6: Evaluation metrics results for the validation dataset in terms of input signal leads. Values are reported in terms of mean ± std.

VALIDATION							
LEAD	**A TARGET (μV)**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
I	0,038 ± 0,017	0,047 ± 0,015	6,259 ± 2,108	6,340 ± 1,789	0,843 ± 0,073	0,028 ± 0,021	0,020 ± 0,009
II	0,045 ± 0,022	0,053 ± 0,018	6,186 ± 2,156	6,184 ± 1,847	0,855 ± 0,065	0,038 ± 0,033	0,023 ± 0,011
V1	0,099 ± 0,021	0,096 ± 0,019	6,336 ± 1,184	6,282 ± 1,073	0,862 ± 0,067	0,119 ± 0,064	0,046 ± 0,017
V2	0,081 ± 0,026	0,078 ± 0,019	6,118 ± 1,726	6,121 ± 1,305	0,838 ± 0,077	0,099 ± 0,079	0,038 ± 0,016
V3	0,060 ± 0,024	0,063 ± 0,017	5,904 ± 1,831	6,076 ± 1,528	0,824 ± 0,083	0,064 ± 0,059	0,030 ± 0,014
V4	0,040 ± 0,015	0,048 ± 0,013	6,031 ± 2,273	6,270 ± 1,795	0,810 ± 0,085	0,033 ± 0,022	0,022 ± 0,009
V5	0,020 ± 0,010	0,030 ± 0,009	6,504 ± 3,080	6,634 ± 2,621	0,780 ± 0,097	0,014 ± 0,010	0,014 ± 0,006
V6	0,016 ± 0,009	0,027 ± 0,009	7,169 ± 3,372	7,057 ± 2,831	0,766 ± 0,111	0,012 ± 0,007	0,013 ± 0,005

Table 3.7: Evaluation metrics results for the testing dataset in terms of input signal leads. Values are reported in terms of mean ± std.

TESTING							
LEAD	**A TARGET (μV)**	**A OUTPUT (μV)**	**DF TARGET (Hz)**	**DF OUTPUT (Hz)**	ρ	**SSD (au)**	**MAD (au)**
I	0,038 ± 0,017	0,046 ± 0,015	6,356 ± 2,149	6,292 ± 1,806	0,835 ± 0,078	0,029 ± 0,024	0,020 ± 0,008
II	0,044 ± 0,022	0,052 ± 0,017	6,047 ± 2,052	6,066 ± 1,665	0,853 ± 0,065	0,038 ± 0,036	0,023 ± 0,011
V1	0,102 ± 0,021	0,098 ± 0,019	6,334 ± 1,190	6,270 ± 1,092	0,857 ± 0,068	0,129 ± 0,069	0,046 ± 0,016
V2	0,083 ± 0,027	0,079 ± 0,020	6,041 ± 1,647	6,227 ± 1,462	0,839 ± 0,079	0,105 ± 0,090	0,038 ± 0,017
V3	0,061 ± 0,027	0,063 ± 0,019	5,842 ± 1,686	6,007 ± 1,421	0,825 ± 0,079	0,068 ± 0,075	0,030 ± 0,016
V4	0,040 ± 0,016	0,048 ± 0,014	6,026 ± 2,291	6,111 ± 1,736	0,807 ± 0,088	0,033 ± 0,024	0,022 ± 0,009
V5	0,020 ± 0,010	0,030 ± 0,009	6,613 ± 3,067	6,716 ± 2,721	0,780 ± 0,106	0,014 ± 0,008	0,014 ± 0,006
V6	0,016 ± 0,008	0,027 ± 0,008	7,096 ± 3,473	7,040 ± 2,916	0,765 ± 0,109	0,012 ± 0,007	0,013 ± 0,005

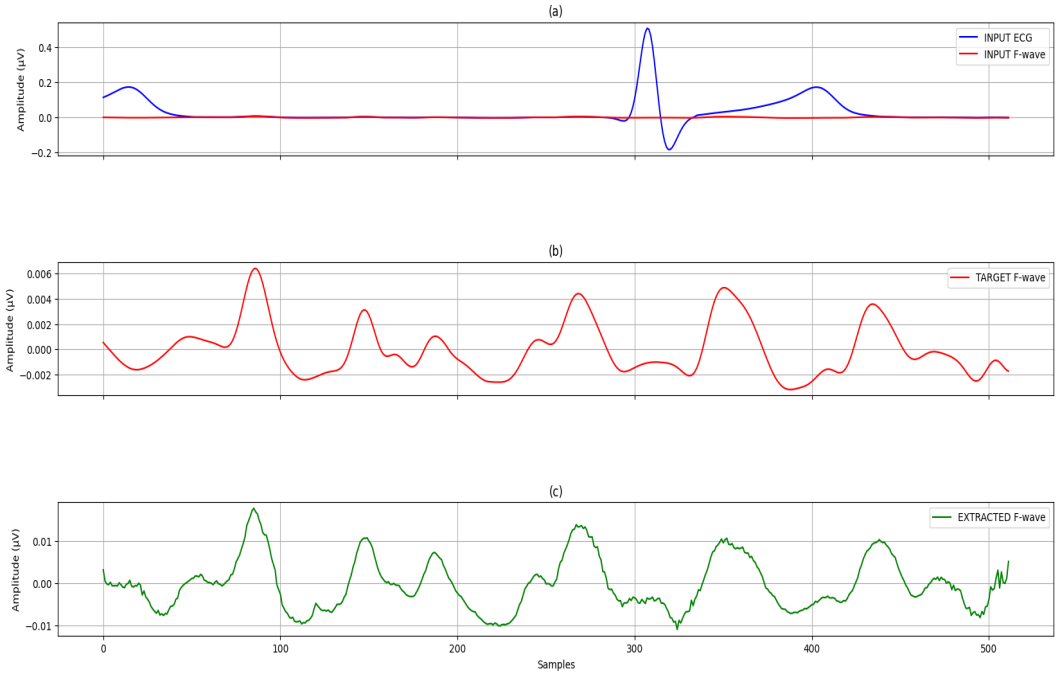

Figure 3.6: Example of F-wave extraction performed by the two-stage deep learning method implemented [1/3]. ECG with F-wave (a). Target F-wave (b). F-wave extracted by the network (c).

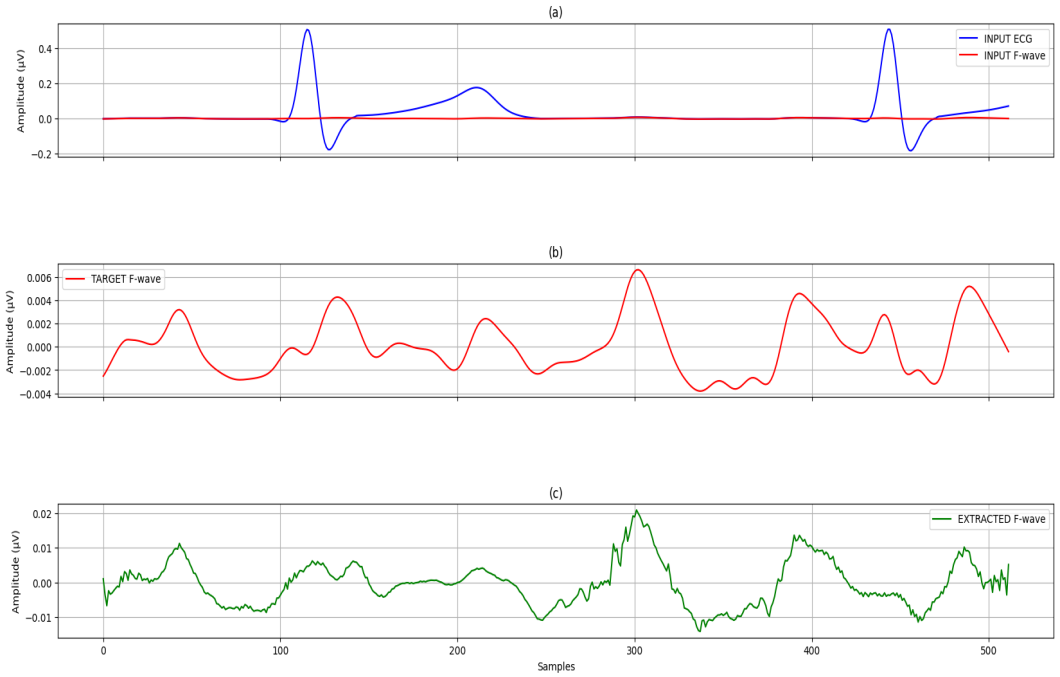

Figure 3.7: Example of F-wave extraction performed by the two-stage deep learning method implemented [1/3]. ECG with F-wave (a). Target F-wave (b). F-wave extracted by the network (c).

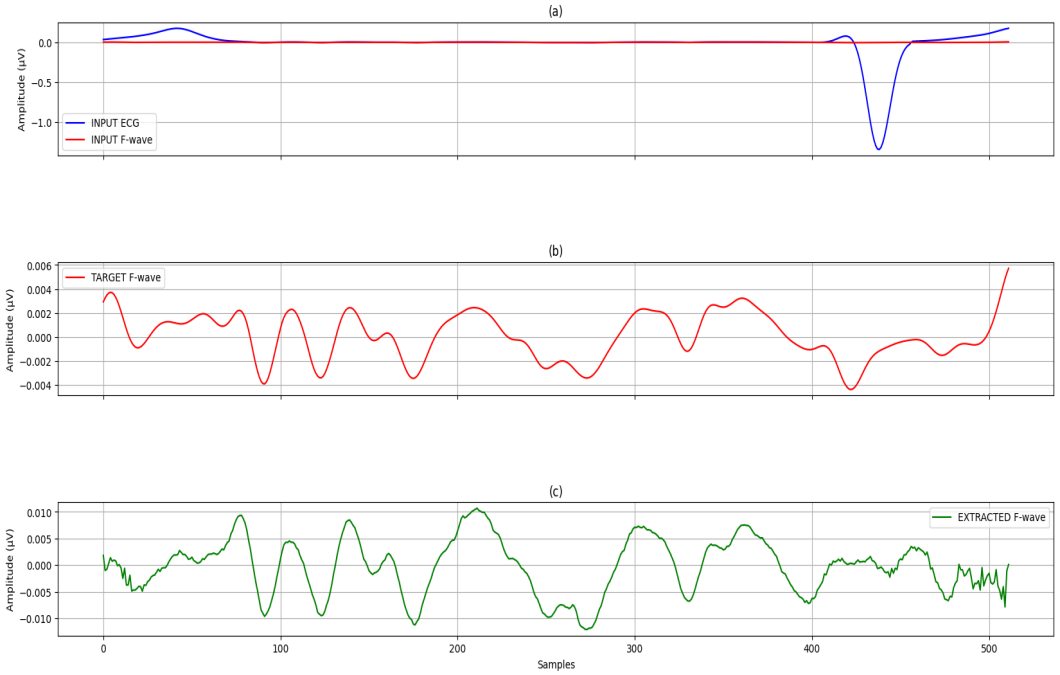

Figure 3.8: Example of F-wave extraction performed by the two-stage deep learning method implemented [1/3]. ECG with F-wave (a). Target F-wave (b). F-wave extracted by the network (c).

3.4 Discussion

In this work it is proposed a two-stage deep learning method to extract F-waves from ECG signals of patients affected by atrial fibrillation. Both the two designed deep learning models have 1D CNN architectures based on multipath modules. The hyperparameters of the network and the loss functions used were defined empirically, training the networks several times and evaluating their performances. The idea behind the use of two networks on cascade, instead of only one, is based on experimental results. It has been seen that, after the first stage, the waveforms of the extracted F-waves are similar in shape to the ground truth F-waves but are inevitably distorted in amplitude. This distortion could be attributed to the downsampling and normalization performed by the first network. For this reason, a second network was introduced, with the same architecture as the first but with a different loss function. By emphasizing the distance between the extracted F-waves and ground truth signals with respect to shape, the second loss function is intended to restore the amplitude of the extracted F-waves from the first network. In other words, this second network doesn't learn how to extract the F-waves from the original signal but rather learns to recover the information lost after the first stage. The second network can be seen as a filter for the amplitude of the signal in input, i.e. the extracted F-waves from the first deep learning model.
According to the results in terms of evaluation metrics analyzed, the performances

of the implemented method are promising. It can be noticed that, independently the signals given as input to the network belong to the training, validation or testing dataset, the extracted F-waves share similar characteristics. This suggests that the implemented method was trained to avoid the problem of overfitting. The input signals differ in terms of both amplitude and lead used to record the ECG. This means that the proposed two-stage deep learning method is strong enough to be able to manage signals that have different shapes and amplitude. The results reported separately for these two different characteristics of the input signal show that between the extracted F-waves and the target signals the correlation coefficient is always high, with the lowest value of 0.74. In particular, it can be noticed that, according to the different amplitude of the input signals, the highest values of correlation were reached for 20 μV amplitude, even if values for 10,30,40,50 μV amplitude are similar, while the lowest values were obtained for 5 μV of input signal amplitude. Analyzing instead the correlation according to the different leads used to record the ECG signal, it can be noticed that, the bipolar leads, together with V1, V2, V3, and V4, show the highest values. V5 and V6, between the precordial leads, are the ones placed farthest from the heart and mainly record the activity of the ventricles; therefore, the recorded R waves and F-waves components are smaller with respect to the one recorded from the other leads. This could explain the lowest values of correlation obtained for their relative input signals.

The F-wave repetition rate, also known as the dominant atrial frequency (DF), helps monitor drug therapy and predict both spontaneous and drug-induced AF termination. Results show that this feature is maintained in the waves extracted by the model, with mean absolute error with respect to the target in the testing dataset of 0.052 Hz. Similar results were obtained for the amplitudes of the output signals. Finally, also the obtained values of SSD and MAD are low, demonstrating the strength of this method and the similarity of the extracted waves in terms of amplitude and shape with respect to the target.

Data used, and the preprocessing stage applied was oriented to increase the dimension of the dataset, together with its randomization. A point of strength of this method was to use only real ECG and F-waves signals, instead of synthetic signals, also available in the original dataset. Furthermore, splitting the original signal in the 1s length windows increased a lot the dimension of the input data and the different shifting of the widows according to the different amplitude of the input signals increased their randomization. The windows were not designed to obtain 1 beat of ECG for each of them. Instead, one window could contain two peaks (R-R) or only one, and it could start with P-waves but also QRS complex or T-waves. Most of the studies on the analysis of ECG records for AF detection are based on heart rate variability, i.e., on the R-R intervals. However, the above-described preprocessing of the input data makes the implemented method completely independent of the RR distance of the input ECG signal.

On the other hand, the main drawback of this work can be related to the data

used. Firstly, according to the European guidelines [23], the diagnosis of AF requires rhythm documentation acquired by 12-lead ECG showing AF analyzed by a physician with expertise in ECG rhythm interpretation with at least 30 s of duration. This method is limited to extracting F-waves from a 1-second ECG window, so, for clinical usage, it requires further improvement in post-processing the output signals.
Moreover, all the ECG signals given as input to train the net contain F-waves. This was the initial hypothesis of this work, but, given the promising results, it might be interesting to try to train the network using as input both ECG signals containing F-waves and signals recorded by healthy patients, not affected by AF.
Thanks to the results obtained using CNN for the classification and analysis of images, nowadays the use of deep learning, and in particular of 1D CNN is increasingly increasing for the study of biomedical signals. For what concerns the study of AF, and particularly the extraction of the characteristic F-waves from recorded ECG signals of diseased patients, no works have been found in the literature using deep learning methods. Classical techniques used to carry out this task involve the use of different signal processing principles such as PCA, ICA, etc...
For this reason, this work, given the promising results obtained, can be considered a forerunner for this branch of research.

Conclusion

This thesis proposed an in-depth analysis regarding a disturbance in the conduction of the electrical impulse, i.e. atrial fibrillation. Specifically, attention has been paid to the characteristic F-waves, which occur, even if with a small amplitude, in the ECG of patients affected by this type of arrhythmia. The detection, diagnosis and extraction of these waves is a complex process; to achieve these aims, the most innovative techniques and methodologies of deep learning were examined. Due to the lack of work with the same goal in the literature, the most recent works in which deep learning techniques were utilized to filter ECG signals and eliminate noisy components were reviewed.

Thus, a two-step deep learning model was proposed to extract the F-wave component from a 1-second length ECG signal window of an AF-affected person. Although the results obtained and evaluated in terms of evaluation metrics are really promising, there is still room for improvement.

Specifically, for a more accurate clinical evaluation of the pathology, longer ECG signal input windows should be used or signal post-processing techniques should be designed to merge the output signals and obtain F-waves that last longer. A further step of improvement could be to modify the network so that it can evaluate, before proceeding with the extraction of the F-waves, if these are contained in the input signal. In other words, the network could be trained with ECG input signals of both healthy and AF-affected patients.

In this thesis, it was demonstrated that deep learning techniques, especially CNNs, can identify and extract patterns in biomedical signals like ECG, and can be exploited, even with simple implementation, to extract stochastic signal components with excellent results.

Despite the use of already known and established deep learning techniques, this work is innovative in their use as filters to extract a component of the principal signal, rather than the recorded one (ECG), and for the application field, i.e. atrial fibrillation and F-waves.

Bibliography

[1] J. Gordon Betts, Kelly A. Young, James A. Wise, Eddie Johnson, Brandon Poe, Dean H. Kruse, Oksana Korol, Jody E. Johnson, Mark Womble, and Peter DeSaix. *Anatomy and Physiology.* OpenStax, Houston, Texas, April 2013.

[2] Patti L and Ashurst JV. *Supraventricular Tachycardia.* . Treasure Island (FL): StatPearls Publishing; 2023 Jan-., StatPearls [Internet], August 2023.

[3] Ludhwani D. *Paroxysmal Atrial Fibrillation.* StatPearls [Internet]., Treasure Island (FL): StatPearls Publishing, June 2023.

[4] Anthony Dupre, Sarah Vincent, and Paul A. Iaizzo. Basic ECG Theory, Recordings, and Interpretation. In Paul A. Iaizzo, editor, *Handbook of Cardiac Anatomy, Physiology, and Devices*, pages 191–201. Humana Press, Totowa, NJ, 2005.

[5] Richard E. Gregg, Sophia H. Zhou, James M. Lindauer, Eric D. Helfenbein, and Karen K. Giuliano. What is inside the electrocardiograph? *Journal of Electrocardiology*, 41(1):8–14, 2008.

[6] Bhumika Chandrakar, OP Yadav, and VK Chandra. A survey of noise removal techniques for ECG signals. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(3):1354–1357, 2013.

[7] Fengru Liu, Yijie Xu, and Yuan Yao. Highly Efficient Low Noise Solutions in ECG Signals. *Journal of Physics: Conference Series*, 2246(1):012030, April 2022.

[8] Aswathy Velayudhan and Soniya Peter. Noise Analysis and Different Denoising Techniques of ECG Signal - A Survey. 2016.

[9] Suranai Poungponsri and Xiao-Hua Yu. An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks. *Neurocomputing*, 117:206–213, October 2013.

[10] M.A.D. Raya and L.G. Sison. Adaptive noise cancelling of motion artifact in stress ECG signals using accelerometer. In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, volume 2, pages 1756–1757, Houston, TX, USA, 2002. IEEE.

[11] Zia-ul-Haque , Rizwan Qureshi, Mehmood Nawaz, Faheem Yar, Nazish Tunio, and Muhammad Uzair. Analysis of ECG Signal Processing and Filtering Algorithms. *International Journal of Advanced Computer Science and Applications*, 10(3), 2019.

[12] V. Fuster, L. E. Rydén, R. W. Asinger, D. S. Cannom, H. J. Crijns, R. L. Frye, J. L. Halperin, G. N. Kay, W. W. Klein, S. Lévy, R. L. McNamara, E. N. Prystowsky, L. S. Wann, D. G. Wyse, R. J. Gibbons, E. M. Antman, J. S. Alpert, D. P. Faxon, V. Fuster, G. Gregoratos, L. F. Hiratzka, A. K. Jacobs, R. O. Russell, S. C. Smith, W. W. Klein, A. Alonso-Garcia, C. Blomström-Lundqvist, G. De Backer, M. Flather, J. Hradec, A. Oto, A. Parkhomenko, S. Silber, A. Torbicki, and American College of Cardiology/American Heart Association/European Society of Cardiology Board. ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation: executive summary. A Report of the American College of Cardiology/ American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines and Policy Conferences (Committee to Develop Guidelines for the Management of Patients With Atrial Fibrillation): developed in Collaboration With the North American Society of Pacing and Electrophysiology. *Journal of the American College of Cardiology*, 38(4):1231–1266, October 2001.

[13] Rachael Zimlich, Debbie Nurmi, and Kristin Shaffer. Types of Atrial Fibrillation: What You Need to Know. November 2023.

[14] Ashish Shukla and Anne B. Curtis. Avoiding permanent atrial fibrillation: treatment approaches to prevent disease progression. *Vascular Health and Risk Management*, 10:1–12, 2014.

[15] Hugh Calkins, Karl Heinz Kuck, Riccardo Cappato, Josep Brugada, A. John Camm, Shih-Ann Chen, Harry J. G. Crijns, Ralph J. Damiano, D. Wyn Davies, John DiMarco, James Edgerton, Kenneth Ellenbogen, Michael D. Ezekowitz, David E. Haines, Michel Haissaguerre, Gerhard Hindricks, Yoshito Iesaka, Warren Jackman, Jose Jalife, Pierre Jais, Jonathan Kalman, David Keane, Young-Hoon Kim, Paulus Kirchhof, George Klein, Hans Kottkamp, Koichiro Kumagai, Bruce D. Lindsay, Moussa Mansour, Francis E. Marchlinski, Patrick M. McCarthy, J. Lluis Mont, Fred Morady, Koonlawee Nademanee, Hiroshi Nakagawa, Andrea Natale, Stanley Nattel, Douglas L. Packer, Carlo Pappone, Eric Prystowsky, Antonio Raviele, Vivek Reddy, Jeremy N. Ruskin, Richard J. Shemin, Hsuan-Ming Tsao, and David Wilber. 2012 HRS/EHRA/ECAS Expert Consensus Statement on Catheter and Surgical Ablation of Atrial Fibrillation: recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design. *Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology: Journal*

of the Working Groups on Cardiac Pacing, Arrhythmias, and Cardiac Cellular Electrophysiology of the European Society of Cardiology, 14(4):528–606, April 2012.

[16] W. Evans and P. Swann. Lone auricular fibrillation. *British Heart Journal*, 16(2):189–194, April 1954.

[17] Alan S. Go, Elaine M. Hylek, Kathleen A. Phillips, YuChiao Chang, Lori E. Henault, Joe V. Selby, and Daniel E. Singer. Prevalence of Diagnosed Atrial Fibrillation in AdultsNational Implications for Rhythm Management and Stroke Prevention: the AnTicoagulation and Risk Factors In Atrial Fibrillation (ATRIA) Study. *JAMA*, 285(18):2370–2375, May 2001. _eprint: https://jamanetwork.com/journals/jama/articlepdf/193807/jcc10004.pdf.

[18] Jelena Kornej, Christin S. Börschel, Emelia J. Benjamin, and Renate B. Schnabel. Epidemiology of Atrial Fibrillation in the 21st Century. *Circulation Research*, 127(1):4–20, 2020. _eprint: https://www.ahajournals.org/doi/pdf/10.1161/CIRCRESAHA.120.316340.

[19] Craig T. January, L. Samuel Wann, Joseph S. Alpert, Hugh Calkins, Joaquin E. Cigarroa, Joseph C. Cleveland, Jamie B. Conti, Patrick T. Ellinor, Michael D. Ezekowitz, Michael E. Field, Katherine T. Murray, Ralph L. Sacco, William G. Stevenson, Patrick J. Tchou, Cynthia M. Tracy, Clyde W. Yancy, and ACC/AHA Task Force Members. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. *Circulation*, 130(23):2071–2104, December 2014.

[20] Jonathan W. Waks and Mark E. Josephson. Mechanisms of Atrial Fibrillation - Reentry, Rotors and Reality. *Arrhythmia & Electrophysiology Review*, 3(2):90–100, August 2014.

[21] Richard E. Klabunde. *Cardiovascular physiology concepts*. Lippincott Williams & Wilkins, Philadelphia, 2005. OCLC: ocm56131691.

[22] Jader Giraldo-Guzman, Sonia H. Contreras-Ortiz, Marian Kotas, Francisco Castells, and Tomasz Moron. Automated Atrial Fibrillation Detection by ECG Signal Processing: A Review. *Critical Reviews in Biomedical Engineering*, 49(3):31–50, 2021.

[23] Gerhard Hindricks, Tatjana Potpara, Nikolaos Dagres, Elena Arbelo, Jeroen J. Bax, Carina Blomström-Lundqvist, Giuseppe Boriani, Manuel Castella, Gheorghe-Andrei Dan, Polychronis E. Dilaveris, Laurent Fauchier, Gerasimos Filippatos, Jonathan M. Kalman, Mark La Meir, Deirdre A. Lane, Jean-Pierre

Lebeau, Maddalena Lettino, Gregory Y. H. Lip, Fausto J. Pinto, G. Neil Thomas, Marco Valgimigli, Isabelle C. Van Gelder, Bart P. Van Putte, Caroline L. Watkins, and ESC Scientific Document Group. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *European Heart Journal*, 42(5):373–498, February 2021.

[24] Martin Grond, Marek Jauss, Gerhard Hamann, Erwin Stark, Roland Veltkamp, Darius Nabavi, Markus Horn, Christian Weimar, Martin Köhrmann, Rolf Wachter, Ludger Rosin, and Paulus Kirchhof. Improved detection of silent atrial fibrillation using 72-hour Holter ECG in patients with ischemic stroke: a prospective multicenter cohort study. *Stroke*, 44(12):3357–3364, December 2013.

[25] Yoav Mintz and Ronit Brodie. Introduction to artificial intelligence in medicine. *Minimally invasive therapy & allied technologies: MITAT: official journal of the Society for Minimally Invasive Therapy*, 28(2):73–81, April 2019.

[26] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.

[27] Mohaiminul Islam, Guorong Chen, and Shangzhu Jin. An Overview of Neural Network. *American Journal of Neural Networks and Applications*, 5(1):7, 2019.

[28] Siddharth Sharma, Simone Sharma, and Anidhya Athaiya. ACTIVATION FUNCTIONS IN NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, 04(12):310–316, May 2020.

[29] Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, April 2021.

[30] Daniel Jakubovitz, Raja Giryes, and Miguel R. D. Rodrigues. Generalization Error in Deep Learning. 2018. Publisher: arXiv Version Number: 3.

[31] Xue Ying. An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, 1168:022022, February 2019.

[32] Supriya O. Rajankar and Sanjay N. Talbar. An optimum ECG denoising with wavelet neural network. In *2015 International Conference on Pervasive Computing (ICPC)*, pages 1–4, Pune, India, January 2015. IEEE.

[33] Yue Qiu, Feng Xiao, and Haibin Shen. Elimination of power line interference from ECG signals using recurrent neural networks. *Annual International Conference*

of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2017:2296–2299, July 2017.

[34] Corneliu T.C. Arsene, Richard Hankins, and Hujun Yin. Deep Learning Models for Denoising ECG Signals. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5, A Coruna, Spain, September 2019. IEEE.

[35] F.P. Romero, D.C. Piñol, and C.R. Vázquez-Seisdedos. DeepFilter: An ECG baseline wander removal filter using deep learning techniques. *Biomedical Signal Processing and Control*, 70, 2021. Publisher: Elsevier Ltd.

[36] Pratik Singh and Gayadhar Pradhan. A New ECG Denoising Framework Using Generative Adversarial Network. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(2):759–764, March 2021.

[37] B. Xu, R. Liu, and Y. Wang. An ECG Sparse Noise Reduction Method based on Deep Unfolding Network. In Xu B., editor, *IMCEC - IEEE Adv. Inf. Manag., Commun., Electron. Autom. Control Conf.*, pages 1078–1082. Institute of Electrical and Electronics Engineers Inc., 2021. Journal Abbreviation: IMCEC - IEEE Adv. Inf. Manag., Commun., Electron. Autom. Control Conf.

[38] Lishen Qiu, Wenqiang Cai, Miao Zhang, Wenliang Zhu, and Lirong Wang. Two-stage ECG signal denoising based on deep convolutional network. *Physiological Measurement*, 42(11):115002, November 2021.

[39] Zixiao He, Xinwen Liu, Hao He, and Huan Wang. Dual Attention Convolutional Neural Network Based on Adaptive Parametric ReLU for Denoising ECG Signals with Strong Noise. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 779–782, Mexico, November 2021. IEEE.

[40] Amir Mohammadisrab, Poorya Aghaomidi, Jalil Mazloum, Mohammad Ali Akbarzadeh, Mahdi Orooji, Nader Mokari, and Halim Yanikomeroglu. Deep Adaptive Denoising Auto-Encoder Networks for ECG Noise Cancellation via Time-Frequency Domain. In *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 1449–1456, Chiang Mai, Thailand, November 2022. IEEE.

[41] Eoin Brophy, Bryan Hennelly, Maarten De Vos, Geraldine Boylan, and Tomas Ward. Improved Electrode Motion Artefact Denoising in ECG Using Convolutional Neural Networks and a Custom Loss Function. *IEEE Access*, 10:54891–54898, 2022.

[42] Raghavendra Badiger and Prabhakar M. ASCNet-ECG: Deep Autoencoder based Attention aware Skip Connection network for ECG filtering. *International Journal of Engineering Trends and Technology*, 71(2):382–398, February 2023.

[43] Huidong Wang, Yurun Ma, Aihua Zhang, Dongmei Lin, Yusheng Qi, and Jiaqi Li. Deep Convolutional Generative Adversarial Network with LSTM for ECG Denoising. *Computational and Mathematical Methods in Medicine*, 2023:1–17, February 2023.

[44] Yanrui Jin, Chengjin Qin, Jinlei Liu, Yunqing Liu, Zhiyuan Li, and Chengliang Liu. A novel deep wavelet convolutional neural network for actual ECG signal denoising. *Biomedical Signal Processing and Control*, 87:105480, January 2024.

[45] Raúl Alcaraz, Leif Sörnmo, and José J Rieta. Reference database and performance evaluation of methods for extraction of atrial fibrillatory waves in the ECG. *Physiological Measurement*, 40(7):075011, August 2019.

[46] M. Lemay, V. Jacquemet, A. Forclaz, J.M. Vesin, and L. Kappenberger. Spatiotemporal QRST cancellation method using separate QRS and T-waves templates. In *Computers in Cardiology, 2005*, pages 611–614, Lyon, France, 2005. IEEE.

[47] Reza Sameni, Gari D Clifford, Christian Jutten, and Mohammad B Shamsollahi. Multichannel ECG and Noise Modeling: Application to Maternal and Fetal ECG Signals. *EURASIP Journal on Advances in Signal Processing*, 2007(1):043407, December 2007.

[48] Andrius Petrenas, Vaidotas Marozas, Andrius Sološenko, Raimondas Kubilius, Jurgita Skibarkiene, Julien Oster, and Leif Sörnmo. Electrocardiogram modeling during paroxysmal atrial fibrillation: application to the detection of brief episodes. *Physiological Measurement*, 38(11):2058–2080, November 2017.

[49] R. Bousseljot, D. Kreiseler, and A. Schnabel. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik/Biomedical Engineering*, pages 317–318, July 2009.

[50] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, Boston, MA, USA, June 2015. IEEE.