International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2022	Issue 53	Volume 14	Number 4	Pages 348-356

ANALYTICAL STUDY OF ESSENTIAL VOCABULARY RELATED TO ECG: A SURVEY

W.R.M. Al-Rawi A.F. Hussein

Biomedical Engineering Department, Faculty of Engineering, Al Nahrain University, Baghdad, Iraq
st.warda.r.mohammed@ced.nahrainuniv.edu.iq, ahmed.f.h.1976@gmail.com

Abstract- The signal on an electrocardiogram (ECG) reflects the electrical activity of the heart. Monitoring heart rate, assessing pulse rhythm, diagnosing cardiac problems, identifying emotions, and biometric identification are just a few of the many uses for the ECG signal that have been explored and implemented in the scientific literature. Several steps, including as processing, extraction of features, classification techniques, feature alteration, and categorization, may be required for ECG analysis, depending on the nature of the investigation. For the accompanying comments, each step must be completed. In addition, the employed success criteria and the adequacy of the ECG signal database play critical roles in the analysis. ECG segmentation and feature extraction play a vital role in detecting the great majority of cardiac disorders. Electrocardiogram (ECG) signals can be acquired for various purposes, including cardiovascular disease diagnosis, arrhythmia recognition, physiological feedback, sleep apnea detection, regular patient monitoring, prediction of sudden cardiac arrest, and vital, emotional, and physical activity recognition systems. The ECG has been widely employed to diagnose and prognosis a variety of heart diseases. Current computerized automated reports can independently diagnose various heart diseases and generate an automatic alert with an acceptable degree of precision. This research will shed light on the most critical electrocardiography topics from a medical and analytic perspective. The second purpose is to facilitate the work of researchers by compiling the most vital data sets. The essential educational methods.

Keywords: Electrocardiogram, Wave, Leads, Signals, Noise, Dataset.

1. INTRODUCTION

Biomedical signals are typically obtained by monitoring (detection or estimation) particular pathological/physiological states in order to assess diagnosis and treatment. They are used to decode and eventually activate certain biological processes in fundamental research. More than that, today's technology allows for many channels of these signals to be recorded. Thus, unique signal conditioning challenges arise when attempting to measure interactions across these channels that are physiologically significant. The goals of signal processing in all these contexts range from eliminating noise to accurately quantifying the signal pattern and its components via analysis (system characterization for modeling and control) to extracting features for use in making decisions (such as in heart replacement devices) [1]. Biological applications utilize signal-processing techniques regularly for the aforementioned reasons. There is always some background noise present in measurements of biological signals. Picking an effective signal-processing approach is necessary for instruments (sensor systems, amplifiers, filtration, etc.), electromagnetic fields (EMI), and any asynchronous, uncorrelated signal having noise characteristics. The purpose of automated ECG analyzers is to aid in diagnosis. Even though they will not be able to take the position of trained medical experts, their points of view may nonetheless be regarded a second, more objective viewpoint. The currently available automated diagnostic techniques are unreliable for identifying ischemic episodes, fibrillation, and arrhythmias. These approaches also have a variety of other limitations and issues [2].

In recent years, the automated identification of cardiac issues has been significantly aided by computer-assisted ECG interpretation, which played a big role in the past. CVD has a substantially higher mortality rate than other common diseases such as cancer; hence, automated methods of evaluating and diagnosing it are necessary. Over the past several decades, numerous researchers have contributed to developing CVD detection algorithms, particularly since the advent of mobile computing devices such as tablets and smartphones [3]. These mobile computer systems enhance the acceptability of wearable diagnostic equipment to detect and analyses cardiovascular diseases. New designs are also motivated by the need to reduce the number of leads, which is accomplished by selecting only the essential leads and reducing the bandwidth size required to transmit the ECG signal during this difficult time and internet congestion brought on by the current COVID-19 pandemic, which is challenging hospital capacity and health care workers. Still looking for a realistic and cost-effective solution to the complex difficulties.

In countries severely affected by the pandemic, healthcare institutions are on the verge of collapse, with hospital and departmental capacities continually being tested. All contemporary devices are intended to record as much information as possible about the electrical activity of the heart with as few leads as possible. In these difficult circumstances, technology alternatives are emphasized more than ever, and the need to change healthcare from antiquated methods to technology-driven solutions is emphasized [4]. Internet of Things (IoT) and other healthcare technologies (machine learning, cloud, edge computing, and deep learning) have been in development for years, but none of the innovations were designed to withstand extreme pressure, such as that experienced during pandemics and other exceptional circumstances. In order to solve the deteriorating global healthcare infrastructure, there must be a fundamental transformation in healthcare technology.

The most typical limitations for Remote Health Monitoring (RM) situations are bandwidth, storage, and data transmission time. Researchers in the physical sciences and engineering are working to address these issues, generate new concepts, and recognize new difficulties. Internet-based telemedicine services are advantageous since they provide rapid and efficient medical treatment. Modern and sophisticated wearable and wireless sensor technologies are frequently utilized by these online healthcare solutions. As a result of rapid technical improvements, a number of remote health monitoring systems have enhanced their capabilities [5]. The primary objective of this paper is to introduce the reader to ECG and to provide a succinct analysis of the most vital issues associated with the notion of cardiology.

2. ELECTROCARDIOGRAM ECG BACKGROUND

Researchers noticed that the heart's electrical currents might have been investigated more than a century ago. Willem Einthoven, a Dutch scientist who won the 1924 Nobel Prize in physiology or medicine, was the inventor of the ECG as it is known today [6]. Consider A cardiac electrocardiogram (ECG) is one of the necessary additional medical tests that can be performed with a cardiograph to detect and assess heart problems through waveform recording, which is the recording of potential electrical changes between two points that occur during the electrical activity of the heart. Figure 1 depicts a healthy person's electrocardiogram (ECG) [1].

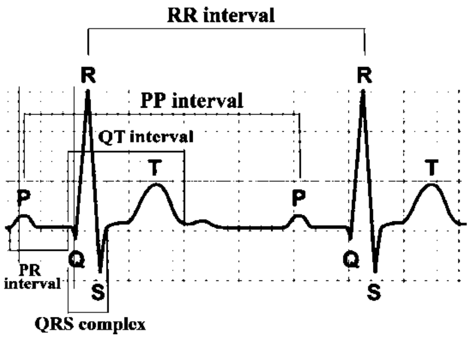


Figure 1. Healthy human electrocardiogram ECG [7]

The information in Figure 1 is describe in Table 1 to illustrate the abbreviation's significance.

Table 1. ECG symbol abbreviations [4, 8]

Topic	Definition	Morphology	Time duration
P wave	The P wave indicates depolarization of the ventricles. Hyperpolarization of the atria travels from the sinoatrial (SA) node to the atrioventricular (AV) nodes and from the right to the left atrium	An aberrant atrial defibrillator implant may be indicated by an atypical P - waves axis (inverted in other leads). Commonly, all leads (except aVR) will show an upright P wave. An enlarged atrium might be the cause of the P wave's abnormally long duration. The P wave produced by a big right atrial tends to have a high peak, whereas the P wave produced by a large left atrium often has two humps	less than eighty minutes
PR interval	A patient's PR interval is the period between the start of their P waves and the start of their QRS complex. In order for electrical activity to go from the septum to the AV nodes, this delay is required	Wolf-Parkinson-White syndrome is a disorder in which the AV node is bypassed, as shown by a PR duration of less than 120 ms. A PR interval greater than 200 milliseconds implies atrioventricular block of the first degree. Pericarditis may cause a decrease in the PR sector (the ECG segment after the P wave and before the QRS complicated)	Start between 120 and 200 minutes
QRS complex	Rapid activation of the right and left ventricular is reflected in the QRS complexes. Because the ventricles have so much more muscle than the atria, the Contraction often has a much higher amplitude than the P waveform	In cardiac conditions like LBBB and RBBB, as well as ventricular arrhythmias like ventricular tachycardia, a prolonged QRS complexes (more than 120 ms) indicates a problem with the heart's atrioventricular node. The QRS complicated may become disproportionately big in response to metabolic problems such acute hyperkalemia or an overdose of tricyclic antidepressants. Abnormally large QRS complexes may suggest left ventricular hypertrophy, whereas abnormally small QRS complexes may indicate peritoneal hemorrhage or infiltrative myocardium. myocardial disease	Start between 80 and 100 minutes
J-point	Between the QRS complexity and the ST segment lies the J-point	As a typical variant, the J-point may be raised. Presence of a distinct J waveform or Bates wave at the J-point is diagnostic of hyperthermia or metabolic alkalosis	null

ST-segment	The ST segment connects the QRS complex to the T wave and represents ventricular depolarization.	Myocardial infarction or ischemia can cause a decrease or an increase. Additionally, LVH and digoxin can cause ST depression. ST-elevation may be attributable to pericarditis, Brugada syndrome, or a spontaneous fluctuation (J-point elevation).	null
T wave	The T wave represents the repolarization of the ventricles. Except for aVR and V1, it is typically vertical in all leads.	Inverted T waves may suggest cardiac ischemia, enlargement of the left ventricle, elevated intracranial pressure, or metabolic problems. T-wave peaks may indicate early hyperkalemia or myocardial infarction.	160 minutes
Corrected QT interval (QTc)	From the starting point of the QRS complicated to the conclusion of the T wave is what is known as the QT interval. Acceptable ranges vary with heart rate, so to derive the QTc, the QT interval must be modified by dividing by the square root of the RR interval.	A prolonged QTc interval raises the likelihood of ventricular tachyarrhythmias and cardiac arrest. Prolonged QT can be caused by a genetic disorder or a side effect of certain medications. An excessively short QTc characterizes extreme hypercalcemia.	Less than 440 minutes
U wave	The U wave is hypothesized to be caused by the repolarization of the interventricular septum. It has an ordinarily low amplitude and is frequently absent.	A pronounced U wave may indicate hypokalemia, hypercalcemia, or hyperthyroidism.	null

3. CG FEATURES

Existing ECG-based identification systems offer both advantages and disadvantages. On the basis of ECG data, inter-subject variability is utilized to characterize an individual (features). Features are generated using the shape of the heartbeat, varied time intervals created by ECG waves, and accurately extracted features. Various factors, like the recognizer's complexity, the necessity for instantaneous recognition, the nature of the recording device, etc., determine the relevant parameters. However, opinions vary widely on the best approach to take and the number and kind of criteria to use in an assessment.

In addition, ECG analysis is typically performed on proprietary datasets, making comparisons between methods unfeasible. And Fiducial-reliant strategies by identifying exact anchor points on ECG recordings, also known as fiducial points or fiducials, multiple recognizer inputs can be retrieved and exploited. The peaks, limits, slopes, and other wave structures function as fiducials. Adaptive thresholds can be attained with detectors. The retrieval precision has a substantial effect on the Fourier synthesis wavelet transform properties [9]. In addition, Figure 2 depicts the ECG taxonomy.

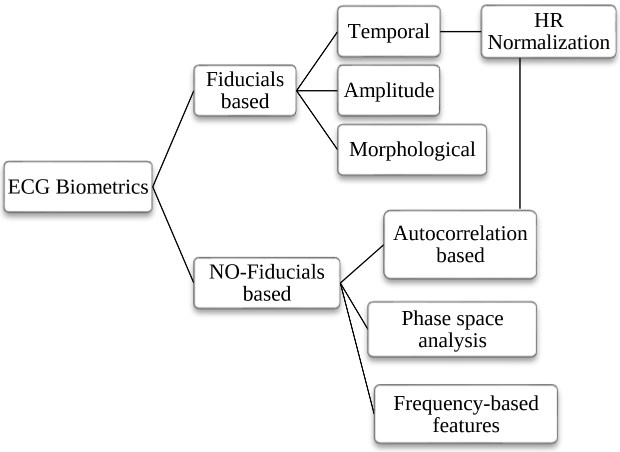


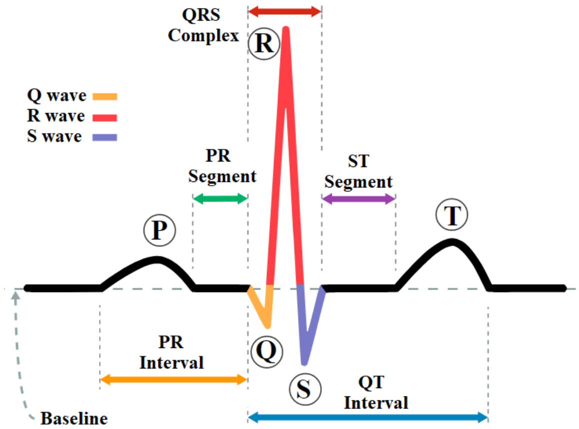
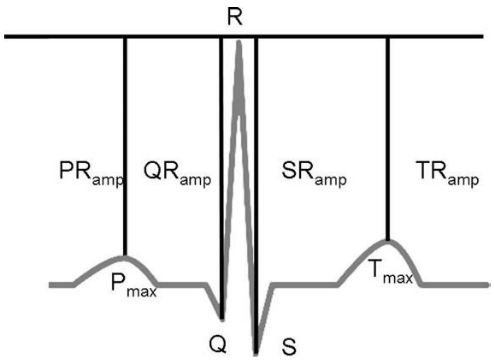
Figure 2. Taxonomy of ECG [10]

Existing ECG-based recognition systems offer both benefits and drawbacks. Inter-subject variability is applied to describe an individual based on ECG values (features). Features are generated using the geometry of the heartbeat, unique periods provided by ECG waves, and precisely extracted features. The employed features vary according to the complexity of the recognizer, the need for real-time identification, the type of recording equipment, etc. Regarding the suitable evaluation method and number/type of attributes, there is no consensus. In addition, ECG analysis is usually performed on proprietary datasets, making it impossible to compare methods. And Utilizing strategies based on fiducials It is possible to extract and use numerous features as recognizer inputs by recognizing certain anchor points, also known as fiducial points or fiducials, on ECG recordings. The peaks, boundaries, and slopes of a wave function serve as fiducials or reference points. Through detectors, adaptive thresholds can be created. Significantly influenced by retrieval precision are the Fourier synthesis wavelet transform features [11].

4. ECG LEADS

It is easier to interpret an ECG if you keep in mind the different perspectives each lead uses to assess the heart. A vertical picture of the heart is obtained from the signals of six "conventional" leads, which are electrodes placed on the extremities (i.e., from the sides of the feet). The inferior area of the heart is evaluated with I and VF, whereas the left lateral area is assessed with II and VL. Using lead VR and the left leg, the right atrium is examined via the right arm and the right leg. The six V leads (V1–V6) provide a horizontal front and left view of the heart. So, V1 and V2 stand for the right side of the heart, V3 and V4 represent the inter-ventricle septum and the anterior walls of the left ventricle, and V5 and V6 stand for the anterolateral walls of the left ventricle.

Table 2. Type of features [4]

Feature Title	Meaning	explanations
Temporal	Temporal correlates between P, QRS, and T waves in an electrocardiogram (ECG) serve as biometrics discriminators, showing electromagnetic pathways from the cross - strait nodes to the Purkinje system during times of cardiac stimulation. Time intervals between events and the length of pulse waves (P, QRS, and T) are the most often used temporal features (PQ, RS, ST, etc.)	
Amplitude	Individual variation in the amplitude of pulse waves is readily observed. The amplitude characteristics of an ECG wave, often evaluated at the R peak, define the comparative amplitude between the peaks of the wave. The ratios of ST segmentation intensity to the peaks of the first and second derivatives are one feature of heart rate amplitude	
Morphological	The overall or interval structure of an ECG may be gleaned from its morphological properties (P-QRS-T). In order to extract morphological components from a pulse, the easiest way is to average the measured values of many synchronized (i.e., R peak) heart rate	$FF=\sqrt{\frac{Var\left(X''\right)/VAR(X)}{Var\left(x'\right)/Var(X)}}$ where: X= QRS complex waveform X'= is the first derivative of the QRS X''= is its second derivative
Autocorrelation based	It draws attention to irregularities in the data and is shift-invariant. Numerous degrees of shape and duration invariance are preserved in the QRS complex. This approach enables the detection of fiducials in samples where their detection would be hampered otherwise	$r[m]=\frac{1}{r\|0\|}\Sigma s[i]s[i+m]$ Where, $r[m]$ = the AC $s[i]$ = the signal at time i and m is chosen greater than the mean QRS duration (in samples)
Phase space analysis	The time-delay technique may be used to evaluate a three-dimensional ECG signal and represent it in two dimensions. Phase-space trajectories and single-lead ECG amplitude normalization vectors may shed light on previously unsuspected aspects of cardiac activity and broaden the range of observable traits	$(s(t), s(t+dt), s(t+2dt))$ Dividing the phase-space into a 30 30 30 grid converts the cross trajectories into a sized particles features space, reducing computation complexity and loop variability caused by noise or ECG irregularities
Frequency-based	We simulated the frequency composition of the ECG data using a linear prediction technique (LPC). Using a least-squares regression method, we combine the first 40 elements of the linear restorations of the ECG spectra with the first 40 points of the ECG spectra to generate a spectrum model for each subject	$x\hat{}[n]=-\sum_{(i=1)}^{p} aix[n-1]$ where error minimization is used to rank the ai coefficients. $e[n]$ where: $e[n]$= $x[n]$-$x\hat{}[n]$ where: $x[n]$ =the actual value

The initials aVR, aVL, and aVF refer to improved leads. The electrical activity between a single electrode and a single leg is evaluated. Lead aVR provides an inaccurate view of the heart. The aVL lead displays electrical activity generated by the lateral heart wall. The aVF lead reveals electrical activity emanating from the inferior wall of the heart. Reduced-lead 12-lead ECG devices are derived from conventional 12-lead ECGs. Different from the standard 12-lead ECG setup, these devices may employ either a reduced-lead set or individualized leads [12]. Choose the leads that will provide the most "data" about the cardiac electrical activity. Based on the information in Table 3, a conventional 12-lead ECG will include three limb inputs (leads I, II, and III), three enhanced limb leads (leads aVR, aVL, and aVF), and six responsible for a large portion lead (V1-V6).

Table 3. Summary of 12-Leads Position [13]

No.	Leads	Position
1	I	From the right arm to the left
2	II	From the right arm to the left leg
3	III	Between the left arm and leg on the left side
4	aVR	right arm
5	aVL	left arm
6	aVF	left leg
7	V1	Right fourth intercostal gap from the sternum
8	V2	Position yourself so that the left side of your chest, between both the fourth and fifth ribs, is exposed
9	V3	involving V2 and V4 conductors.
10	V4	Fifth intercostal space along the line of the midclavicular
11	V5	Aligned with the left axillary line and parallel to V4
12	V6	Right beside V5 on the midaxillary meridian. At the exact center of your lower armpit.

5. CONTIGUITY OF LEADS

In a standard 12-lead configuration, successive leads of the same color are consecutive. Every one of the 12 lines on an electrocardiogram (ECG) records the electrical activity of the heart from a distinct vantage point and corresponds to a distinct anatomical area of the heart [14]. The Table 4 depicts the distribution of leads.

Table 4. Contiguity of leads [15]

I Lateral	aVR	V1 Septal	V4 Anterior
II Inferior	aVL Lateral	V2 Septal	V5 Lateral
III Inferior	aVF Inferior	V3 Anterior	V6 Lateral

6. ACQUIRED ECG QUALITY CRITERIA

As the principal method for heart research and the clinical diagnosis of cardiovascular disorders, the ECG signal is a critical human physiological signal that contains information from the human cardiac conduction system. The electrical impulses generated by the heart's myocardium serve as a variable voltage source during the cardiac cycle. Peaks in the ECG wave are identified with the skin potentials in the chest (P, Q, R, S, as well as T) that are generated by this process. This signal is applied to diagnose a range of cardiac function and location abnormalities. A wave's amplitudes, periods, and polarity are analyzed to formulate a diagnosis. A transducer recognizes this bioelectric signal from the patient's body and translates the flow of ionic current in the body into the flow of electronic currency in the lead wires.

For this experiment, Eithoven's three bipolar limb lead system was employed. I stand for the left and right arms, II for the left and right legs, and III for the left and right legs [16]. A lead is a differential recording between the electrodes of a biopotential device. The right leg acts as a reference point for the ground. The captured signal is analyzed to extract frequency (heart rate) and various time intervals and amplitudes. A comparison between the extracted and typical values for a healthy individual will facilitate the development of a medical opinion. The ECG signal is an AC signal with an amplitude of around 2.5 mV. The objective of the hardware module is to amplify and digitize the signal. The signal is transmitted serially to the computer for display and processing [17].

7. BEST POPULAR DATA SETS

There are five essential databases commonly used to evaluate electrocardiogram-related activities, with the bulk of databases originating from medical devices and a small number from healthcare devices. Healthcare device data often contains fewer leads than medical device data, suggesting that healthcare device data is typically less informative [8]. In comparison, medical device data is harder to acquire. The prevalence and quantity of healthcare ECG monitors, such as intelligent gadgets and wristbands, are increasing Traditional 12- or 15-lead ECG devices can detect more cardiac irregularities than a single-lead ECG. Explanation of the most frequent datasets presented in Table 5.

Table 5. Description of Notable Statistics

RE	Title of data	Description of information
Goldberger, et al., 2000 [20]	The PTB Diagnostic ECG	This collection contains 549 ECG data with 15 channels from 290 individuals, spread across 549 records The sampling rate can exceed 10 kHz at most. There are 216 individuals with one of eight forms of cardiac disease, 52 individuals in good health, and 22 individuals with unknown conditions
Goldberger, et al., 2000 [20]	PTB XL	The collection includes 21837 clinical 12-lead ECGs with a 10-second duration from 18885 individuals. Two cardiologists could have interpreted the raw waveform data and assigned several ECG statements per record. The 71 distinct ECG statements that adhere to the SCP-ECG standard include diagnostic, form, and rhythm statements. Training and test set partitioning are advised to verify the comparability of machine learning algorithms learnt on a dataset. In conjunction with the detailed annotation, this transforms the dataset into a valuable resource for constructing and evaluating ECG interpretation algorithms. The dataset includes rich metadata concerning demographics, infarct features, the probability for diagnostic ECG statements, and signal quality.
Moody and Mark, 2001 [18]	MIT-BIH Arrhythmia Database	Data from 48 30-minute ECG recordings from 47 individuals at Boston's Beth Israel Hospital (now the "Beth Israel Deaconess Medical Center"). Each Automatic data series samples at 360 Hz and offers 11-bit accuracy across a 10-mV range. There are several diagnoses at the metric level and beyond in this dataset.

Clifford, et al., 2017 [19]	The PhysioNet Computing in Cardiology	Includes 8,528 anonymized ECG records from AliveCor medical devices, recorded at 300 Hz for 9-60 seconds. Five thousand one hundred fifty-four of the tapes are standard. There are an additional 2,557 recordings in addition to the 717 AF and 717 AF recordings. Forty-six audio files are noisy. Additionally, the evaluation of 3,658 test recordings is kept private. These details have been gathered. from medical equipment
Brito, et al., 2019 [21]	The MIT-BIH Atrial Fibrillation Database	consist of 25 10-hour, 2-lead, long-term ECG recordings at 250 hertz for persons with atrial fibrillation (mostly paroxysmal). Beth Israel Hospital in Boston used the 2017 PhysioNet Computing in Cardiology Challenge data set to conduct an in-depth study of deep learning techniques. Only strategies having a one-score greater than 0.80 were considered for inclusion in the list of covert methods.

8. NOISE IN ECG AND PROCESSING

Electrocardiogram ECG signals are an electrical representation of heart function produced by a transducer component or device that detects and converts mechanical energy (vibration) into an electrical signal for filtration, processing, and additional analysis to determine the patient's health condition. Through the movement of a diaphragm placed on a patient's skin as part of a balanced bridge arrangement, the equipment captures voltage variations generated by the diaphragm's detection of the heart's movements. ECG signals are essentially tied to multiple types of noise [22]. These sounds are classified as shown in Figure 3. Table 6 showing each type of noise, what causes it and how it occurs.

ECG Noise Types	Baseline wanders (low-frequency range)
	Power line interference (Medium frequency range)
	Electromyogeaphy noise (high-frequency range)
	Burst noise
	Electrode contact noise
	Muscle contraction
	Motion artifacts

Figure 3. Summary of Noise Type [23]

Table 6. Most Common Types of Noise Prevalent in ECG Signals [22]

Type	Reasons	Maximum amplitude and time	Rainbow range	Belongings
baseline wander	Breathing, physical movement, poor electrode contact, and skin-electrode impedance can alter ECG readings	Depend on electrode arrangement, electrolyte properties, skin impedance, and subject motion	range between 0.05 and 1 Hz	segment and other LF ECG signal components are altered
power-line interference	The ECG signal acquisition circuit is linked inductively and capacitively to ubiquitous power lines	50 percent of the ECG signal's peak-to-peak amplitude	50/60 Hz-centered, 1 Hz-broadband, narrowband noise	Modifies the amplitude, duration, and form of low-amplitude ECG local waves
muscle artefacts	Electrical activity of the muscle during contraction or in response to a rapid body movement	10 per cent of the ECG signal's peak-to-peak amplitude	bandwidth ranges between 20 and 1000 Hz	modify the local waveforms of the ECG signal

9. LEARNING TECHNIQUES

AI is the reproduction of human intelligence in robots trained to mimic human cognition and behavior. AI systems typically function by ingesting enormous quantities of training data, analyzing the data for correlations and patterns, and employing these patterns to predict future states. The rapid progress of artificial intelligence technology is primarily attributable to the fact that AI processes massive volumes of data substantially faster than humans and provides more accurate forecasts. Superior artificial intelligence would be capable of thinking and acting with the greatest likelihood of achieving a predefined goal. Machine learning, a subfield of artificial intellectual ability, is the practice of teaching computers how to use and learn from fresh data with little to no human interaction. This independent learning is enabled by absorbing vast volumes of unstructured data, including text, images, and videos. Artificial intelligence is often associated with machine learning techniques. These disciplines develop expert systems that can anticipate outcomes or categorize information with the use of AI-based algorithms. With amplitudes and durations measured in millivolts and milliseconds, the interpretation of ECG signals may be subject to substantial inter-and intra-observer variation. By utilizing machine learning approaches, automated diagnostic systems can avoid these limits [16].

- Machine Learning

Segmentation, clustering algorithms, clustering algorithms, time - series data identification, regress, and forecasting are all examples of data mining's specialized methods. According to [11], there are three primary groups of machine learning algorithms (Figure 4):

- Supervised Learning: These models require individual training for the data analyst to contribute input data and anticipate results during the activity of a particular machine learning model. To create predictions, the data analyst will identify the variables or characteristics that the model will use. The model may be able to predict new data based on what it has learned during training. The model is displayed throughout training, and new data can be forecasted after training.

The output of classification algorithms is restricted to specified values, such as classifying emails as "spam" or "not spam." Consequently, the results of a regression method are continuous and may lie within a range, such as a temperature or a commodity's price.

• Unsupervised Learning: These models conclude from incoming data without tagged responses. This technique is used to identify patterns and other data structures, including data point clustering, which is used to identify data groupings or hidden patterns.

• Semi-supervised learning: This machine learning combines the two strategies described above. Even if data scientists provide a model with heavily annotated training data, it is permitted to explore the data and develop their understanding of the set.

Cluster analysis, the most common technique for unsupervised learning, is applied to exploratory data to find hidden patterns or group them into data clusters. The similarity is measured using Euclidean and probabilistic metrics to determine the clusters.

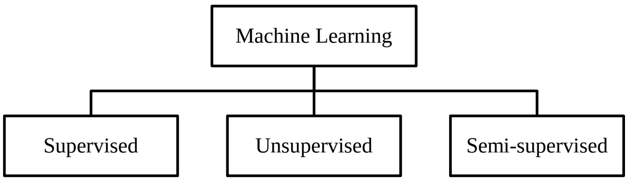


Figure 4. Machine Learning Models type

• Deep learning: Deep learning requires neural networks consisting of at least three layers. A neural network is a computer program trained to learn from data in the same manner that the human brain does, by discovering patterns and drawing conclusions.

Approximations may be obtained using a neural network with just one hidden layer, but the accuracy of the network can be improved by adding more hidden layers. Deep learning is used in a wide range of AI applications and services because it improves automation by allowing analytical and physical activities to be carried out without the need for human interaction [30]. With the use of data inputs, weight, and prejudice, data augmentation neural networks (also known as artificial networks) mimic the human brain. These subsystems can accurately identify, sort, and describe collections of data items. In deep neural networks, multiple layers of interconnected nodes optimize the prediction or classification.

Forward propagation is the name for this mathematical process. Typically, just the input and output layers of a deep training algorithm are shown to the user. A deep learning model's input layer receives data for processing, while the model's output layer generates a prediction or classification based on the processed data. Backpropagation is a process that utilizes gradient descent to measure errors in predictions before altering the weights of a function by traveling backwards through the layers to train the model. Backpropagation and forward propagation enable neural networks to make predictions and fix errors. The algorithm's accuracy continually improves over time. There are a variety of neural network models to address specific problems or datasets, and deep learning techniques are quite advanced [24]. Figure 5 provides an overview of the most prevalent method used in the ECG position, which comprises deep and automated learning techniques:

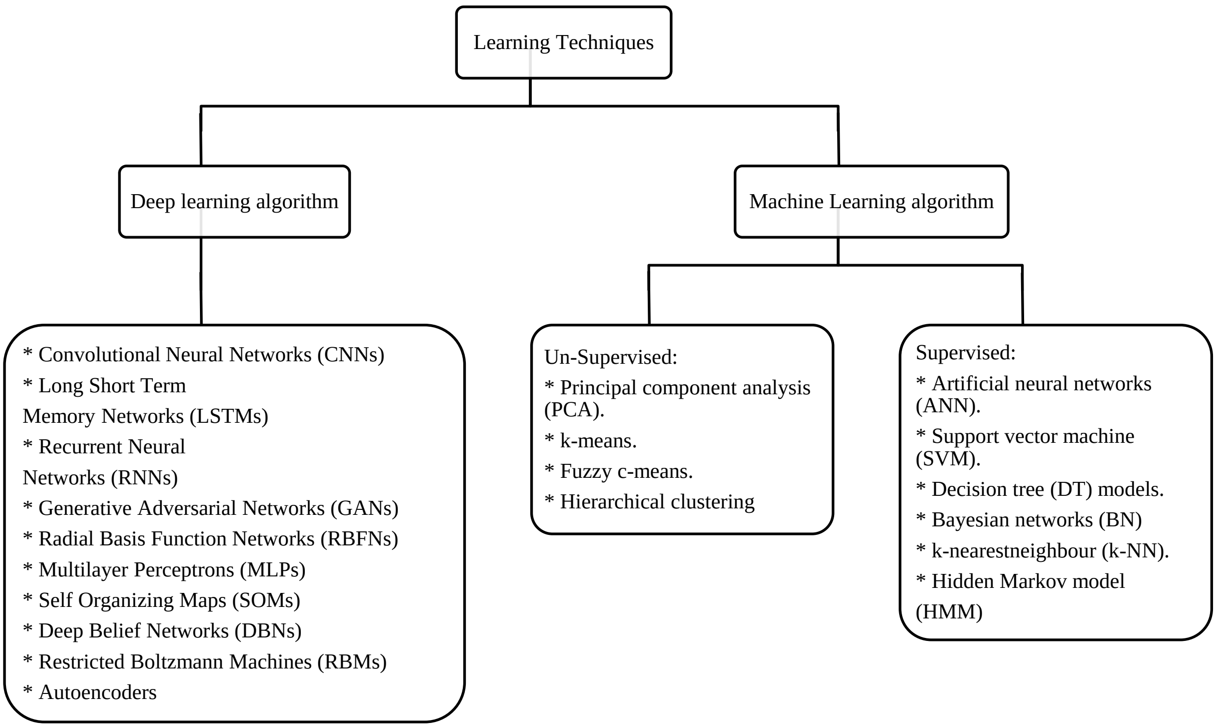


Figure 5. Summarized Method [16]

Table 7 shows some of the previous studies related to deep learning, which were listed analytically in terms of the field used, the data used, the algorithms, and the results it reached to know the importance of deep learning in the subject of electrocardiography.

Table 7. A review of a group of studies specialized in deep learning

Year and RF.	Application	Algorithm	Database	Results
Herry, et al., 2017 [25]	Heartbeat classification	synchro squeezing transform (SST) support vector machine (SVM) classifier	(MIT-BIH) arrhythmia database	Sensitivity, specificity and accuracy were 92.12%, 86.69% and 87.29%, respectively
Mathews, et al., 2018 [26]	Heartbeat classification	Boltzmann constrained machine Deep learning using deep belief networks (DBN)	MIT-BIH Arrhythmia Database	SVEB are: accuracy =93.78%, Sensitivity =88.39% respectively=6.68%, VEB are: accuracy =96.94%, Sensitivity =85.22% respectively=4.11%,
Ortin, et al., 2019 [27]	ventricular heartbeat classifier	Echo State Network (ESN)	MIT-BIH arrhythmia database INCART database	MIT-BIH arrhythmia database sensitivity and precision: modified lead II: 95.3 and 88.8 V1 lead: 90.9 and 89.2
Kamiri and Mariga, 2021 [28]	Premature Ventricular Contractions [PVCs] beat classification	discrete wavelet transforms (DWT) KNN and SVM algorithms	MIT-BIH Arrhythmia Database	99.75% accuracy
Houssein, et al., 2021 [1]	ECG beat classification	two-event related moving-averages (TERMA) algorithm fractional-Fourier-transform (FrFT) algorithm	MIT-BIH Database St. Petersburg INCART Shaoxing People's Hospital (SPH) database	99.85% for INCART 68% for SPH

Table 8. Previous studies

Survey	Year	ECG Lead	Learning Techniques	Noise	Data	Features	Method
Berkaya, et al., 2018 [11]	2018	✓	×	×	×	×	×
Thiyagarajan and Chakravarthy, 2019 [12]	2019	×	✓	×	×	×	✓
Uwaechia and Ramli, 2021 [2]	2021	✓	×	✓	×	×	✓
Shabaan, et al., 2020 [16]	2020	✓	×	×	×	✓	✓
Merdjanovska and Rashkovska, 2022 [4]	2022	×	✓	×	✓	×	✓
Bhirud and Pachghare, 2020 [29]	2022	×	✓	×	×	✓	✓
This survey	2022	✓	✓	✓	✓	✓	✓

10. COMPARISON OF SURVEY STUDY

Numerous studies have been undertaken on the issue of ECG; in the Table 8 are some of the previous studies of type (survey) on the subject of ECG.

11. CONCLUSION

The electrocardiogram (ECG) has garnered substantial attention as a prospective biometric feature and is a crucial diagnostic tool for assessing a patient's cardiac state. The ECG possesses strong discriminatory properties in the field of biometrics. Considering the frequency with which ECGs are obtained, automated ECG interpretation algorithms for diagnostic support systems hold considerable promise for medical staff. The idea of ECG has been questioned from a medical standpoint; what is its significance, and how does its initiation play a vital part in early disease diagnosis? ECG has advantages that enable specialists to diagnose the disease, and the essential methods used by previous researchers were reviewed. This paper is an analytical review and summary of all the cardiogram data, a comparison between the previous survey pipes, and a review of the points raised in them.

REFERENCES

[1] M. Shabaan, et al., "Survey: Smartphone-Based Assessment of Cardiovascular Diseases using ECG and PPG Analysis", BMC Med. Inform. Decis. Mak., Vol. 20, No. 1, pp. 1-16, 2020.

[2] A.N. Uwaechia, D.A. Ramli, "A Comprehensive Survey on ECG Signals as New Biometric Modality for Human Authentication: Recent Advances and Future Challenges", IEEE Access, 2021.

[3] H.M. Lynn, S. Yeom, P. Kim, "ECG-Based Biometric Human Identification Based on Backpropagation Neural Network", The 2018 Conference on Research in Adaptive and Convergent Systems, pp. 6-10, 2018.

[4] E. Merdjanovska, A. Rashkovska, "Comprehensive Survey of computational ECG analysis: Databases, Methods and Applications", Expert Syst. Appl., p. 117206, 2022.

[5] S. Aziz, M.U. Khan, Z.A. Choudhry, A. Aymin, A. Usman, "ECG-based Biometric Authentication Using Empirical Mode Decomposition and Support Vector Machines", The 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 906-912, 2019.

[6] M.N. Teferra, et al., "Electronic Textile Electrocardiogram Monitoring in Cardiac Patients: A Scoping Review Protocol", JBI Evid. Synth., Vol. 17, No. 2, pp. 147-156, 2019.

[7] H. Li, P. Boulanger, "A Survey of Heart Anomaly Detection Using Ambulatory Electrocardiogram (ECG)", Sensors, Vol. 20, No. 5, p. 1461, 2020.

[8] H. Li, P. Boulanger, "A Survey of Heart Anomaly Detection Using Ambulatory Electrocardiogram (ECG)", Sensors, Vol. 20, No. 5, p. 1461, 2020.
[9] K.K. Tseng, et al., "Human Identification with Electrocardiogram", Enterp. Inf. Syst., Vol. 12, No. 7, pp. 798-819, 2018.
[10] S. Hong, Y. Zhou, J. Shang, C. Xiao, J. Sun, "Opportunities and Challenges of Deep Learning Methods for Electrocardiogram data: A Systematic Review", Comput. Biol. Med., Vol. 122, p. 103801, 2020.
[11] S.K. Berkaya, A.K. Uysal, E.S. Gunal, S. Ergin, S. Gunal, M.B. Gulmezoglu, "A Survey on ECG Analysis", Biomed. Signal Process. Control, Vol. 43, pp. 216-235, 2018.
[12] S. Thiyagarajan, T. Chakravarthy, "A Survey on ECG Data Classification Using Machine Learning Approaches", Adalya Journal, Vol. 8, No. 8, pp. 401-408, 2019.
[13] H.H.M. Draisma, et al., "LEADS: An Interactive Research Oriented ECG/VCG Analysis System", Computers in Cardiology, pp. 515-518, 2015.
[14] P.H.B. de Melo, "A Brief Review on Electrocardiogram Analysis and Classification Techniques with Machine Learning Approaches", U. Porto J. Eng., Vol. 7, No. 4, pp. 153-162, 2021.
[15] A. Perron, T. Lim, U. Pahlm Webb, G.S. Wagner, O. Pahlm, "Maximal Increase in Sensitivity with Minimal Loss of Specificity for Diagnosis of Acute Coronary Occlusion Achieved by Sequentially adding Leads from the 24-Lead Electrocardiogram to the Orderly Sequenced 12-Lead Electrocardiogram", J. Electrocardiol., Vol. 40, No. 6, pp. 463-469, 2007.
[16] M. Zile, "Smart Energy Management in Solar/Wind Power Stations using Artificial Neural Networks", International Journal on Technical and Physical Problems on Engineering (IJTPE), Issue 48, Vol. 13, No. 3, pp. 74-79, September 2021.
[17] J.R. Pinto, J.S. Cardoso, A. Lourenco, C. Carreiras, "Towards a Continuous Biometric System Based on ECG Signals Acquired on the Steering Wheel", Sensors, Vol. 17, No. 10, p. 2228, 2017.
[18] G.B. Moody, R.G. Mark, "The Impact of the MIT-BIH Arrhythmia Database", IEEE Eng. Med. Biol. Mag., Vol. 20, No. 3, pp. 45-50, 2001.
[19] G.D. Clifford, et al., "AF Classification from a Short Single Lead ECG Recording: The PhysioNet/Computing in Cardiology Challenge 2017", Computing in Cardiology (CinC), pp. 1-4, 2017.
[20] A.L. Goldberger, et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", Circulation, Vol. 101, No. 23, pp. e215-e220, 2000.
[21] C. Brito, A. Machado, A.L. Sousa, "Electrocardiogram Beat-Classification Based on a ResNet Network", MedInfo, pp. 55-59, 2019.
[22] S. Chatterjee, R.S. Thakur, R.N. Yadav, L. Gupta, D.K. Raghuvanshi, "Review of Noise Removal Techniques in ECG Signals", IET Signal Process, Vol. 14, No. 9, pp. 569-590, 2020.
[23] P. Bing, W. Liu, Z. Wang, Z. Zhang, "Noise Reduction in ECG Signal Using an Effective Hybrid Scheme", IEEE Access, Vol. 8, pp. 160790-160801, 2020.
[24] X. Liu, H. Wang, Z. Li, L. Qin, "Deep Learning in ECG Diagnosis: A Review", Knowledge-Based Syst., Vol. 227, p. 107187, 2021.
[25] C.L. Herry, M. Frasch, A.J.E. Seely, H. Wu, "Heart Beat Classification from Single-Lead ECG Using the Synchro Squeezing Transform", Physiol. Meas., Vol. 38, No. 2, p. 171, 2017.
[26] S.M. Mathews, C. Kambhamettu, K.E. Barner, "A Novel Application of Deep Learning for Single-Lead ECG Classification", Comput. Biol. Med., Vol. 99, pp. 53-62, 2018.
[27] S. Ortin, M.C. Soriano, M. Alfaras, C.R. Mirasso, "Automated Real-Time Method for Ventricular Heartbeat Classification", Comput. Methods Programs Biomed., Vol. 169, pp. 1-8, 2019.
[28] J. Kamiri, G. Mariga, "Research Methods in Machine Learning: A Content Analysis", Int. J. Comput. Inf. Technol., Vol. 10, No. 2, pp. 78-91, 2021.
[29] B. Bhirud, V.K. Pachghare, "Arrhythmia Detection using ECG Signal: A Survey", International Conference on Computational Science and Applications, pp. 329-341, 2020.
[30] M. Mbida, A. Ezzati, "Artificial Intelligence Auscultation System for Physiological Diseases", International Journal on Technical and Physical Problems on Engineering (IJTPE), Issue 49, Vol. 13, No. 4, pp. 97-103, December 2021.

BIOGRAPHIES

Wardah Rajab Mohammed Al-Rawi was born in Al-Anbar, Iraq on March 1, 1990. She received the B.Sc. degree in Biomedical Engineering from Al-Nahrain University, Baghdad, Iraq in 2013. Now, she is pursuing the M.Sc. in Research Phase in Biomedical Engineering at the same university. Her research interests are deep learning applications of biomedical engineering.

Ahmed Faeq Hussein was born in Baghdad, Iraq in June 1976. He received the B.Sc. degree in Electrical Engineering from Al-Mustansiriyah University, Baghdad, Iraq in 1998, and the M.Sc. degree in Computer Engineering from University of Technology, Baghdad, Iraq, in 2004, and the Ph.D. degree in Computer and Embedded System Engineering from University of Putra Malaysia, Seri Kembangan, Malaysia in 2018. He was a Senior Engineer with Medical Department, Ministry of Health, Iraq, until 2009. He has been an Assoc. Prof. with the Bio-Medical Engineering Department, Al-Nahrain University, Baghdad, Iraq. His research interests include bio-medical signal processing, low-energy Bluetooth communication, and cloud-based applications.