Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 04 Issue 02
2023

Review Article
https://doi.org/10.56979/402/2023

ECG Models Predict AFib From 2012-2022: A Systematic Literature Review

Abdul Majid Soomro[3]*, Awad Bin Naeem[1], Fridous Ayub[5], Biswaranjan Senapati[2], Ojas Prakash-bhai Doshi[4], and Nimra Bari[5]

[1]Department of Computer Science, National College of Business Administration & Economics, Multan, Pakistan.
[2]Department of Computer Science and Data Science, Parker Hannifin Corp, USA.
[3]Department of Computer Science (FSKTM), University Tun Hussein Onn Malaysia, Malaysia.
[4]Department of Pharmaceutical Sciences, Arnold & Marie Schwartz College of Pharmacy and Health Sciences Brooklyn, NY, USA.
[5]Department of Computer Science, Women University Swabi, Pakistan.
*Corresponding Author: Abdul Majid Soomro. Email: gi180004@siswa.uthm.edu.my

Received: August 08, 2022 **Accepted:** March 05, 2023 **Published:** March 29, 2023.

Abstract: The most prevalent arrhythmia in the world is atrial fibrillation (AFib). Every year, 4.7 million individuals are diagnosed with atrial fibrillation, and it affects more than 33 million people globally. While symptoms vary by individual, the most prevalent ones are a fast heartbeat and chest discomfort. The heart rate might reach 101-175 beats per minute when atrial fibrillation develops, although the usual heart rate is 60-100 beats per minute. There are four kinds, two of which are difficult to identify using normal procedures such as an EKG (ECG). Nevertheless, as smart wearable devices have grown more commonplace, there are various techniques to identify and forecast the beginning of AF using merely ECG tests, making physicians' diagnoses simpler. By searching several databases, this study reviewed articles published in the past decade (2012 to 2021), focusing on patients who used DL(DL) for AF prediction research. The results showed that only 23 studies were selected as systematic reviews, of which 4 applied Artificial Intelligence techniques (21%), 12 of which used DL methods (52%), and the other 7 focused on the application of the general Machine Learning Model (36%). All in all, this study shows that in the context of AI, AF prediction is still an untapped field and deep learning techniques are improving accuracy, but these applications are not as frequent as expected. In addition, since 2016, more than half of the selected studies have been published, which confirms that the topic is recent and has great potential for further research.

Keywords: Electrocardiogram waveform; Electrocardiogram (ECG); Deep Learning model; classification algorithms; Atrial Fibrillation (AFib).

1. Introduction

In clinical practice, the most common type of persistent arrhythmia is atrial fibrillation (AFib)[1]. The atrium loses its normal and effective contraction function and enters a condition of rapid and disorganised fibrillation, with a frequency of 300-600 beats per minute, while the ventricle beats swiftly and erratically, with a frequency of 100-200 beats per minute[2]. The real incidence of atrial fibrillation may be greater since some individuals were not recognised and treated because they did not exhibit clear symptoms. If not treated effectively, atrial fibrillation may lead to major consequences such as blood clots, stroke, heart failure, and other heart-related issues, including death[3]. Only the most severe Long-term Persistent and Permanent may be readily diagnosed with an ECG check, according to, with the other categories being more difficult to spot owing to the irregularity of their symptoms[4]. Paroxysmal atrial fibrillation, persistent atrial fibrillation, and long-standing atrial fibrillation Atrial fibrillation are classified into three types: long-standing, persistent, and permanent. These three types (long-standing, persistent, and permanent) may be readily identified with an electrocardiogram test. Because of the irregularity of its symptoms, paroxysmal atrial fibrillation is very difficult to diagnose[5]. As a result, many prediction algorithms for the identification of AF were created, and these models enable the diagnosis of a patient's AF status based just on a brief ECG signal, eliminating extra-long and invasive equipment and procedures[6].

Nowadays, there are several smart intelligent gadgets are available that revolutionize the diagnosing procedure and reduce the pain of patients while their results are too quick and considerable[7]. Using these smart and intelligent gadgets, patients' ECG and pulse signals have been collected at the time of performing daily routine activities and their emotional state. Although, In recent years, different organizations and well-known professionals have developed different advanced techniques for the detection of various kinds of AFib[8]. Most of these are integrated with sensors and effective Deep Learning algorithms and automated the process of diagnosis in real-time at any place at any time.

This study conducts a systematic literature review on ECG models to predict AFib with the empowerment of DeepLearning methods during the time frame of 2012 to 2021. Previously, We did not find any organized study that cover the same topic within the same time frame as we mentioned. From the several selected studies, we identified the three major parameters as input for this:

1. For indicating which features databases, pre-processing, and prediction methods have been and are currently applied in these systems.
2. This study also contains a collection of significant methods which perform better in the classification of AFib.
3. This study consists of a comprehensive discussion on the prediction of AFib that has been produced and its current challenges.

A detailed discussion of this introduction is explained in Part I sections; Part II contains a descriptive overview of the research methodology that defines the eligibility selection for the extraction of knowledge. Part III is the core of this study which describes the experimental results of SLR by representing the chosen articles and their features in summary tables while Part IV is covering discussions and in terms of the answer to the research questions. In the last section Part V, we conclude this paper with its major findings including highlights and limitations.

2. Research Methodology

This SLR is carried out under the recommendations of the Cochrane Manual for Interventions of a systematic review. The recommended resulting items for (PRISMA) description consist of the different guidelines collected. This section includes the descriptive detail of the methodology to conduct this SLR.

2.1 Research Strategy

The following databases Pub Med, Springer Links, Ovid-Medline, Taylor, Francis, IEEE Explore, Science Direct, Scopus, Web of Science, and ACM Digital Library were approaches for searching relevant publications peer-reviewed that started from 1st January 2012 to September 2021[9]. From the above-mentioned databases, two interdisciplinary databases are Scopus and Web of Science, while ACM Digital library was according to, the predominant database that is relevant to academic databases of Computer science, as well as IEEE Xplore that was selected for its greater number of studies, took from the computer science field. PubMed is used as it has provided the most relevant content[10].

The articles are being searched based on different keywords:

(DLOR Machine Learning OR Transfer Learning OR Artificial Intelligence) AND (ECG OR Electrocardiogram OR electro cardio) AND (AFib OR arrhythmia) AND (classification OR prediction OR prognosis OR foresee).

2.2 Study Design

The purpose of this SLR is to classify and present a systematic study on ECG-based models for AFib Prediction applying DL techniques we are covering in the previous ten years from 2012 to 2021. By using the online tool Rayyan QCRI, we assessed the titles and abstracts of all recognized publications for appropriacy. For increasing the search sensitivity, the insertion criteria are highly explained. The basic goal is to find publications that used any DL algorithm with ECG data to predict AFib in individuals who had never been treated before. Studies that used approaches to identify the existence of an actual AFib episode in a patient are excluded from consideration because the main focus of this review is on predicting AFib rather than detecting it, that was, predicting AFib onset before it happens rather than detecting it. Table 1 summarizes further inclusion and exclusion criteria. All of the publications that are not excluded following the analysis have examined the complete texts for eligibility by using the inclusion and exclusion criteria given in Table 1.

Table 1. Inclusion and exclusion benchmarks applied in this SLR as explained

Type	Inclusion	Exclusion
Publishing[11]	Conference and journal	All except Conference and journal
Study design	All	Non
Setting[12]	All	Non
Reported Results[13]	However one: efficiency sensitivity specificity confusion matrix	All others that did not result in any metric
Nobel review[14]	Journal and Conference	All except Conference and journal
Participators[15]	With no current medical processes or drugs results between the ECG accumulation	With any current surgical processes and ingestion or drugs effects between the ECG collective data
Language[16]	English	All other
The geographic location of studies[17]	All	Non
Exposure of interest	All	Non
Data[18]	All	Non

2.3 Data extraction

The following data is retrieved from the chosen articles using pre-defined categories to gather relevant data for assessing, analyzing, and evaluating the model's attributes and experimental setups:

1. Study Knowledge: This setup is identifying the citation of studies and publication year.
2. Data Insertions: This setup Examines the inputs used to create the algorithm, such as the dataset, the total number of participants, and the participant's ages in the dataset.
3. Signal treatment: It specifies the characteristics retrieved from the ECG signals received as input, as well as the signal duration, applied for training and the tools applied to the process.

4. Methodologies: This setup is specifying the algorithms used for ECG signal pre-processing also AFib prediction as well as model assessment, the total amount of iterations, and during the training and testing phase separation of data.
5. Performance: it specifies the evaluation metrics applied for prediction assessment.

2.4 Research questions

Here is a list of the research questions of this study.

- (RQ1) How do the research studies address classification problems for AFib?
- (RQ2) What sort of datasets and feature extraction techniques are used?
- (RQ3) Which preprocessing techniques are used?
- (RQ4) What type of Deep Learning models are applied for the prediction of AFib?
- (RQ5) Which metrics are mainly applied to measure the accuracy of Deep Learning algorithms?

In Table 2 we are presenting the above five RQs with their corresponding motivations. This step is classifying the tendency and possible convenience for the use of research methods in this field.

Table 2. The following table is describing the research questions.

RQ NO	Motivations
RQ1	For classifying the tendency and achievable chances to focus on the research topic.
RQ2,3	For identifying databases, preprocessing approaches and new advancements in features applied to predict AFib, are common.
RQ4	To find the new prediction techniques that are applied for predicting AFib with ECG data based on current articles.
RQ5	We proposed it to classify the methods that might predict AFib episodes very accurately.

3. Experimental Results

After deleting duplicates, our investigation yielded 382 unique entries in the beginning. After examining the title and abstract, 293 items are deleted using the inclusion and exclusion criterion shown in Table 1, where 84 full-text articles are examined for eligibility and 72 records are discarded following full-text revision. The following is a list of the records that are not included in this SLR. Sixty-four articles reported AFib research; however, no AF prediction is made at the time of execution. Two articles cannot be fully studied because the authors of the purposed SLR are unable to access the full paper. Two publications do not describe the assessment metrics as shown in Table 1 inclusion criteria. Two articles are concentrated on reviewing the art state in the detection of AFib. One research is published before 2009 while another used ECG data obtained from patients undergoing surgery (prophylactic ICD execution). Pursuing reference is done on the last 12 records moreover eleven articles are included and bringing the total number of 23 articles that are added in this SLR for data elimination and the approximate synthesis phase. Figure 1 depicts the flow diagram for article identification and inclusion.

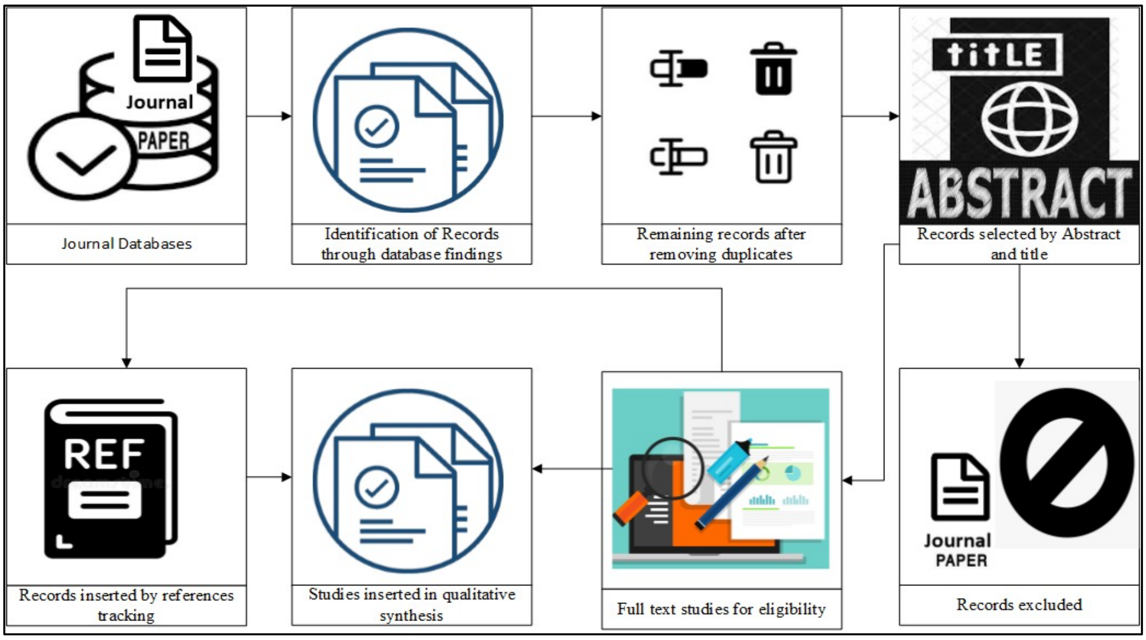

Figure 1. Flow diagram of recognition and inclusion articles

3.1. Eligibility of the studies

Although all the articles that are chosen in this SLR fulfilling the insertion and elimination criteria which are necessary to explain why particular studies are chosen. The article offers a method for short-term predicting Persistent AFib. However, the ECG data utilized is gathered from sheep rather than humans. Despite this, we believe this research is not too suitable due to the ECG signal's nature, but usually because of the methods and algorithms presented in the work. AFib Prediction is exclusively done between pre and post-AFib stages in the following articles, excluding situations where no prediction of AFib is made. Despite they are admitted due to publications that provide useful information for our research. At last, neither the tables nor the article text matched the stated analysis using the single-fold technique in the research. This last research is selected to be admitted, just for considering the perfect measurements for the 10-fold technique, which contains accurate outcomes of values in the tables and text of the study.

3.2 Source of Evidence:

Of the selected twenty-three types of research that are admitted in this SLR half of them are published in 2018 starts and the papers that are published in 2021 are also included in this research as shown in Fig.2.

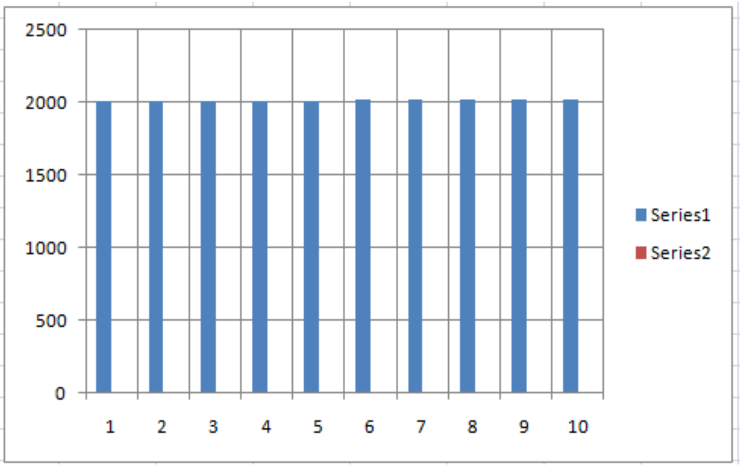

Figure 2. Number of studies from 2012 to 2021

Table 3 is showing the chosen article's classification on the kind of publication and their major focusing area for publishing citations places.

Table 3. This table is presenting the number of publications chosen from journal and conference types.

Name of publications	Descriptions	No. of articles	Total sections
Journal and conference	Hindawi	1	8%
	M dpi	4	33%
	Lancet Digital Health	1	8%
	Springer	2	25%
	Kim et al. Int J Arrhythm	1	8%
	Medicine Journal	1	8%
	Eesss. sp	2	16%
	IEEE Conference and promoted by IEEE	3	25%
	Computer Science Journal	3	25%
	Bioinformatics Journal	4	8%
	ESC	1	

3.3 Study participators and design

In this section, we are describing and explain the selected studies with their results. Two types of research (8.69%) do not provide the size of a sample. One research (4.34%) is based on a sample of about 126000 individuals. Seven kinds of study (30.43%) are based on tiny samples of persons (less than 100) with databases, whereas two studies (8.69%) are based on samples of 100 to 25 000 people[19]. Three types of research (13.04%) used a private ECG records database, one study with (4.34%) result used a database of UCI Repository Warehouse, one with the Mayo Clinic database of ECG laboratories, one based on the database of Medical Information Mart for Intensive Care III, five studies with (21.73%) results used the AFib Prediction Database, and one study with (4.34%) result used the database of China KadoorieBi-obank[20]. The databases may be accessed via the Physio net Repository. Only the articles are accessible to the general audience.

3.4 Prediction techniques

The proposed studies employed a variety of DL techniques to predict AFib are given below:

1. Two types of research applied CNN.
2. Five types of research including] used support vector machine SVM.
3. One article used its study by applying Arrhythmia Fuzzy Hybrid Classifier.
4. Another article applied Markov Chain;
5. Two types of research applied statistical AI techniques.
6. The following studies also applied DL techniques.
7. Last, the final article applied the Expert's Mixture technique to predict AFib.

3.5 Collection of Data from selected articles

It is critical to obtain as much knowledge as desirable from the proposed research throughout the quality synthesis process. Although the fact that all of the research included some additional data, most of the data from these studies is not comparable, due to this reason, they are not included in upcoming data tables. The numbers of datasets applied in selected articles and the number of participants, from where the data was gathered and the age of participants, if it's given are shown in Table 4.

Table 4. This table explains the Input knowledge collected from selected studies.

NM=Not Mentioned.

Year	Datasets used	No. of participators	Age of participators
	MIT-BTH d/Dense Nets	NM	NM
	MIT-BIH AFib Database	NM	NM
	MIT-BIH, IN CART, BIDMC Databases	NM	NM
2021[19]	MIT-BIH ECG	NM	NM
	Dataset	NM	NM
	Meta-Analysis/DL Algorithm	NM	NM
	MIT-BIHA	NM	NM
	MIT-BIHNSR	NM	NM
		NM	NM
	Physio Net	NM	NM
	MIT-BIH Atrial Fibrillation Dataset,	NM	NM
	MIT-BIH Atrial Fibrillation Database	NM	NM
2020[21]	Or AFDB	NM	NM
	Datasets of ECGs (standers 10s 12 channel format) datasets of 962 ECGs	NM	NM
	MIT-BIH AF and MIT-BIH arrhythmia database	NM	NM
	AI model		
		NM	NM
2019[22]	Mayo Clinic ECG Laboratory.	126526	>18, average60.3
	Medical details Mart for Intensive Care III database.	246	NM
	Atrial Fibrillation Prediction Database.	53	NM
2019[23]	Atrial Fibrillation Prediction Database.		NM
	Own collected dataset.	53	NM
	CI Repository Warehouse.	5	NM
2018[24]	Own collected dataset	33	NM
2016[25]	Atrial Fibrillation Prediction Database	53	NM
	China Kadoorie Biobank. Own collected dataset.	24369	NM
2013[23]	Atrial Fibrillation Prediction Database	75	NM
2012[26]	Atrial Fibrillation Prediction Database.	NM	NM

Deep Learning Techniques are not used in all of the chosen twenty-23 types of research that did not need chosen features and elimination from the original ECG signals. Table 5 showed the features extracted from the ECG signal they include the total number of the frequency domain, time domain, space domain, and non-linear characteristics retrieved from a particular study, however, as an input the signal duration is utilized to the DL method and the tools used throughout the data collection and pre-processing study phases. There is a notable variation among the lowest and highest duration of signals used in every selected article. Three studies used a 300-second signal, with 30 and 10 seconds following every two studies. In other publications, signals of lengths 120, 180, 3600, and 0 seconds are reported to be used in each study.

Table 5. This table is explaining the Signal treatment information gathered from selected articles.

NM= Not Mention

Years	Features extracted from ECG signals	Signal timeline	Tools and techniques applied
	NM	9/61	Dense ECG, without class weight, Dense ECG, with class weight
2021[19]	NM	NM	DL tool
	RR-Intervals	NM	Deep Learning
	NM	NM	DL techniques
	NM	NM	CNN, LSTM
	83 features R-R Intervals, single-wave features, and full-wave features. Deep features	300 and 250	ML (SVM, LR, MLP, XG Boost)
	NM	Random	DL (CNN, LSTM, Bilstm, Rest Net)
	Entropy feature	60 MINUTES,30 MINUTES With both DBS 4sec in preprocessing	The Deep Learning techniques are known as H-ELM
	NM	30	CNN
2020[21]	NM	10	GE-Marquette ECG machine MUSE system Kera's, Tensor Flow Python, R
	5 time-domain		
	NM	NM	Excel, MATLAB 2015, Rapid Miner
	3 frequency-domain 2 time-domain 2 non-linear	120	NM
		900	C++, Lib SVM library
	frequency-domain 5 time-domain 8 non-linear 11 time-frequency	300	NM

2019[22]	frequency-domain 5 time-domain 3 non-linear	30	MATLAB, Lib SVM toolbox
	NM 1time-domain	30	Caffe DL framework.
	1 space-domain	10	NM
	8 frequency-domain 1 time-domain	1800	NM
	frequency-domain1 time-domain	300	Pan-Tompkin's algorithm, MATLAB 2008
2019[23]	frequency-domain 6 time-domain 4 non-linear	300	NM

Table 6 contained different types of methods that are applied in every chosen study. These various types of methods are pre-processing methods, prediction methods, and evaluation methods. Also, the number of training iterations applied and the data split with training and testing groups are all shown in Table 6. Table 7 shows the information on Accuracy, sensitivity, specificity, F-score, and Area under the curve. In all of the preferred studies, the recognized algorithms are compared to the reported accuracy in which most of the studies do not provide information on sensitivity and specificity also not the F-Score nor the Area Under the Curve are given. Almost all of the studies that are chosen, as shown in Table 5, used feature extraction. The studies that simply used a Deep Learning approach did not accomplish feature extraction. Table 8 shows all of the features picked and retrieved from each of the selected studies as we go deeper into the research analysis and comparison of the chosen articles. Despite, all the articles are not applied the duplicate dataset, the findings of every selected study may demonstrate that its techniques can predict AFib which is a fundamental achievement, all the studies have in common. Table 9 illustrates the prediction horizon for each of the studies that are chosen, i.e., how long can the resultant models predict AFib episodes throughout the time?

Table 6. This table is representing the methods that are applied in the selected articles
Where Not Mention=NM.

Year	Data processing	Prediction methods	Validation methods	No.of iterations	Data spilt Training/testing
	QRS detection	RNN, LSTM	5-fold cross-validation and 10-fold cross-validation	NM	50/75
2021[19]	Noise removal ECG signal, high-frequency	CNN, LSTM RNN, GRU	NM	NM	NM
	NM	ANN, CNN	NM	NM	NM
	NM	NM	NM	NM	NM
	NM	COMD, READ	NM	NM	NM

	removing noises from ECG signals, used zero pending	DL (CNN, LSTM)	10-fold cross-validation	NM	70/20
	ECG signal, train and evaluate the automatic AF	CNN, LSTM	NM	6000	NM
	NM	Deep Learning	NM	NM	80%
	Filtering baseline wandering noise, elimination of high-frequency artefacts	Health Control (HC) ECG signals and AF ECG signals	10-fold cross-validation	NM	NM
	NM	Deep Learning	NM	NM	0.61
2020[21]	Noise removal, QRS detection Normalization, Missing values removal	Belief Functions Theory Mixture of Experts	10-fold cross-Validation	NM	47/53
	Noise removal, QRS detection	Novel Arrhythmia Fuzzy Hybrid Classifier Algorithm	10-foldCross-Validation	5	90.6/9.4
	Noise removal	Weighted	10-fold cross-Validation	100	75/25
	NM	SVM	NM	30000	90/10
2019[22]	NM	CNN with ON/OFF ReLU	5-fold cross-Validation	NM	NM
	Hamilton and Tompkins algorithm, Mc Names algorithm	SVM	10-fold cross-Validation	10	90.6/9.4
	Noise removal	HRV identification and Morphologic Variability of QRS difficulties	NM	NM	50/50
2019[23]	Noise removal QRS identification	SVM	NM	NM	47/53

Table 7. This table is representing the evaluation of the selected articles.

NM=Not mention

Year	Accuracy	Sensitivity	Specificity	F-Score	Area under curve
2021[19]	94.70	NM	NM	0.929,	0.94/0.96
	NM	97.87%	99.29%	0.929	NM
	NM	NM	98.94%	And	98.82
	95.70%	NM	97.57%,	0.931,	97.03,
	and	NM	97.18%	0.931	97.39%
	97.10%	NM	99.04,	NM	99.1,
	98%,87	NM	99.21%	97.25,	98.74%
	%,83%,	NM	NM	96.73%	NM
	99.09	NM	NM	99.23,	NM
	97.1%,	98.6%,	96.04%,	96.94%	NM
	96.04%	99.1%	96.2%	NM	NM
2020[18]	86.5	0.85	0.86	0.86	NM
	97.21	97.34	97.08	NM	0.97
	NM	0·867	0·995	0.887	0.983
	99.93	99.86	100	NM	NM
	66.3	53.3	73.2	52.5	0.61
2020[21]	83.30%	82.30%	83.40%	45.40%	90.00%
	70.49%	77.07%	63.90%	NM	NM
	98.21%	100.00%	96.55%	NM	NM
	82.80%	0.40%	0.43%	1.21%	NM
	82.00%	86.00%	80.00%	74.51%	90.88%
	87.70%	86.80%	88.70%	NM	NM
2019[22]	84.90%	66.70%	97.00%	NM	93.50%
	80.20%	81.10%	79.30%	NM	NM
	83.58%	NM	NM	NM	NM
	75.60%	NM	NM	NM	83.00%
2019[23]	90.00%	89.44%	89.29%	NM	89.40%
	96.64%	96.30%	93.10%	NM	NM

4. Discussion and Implications

In this section, we demonstrate and summarize the research questions related to this SLR that were previously defined. The basic goal of this SLR is to discover, evaluate, and examine the current art stat of ECG models to predict the AFib by applying DL(DL) techniques.

4.1 (Q1) How do the research studies address prediction problems for AFib?

The majority of articles from the preferred studies address the AFib prediction problem i.e., their major concentration is to predict AFib and not any other heart pathologies and arrhythmias types. All of the studies that were chosen employed classification prediction, which implies that they all classified the prediction using discrete labels. Only one of the twenty-three studies that were selected in this SLR used a risk-based technique to predict problems which means the maximum articles used AFib time series prediction. Only two papers stated a study in which various class techniques are used, as regards the total number of classes applied to predict the process while the other remaining studies applied binary among ''pre-AFib" and ''not pre-AFib" episodes. Despite this, all the selected articles are not reporting predictions of event horizon except some of them where, two studies utilized a 30-minute horizon, while the others used fourteen days, sixty minutes, five minutes, two minutes, and less than a zero-minute horizon (instantly earlier the AFib episodes).

From the twenty-three selected studies, eight studies performed AFib prediction by using input signals that are less than or equal to 300 seconds and 5 min long, being the most common duration of signal employed by the studies. The datasets that are applied in selected studies when looking at them, three of the five studies that utilized the dataset have the three most accurate models, indicating that this is a solid choice for further research on evaluating the problem.

4.2 (RQ2) What sort of datasets and feature extraction techniques are used?

Despite, most of the selected articles do not extract ECG signal features when they perform ECG signal features extraction the chosen features have a direct influence on the model's capacity to predict the presence of AFib with greater accuracy. The various features selected from publications examined in this SLR are given in Table 8.

Table 8. Table 8 contains the Features collected from each of the chosen articles' input.

Domain	Features
Nonlinear	SD1/SD2 ratio Approximate Entropy Standard Variation 1 (SD1) Standard Variation 2 (SD2)
Space	R wave Dimension S wave Dimension T wave Dimension P wave Dimension Q wave Dimension
Time	Average Standard Deviation of all NN intervals to every five-minute duration of the entire recording (SDANN) ST level RR Standard Deviation interruption (RRSD). RR interval Mean. RR interval Skewers. RR interval Kurtosis. Adjacent RR intervals numbers that vary by greater than 50 milliseconds (NN50) NN50 separated by the RR intervals total number (PNN50) The mean square root of the squares of differences among adjacent RR intervals (RMSSD). Differences Standard deviation among adjacent RR intervals (SDSD) Smoothed Pseudo Winger Ville distribution (SPWVD).
Frequency	Low-Frequency band power (LF) High- High band power (HF) LF/HF ratio Low-Frequency Fast Fourier Transforms component (FFT-LF). High-Frequency Fast Fourier Transforms component (FFT-HF). LL-H1.

	LL-H2. HH-H3. ROI-H1. ROI-H2. ROI-H3. QRS segment periods. Intervals of P-R waves. Intervals of Q-T waves. Intervals of the T wave. Intervals of the P wave. spectrum Weighted center (ROI-WCOB. Very Low-Frequency band power (VLF).

Standard Deviation of RR Intervals, lowest frequency band power, Standard Deviation, and Mean of RR Intervals are the most commonly used characteristics, applied in at least three various chosen articles. The majority of methods are using ECG signals; however, with heart morphology data one article is linked with ECG signals data. Almost half of the selected publications used the AFib prediction database, while the remaining four used their own ECG signals dataset. Other studies make use of the UCI Repository Warehouse dataset, the Medical Information Mart for Critical Care III database, a China KadoorieBiobank database, and the Mayo Clinic ECG Laboratory database. The UCI Repository Warehouse dataset contains 452 occurrences and 279 attributes, including ECG records provided in image format for all patients aged eighteen and older between 1993 and 2017. Thereafter, eighteen patients with at least one digital, regular sinus rhythm, standard ten-second, and twelve-lead ECG obtained in a supine posture are admitted to the Mayo Clinic ECG Laboratory's database, which is used in the article.
The signals are captured using a GE-Marquette ECG Machine at a sampling rate of 500 Hz and stored with the MUSE data management system. All completed data records are evaluated by a qualified technician and physician-supervised, with diagnostic label modifications made as needed. According to the report, the Medical Information Mart for Critical Care III database was compiled between 2001 and 2012. It collects data on heart rate, arterial blood pressure, and respiratory rate from over 40 thousand ill patients. Higher frequency charts, such as ECGs and continual blood pressure readings from Intensive Care Unit patients, are also included in the Medical Information Mart for Intensive Care III database. Only patients with AFib at the time of reporting are considered for the study.
The AFib Prediction Database is applied in the following articles, and this database contains two-channel long-term ECG recordings excerpts and it is also excerpted inside the training set and equal-size test set. As described, the database contains digitized ECG signals (sampled at one hundred and twenty-eight Hz for each signal using twelve-bit resolution) and a collection of unaudited, automatically produced QRS explanations. Although the chosen studies consistently refer to fifty-three or seventy-five participants, the records were obtained from forty-eight people, as shown in Table 3.
The 10 seconds length twelve-lead ECG signals data at five hundred Hz is used for the research, also blood pressure data is obtained from 24 369 individuals with the Mortara ELIx50 device between 2013 and 2014.

4.3 (RQ3) Which preprocessing techniques are used?

Table 9 summarizes the pre-processing techniques utilized in each of the twenty-three proposed articles. The majority of the studies were amplified using prediction techniques, which didn't require any signal pre-processing. Each article then used normalization and missing value removal techniques. Noise removal (5 articles) is the most common pre-processing technique, followed by QRS Detection (4 articles), Signal Correction, and Signal Correction, respectively. (2 articles)

Table 9. This table is describing the Prediction horizon for every chosen study where NM stands for "Not Mentioned."

Years	Prediction horizon
2021	NM
	NM
	NM
	NM
	NM
	NM
2020	NM
	NM
	NM
	60 minutes
	NM
2019	NM
	60 minutes
2018	5 minutes
	NM
	2 minutes
	NM
2017	14 days
2016	30 minutes
	NM
	NM
2013	30 minutes
2012	NM

4.4 (RQ4) What type of Deep Learning models are applied for the prediction of AFib?

SVM is the most widely used prediction algorithm which is pursued by CNN. Other chosen research used statistical DL approaches like Morphologic QRS variability complications and HRV to examine the faithful theory of Functions as well as the Markov Chain, Classifier of Arrhythmia Fuzzy Hybrid and Expertise Mixture. We can identify the prediction approach types utilized in a specific study by categorizing predictive models into three categories: Machine Learning (ML), Deep Learning (DL), and Artificial Intelligence (AI).

4.5 (RQ5) Which metrics are mainly applied to measure the accuracy of Deep Learning algorithms?

For labelling a model comparative assessment applied in selected articles, this SLR researchers cluster the deliberation in the following terms:

1. The selected articles have the same duration of signals.
2. The same datasets apply to all selected articles.
3. Articles that used the same DL approach.

4. All of the selected research.
5. Selected research using the same prediction techniques.

4.5.1 The selected articles have the same duration of signals.

Studies that used a variety of Experts as a method according to, as well as associating the two greatest executions for using the features SD1, LF, SD2 HF, SD1/SD2, and Samples of Entropy in articles. The following two studies utilize 30 seconds of signal duration. These two experiments similarly give 85% accuracy, however owing to a higher number of people from whom data is acquired and the use of SVM rather than CNN in the first study, the maximum performance was reached. Two papers state work completed with the duration of 10-second signals. The executed stronger rather than the because it used CNN techniques instead of SVM with the greatest number of individuals from whom the data is obtained at least 525 times.

4.5.2 The same datasets apply in all selected articles.

Only five studies from the chosen research utilized the same dataset, all the other remaining articles are left to work with mostly used datasets. Although by analyzing sing in the article, by using a variety of expertise, Statistical DL techniques, and SVM These three articles obtained accuracies over or less than 99.93%. The two lowest-executing studies utilized SVM, making it impossible to show, which approach is the best to use.

4.5.3 Articles that used the same DL approaches.

The accuracy rates of the two articles that employed ML approaches are quite close. Searching for articles that used DL techniques while excluding those that used AL methods, all of these articles achieving 99.93% accuracy by using the DL Prediction Database and based on 3600-second signal durations, as well as time-domain, non-linear features, and frequency-domain extracted from the input ECG signal.

4.5.4 All of the selected research.

The research findings indicated that increasing the ECG signals duration time and sending these signals for prediction is not compulsorily to increase the developed model's accuracy. The highest prediction accuracies achieved in the following articles are shown here with their results. The article with 90.00%, 83.30%, 99.93%, 96.64% and 98.21%, inside these entire articles 300-second signal portions are applied. The following studies from the chosen article’s models acquired the poorest accuracies. With a 70.49% result, with 75.60%result, and the article with 80.20%result, these studies are also respectively based on 3600 seconds, 10 seconds, and 1800 seconds. These results may show that signals that are too short 10 seconds or less and too lengthy 1800 seconds or more are not the ideal techniques in this SLR for the problem being evaluated.

4.5.5 Selected research using the same prediction techniques.

LF, SD2, and Sample Entropy are utilized as features inside the research that used Support Vector Machine to perform the best. On the other hand, for the selected articles that used CNN as a prediction technique, these two studies achieved too much similar accuracy rates. Finally, the three articles are compared in this SLR according to their accuracy results that are using DL techniques and the authors acquired the highest accuracy in the article that used AFib Prediction Database, with 3600 seconds and 1800 second signals duration and the article [did not mention its signal duration, on the other hand, the article had 10-second signal duration. Finally, the article has the highest performance of the three of them. Finally, the researchers emphasize the article's accuracy using the DL Prediction Database and ECG signals with 3600 seconds and 1800 seconds durations, also the study does not highlight its signal duration whereas the study has the duration of a 10-second signal. At last, the DL technique is used in all of these three articles.

5. Conclusion and Future Work

This SLR provides and encapsulates the current data-based work to predict AFib with input ECG data and DL techniques. The followings are the major conclusions from the twenty-three selected articles that are reviewed:

5.1 RQ1

Despite this, there is no existence of the present highest amount of published studies consisting of those articles that concentrated to predict AFib with DL and ECG signals. The majority of articles that do exist evaluate the problem by predicting the exclusive AFib cases and as a bigger collection of articles binary prediction system, at the same time these are not concentrating on other cardiovascular cases predictions, the majority of prior research used ECG signals with five minutes 300 seconds duration. However, most of the articles have focused on escalating the ECG signals period length applied as input to predict the models. It does not automatically expand the achieved final model accuracy.

5.2 RQ2

For training, the majority of accurate methods are obtained with AFib Prediction Database among all of the studies that are chosen in this SLR. Approximately, half of the chosen articles are using this database where Standard Deviation, RR Standard Deviation, Lowest Frequency band power and RR Intervals Mean are the greatest applied features that are obtained from inputted ECG signals. This might imply that RR intervals inside the ECG examinations have the highest importance in terms of predicting AFib

5.3 RQ3

Various pre-processing procedures are utilised in all of the chosen research, with Noise Removal being the most popular by detecting the QRS complex and collecting the most widely used characteristics relating to climax and RR intervals.

5.4 RQ4

The tendency in predictive techniques with DL approaches is increasing where SVM and CNN are the most two commonly applied techniques across all of the chosen articles, showing the tendency of DL approaches inside this field. However, when compared to the simplest SVM techniques, the researcher of this SLR found that the use of AI methods is less authentic which resulted in the greatest inconsistency of findings and greater difficulty in obtaining the required conclusions from the analyzed data.

5.5 RQ5

Usually, the greatest accuracy rates are acquired from DL techniques-based studies. This higher accuracy is achieved using Expertise Mixtures techniques with the implementation of SVM where sample Entropy, LF, and SD2 are chosen features that are administrated for greatest accuracy, as well as ECG signals, are used with 3600 and 1800 durations as an input to train the methods to derive for predicting accuracy with higher rates. AFib prediction Database is one of the most used databases that achieved the three highest accurate models in this SLR.

Fig 2 presents the reference Time window to this SLR from 2012 to 2021 where less than 90% of the selected articles are executed before 2018 and 50% belong to the following last three years 2019, 2020, and 2021. The total collection of work for predicting AFib episodes is continuously expanding and presenting unique outcomes. However, the impressive outcomes achieved from DL techniques to predict various areas like health care are not used in some studies which are concentrating on AFib prediction with ECG signals. So, the highest results are obtained with DL approaches like Expertise Mixture and SVM.

Finally, this SLR has some limitations given here:

- This SLR only prefers research in articles that are written in English.
- The research in studies returned some studies based on selected articles cross-reference.

All of the studies which included data gathered from sick people based on current surgical proceedings and using known cardiovascular problems and might be used to infer ECG examination outcomes are eliminated from this SLR.

At last, the chosen research has to provide the following assessment metrics accuracy, sensitivity, specificity, and the confusion matrix by eliminating any article without its evaluations.

Funding: No funding was received.
Data Availability Statement: The authors declare that all data supporting the findings of this study are available within the article.
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mhamdi, L., et al., Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems. Biomedicines, 2022. 10(8): p. 2013.
2. Javaid, A., et al., Medicine 2032: The future of cardiovascular disease prevention with machine learning and digital health technology. American Journal of Preventive Cardiology, 2022. 12: p. 100379.
3. Buelga Suárez, M., et al., Smartwatch ECG Tracing and Ischemic Heart Disease: ACS Watch Study. Cardiology, 2022.
4. Kim, M.-J., Building a Cardiovascular Disease Prediction Model for Smartwatch Users Using Machine Learning: Based on the Korea National Health and Nutrition Examination Survey. Biosensors, 2021. 11(7): p. 228.
5. Chen, Y., et al., A Single-Center Validation of the Accuracy of a Photoplethysmography-Based Smartwatch for Screening Obstructive Sleep Apnea. Nat Sci Sleep, 2021. 13: p. 1533-1544.
6. Stark, K., et al., Watch out for ST-elevation myocardial infarction: a case report of ST-elevation in single-lead electrocardiogram tracing of a smartwatch. Eur Heart J Case Rep, 2020. 4(6): p. 1-4.
7. Sengupta, A., et al., A Mobile Health Intervention System for Women With Coronary Heart Disease: Usability Study. JMIR Form Res, 2020. 4(6): p. e16420.
8. Samol, A., et al., Recording of Bipolar Multichannel ECGs by a Smartwatch: Modern ECG Diagnostic 100 Years after Einthoven. Sensors, 2019. 19(13): p. 2894.
9. Samol, A., et al., Single-Lead ECG Recordings Including Einthoven and Wilson Leads by a Smartwatch: A New Era of Patient Directed Early ECG Differential Diagnosis of Cardiac Diseases? Sensors, 2019. 19(20): p. 4377.
10. Wall, H.K., et al., Vital Signs: Prevalence of Key Cardiovascular Disease Risk Factors for Million Hearts 2022 - United States, 2011-2016. MMWR Morb Mortal Wkly Rep, 2018. 67(35): p. 983-991.
11. Islam, S.S., et al., Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study. PeerJ Computer Science, 2022. 8: p. e898.
12. Hu, M., et al., Development and preliminary validation of a machine learning system for thyroid dysfunction diagnosis based on routine laboratory tests. Communications Medicine, 2022. 2(1): p. 9.
13. Guleria, K., et al., Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning. Measurement: Sensors, 2022. 24: p. 100482.
14. Alyas, T., et al., Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach. BioMed Research International, 2022. 2022: p. 9809932.
15. salman, K. and E. Sonuç, Thyroid Disease Classification Using Machine Learning Algorithms. Journal of Physics: Conference Series, 2021. 1963(1): p. 012140.
16. Li, Y., et al., Serum Raman spectroscopy combined with Deep Neural Network for analysis and rapid screening of hyperthyroidism and hypothyroidism. Photodiagnosis and Photodynamic Therapy, 2021. 35: p. 102382.
17. Hosseinzadeh, M., et al., A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things. The Journal of Supercomputing, 2021. 77(4): p. 3616-3637.
18. Chaubey, G., et al., Thyroid Disease Prediction Using Machine Learning Approaches. National Academy Science Letters, 2021. 44(3): p. 233-238.
19. Aversano, L., et al., Thyroid Disease Treatment prediction with machine learning approaches. Procedia Computer Science, 2021. 192: p. 1031-1040.
20. Abbad Ur Rehman, H., et al., Performance Analysis of Machine Learning Algorithms for Thyroid Disease. Arabian Journal for Science and Engineering, 2021. 46(10): p. 9437-9449.
21. Chai, X., Diagnosis Method of Thyroid Disease Combining Knowledge Graph and Deep Learning. IEEE Access, 2020. 8: p. 149787-149795.
22. Zhang, B., et al., Machine Learning–Assisted System for Thyroid Nodule Diagnosis. Thyroid, 2019. 29(6): p. 858-867.
23. Shahid, A.H., et al. A Study on Label TSH, T3, T4U, TT4, FTI in Hyperthyroidism and Hypothyroidism using Machine Learning Techniques. in 2019 International Conference on Communication and Electronics Systems (ICCES). 2019.
24. Tyagi, A., R. Mehra, and A. Saxena. Interactive Thyroid Disease Prediction System Using Machine Learning Technique. in 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC). 2018.
25. Prasad, V., T.S. Rao, and M.S.P. Babu, Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms. Soft Computing, 2016. 20(3): p. 1179-1189.
26. Asif, M.A.A.R., et al. Computer Aided Diagnosis of Thyroid Disease Using Machine Learning Algorithms. in 2020 11th International Conference on Electrical and Computer Engineering (ICECE). 2020.