A Focused Survey of ECG Datasets for Artificial Intelligence-Based Atrial Fibrillation Detection

ASSALHI Imane[1], Bybi Abdelmajid[2], Oulad Hamdaoui Hanaa[3],
Ebobisse Djene Yves Frederic[4], Drissi Lahssini Hilal[5]
MEAT Laboratory, EST Sale, EMI Rabat, Mohammed V University in Rabat, Sale, Rabat, Morocco[1, 2, 3]
SmartiLab Laboratory, EMSI Rabat, Rabat, Morocco[1, 4]
LASTIMI Laboratory, EST Sale, EMI Rabat, Mohammed V University in Rabat, Sale, Rabat, Morocco[5]

***Abstract*—Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia and increases the risk of stroke, heart failure, and mortality. Electrocardiography (ECG) is the most important technology for AF detection because it is inexpensive, non-invasive, and provides clinically useful information. However, the variability of ECG patterns, particularly during paroxysmal AF creates challenges in detecting AF. Artificial Intelligence (AI) offers a promising opportunity to improve AF recognition. However, AI performance is contingent on obtaining high-quality and diverse ECG datasets. This paper presents a focused survey of 15 publicly available and clinical ECG datasets used in AI-driven AF detection research between 2023 and 2025. We analyze the datasets based on acquisition methods, ECG type, format, lead configurations, annotation richness, and their application in AI models. Our comparative analysis reveals major trends, challenges such as data imbalance and motion artifacts, and gaps in current datasets including limited demographic diversity and underrepresentation of wearable ECG data. This study aims to guide future research toward more robust, interpretable, and inclusive AF detection models.**

Keywords*—*Atrial fibrillation; ECG datasets; Artificial Intelligence; AI-ECG; dataset survey; AF detection

I. Introduction

Atrial fibrillation (AF) is a serious and prevalent cardiac condition characterized by irregular electrical activity in the atria, which can lead to blood clots, stroke, and heart failure [1]. As the global burden of AF rises, early and accurate detection has become critical for timely treatment and risk management. Electrocardiograms (ECGs) serve as the gold standard for detecting AF due to their ability to capture cardiac electrical activity in a fast, non-invasive, and cost-effective manner [2].

In recent years, Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), has shown remarkable performance in analyzing ECG signals to automatically detect AF. However, these AI models depend heavily on the quality, structure, and diversity of the ECG datasets used for training and validation. While previous surveys often focus on AI architectures, little attention has been paid to the ECG datasets themselves.

This paper addresses that gap by providing a focused survey of ECG datasets used in AI-based AF detection. The survey is structured as follows: First, a technical background discusses ECG types for AF diagnosis and highlights commonly used AI technologies. Second, the related work section emphasizes the importance of our survey and outlines our added value compared to existing reviews, which often overlook the role of ECG datasets in AI-based AF diagnosis. Third, we present our methodology for assembling recent studies from 2023-2025 and extracting the datasets used. The survey results follow, organized in two tables: one summarizing the extracted datasets, and the second providing an analysis grid of 12 selected studies detailing preprocessing techniques, AI models, architectures, and performance metrics. Finally, we discuss insights from these results, highlighting trends, challenges, and future research directions.

II. Background

Electrocardiograms (ECGs) record the heart's electrical activity and can be captured using different lead configurations, such as the standard 12-lead or single-lead used in wearable devices. The duration of ECG recordings can vary from seconds to days, with longer recordings, known as Ambulatory ECGs, often obtained via wearable devices. These continuous recordings are particularly useful for detecting paroxysmal atrial fibrillation (AF), which occurs intermittently and may not be captured in shorter ECG sessions. To address this, AI models like Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Convolutional Neural Networks (CNNs), and Long Short-Term Memory Networks (LSTMs) are trained on diverse ECG data, including ambulatory recordings, with recent models incorporating explainable AI (XAI) techniques to improve interpretability and clinical trust [3]. The data can be stored as raw signals or transformed into images for visual or AI-based analysis [4].

III. Related Work

Recent studies have explored AI-based tools for assessing atrial fibrillation (AF) burden, showing high correlation with manual physician assessments. Notably, AI demonstrated strong agreement and minimal bias, offering an efficient and accurate alternative for AF burden evaluation [5]. However, when it comes to AI-driven AF detection, the dataset plays a crucial role in the precision of the diagnosis. The quality, variety, and characteristics of ECG datasets significantly impact the performance of AI models. Despite the growing importance of these datasets, there is a noticeable gap in existing reviews and surveys. Many reviews discuss ECGs from a biometric perspective [6] or focus on ECGs for arrhythmia detection in general, but few address the specifics of AI-based AF detection[7], [8]. My previous review focused on the latest AI technologies for AF detection [9], but this current work expands by compiling and analyzing the ECG datasets specifically used for AI-driven AF diagnosis. This review aims to fill the gap by highlighting the diverse datasets used in AF detection, which will aid future research in understanding the strengths and limitations of various data sources and model implementations in the field.

IV. Methodology

To conduct this dataset-oriented survey on atrial fibrillation (AF) detection using artificial intelligence (AI) models applied to electrocardiogram (ECG) data, we followed systematic review principles, inspired by PRISMA guidelines as shown in Fig. 1, to ensure transparency and reproducibility in the selection and analysis of studies. Bibliographic references were collected using the following keywords: "Artificial Intelligence", "AI", "Machine Learning", "ML", "Deep Learning", "DL", "Atrial Fibrillation", "AF", "AFib", "Electrocardiogram", and "ECG". These were queried in several major scientific databases, including Scopus, Science Direct, Web of Science, and IEEE Xplore. As part of the inclusion criteria, we considered only papers that had the core terms in the title, ensuring a strict focus on AF detection using AI applied to ECG data with a total of 304 papers. Additionally, the time frame was restricted to 2023–2025 and limited to 25 June 2025, and only articles published in English were considered. We included only original research articles, conference papers, and technical studies, while review and survey papers were examined separately to identify potential overlaps or gaps but were ultimately excluded from the analysis. All retrieved references were exported into Mendeley for management and screening. After removing duplicates, 81 unique entries remained. During the screening process, papers were excluded if they:

- Focused on physiological signals other than ECG (e.g., PPG or ECHO).
- Studied diseases related to AF rather than AF itself (e.g., stroke outcomes or AF recurrence after catheter ablation).
- Did not report the use of AI techniques for AF detection.

Following the title and abstract screening, 48 papers met the inclusion criteria. In the full-text screening phase, 33 full-text papers were successfully retrieved. After reviewing the literature, it became clear that a lack of standardization in ECG datasets used across AI models was a significant limitation. This prompted a focus on compiling a dataset-oriented survey of ECG datasets for AI-based AF detection. We've selected 33 studies from 2023 to 2025, extracting key information about 15 different ECG datasets. The dataset attributes included ECG type, acquisition method, sample size, format, annotation quality, and availability. To gather this information, we conducted online searches for dataset sources and details, including any risk factors or clinical labels associated with each dataset. The extracted data were organized in Table I, summarizing each dataset's key characteristics and their relevance to AI-based AF detection. This compilation provides a comprehensive overview of the datasets, filling a gap in the current literature.

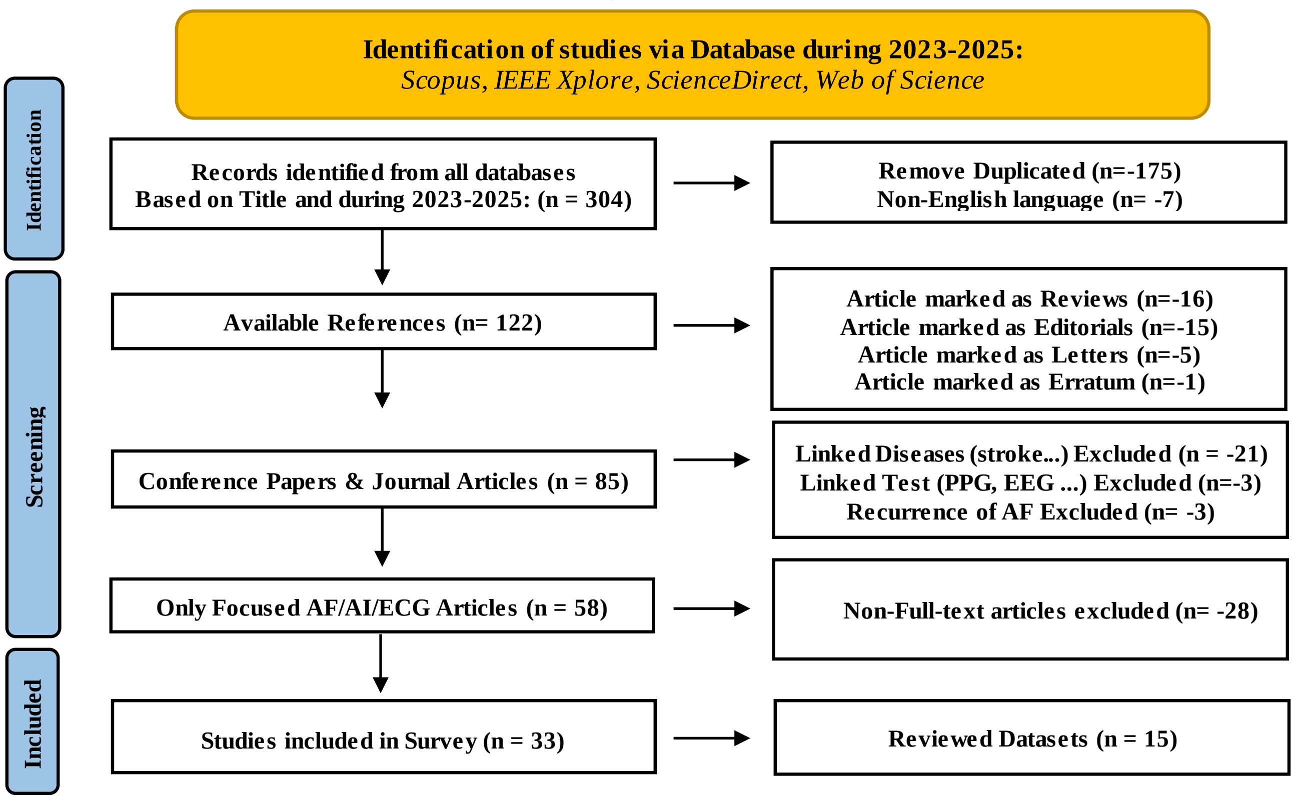

Fig. 1. PRISMA flow diagram representing the paper selection process for the dataset-oriented survey on AI-based Atrial Fibrillation detection.

TABLE I. SUMMARY OF ECG DATASETS USED FOR ATRIAL FIBRILLATION DETECTION

Ref	ABR	Name of Dataset	Year	Source type	Acquisition device	Sample Size	Duration	ECG Type	ECG Format	Number of ECG leads	Annotation	Link	Studies
[10]	CPSC 2018	- The China Physiological Signal Challenge Database - CPSC-Extra Database	2018	Public	Clinical devices from 11 hospitals	10 330 records: - Set 1: 6877 Records - Set 2: 3453 Records	6s to 60s	Standard ECG	Signal	12	- ECG only - Age & Sex only	The China Physiological Signal Challenge 2018	[11] [12] [13]
[14]	INCART	St. Petersburg Institute of Cardiological Technics 12lead Arrhythmia Database	2008	Public	Holter Monitor	75 Records from 32 Patient	30 min	Ambulatory ECG	Signal	12	Over 175,000 beat annotations; auto-detected and manually corrected; includes age, sex, diagnosis.	St Petersburg INCART 12-lead Arrhythmia Database v1.0.0	[12] [15]
[16]	PTB	PTB Diagnostic ECG Database	2004	Public	PTB custum prototype Recorder	549 Records from 290 Patient	Varies up to minutes	Resting ECG	Signal	15 leads: • 12 • 3 Frank	Diagnosis, age, gender, history, medications, interventions, and clinical summaries **(some records missing Annotations).**	PTB Diagnostic ECG Database v1.0.0	[11] [12] [13] [17] [18]
[19]	PTB-XL	PTB-XL, a large publicly available electrocardiography dataset	2019	Public	Device from Achiller AG	21 799 Records from 18 689 Patients	10s	Resting ECG	Signal	12	SCP-ECG codes (71 statements), multi-label: diagnostic, form, rhythm; metadata includes age, sex, height, weight, heart axis, infarction stage, signal noise info.	PTB-XL, a large publicly available electrocardiography dataset v1.0.3	
X	Georgia	The Georgia 12-lead ECG Challenge (G12EC) Database	2000	Public	Not specified (Assuming the clinical devices)	10 344 Records	10s	Resting ECG	Signal	12	Labels not clearly defined in brief, Includes diagnosis/class	Georgia 12-Lead ECG Challenge Database	[11] [12]
[20]	2017 PhysioNet Challenge	The PhysioNet/Computing in Cardiology Challenge 2017	2017	Public	Kardia Mobile: AliveCor	8528 Records	9s to 60s (Varies)	Minimal ECG	Signal	Single lead	Rhythm Labels : • Normal • Atrial • Fibrillation • Other Rhythms • Noisy	AF Classification from a Short Single Lead ECG Recording: The PhysioNet/Computing in Cardiology Challenge 2017 v1.0.0	[21] [22] [23] [24] [25] [26] [27] [28]
X	MUSE	the MUSE cardiology information system (GE,Healthcare, Chicago, IL, USA)	X	Clinical Not Available	Clinical	Diverse and Over 300 000 Records	Usially 10s	Resting ECG	Signal	12	• Raw waveform signals • Demographics (age, sex, etc.) • Clinical diagnoses and physician interpretations • ECG measurements (HR, PR interval, QT, etc.) • Not publicly available,	MUSE Cardiology Information System | GE HealthCare (United Kingdom)	[3] [29]

Ref	ABR	Name of Dataset	Year	Source type	Acquisition device	Sample Size	Duration	ECG Type	ECG Format	Number of ECG leads	Annotation	Link	Studies
											• Access typically requires collaboration with a hospital using the MUSE system.		
X	PHYJ	Hospital of Yangjiang (PHYJ),Yangjiang, China, Database	X	Clinical Not Available	Clinical	100 Records from 100 Patients	15s	Resting ECG	Array	12	• Labels confirmed by Holter monitoring and cardiologist review • Not publicly available.	X	[30]
X	EHR	the ECG and electronic health record (EHR) databases from AZ Delta	2022	Clinical Not Available	GE MUSE Cardiology System	173 537 ECGs from 68 880 Patients	10s	Resting ECG	MUSE Format	• 12 • Single	• AF diagnosis from structured EHR + MUSE ECG labels • Not publicly available	Best Electronic Health Record (EHR) Datasets & Databases 2025 | Datarade	[31]
[32]	Chapman	Chapman University and Shaoxing People's Hospital database,	2020	Clinical Available	GE MUSE Cardiology System	10 646 Patients	10s	Resting ECG	Digital	12	• Expert-labeled: - 11 rhythms + 67 conditions	A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients	[13] [33] [34]
X	CNUH	Chonnam National University Hospital databases	2022	Clinical Not Available	MobiCARE-MC100	3059 ECGs from 6720 patients	60s	Mobile ECG	Digital	Single	• Labeled as "AF" (masked AF) or "Healthy" based on 12-lead ECG history • Not publicly available	X	[35]
[36]	AFDB	MIT-BIH Atrial Fibrillation Database	2000	Public	Clinical Ambulatory ECG recorders	25 records	10h	Ambulatory ECG	Digital	2 leads	• Rhythm: AFIB, AFL, J, N (manual) • Beat: .qrs (auto), .qrsc (manually corrected, some records)	MIT-BIH Atrial Fibrillation Database v1.0.0	[37] [28] [38] [39]
[40]	LTAF	The Long Term AF Database	2000	Public	Clinical Ambulatory ECG recorders	86 Records from 80 Patients	21h to 24h	Ambulatory ECG	Digital	2 lead or 3 leads	• ST episode annotations • beat annotations • ST level measurements	Long Term ST Database v1.0.0	[38]
[41]	MITDB	The MIT-BIH Arrhythmia Database	2001	Public	Clinical Ambulatory ECG recorders	48 Records from 47 Patients	30min	Ambulatory ECG	Binary	2 leads	• Beat-by-beat annotations by cardiologists (~110,000 beats annotated)	MIT-BIH Arrhythmia Database v1.0.0	[26] [28] [38]
[42]	NSRDB	MIT-BIH Normal Sinus Rhythm Database	2000	Public	Clinical Ambulatory ECG recorders	18 Subjects	24h	Ambulatory ECG	Binary	2 leads	• No significant arrhythmias; • Rhythm annotations only • Healthy volunteers	The MIT-BIH Normal Sinus Rhythm Database	[38]

V. ECG Datasets Landscape

Over the past decade, a wide range of ECG datasets have been utilized to develop AI models for Atrial Fibrillation (AF) detection. Table I presents a comparative analysis of 15 ECG datasets that were extracted following the used methodology in section before. These datasets differ significantly in terms of source (public vs. clinical), ECG type (resting vs. ambulatory), number of leads (from single-lead to 15-lead), signal duration(6 seconds to 24 hours), and data format (digital signals, WFDB, MUSE, or image-based ECG scans). Publicly accessible datasets like PTB-XL, AFDB, MITDB, and the PhysioNet 2017 Challenge have been widely adopted due to their availability and rich rhythm annotations [19], [20], [36], [41]. Clinical datasets such as MUSE, PHYJ, and CNUH, while offering more realistic clinical scenarios, are often inaccessible, limiting reproducibility and external validation[44]. There is growing interest in single-lead and ambulatory ECG datasets, especially for wearable and real-time monitoring applications [36]. In contrast, 12-lead ECGs remain the most common in clinical datasets due to their diagnostic richness [19], [32]. Regarding duration, some datasets contain short snapshots (typically 10 seconds) while others, such as LTAF and MITDB, include long-term ambulatory ECGs suitable for detecting paroxysmal or transient AF episodes. Datasets vary considerably in annotation granularity. Some, like the PhysioNet 2017 Challenge and AFDB, provide beat-level or rhythm-level labels, which are essential for supervised learning. Others, like PTB and CNUH, contain diagnostic summaries but lack detailed temporal annotation, which restricts their use in time-series AF detection tasks. A few datasets offer clinical metadata such as age, sex, and comorbidities, but this information remains inconsistent and often incomplete, hindering the study of model fairness and demographic generalization. A variety of acquisition devices are represented, including hospital-grade machines, Holter monitors, and portable solutions like KardiaMobile [5], [45]. Additionally, several large-scale, non-public clinical datasets (e.g., from hospitals in the US, Israel, Japan, and the UK Biobank) have been used in recent studies to validate generalization performance across diverse patient populations [46], [47]. One notable example involved over 320,000 ECGs from 130,000+ patients using both Philips and GE systems [48]. These real-world datasets, though inaccessible to the broader community, are essential for robust AF detection, especially in challenging cases such as paroxysmal or asymptomatic AF [40], [41], [49], [50], [51]. Despite recent advancements, the lack of standardized annotations, inconsistent demographic diversity, and limited access to prospective real-world datasets remain key barriers to model development and evaluation

VI. AI Models Landscape for AF Diagnosis

Table II provides a comprehensive overview of 12 representative studies from 2024–2025 that apply AI techniques to ECG datasets for atrial fibrillation (AF) detection. Each entry outlines the dataset used, preprocessing pipeline, model architecture, performance, and key innovations. Together, they illustrate the current state of applied AI in this domain and support a synthesis of prevailing trends, methods, and research directions.

Key Observations:

1) Dominance of Deep Learning (DL): CNNs remain the most widely adopted models, often serving as backbone architectures in hybrid setups (e.g., CNN-LSTM, CNN-RNN). Lightweight architectures (e.g., MobileNet, ultra-compact CNNs) are increasingly favored in resource-constrained environments like wearables and mobile devices [3][11][21].

2) Rise of hybrid and transformer models: Hybrid models combining CNNs with LSTMs or GRUs enhance temporal pattern recognition. Transformer-based models are also emerging, particularly in multimodal settings (e.g., ECG + HRV + demographics), demonstrating high AUROC scores (e.g., 0.9668 with VGG-16) [24] [33] [38].

3) Preprocessing pipelines vary widely: Signal preprocessing includes traditional filtering (e.g., Butterworth, Pan-Tompkins), wavelet transforms (e.g., CWT), and advanced time–frequency representations (e.g., STFT, spectrograms). Some models integrate domain-specific techniques like ECG segmentation (PQRST) or RR interval extraction for feature fusion. [22] [30] [33] [35] [37].

4) Explainability and interpretability: Explainable AI (XAI) is increasingly used through techniques like Layer-wise Relevance Propagation (LRP), anomaly scoring, and saliency maps. Studies applying LRP or segment-wise analysis provide insights into model behavior and potential diagnostic biomarkers, such as T-wave/ST-segment deviations in paroxysmal AF[3] [33] [34].

5) Dataset-specific adaptations: Several studies tailor their models to specific datasets public or clinical accounting for noise, class imbalance, and sampling frequency. For example, the use of DBSCAN-GAN for denoising and synthetic augmentation on noisy wearable ECG datasets shows promise for generalization [24].

6) Clinical integration and risk factors: A subset of studies incorporates demographic features (e.g., age, sex) and electronic health record (EHR)-derived risk scores. This multimodal fusion enhances AF prediction, particularly in identifying paroxysmal AF from normal sinus rhythm, and aligns with real-world deployment needs [31].

7) Performance benchmarks: Reported metrics span AUROC (up to 0.98), F1-scores (up to 0.99), and accuracy (>96% in some cases). However, performance varies across datasets and is often affected by noise, signal duration, or sampling frequency. Lightweight models maintain high performance with reduced computational cost, enabling deployment in real-time monitoring systems [18].

8) Cross-dataset and external validation: Some studies emphasize generalization by testing across multiple public datasets (e.g., PTB-XL, MITDB, PhysioNet 2017). However, only a few employs external validation on private clinical data an essential step for real-world readiness [11].

This comparative summary not only emphasizes the practical applications of each dataset, but also reveals methodological trends, current performance benchmarks, and emerging best practices in AI-driven AF detection.

TABLE II. OVERVIEW OF AI MODELS FOR ATRIAL FIBRILLATION DETECTION FROM ECG SIGNALS

Papers	Year	Dataset	Preprocessing Techniques	AI Models Type	Architecture details	Performance Metrics	Key Contribution	Notes
[11]	2025	• CPSC2018 • PTB-XL • Georgia	• Raw ECG; • Feature extraction; • Demographic data (age, sex); • Saliency maps	• Deep Learning: ○ CNNs, ○ RNNs (GRU), ○ Transformers	• AlexNet, • VGG-16, • LeNet, • ResNet, • Inception, • FCN, • GRU, • Transformers	• AUROC: ○ VGG-16 (0.9668), ○ AlexNet (0.9617); • Sensitivity: ~0.9225 with HRV+demo	• Multimodal input (ECG + HRV + demographics) • Improves AF detection; • Simple models outperform complex ones	• Final best performers: AlexNet and VGG-16 due to efficiency, performance, and interpretability emphasises • Best performance on PTB-XL; generalizability and label quality vary across datasets
[21]	2025	Physionet challenge 2017	• Bandpass filtering • Short-Time Fourier Transform (STFT) • Reverse polar transform • PanTompkins (P-T) algorithm	• CNN • Ensemble Voting	• Pretrained MobileNet, ResNet50,DenseNet121 • 5-fold cross-validation • Soft and hard voting ensemble	• Accuracy, • Precision, • Recall, • F1-score	• Introduced reverse polar-transformed spectrograms for ECG • Improved AFib detection with better visual representation • Effective use of compact square matrices for CNN input	• Sensitive to signal amplitude and filter type • Well-suited for real-time and wearable ECG analysis
[3]	2025	MUSE	• Feature extraction; • Demographic data (age, sex);	• Deep Learning (CNN)	Not fully disclosed • CNN • Layer wise Relevance Propagation	• AUROC: 0.905 ± 0.00	• Demonstrated ability to predict paroxysmal AF onset from normal sinus rhythm ECGs using deep learning. • Used LRP to identify T-wave/ST abnormalities as key predictors	• Focused on paroxysmal AF prediction. • Applied LRP (Layer-wise Relevance Propagation) for XAI.
[22]	2025	Physionet challenge 2017	• Bnechmark Algorithm • Expert cardiologist review • removal of artifacted ECGs	• Deep Learnig	• Cloud-based AI platform: "The Willem Artificial Intelligence platform" [43] • Trained on >520,000 patients' ECGs; • detects 23+ arrhythmias; • Uses data from: ○ PTB-XL, ○ MIT-BIH, ○ Georgia DB, ○ ESC DB, ○ AHA DB, etc.	• Accuracy: 96.4% • Sensitivity: 84.2% • Specificity: 97.6% • PPV: 78.0% • NPV: 98.4%	• High-performance AF detection from 1-lead ECGs, • Outperforming rule-based algorithms • Real-world deployability	• Detected additional arrhythmias (PVC, PAC, AV block); • Platform was not trained on test ECGs (external validation) • Performance degraded with noise/artifacts
[23]	2025	Physionet challenge 2017	• Raw ECG used directly;	• Hybrid DL ○ CNN	• Ultra-lightweight CNN for feature extraction	• F1 Score: 99.56%,	• Introduces a compact, real-time	• Optimized for speed and embedded applications;

Papers	Year	Dataset	Preprocessing Techniques	AI Models Type	Architecture details	Performance Metrics	Key Contribution	Notes
			• Class imbalance addressed using Neurokit, BioSPPy, and EngZeeMod filters for synthetic data augmentation	○ LSTM	• LSTM for temporal patterns; • 64.9K parameters; • 0.48 ms inference time	• Size gain: 102.25 dB, • 1.06% Gain over SOTA	architecture with high accuracy on imbalanced ECG data	• Avoids complex preprocessing
[17]	2025	PTB-XL	• Coarse graining (1s/0s) • Complexity feature extraction from all 12 leads; • Evaluated each lead and combinations.	• Traditional ML	• Multiple models optimized using Bayesian optimization. • KNN was tuned over 1–99 neighbors (odd values) with 10 distance functions; • SVM models tested with linear, • Quadratic polynomial, • Gaussian kernels. • Leave-one-out cross-validation was used for robust evaluation.	• Peak accuracy ~0.69 at 125 Hz sampling (using lead V6).	• First human application of complexity analysis for detecting a history of PAF from normal sinus rhythm recordings.	Additional experiments at 500 Hz were conducted using various coarse graining techniques: • LZ76 (A), • LZ78 (B), • Titchener (C); also reporting results with configurations: ○ BD (Beat Detection), ○ FD (Feature Detection), KM (K-means), ○ TC (Threshold Crossing) across individual leads and their combination. ○ KNN outperformed SVM overall.
[30]	2024	PHYJ	• Signal Preprocessing (Filter) • Signal Dimensionality Reduction (SDR); • Kernal Principal Component Analysis (KPCA) for dimensionality reduction; • Wavelet Transform (CWT) for time–frequency domain;	• Deep Learning (CNN)	• AlexNet, • VGG19, • ResNet152, • Inception-v3, • Inception ResNet-v2	• Accuracy (ACC), • Sensitivity (SEN), • Specificity (SPF) • Cross Validation	• First approach to detect PAF from sinus rhythm ECGs using full 12-lead CWT + KPCA + deep network fusion	• Highlights potential to detect PAF from non-diagnostic ECGs; • Future work aims to scale validation and involve open-access datasets for pretraining
[31]	2024	EHR MUSE	• ECG converted to OMOP-CDM; • Risk factor extraction from ICD codes, Rx, and measurements	• Multiple residual neural network	• ResNet • Random Forest	• AUC = 0.74 (ECG only), • AUC = 0.76 (ECG + 6 RFs); • Stable across age and sex	• Demonstrates that AF can be predicted during sinus rhythm using 1-lead ECG + 6 key risk factors • Model matches 12-lead performance.	• Designed for **real-world deployment** • Reduces **age bias** seen in prior studies Prospective clinical validation planned

Papers	Year	Dataset	Preprocessing Techniques	AI Models Type	Architecture details	Performance Metrics	Key Contribution	Notes
[33]	2024	• Chapman • PTB-XL	• ECGs segmented into PreQ, QRS, and PostS • Only Lead II used	• Unsupervised DL: ○ LSTM Autoencoder ○ XGBoost	• LSTM-based autoencoder trained on normal ECG segments • Anomaly detection via MSE per segmenT • Postprocessing with XGBoost classifier on anomaly scores	• AUROC (PreQ): ○ Experiment A: 0.96, ○ Experiment B: 0.90, ○ Experiment C: 0.95 • AUROC (XGBoost): 0.98 • F1 Score: 0.94	• Demonstrated explainable, segment-wise anomaly detection of AF without supervision • PreQ segment most predictive of AF	• Cross-dataset validation (Germany/China) • Highlights clinical interpretability by segment-level scores • Addresses physician concerns about DL "black-boxing"
[24]	2024	Physionet challenge 2017	• Two-stage ECG denoising and filtering • ECG-SQE (signal quality evaluation using Pan-Tompkins based ML) • 9-sec segmentation • Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for outlier filtering	• Deep learning (hybrid)	• MuDANet: Multi-stream CNN-RNN with Dense Attention ○ Dual-stream Conv-RNN (enhances signal representation ○ Attention mechanism improves focus on discriminative features ○ Final classification into AFR, NSR, and Other • DBSCAN-GAN: ○ Density-based clustering to isolate clean data ○ GAN trained only on clean samples for synthetic augmentation	• F1 Score: 0.876 (baseline) • F1 Score: 0.962 (with DBSCAN-GAN and 10-fold CV)	• Introduced MuDANet: dual-stream CNN-RNN with dense attention for AF detection. • Proposed DBSCAN-GAN for noise-aware, class balanced synthetic ECG generation. • Demonstrated strong generalization using synthetic data in a noisy real-world dataset.	• Real-time capable and wearable-device friendly • Combines traditional ECG signal cleaning (Pan-Tompkins) with advanced DL and GAN-based augmentation • Overcomes common data imbalance and outlier noise issues in wearable ECG datasets
[35]	2024	CNUH	• PQRST detection and trimming • Baseline correction • 0.5 Hz high-pass Butterworth filter • Segmentation into 10-second intervals • Random under-sampling to reduce class imbalance	• Deep Learning ○ CNN ○ LSTM ○ RNN	• ResNet50: deep CNN architecture with skip connections • RNN and LSTM models for comparison • Training with AdamW optimizer, early stopping, batch size = 32	• ResNet50: ○ F1 = 71.9% ○ Recall = 79.3% ○ Precision = 65.8% ○ Accuracy = 70.5% ○ AUC = 0.79 • External set" ○ F1 = 64.1%, ○ AUC = 0.68	• Demonstrated that deep learning can detect masked AF from NSR in single-lead mobile ECG	• Study used a private clinical dataset; • High potential for real-world wearable device applications
[38]	2024	• AFDB • LTAF • MITDB • NSRDB	• R-peak detection to compute RR intervals (RRIs) • Removal of poor-quality or noisy data	• Hybrid CNN-LSTM	• Multi-input fusion network combining ECG and RRI features • Res-CNN for morphological features	• High accuracy across 4 external datasets • Ablation study	• Proposed MIF-AFNet: a multi-input fusion model integrating RR interval + ECG features for robust AF detection	• - Limitations: requires ≥30s input; R-peak detection accuracy affects performance; lacks diverse arrhythmias • Future work: clinical validation, enhanced

Papers	Year	Dataset	Preprocessing Techniques	AI Models Type	Architecture details	Performance Metrics	Key Contribution	Notes
			• ECG augmentation via vertical flipping		• Bi-LSTM for temporal sequence modeling of RRI • Low complexity design for real-time application	validated robustness • Maintains performance with ectopic beats	• High generalization capability across datasets and rhythms	R-peak detection, training on multi-arrhythmia data, improved noise handling

VII. Analysis and Critical Evaluation

Building on the comparative review of ECG datasets and applied AI models, this section critically evaluates key trends, limitations, and methodological patterns observed across recent studies on AF detection.

1) Patterns and bias in dataset usage: Despite the growing number of ECG datasets, the majority of recent studies rely heavily on PTB-XL, PhysioNet 2017, and AFDB. This over-reliance introduces dataset bias, limiting the generalizability of models to broader clinical populations. Less common datasets, especially clinical ones like MUSE or EHR, are underutilized due to accessibility barriers, despite offering more realistic and varied signals.

2) Dataset limitations affecting model performance: Several datasets exhibit class imbalance (e.g., fewer AF vs. non-AF samples), requiring oversampling, Synthetic Minority Over-sampling Technique (SMOTE), or weighted loss functions. Wearable ECG datasets often contain motion artifacts and noise, while many collections lack demographic metadata (e.g., age, sex, comorbidities), which hampers model personalization. Only a few datasets integrate clinical risk factors, despite their value in enhancing prediction.

3) Clinical vs. Wearable datasets: Clinical datasets (e.g., PTB-XL, Chapman) offer clean, high-resolution ECGs from controlled environments. In contrast, wearable datasets enable long-term monitoring but introduce higher noise levels and data variability. Many models perform well on clean datasets but degrade on wearable ones unless robust preprocessing or augmentation is applied. This creates a gap between lab performance and real-world application [33]..

4) Single-lead vs. Multi-lead ECGs: 12-lead ECGs remain the standard for model development and benchmarking due to their diagnostic richness. However, single-lead ECGs, especially from wearables, are gaining traction for screening. While multi-lead ECGs generally outperform in accuracy, well-designed models using single-lead inputs (e.g., AZ Delta study) demonstrate that simpler signals combined with clinical features can achieve comparable performance, especially in ambulatory settings.

5) Reproducibility and transparency gaps: Many studies omit details about signal preprocessing, segmentation, or lead selection. Some apply noise filters or augmentation without reproducible descriptions. Additionally, few publish code or preprocessing pipelines, making it difficult to replicate findings or benchmark new approaches. This lack of transparency reduces trust and slows progress.

6) Standout approaches: A few studies stand out for their innovation. For example, the AZ Delta AF-SR study combined a single-lead ECG during sinus rhythm with six clinical risk factors, achieving strong predictive performance while reducing age-related bias ideal for wearable screening [31]. Another notable work used LSTM autoencoders to detect anomalies in specific ECG segments (PreQ), enhancing interpretability and requiring no labeled AF samples.

Overall, while current approaches show promise, overcoming dataset biases, improving reproducibility, and integrating richer clinical context remain essential for advancing reliable and generalizable AF detection models.

VIII. Discussion and Feature Directions

A. Key Takeaways:

Our analysis of recent AI-based Atrial Fibrillation (AF) detection studies reveals consistent patterns across dataset usage and model development. A key concern is the heavy dependence on a limited set of datasets particularly PTB-XL, PhysioNet 2017, and AFDB while more clinically representative datasets remain underutilized. This trend narrows the scope of model validation and hinders generalizability to real-world scenarios. Moreover, datasets with limited demographic diversity and missing metadata restrict the ability to evaluate models for fairness, particularly across age groups, sex, or comorbidities. Additionally, wearable ECGs, while increasingly relevant for remote monitoring and early AF screening, are still underexplored due to their noisy nature and signal variability. As a result, models often perform well in controlled, clinical conditions but falter when deployed in real-world or ambulatory environments.

B. Gaps and Limitations in Current Approaches:

1) Overfitting to benchmarks: Many models are optimized for performance on a few benchmark datasets, which inflates accuracy metrics but limits cross-dataset robustness.

2) Lack of dataset standardization: Inconsistent data formats, label definitions (e.g., AF types), and missing preprocessing documentation make reproducibility difficult and comparisons unreliable.

3) Fairness blind spots: Absence of demographic and clinical metadata (e.g., sex, age, comorbidities) prevents analysis of bias, reducing the trustworthiness of AI models for sensitive applications.

4) Underutilization of wearable data: Despite their value for long-term monitoring, wearable ECG datasets are rarely used in training or evaluation due to signal quality issues.

5) Reproducibility challenges: Very few studies share their code, training configurations, or preprocessing pipelines, making model replication and clinical translation problematic.

C. Future Research Directions

To address these challenges and unlock the full potential of AI in AF detection, we recommend the following research directions:

1) Standardization and benchmarking needed: Most public datasets vary in signal format, label structure, and metadata availability. Establishing a benchmark framework with unified preprocessing and labeling would allow for fair model comparisons and reproducibility.

2) Underuse of wearable and real-time ECGs: Wearable devices are central to future AF screening strategies, yet few studies utilize real-world data from such sources. Research should focus on noise-robust algorithms and domain adaptation techniques to bridge the performance gap.

3) Explainability still lacking: Explainable AI remains an exception rather than the rule. Future models should incorporate interpretability methods (e.g., attention maps, ECG segment analysis, SHAP values) to support clinical trust and decision-making.

4) Missing demographics and clinical metadata: Datasets should include age, sex, risk factors, and comorbidities to support personalized predictions and fairness assessment. This metadata is crucial for clinical deployment and regulatory validation.

5) Hybrid modeling with clinical context: Few models integrate non-ECG features like EHR data, medication history, or symptom logs. Incorporating these modalities can improve predictive performance and reduce overreliance on limited signal features.

D. Proposed Dataset Guidelines for Future Research

Based on our review, the following dataset practices are strongly encouraged to support future innovation:

- Provide at least 3 leads when possible (e.g., lead I, II, V1) for improved feature richness.
- Share both raw and preprocessed signal versions.
- Use consistent and annotated labels, including AF types (e.g., paroxysmal, persistent).
- Include demographic and clinical metadata (age, sex, comorbidities).
- Ensure public access under a research license to foster collaboration and reproducibility.

IX. CONCLUSION

This survey examined 15 ECG datasets from 33 recent AI models for Atrial Fibrillation (AF) detection studies, revealing a strong reliance on PTB-XL, PhysioNet 2017, and AFDB. While these resources have advanced the field, their overuse, coupled with missing demographic and clinical metadata, limits generalizability and fairness. Our analysis highlights the growing gap between model performance on clean benchmark datasets and the noisy reality of wearable ECGs. We emphasized the importance of dataset quality, transparency in preprocessing, and the inclusion of patient context for real-world readiness. This work represents the first survey focused specifically on datasets in AI-based AF detection. By mapping current usage patterns and limitations, we provide a guide for future research toward more inclusive, robust, and clinically meaningful AI tools. A collaborative push for standardized, annotated, and diverse datasets including wearable ECGs and multimodal data is essential to bridge the gap between lab success and clinical impact.

REFERENCES

[1] "The Clinical Profile and Pathophysiology of Atrial Fibrillation | Circulation Research." Accessed: Jun. 25, 2025. [Online]. Available: https://www.ahajournals.org/doi/full/10.1161/CIRCRESAHA.114.303211

[2] S. S. Chugh et al., "Worldwide Epidemiology of Atrial Fibrillation: A Global Burden of Disease 2010 Study," Circulation, vol. 129, no. 8, pp. 837–847, Feb. 2014, doi: 10.1161/CIRCULATIONAHA.113.005119.

[3] Y. Jin, W. Chang, B. Ko, K.-H. Choi, and K. H. Lee, "Explainable paroxysmal atrial fibrillation diagnosis using an artificial intelligence-enabled electrocardiogram," Korean Journal of Internal Medicine, vol. 40, no. 2, pp. 251–261, 2025, doi: 10.3904/kjim.2024.130.

[4] F. M. Dias et al., "Artificial Intelligence-Driven Screening System for Rapid Image-Based Classification of 12-Lead ECG Exams: A Promising Solution for Emergency Room Prioritization," IEEE Access, vol. 11, pp. 121739–121752, 2023, doi: 10.1109/ACCESS.2023.3328538.

[5] E. Hennings et al., "Assessment of the atrial fibrillation burden in Holter electrocardiogram recordings using artificial intelligence," Cardiovascular Digital Health Journal, vol. 4, no. 2, pp. 41–47, Apr. 2023, doi: 10.1016/j.cvdhj.2023.01.003.

[6] D. Meltzer and D. Luengo, "ECG-Based Biometric Recognition: A Survey of Methods and Databases," Sensors, vol. 25, no. 6, Mar. 2025, doi: 10.3390/s25061864.

[7] Y. Ansari, O. Mourad, K. Qaraqe, and E. Serpedin, "Deep learning for ECG Arrhythmia detection and classification: an overview of progress for period 2017–2023," Frontiers in Physiology, vol. 14, 2023, doi: 10.3389/fphys.2023.1246746.

[8] J. Rahul and L. D. Sharma, "Advancements in AI for cardiac arrhythmia detection: A comprehensive overview," Computer Science Review, vol. 56, May 2025, doi: 10.1016/j.cosrev.2024.100719.

[9] ASSALHI Imane and BYBI Abdelmajid, "Artificial Intelligence for Atrial Fibrillation Detection: A Systematic Review of Recent Advances in ECG-Based Deep Learning Models," in 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), IEEE, 2025.

[10] F. Liu et al., "An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection," j med imaging hlth inform, vol. 8, no. 7, pp. 1368–1373, Sep. 2018, doi: 10.1166/jmihi.2018.2442.

[11] A. Rawshani et al., "Integrating deep learning with ECG, heart rate variability and demographic data for improved detection of atrial fibrillation," Open Heart, vol. 12, no. 1, 2025, doi: 10.1136/openhrt-2025-003185.

[12] B. B.-S. Chuang and A. C. Yang, "Optimization of using multiple machine learning approaches in atrial fibrillation detection based on a large-scale data set of 12-lead electrocardiograms: Cross-sectional study," Jmir Formative Research, vol. 8, 2024, doi: 10.2196/47803.

[13] A. Petroni, F. Cuomo, G. Scarano, S. Colonnese, P. Francia, and M. Pediconi, "MUSE: MUlti-lead Sub-beat ECG for remote AI based atrial fibrillation detection," Journal of Network and Computer Applications, vol. 212, 2023, doi: 10.1016/j.jnca.2022.103544.

[14] V. Tihonenko, A. Khaustov, S. Ivanov, and A. Rivin, "St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database." physionet.org, 2007. doi: 10.13026/C2V88N.

[15] A. Makhir, M. H. E. Alaoui, A. Jilbab, and L. Bellarbi, "Classification of Atrial Fibrillation and Cardiac Arrhythmias by a CNN-BiLSTM Hybrid Model with DWT Preprocessing," presented at the 2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology, IRASET 2024, 2024. doi: 10.1109/IRASET60544.2024.10549460.

[16] R.-D. Bousseljot, D. Kreiseler, and A. Schnabel, "The PTB Diagnostic ECG Database." physionet.org, 2004. doi: 10.13026/C28C71.

[17] S. Creasy, V. Alexeenko, K. Jeevaratnam, P. J. Aston, G. Y. H. Lip, and G. Tse, "Electrocardiogram sampling frequency for the optimal performance of complexity analysis and machine learning models: Discrimination between patients with and without paroxysmal atrial fibrillation using sinus rhythm electrocardiograms," Heart Rhythm O2, vol. 6, no. 1, pp. 48–57, 2025, doi: 10.1016/j.hroo.2024.11.002.

[18] T. Bender, P. Gemke, E. Idrobo-Avila, H. Dathe, D. Krefting, and N. Spicher, Benchmarking the impact of noise on deep learning-based classification of atrial fibrillation in 12-lead ECG, vol. 302. 2023, pp. 977–981. doi: 10.3233/SHTI230321.

[19] P. Wagner et al., "PTB-XL, a large publicly available electrocardiography dataset," Sci Data, vol. 7, no. 1, p. 154, May 2020, doi: 10.1038/s41597-020-0495-6.

[20] G. Clifford et al., "AF Classification from a Short Single Lead ECG Recording: the Physionet Computing in Cardiology Challenge 2017," presented at the 2017 Computing in Cardiology Conference, Sep. 2017. doi: 10.22489/CinC.2017.065-469.

[21] D. Kwon, H. Kang, D. Lee, and Y.-C. Kim, "Deep learning-based prediction of atrial fibrillation from polar transformed time-frequency electrocardiogram," Plos One, vol. 20, no. 3 March, 2025, doi: 10.1371/journal.pone.0317630.

[22] F. D. Guio et al., "Enhanced detection of atrial fibrillation in single-lead electrocardiograms using a Cloud-based artificial intelligence platform," Heart Rhythm, 2025, doi: 10.1016/j.hrthm.2024.12.048.

[23] A. Khan et al., "A deep learning-based ultra-lightweight architecture for atrial fibrillation detection using single-lead ECG recordings," IEEE Access, vol. 13, pp. 86474–86486, 2025, doi: 10.1109/ACCESS.2025.3567093.

[24] G. Msigwa, E. Ntambala, and J. Yun, "Enhancing atrial fibrillation classification from single-lead electrocardiogram signals using attention-based networks and generative adversarial networks with density-based clustering," Engineering Applications of Artificial Intelligence, vol. 133, p. 108607, Jul. 2024, doi: 10.1016/j.engappai.2024.108607.

[25] H. Zhang, H. Zhao, and Z. Guo, "Artificial Intelligence-Based Atrial Fibrillation Recognition Method for Motion Artifact-Contaminated Electrocardiogram Signals Preprocessed by Adaptive Filtering Algorithm," Sensors, vol. 24, no. 12, p. 3789, Jun. 2024, doi: 10.3390/s24123789.

[26] L. Xia, S. He, Y. F. Huang, and H. Ma, "Multiscale Dilated Convolutional Neural Network for Atrial Fibrillation Detection," Plos one, vol. 19, no. 6, p. e0301691, 2023.

[27] M. E. A. Bourkha, A. Hatim, D. Nasir, and S. El Beid, "Enhanced Atrial Fibrillation Detection-based Wavelet Scattering Transform with Time Window Selection and Neural Network Integration," International Journal of Advanced Computer Science and Applications, vol. 14, no. 12, pp. 494–503, 2023, doi: 10.14569/IJACSA.2023.0141252.

[28] Y. Hu, M. Wang, H. Tang, T. Feng, and C. Liu, "Detection of paroxysmal atrial fibrillation from dynamic ECG recordings based on a deep learning model," Journal of Personalized Medicine, vol. 13, no. 5, 2023, doi: 10.3390/jpm13050820.

[29] H. Gruwez et al., "Detecting paroxysmal atrial fibrillation from an electrocardiogram in sinus rhythm: External validation of the AI approach," Jacc Clinical Electrophysiology, vol. 9, no. 8P3, pp. 1771–1782, 2023, doi: 10.1016/j.jacep.2023.04.008.

[30] M. Deng et al., "An intelligent computer-aided diagnosis method for paroxysmal atrial fibrillation patients with nondiagnostic ECG signals," Biomedical Signal Processing and Control, vol. 88, p. 105683, Feb. 2024, doi: 10.1016/j.bspc.2023.105683.

[31] S. Dupulthys et al., "Single-lead electrocardiogram Artificial Intelligence model with risk factors detects atrial fibrillation during sinus rhythm," EUROPACE, vol. 26, no. 2, 2024, doi: 10.1093/europace/euad354.

[32] J. Zheng, J. Zhang, S. Danioko, H. Yao, H. Guo, and C. Rakovski, "A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients," Sci Data, vol. 7, no. 1, p. 48, Feb. 2020, doi: 10.1038/s41597-020-0386-x.

[33] S. Choi et al., "Diagnosis of atrial fibrillation based on AI-detected anomalies of ECG segments," Heliyon, vol. 10, no. 1, p. e23597, Jan. 2024, doi: 10.1016/j.heliyon.2023.e23597.

[34] A. Mohan, D. Elbers, O. Zilbershot, F. Afghah, and D. Vorchheimer, "Deciphering Heartbeat Signatures: A Vision Transformer Approach to Explainable Atrial Fibrillation Detection from ECG Signals," Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.09474

[35] J. Kim, B. Ko, M. Lee, K. H. Lee, S. J. Lee, and Y.-S. Lee, "Identification of atrial fibrillation with single-lead mobile ECG during normal sinus rhythm using deep learning," Journal of Korean Medical Science, vol. 39, no. 5, 2024, doi: 10.3346/jkms.2024.39.e56.

[36] G. B. Moody and R. G. Mark, "MIT-BIH Atrial Fibrillation Database." physionet.org, 1992. doi: 10.13026/C2MW2D.

[37] S. Mandala, A. R. P. Wibowo, Adiwijaya, Suyanto, M. S. M. Zahid, and A. Rizal, "The effects of daubechies wavelet basis function (DWBF) and decomposition level on the performance of artificial intelligence-based atrial fibrillation (AF) detection based on electrocardiogram (ECG) signals," Applied Sciences (Switzerland), vol. 13, no. 5, Mar. 2023, doi: 10.3390/app13053036.

[38] Y. Zou et al., "A generalizable and robust deep learning method for atrial fibrillation detection from long-term electrocardiogram," Biomedical Signal Processing and Control, vol. 90, 2024, doi: 10.1016/j.bspc.2023.105797.

[39] M. M. Rahman, M. W. Rivolta, R. Sassi, and F. Badilini, "Quantifying uncertainty of a deep learning model for atrial fibrillation detection from ECG signals," in Computing in cardiology, 2023. doi: 10.22489/CinC.2023.340.

[40] F. Jager et al., "The Long-Term ST Database." physionet.org, 1995. doi: 10.13026/C2G01T.

[41] G. B. Moody and R. G. Mark, "MIT-BIH Arrhythmia Database." physionet.org, 1992. doi: 10.13026/C2F305.

[42] T. A. L. The Beth Israel Deaconess Medical Center, "The MIT-BIH Normal Sinus Rhythm Database." physionet.org, 1990. doi: 10.13026/C2NK5R.

[43] Idoven 1903 S.L., "Evaluation of Electrocardiographic Data From High-risk Cardiac Patients Using WillemTM Cardiologist-level Artificial Intelligence Software. WILLEM Trial.," clinicaltrials.gov, Clinical trial registration NCT05890716, May 2025. Accessed: Jun. 25, 2025. [Online]. Available: https://clinicaltrials.gov/study/NCT05890716

[44] "MUSE Cardiology Information System." Accessed: Jun. 25, 2025. [Online]. Available: https://www.gehealthcare.com/en-gb/products/diagnostic-cardiology/muse-v9?npclid=CjwKCAjw0a-SBhBkEiwApljU0nguvZYAZ1DaaIftQ4qgx-ruc99u6H5Q2Rc63w7lgTIGvqD964bblRoC8TkQAvD_BwE

[45] D. Mannhart et al., "Clinical validation of an artificial intelligence algorithm offering cross-platform detection of atrial fibrillation using smart device electrocardiograms," Archives of Cardiovascular Diseases, vol. 116, no. 5, pp. 249–257, May 2023, doi: 10.1016/j.acvd.2023.04.003.

[46] X. Wang et al., "Genetic Susceptibility to Atrial Fibrillation Identified via Deep Learning of 12-Lead Electrocardiograms," Circulation Genomic and Precision Medicine, vol. 16, no. 4, pp. 340–349, 2023, doi: 10.1161/CIRCGEN.122.003808.

[47] N. Ben-Moshe, S. B. Brimer, E. Zvuloni, J. A. Behar, K. Tsutsui, and L. Sornmo, "RawECGNet: Deep Learning Generalization for Atrial Fibrillation Detection from the Raw ECG," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 9, pp. 5180–5188, 2024, doi: 10.1109/JBHI.2024.3404877.

[48] Y. Kim et al., "Predicting Future Incidences of Cardiac Arrhythmias Using Discrete Heartbeats from Normal Sinus Rhythm ECG Signals via Deep Learning Methods," Diagnostics, vol. 13, no. 17, p. 2849, Sep. 2023, doi: 10.3390/diagnostics13172849.

[49] A. Raghunath et al., "Artificial intelligence–enabled mobile electrocardiograms for event prediction in paroxysmal atrial fibrillation," Cardiovascular Digital Health Journal, vol. 4, no. 1, pp. 21–28, Feb. 2023, doi: 10.1016/j.cvdhj.2023.01.002.

[50] G. Baj et al., "Comparison of discrimination and calibration performance of ECG-based machine learning models for prediction of new-onset atrial fibrillation," BMC Medical Research Methodology, vol. 23, no. 1, 2023, doi: 10.1186/s12874-023-01989-3.

[51] T. Hygrell et al., "An artificial intelligence–based model for prediction of atrial fibrillation from single-lead sinus rhythm electrocardiograms facilitating screening," EP Europace, vol. 25, no. 4, pp. 1332–1338, Apr. 2023, doi: 10.1093/europace/euad036.