Heliyon 10 (2024) e27200

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

An Arrhythmia classification approach via deep learning using single-lead ECG without QRS wave detection

Liong-Rung Liu [a,b,f], Ming-Yuan Huang [b,f], Shu-Tien Huang [a,b,f], Lu-Chih Kung [b,f], Chao-hsiung Lee [b,f], Wen-Teng Yao [c,f], Ming-Feng Tsai [a,c,f], Cheng-Hung Hsu [a], Yu-Chang Chu [a], Fei-Hung Hung [g], Hung-Wen Chiu [a,d,e,*]

[a] Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan
[b] Department of Emergency Medicine, Mackay Memorial Hospital, Taipei, Taiwan
[c] Division of Plastic Surgery, Department of Surgery, Mackay Memorial Hospital, Taipei, Taiwan
[d] Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei, Taiwan
[e] Bioinformatics Data Science Center, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan
[f] Department of Medicine, Mackay Medical College, New Taipei City, Taiwan
[g] Health Data Analytics and Statistics Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan

ARTICLE INFO

Keywords:
Arrhythmia
Deep learning
Convolution neural network (CNN)
Single-lead electrocardiogram (ECG)
Computer-aided diagnosis (CAD)

ABSTRACT

Arrhythmia, a frequently encountered and life-threatening cardiac disorder, can manifest as a transient or isolated event. Traditional automatic arrhythmia detection methods have predominantly relied on QRS-wave signal detection. Contemporary research has focused on the utilization of wearable devices for continuous monitoring of heart rates and rhythms through single-lead electrocardiogram (ECG), which holds the potential to promptly detect arrhythmias. However, in this study, we employed a convolutional neural network (CNN) to classify distinct arrhythmias without QRS wave detection step. The ECG data utilized in this study were sourced from the publicly accessible PhysioNet databases. Taking into account the impact of the duration of ECG signal on accuracy, this study trained one-dimensional CNN models with 5-s and 10-s segments, respectively, and compared their results. In the results, the CNN model exhibited the capability to differentiate between Normal Sinus Rhythm (NSR) and various arrhythmias, including Atrial Fibrillation (AFIB), Atrial Flutter (AFL), Wolff-Parkinson-White syndrome (WPW), Ventricular Fibrillation (VF), Ventricular Tachycardia (VT), Ventricular Flutter (VFL), Mobitz II AV Block (MII), and Sinus Bradycardia (SB). Both 10-s and 5-s ECG segments exhibited comparable results, with an average classification accuracy of 97.31%. It reveals the feasibility of utilizing even shorter 5-s recordings for detecting arrhythmias in everyday scenarios. Detecting arrhythmias with a single lead aligns well with the practicality of wearable devices for daily use, and shorter detection times also align with their clinical utility in emergency situations.

* Corresponding author. Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan.

E-mail addresses: lewis0815@gmail.com (L.-R. Liu), m.c.yuan928@gmail.com (M.-Y. Huang), doc.stanley@gmail.com (S.-T. Huang), ryan0805@gmail.com (L.-C. Kung), leeruai@gmail.com (C.-h. Lee), wentengyao@gmail.com (W.-T. Yao), vacaflower@gmail.com (M.-F. Tsai), edward19931119@gmail.com (C.-H. Hsu), nono1688s@gmail.com (Y.-C. Chu), fionhung@tmu.edu.tw (F.-H. Hung), hwchiu@tmu.edu.tw (H.-W. Chiu).

https://doi.org/10.1016/j.heliyon.2024.e27200
Received 7 January 2024; Received in revised form 18 February 2024; Accepted 26 February 2024
Available online 29 February 2024

1. Introduction

1.1. Research background

Cardiovascular diseases (CVDs) are the leading cause of death worldwide. In 2019, an estimated 17.9 million people died from CVDs, representing 32% of all global deaths [1]. Arrhythmia refers to an abnormal heartbeat, which can be either tachycardia, bradycardia, or irregular rhythm. It may lead to CVDs development or progression of symptoms or complications of underlying heart disease [2]. The standard 12-lead electrocardiogram (ECG) is used for clinical examination of patients in hospitals [3]. However, the 12-lead ECG has the disadvantage of being a bulky device that is not practical for daily use in a home setting. Although portable multi-lead ECG are available on the market, they are not suitable for long-term use by the general public, and the results still require expert interpretation [4]. Potentially life-threatening heart rhythms can occur at any time during daily activities, sometimes lasting only for a few seconds or minutes. Therefore, such rhythms may go undetected during routine examinations or even when a patient is sent to the hospital. For example, nonsustained ventricular tachycardia *(*NSVT*)* occurs within a few seconds and then returns to a normal rhythm [5]; thus, it is rarely recorded on a 12-lead ECG. Traditionally, the Holter examination is arranged for high-risk patients, wherein patients must wear a device for 24 h, which is inconvenient and inefficient in an emergency setting [6]. While ECG is currently the best way to diagnose arrhythmias, early detection can be challenging due to the possibility of it being asymptomatic.

To address these challenges, many wearable devices and smart phone applications are now being studied to continuously track heart rates and rhythms using photople-thysmography (PPG) or single-lead ECG, which has the potential to detect arrhythmia immediately [7–10]. The traditional method of ECG detection relies on manually selecting features and making subjective judgments to determine heart rhythms. Convolutional neural networks (CNNs) utilize an operation method that eliminates the necessity for manual feature selection, rendering them highly effective in the domain of image processing [11]. Originally introduced in 1990 for the recognition of handwritten digits, researchers are currently incorporating artificial intelligence powered Computer-Aided Diagnosis (CAD) systems into medical imaging to facilitate early diagnosis and intervention by physicians [12]. CNNs have recently demonstrated their potential in accurately classifying various types of arrhythmias and ECG signals [13–16]. The utilization of CNN architecture enables the automatic extraction of ECG signal features in a non-linear manner, providing an objective approach for detecting and classifying heart rhythms, even those that present challenges or might lead to misjudgments when using traditional methods [17].

1.2. Research hypothesis and objectives

In this study, we trained one-dimensional CNNs on randomly selected single-lead ECG data with correct rhythm labels. In contrast to previous literature [18–21], which focused on the extraction of QRS waves from ECG signals for heart rhythm analysis, we employed a CNN to directly analyze single-lead ECG images. Compared with the existing literature, our investigation faces more complex challenges. Specifically, we aim to train a model that can recognize nine different heart rhythms simultaneously. Notably, we added

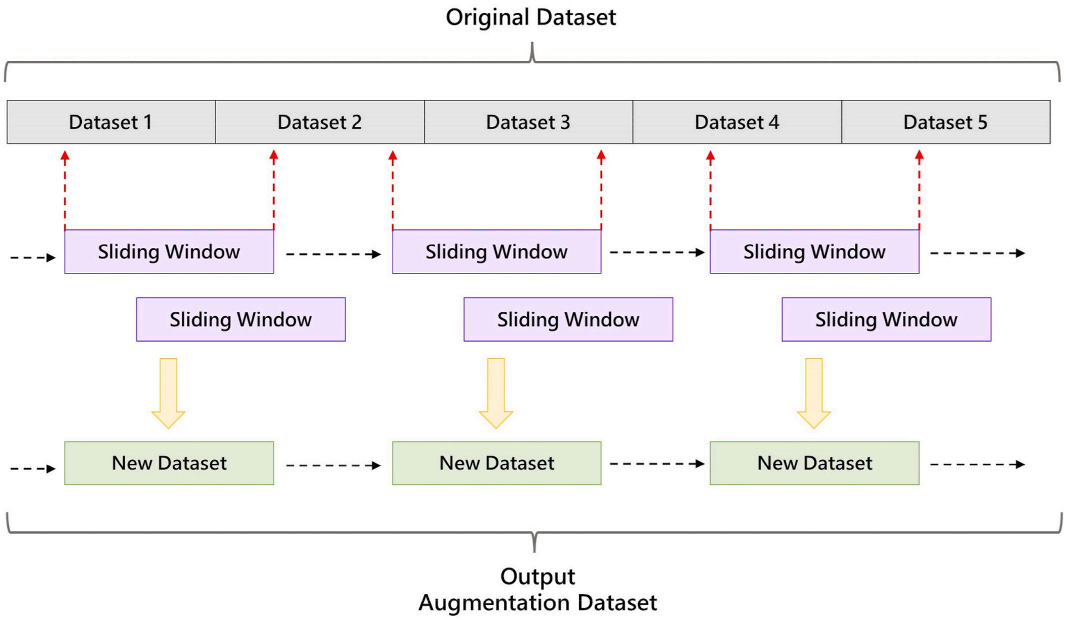

Fig. 1. The method for increasing the number of samples through overlapping sampling illustrates the 'overlap method' used in our experimental design to extract data for rhythms with an adequate number of samples based on the time interval. This method demonstrates how segments were selected and connected to create new samples, with a paragraph containing two concurrent segments designated as a distinct sample.

complexity by including rhythms that exhibit similarities (e.g., VT, VF, and VFL in ventricular arrhythmias), and validated with different datasets. Furthermore, we embarked on two distinct experiments involving model training with varying durations, specifically 5 and 10 s. This selection is rooted in the customary practice of employing a 10-s temporal window in traditional 12-lead ECGs. Our overarching objective was to gauge whether comparable outcomes could be achieved within a truncated timeframe, thereby potentially facilitating early arrhythmia detection in exigent scenarios. Our study aims to determine the model's ability to accurately distinguish rhythms given a sufficient training database, and the ability of CNNs to identify heart rhythms under different time constraints.

2. Materials and methods

2.1. Datasets and samples

We extracted nine heart rhythms from three databases - the MIT-BIH Arrhythmia Database (MITDB), Atrial Fibrillation Database (AFDB), and MIT-BIH Ventricular Fibrillation Database (VFDB) - using the Waveform Database Software Package (WFDB) provided on PhysioNet. All of the rhythms are in single-lead format. The MITDB data was originally sampled at a frequency of 360 Hz, while AFDB and VFDB were sampled at 250 Hz. To standardize the sampling frequency across all three databases, we down sampled the MITDB data to 250 Hz prior to training. Our sample segments do not use the signal as the starting and ending point, which is commonly done in previous experiments. Instead, we have captured a 5-s or 10-s segment of the cardiac rhythm. Direct segmentation was used to extract experimental data for rhythms with sufficient samples based on the time interval. In our experimental design, we employed the "overlap method" as shown in Fig. 1, in which segments were selected and connected in a manner that resulted in the creation of new samples. Specifically, we designated a paragraph encompassing two concurrent segments as a distinct sample. For the 5-s intervals, we used a total of 8302 samples for testing, while for the 10-s intervals, we used 8459 samples. Out of these samples, 5734 were used for training and 2568 for validation for the 5-s intervals, while for the 10-s intervals, 5934 were used for training and 2525 for validation (Table 1).

2.2. CNN architecture

Our model structures referred to a previous CNN-based heart rhythm classification studies [22].The CNN model used 1250 and 2500 input samples for 5- and 10-s inputs, respectively. The first convolutional layer had a size of 27, and the output feature map size was 1224×3 and 2474×3 for the 5- and 10-s inputs, respectively. The Max Pooling size for the entire model was set to 2. After the initial convolutional layer, the model applied Max Pooling with a 24-size convolutional filter. Another Max Pooling was performed after a 15-size convolutional layer, followed by a 10-size convolutional layer and Max Pooling. The final output layer consisted of a fully connected layer with a size of 1250, 10, and 8 for the 5-s input, 10-s input, and classification layers, respectively. A schematic representation of the CNN architecture used in this study is shown in Fig. 2. To train the CNN model, the stochastic gradient descent momentum (sgdm) algorithm was used with a minibatch size of 20 and a validation frequency of 30. The training was conducted for 200 epochs with GPU acceleration to speed up the process. The model was evaluated on a test dataset to measure its classification accuracy, precision and recall.

2.3. External model testing

The Creighton University Ventricular Tachyarrhythmia Database (CUDB) from PhysioNet is comprised of over 7000 ECG recordings featuring various arrhythmias at a sampling rate of 250 Hz, encompassing all the rhythms in our model. The CUDB database contains ECG recordings that were gathered from real-world clinical settings, allowing for a more authentic portrayal of arrhythmia detection in actual practice. In this portion of the experiment, we obtained the complete rhythm from CUDB without using any overlapping sampling techniques. Subsequently, we employed our CNN model to classify these rhythms. The purpose was to use external model to assess how well the CNN model performs on unseen data. This helps us analyze the likelihood of using the model in

Table 1
5-s and 10-s ECG datasets. A total of 8302 samples for the 5-s interval, and 8459 samples for the 10-s interval.

Rhythms	5-s		10-s	
	Training	Validation	Training	V V alidation
NSR	667	333	702	298
AFIB	714	286	693	307
AFL	687	313	712	288
WPW	441	167	675	241
MII	617	259	512	256
SB	715	285	711	289
VT	718	282	698	302
VF	500	318	500	275
VFL	675	325	731	269
Total	5734	2568	5934	2525

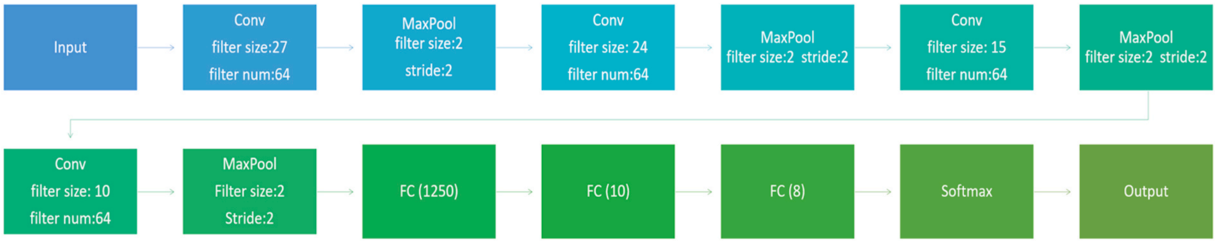

Fig. 2. Schematic diagram of the convolutional neural network model, which includes convolutional layers, Max Pooling operations, and a final fully connected layer for classification.

real-world scenarios and identify potential issues it may face. Additionally, we remain cautious about the possibility of overfitting. As we increase the data using the overlap method, we are mindful of the potential risk of overfitting the model. It is crucial to ensure that the new samples are not too similar to the original ones to avoid bias in the analysis. External model testing can also be utilized to detect overfitting. To ensure the accuracy of the results, the testing samples used were not previously employed in any experiments. A research workflow summarizing all research steps is presented in Fig. 3.

3. Results

3.1. Ten seconds CNN model

The validated results of the 10-s neural network model are presented in Table 2, showcasing remarkable accuracy ranging from 95.3% to 99.8% for each category. Notably, aside from NSR and VF, which displayed sensitivities of 78.8% and 70.9%, respectively, all other seven heart rhythms exhibited sensitivities surpassing the 90% threshold. Furthermore, the specificities for all heart rhythms demonstrated consistently satisfactory performance, ranging between 95% and 99.8%. Upon examining the confusion matrix depicted in Fig. 4, we observed that out of the 80 misclassified VF samples, a remarkable 79 were erroneously categorized as VT, with the remaining 1 sample being classified as SB. Furthermore, among the 63 misclassified NSR samples, a significant proportion, specifically 35 samples, were misclassified as VFL. The remainder of the misclassified samples consisted of 16 cases of AFIB, 1 case of AFL, 4 cases of MII, and 7 cases of VT. Based on these findings, it can be inferred that the diminished sensitivity primarily stems from the misclassification of VF and VFL.

3.2. Five seconds CNN model

The 5-s neural network model demonstrated outstanding performance in accurately identifying all nine types of heart rhythms, consistently achieving high accuracy rates, as summarized in Table 3. Sensitivity values ranged from 89.4% to a perfect 100%, while specificity was consistently excellent, ranging from 98.7% to 100%. Notably, VT exhibited the lowest sensitivity among all categories at 89.4%. Upon closer examination of the confusion matrix presented in Fig. 5, it was observed that instances where VT was misclassified, three cases were identified as AFL, one as MII, and two as SB. Interestingly, there were 27 cases of VT misclassified as VF,

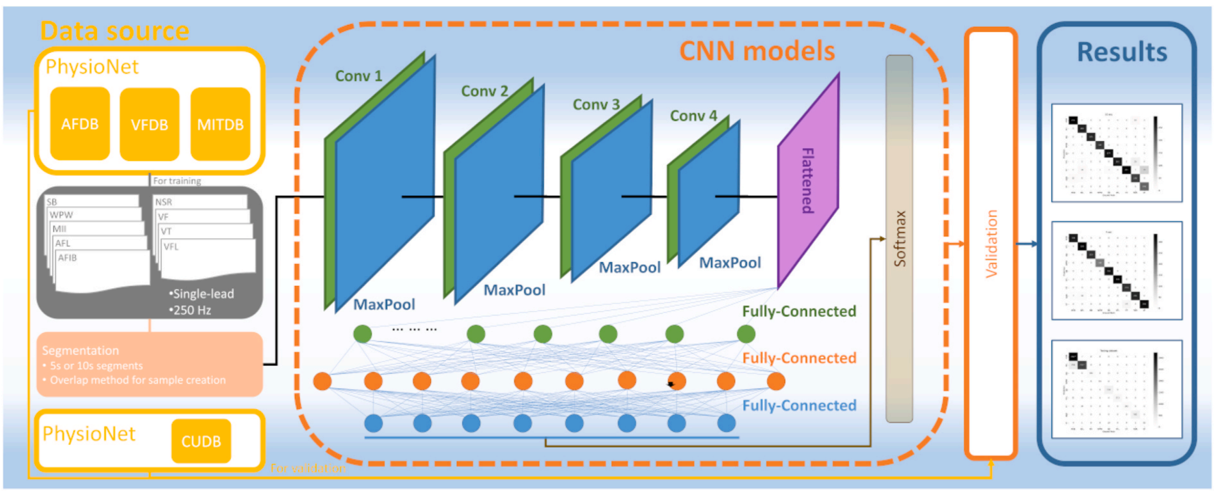

Fig. 3. Overview of the research steps, including data extraction, CNN model architecture, and external model testing using the CUDB for real-world arrhythmia detection assessment.

Table 2
Validated results of the 10-s neural network model, illustrating high accuracy levels ranging from 95.3% to 99.8%. Sensitivities for all rhythms, except NSR and VF, exceeded 90%. Additionally, specificities for all rhythms remained consistently satisfactory.

	AFIB	AFL	MII	WPW	SB	VFL	VT	VF	NSR
ACC	98.7%	98.7%	99.7%	99.8%	99.4%	98.2%	95.3%	96.7%	96.7%
SEN	95.1%	90.2%	99.6%	99.5%	95.8%	99.6%	97.6%	70.9%	78.8%
SPEC	99.2%	99.8%	99.7%	99.8%	99.9%	98.0%	95.0%	99.8%	99.1%

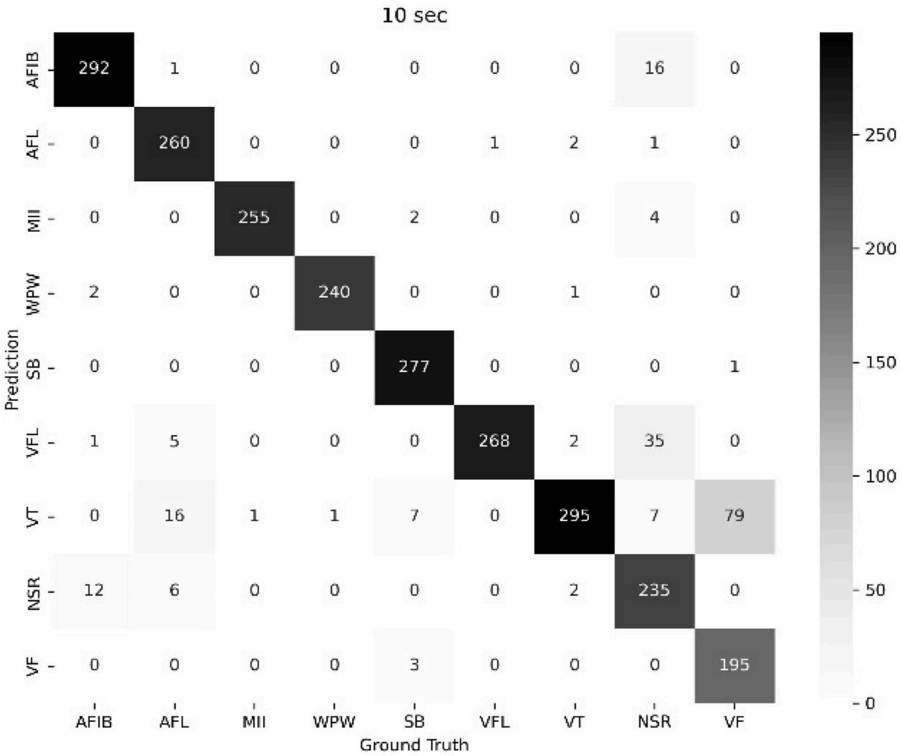

Fig. 4. 10-s confusion matrix displaying misclassifications, with most VF samples classified as VT, while NSR samples are misclassified as VFL and other arrhythmias.

Table 3
The 5-s neural network model accurately presents the identification of all nine rhythms, with sensitivity values ranging from 89.4% to 100%. Specificity consistently remained at high levels, ranging from 98.7% to 100%.

	AFIB	AFL	MII	WPW	SB	VFL	VT	VF	NSR
ACC	99.1%	98.8%	99.6%	99.9%	99.4%	99.7%	98.5%	98.8%	98.9%
SEN	94.4%	94.9%	99.3%	98.9%	96.0%	100%	89.4%	100%	96.6%
SPEC	99.7%	99.3%	99.6%	100%	99.8%	99.7%	99.7%	98.7%	99.1%

indicating a relatively higher rate of false-positive VF detections. Remarkably, the experimental findings suggest that the 5-s model performed as effectively, if not better than, the 10-s model. Although some categories displayed slightly improved sensitivity and specificity in the 5-s model, both neural network models exhibited overall comparable performance, with no statistically significant distinctions.

3.3. Realistic scenario simulation

Given the similarity in performance between the two models, we opted to prioritize the 5-s model as it offers a faster detection of potentially fatal heart rhythms. As such, we utilized the 5-s model for testing against the CUDB database and simulated real-world conditions. The overall average accuracy achieved was 91.6%, with the majority of arrhythmias being correctly classified. Based on the CUDB testing results, the model appears to possess a certain level of classification capability when presented with unseen data.

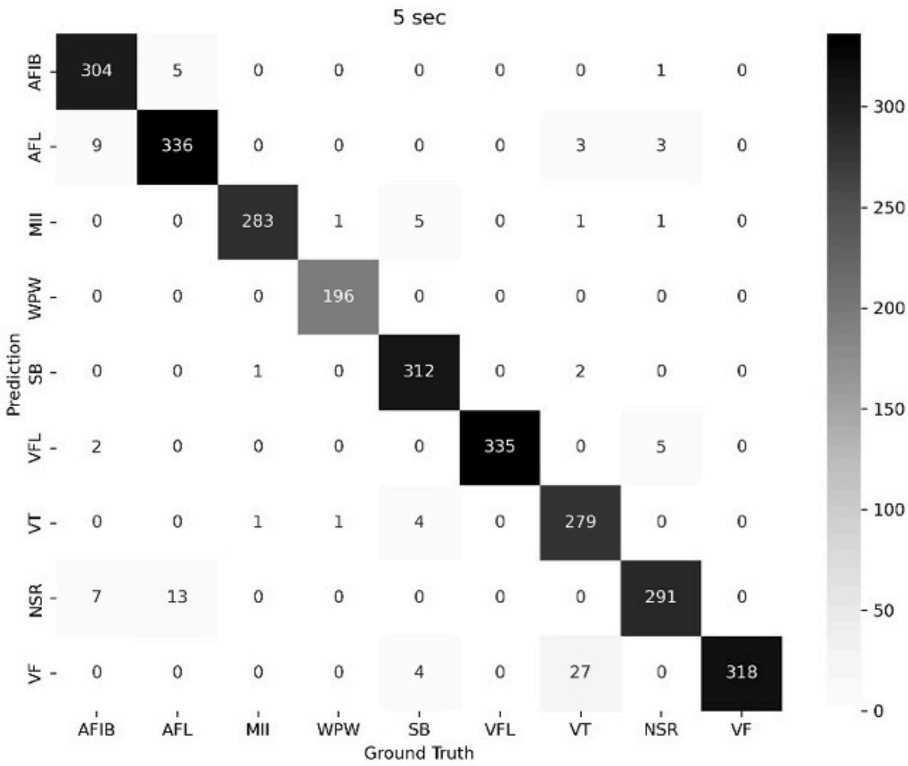

Fig. 5. 5-s confusion matrix illustrates the misclassifications, with VT often misidentified as AFL, MII, SB, and VF, pointing to a notable false-positive rate for VF detections.

Nevertheless, the results from Fig. 6 highlight an important finding in the confusion matrix analysis. The most prevalent group of misclassified samples were AFIB samples being mistakenly labeled as AFL, with a total of 458 instances. In comparison, the other misclassified waveforms only amounted to 184 instances, with the highest being AFL misclassified as VT with only 34 instances. One possible explanation for this high number of misclassified AFIB samples is the uneven distribution of waveforms in the CUDB database. Specifically, the database contains a total of 3747 AFIB samples, making it the most represented waveform, accounting for approximately 49% of all samples. On the other hand, the least represented waveform in the database is WPW, with only 20 instances, all of which were correctly identified by the model. Therefore, it's possible that the imbalanced representation of waveforms in the database contributed to the misclassification of AFIB samples as AFL.

4. Discussion

4.1. Comparison of 5-s and 10-s models in detecting arrhythmias

The experimental findings reveal that both 5-s and 10-s intervals yielded strong performance in arrhythmia detection. The validation accuracies for the 5-s and 10-s models were 96.51% and 98.1%, respectively, signifying high levels of accuracy for both durations. In fact, for most heart rhythms, both models consistently achieved accuracies exceeding 95%. Interestingly, the 10-s model exhibited lower sensitivity in detecting VF and NSR, at 70.9% and 78.8%, respectively. This could be attributed to movement or interference during longer recordings, potentially causing waveform variations. Additionally, the extended interval may capture more intricate heart rhythm dynamics, posing a more challenging pattern recognition task for the model. Fig. 4 highlights that out of the 80 instances where VF was misclassified, a notable 79 were erroneously identified as VT. However, it's important to note that this misclassification doesn't affect clinical practice. According to Advanced Cardiac Life Support (ACLS) guidelines [23], the clinical management of VT and VF patients entails the same immediate defibrillation. Furthermore, the reduced sensitivity in detecting NSR is primarily due to 35 cases being classified as VFL. In practice, VFL is not typically confused with NSR since VFL patients lack a pulse and require immediate cardiopulmonary resuscitation (CPR) [24], while NSR represents a normal heart rhythm. Consequently, such confusion is unlikely in real-world applications. Notably, specificity did not significantly differ between the 5-s and 10-s models. Thus, the results suggest that a 5-s duration may be sufficient for arrhythmia detection. It's important to acknowledge that both models exhibited misclassification errors between VT and VF, partly due to the similarities between polymorphic VT and VF, which can also challenge cardiologists [25].

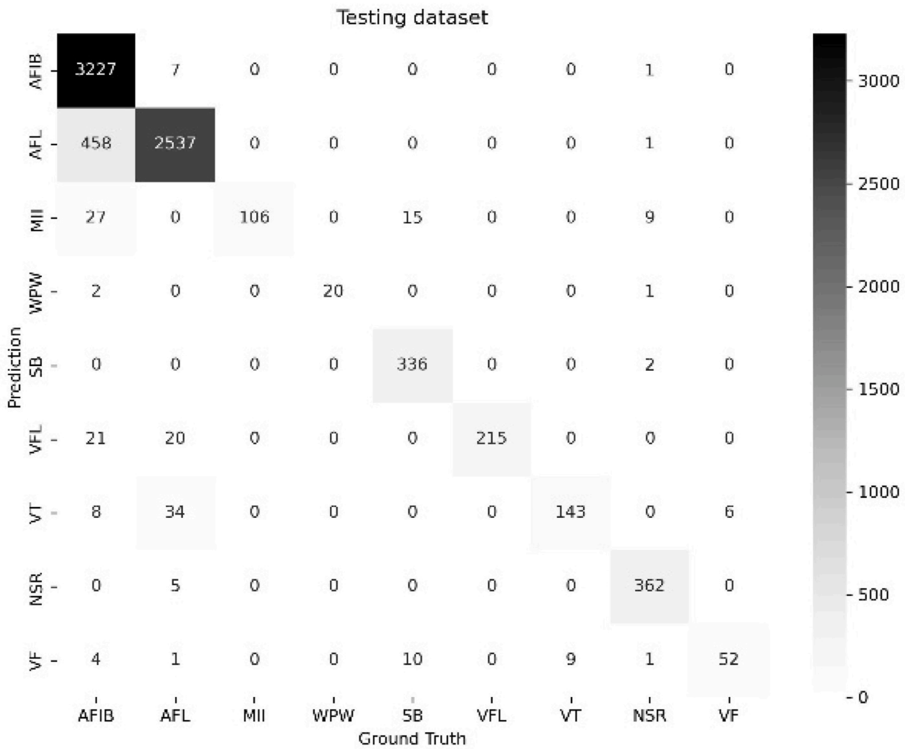

Fig. 6. Confusion matrix highlighting the frequent misclassification of AFIB samples as AFL, totaling 458 instances. In contrast, other waveform misclassifications amounted to only 184 instances, with the highest being AFL mistakenly classified as VT in just 34 instances.

4.2. *The impact of data imbalance on CNNs performance*

In the CUDB test, approximately 6% of cases were incorrectly identified as AFL, resulting in lower sensitivity for AFIB compared to the other eight arrhythmias. One primary reason for this could be the data imbalance in CUDB, where AFL samples constitute approximately 34% of the dataset. Data imbalance is a common challenge in various fields, impacting the training and evaluation of machine learning models [26]. When dealing with imbalanced datasets, characterized by substantially fewer samples in certain classes than others, CNN models encounter difficulties in accurately capturing the genuine data patterns. They tend to exhibit a bias towards the more prevalent samples, potentially resulting in overfitting on the majority class and subsequent suboptimal performance on the minority class. Furthermore, imbalanced data can lead to mislabeling or misclassification of rare minority class samples, further exacerbating dataset imbalances. Balancing the dataset becomes crucial to ensure that CNN models can effectively learn to classify all types of heartbeats, thereby enhancing their performance and reliability. Due to the distinctive characteristics of ECG signals, accurately recording certain types of heartbeats can be inherently more challenging, resulting in an uneven distribution of data. Nevertheless, this disparity is considered acceptable in our experiments, as it is not anticipated to significantly affect clinical utility. One primary rationale for this assertion is our belief that it would not substantially impact clinical practice. AFIB episodes that occur outside of a hospital setting are frequently missed, leading to a missed opportunity for early treatment and an increased risk of long-term cardiovascular disease. However, these episodes seldom pose an immediate threat. Our primary objective is to facilitate early arrhythmia detection through a wearable device, issuing timely alerts irrespective of whether they pertain to AFIB or AFL. This capability empowers users to promptly seek medical attention, and once the doctor confirms the rhythms, appropriate treatment can be administered. In future experiments, it is imperative that we place greater emphasis on the collection and integration of diverse data into these databases to enhance the accuracy of our models for future applications.

4.3. *Comparative analysis models for arrhythmia detection*

When comparing our proposed model with state-of-the-art models validated in the same public dataset, it is important to note that few CNN studies have reported using single lead to identify multiple types of arrhythmias, regardless of whether they utilized open datasets. However, there are some similar experiments, such as those focusing on using single-lead ECG to detect AFIB [27] and using CNN to detect VT and VF [28]. Yun et al. utilized 11 open-source databases from PhysioNet and developed a deep learning model based on transformer for AFIB/AFL segmentation in single lead ECG using self-supervised learning with masked signal modeling [27]. Their model achieved a precision of 0.964, recall of 0.948, and F1 score of 0.956. Gadaleta et al. predicted the occurrence of AFIB in single lead ECG recordings from home settings using a deep learning model [29]. In their experiment, AFIB episodes needed to last for more than 30 s, and the model achieved a sensitivity of 0.80, specificity of 0.65, precision of 0.09, and F1 score of 0.17. Daydulo et al. utilized MIT-BIH and BIDMC databases from PhysioNet and employed ResNet 50 and AlexNet to discern arrhythmias, heart failure, and NSR [30]. Their model achieved an accuracy of 0.992, sensitivity of 0.992, specificity of 0.996, and an average recall, precision, and F1 score of 0.992. It is worth noting that these arrhythmias were not further classified. A. Mjahad et al..used MIT-BIH and AHA databases and employed four different CNN architectures (InceptionV3, MobileNet, VGGNet, and AlexNet) to discern VF and VT features [28]. The results showed a sensitivity of 98.16% and specificity of 99.07% for VF, sensitivity of 90.45% and specificity of 99.73% for VT, sensitivity of 99.34% and specificity of 98.35% for NSR, and sensitivity of 96.98% and specificity of 99.68% for other rhythms, with corresponding accuracies. A. Argha et al. utilized the PhysioNet/CinC Challenge 2017 dataset and developed a hybrid deep learning model capable of classifying ECG recordings into four classes by detected QRS wave: NSR, AFIB, other rhythms, and too noisy recordings [31]. Their proposed model achieved an average test F1-score of 0.892. We want to emphasize the uniqueness of our model, which utilizes a single lead and does not require the extraction of QRS wave features, enabling the detection of up to 8 types of arrhythmias within 5 s. Furthermore, we have achieved performance comparable to previous experimental models (Table 4.).

4.4. The role of single-lead ECG detection

The 12-lead ECG is valuable for diagnosing cardiovascular diseases, like myocardial infarction, in specific cardiac regions. However, in our experiment, we opted for a single-lead approach for machine learning on nine different heart rhythms. This choice is grounded in the fact that these rhythms don't produce unique abnormal waveforms confined to a single lead. When arrhythmias occur, the waveforms appear consistently across all leads. This characteristic enables arrhythmia detection using a wearable device equipped with a single-lead sensor, providing real-time heartbeat monitoring. Presently, several companies have developed single-lead ECG wearable devices. However, their widespread use in arrhythmia detection remains limited, primarily due to FDA approval being granted solely for the detection of atrial fibrillation (AFIB) [32]. Large-scale studies, such as the Apple Watch study and the Huawei heart study, have assessed AFIB screening using wearable devices and involved extensive participant cohorts [33,34].These studies reported satisfactory positive predictive values. Arrhythmias, especially ventricular types, can pose a risk of sudden death. Nevertheless, some individuals with inherited heart conditions, like Brugada syndrome, may naturally revert to sinus rhythm after a brief period of NSVT [35]. Early detection of such situations through single-lead wearable devices can promptly trigger alerts for medical intervention, facilitating timely treatment. This capability underscores the potential of wearable devices in improving patient outcomes. Our study has demonstrated that the application of single-lead ECG for arrhythmia detection goes beyond its current use in AFIB alone. Moreover, harnessing the capabilities of CNN enables the swift identification of various arrhythmias, encompassing ventricular arrhythmias and bradycardias. Additionally, our experiments have highlighted the potential for CNN to deliver rapid assessments within a mere 5-s timeframe, which holds promise for the future advancement of wearable devices designed for heart rhythm monitoring. These findings highlight the potential of single-lead ECG devices in arrhythmia detection. However, wider adoption of these devices for arrhythmia screening and monitoring requires regulatory approvals and further validation in large-scale clinical studies. Collaborations between researchers, healthcare providers, and regulatory bodies will be essential in advancing the use of single-lead wearable ECG devices for improved patient outcomes.

4.5. Limitataions

In our experiment, we selected heart rhythm data from the AFDB, VFDB, MITDB, and CUDB databases on PhysioNet, covering a wide spectrum of Supraventricular arrhythmias, Ventricular arrhythmias, and bradycardias. However, it's important to acknowledge that certain other arrhythmias, such as Paroxysmal supraventricular tachycardia (PSVT), Premature ventricular beats (PVCs), and sick sinus syndromes beyond Mobitz type II were not included in this study. While our experiment validated CNN's high accuracy in identifying these selected arrhythmias, future clinical practice may benefit from incorporating these additional arrhythmias into our research to enhance its clinical relevance. One limitation of ECG signals is their vulnerability to noise interference [36]. Even minor movements can cause a baseline drift that appears as low-frequency interference in the signal. Various sources of noise, including electromagnetic interference, muscle artifacts, and environmental factors such as temperature and humidity, can cause this drift. Additionally, the magnitude of the measured electric potential can vary with electrode placement, resulting in variations in recorded signals. The complex structure of ECG signals makes them difficult to interpret accurately. They often exhibit high levels of noise and complexity, leading to false-positive or false-negative diagnoses. Correctly interpreting ECG waveforms that have been affected by interference can be challenging, and labeling certain conditions may require a significant amount of time and effort. To develop our CNN model, we used the publicly available database PhysioNet, which provided a valuable resource for comparing our results with established standards. However, using clinical data for model development and testing can be challenging, as it requires precise labeling of the data and accounting for the effects of noise and artifacts [37].However, the samples were from public databases, in which ECG images may have already been screened or standardized. In future studies, the utility of the model will be tested by validating clinical ECG results using sufficient data.

5. Conclusions

Arrhythmias pose an immediate risk of fatality or can elevate the risk of cardiovascular diseases over time. CNN has emerged as a promising tool for automated arrhythmia detection from ECG signals. Detecting arrhythmias through the single-lead devices could enable early intervention for asymptomatic cases and enhance prehospital rescue chances for life-threatening rhythms, ultimately

Table 4
Comparison of performance metrics between our 5-s model and previous models.

	Sensitivity	Specificity	Precision	F1 score
Yun et al.	0.948	–	0.964	0.956
Gadaleta et al.	0.8	0.65	0.09	0.17
Daydulo et al.	0.992	0.996	–	0.992
Mjahad et al.	0.981	0.991	–	–
Argha et al.	–	–	–	0.892
Our study	0.965	0.997	0.991	0.983

improving patient survival rates. This model aims to be applied using a single-lead monitor in the clinical care of critically ill patients. With the acquisition of additional data and results in the future, there is a possibility for further development into a wearable device for out-of-hospital patients. Preventive medicine aims to utilize wearable devices for vital sign monitoring Developing an effective CNN capable of analyzing intricate heart rhythms in individual patients will lead to a cost-effective, automated, and safe arrhythmia detection and monitoring system. While addressing CNN limitations and enhancing interpretability through further research is essential, the potential for CNNs to transform ECG analysis and enhance patient outcomes is evident.

Funding

This research was funded by the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. Grant No. DP2-TMU-112-N-06 and the National Science and Technology Council, Taiwan, under Grant No. 112-2221-E−038-008. Hung-Wen Chiu has received research support from MOE and National Science and Technology Council. Ming-Feng Tsai has received research support from National Science and Technology Council.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

As the article does not involve Human participants, thus, there is no such informed consent.

CRediT authorship contribution statement

Liong-Rung Liu: Writing – original draft. **Ming-Yuan Huang:** Conceptualization. **Shu-Tien Huang:** Writing – review & editing. **Lu-Chih Kung:** Formal analysis. **Chao-hsiung Lee:** Validation. **Wen-Teng Yao:** Resources. **Ming-Feng Tsai:** Funding acquisition. **Cheng-Hung Hsu:** Methodology. **Yu-Chang Chu:** Visualization. **Fei-Hung Hung:** Visualization. **Hung-Wen Chiu:** Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hung-Wen Chiu reports financial support was provided by Ministry of Education. Ming-Feng Tsai reports financial support was provided by National Science and Technology Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We gratefully acknowledge the invaluable resources and support provided by PhysioNet, which have been essential to the completion of this work.

References

[1] G.A. Roth, et al., Global Burden of cardiovascular diseases and risk factors, 1990-2019: Update from the GBD 2019 study, J. Am. Coll. Cardiol. 76 (25) (Dec. 2020) 2982–3021, https://doi.org/10.1016/j.jacc.2020.11.010.
[2] T.O. Mene-Afejuku, A.O. Bamgboje, M.O. Ogunniyi, O. Akinboboye, U.N. Ibebuogu, Ventricular arrhythmias in Seniors with heart failure: present Dilemmas and Therapeutic Considerations: a Systematic review, Curr. Cardiol. Rev. 18 (2) (2022) e181021197279, https://doi.org/10.2174/1573403X17666211018095324.
[3] P. Gouda, P. Brown, B.H. Rowe, F.A. McAlister, J.A. Ezekowitz, Insights into the importance of the electrocardiogram in patients with acute heart failure, Eur. J. Heart Fail. 18 (8) (2016) 1032–1040, https://doi.org/10.1002/ejhf.561.
[4] A. Bansal, R. Joshi, Portable out-of-hospital electrocardiography: a review of current technologies, J. arrhythmia 34 (2) (Apr. 2018) 129–138, https://doi.org/10.1002/joa3.12035.
[5] J.E. Marine, Nonsustained ventricular tachycardia in the normal heart: risk Stratification and management, Card. Electrophysiol. Clin. 8 (3) (Sep. 2016) 525–543, https://doi.org/10.1016/j.ccep.2016.04.003.

[6] D.J. Miller, K. Shah, S. Modi, A. Mahajan, S. Zahoor, M. Affan, The Evolution and application of cardiac monitoring for Occult atrial fibrillation in Cryptogenic Stroke and TIA, Curr. Treat. Options Neurol. 18 (4) (2016) 17, https://doi.org/10.1007/s11940-016-0400-y.
[7] S.A. Lubitz, et al., Rationale and design of a large population study to validate software for the assessment of atrial fibrillation from data acquired by a consumer tracker or smartwatch: the Fitbit heart study, Am. Heart J. 238 (Aug. 2021) 16–26, https://doi.org/10.1016/j.ahj.2021.04.003.
[8] F. Sana, E.M. Isselbacher, J.P. Singh, E.K. Heist, B. Pathik, A.A. Armoundas, Wearable devices for ambulatory cardiac monitoring: JACC state-of-the-art review, J. Am. Coll. Cardiol. 75 (13) (Apr. 2020) 1582–1592, https://doi.org/10.1016/j.jacc.2020.01.046.
[9] D. Marinucci, A. Sbrollini, I. Marcantoni, M. Morettini, C.A. Swenne, L. Burattini, Artificial neural network for atrial fibrillation identification in portable devices, Sensors 20 (12) (2020), https://doi.org/10.3390/s20123570.
[10] P.-H. Chan, et al., Head-to-Head comparison of the AliveCor heart monitor and Microlife WatchBP Office AFIB for atrial fibrillation screening in a primary care setting, Circulation 135 (1) (Jan. 2017) 110–112, https://doi.org/10.1161/CIRCULATIONAHA.116.024439. United States.
[11] L. Liong-Rung, et al., Using artificial intelligence to establish chest X-Ray image recognition model to assist crucial diagnosis in Elder patients with Dyspnea, Front. Med. 9 (2022) 893208, https://doi.org/10.3389/fmed.2022.893208.
[12] S.-T. Huang, et al., Deep learning-based clinical Wound image analysis using a Mask R-CNN architecture, J. Med. Biol. Eng. (2023), https://doi.org/10.1007/s40846-023-00802-2.
[13] M. Hammad, S. Meshoul, P. Dziwiński, P. Pławiak, I.A. Elgendy, Efficient Lightweight Multimodel deep Fusion based on ECG for arrhythmia classification, Sensors 22 (23) (2022), https://doi.org/10.3390/s22239347.
[14] S. Somani, et al., Deep learning and the electrocardiogram: review of the current state-of-the-art, Eur. Eur. pacing, arrhythmias, Card. Electrophysiol. J. Work. groups Card. pacing, arrhythmias, Card. Cell. Electrophysiol. Eur. Soc. Cardiol. 23 (8) (Aug. 2021) 1179–1191, https://doi.org/10.1093/europace/euaa377.
[15] X. Liu, H. Wang, Z. Li, L. Qin, Deep learning in ECG diagnosis: a review, Knowledge-Based Syst. 227 (2021) 107187, https://doi.org/10.1016/j.knosys.2021.107187.
[16] Y. Chen, C. Zhang, C. Liu, Y. Wang, X. Wan, Atrial fibrillation detection using a Feedforward neural network, J. Med. Biol. Eng. 42 (1) (2022) 63–73, https://doi.org/10.1007/s40846-022-00681-z.
[17] Y. Ji, S. Zhang, W. Xiao, Electrocardiogram classification based on faster regions with convolutional neural network, Sensors 19 (11) (2019), https://doi.org/10.3390/s19112558.
[18] M. Hammad, et al., Automated detection of myocardial infarction and heart conduction disorders based on feature selection and a deep learning model, Sensors 22 (17) (2022), https://doi.org/10.3390/s22176503.
[19] P. de Chazal, R.B. Reilly, A patient-Adapting heartbeat classifier using ECG Morphology and heartbeat interval features, IEEE Trans. Biomed. Eng. 53 (12) (2006) 2535–2543, https://doi.org/10.1109/TBME.2006.883802.
[20] T.-R. Wei, S. Lu, Y. Yan, Automated atrial fibrillation detection with ECG, Bioeng. (Basel, Switzerland) 9 (10) (2022), https://doi.org/10.3390/bioengineering9100523.
[21] S. Banerjee, Identification of elevated ST segment and deep Q type MI variant using Cross wavelet transform and Hierarchical classification from ECG signals, J. Med. Biol. Eng. 37 (4) (2017) 492–507, https://doi.org/10.1007/s40846-017-0280-y.
[22] U.R. Acharya, H. Fujita, O.S. Lih, Y. Hagiwara, J.H. Tan, M. Adam, Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network, Inf. Sci. 405 (2017) 81–90, https://doi.org/10.1016/j.ins.2017.04.012.
[23] K. Hanneken, D. Gaieski, A. Vempati, R. Hall, Extracorporeal Membrane Oxygenation (ECMO) for Refractory cardiac arrest, J. Educ. Teach. Emerg. Med. 5 (4) (Oct. 2020) S28–S58, https://doi.org/10.21980/J88W69.
[24] C. Hawkes, et al., Epidemiology and outcomes from out-of-hospital cardiac arrests in England, Resuscitation 110 (Jan. 2017) 133–140, https://doi.org/10.1016/j.resuscitation.2016.10.030.
[25] A.Y. Hannun, et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (1) (2019) 65–69, https://doi.org/10.1038/s41591-018-0268-3.
[26] A.W. Olthof, P.M.A. van Ooijen, L.J. Cornelissen, Deep learning-based natural Language processing in Radiology: the impact of report complexity, disease prevalence, dataset size, and algorithm type on model performance, J. Med. Syst. 45 (10) (2021) 91, https://doi.org/10.1007/s10916-021-01761-4.
[27] D. Yun, et al., Automatic segmentation of atrial fibrillation and flutter in single-lead electrocardiograms by self-supervised learning and Transformer architecture, J. Am. Med. Inform. Assoc. 31 (1) (Dec. 2023) 79–88, https://doi.org/10.1093/jamia/ocad219.
[28] A. Mjahad, M. Saban, H. Azarmdel, A. Rosado-Muñoz, Efficient extraction of deep image features using a convolutional neural network (CNN) for detecting ventricular fibrillation and tachycardia, Journal of Imaging 9 (9) (2023), https://doi.org/10.3390/jimaging9090190.
[29] M. Gadaleta, et al., Prediction of atrial fibrillation from at-home single-lead ECG signals without arrhythmias, npj Digit. Med. 6 (1) (2023) 229, https://doi.org/10.1038/s41746-023-00966-w.
[30] Y.D. Daydulo, B.L. Thamineni, A.A. Dawud, Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals, BMC Med. Inform. Decis. Mak. 23 (1) (2023) 232, https://doi.org/10.1186/s12911-023-02326-w.
[31] A. Argha, et al., Assessing the Generalizability of a deep learning-based automated atrial fibrillation algorithm, Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf. 2023 (2023) 1–6, https://doi.org/10.1109/EMBC40787.2023.10341108. Jul.
[32] N. Isakadze, S.S. Martin, How useful is the smartwatch ECG? Trends Cardiovasc. Med. 30 (7) (Oct. 2020) 442–448, https://doi.org/10.1016/j.tcm.2019.10.010.
[33] M. V Perez, et al., Large-scale assessment of a smartwatch to identify atrial fibrillation, N. Engl. J. Med. 381 (20) (Nov. 2019) 1909–1917, https://doi.org/10.1056/NEJMoa1901183.
[34] Y. Guo, et al., Mobile Photoplethysmographic Technology to detect atrial fibrillation, J. Am. Coll. Cardiol. 74 (19) (Nov. 2019) 2365–2375, https://doi.org/10.1016/j.jacc.2019.08.019.
[35] A.D. Krahn, E.R. Behr, R. Hamilton, V. Probst, Z. Laksman, H.-C. Han, Brugada syndrome, JACC. Clin. Electrophysiol. 8 (3) (Mar. 2022) 386–405, https://doi.org/10.1016/j.jacep.2021.12.001.
[36] F. Fu, et al., Comparison of machine learning algorithms for the Quality assessment of wearable ECG signals via Lenovo H3 devices, J. Med. Biol. Eng. 41 (2) (2021) 231–240, https://doi.org/10.1007/s40846-020-00588-7.
[37] R.S. Andersen, A. Peimankar, S. Puthusserypady, A deep learning approach for real-time detection of atrial fibrillation, Expert Syst. Appl. 115 (2019) 465–473, https://doi.org/10.1016/j.eswa.2018.08.011.