AUTOMATED HEART ARRHYTHMIA DETECTION FROM ELECTROCARDIOGRAPHIC DATA

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

College of Engineering and Science

Victoria University

by

Jinyuan He

August 2020

ABSTRACT

AUTOMATED HEART ARRHYTHMIA DETECTION FROM ELECTROCARDIOGRAPHIC DATA

Jinyuan He, Ph.D.

Victoria University 2020

Heart arrhythmia is a severe heart problem, which threatens people's lives by preventing their hearts from pumping enough blood into vital organs. Arrhythmia has been a major worldwide health problem for years, accounting for nearly 12% of global deaths every year. The research of automated heartbeat classification is highly demanded, which provides a cost-effective screening for heart arrhythmia and allows at-risk patients to receive timely treatments. To construct an effective automated heartbeat classification model from ECG recordings for arrhythmia detection, several key challenges must be addressed, including data quality, heartbeat segmentation range, data imbalance problem, intra and inter-patients variations, identification of supraventricular ectopic heartbeats from normal heartbeats, and model interpretability. This thesis comprehensively discusses these challenges and proposes four practical models to gradually tackle the heartbeat classification task.

Specifically, in Chapter 3, a model named D-ECG is proposed to solve the problems suffered by previous methods of applying a standalone classifier and using a static feature set to classify all heartbeat types. D-ECG introduces the *dynamic ensemble selection* techniques in heartbeat classification for the first time and incorporates a result regulator to improve the disease heartbeats detection performance. Although the dynamic ensemble selection technique has introduced visible improvements in the heartbeat classification task, they also brought some

disadvantages. The dynamic selection nature, which determines the best classifiers according to the sample to be predicted, can result in a delay of the model prediction, making the model less practical in online detection scenarios. In Chapter 4, the author proposes a novel pyramid-like model to tackle this problem. The model adopts a dual-channel classification strategy and customizes a binary classification algorithm that takes neighbor-related information into account to assist disease heartbeats detection. Compared to the D-ECG framework, the pyramid-like model can provide more timely response to an unknown heartbeat while maintaining a good classification performance as the D-ECG framework. It has the potential to be applied in online detection scenarios.

In Chapter 5, the author examines the recent advances brought by deep neural networks and proposes a DNN-based solution named Multi-channels Convolution Neural Network (MCHCNN) to solve the problems of current deep-learning based heartbeat classification models. As an improvement, the proposed network accepts raw ECG heartbeat and heart rhythm (RR-intervals) as inputs and uses different sizes of convolution filters in parallel to capture temporal and frequency patterns from ECG signals. The experimental results have shown visible improvements brought by MCHCNN. However, there is still a long way before MCHCNN can make practical impacts because its performance of S-type heartbeats detection is still relatively low. To tackle this problem, the author investigates the potential causes to the problem and proposes an advanced two-step DNN-based classification framework in Chapter 6. Due to the observed difficulty of detecting S-type heartbeats from N-type heartbeats, the proposed framework trains a deep dual-channel convolutional neural network (DDCNN) which accepts segmented heartbeats as input in the first step to classify V-type, F-type and Q-type heartbeats. At this stage, S-type and N-type heartbeats are not the targets, so they are put into one bundle to be studied in the next step. In the second step, a central-towards LSTM supportive model (CLSM) is specially designed to distinguish S-type heartbeats from N-type ones. The RR-intervals of a heartbeat and its neighbors are arranged in sequence form, serving as the input to CLSM. In particular, CLSM learns and extracts hidden temporal dependency between heartbeats by processing the input RR-interval sequence in central-towards directions. Instead of using raw individual RR-intervals, the abstractive, mutual-connected temporal information provides stronger and more stable support for identifying the problematic S-type heartbeats. Besides, as an improvement as well as a necessary driver for activating the CLSM, a rule-based data augmentation method is also proposed to supply high-quality synthetic samples for the under-represented S-type RR-interval sequences.

Extensive experiments are conducted to provide a comprehensive evaluation for each proposed model. The results prove that the research of heartbeat classification presented in this thesis brings practical ideas and solutions to the arrhythmia detection problem.

DOCTOR OF PHILOSOPHY DECLARATION

I, Jinyuan He, declare that the PhD thesis entitled *Automated Heart Arrhythmia Detection from Electrocardiographic Data* is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature [redacted] Date 02/02/2020

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere thanks to my principal supervisor, Prof. Hua Wang, for his excellent supervision and continuous support. I am also grateful to my associate supervisor, Prof. Yanchun Zhang, who is an expert in *Applied Informatics* and *Medical Data Mining*. He inspired my current research and provided an opportunity for me to join his research team. I am indebted to Dr. Jia Rong who is also my associate supervisor. She shared with me her tremendous experience on how to do research and gracefully gave me encouragement, guidance and invaluable advice to support my work.

Besides my supervisors, I would like to thank my lab mates, Dr. Le Sun, Dr. Ye Wang, Dr. Rui Zhou, Dr. Jiangang Ma, PhD Candidates Alice, Peng Zhang, Xinyu Cao, and Jiahua Du, and Mr. Chuanchuan Zheng. I would like to thank all members and visitors in our research team.

Last but not least, I would like to express my deepest gratitude to my family and friends. Without their warm love and endless support, this dissertation would not have been possible.

PUBLICATIONS

1. **Jinyuan He**, Jia Rong, Le Sun, Hua Wang and Yanchun Zhang. ”An Advanced Two-step Deep Neural NetworkBased Classification Model for Arrhythmia Detection”. The 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) (2020) (**Accepted**)
2. **Jinyuan He**, Jia Rong, Le Sun, Hua Wang and Yanchun Zhang. ”Cardiac Arrhythmia Detection from IoT-based ECGs Using Ensemble Methods”. World Wide Web (2020). (**Accepted**)
3. **Jinyuan He**, Le Sun. ”An extensible framework for ECG anomaly detection in wireless body sensor monitoring systems”. IJSNet, 2019, 29 (2), 101-110.
4. Le Sun, **Jinyuan He**, Chen Wang, Hai Dong, Jiangang Ma and Yanchun Zhang. ”Survey of Cloud SLA Assurance in Pre-interaction and Post-interaction Start Time Phases”. Journal of Computers, 2019, 30 (1), 23-30.
5. Le Sun, **Jinyuan He**, Chen Wang, Jiangang Ma, Hai Dong and Yanchun Zhang. ”A Framework for Discovering Variable-length Motifs in Medical Data Streams”. Journal of Computers, 2019, 30 (1), 105-113.
6. **Jinyuan He**, Jia Rong, Le Sun, Hua Wang and Yanchun Zhang. ”A pyramid-like model for heartbeat classification from ECG recordings”. PloS one, 2018, 13 (11).
7. **Jinyuan He**, Jia Rong, Le Sun, Hua Wang, Yanchun Zhang and Jiangang Ma. ”D-ECG: a dynamic framework for cardiac arrhythmia detection from IoT-based ECGs”. International Conference on Web Information Systems Engineering. Springer, Cham, 2018: 85-99.
8. Le Sun, **Jinyuan He**, Xiaoxia Yin, Yanchun Zhang, Jeon-Hor Chen, Tomas Kron and Min-Ying Su. ”An image segmentation framework for extract-

ing tumors from breast magnetic resonance images". Journal of Innovative Optical Health Sciences, 2018, 11 (04), 1850014.

9. Le Sun, Chen Wang, **Jinyuan He**, Hai Dong, Jiangang Ma and Yanchun Zhang. "User-oriented Cloud SLA Assurance Framework". Journal of Computers, 2018, 29 (6), 235-242.

10. Le Sun, **Jinyuan He**, Jiangang Ma, Hai Dong and Yanchun Zhang. "Limited-length suffix-array-based method for variable-length motif discovery in time series". Journal of Internet Technology, 2018, 19 (6), 1841-1851.

11. **Jinyuan He** and Le Sun. "Study of Different Types of Fuzzy Sets and Fuzzy Decision Making Methods". 2018 1st International Cognitive Cities Conference (IC3), 92-97.

12. Le Sun and **Jinyuan He**. "Interactive Construction of Criterion Relations for Multi-criteria Decision Making". International Conference on Cloud Computing and Security. Springer, Cham, 2018: 589-600.

13. Le Sun and **Jinyuan He**. "Criteria Interdependence in Fuzzy Multi-criteria Decision Making: A Survey". International Conference on Cloud Computing and Security. Springer, Cham, 2018: 397-407.

14. Le Sun, **Jinyuan He**, Xiaoxia Yin, Yanchun Zhang, Jeon-Hor Chen, Tomas Kron and Min-Ying Su. "Deep Learning based Anomaly Detection". Journal of Computers, 2018, 29 (6), 148-157.

15. **Jinyuan He** and Le Sun. "An Interactive Service Composition Model Based on Interactive POMDP". 2018 1st International Cognitive Cities Conference (IC3), 103-108.

16. **Jinyuan He** and Le Sun. "A Review on SLA-Related Applications in Cloud

Computing". 2018 1st International Cognitive Cities Conference (IC3), 87-91.

17. **Jinyuan He** and Le Sun. "An Interactive Interpretive Structural Modeling for Determining Criteria Relations". 2018 1st International Cognitive Cities Conference (IC3), 98-102.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

As a powerful tool to discover values from data, artificial intelligence has attracted an increasing number of attentions in the era of big data. Currently, the research and applications of artificial intelligence have covered a wide range of industry fields, including Healthcare [78], Education [117], Agriculture [110], Government Management [132], Finance and Economics [43], Military [161], Automotive industry [171], etc. The field of artificial intelligence moves extremely quickly. It has unconsciously changed our lives, helping us to be healthier, happier, more productive and more creative. In this thesis, the author presents another medical application of artificial intelligence: automated heart arrhythmia detection from electrocardiographic data. The thesis starts by introducing the research background and motivation, followed by clarification of the research problems. After that, the author presents all the efforts that have been made to tackle the research problems in details. Finally, the thesis is summarized with the main contributions and future research plans.

1.1 Research Background and Motivation

Heart arrhythmia, also known as irregular heart rhythms, is a group of conditions in which the body electrical impulse varies from the normal sequence, causing the heart to beat erratically (too fast or too slowly). Heart arrhythmia has been a major worldwide health problem for years. It threatens people's lives by preventing their hearts from pumping enough blood into the vital organs. Statistics [133] show that about 12% of deaths around the world are because of heart arrhythmia each year.

Arrhythmia can occur at all age groups. Early detection and timely treatment are the keys to survival from arrhythmia. Since arrhythmia may not cause any noticeable symptoms or signs, it is difficult for people to realize whether they have been affected by the heart problem by themselves. The electrocardiogram (ECG) plays a pivotal role in the diagnosis of arrhythmia because ECG captures heart rate, rhythm, and vital information regarding the electrical heart activities and related conditions. Clinically, detection of arrhythmia is conducted by interpretation of individual heartbeats in patients' ECG recordings. However, the manual interpretation of ECG recordings is time-consuming and error-prone, especially for the long-term ECG recording, which is essential for capturing the sporadically occurred arrhythmia [209]. Therefore, an automated method to assist clinicians in detecting arrhythmia heartbeats from ECG is highly demanded.

1.2 Challenges

Heartbeat classification on ECG is a core step towards identifying arrhythmia. As reported by the Association for Advancement of Medical Instrumentation (AAMI) [8], there are 15 original types of heartbeats which are further categorized into five super classes: Normal (N), Supraventricular (S) ectopic, Ventricular (V) ectopic, Fusion (F) and Unknown (Q). In particular, problematic arrhythmias are mostly found in $\boldsymbol{S}$**-type** and $\boldsymbol{V}$**-type** heartbeats [38]. Table 1.1 presents the 15 types and the hierarchy of the 5 super classes. The author presents several ECG samples of different heartbeat classes in Fig.1.1. It can be observed that the V-type heartbeat exhibits a huge morphological difference against other heartbeats, while the *normal* (N-type) and the S-type heartbeats are similar in shape.

To construct an effective automated heartbeat classification model from ECG recordings for arrhythmia detection, several key challenges must be addressed. The first challenge is data quality. In real-world practices, ECG signals usually come

Table 1.1: ECG heartbeat types

AAMI class	Original class	Type of beat
Normal (N)	N	Normal beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
	e	Atrial escape beat
	j	Nodal (junctional) escape beat
Supraventricular ectopic beat (S)	A	Atrial premature beat
	a	Aberrated atrial premature beat
	J	Nodal (junctional) premature beat
	S	Supraventricular premature beat
Ventricular ectopic beat (V)	V	premature ventricular contraction
	E	Ventricular escape beat
Fusion beat (F)	F	Fusion of ventricular and normal beat
Unknown beat (Q)	/	Paced beat
	f	Fusion of paced and normal beat
	Q	Unclassifiable beat

with serious background noise and baseline wanders. Baseline wanders is the effect that the base axis (X-axis) of individual heartbeats appear to move up or down, rather than being straight all the time. Fig.1.2 shows the baseline wanders effect. Besides, there may be also some spikes due to the sensing error that could mess up the signal normalization process. The quality of ECG recordings has a direct impact on the heartbeat classification performance.

The second challenge is heartbeat segmentation. Since classification is performed on individual heartbeats, the original ECG recordings need to be segmented to a collection of successive heartbeats. Fig. 1.3 presents the fiducial points and key intervals of a normal heartbeat. The R peak is a good indicator to locate a heartbeat and the Pan-Tompkins algorithm [146] provides an accurate method for identifying R peak locations of ECG recordings, which save some troubles. However, it is still needed to define a proper range of a heartbeat in consideration of the designed classification methods. A too large range will lead to cross-overs of successive heartbeats and increase the computation burden, whereas a too small

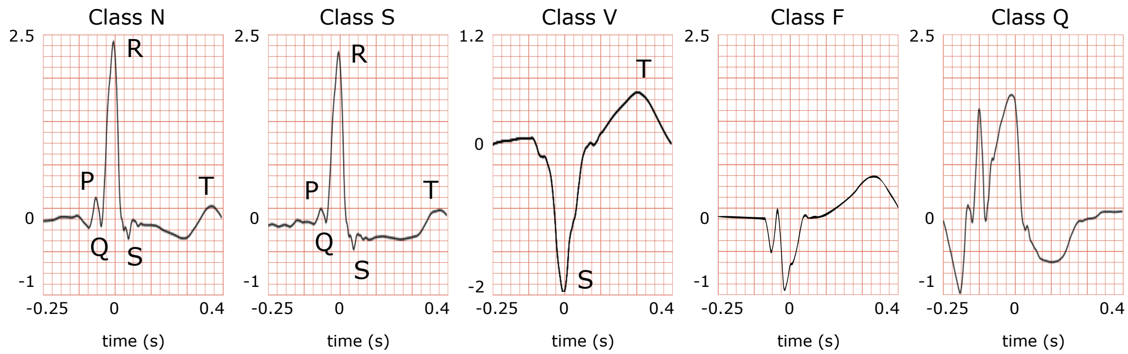


Figure 1.1: Examples of different types of heartbeats. Letters indicate the P-waves, R-peaks, QRS-complexes and T-waves, corresponding to their references in the medical literature. Time gap between two successive R peaks is known as RR-interval. Specifically, **previous-*RR*-interval** denotes the interval between the current R peak and the previous R peak. In comparison to the normal heartbeat (class N), the S-type heartbeat has a less obvious P-wave which is due to *junctional premature beating*. The V-type heartbeat exhibits a deep and capacious S-wave caused by *left bundle branch block*. Class F is a fusion of paced and normal heartbeats. The unclassifiable beat is denoted as class Q.

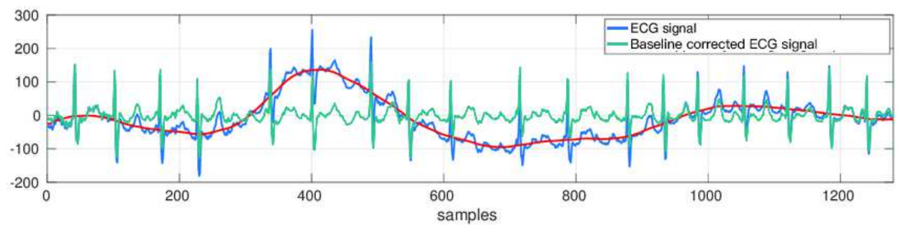


Figure 1.2: The baseline wanders effect.

range will result in the exclusion of key fluctuations, causing impacts to classify heartbeats.

The third challenge is data imbalance problem. As mentioned, in Sec. 1.1, the occurrence of arrhythmia-related heartbeats, especially the S-type heartbeats, is a sporadic event. For most patients, more than 90% of their heartbeats are normal heartbeats. The under-supplied disease heartbeats impose an obstacle for learning models to recognize the abnormal patterns. Moreover, the severe unevenly

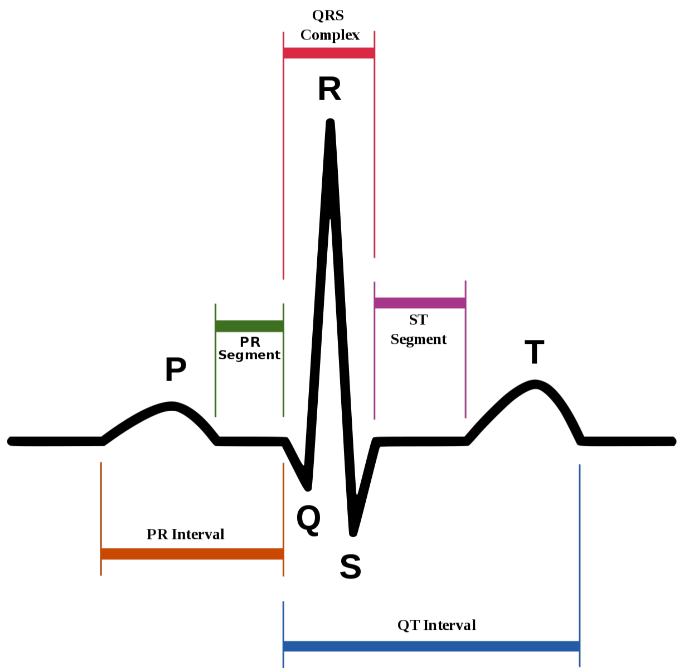


Figure 1.3: Fiducial points and key intervals of a segmented normal ECG heartbeat [189].

distributed heartbeats tend to bias an automated heartbeat classification method.

The fourth challenge is the intra and inter-patients variations. The intra-patient variations is that a patient's heartbeats (of the same type) may exhibit different patterns at different timestamps. For example, given a male patient X, the previous-RR-intervals of X's normal heartbeats in the morning and in the evening can vary significantly. By contrast, the inter-patient variations is that heartbeats (of the same type) of different patients may exhibit different patterns. The intra and inter-patients variations have caused a lot of troubles to the automated heartbeat classification task.

The fifth challenge is the identification S-type heartbeats from normal heartbeats, which is one of the most problematic tasks for existing arrhythmia detection methods. As shown in Fig. 1.1, it is less likely to provide accurate identification of the S-type heartbeats from the normal ones merely based on the morphology. In clinical practice, special rhythm information between two heartbeats, known as the RR-interval, is needed to help identify the S-type heartbeats because the S-type heartbeats are premature heartbeats and they normally have shorter previous-RR-intervals than the normal heartbeats. However, the inter- and intra-patients variations existing in the heart rhythms still impose great challenges to the detection tasks.

The sixth challenge is interpretability. Model explainability is important for a machine learning model to be applied in clinical practices, which is essential for clinicians to rationalize the model prediction [11]. Basically, existing solutions for automated heartbeat classification can be roughly divided as feature-engineering based and deep-learning based methods. In fact, the feature-engineering based methods have incorporated medical knowledge during the feature extraction stage. Therefore, they are easy to interpret. However, the deep-learning based methods perform *black box* operations, in which the features are learned from data automatically. A comprehensive investigation into the deep-learning models should be conducted to explain the intermediate processes of the models and to improve clinicians' trust in the models.

It is worth to note that, the challenges presented in this section are not mutual-independent. In fact, they are inter-twisted with each other, making the heartbeat classification task more complicated.

1.3 Research Questions

The author formalizes the above-discussed challenges into specific research questions, which will then be addressed one by one in the following chapters.

- What are potential quality problems of real-world collected ECG signals and how to improve the quality of these signals to facilitate the follow-up anomaly detection tasks?

- How to perform accurate heartbeat segmentation on the real-world collected ECG recordings and how to evaluate the segmentation performance? Does the segmentation provide enough information for an algorithm to recognize the anomalies?
- How to tackle the data imbalance problem, especially when abnormal cases account for less than 10% of the total data? How to provide an effective and reasonable data augmentation method that can catch the real distribution of the abnormal samples?
- How to deal with the intra and inter-subjects variations? How to effectively catch the general patterns of different subjects to improve the arrhythmia detection performance?
- What are the key differences between the S-type heartbeats and normal heartbeats? How to extract temporal information to help improve the classification performance of the S-type heartbeats?
- How to explain and interpret an arrhythmia detection model to enhance its application in practice?

1.4 Contributions

The author has made several attempts to tackle the six challenges and the specific research questions discussed in the previous sections. The achievements are summarized as follows.

It is worth noting that the proposed models are trained and evaluate on the baseline MIT-BIH-AR database. To provide a better explanation of the achievements made, the author provide a brief introduction of the database first. MIT-

BIH-AR is the benchmark database for arrhythmia detection, which is used in most published research [119]. The database contains 48 two-lead ambulatory ECG recordings from 47 patients (including 22 females and 25 males). Each recording is denoted by a 3-digits number. The recordings were digitized at 360 Hz per second per channel with 11-bit resolution over a 10-mV range. For most of them, the first lead is modified limb lead II (except for the recording 114). The second lead is a pericardial lead (usually *V1*, sometimes are *V2*, *V4* or *V5*, depending on subjects). In this study, only the modified limb lead II is used.

1.4.1 D-ECG: A Dynamic Framework for Automated Cardiac Arrhythmia Detection

The author notices that many existing methods for heartbeat classification are facing a bottleneck of applying a standalone classifier and using a static feature set to classify all heartbeat samples. This has been shown to cause huge impacts on identification of the problematic heartbeats because of the intra and inter patients variations. However, a single classifier is believed to be an expert only in certain local regions of the feature space.

In this work, the author propose a dynamic framework named D-ECG, which introduces the *dynamic ensemble selection* (DES) techniques to solve the problems. Technically, the proposed D-ECG consists of five phases: preprocessing, feature extraction, classifier pool training, dynamic selection classification and result refinement. In the classifier pool training stage, the synthetic minority oversampling as well as the edited nearest neighbors technique (SMOTEENN) are adopted to remedy the negative effect caused by the data imbalance of the ECG database. Besides, a result regulator is creatively developed to refine the result from the previous phase. Specifically, the result regulator is a *Support Vector Machine* trained with an adjusted feature set which is different from the one used for DES training. The rationale of such a classification strategy is that the sensitivities to certain feature varies with heartbeat types [209].

To the best of our knowledge, this is the first time that the DES techniques are introduced in the heartbeat classification task. The author conducts extensive experiments to examine the performances of a wide range of DES techniques and the performance of the proposed framework. The result shows that the proposed framework has brought visible improvements on overall heartbeat classification accuracy as well as the sensitivity of disease heartbeats.

1.4.2 A Pyramid-like Model to Improve Heartbeat Classification Performance

Although the *dynamic ensemble selection* (DES) techniques has introduced visible improvements on the heartbeat classification task, they also brought some disadvantages. The dynamic selection nature, which determines the best classifiers according to the sample to be predicted, can result in a serious delay of the model prediction, making the model less practical in online detection scenarios.

In this work, the author proposes a novel pyramid-like model to tackle the challenges discussed in Sec. 1.2. Compared to the D-ECG framework, the proposed pyramid-like model can provide more timely response to an unknown heartbeat while maintaining a good classification performance as the D-ECG framework. It has the potential to be applied in online detection scenarios.

Specifically, by noticing that most of the existing works take heartbeats as mutual-independent data samples and ignore the neighbor-related information of heartbeats, the author specially designs a pyramid-like structure to takes advantage of the information provided by the surrounding heartbeats to assist identification of disease heartbeats. The proposed pyramid-like model is made up of the nsDispatcher, nRefiner and sRefiner. Fig. 1.4 presents the entire framework. The classification process has two stages, known as *level-1* and *level-2* classification. In *level-1* classification, the raw heartbeat data is processed by the *nsDispatcher* at first, where each heartbeat is categorized into the N or S group. After that, in the *level-2* classification, the *nRefiner* classifies the heartbeats in the upper N group to the N, V, F or Q group. Simultaneously, the *sRefiner* classifies the heartbeats in the upper S group to the S, V, F or Q group. The *nsDispatcher* is an algorithm that specially design to consider neighbor-related information to assist disease heartbeats detection by incorporating medical rules, whereas *nRefiner* and *sRefiner* are two classification strategies to tackle the data imbalance problem and the static features problem, respectively. Details are given in Chapter 4.

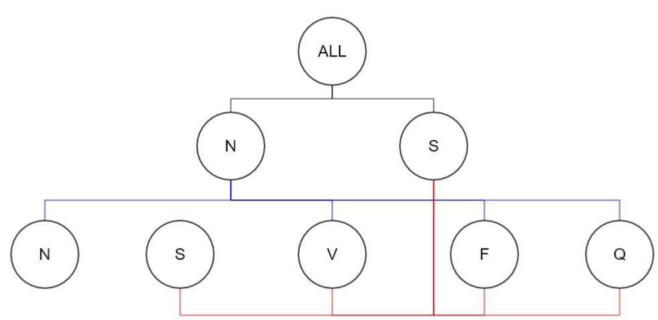


Figure 1.4: Overall structure of the proposed pyramid-like model.

On the basis of the proposed pyramid-like model, we also propose a completed solution for real-time arrhythmia detection from IoT (*Internet-of-Things*) ECGs. Fig. 1.5 depicts the solution. The proposed solution covers the whole life-cycle of online arrhythmia detection: (1) IoT ECG collection and transfer; (2) ECG recording cleaning; (3) heartbeat segmentation and featurization; (4) heartbeat classification; (5) Result Notification.

1.4.3 Inspecting the advances brought by DNN in Heartbeat Classification

Recent advances in heartbeat classification are largely driven by deep neural networks (DNNs). A DNN is a computational model consisting of multiple processing layers, which can automatically learn the high-level representations of the raw ECG recordings without extensive data preprocessing. In consideration of the sporadic occurrence of S-type heartbeats, which imposes a great challenge to DNN training, many DNN-based studies used synthetic heartbeats for model training and evaluation [3,113,114,203,207]. However, these efforts suffer from data leakage because, after augmentation, data is not partitioned patient-wise into training and test sets. So that beats from the same patient may appear in training and test, and the deep learning algorithms may learn patient-specific characteristics during training which then appear on test data. Additionally, the over-optimistic results obtained from data leakage have hidden a potential limitation of these DNN models in which only the ECG segmented heartbeats are accepted as inputs. The inter-heartbeat rhythm information is not well considered in these models. As mentioned in Sec.1.2, the rhythm provides indispensable information to distinguish the S-type arrhythmia heartbeats. Without such information, a high misclassification rate is probably obtained on S-type heartbeats. The problem is still open.

The author re-implements some DNN-based works [3, 113, 114] and evaluates these models on the benchmark MIT-BIH Arrhythmia database [137] following the well-recognized inter-patient evaluation scheme [38]. As compared against the reported performances in literature, these models' performances measured in our experiments have some degradation. The result confirms that, without considerations of heart rhythm, a DNN is less likely to identify S-type heartbeats.

To solve the above problem, the author propose a DNN-based solution named

Multi-channels Convolution Neural Network (MCHCNN). As an improvement, the proposed network accepts raw ECG heartbeat and heart rhythm (RR-intervals) as inputs and uses different sizes of convolution filters in parallel to capture temporal and frequency patterns from ECG signals. Fig. 1.6 presents the proposed network architecture. Although the experimental results has shown visible improvements brought by MCHCNN, there is still a long way before MCHCNN can make practical impacts because its performance of S-type heartbeats detection is still relatively low. Further improvements are necessary.

1.4.4 An Advanced Two-step Deep Neural Network-based Classification Framework for Arrhythmia Detection

Deep neural networks (DNNs) have brought noticeable advances to the field of arrhythmia detection, but to identify the problematic supraventricular ectopic (S-type) heartbeats is still a bottleneck in most of the existing studies. This is mainly due to morphological similarity between the S-type heartbeats and the normal ones, the imbalanced heartbeat occurrence rate, and both the inter- and intra-patients variations in heart rhythms. Although the MACHCNN proposed in Sec. 1.4.3 has made some improvements, it still suffers from this bottleneck.

In this work, the author presents a two-step DNN-based classification framework to identify problematic heartbeats for arrhythmia detection. Due to the observed difficulty of detecting S-type heartbeats from N-type heartbeats, the proposed framework trains a deep dual-channel convolutional neural network (DD-CNN) which accepts segmented heartbeats as input in the first step to classify V-type, F-type and Q-type heartbeats. At this stage, S-type and N-type heartbeats are not the targets, so they are put into one bundle to be studied in the next step. In the second step, a central-towards LSTM supportive model (CLSM) is specially designed to distinguish S-type heartbeats from N-type ones. The RR-intervals of a heartbeat and its neighbors are arranged in sequential form, serving as the input to CLSM. In particular, CLSM learns and extracts hidden temporal dependency between heartbeats by processing the input RR-interval sequence in central-towards directions. Instead of using raw individual RR-intervals, the abstractive, mutual-connected temporal information provides stronger and more stable support for identifying the problematic S-type heartbeats. Besides, as an improvement as well as a necessary driver for activating the CLSM, a rule-based data augmentation method is also proposed to supply high-quality synthetic samples for the under-represented S-type RR-interval sequences. To avoid data leakage, the benchmark evaluation dataset is split into training and test sets at patient level following the well-recognized inter-patient division paradigm proposed in [38]. The synthetic training samples are generated from the training set only.

Extensive experiments on three real-world ECG databases are implemented to evaluate the proposed framework and the rule-based data augmentation method. The experimental results show that the proposed framework has the potential to make a substantial clinical impact. Besides, Although CLSM is initially designed as the second-step structure in the proposed framework, it is a general and flexible binary classifier. For those works suffering from the confusion of the S-type and the normal heartbeats, CLSM can be easily integrated as a complement without changing their original structures. This is why we define CLSM as a supportive model.

1.5 Thesis Structure

The rest of this thesis is structured as follows. Chapter 2 introduces background knowledge of ECG signals, reviews previous research efforts to tackle the heartbeat classification task and discusses the evaluation paradigm for an automated heartbeat classification model. Chapter 3 to Chapter 6 present my main research outcomes. Specifically, Chapter 3 propose a dynamic framework named D-ECG, which introduced the dynamic ensemble selection techniques to solve the heartbeat classification problem. Chapter 4 proposes a novel pyramid-like model to tackle the problem. Compared to the D-ECG framework, the pyramid-like model can provide more timely response to an unknown heartbeat. Chapter 5 inspects the recent advances brought by deep learning and proposes a DNN-based solution named MCHCNN to solve the problems of existing deep learning based methods. Chapter 6 proposes an advanced two-step deep neural network-based classification framework for arrhythmia detection. The entire thesis is summarized with our achievements and future work discussion in Chapter 7.

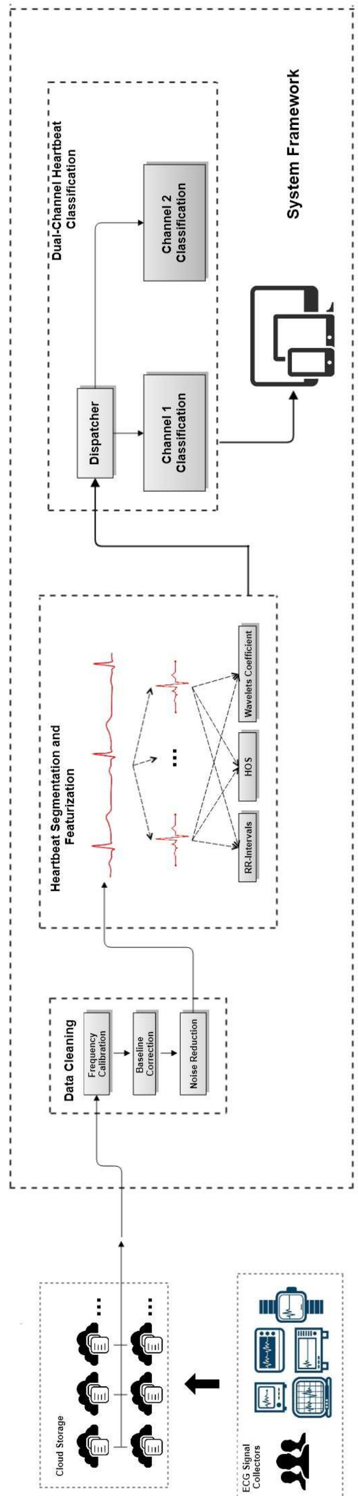


Figure 1.5: A solution for online arrhythmia detection.

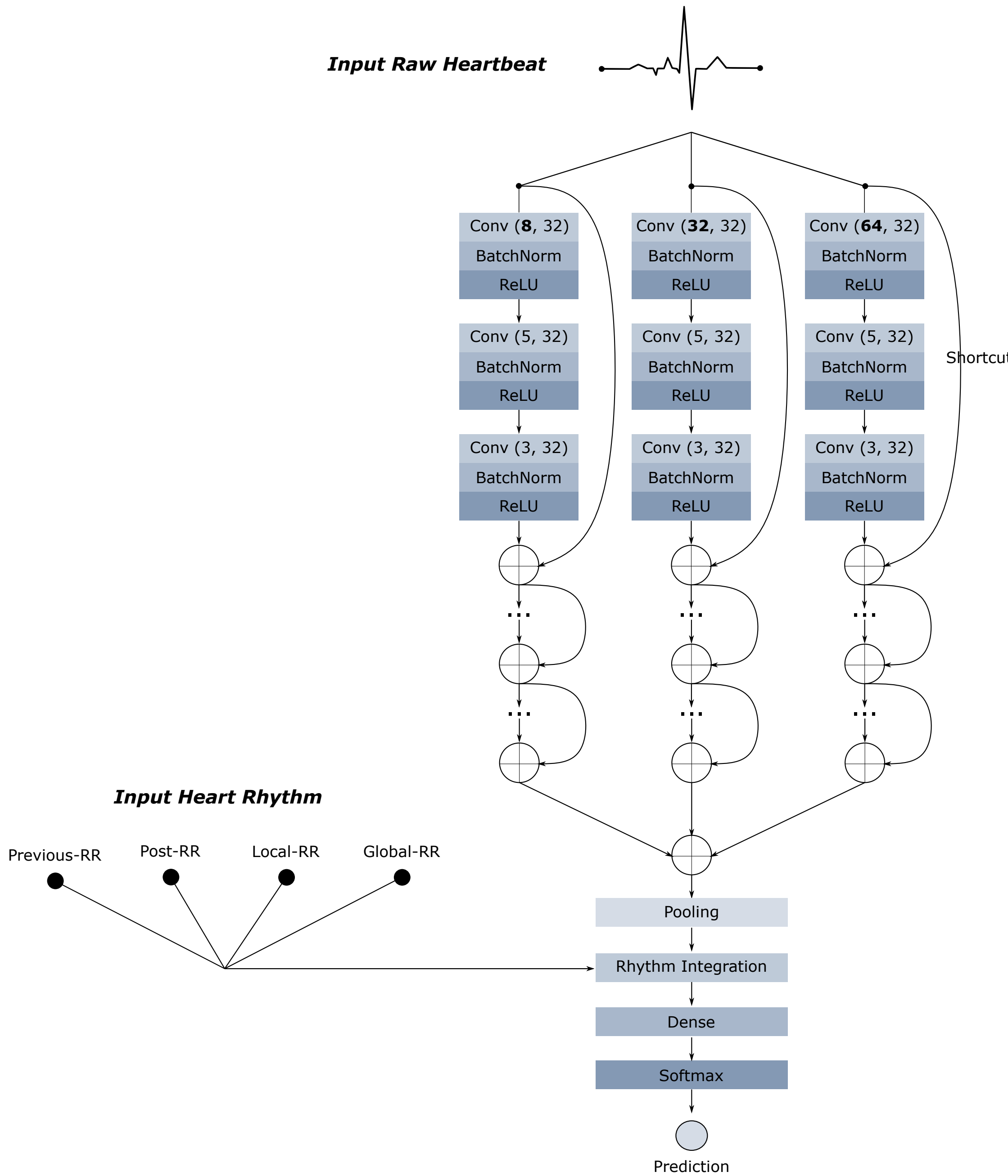


Figure 1.6: Architecture of the proposed Multi-channels Convolution Neural Network.

CHAPTER 2

LITERATURE REVIEW

This chapter presents related works to tackle the heartbeat classification task. In particular, Sec. 2.1 introduces primal aspects of ECG signals. Sec. 2.2 reviews previous research efforts to tackle the heartbeat classification task. Sec. 2.3 discusses the paradigm and metrics used to evaluate a heartbeat classification model. In Sec. 2.4, an overview of IoT-based ECG monitoring system is provided.

2.1 ECG Signal

ECG records the electrical activity of the heart via measuring the heart voltage potential differential. It captures heart rate, rhythm, conditions of heart muscles and other vital information regarding the electrical heart activities.

Analysis of ECG has a wide range of applications, covering fields of heart disease classification [4, 7, 18, 23, 28, 48–51, 56, 66, 77, 84, 87, 88, 97, 100, 105, 108, 118, 122, 124, 126–128, 135, 136, 138, 145, 148, 151, 175, 179, 185, 201, 202, 211], sleep apnea detection [22, 61, 67, 85, 134, 163], biometric identification [62, 122, 172, 182], emotion recognition [5, 116], driver drowsiness classification [29, 89, 90], and others [82, 107].

Currently, there are multiple approaches to measure ECGs. Based on the measurement principles, these approaches can be classified into three categories: *In-the-person*, *On-the-person*, and *Off-the-person* [34].

- *In-the-person.* Measuring devices are implanted inside the body of the person via surgical implanation or ingestion of pill-shaped system. Such kind of approaches are mainly for chronic patients and in extreme clinical scenarios.
- *On-the-person.* Most of the measuring approaches come from this category. Such approaches use a device that is directly attached to body surface to ob-

tain the signal. Bedside monitors and Holter machines are typical examples, in which 12 leads electrodes are placed on a patient's arms, legs and chest.

- *Off-the-person*. This approach measures ECG by sensors which are embedded into daily used objects. In recent years, applications of such approaches presents an increasing trend. In comparison to the *In-the-person* and the *On-the-person* approaches, this approach is more human-friendly and less invasive. It is aligned with the future trend of *Internet-of-Things* and *Artificial Intelligence*.

2.2 Previous Efforts in Heartbeat Classification

Heaps of research efforts have been made to tackle the heartbeat classification task over the years. The existing solutions can be roughly allocated to either the feature-engineering based or the deep-learning based category. Table 2.1 summarizes their differences.

2.2.1 The Feature-engineering Based Methods

The feature-engineering based methods focus on signal feature extraction, feature selection, and classifier selection.

Feature extraction and selection

The feature extraction stage plays a core role in feature-engineering based methods. It causes direct impacts on the final classification performance. In raw ECG signals, a heartbeat is denoted as a high-dimensional time-series sequence, which is difficult for a classifier to interpret and discover important information to distinguish different heartbeat types. Therefore, some patterns needs to be sum-

Table 2.1: Comparison between Feature-engineering based and Deep-learning based methods

	Feature Engineering	Deep Learning
Work flow	Feature extraction, selection and classifier determination	End-to-end processing
Commonly used features	Mainly handcrafted, including RR-intervals, higher-order statistics, wavelet, signal energy coefficients, etc.	Learned by networks, including CNN, RNN, LSTM, etc.
Feature selection	PCA, floating sequential search, weighted LD model	N.A.
Commonly used classifiers	SVM, nearest neighbors, artificial neural networks, weighted linear discriminant, optimum-path forest	N.A.
Training data	Less	More
Parameters	Less	More
Explainability	High	Low
Current limitations	Use of fixed features for all heartbeat types classification; Limitation of static classifiers to handle both intra- and inter patients variations	Lack of considerations of frequency patterns and heart rhythms; A biased evaluation is followed.

marized and extracted from the time-series sequence to represent the heartbeats. Such patterns are known as features. Feature extraction is normally performed under the guidance of medical knowledge, which helps to improve classification performance and increase the model explainability. For example, RR-intervals is one of the most commonly used features because it contains indispensable rhythm information to distinguish the premature S-type heartbeats from the normal ones [4, 23, 54, 55, 99, 151, 209].

Most of the heartbeat features found in literature are extracted in the time or frequency domain. The author explains and summarizes the most widely used features as follows.

- RR-interval. This feature denotes the heart rate. In particular, the interval between the current R peak and the previous R peak is known as previous-RR, while the interval between current R peak and the following R peak is post-RR. Local-RR is the average of local RR-intervals. The range of local-RR needs to be determined based on the model. Global-RR is defined as the average of RR-intervals on the entire ECG recording.

- QRS complex statistics [38, 96, 209]. This feature is presented as statistics to describe a certain interval in the QRS complex of a heartbeat. It is widely used in clinical practices. Depending on different intervals, as shown in Fig. 1.3, this feature can be P-duration, QRS-duration, T-duration, PR-interval, or QT interval. It is worth to mention that the use of this feature requires an extra algorithm to detect the fiducial points, such as the one proposed by Laguna et al. [103]. The accuracy of the point detection can be a potential influence factor of the heartbeat classification performance.

- Morphology amplitudes [14, 38, 131, 209]. This feature uses a group of values obtained from down-sampling of the heartbeat segment amplitude to depict the ECG morphology. Depending on different segment, this feature can be P-morphology, QRS-morphology, or ST-morphology. Similar to the QRS complex statistics, the morphology amplitudes is also widely used in clinical practices. However, this feature is not very efficient for an automated model because it suffers from high dimensions.

- Principal components [17, 91, 177]. This feature is the coefficients obtained by performing the principal component analysis (PCA) on the original ECG heartbeats or the key segments in the heartbeats. It avoids the high dimension problem of including the raw or down-sampled sequence in the feature vector. In the theory of PCA [192], an ideal feature set has three characters: (1) high variance of individual feature; (2) mutual-uncorrelated; (3) not too many. PCA produces an ideal set of features by creating a set of principal components. The first component (C_1) is the strongest underlying trend which captures the highest variance of the data. The second component (C_2) is the second strongest underlying trend that happens to be perpendicular to C_1. The third component (C_3) is the third strongest underlying trend which is also perpendicular to C_1 and C_2, and so on. It is worth to note that PCA is an unsupervised process. That means, it is not guaranteed that the principal components contribute significantly to the classification task.

- Independent components [20, 155, 205, 206]. Similar to the principal components, this feature is the coefficients obtained by performing the independent component analysis (ICA) on the original ECG heartbeats or the key segments in the heartbeats. In particular, ICA [75] is a statistical method to find mutual-independent components from multivariate data. In the heartbeat classification task, ICA is applied to calculate the independent components from the ECG signal in the Fourier transformed domain or the time domain. The independent components then serve as important features to distinguish different heartbeat types. It is reported that the use of independent components helps to identify normal heartbeats [142].

- Higher-order statistics [41, 42, 77, 109, 130, 163, 172]. The *skewness* (3rd order statistics) and the *kurtosis* (4th order statistics) are two commonly used Higher-order statistics. They are effective in estimating shape parameters of ECG signals. For an input signal, assume $X_{1...,N}$ denotes all the data samples, $\bar{X}$ is the mean and s is the standard deviation, the *skewness* and *kurtosis* can be defined respectively as below.

$$Skewness = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^3/N}{s^3} \tag{2.1}$$

$$Kurtosis = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^4/N}{s^4} \tag{2.2}$$

Application of the higher-order statistics helps to distinct V-type heartbeats because the major difference of V-type heartbeats against other types is the heartbeat shape. Moreover, it also helps to reduce the variability among individual heartbeats of the same type [141].

- High-order accumulates [140,141]. This feature provides a statistical description of QRS complex. It helps to amplify the differences between different types of heartbeats. The second- ($C_{2x}(k)$), third- ($C_{3x}(k,l)$) and fourth-order (($C_{4x}(k,l,m)$)) accumulates are defined as follows.

$$C_{2x}(k) = E\{x(n)x(n+k)\} \tag{2.3}$$

$$C_{3x}(k,l) = E\{x(n)x(n+k)x(n+l)\} \tag{2.4}$$

$$\begin{aligned} C_{4x}(k,l,m) =& E\{x(n)x(n+k)x(n+l)x(n+m)\} \\ & - C_{2x}(k)C_{2x}(m-l) - C_{2x}(l)C_{2x}(m-k) \\ & - C_{2x}(m)C_{2x}(l-k) \end{aligned} \tag{2.5}$$

where E denotes expectation, and k, l, and m are the time lags.

- Wavelet coefficients [60, 101, 112]. *Discrete wavelet transform* (DWT) provides a time-frequency representation of a signal, which is widely used in data compression, noise reduction and multi-frequency-bands signal analysis. DWT iteratively decomposes a signal to different frequency bands with a scaling function and a wavelet function. The high-frequency component provides the detail information of the upper-level signal whereas the low-frequency component is a coarse approximation of the upper-level signal. Each component is represented by a collection of wavelet coefficients, which is obtained by the inner products of mother wavelet function and the upper-level signal. As reported in [119], wavelet coefficients is believed to be the best feature for heartbeat classification. The choice of the mother wavelet function is the key of the *discrete wavelet transform*, which heavily depends on applications. For heartbeat features extraction, the Haar wavelet is always chosen because of its simplicity. Besides, it has been demonstrated as the ideal wavelet for short time signal analysis [204]. The Haar function can be represented as

$$\psi(t) = \begin{cases} 1 & 0 \leq t<1/2, \\ -1 & 1/2 \leq t<1, \\ 0 & otherwise. \end{cases}, \tag{2.6}$$

and its corresponding scaling function is

$$\phi(t) = \begin{cases} 1 & 0 \leq t<1, \\ 0 & otherwise. \end{cases}, \tag{2.7}$$

where t denotes sample values.

- Statistics of wavelet coefficients. Instead of using the wavelet coefficients directly, some statistical features extracted from the wavelet coefficients have also been proposed. Examples include mean, standard deviation, energy [60] and coefficient variance [204]. These features are less sensitive to the variations of marked fiducial points.

- Random projection [13, 74]. This feature is obtained from projecting the original ECG signals onto a random matrix. It denotes a low dimensional representation to the original signal. The feature has a low computational complexity which is suitable for real-time detection scenarios.

- Maximal Lyapunov exponents [143, 177, 178]. The Lyapunov exponent is a measure to distinguish different types of trajectories based on their sensitive dependence on the initial conditions. The maximal Lyapunov exponent λ can be defined as follows:

$$\lambda = \lim_{t\to\infty} \lim_{|\delta \mathbf{Z}_0|\to 0} \frac{1}{t} \ln \frac{|\delta \mathbf{Z}(t)|}{|\delta \mathbf{Z}_0|} \tag{2.8}$$

where $\delta \mathbf{Z}_0$ denotes the initial separation and $|\delta \mathbf{Z}(t)| \approx e^{\lambda t}|\delta \mathbf{Z}_0|$.

- DTW distance [188, 209]. This feature denotes the similarity between a given beat and the median beat of a recording. It helps to distinguish the V-type beats which exhibit significant morphological differences against other heartbeats. However, the feature suffers from high computation complexity. It is less likely to be included in a real-time detection system.
- Other used features found in the literature include Fuzzy clustering coefficients [18, 139, 144], Hermite transformed representation [79], Linear predictive codes [64], Local fractal dimension [135], Morphological areas [209], and etc.

These features have been experimentally proven to be able to make contributions to the distinction of different types of heartbeats. However, it is impossible to include all these features into consideration as this will lead to model overfitting. Therefore, feature selection is highly demanded. Most of the existing works decide which features to be used based on their experiences and understanding of the task, which does not guarantee the used features are optimal. Only some works have investigated the feasibility of automated feature selection techniques to select the most representative features. Llamedo and Martinez [115], and Mar et al. [124] conduct feature selection by applying the *floating sequential search* strategy. Doquire et al. [46] search for the most representative features by using the *filter* and the *wrapper* feature selection techniques, respectively, and they find that

RR-intervals and higher order statistics have a strong discrimination power. To have a better understanding of how techniques work, the author provides a brief overview of feature selection as follows.

A feature selection technique consists of a search strategy to select candidate feature subsets and an objective function to evaluate these candidates. There are three main categories of search strategies: exponential, sequential and randomized. The *floating sequential search* (FSS) mentioned above is a sequential search strategy [149]. It is an extension to the *Plus-L minus-R selection* (LRS) strategy. The basic idea of LRS is presented in Algorithm 1. Given values of L and R, if L is larger than R, LRS starts from an empty set and repeatedly adds L features and removes R features to optimize the classification performance; otherwise, LRS starts from the full set and repeatedly removes R features followed by L additions to optimize the classification performance.

Algorithm 1 The workflow of LRS

Require: Numbers of features to be added, L; Numbers of features to be removed, R; A feature full set, $fullFeas$; A classification performance metric, J.

Ensure: The optimal feature subset, $optFeas$.

if L >R **then**
    $optFeas = \emptyset$
else
    $optFeas = fullFeas$
end if
$K = 0$
while K <a Preset Value **do**
    if L >R **then**
        for repeat L times **do**

$x^+ = \arg\max_{x \notin optFeas} J(optFeas + x)$

$optFeas = optFeas + x^+$

end for

for repeat R times **do**

$x^- = \arg\max_{x \in optFeas} J(optFeas - x)$

$optFeas = optFeas - x^-$

end for

else

for repeat R times **do**

$x^- = \arg\max_{x \in optFeas} J(optFeas - x)$

$optFeas = optFeas - x^-$

end for

for repeat L times **do**

$x^- = \arg\max_{x \in optFeas} J(optFeas - x)$

$optFeas = optFeas + x^+$

end for

end if

end while

FSS improves LRS by allowing the values of L and R to be determined from the data instead of fixing them. In terms of the objective functions, they can be classified into two categories: *filters* and *wrappers* [63]. In *filters* objective function, feature subsets are evaluated based on their information content, such as inter-class distance, mutual-correlations, statistical dependence or information-theoretic measures. By contrast, in *wrappers* objective functions, feature subsets are evaluated based on their actual classification performance on test data. Therefore, when using *wrappers* objective functions, a classifier and a performance metric need to be determined. This leads to a disadvantage of *wrappers* objective functions: the optimal feature subset will be specific to the classifier and classification metric under consideration.

Classifier selection

Regarding the classifiers used in the heartbeat classification task, the *support vector machine* (SVM) is the most widely used for its robustness, good generalization and computationally efficiency [1, 23, 35, 39, 74, 147, 154]. Besides, *decision trees* (DT) [54, 121, 152], *K-nearest neighbors* (KNN) [9, 97, 104, 143, 173] and *artificial neural networks* (ANN) [52, 60, 131, 141, 144] are also frequently found in the literature. Other used classifiers include *optimum-path forest* (OPF) [37], *linear discriminants*(LD) [38], *conditional random field* [40], and *reservoir computing with logistic regression* [53], *hidden Markov models* [30, 58] etc. The author provides an overview of the frequently used classifiers as below.

- SVM. *Support vector machine* [31] is a binary classifier. It is an extension to *maximal margin classifier*. Given training observations $x_1, \ldots, x_n \in R^p$ and associated class labels $y_1, \ldots, y_n \in \{-1, 1\}$. Briefly, the training objective of *maximal margin classifier* is to construct a separating hyper-plane that has the farthest minimum distance to the training observations. The training process is formulated as below.

$$\begin{aligned}
&\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{maximize}}\, M \\
&\text{subject to } \textstyle\sum_{j=1}^{p} \beta_j^2 = 1 \\
&y_i\left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}\right) \geq M \forall i = 1, \ldots, n.
\end{aligned} \tag{2.9}$$

  where $\beta_1, \ldots, \beta_n$ are the parameters of the hyper-plane.

  Maximal margin classifier classifies unknown observations depending on which side of the hyper-plane the observation is located. However, this classi-

Table 2.2: Non-linear kernels

Kernels	Mathematical representation
Polynomial Kernel	$K(x_i, x_{i'}) = (x_i \cdot x_{i'} + 1)^p$
Gaussian Kernel	$K(x_i, x_{i'}) = e^{\frac{-1}{2\sigma^2}(x_i - x_{i'})^2}$
Radical Kernel	$K(x_i, x_{i'}) = \exp^{-\gamma \sum_{j=1}^{p} \left(x_{ij} - x_{i'j}\right)^2}$
Sigmoid Kernel	$K(x_i, x_{i'}) = \tanh(\eta x_i \cdot x_{i'} + \nu)$

fier is limited by linearity. It cannot separate classes with non-linear decision boundaries. SVM improves the *maximal margin classifier* by introducing the *kernel* trick to enlarge the feature space. A *kernel* is a function that quantifies the similarity of two observations. Let x_i and $x_{i'}$ denote two observations, a *kernel* function K can be represented as

$$K(x_i, x_{i'}) = \sum_{j=1}^{p} x_{ij} x_{i'j}. \tag{2.10}$$

It is equivalent to the inner product of the two observations, but it is more computational efficient than calculating the inner product directly. Unlike using quadratic and cubic terms to enlarge the feature space which involves heaps of inner products calculations, using *kernel* functions avoids complex calculations in high-dimensional space. There are some commonly used kernels, including the Polynomial Kernel, the Gaussian Kernel, the Radical Kernel and the Sigmoid Kernel. The author presents these kernels in Table 2.2. Different kernels lead to different decision boundaries. It is hard to say which one is the best. It really depends on specific tasks and data distributions.

It is worth to note that SVM is sensitive to imbalanced data. Therefore, heartbeats balancing should be considered when applying SVM on the heartbeat classification task.

- DT. *Decision Tree* [150] is a flowchart-like classification model in which internal nodes denote selected features and terminal nodes represent labels. It classifies unknown samples based on the *if-then-else* rules acquired on training observations. A DT is built by iteratively selecting the most discriminatory features as internal nodes. The *Gini index* or the *entropy* is used to quantify the discrimination ability of a feature, which are defined by

$$\begin{aligned} &Gini = \textstyle\sum_{k=1}^{K} \hat{p}_k \left(1 - \hat{p}_k\right), \\ &and \\ &Ent = -\textstyle\sum_{k=1}^{K} \hat{p}_k \log \hat{p}_k. \end{aligned} \tag{2.11}$$

Here $\hat{p}_k$ denotes the proportion of observations in the data set that are from the k-th class.

Compared to other classifiers, decisions given by DT are easier to interpret and rationalize, which is important in medical applications. However, DT is not efficient for features with continuous value, such as the RR-interval, HOS, and wavelet coefficients. Besides, DT tends to overfit training data if a high-dimensional feature vector is presented, but this can be relieved by applying tree pruning [123] or constructing random forest [72].

- KNN. The K-nearest neighbors is a simple yet effective classification method. Its basic idea is presented in Algorithm 2. For each unknown observation, KNN calculates its distances to all training observations. The top K nearest training observations are then selected to vote on the label of the unknown observation. KNN is efficient in connecting previous knowledge which is represented by training observations to predict an unkonwn sample. However, it has high computational cost. Therefore, KNN cannot be applied on real-time heart disease detection scenarios.

Algorithm 2 The K-nearest neighbors

Require: Numbers of neighbors, K; Traing set, D; Test set, T.

Ensure: Label predictions for each test sample.

1: **for** each sample $z' = (x', y')$ in T **do**
2:     **for** each sample $z = (x, y)$ in D **do**
3:         Calculate the distance of z' to z, $d(z', z)$
4:     **end for**
5:     $D_k =$ [the K nearest neighbors of z' in D]
6:     $y' = \text{argmax} \sum_{(x',y') \subseteq D_k}^{k} I\,(v = y_i)$
7: **end for**
8: Return

- ANN. The Multilayer Perceptrons (MLP) and the Probabilistic Neural Networks (PNN) are two widely used artificial neural network structures. MLP is a feedforward multilayer network which is shown in Fig.2.1. A MLP network consists of an input layer, multiple hidden layers and an output layer. Layers are made up of processing units (known as neurons) which receive and transform information. Neurons in different layers are connected by weighted edges. The network weights can be learned by applying backpropagation [70] during network training. Specifically, the hidden layers are also known as processing layers in which the output of a neuron is a nonlinear mapping of the weighted sum of the outputs of neurons in the previous layer. The non-linear transformations are achieved through activation functions. The commonly used activation functions include ReLU, Tanh, and sigmoid.

They can be represented as

$$ReLU(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases} \tag{2.12}$$

$$Tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \tag{2.13}$$

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{2.14}$$

It is worth to note that the numbers of total layers in a network and the number of neurons in a layer have direct influences on the classification performance of the network. Therefore, these parameters need to be carefully determined based on specific tasks. According to Mar et al. [124], MLP is reported to be superior to the linear discriminants classifier in the heartbeat classification task.

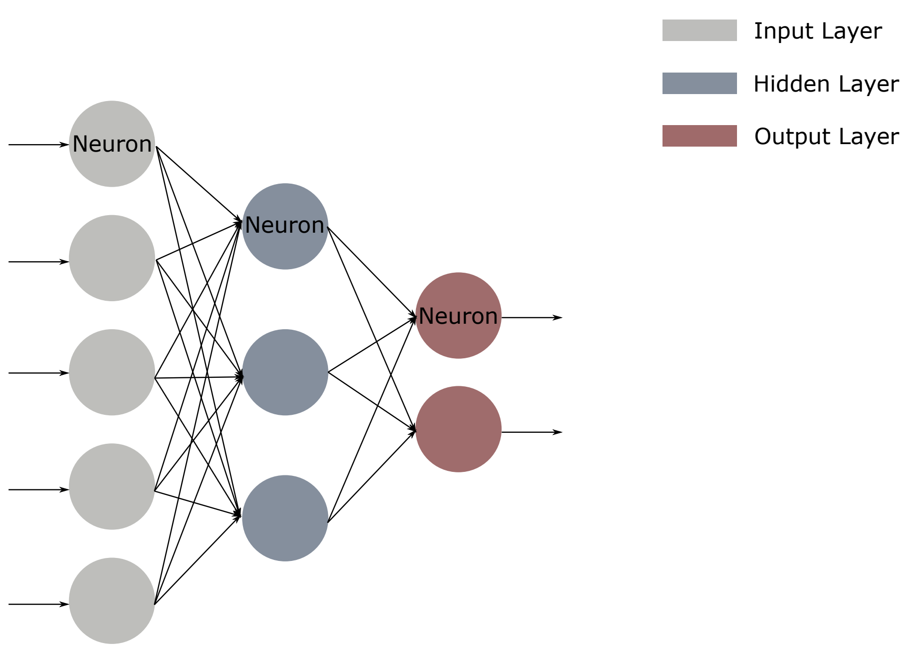


Figure 2.1: Structure of Multilayer Perceptrons.

PNN [166] is a radial basis function network, a special type of feedforward network. PNN consists of an input layer, a pattern layer, a summation layer and an output layer.

- The input layer receives values of training observations. Each neuron in the input layer takes in the value of a feature.
- The pattern layer is also called radial basis layer. It contains one neuron for each training observation. The neurons compute the Euclidean distances of the neurons' centers to the test cases and then perform nonlinear transformations via applying the radial basis function.
- The summation layer is used to connect the pattern neurons of different categories to vote for the most possible category of the test observation. It contains one neuron for each category. That is, the number of neurons in the summation layer is equals to the number of categories.
- The output layer outputs the category that has the most votes.

As compared to the traditional MLP, PNN is easier to train and converge [204]. Therefore, PNN is very suitable for the real-time processing scenarios. Besides, PNN is less sensitive to outliers, making it less distracted by the noises in the real-world collected ECG signals.

Critical analysis of existing works

One of the most significant works in heartbeat classification field is proposed at 2004 by De Chazal et al. [38]. By noticing that the improper training and test data separation can bias the classification results, this work proposes an inter-patients evaluation paradigm to prepare the benchmark MIT-BIH-AR database for heartbeat classification model evaluation. The inter-patients paradigm has made great impacts and it is widely followed by later works. Features used in this work are the ECG morphology and heartbeat interval from both ECG leads. De Chazal et al. adopt a weighted LD classifier design to process features from different ECG leads. However, feature examination and feature selection are not performed in this work and a less satisfied performance (overall accuracy 83%) is reported.

Llamedo and Martinez [165] improve De Chazal's work [38] by including frequent domain features into consideration and applying the floating feature selection algorithm to reduce the risk of overfitting. The wavelet transform is used to allow extraction of multi-scale frequent features. However, the floating feature selection algorithm does not perform as well as expected, but this work can still achieves an overall accuracy of 90%. Later, this work is extended by Tanis Mar et al. [124]. Tanis Mar et al. consider a feature combination of temporal, morphological and statistical features and use the sequential forward floating search (SFFS) algorithm for feature selection. As compared to exhaustive search method which is computational heavy, the SFFS works a lot more efficiently. More importantly, it has been reported that for most of the time SFFS is able to find solutions very close to the optimal one. This work reports an improved the disease heartbeat detection performance. However, the feature selection is limited in the extracted features only, which can be a potential issue to limit the model performance because there might be other more significant features.

Yakoub Bazi et al. [12] propose to use SVM to take in morphological features and wavelet coefficients to classify heartbeats for the first time. The proposed model is evaluated with the inter-patients paradigm. However, this work only report an overall accuracy of 92%. Recall and precision rates of the disease heartbeats are not found. It is difficult to tell the true performance of this model because of the imbalance nature of the test data.

Luz et al. [118] propose an efficient arrhythmia classification model by applying and analyzing a robust supervised graphbased pattern recognition technique, the optimum-path forest (OPF) classifier. As reported bu Luz et al., this is the first time that OPF classifier is used to the ECG heartbeat signal classification task. In order to demonstrate the effectiveness of OPF classifier, this work comprehensively compare its performance (in terms of training and testing time, accuracy, specificity, and sensitivity) of to the three well-known expert system classifiers, i.e., support vector machine (SVM), Bayesian and multilayer artificial neural network (MLP). The features used here come from six published works [38, 60, 164, 198, 204, 205]. However, the reported disease heartbeats detection performance is less satisfying, with the best sensitivity of the S-type and the V-type heartbeats being merely 18.3% and 82.4%, respectively.

Ye et al. [199] propose a novel ECG heartbeat classification model that combines general multi-class and specific two-class classifiers. The general classifier is trained on the global training data set with 5 heartbeat classes, whereas the specific classifier leverage the individual information to help detect abnormal heartbeats from the normal ones. The proposed models achieve a decent performance on S-type and V-type heartbeats detection. However, this model is less practical in real-world applications, especially in IoT devices. The reason is that the model requires fine-tuning for each user, including obtaining individual information and retraining the specific two-class classifier, which increase the computational and power burden of the IoT devices.

A weighed SVM model is proposed by De Lannoy et al. [39]. To avoid the impact of the data imbalance, a convex approximation of the balanced classification rate rather than the standard accuracy is used to optimize the SVM model. Besides, assessment of features is performed to select the best feature set. The model achieve a good performance on disease heartbeat detection, but the accuracy of the normal heartbeat is 80% only. That means, a large portion of normal heartbeats are misclassified as abnormal. It is difficult to apply this model in real-world scenarios because it can cause a lot of fake alarms and waste medical resources.

Feature extraction and selection play a vital role in a heartbeat classification model. Zhang et al. [209] introduce a novel disease-specific feature selection method to investigate the significance of extracted features in catching the differences of various types of heartbeats. The results show that RR-interval is the most powerful feature in distinguishing disease heartbeats from normal heartbeats, while the morphological distance (DTW distance between a beat and the median beat of a recording) is the No.1 useful feature to distinguish different disease heartbeats. In this work, the selected features from both ECG leads are input into combining SVMs for heartbeat classification. The model achieves an overall accuracy of 86%, and sensitivities of 89%, 79% and 85% for normal, S-type and V-type heartbeats, respectively. Still, the disease heartbeat detection performance is obtained at the expense of a high misclassification rate of normal heartbeats.

To summarize, the feature-engineering based methods are easy to interpret and rationalize. However, such methods often experience difficulty in achieving satisfactory performance on abnormal heartbeat detection while keeping a good overall classification accuracy, especially when S-type arrhythmia heartbeats are involved. Besides, the effectiveness of extracted features, the mutual-influences among features, and the compatibility between the feature distribution and the classifiers are three major factors that lead to a solid upper-bound on model performance.

2.2.2 The Deep-learning Based Methods

Recent advances in heartbeat classification are largely driven by deep neural networks (DNNs). A DNN is a computational model consisting of multiple processing layers, which can automatically learn the high-level representations of the raw ECG recordings without extensive data preprocessing. Convolution neural network (CNN) [98], Recurrent neural network (RNN) [156], and Long short-term memory network (LSTM) [73] are representative DNN structures. Most of the existing deep-learning based heartbeat classification methods are extended from these structures. To have a better understanding of the existing methods, the author presents the preliminary knowledge of CNN, RNN and LSTM as follows.

- **CNN**

  CNN is useful in learning representations of data. It is commonly applied in image and video recognition, recommend systems, image classification, and natural language processing. Recently, CNNs have attracted more and more attention in the applications in ECG signal classification because they have been proven effective in recognizing key patterns and learn useful features, such as P-waves and QRS-complexes of ECG heartbeats [65].

  A convolutional neural network is normally made up of an input layer, an output layer, multiple convolutional layers, pooling layers, and dense Layers. The convolutional layer is the core building block of CNN, in which most of the computational heavy lifting is done. Specifically, we define an input tensor I as a multidimensional array of data and the kernel K as a multidimensional array of parameters. The convolution operation S is actually a weighted average of an input tensor and a kernel at every position:

$$S(p) = \int I(p-a)K(a)da \tag{2.15}$$

where p denotes the position index. Fig. 2.2 present an example of 2D convolution. In this case, the input tensor is 4×3 and the kernel is 2×2. The output of the convolution operation is called a feature map, which is obtained by applying matrix multiplication between the kernel and every equal-sized sub-matrix in the input tensor. In practice, the number of kernels in a CNN need to be determined depending on specific tasks. Each kernel outputs a feature map. The convolution operation is usually followed by a ReLU activation to enable the network to learn non-linear patterns from the input data.

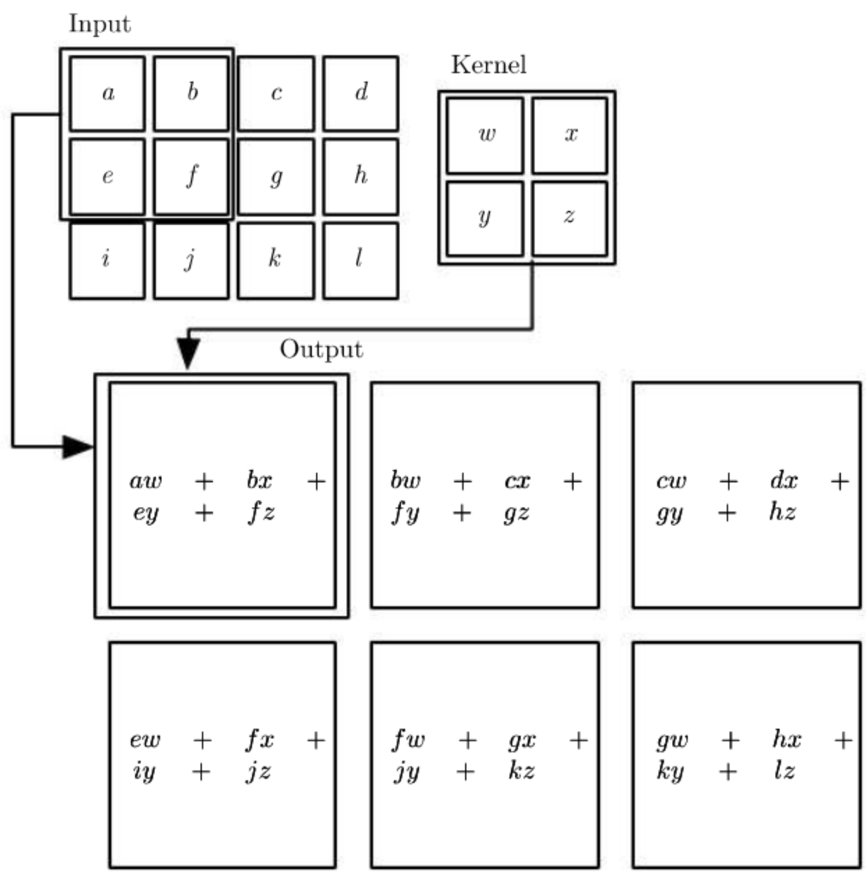


Figure 2.2: An example of 2D convolution [59].

The pooling layer replaces the output of the previous layer with a summary statistic. It is used to reduce the size of the learned representations to prevent overfitting. Commonly applied pooling methods include max pooling, average pooling, etc. The purpose of the dense layer is to provide an overall regulation of the previously learned representations. Finally, the output layer applies a softmax function on the outputs of the dense layer to calculate the categorical probability distribution. The softmax function can be mathematically denoted as

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{2.16}$$

where z is the input vector to the softmax function and $j = 1, 2, ..., K$ indexes the output categories.

- **RNN**

RNN is used for processing sequential data. A basic RNN structure consists of an input layer, a hidden layer and an output layer. Fig. 2.3 shows how a RNN maps an input sequence x to an output sequence o. In this case, U, V, and W denote weight matrix. The author mathematically presents the mapping process with the equations below.

$$h_t = f\left(U \cdot X_t + W \cdot h_{t-1}\right) \tag{2.17}$$

$$O_t = g\left(V \cdot h_t\right) \tag{2.18}$$

where f and g denote activation functions (sigmoid, ReLU, or tanh). It can be clearly seen that RNN produces an output at each time step. The current output is not only related to the current input, but also related to the hidden state in the previous time step. As compared to CNN, RNN is more effective in capturing mutual-dependence of sequential data. However, RNN has problems in dealing with long data sequences because the gradients which carry important information for optimizing network parameters get smaller with the increase of time steps. The vanishing gradient problem imposes a great challenge to the RNN training process.

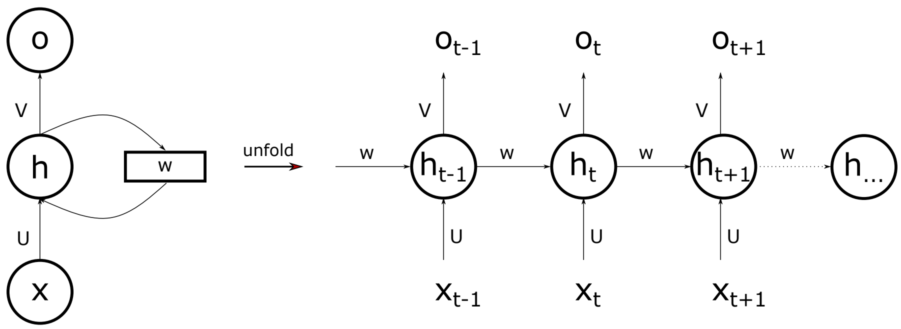


Figure 2.3: A graphical representation of RNN.

- **LSTM**

The long short-term memory network (LSTM) [73] is a variation of recurrent neural networks (RNN). It alleviates the vanishing gradient problem presented in the ordinary RNNs and it is able to learn temporal relationship across long periods of time.

A common LSTM unit consists of a *cell* c_t, an *input* gate i_t, a *forget* gate f_t and an *output* gate o_t, as shown in Fig.2.4. The *cell* remembers the time dependency between elements in the input sequence. Its memory can be effectively conveyed along the entire processing chain with just limited linear interactions. The *input* gate controls the new information to be stored in the *cell*. The *forget* gate decides the information to be thrown away from the *cell*. The unit output is managed by the *output* gate based on the current *cell*'s memory.

As an illustration, let W_n and U_n be the weights of inputs and recurrent

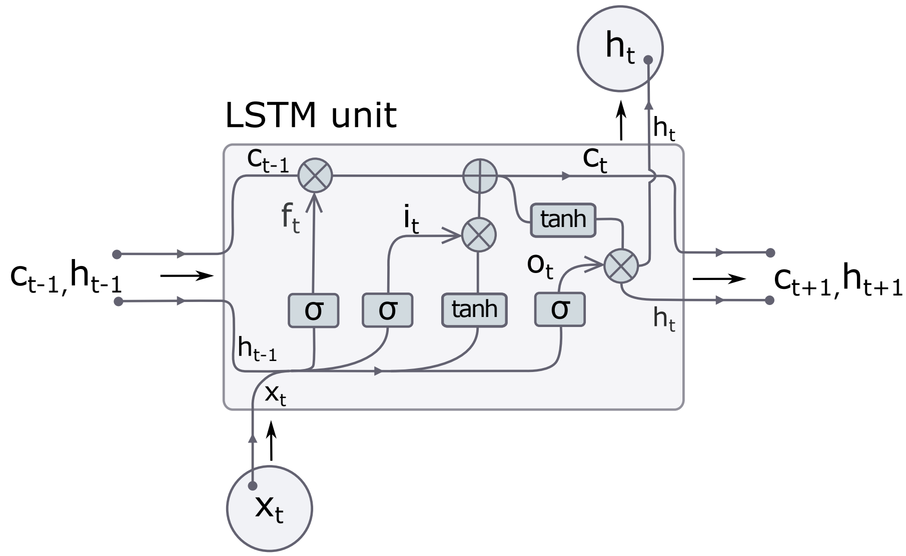


Figure 2.4: A LSTM Unit.

connections respectively, and b_n be the bias. The subscript n can be the *forget* gate f, *input* gate i, *output* gate o or the *cell* c. Given the input x_t, the LSTM unit at time t is updated as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{2.19}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{2.20}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{2.21}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{2.22}$$

$$h_t = o_t \circ tanh\left(c_t\right) \tag{2.23}$$

where σ represents the sigmoid function and the operator $\circ$ denotes the element-wise product. Apparently, in LSTM, the information stored in cells is not directly processed by any activation function. Therefore, the gradients of cell states can be passed over a long distance without being vanishing. Applying LSTM in a heartbeat sequence helps to capture mutual-relationships of fiducial points in the heartbeat. Moreover, LSTM can be used to explore temporal dependencies between heartbeats.

Critical analysis of existing works

Kiranyaz et al. [93] apply an adaptive 1D-CNN on ECG signals for the first time. They argue that the proposed method helps to avoid the use of hand-crafted features and reduce the computation cost of applying PCA on features to avoid the 'Curse of Dimensionality'. The proposed 1D-CNN is easy and efficient to implement on hardware, making it eligible for real-time heart monitoring and disease warning. However, the method is patient-specific, which means that it can not be applied in unknown patients. This limits the application of this method in many real-world scenarios.

Acharya et al. [3] propose a 9-layer CNN for the general heartbeat classification task. The model accepts segmented heartbeats as input. It calculates the probability of each heartbeat type that the input heartbeat may belong to. Chauhan and Vig [19] apply LSTM to detect abnormal heartbeats in ECG signals. They claim that the stack of LSTM layers help to extract useful temporal features of heartbeats. Yildirim [203] proposes a deep bidirectional LSTM and wavelet sequences based heartbeat classification model. In this model, the input heartbeats are firstly decomposed to multiple frequency resolutions via discrete wavelet transform (DWT). Each component is then processed by the stacked bidirectional LSTM to model the fluctuations in both directions. On the basis of Yildirim 's work, Liu et al. [113] propose a heartbeat classification model by integrating stacked bidirectional LSTM and 2D CNN. Liu et al. acknowledge the advantages of DWT decomposition and stacked bidirectional LSTM, and they accept this idea in their model. Moreover, they apply a 2D convolution layer to summarize the fluctuation patterns learned by the stacked bidirectional LSTM layers. There is also a work proposed by Liu et al. on integrating LSTM and CNN for heartbeat classification [114]. In this work, instead of using DWT, Liu et al. apply the ensemble empirical mode decomposition to decompose the segmented heartbeats into N components. The first $N/2$ components are processed by 2D CNN layers and the remaining components are processed by stacked bidirectional LSTM layers. The output of both the CNN and LSTM layers are gathered in the fusion layer to give a classification decision.

All models discussed in the above paragraph are evaluated with the MIT-BIH arrhythmia database in their original studies and reported to have supremely good heartbeat classification performances. However, the reported results are over-optimistic because the evaluations suffer from data leakage. It is worth to note that a large dataset is highly requested to guarantee enough data samples for training, validating, and testing a DNN model. Since ECG signals are very imbalanced in heartbeat types, as mentioned in Sec. 1.2, the discussed DNN-based works use synthetic heartbeats to complement the shortage of abnormal heartbeats in MIT-BIH arrhythmia database. The synthetic samples are mixed with the original ones to be allocated into any of the training, validating, or testing set. That means, the heartbeat samples from the same patients have high chances to appear in both training and test datasets. In this case, the discussed models would learn the particularities of the patient's heartbeats during the training and obtain over-optimistic results on test heartbeats. Therefore, the reported results can not truly reveal how these models will perform in real-world scenarios. Additionally, the biased evaluation process may hide a possible limitation of the discussed DNN-based works, since these works accept segmented raw ECG heartbeats as the only input. The heartbeat rhythm information is not properly considered. As mentioned in Sec. 1.2, the previous-RR-interval provides indispensable rhythm information to indicate the difference between the S-type arrhythmia heartbeats from the normal ones. If such information is not counted into consideration, a high misclassification rate is probably obtained for the S-type arrhythmia heartbeat detection.

A recent research attempt to solve the aforementioned problems is done by Sellami et al. [157] in which a 5-layer CNN with residual connections is proposed. Unlike Acharya's work that generates synthetic heartbeats to overcome the class imbalance problem, this work proposes a batch-weighted loss function for imbalanced data. Moreover, the model considers the rhythm information by including the neighbor heartbeats into the analysis. To reveal the true model performance, the model is evaluated on the MIT-BIH-AR database following the inter-patient paradigm proposed in [38]. This evaluation paradigm balances the heartbeat distribution in training and test set. More importantly, it alleviates the data leakage problem. Details of this evaluation paradigm will be discussed in Sec. 2.3. The experimental results demonstrate the improvements brought by Sellami's model in detection of the disease heartbeats. However, this is at the cost of the classification performance of normal heartbeats. This implies that the model misclassifies a large portion of normal heartbeats as the disease heartbeats. In real-world practice, the erroneous classification of normal heartbeats could lead to unnecessary additional tests, unnecessary patient treatments, expensive costs, and risks for patients. Moreover, in Sellami's work, model evaluation is only performed on the benchmark database. There is a lack of a model generalization ability evaluation by applying the proposed model on a broader range of ECG databases.

Although deep learning has introduced both new techniques and ideas to solve the heartbeat classification problem, there is still a long way to go before developing a practical DNN-based model that can create substantial impacts on clinical practices.

2.3 Evaluation of A Heartbeat Classification Model

The Association for Advancement of Medical Instrumentation (AAMI) specifies several ECG databases to evaluate heartbeat classification models. The MIT-BIH Arrhythmia database (MIT-BIH-AR) [137] is the most representative one which is presented in most of the automated arrhythmia detection works. It is unique because it contains all the five arrhythmia-related heartbeats groups proposed by AAMI. The database contains 48 two-leads ambulatory ECG records from 47 patients (22 females and 25 males). Each record has approximately 30 minutes in length. These recordings were digitized at $360Hz$. For most of them, the first lead is *modified limb lead II* (except for the recording 114). The second lead is a pericardial lead (usually $V1$, sometimes are $V2$, $V4$ or $V5$, depending on subjects).

Performance metrics have also been specified by the AAMI. Unlike many classification problems in which only the overall accuracy is used as a performance indicator, multiple metrics are needed to comprehensively evaluate a heartbeat classification model. The used metrics include Sensitivity (also known as recall), Positive predictivity (also known as precision), and Overall accuracy, etc. This is because most peoples' ECG recordings are very imbalanced in heartbeats distribution, with more than 90% being normal heartbeats. The imbalance can lead to a strong distortion of the overall accuracy. Moreover, how many arrhythmia-related heartbeat are correctly recognized and how many predictions are correct among all suspected arrhythmia-related heartbeats are of interest to medical professionals because missing a disease heartbeat can cause serious consequence to patients. Therefore, the Sensitivity and the Positive predictivity are particular important in the heartbeat classification task. The metrics are formally defined as follows.

$$Recall = \frac{TP}{TP + FN} \tag{2.24}$$

$$Precision = \frac{TP}{TP + FP} \tag{2.25}$$

$$Accuracy = \frac{TP + TN}{\sum} \tag{2.26}$$

where TP, TN, FP and FN denotes true positive, true negative, false positive and false negative, respectively, and $\sum$ represents the amount of instances in the data set.

Although benchmark databases and evaluation metrics are specified, there is a lack of a specific protocol which clearly indicates how to perform evaluation with the databases. That is, the AAMI does not specify which recordings or heartbeats should be used to train a model and which should be used to test the method. This problem can not be simply solved by performing cross-validation on recordings or individual heartbeats. Performing cross-validation on recordings is impractical because it is difficult to guarantee even heartbeat distributions of the training and the test sets. It is possible that the training set does not contain sufficient arrhythmia-related heartbeats for a model to learn the abnormal patterns, or that the test set does not include enough arrhythmia-related heartbeats to test the model disease detection ability. Performing cross-validation on individual heartbeats is referred as the *intra-patient* paradigm in the literature [38]. The paradigm is adopted in many works [10, 18, 24, 26, 44, 52, 60, 81, 91, 102, 141, 145, 176, 187, 198, 205, 206]. However, this paradigm can lead to data leakage because it makes heartbeats from the same recording, which are highly correlated, possibly be used for training and testing a model at the same time. Therefore, the paradigm tends to over-estimate a heartbeat classification model.

To reveal the true performance a heartbeat classification model, de Chazal et al. [38] propose an evaluation protocol, known as *inter-patient* paradigm, to divide the MIT-BIH Arrhythmia database into a training set and a test set. The paradigm balances the heartbeat distribution on both training and test sets. More

Table 2.3: The inter-patient paradigm

Data set	N	S	V	F	Q	Recordings (Patient ID)[1,2]
DS1	45808	943	3786	414	8	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	44198	1836	3219	388	7	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

[1] Each recording in MIT-BIH-AR is denoted by a 3-digits number and the numbers are originally discontinuous.
[2] Recordings 102, 104, 107 and 217 containing paced beats are excluded from analysis [8].

importantly, it avoids the training and the test heartbeats coming from the same patient. The *inter-patient* paradigm simulates arrhythmia detection scenarios in real world, in which test heartbeats are usually unknown to the existing automated algorithms or methods. It makes the heartbeat classification task more challenging. Table 2.3 presents the *inter-patient* paradigm in details, where **DS1** is the training set and **DS2** is the test set.

Compared to the *intra-patient* paradigm, those works adopting the *inter-patient* paradigm report less promising but more realistic results in the literature. In order to develop a practical automated arrhythmia detection model, the author follows the *inter-patient* paradigm in all research works presented in this thesis.

2.4 IoT-based ECG Monitoring System

The rapid development of the *off-the-person* ECG sensing approaches facilitates the growth of IoT-based ECG monitoring systems. Basically, an IoT-based ECG monitoring system consists of three parts: the ECG sensing network, IoT cloud, and graphical user interface [197].

The ECG sensing network is responsible for tracking patients' heart status, generating ECG recordings correspondingly, and transmitting the produced data to the IoT cloud.

The IoT cloud provides four functions including data cleaning, data storage, ECG analysis and disease warning. In particular, in data cleaning module, ECG signals are preprocessed to correct the anomalies and removes the noises that are introduced during the ECG collection and transmission. Data storage provides a storing service to store the large volume of ECG signals and a fast-access service to the stored ECG signals. ECG analysis plays a vital role in an IoT cloud. It performs ECG signal interpretation and heart disease detection, such as arrhythmia detection [4, 7, 23, 25, 83, 119] and sleep apnea detection [158]. If any anomaly is detected during the ECG analysis stage, the analytical results will be summarized and prompted to medical professionals for further conformation.

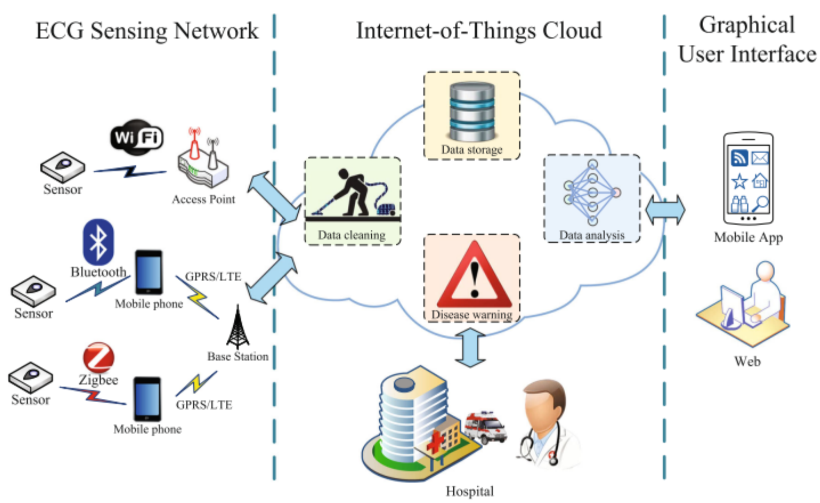


Figure 2.5: Overview of an Iot-based ECG monitoring system [197].

Finally, regarding the GUI module, it is used for data visualization and patients' self-management. It allows patients to track their health status conveniently.

It is worth noting that data security and privacy is an important concern for a IoT-based ECG monitoring system [184]. The ECG recordings stored in the cloud may suffer from various types of security threats such as linking attacks [168], unauthorized access. Therefore, in order to ensure the security and privacy of IoT-based ECG recordings, a proper data access control module, such as [80, 106], should be integrated when constructing a IoT-based ECG monitoring system.

CHAPTER 3

D-ECG: A DYNAMIC FRAMEWORK FOR AUTOMATIC CARDIAC ARRHYTHMIA DETECTION FROM IOT-BASED ECG RECORDINGS

3.1 Chapter Abstract

Cardiac arrhythmia has been identified as a type of cardiovascular diseases (CVDs) that cause approximately 12% of all deaths globally. The current progress on arrhythmia detection is facing a bottleneck for adopting single classifier and static ensemble methods. Besides, most of the work tends to use a static feature set for characterizing all types of heartbeats, which may limit the classification performance. To fill in the gap, a novel framework called D-ECG is proposed in this chapter to introduce dynamic ensemble selection technique (DES) to provide accurate detection of cardiac arrhythmia. In addition, the proposed D-ECG develops a result regulator that uses different features to refine the classification result from the DES technique. The results reported in this paper have shown visible improvements on the overall heartbeat classification accuracy as well as the sensitivity of disease heartbeats.

3.2 Introduction

Cardiac arrhythmia is a type of cardiovascular diseases (CVDs) that seriously affects millions of people around the world and accounts for approximately 80% of the sudden cardiac death [190]. At critical levels, arrhythmia can be categorized as life-threatening and non-life-threatening arrhythmia [209]. Life-threatening arrhythmia imposes an imminent risk to patients' lives and emergency treatment is required, whereas non-life-threatening arrhythmia just presents a long-term health threat to patients but special care is still needed to avoid further deterioration of heart function.

ECG provides a noninvasive and inexpensive way to study arrhythmia. It records the electrical activities of the heart. The Advancement of Medical Instrumentation (AAMI) categorizes ECG heartbeats into 5 super-classes: Normal (N), Supraventricular ectopic beat (S), Ventricular ectopic beat (V), Fusion beat (F) and Unknown beat (Q) [8]. Most arrhythmia are found related to the S and V-type heartbeats.

With a rapid growth in Internet-of-Things (IoT) techniques, more and more wearable ECG monitoring devices are developed to produce high-quality ECG recordings [197]. In comparison to the traditional Holter device, the IoT-based ECG monitoring devices provide a more human-friendly way for heart status tracking. They produce ECG recordings continuously and upload the recordings to the IoT cloud in real time, making it easy and convenient for both clinicians and patients to access the ECG recordings. However, manual interpretation of the large amount of continuously generated ECG recordings could be very time-demanding and error-prone. Therefore, a computer-assisted method is always needed to help analyze and interpret the ECG signals. This work aims to develop an automatic method to help detect arrhythmia-related heartbeats from the IoT-based ECG recordings.

Many research attempts have been made to address this problem. Current methods are mainly facing a bottleneck for adopting a single classifier trained with a predefined feature set [23, 37, 38, 204], which may bias the classification and lead to a relatively low generalization performance. This is because that the value of features, such as the RR interval, could vary significantly from patients to patients (even in the same patient) while a single classifier is believed to be an expert only in certain local regions of the feature space [32, 210]. Although some ensemble methods, such as *random forest* [4] and *ensemble of support vector machine* [74], have been employed to remedy the disadvantages, the problem is only partly solved because the diversity of the traditional ensembles is relatively low. Moreover, similar to the use of a single classifier, the classifiers used to construct an ensemble are determined in the training phase. It is not guaranteed that the selected classifiers are suitable for making predictions for the input heartbeat. Besides, using a static set of features for classifying all types of heartbeats can also limit the classification performance because the discriminatory power of a feature depends on the types of the heartbeats involved [209]. The inclusion of less discriminatory features can introduce difficulties in recolonization of disease heartbeats. For example, RR-intervals help to distinguish the disease heartbeats from the normal ones. However, the presence of multiple RR-intervals can also confuse the classification between the S-type and the V-type heartbeats.

Apart from the above mentioned problems, the performance of arrhythmia detection also relates heavily to the training data preparation, since ECG heartbeat data set is naturally extremely imbalanced, where normal heartbeat is the dominant. A proper measure is a necessity to eliminate the negative effect cause by the imbalance [186]. Otherwise, the trained classifier may have difficulties in correctly recognizing the disease heartbeats which only accounts for a small portion of the whole heartbeat data set [23].

This chapter aims to solve the above problems by proposing a dynamic framework called D-ECG for cardiac arrhythmia detection. The D-ECG introduces the *dynamic ensemble selection* (DES) technique to improve the disease detection accuracy. The DES technique works by estimating a competence level of each classifier from a pool of classifiers in the training phase. When testing, according to each heartbeat to be classified, the DES technique selects the most competent classifiers to predict the heartbeat label. To the best of our knowledge, this is the first time that the DES technique is used in cardiac arrhythmia detection scenario. Specifically, the proposed D-ECG consists of five phases: preprocessing, feature extraction, classifier pool training, dynamic selection classification and result refinement. In the classifier pool training stage, the *synthetic minority oversampling* as well as the *edited nearest neighbors* technique (*SMOTEENN*) is adopted to remedy the negative effect caused by the data imbalance of the ECG database. Besides, a result regulator is creatively developed to refine the result from the previous phase. Essentially, the regulator is a classifier trained with a feature set which is different from the one used for DES training.

As a summary, this chapter makes the following contributions:

- Proposing a dynamic framework named D-ECG, which first introduces the DES techniques for automatic cardiac arrhythmia detection.
- Customizing a result regulator to improve the heartbeat classification performance.
- Experimentally comparing the performance of various DES techniques in ECG-based heartbeats classification.
- Experimentally evaluating the feasibility of D-ECG in arrhythmia detection, with the results being compared against the stat-of-the-art methods in the same field in terms of overall accuracy, sensitivity and positive predictive.

The rest of this work is structured as follows. Section 3.3 analyzes the pros and cons of current methods in arrhythmia detection and introduces the background knowledge of IoT-based ECG monitoring system and Dynamic ensemble selection techniques. Section 3.4 details the proposed D-ECG. Experiment results and discussion are presented in Section 3.5. Section 3.6 concludes this work and discusses the future work.

3.3 Related Work

This section firstly introduces the IoT-based ECG monitoring system and then reviews current methods in arrhythmia detection. After that, the dynamic ensemble technique is presented in detail.

3.3.1 IoT-based ECG Monitoring System

An IoT-based ECG monitoring system mainly consists of three parts: an ECG sensing network, an IoT cloud, and a graphical user interface (GUI) [197]. The ECG sensing network is responsible for generating ECG recordings for patients and transmitting the produced data to the IoT cloud. The IoT cloud is mainly used for data storage and analysis. Specifically, the IoT cloud performs ECG signal interpretation and heart disease detection, such as arrhythmia detection [4, 23, 37, 60, 119, 154, 209] and sleep apnea detection [4, 7, 23, 25, 83, 158]. The GUI module is often used for data visualization and management.

Data security and privacy are important concerns for a IoT-based ECG monitoring system [184], since the ECG recordings stored in the cloud may suffer from various types of security threats, such as linking attacks and unauthorized access [168]. In order to ensure the safety and privacy of the ECG recordings, security mechanisms, such as anonymization [169, 208] and access control [80, 106, 167, 180, 181, 183], should be considered when designing an IoT-based ECG monitoring system.

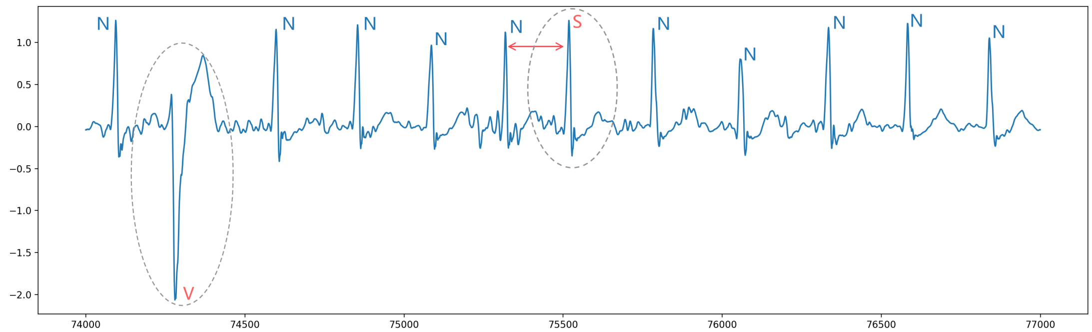


Figure 3.1: A sample ECG recording that contains normal beats (N), supraventricular (S) and ventricular ectopic (V) beats.

3.3.2 Current Methods, Achievements and Problems

Since the life-threatening arrhythmia has been well studied [25], this study focuses on the investigation of non-life-threatening arrhythmia (Supraventricular ectopic beat (S-type)) and the related ectopic heartbeats (Ventricular ectopic beat (V-type)). Fig.3.1 presents a sample ECG segment that contains normal heartbeats (N), supraventricular and ventricular ectopic beats. It can be clearly seen that N-type and S-type beats are difficult to tell from their shapes. In fact, N-type and S-type beats are similar in most of the heartbeat characteristics except for the RR intervals, where S-type beats generally have a shorter *previous*-RR compared to the N-type beats from the same patient. On the contrast, V-type beats significantly differ from N-type and S-type beats in terms of the heartbeat morphology.

Current studies on cardiac arrhythmia detection mainly focused on signal feature extraction and selection and the use of various classifiers combinations [119]. In terms of feature extraction, features extracted from cardiac rhythm or the time and frequency domain are usually combined to characterize a heartbeat. The commonly used features include RR intervals [4, 23, 45, 209], samples or segments of ECG curves [145,187], higher-order statistics [4,40], wavelet coefficients [37,60,154], signal energy [204]. In consideration of the negative impact of having irrelevant features, some works have applied techniques to reduce the feature space, such as the *floating sequential search* [115,124] and the *weighted LD model* with a forward-backward search strategy [46]. Although the extracted features in previous work have been proven to be effective in characterizing an ECG-based heartbeat, the classification performance is limited by using a static set of features for classifying all types of heartbeats.

Regarding the choice of classifiers for arrhythmia classification, the *support vector machine* (SVM) is the most widely used for its robustness, good generalization and computationally efficiency [1, 35]. Besides, the *nearest neighbors* (NN) and *artificial neural networks* (ANN) are also frequently found in the literature. Other classifiers include *weighted linear discriminant* (WLD), *decision tree*, *optimum-path forest* (OPF). Nevertheless, the use of a single classifier can bias the classification and lead to a relatively low generalization performance. Although some ensemble methods have been employed to remedy the disadvantages, the problem can only be partly solved.

3.3.3 Dynamic Ensemble Selection

DES is one of the promising approaches to construct a multiple classifier system (MCS). Recently, more and more works are reporting the superior performance of the DES over the static methods [15,32]. A DES-based system is composed of three stages: generation, selection and aggregation [15]. In the generation stage, a collection of classifiers are trained to create an accurate and diverse classifier pool. In the selection stage, an ensemble containing the most competent classifiers is selected. Finally, in the aggregation stage, the output of each classifier in the selected ensemble are aggregated to give the final decision of the system.

The core issue of DES techniques is the selection of the most competent classifiers for any testing sample [33]. Usually, the competence of a classifier in the pool is measured by its performance over a local region of the feature space where the testing sample is located. This means that, the competence level of a base classifier not only relates to the performance metrics but also depends on the testing sample and its neighbors in the feature space. Methods for defining a local region includes clustering [111], K-nearest neighbors [160], potential function model [194, 195] and decision space [16]. The criterion for measuring the performance of a base classifier can be divided as individual-based and group-based criterion. In individual-based criterion, each base classifier is independently measured on use of the metrics such as ranking, accuracy, probabilistic, behavior [16], meta-learning [32], etc., whereas in the group-based criterion, the performance of a base classifier relates to its iterations with other classifiers in the pool. For example, diversity, data handling [196] and ambiguity [47] are widely used as group-based performance metrics.

Regarding the aggregation approaches, there are three main strategies for results combination: static combiner, trained combiner and dynamic weighting. The majority voting scheme is a representative static combiner, which is also commonly used in the traditional ensemble methods. In trainable combiners, the outputs of the selected based classifiers are used as the input features for another learning algorithm, such as [15, 129]. The advantage of this strategy is that the combiner can be well modified to adapt to the specificity of different classification problems. In dynamic weighting, higher weight value will be allocated to the most competent classifier and then the outputs of all the weighted classifiers are aggregated to give the final decision.

Currently, the most prevalent DES techniques include the DES-KL [195], DES-KNN [162], KNORA-E [95], KNORA-U [95], KNOP [16], DES-P [195], DES-RRC

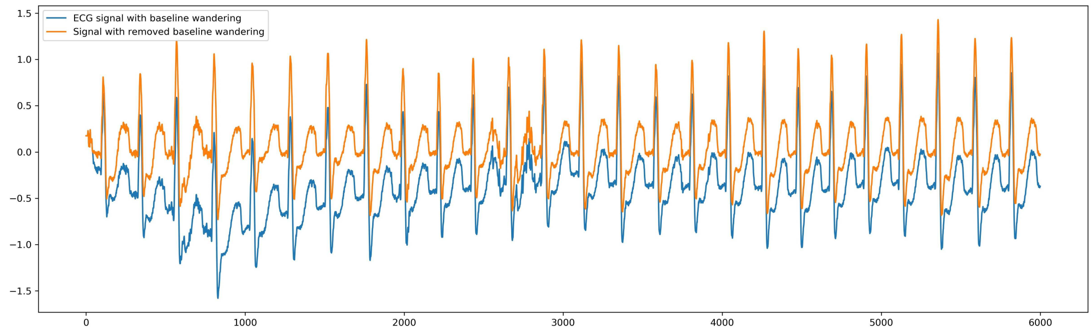


Figure 3.2: A sample ECG record with background noise and baseline wandering. The background noise imposes a strong vibration to the original signal, whereas the baseline wandering makes the individual heartbeats move up and down instead of being strait in the X-axis.

[194], META-DES [16], etc. The differences between these DES techniques are mainly in the methods for defining a local region and the classifier competence measure. In this work, we experimentally evaluate the effectiveness of these DES techniques in cardiac arrhythmia detection.

3.4 Dynamic ECG Framework

Details of the proposed D-ECG are presented in this section. Essentially, the D-ECG is composed of five phases: preprocessing, feature extraction, classifier pool training, dynamic selection classification and result refinement.

3.4.1 ECG Data Preprocessing

Given the fact that ECG signals usually come with serious background noise and baseline wandering, as shown in Fig.3.2, proper measures should be taken to eliminate the negative effect caused by the noises on cardiac arrhythmia detection.

To remove the baseline wandering, each ECG signal is processed with a 200-ms width median filter followed by a 600-ms median filter to obtain the signal baseline. The baseline is then subtracted from the raw ECG signal to get the baseline corrected data.

After that, the discrete wavelet transform is employed to remove the background noise from the baseline corrected signals. The signals are decomposed to different frequency bands with various resolutions on use of the Daubechies-4 (DB4) mother wavelet function [36]. The reason for choosing DB4 is that its short vanishing moment is ideal for analyzing signals like ECG with sudden changes. The coefficients of detail information (cD_x) in each frequency band are then processed by a high-pass filter with a threshold value

$$T = \sqrt{2 * log(n)} \tag{3.1}$$

,where n is the length of the input signal. Those coefficients that failed by the filter are set to zero. Finally, the clean signals are obtained by employing inverse discrete wavelet transform on the coefficients.

The clean signals are segmented to individual heartbeats by taking advantage of the R peak locations. For each R peak, 90 samples (250 ms) before R peak and 144 samples (400 ms) after R peak are taken to represent a heartbeat, which is long enough to catch the samples representing the re-polarization of ventricles and short enough to exclude the neighbor heartbeats [4].

3.4.2 Feature Extraction

Noticing a fact that disease heartbeats can cause disorders to heartbeat shape and heart rhythms in ECG signal. In order to effectively catch these anomalies, three types of features are used to represent individual heartbeats: RR-intervals, higher order statistics and wavelet coefficients.

RR-interval is the time distance between two successive R peaks. As experimentally proven in [209], the RR interval is one of the most indispensable features for heartbeat classification and it has great capacity to tell the supraventricular premature beats and the ectopic beats from the normal beat. In this work, four types of RR intervals are extracted from ECG signals: pre_RR, $post\_RR$, $local\_RR$ and $global\_RR$ [119].

However, it should be noted that the RR intervals can significantly vary from patient to patient, known as the inter-patient variations. To reduce the negative impacts, we normalize the RR intervals for each heartbeat in the way below:

$$nomalized\_pre\_RR = \frac{pre\_RR}{mean(ds.pre\_RR)} \tag{3.2}$$

$$nomalized\_post\_RR = \frac{post\_RR}{mean(ds.post\_RR)} \tag{3.3}$$

$$nomalized\_local\_RR = \frac{local\_RR}{mean(ds.local\_RR)} \tag{3.4}$$

$$nomalized\_global\_RR = \frac{global\_RR}{mean(ds.global\_RR)} \tag{3.5}$$

where $ds.pre\_RR$ denotes the average of all pre_RRs in the ds which a heartbeat belongs to, and so on.

Regarding the *higher order statistics* (HOS), it is reported to be useful in catching subtle changes in ECG data [125]. In this work, the skewness (3rd order statistics) and kurtosis (4th order statistics) are calculated for each heartbeat. They can be mathematically defined as follows, where $X_{1...,N}$ denotes all the data samples in a signal, $\bar{X}$ is the mean and s is the standard deviation.

$$Skewness = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^3/N}{s^3} \tag{3.6}$$

$$Kurtosis = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^4/N}{s^4} - 3 \tag{3.7}$$

The wavelet coefficients provide both time and frequency domain information of a signal, which is claimed to be the best features of ECG signal [119]. The choice of the mother wavelet function used for coefficients extraction is crucial to the final classification performance. In this work, the Haar wavelet function is chosen because of its simplicity and that it has been demonstrated as an ideal wavelet for short time signal analysis [204]. The Haar function can be represented as

$$\psi(t) = \begin{cases} 1 & 0 \leq t<1/2, \\ -1 & 1/2 \leq t<1, \\ 0 & otherwise. \end{cases} \tag{3.8}$$

and its corresponding scaling function is

$$\phi(t) = \begin{cases} 1 & 0 \leq t<1, \\ 0 & otherwise. \end{cases} \tag{3.9}$$

where t denotes sample values.

3.4.3 Classifier Pool Training

A classifier pool, containing a set of base classifiers that are trained both accurately and diversely, is created in this stage. First of all, the SMOTEENN technique [6,21,191] is adopted to remedy the training data imbalance problem. The minority classes (the S-type and V-type heartbeats) are up-sampled to the same amount of normal heartbeats.

In order to increase the diversity of the classifier pool, we take 6 different classifiers into consideration, including multi-layers perceptron, support vector machine (SVM), linear SVM, Bayesian model with Gaussian kernel, decision tree, and K-nearest neighbors model. The classifiers are trained using different training subsets.

A dynamic selection dataset named DSEL is split from the training set for ensemble selection. Given an unknown heartbeat, we calculate its local region within the scope of DSEL. The competence of each base classifier is then measured and compared in the calculated local region.

3.4.4 Dynamic Selection Classification

In this stage, a dynamic ensemble selection technique is equipped into the framework. There are several prevalent DES techniques which have been proven their effectiveness in some classification problems. The differences between them are summarized in Table 3.1. Given a well-defined local region in DSEL,

- **DES-KNN** selects the top N accurate classifiers and top J diverse classifiers to compose the ensemble.
- **DES-KL** measures the competence level using

$$\sigma_{i,j} = \sum_{x_k \in DSEL} C_{src} exp(-d(x_k, x_j)^2), \tag{3.10}$$

  where x_j is a query sample, C_{src} is the KL divergence between the uniform distribution and the vector of class supports.
- **DES-P** selects the classifiers that have a better accuracy than a random classifier into the ensemble.
- **DES-RRC** measures the competence level using

$$\sigma_{i,j} = \sum_{x_k \in DSEL} C_{src} K(x_k, x_j), \tag{3.11}$$

  where x_j is a query sample, C_{src} denotes the source competence proposed in [193] and $K(x_k, x_j)$ is a Gausssian potential function. The classifiers that have a higher competence level than a random classifier are selected.

Table 3.1: A brief comparison between prevalent DES techniques

Technique	Local region definition	Competence measure	Reference
DES-KNN	K-NN	Accuracy & Diversity	Soares et al. [162]
DES-KL	Potential function	Probabilistic	Woloszynski et al. [195]
DES-P	Potential function	Probabilistic	Woloszynski et al. [195]
DES-RRC	Potential function	Probabilistic	Woloszynski et al. [194]
KNORA-E	K-NN	Oracle	Ko et al. [95]
KNORA-U	K-NN	Oracle	Ko et al. [95]
KNOP	K-NN	Behavior	Cavalin et al. [16]
META-DES	K-NN	Meta-Learning	Cruz et al. [32, 32]

- **KNORA-E** selects the classifiers that correctly recognize all samples in the local region. If no base classifier is qualified, the local region shrinks.

- **KNORA-U** selects all classifiers that correctly recognize at least one sample in the local region. A majority voting scheme is used to give the final result. In the voting stage, a selected classifier is allowed to vote more than once. The number of votes depends on the amount of samples the classifier correctly recognizes.

- **KNOP** works similarly to KNORA-U. The difference is that KNORA-U works in the feature space, whereas KNOP works in the decision space [16]. That is, in KNOP, all samples in MIT-BIH-AR are transform from the feature space to the decision space in advance.

- **Meta-DES** considers a base classifier as competent or incompetent by defining a set of meta features for the classifier and training a meta-classifier which takes in the meta features to predict if a base classifier is competent.

In this work, we individually evaluate the performance of these DES techniques in cardiac arrhythmia detection and choose the best one to fit in the proposed framework.

3.4.5 Result Refinement

A result regulator is developed in this stage. It is used for refine the classification decisions made on disease heartbeats. That is, if an unknown heartbeat is classified as S-type or V-type in the dynamic classification stage, it is then passed through the regulator for decision refinement.

Specifically, the result regulator is a SVM classifier trained with a feature set which is different from the one used for DES training. The motivation behind constructing such a regulator is two-folds: (1) it is reported that the discriminatory power of a feature depends on the types of the heartbeats involved [209]; (2) the S-type heartbeats tend to be recognized as V-type beats under the presence of multiple heart rhythm features. In order to precisely catch the differences between S-type and V-type heartbeats, we exclude the RR-intervals from the feature set and only use the S-type and V-type beats to train the result regulator.

3.5 Experiment Result Analysis and Discussion

This section starts with the introduction of the MIT-BIH Arrhythmia benchmark database, followed by the specification of the experiment settings. Next, we individually evaluate the heartbeat classification performance of each DES technique presented in Table 3.1. The performances are then compared to the single classifier and the traditional ensembles. After that, we incorporate the best performing DES technique into the proposed D-ECG and evaluate the whole model on the MIT-BIH Arrhythmia database to see whether the result regulator helps improve the disease heartbeats classification performance. Finally, we compare the proposed D-ECG against the stat-of-the-art methods.

3.5.1 The MIT-BIH Arrhythmia Database

The MIT-BIH arrhythmia database [137] contains 48 two-lead ambulatory ECG recordings from 47 patients (22 females and 25 males), with each recording approximately 30 minutes in length. These recordings were digitized at 360 Hz per second per channel with 11-bit resolution over a 10-mV range. For most of the recordings, the first lead is modified limb lead II (except for the recording 114 which used $V5$ as the first lead and MLII as the second). The second lead is a pericardial lead (usually $V1$, sometimes are $V2$, $V4$ or $V5$, depending on subjects).

Intra-patient and inter-patient paradigm [4, 38, 200, 209] are two different types of paradigms concerning the use of the MIT-BIH-AR database for performance evaluation. The intra-patient paradigm separates the entire data set into a training set and a testing set merely according to the heartbeat labels, whereas inter-patient paradigm groups the heartbeats by patients and partitions the patients into a training set (DS1) and a testing set (DS2), as shown in Table 3.2. It has been empirically proven that the intra-patient paradigm can produce an over optimistic classification result by allowing training and testing heartbeats coming from the same patient [119]. Therefore, in order to reveal the true performance of the proposed model and have a fair comparison with the state-of-the-art rivals, the inter-patient paradigm is strictly followed in this work.

On the basis of inter-patient paradigm, we further split DS1 into a training set and a dynamic selection data set (DSEL), which account for 70% and 30%, respectively. In addition, we generate six different subsets from the training set to train the base classifiers in the ensemble pool.

Table 3.2: Recording distributions and class proportions in DS1 and DS2

Dataset	N	S	V	F	Q	Recordings[1]
DS1	45808	943	3786	414	8	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	44198	1836	3219	388	7	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

[1] Each recording is denoted by a 3-digits number and the numbers are originally discontinuous.

3.5.2 Experiment Settings

The experiments presented in this work are programmed in Python 3.63 and run in a 64-bits Ubuntu 16.04 platform with an $i7-7700k$ CPU and 16GB memory.

The metrics used for classifier performance evaluation are sensitivity (Se), positive predictive value ($+P$) and accuracy value (Acc). It is worth to note that penalties should not be applied for the misclassification of class F-type and Q-type, as recommended by the AAMI standard [38].

3.5.3 Comparative Analysis of DES Techniques

The effectiveness of the DES techniques in arrhythmia detection is evaluated in this subsection. They are trained on the same classifier pool as mentioned in section 3.4.3. The used features are RR-intervals, HOS and wavelet coefficients. The results are summarized in Table 3.3.

It can be seen from Table 3.3 that all DES techniques have a similar performance in terms of the overall accuracy and positive predictive of normal heartbeats except DES-KNN. META-DES obtains the best result in overall accuracy, sensitivity of normal heartbeats and positive predictive of S-type heartbeats. KNORA-U

Table 3.3: Heartbeats classification performance of DES techniques, single classifiers and traditional ensembles.

Technique	Classifiers	Acc	N		S		V	
			Se	+P	Se	+P	Se	+P
DES	DES-KNN	81.79	82.15	99.28	68.7	28.1	93.72	35.9
	DES-KL	90.39	91.86	98.98	65.62	32.44	94.78	61.38
	DES-P	90.42	91.87	99.01	65.56	32.46	95.22	61.44
	DES-RRC	89.4	90.81	99.01	64.37	28.58	94.81	60.4
	KNORA-E	89.43	90.74	99.04	65.11	30.72	95.65	57.85
	KNORA-U	89.97	91.1	98.77	**74.86**	35.69	93.72	59.67
	KNOP	90.54	92.02	98.88	64.37	35.56	95.62	60.0
	META-DES	**90.77**	**92.35**	98.8	63.63	**36.06**	95.0	60.8
Single Classifier	SVM	81.83	81.35	**99.48**	88.08	30.17	94.91	36.95
	KNN	84.22	84.84	98.5	72.01	22.01	92.73	52.89
	Perceptron	80.22	81.75	98.93	40.99	12.85	89.9	39.53
	Linear SVM	85.08	86.9	99.01	48.52	18.14	89.69	48.28
	Bayesian	79.13	80.52	97.59	64.37	31.06	77.63	28.65
	Decision Tree	78.37	80.38	96.56	52.74	18.37	74.37	39.95
Homogeneous Ensemble	E-SVM (50) [1]	82.76	82.87	99.35	72.92	30.89	**96.68**	36.76
	E-KNN (10)	84.27	84.9	98.62	72.06	22.86	92.54	51.46
	E-Perceptron (50)	85.95	87.98	99.0	43.9	16.44	91.61	52.79
	E-LSVM (50)	86.7	88.96	98.95	46.12	18.55	88.51	52.74
	E-GB (50)	81.58	83.89	96.58	36.66	17.9	84.28	40.19
	Forrest (50)	87.53	89.2	98.59	56.04	33.6	90.37	45.86
Heterogeneous Ensemble	Mixture[2]	89.48	91.13	99.02	61.63	27.0	92.98	**61.7**

[1] The number denotes the amount of classifiers in the ensemble.
[2] The mixture ensemble contains multi-layers perceptron, SVM, linear SVM, Logistic regression, Bayesian model with Gaussian kernel, decision tree, and K-nearest neighbors model.

achieves the best sensitivity of normal heartbeats, which is approximately 6.16% higher than the second best result obtained by DES-KNN. Regarding the detection of V-type heartbeats, all DES techniques have a promising performance, with the lowest sensitivity higher than 93%. However, the corresponding positive predictive values are struggling around 60%.

In order to demonstrate the advantages of DES techniques, we compare them against single classifier and traditional ensembles methods. We categorize the traditional ensemble methods into homogeneous ensemble and heterogeneous ensemble. The former generates an ensemble with a numbers of certain classifiers that trained with different training sets, while the later contains various types of classifiers. In traditional ensemble methods, the final decision is given based on the majority voting scheme. The experimental results are presented in Table 3.3 as well. Obviously, the average performance of DES techniques is significantly better than that of single classifier. Though the homogeneous ensemble methods have make some improvements to the corresponding single classifier, the overall performance is still lower than that of DES techniques, with no method achieving higher than 88% in accuracy and higher than 90% in detection sensitivity of normal heartbeats. Surprisingly, the mixture ensemble performs closely to the average performance of DES techniques. However, since the number of classifiers in the ensemble is small, the performance is extremely sensitive to the choices of classifiers.

3.5.4 D-ECG Performance Evaluation

Since META-DES obtains the best results in overall accuracy, sensitivity of normal heartbeats and positive predictive of S-type heartbeats, and performs closely to other DES techniques in the rest of the metrics, we incorporate META-DES into the proposed D-ECG structure.

Table 3.4 shows the arrhythmia detection results of META-DES and D-ECG in DS2. The classification decisions of D-ECG are given by applying the result

Table 3.4: Arrhythmia detection result of **META-DES** and **D-ECG** in DS2 of the MIT-BIH-AR database

		Predicted class (**META-DES**)					Predicted class (**D-ECG**)				
		N	S	V	F	Q	N	S	V	F	Q
True class	N	40698	1921	1331	0	119	40698	2053	1199	0	119
	S	125	1116	513	0	0	125	1472	157	0	0
	V	102	58	3058	0	1	102	103	3013	0	1
	F	3	0	4	0	0	3	0	4	0	0
	Q	263	0	124	0	1	263	2	122	0	1

Table 3.5: Arrhythmia detection performance comparison between before- and after-refinement

Refinement	Acc	N		S		V	
		Se	+P	Se	+P	Se	+P
Before	90.77	92.35	98.8	63.63	36.06	95.0	60.8
After	**91.4** ↑	92.35	98.8	**83.92** ↑	**40.55** ↑	93.6	**67.03** ↑

regulator on the decisions of META-DES. We summarize the results and make a straight-forward comparison between them in Table 3.5. It is apparent that the regulator component has made visible improvements to the results obtained by META-DES, with overall accuracy increasing from 90.77% to 91.4%, sensitivity of S-type heartbeats increasing by more than 20%, positive predictive of S-type and V-type increasing from 36.06% and 60.8% to 40.55% and 67.03%, respectively.

The proposed D-ECG is compared to the state-of-the-art rivals in cardiac arrhythmia detection. The comparative results are summarized in Table 3.6. It is clear that the proposed D-ECG achieves the best sensitivity of both S-type and V-type heartbeats, being significantly better than the second highest. In the meantime, the proposed D-ECG maintains the second best overall accuracy and sensitivity of normal heartbeats. Shan's model [23] obtains the highest accuracy and sensitivity on normal heartbeats, but it fails in detection of S-type heartbeats,

Table 3.6: Arrhythmia detection comparison between the proposed D-ECG and the stat-of-the-art rivals in DS2 of the MIT-BIH-AR database

Method	Acc	N		S		V	
		Se	+P	Se	+P	Se	+P
Proposed D-ECG	91.4	92.35	98.8	**83.92**	40.55	**93.6**	67.03
De Chazal [38]	81.9	86.9	99.2	75.9	38.5	77.7	81.9
Ye C [200]	86.4	88.5	97.5	60.8	**52.3**	81.5	63.1
Zhang Z [209]	86.7	88.9	99.0	79.1	36.0	85.5	**92.8**
Shan C [23]	**93.1**	**98.4**	95.4	29.5	38.4	70.8	85.1
Mariano L [115]	78.0	78.0	99.0	76.0	41.0	83.0	88.0

with the sensitivity being merely 29.5%. Although the proposed D-ECG has a relative low positive predictive of both S-type and V-type heartbeats, it is still a more appropriate choice than other listed works for cardiac arrhythmia detection from a clinical point of view. This is because in a clinical environment, misclassification of a normal heartbeat would not lead to a disaster, but missing a disease heartbeat can kill.

3.6 Chapter Conclusion

In this chapter, a dynamic framework named D-ECG for automatic cardiac arrhythmia detection from IoT-based ECG recordings is proposed. The D-ECG introduces the dynamic ensemble selection techniques to improve the heartbeat classification performance for the first time. Specifically, the proposed D-ECG is made up of five phases: preprocessing, feature extraction, classifier pool training, dynamic selection classification and result refinement. In the training stage, we use the SMOTEENN technique to remedy the negative effect of data imbalance problem in ECG signals. Moreover, we propose a result regulator in the last stage of D-ECG to use different features to improve the disease heartbeats classification performance. We conduct extensive experiments to investigate the feasibility of DES techniques in heartbeat classification. The results show that the DES techniques generally have a superior performance than single classifier and traditional ensemble methods in majority evaluation metrics. We also experimentally evaluate the effectiveness of the proposed D-ECG using the benchmark MIT-BIH-AR database and compare the result against the state-of-the-art methods. The results show that the proposed D-ECG brings visible improvements to the heartbeat classification task in terms of overall classification accuracy and the sensitivity of disease heartbeats.

Notwithstanding all contributions that have been made in this work, the proposed D-ECG is still far from perfect. The dynamic nature trades high consuming of computation resources with performance increase. This limits the applications and deployments of D-ECG in a small IoT cloud and in scenarios with high concurrent connections. In the next study, we try to investigate a more efficient model by considering the neighbor information of heartbeats.

CHAPTER 4

A PYRAMID-LIKE MODEL FOR HEARTBEAT CLASSIFICATION FROM ECG RECORDINGS

4.1 Chapter Abstract

Heartbeat classification is an important step in the early-stage detection of cardiac arrhythmia, which has been identified as a type of cardiovascular diseases (CVDs) affecting millions of people around the world. The current progress on heartbeat classification from ECG recordings is facing a challenge to achieve high classification sensitivity on disease heartbeats with a satisfied overall accuracy. Most of the works take individual heartbeats as independent data samples in processing. Furthermore, the use of a static feature set for classification of all types of heartbeats often causes distractions when identifying *Supraventricular Ectopic beats*. In this work, a pyramid-like model is proposed to improve the performance of heartbeat classification. The model separates the classification of normal and supraventricular ectopic beats from the overall heartbeat classification, and takes advantage of the neighbor-related information to assist identification of supraventricular ectopic bests. We evaluate the proposed model with the benchmark *MIT-BIH-AR* database and the *St. Petersburg Institute of Cardiological Technics*(INCART) database. The results show that the proposed pyramid-like model exhibits higher performance than the state-of-the-art methods in the identification of disease heartbeats.

4.2 Introduction

An electrocardiogram (ECG) is a recording of the electrical activity of the heart over a period of time. It provides a noninvasive and inexpensive way for studying the heart. Heartbeat classification is one of the important fields in ECG analysis. The Association for Advancement of Medical Instrumentation (AAMI) categorized heartbeats into 5 super classes: Normal(N-type), Supraventricular (S-type) ectopic, Ventricular (V-type) ectopic, Fusion (F-type) and Unknown (Q-type) beats [8]. Heartbeat classification is an essential step toward identifying arrhythmia. Arrhythmia affects the body by impacting heart's ability to pump blood. Critically, arrhythmia can be divided as life-threatening and non-life-threatening arrhythmia [209]. For example, ventricular fibrillation and tachycardia are life-threatening arrhythmia, which are fatal and require medical attention immediately. Non-life-threatening arrhythmia, such as atrial fibrillation, just presents a chronic health threat to patients, but special care is still needed to avoid further deterioration of heart function.

Although to perform an electrocardiography test is simple, the manual interpretation of ECG recordings could be time-consuming and error-prone, especially for the long-term ECG recordings. Hence, an intelligent approach on automatic heartbeat classification from ECG recordings is highly demanded, which would be of great assistance for clinicians in heart diseases diagnosis.

Many research attempts have been made to address the heartbeat classification problem. The current process has difficulties in guarantying a high detection sensitivity of disease heartbeats as well as maintaining a good overall classification accuracy. Most of the existing works take heartbeats as mutual-independent data samples, with no connections to their predecessors or successors [38, 115, 120, 200, 209]. Therefore, the neighbor-related information is ignored in their models. In addition, the use of a single static feature set to classify all types of heartbeats together may cause a high misclassification rate on S-type beats in particular.

To develop a practical heartbeat classification model, a number of factors need to be considered: (1) ECG recordings are imbalanced and usually dominated by the normal heartbeats; (2) Some shape-related features must be included to distinguish the V-type heartbeats from the normal heartbeats because these two heartbeat types have different QRS complexes; (3) The normal and S-type heartbeats are similar in QRS complex morphology, but the S-type heartbeats have a fast heart rhythm. In other words, the existence of the shape-related features makes a S-type heartbeat be easily misidentified as a normal beat. In this study, we aim to propose a pyramid-like model to solve these problems and improve the heartbeat classification performance.

The rest of this work is structured as follows. Section 4.3 reviews current methods in arrhythmia detection and introduces related techniques used in this work. Section 4.4 presents the proposed pyramid-like model. Section 4.5 introduces the experimental ECG databases. Experiment results and discussion are presented in Section 4.6. Section 4.7 concludes this work and discusses the future work.

4.3 Related Work

In this section we review the related studies in heartbeat classification from ECG recordings and introduce two feature extraction techniques: the Higher-order statistics and the Discrete wavelet transformation. The *Earth mover's distance* (EMD) is also discussed for measuring the dissimilarity of two multi-dimensional distributions.

4.3.1 Literature Review

Many machine-learning approaches have been proposed for automatic heartbeat classification for the last two decades. The differences between these approaches are mainly the features and the classifiers.

The features used to represent a heartbeat are usually extracted from cardiac rhythm or time/frequency domains, in which the RR-Interval is reported as one of the most widely used feature [4, 23, 38, 119, 174, 209]. RR-Interval holds indispensable information about heart rhythms and has capacity to discriminate the disease heartbeats from the normal ones. Other features, such as the *higher order statistics* (HOS) [4, 40], *wavelet coefficients* [10, 60, 101, 112, 198, 204], *morphological amplitudes* [86,209], *signal energy* [204], and *random projection features* [13,74], can also be commonly found in the literature. As irrelevant features could cause negative impacts to the classification performance and decrease the generalization power, different feature selection techniques have been applied to clear up the noise and reduce the feature dimension, such as the *floating sequential search* [115] and the *weighted linear discriminant model with a forward-backward search strategy* [46].

Regarding the classifiers, the *support vector machine* (SVM) [23, 39, 74, 147, 154], *nearest neighbors* (NN) [104, 173], *artificial neural networks* (ANN) [60, 124], *optimum-path forest* (OPF) [37], *linear discriminants*(LD) [38], *conditional random field* [40], and *reservoir computing with logistic regression* [53] are common choices for the heartbeat classification problem. However, using a single classifier can bias the classification and lead to a relatively low generalization performance. Some ensemble methods, such as random forest [4] and ensemble of SVM [74], have been employed to remedy the disadvantages.

Although some promising results have been achieved, the current methods on heartbeat classification still have some problems. The internal connections among heartbeats are often ignored in existing classification process. All types of heartbeats are presented using a same set of static features. This could limit the classification performance and possibly lead to a failure in identification of S-type heartbeats.

4.3.2 Higher-order Statistics

The *higher-order statistics* (HOS) methods are commonly used to estimate signal shape. They contain both amplitude and phase information of non-Gaussian linear processes and high immunity to the Gaussian background noise in comparison to the lower-order statistics [27]. In this work, we count the *skewness* (3rd order statistics) and the *kurtosis* (4th order statistics) into our feature set.

The *skewness* measures the symmetry of a distribution. The *kurtosis* denotes whether the distribution is heavy-tailed or light-tailed, as compared to the normal distribution. For an input signal, assume $X_{1\ldots,N}$ denotes all the data samples, $\bar{X}$ is the mean and s is the standard deviation, the *skewness* and *kurtosis* can be defined respectively as below.

$$Skewness = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^3/N}{s^3} \tag{4.1}$$

$$Kurtosis = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^4/N}{s^4} \tag{4.2}$$

4.3.3 Discrete Wavelet Transform

The *discrete wavelet transform* (DWT) provides a time-frequency representation of a signal, which is widely used in data compression, noise reduction and multi-frequency-bands signal analysis. The DWT iteratively decomposes a signal to different frequency bands with a scaling function and a wavelet function. The high-frequency component provides the detail information; while the low-frequency components is a coarse approximation of the upper-level signal. Each component is represented by a collection of wavelet coefficients, which is obtained by the inner products of mother wavelet function and the upper-level signal. Fig 4.1 presents the whole decomposition process. Only the low-frequency components are decomposed.

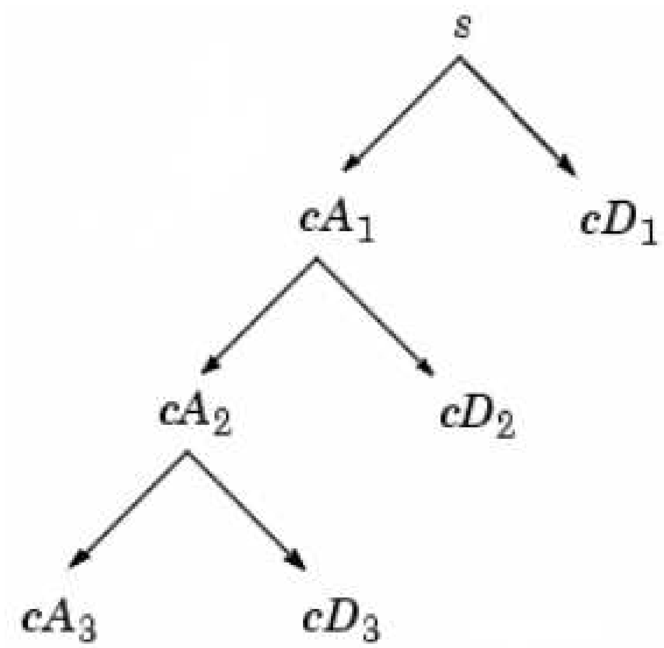


Figure 4.1: A demonstration of discrete wavelet decomposition. cA_x and cD_x denote the wavelet coefficients of coarse approximation and detail information at x level, respectively.

The choice of the mother wavelet function is the key of the *discrete wavelet transform*, which heavily depends on applications. In term of noise reduction on raw ECG signals, we use the Daubechies-4 wavelet for its good orthogonality and short vanishing moment. For morphology features extraction, the Haar wavelet is chosen because of its simplicity. Besides, it has been demonstrated as the ideal wavelet for short time signal analysis [204]. The Haar function can be represented as

$$\psi(t) = \begin{cases} 1 & 0 \leq t{<}1/2, \\ -1 & 1/2 \leq t{<}1, \\ 0 & otherwise. \end{cases}, \tag{4.3}$$

and its corresponding scaling function is

$$\phi(t) = \begin{cases} 1 & 0 \leq t{<}1, \\ 0 & otherwise. \end{cases}, \tag{4.4}$$

where t denotes sample values.

4.3.4 Earth Mover's Distance

The *Earth mover's distance* (EMD) is a metric of dissimilarity between two multi-dimensional distributions [153]. A distribution can be represented by a set of clusters. Such a representation is called the *signature* of the distribution. Data points from a distribution are grouped into a set of clusters, with each cluster denoted by its mean (or mode) and the fraction of the distribution that belongs to the cluster. Thus, one cluster can be regarded as a single feature in a signature. The distance between the features is called the *ground distance*. Signatures could be different in length. For example, simple distributions have shorter signatures than the complex ones.

The *Earth mover's distance* can be formulated and solved as a *transportation problem* [71]. Assume that there is a signature P with m cluster:

$$P = \{(p_1, w_{p1}), ..., (p_m, w_{pm})\}, \tag{4.5}$$

and a signature Q with n cluster:

$$Q = \{(q_1, w_{q1}), ..., (q_n, w_{qn})\}, \tag{4.6}$$

where p and q are the cluster representatives (mean or mode), and w denotes the cluster weight.

Let $D = \left[d_[i, j]\right]$ be the ground distance between p_i and q_j and $F = [f_{i,j}]$ be the flow between p_i and q_j. The optimal F is obtained by minimizing the overall work:

$$W = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{i,j} d_{i,j}, \tag{4.7}$$

subject to the following constrains:

$$0 \leq f_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n, \tag{4.8}$$

$$\sum_{j=1}^{n} f_{i,j} \leq w_{pi}, 1 \leq i \leq m, \tag{4.9}$$

$$\sum_{i=1}^{m} f_{i,j} \leq w_{qj}, 1 \leq j\ leqn, \tag{4.10}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{i,j} = min \left\{ \sum_{i=1}^{m} w_{pi}, \sum_{j=1}^{n} w_{qj} \right\} \tag{4.11}$$

The *Earth mover's distance* is defined as the work normalized by the total flow:

$$EMD(P, Q) = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{i,j} d_{i,j}}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{i,j}} \tag{4.12}$$

4.4 Methodology

This section presents the proposed methodology. Firstly, we introduce the preprocessing method. Then we discuss the appropriate features for heartbeat classification. After that, we present the pyramid-like model in detail.

4.4.1 Preprocessing

The raw ECG signals always come with Gaussian white noise and baseline wanders. The baseline wanders is the effect that the base axis (X-axis) of individual heartbeats appear to move up or down rather than being straight all the time, as shown in Fig 4.2. In order to avoid propagation of the negative impact of these two problems to the classification stage, an effective method for cleaning up the ECG recordings is indispensable.

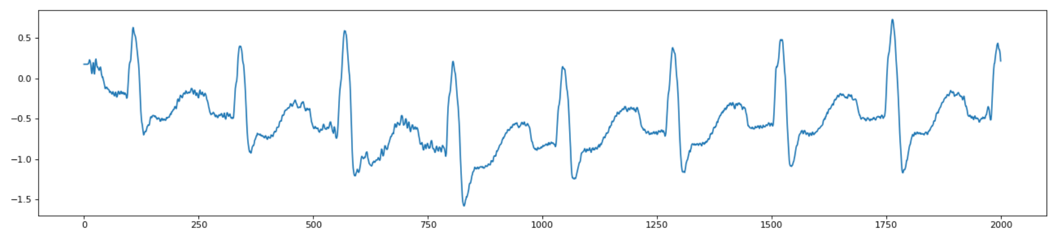


Figure 4.2: A sample ECG recording with Gaussian white noise and baseline wanders.

To correct the baseline wanders, each ECG recording is processed with a 200-ms width median filter followed by a 600-ms median filter to obtain the signal baseline, which is then subtracted from the raw ECG signal to get the baseline corrected data. Then, a discrete wavelet transform is applied to remove the Gaussian white noise. The baseline corrected recordings are decomposed to different frequency bands with various resolutions. We select the *Daubechies-4* as the mother wavelet function because its short vanishing moment is ideal for analyzing signals like ECG with sudden changes. The coefficients of detail information (cD_x) in each frequency band are then processed by a high-pass filter with a threshold value

$$T = \sqrt{2 * log(n)}, \tag{4.13}$$

where n indicates the length of the input signal. The coefficients that failed by the filter are set to zero. Finally, the clean recordings are obtained by employing inverse discrete wavelet transform on the coefficients.

After noise reduction, The ECG recordings are segmented to individual heartbeats using the R locations provided by the databases. For each R peak, 90 samples (250-ms) before R peak and 144 samples (400-ms) after R peak are taken to represent a heartbeat. This is long enough to catch the samples representing the re-polarization of ventricular and short enough to exclude the neighbor heartbeats [4].

4.4.2 Feature Extraction

Three types of features are used to characterize a heartbeat in this work: RR-interval, HOS and *wavelet coefficients*. Table 4.1 and Table 4.2 summarize the statistics of these features and give their p-values among normal, S-type and V-type beats. Fig 4.3 gives a visual demonstration on the feature significance via boxplots.

The RR-interval is the time distance between two successive R peaks. Specifically, the interval between the current R peak and the previous R peak is known as *PreRR*, while the interval between current R peak and the following R peak is *PostRR*. The RR-interval is one of the most indispensable features used for heartbeat classification. Zhancheng et al. [209] have done extensive work to prove that *PreRR* is the top distinguishing feature for recognizing S beats. Table 4.1 shows the p-value of *PreRR* between normal and S-type heartbeats is 2.16e$-$58, which means that *PreRR* leads to a significant difference between the normal and S-type beats.

The *skewness* (3rd order statistics) and the *kurtosis* (4th order statistics) are effective in estimating shape parameters of ECG signals. They are able to well distinguish V beats because the major difference of V beats against other types of heartbeats is the shape. The corresponding p-values in Table 4.1 justify this statement.

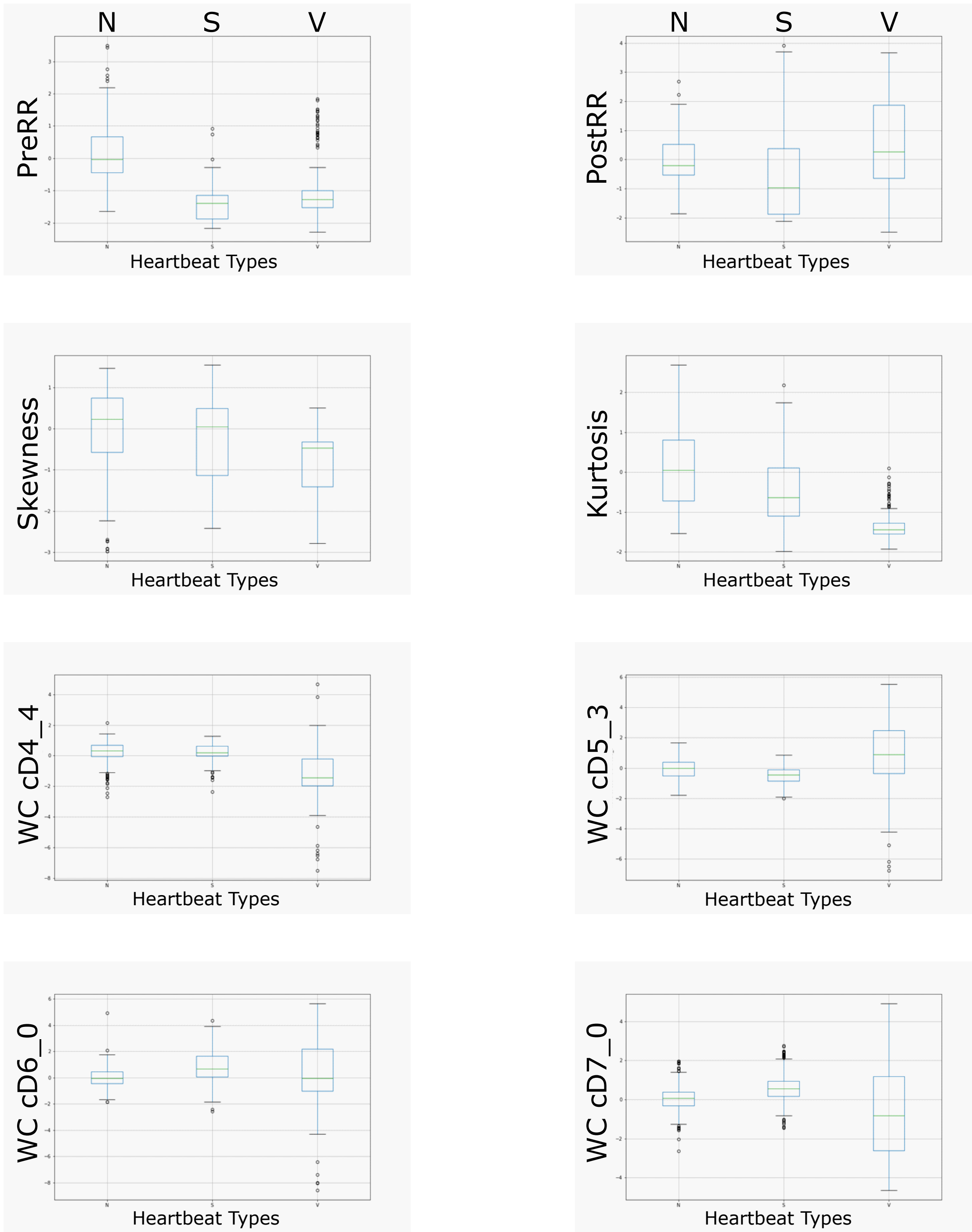


Figure 4.3: Boxplots for the extracted features of ECG signals. Notes: for each frequency component, we only pick one coefficient as example.

Table 4.1: Feature statistics and the corresponding p-values between heartbeat classes

Feature	Statistics (mean ± std)			P-values		
	N	S	V	N - S	N - V	S - V
preRR	[-0.81, 1.17]	[-1.98, -0.79]	[-1.86, -0.33]	2.16e−58	2.31e−38	4.27e−05
postRR	[-0.88, 0.89]	[-2.02, 0.79]	[-1.17, 1.99]	1.63e−07	1.94e−03	2.69e−11
skewness	[-0.99, 1.01]	[-1.36, 0.58]	[-1.63, -0.13]	8.48e−05	2.71e−21	2.33e−08
kurtosis	[-0.91, 1.09]	[-1.29, 0.35]	[-1.63, -0.91]	1.61e−09	5.42e−54	2.05e−30
cD4.0	[-0.82, 0.98]	[-0.95, 1.98]	[-2.38, 1.88]	3.47e−04	4.53e−02	3.33e−05
cD4.1	[-0.98, 0.7]	[-0.67, 1.57]	[-1.0, 2.29]	4.24e−09	3.15e−09	1.65e−01
cD4.2	[-0.98, 1.01]	[-0.6, 1.44]	[-1.2, 1.24]	7.40e−05	9.42e−01	4.77e−04
cD4.3	[-0.77, 0.84]	[-0.86, 0.35]	[-2.52, 0.83]	6.81e−05	7.24e−11	3.35e−06
cD4.4	[-0.54, 0.96]	[-0.32, 0.74]	[-2.71, 0.62]	9.42e−01	4.76e−20	1.08e−21
cD4.5	[-1.0, 0.96]	[-0.76, 1.23]	[-0.55, 1.83]	9.84e−03	3.52e−09	2.80e−04
cD4.6	[-0.97, 1.15]	[-1.37, 0.35]	[-1.25, 1.32]	1.35e−09	6.50e−01	8.73e−07
cD4.7	[-1.03, 0.74]	[-1.57, 0.84]	[-1.18, 3.22]	3.94e−02	1.76e−11	6.51e−14
cD4.8	[-0.79, 0.86]	[-1.09, 0.75]	[-1.97, 2.47]	1.86e−02	2.05e−01	1.40e−02
cD4.9	[-0.96, 0.88]	[-1.06, 0.81]	[-1.99, 2.04]	3.69e−01	6.57e−01	3.30e−01
cD4.10	[-0.82, 0.87]	[-0.39, 1.04]	[-2.17, 2.06]	1.46e−04	6.14e−01	1.63e−02

The *wavelet coefficients* provide multi-frequency-bands information of signals. Since each heartbeat only contains 235 data samples, the maximum level of wavelet decomposition is up to 7. As reported by Asl et al. [10], each type of heartbeats can find its own representative and distinct components in the detail information at

Table 4.2: Feature statistics and the corresponding p-values between heartbeat classes

Feature	Statistics (mean $\pm$ std)			P-values		
	N	S	V	N - S	N - V	S - V
cD4.11	[-0.78, 0.89]	[-0.48, 1.23]	[-2.19, 1.47]	1.81e−04	3.85e−03	4.43e−07
cD4.12	[-0.74, 0.73]	[-0.44, 0.93]	[-2.39, 2.49]	5.38e−04	7.52e−01	2.85e−01
cD4.13	[-0.52, 0.49]	[-3.02, 1.43]	[-2.06, 1.36]	1.97e−06	7.83e−03	2.64e−02
cD4.14	[-0.51, 0.51]	[-3.52, 4.09]	[-1.17, 0.95]	2.96e−01	1.87e−01	1.60e−01
cD5.0	[-0.7, 0.73]	[-0.74, 1.76]	[-2.21, 2.32]	1.38e−06	8.10e−01	1.25e−02
cD5.1	[-0.91, 0.93]	[-0.71, 0.96]	[-2.25, 0.84]	1.93e−01	4.01e−08	9.97e−11
cD5.2	[-1.0, 0.96]	[-0.42, 1.38]	[-1.37, 2.27]	2.61e−07	1.58e−03	8.47e−01
cD5.3	[-0.83, 0.56]	[-1.12, 0.1]	[-1.3, 3.72]	1.93e−08	1.54e−12	3.84e−19
cD5.4	[-0.78, 0.81]	[-1.16, 0.63]	[-2.02, 2.37]	1.01e−03	3.26e−01	8.50e−03
cD5.5	[-0.74, 0.85]	[-0.37, 1.29]	[-2.82, 2.53]	7.77e−07	3.18e−01	2.39e−03
cD5.6	[-1.03, 0.98]	[-1.23, 0.94]	[-2.46, 2.23]	2.64e−01	6.34e−01	8.66e−01
cD6.0	[-0.7, 0.56]	[-0.45, 1.95]	[-2.33, 2.91]	2.59e−16	5.87e−02	2.53e−02
cD6.1	[-1.0, 0.86]	[-1.36, 0.75]	[-2.11, 1.5]	1.92e−02	1.01e−01	9.86e−01
cD6.2	[-0.84, 0.83]	[-0.5, 0.91]	[-1.8, 2.12]	6.28e−03	2.59e−01	7.75e−01
cD6.3	[-0.75, 0.73]	[-2.65, 1.1]	[-2.01, 1.77]	1.23e−07	4.17e−01	6.06e−04
cD7.0	[-0.73, 0.85]	[-0.22, 1.43]	[-2.88, 1.84]	6.75e−11	9.81e−04	4.91e−10
cD7.1	[-0.85, 0.88]	[-0.94, 1.11]	[-2.46, 2.09]	4.67e−01	2.36e−01	1.22e−01

level 4-7. In this study, the detail information at these levels are used to represent morphology-related features of an ECG signal. P-values of the coefficients are presented in Table 4.2.

In conclusion, each of the above-mentioned features has been proven to be able to distinguish at least one certain type of heartbeats from the others. However, as discussed in Sec.4.2, using all these features for all heartbeat types classification can lead to a poor classification performance. Therefore, a pyramid-like model is proposed to select and organize these features to improve performance.

4.4.3 Pyramid-like Classification Model

The proposed pyramid-like model is made up of the *nsDispatcher*, *nRefiner* and *sRefiner*. Fig 4.4 present the entire framework. The classification process has two stages, known as *level-1* and *level-2* classification. In *level-1* classification, the raw heartbeat data is processed by the *nsDispatcher* at first, where each heartbeat is categorized into the N or S group. After that, in *level-2* classification, the *nRefiner* classifies the heartbeats in the upper N group to the N, V, F or Q group. Simultaneously, the *sRefiner* classifies the heartbeats in the upper S group to the S, V, F or Q group.

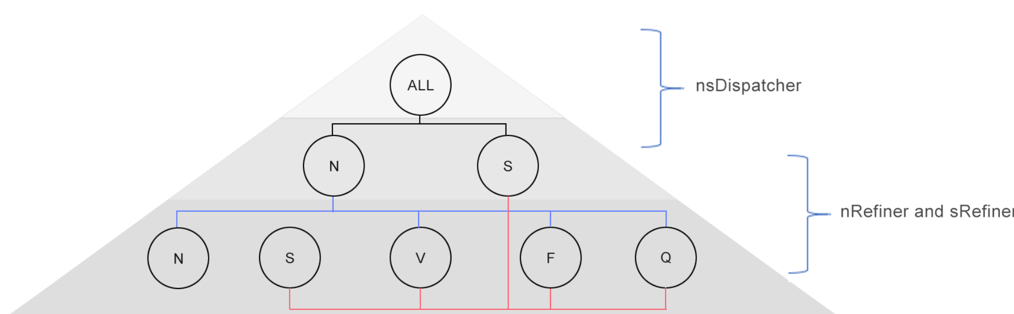


Figure 4.4: Overall structure of the proposed pyramid-like model.

When the shape-related features are included in consideration, N-type and S-type heartbeats are difficult to distinguish, because these two heartbeat types share a similar QRS complex. Therefore, we focus on classification of N and S

beats specially. In *nsDispatcher*, only the heart rhythm information (RR-interval) is considered.

Model Training

Algorithm 3 presents the training process of *nsDispatcher*. The input training data is denoted as $DS_{training}$, where each ECG recording represents a patient.

Algorithm 3 nsDispatcher Training

Require: A training ECG recordings database, $DS_{training}$

Ensure: Threshold values for each patient, $trsValues$

$step \leftarrow 0.05$
for patient in $DS_{training}$ **do**
  $heartbeats \leftarrow$ Nomalize($patient.heartbeats$)
  $pid \leftarrow patient.id$
  for hb in heartbeats **do**
    if hb.label $\in$ N **then**
      $labelTrue[pid]$.append($hb.label$)
      $normalBeats[pid]$.append($hb.preRR$)
    else if hb.label $\in$ S **then**
      $labelTrue[pid]$.append($hb.label$)
    else
      continue
    end if
  end for
  $normalPreRR \leftarrow$ median($normalBeats[pid]$)
  $t \leftarrow 0$
  while t > -1 **do**

18: **for** hb in heartbeats **do**
19: **if** $hb.label \notin$ (N $\cup$ S) **then**
20: continue
21: **else if** (hb.preRR - hb.postRR) / normalPreRR < t **then**
22: $labelPred[pid]$.append('S')
23: **else if** (hb.preRR - normalPreRR / normalPreRR < t **then**
24: $labelPred[pid]$.append('S')
25: **else**
26: $labelPred[pid]$.append('N')
27: **end if**
28: **end for**
29: $N\_Sen[t] \leftarrow$ getSensitivity($'N'$, $labelTrue[pid]$, $labelPred[pid]$)
30: $S\_Sen[t] \leftarrow$ getSensitivity($'S'$, $labelTrue[pid]$, $labelPred[pid]$)
31: $t \leftarrow t - step$
32: $labelPred[pid] \leftarrow$ NULL
33: **end while**
34: $trsValues[pid] \leftarrow \arg\max_t (N\_Sen[t] + S\_Sen[t])$
35: **end for**

The core of the *nsDispatcher* is the decision rules shown between line 18 - 28 in Algorithm 3. They determine which group (N or S) a heartbeat belongs to. Let *hb* denote a heartbeat and t be the threshold value, the decision rules can then be mathematically expressed as

rule 1:

$$\frac{hb.preRR - hb.postRR}{normalRreRR} < t, \tag{4.14}$$

and **rule 2**:

$$\frac{hb.preRR - normalRreRR}{normalRreRR} < t, \tag{4.15}$$

where $normalRreRR$ represents the median value of the $PreRR$ values of the normal heartbeats.

The rules are motivated by two observations: (1) a S-type beat generally has a shorter $PreRR$ value than that of a surrounding N-type beat; and (2) the gap of the $PreRR$ value between a S-type beat and a N-type beat varies with patients. Therefore, a heartbeat should not be treated as an independent data sample, but be associated with the surrounding beats as well as the patient-specific information. The **rule 1** uses the surrounding beats to help classification. Suppose that in an ECG recording, there is a S-type beat followed by a N-type beat. The S-type beat can be easily caught by the **rule 1**. However, when there are two successive S-type or N-type beats, the **rule 1** can fail because there is not enough information. As such, the **rule 2** is applied to complement the **rule 1** by taking advantage of the patient-specific information ($normalPreRR$).

If any of the rules is satisfied, the heartbeat is categorized as class S, otherwise as class N. The goal of the training process is to find out the best threshold value (t) that helps to achieve a high detection sensitivity of both the N-type and S-type heartbeats for the decision rules of each patient. We traverse every possible t in the range of $(-1, 0)$. Values beyond this range is practically impossible so far. The parameter $step$ is used to control the precision of t. The smaller the $step$, the more precise the t but the more time-consuming the training process. Formally, the objective function (line 34 in Algorithm 3) is formulated as:

$$\arg\max_t (N\_Sen[t] + S\_Sen[t]). \tag{4.16}$$

The trained threshold values are stored in $trsValues$ (line 34 in Algorithm 3).

Table 4.3: The nRefiner and the sRefiner

	Classifier	Features
nRefiner	Mix Ensemble(Linear SVM, SVM, Decision Tree, KNN, Logistic Regression, Perceptron, and Bayes)	heartbeat rhythm, HOS, and wavelet coefficients
sRefiner	SVM	HOS and wavelet coefficients

In terms of the *nRefiner* and the *sRefiner*, Table 4.3 summarizes their compositions and the training features. Notice that the N group is seriously imbalanced and dominated by the normal heartbeats. To reduce the impact caused by the imbalance problem, a mix classifier ensemble method is applied in the *nRefiner*. The reason for excluding the heartbeat rhythm for training the *sRefiner* is that the V-type heartbeats could also have irregular RR-interval values as the S-type heartbeats.

Classification

The details of *level-1* and *level-2* classification are presented in Algorithm 4 and Algorithm 6, respectively.

In *level-1* classification, one important step is the estimation of the normal $PreRR$ value of a patient (line 4 - 11 in Algorithm 4). For each patient p_a in DS_{test}, we perform a statistical analysis on p_a's heartbeat $PreRR$ values via *boxploting*. If less than 10% of the data are considered as outliers, we assume that the ECG recording is dominated by the normal heartbeats and use

$$E(normalRreRR) \leftarrow median(heartbeats.preRRs) \tag{4.17}$$

to estimate the normal $PreRR$ value. Such an assumption is practical and reasonable because S-type heartbeats occur sparsely in real-world applications.

On the other hand, if more than 10% of the data are considered as outliers, the ECG recording is likely to be distorted by the S-type heartbeats and $median(heartbeats.preRRs)$ could represent the $PreRR$ value of a S-type beat. In such a case, we use

$$E(normalPreRR) \leftarrow mean(mean(outliers), median(heartbeats.preRRs)) \tag{4.18}$$

to estimate the normal $PreRR$ value. This guarantees that the $E(normalPreRR)$ is not representing an irregular value.

Algorithm 4 Level-1 Classification

Require: A test ECG recordigns database, DS_{test}; The trained threshold values, $trsValues$.

Ensure: The result of level-1 classification, $lev1Result$.

```
for patient in DS_test do
    pid ← patient.id
    heartbeats ← Nomalize(patient.heartbeats)
    stats ← boxplot(heartbeats.preRRs)
    outliers ← stats.outliers
    if len(outliers) / len(heartbeats) >0.1 then
        E(normalPreRR) ← mean(mean(outliers), median(heartbeats.preRRs))
    else
        E(normalPreRR) ← median(heartbeats.preRRs)
    end if
    neighbor ← getNeighbor(patient)
    t ← trsValues[neighbor]
    if t equals to 0 then
        t ← min(trsValues)
```

```
15:     end if
16:     for hb in heartbeats do
17:         if (hb.preRR - hb.postRR) / E(normalPreRR) < t then
18:             lev1Result[pid].append('S')
19:         else if (heartbeat.preRR - E(normalPreRR)) / E(normalPreRR) < t
    then
20:             lev1Result[pid].append('S')
21:         else
22:             lev1Result[pid].append('N')
23:         end if
24:     end for
25: end for
```

The algorithm goes on by looking for a patient p_b in $DS_{training}$ who has the most similar *PreRR* values distribution with p_a, and assign p_b's threshold value to p_a (line 12 - 13 in Algorithm 4). We implement a function named $getNeighbor$ (Algorithm 5) to perform the task. The function uses the *Earth mover's distance* (EMD) to measure the dissimilarity of two distributions. Notice that if p_b's threshold value equals to 0, which means that no S-type beat is found in p_b, it is believed that there is also a low probability to find S beats in p_a. However, we never want to miss a potential S-type beat, which may lead to a serious consequence to a patient. In such a case, we assign the smallest value in $trsValues$ to p_a (line 14 - 16 in Algorithm 4). This implies that the algorithm tries to search for the potential S-type beats while avoiding classifying the N-type as S-type beats.

Once the $E(normalPreRR)$ as well as the t are ready, the heartbeats are processed by the decision rules.

Algorithm 5 Find the nearest neighbor of a patient

Require: An ECG recording of a patient, $testPatient$; The training ECG recordings database, $DS_{training}$.

Ensure: A patient in $DS_{training}$ who has the most similar *previour-RR* values distribution of the $testPatient$, *neighbor*.

function GETNEIGHBOR(testPatient)
    $data1 \leftarrow$ Normalize($testPatient.heartbeats.preRRs$)
    for trainPatient in $DS_{training}$ **do**
        $pid \leftarrow trainPatient.id$
        $data2[pid] \leftarrow$ Normalize($trainPatient.heartbeats.preRRs$)
    end for
    $neighbor \leftarrow \arg\max_{pid}(\text{EMD}(data1, data2[pid]))$
    return *neighbor*
end function

In *level-2* classification (Algorithm 6), each heartbeat in the *N* group is further classified by the *nRefiner* to class *N*, *V*, *F* or *Q*. Similarly, the *sRefiner* reclassified the *S* beats to class *S*, *V*, *F* or *Q*.

Algorithm 6 Level-2 Classification

Require: The test ECG recordings database, DS_{test}; The level-1 classification result, $lev1Result$.

Ensure: The final result of the pyramid-like model, $finalResult$.

for patient in DS_{test} **do**
    $pid \leftarrow patient.id$
    $heartbeats \leftarrow$ Nomalize($patient.heartbeats$)
    for hb in heartbeats **do**

```
    if lev1Result[pid][hb.id] ∈ N then
        finalResult[pid].append(nRefiner(hb))
    else if lev1Result[pid][hb.id] ∈ S then
        finalResult[pid].append(sRefiner(hb))
    else
        continue
    end if
  end for
end for
```

4.5 Experimental ECG Databases

In this section, two ECG databases are introduced, namely the *MIT-BIH-AR* database and the *INCART* database. They are public-accessible from the *Physiobank* [57].

Most of the works on heartbeat classification are trained and evaluated their models on the *MIT-BIH-AR* database. In order to have a fair comparison, both the training and the evaluation of the pyramid-model are performed on the *MIT-BIH-AR* database as well. Besides, we use the *INCART* database to assess the generalization performance of the proposed model.

All ECG recordings in these databases have an equal length of 30 minutes, but they are not sampled in the same frequency. They need to be re-sampled to $360Hz$ before use. The recordings are well-labeled at heartbeat level. The original heartbeat annotations include 15 classes, which are further grouped into 5 super-classes by the $AAMI$ [8], as shown in Table 4.4.

Details of these databases are respectively given below.

Table 4.4: ECG-based heartbeat annotations

AAMI class	Original class	Type of beat
Normal (N)	N	Normal beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
	e	Atrial escape beat
	j	Nodal (junctional) escape beat
Supraventricular ectopic beat (S)	A	Atrial premature beat
	a	Aberrated atrial premature beat
	J	Nodal (junctional) premature beat
	S	Supraventricular premature beat
Ventricular ectopic beat (V)	V	premature ventricular contraction
	E	Ventricular escape beat
Fusion beat (F)	F	Fusion of ventricular and normal beat
Unknown beat (Q)	$/$	Paced beat
	f	Fusion of paced and normal beat
	Q	Unclassifiable beat

4.5.1 MIT-BIH-AR Database

The database contains 48 two-lead ambulatory ECG recordings from 47 patients (including 22 females and 25 males). Each recording is denoted by a 3-digits number. The recordings were digitized at $360Hz$ per second per channel with 11-bit resolution over a $10 - mV$ range. For most of them, the first lead is modified limb lead II (except for the recording 114). The second lead is a pericardial lead (usually *V1*, sometimes are *V2*, *V4* or *V5*, depending on subjects). In this study, only the modified limb lead II is used.

The database is seriously imbalanced. The N beats dominate most of the recordings. Therefore, the k-fold validation scheme cannot be applied to split the database for training and testing. Two different paradigms are found in the literature to solve this problem [4,38,200,209]. One is the intra-patient paradigm, which first mixes up the heartbeats from all recordings and then evenly allocates each category of heartbeats into two groups. The other one is the inter-patient paradigm. In this paradigm, the ECG recordings are divided into two datasets ($DS1$ and $DS2$) with each dataset containing approximately the same portion of heartbeat classes. Table 4.5 shows the division and the corresponding heartbeat classes distribution. The $DS1$ is used for model training and the $DS2$ is used for model performance evaluation.

It has been empirically proven that the intra-patient paradigm can bias the classification result by allowing training and testing heartbeats coming from the same patient [119]. By contrast, the inter-patient paradigm is more objective. In order to reveal the true performance of the pyramid-like model and have a fair comparison with the state-of-the-art rivals, the inter-patient paradigm is strictly followed in this work.

4.5.2 INCART 12-leads Arrhythmia Database

This database consists of 75 ECG recordings sampled at $257Hz$. Each recording contains 12 standard leads. Similarly, only the modified limb lead II is used in this study. The annotations were first produced by an automatic algorithm and then corrected manually based on the standard PhysioBank beat annotation definitions. None of the recordings contains pacemakers, but most of them have ventricular ectopic beats. The heartbeat distribution of the $INCART$ database is shown in

Table 4.5: The inter-patient division paradigm.

Data set	N	S	V	F	Q	Recordings[1,2]
DS1	45808	943	3786	414	8	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	44198	1836	3219	388	7	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

[1] Each recording is denoted by a 3-digits number and the numbers are originally discontinuous.
[2] As recommended by the AAMI, the four recordings (102, 104, 107 and 217) containing paced beats are excluded from the analysis.

Table 4.6: Heartbeat distributions in the INCART database

Database	N	S	V	F	Q
INCART	153491	1958	19993	219	6

Table 4.6.

4.6 Experimental Evaluation

In this section, we conduct a benchmark evaluation for the proposed pyramid-like model on the *MIT-BIH-AR* database, with the result being compared to the state-of-the-art methods. Besides, we use the *INCART* database to assess the model generalization performance.

All the experiments presented in this work are programmed in Python 3.63 and done in a 64-bits Windows 10 PC, with $i5 - 4590$ CPU and 12 GB memory.

4.6.1 Evaluation Metrics

In this work, the performance is evaluated by sensitivity (Se), positive predictive value ($+P$) and accuracy value (Acc). It should be noted that penalties would not be applied for the misclassification of class F and Q, as recommended by the $AAMI$ standard.

4.6.2 Classification Result and Discussion

Table 4.7 shows the result of the *level-1* classification. The majority of the N-type and S-type beats are correctly classified by the *nsDispatcher*. Although 3153 N-type beats are misclassified as S-type beats, they only account for a small portion of the total N-type beats. A good classification sensitivity and positive predictive value of the N-type beat are still achieved. On the other hand, the misclassified N-type beats lead to a decrease of the positive predictive value of the S-type beats. However, as the heartbeat classification plays an important role toward identifying the cardiac arrhythmia, the accuracy over the S-type heartbeats is considered the most important [37]. From an overall point of view, the *nsDispatcher* does a decent job.

Table 4.8 gives the final classification results of the proposed pyramid-like model in detail. It is worth noting that, from *level-1* to *level-2* classification, only 164 N-type beats and 87 S-type beats are misclassified by the *nRefiner* and the *sRefiner*. In addition, the *level-2* classification achieves superior performance in the detection of V-type beats. The results indicate the effectiveness of the *nRefiner* and the *sRefiner*. In terms of the F-type and Q-type beats, a poor performance is obtained, which is a normal phenomenon because both F-type and Q-type beats are originally unclassifiable. The same issue is commonly found in all the existing research works.

Table 4.7: The result of level-1 classification of the proposed model on DS2

		Predicted class	
		N	S
True class	N	40918	3151
	S	74	1680
	V	872	2347
	F	383	5
	Q	5	2

Table 4.8: The result of level-2 classification of the proposed model on DS2

		Predicted class				
		N	S	V	F	Q
True class	N	40754	2762	508	45	0
	S	71	1593	87	3	0
	V	125	151	2856	87	0
	F	317	1	62	8	0
	Q	2	0	4	1	0

The proposed model is compared to the state-of-the-art methods over the same test set ($DS2$). Table 4.9 summarizes the comparative result. The proposed model exhibits higher performance in terms of the positive predictive value of N-type beats and the sensitivity value of the disease heartbeats (S-type and V-type). In addition, it takes the second best place in global accuracy (91.5%) and the sensitivity value of normal heartbeats (99.0%).

Although our model has the lowest positive predictive value of the S-type beats, we make a breakthrough in the sensitivity value (91.0%). Actually, as we can see, the positive predictive values of S-type heartbeats are commonly low in most of the existing methods. The best one is obtained by Ye C et al. [200], which is just 17% better than ours, but we beat it in the sensitivity value by more than 30%.

Table 4.9: Performance comparison of the proposed model and the state-of-the-art methods on DS2

Method	Acc(%)	N		S		V	
		Se(%)	+P(%)	Se(%)	+P(%)	Se(%)	+P(%)
Proposed	91.5	92.0	**99.0**	**91.0**	35.0	**89.0**	81.0
De Chazal [38]	81.9	86.9	99.2	75.9	38.5	77.7	81.9
Ye C [200]	86.4	88.5	97.5	60.8	**52.3**	81.5	63.1
Zhang Z [209]	86.7	88.9	99.0	79.1	36.0	85.5	**92.8**
Shan C [23]	**93.1**	**98.4**	95.4	29.5	38.4	70.8	85.1
Mariano L [115]	78.0	78.0	99.0	76.0	41.0	83.0	88.0

Table 4.10: Classification result of the proposed pyramid-like model in the INCART database

		Predicted class		
		N	S	V
True class	N	138620	6871	8000
	S	106	1554	298
	V	792	1643	17783

Table 4.11: Generalization performance comparison between the proposed model and the stat-of-the-art rival in the INCART database

Method	Acc(%)	N		S		V	
		Se(%)	+P(%)	Se(%)	+P(%)	Se(%)	+P(%)
Proposed	90.0	90.3	**99.3**	79.4	**15.4**	**87.0**	72.7
Mariano L [115]	**91.0**	**92.0**	99.0	**85.0**	11.0	82.0	**88.0**

4.6.3 Generalization Result and Discussion

The classification result on the *INCART* database is summarized in Table 4.10. The performance is compared to the latest work by Mariano L. and Juan P. [115], which is the only work can be found performing model evaluation on both the *MIT-BIH-AR* and the *INCART* database. Table 4.11 presents the comparative result.

Notice that the compared method [115] follows the $AAMI2$ labeling, where class F and Q are merged into class V. In order to have a fair comparison, we adapt the proposed model to the $AAMI2$ labeling.

As seen from Table 4.11, the proposed model has a comparable performance with the rival on the *INCART* database. Both the works achieve similar values in all metrics. However, if we look back at Table 4.9, the proposed pyramid-like model presents better performance on $DS2$.

It is worth noting that, from $DS2$ to the *INCART* database, the proposed model maintains a stable heartbeat classification performance. This is very important, as robustness is indispensable for an algorithm to be applied in a clinical practice.

4.7 Chapter Conclusion and Future Work

Millions of people around the world are suffering from the cardiac arrhythmia. Automatic heartbeat classification helps early identify this issue, making it possible for people to get the right treatment sooner. In this paper, a pyramid-like model has been proposed for automatic heartbeat classification. The model integrates three components, namely *nsDispatcher*, *nRefiner* and *sRefiner*. During the classification process, the *nsDispatcher* first allocates the heartbeats into the N or S group. The *nRefiner* and the *sRefiner* then further classify the heartbeats in the N and S group respectively to give the final decision. The significance of the proposed model is that it takes the surrounding heartbeats as well as the patient-specific information into consideration to help identification of a S-type heartbeat. Besides, the *nRefiner* and the *sRefiner* are customized with different classifier structure and training features to adapt to the classification requirements in the N and S group.

The proposed model has been evaluated on the *MIT-BIH-AR* database, with the performance being compared against the state-of-the-art methods. In addition, the *INCART* database is used to measure the generalization performance of the proposed model. The experimental results have proven the effectiveness and robustness of the proposed model in heartbeat classification.

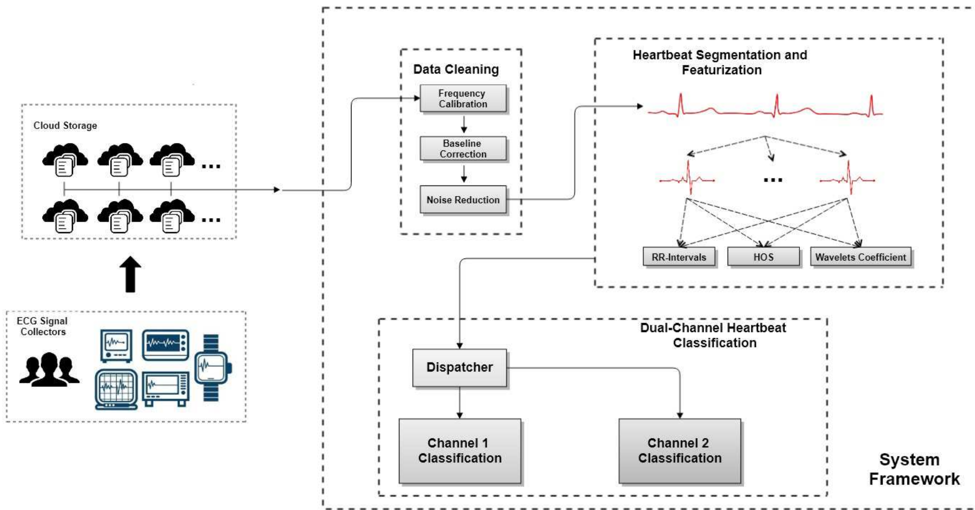


Figure 4.5: Extension of pyramid-like model for online-detection scenarios.

In our next study, we aim to extend the proposed pyramid-like model to online-detection scenarios. We outline the framework design in fig.4.5. The framework comprises four modules: a raw ECG cleaning module, a heartbeats segmentation and featurization module, a heartbeat classification module and a result notification module.

CHAPTER 5

A FRAMEWORK FOR CARDIAC ARRHYTHMIA DETECTION FROM IOT-BASED ECGS AND INSPECTION OF LATEST ADVANCES BROUGHT BY DEEP LEARNING

5.1 Chapter Abstract

Cardiac arrhythmia has been identified as a type of cardiovascular diseases (CVDs) that causes approximately 12% of all deaths globally. The development of *Internet-of-Things* has spawned novel ways for heart monitoring but also presented new challenges for manual arrhythmia detection. An automated method is highly demanded to provide support for physicians. Current attempts for automatic arrhythmia detection can roughly be divided as feature-engineering based and deep-learning based methods. Most of the feature-engineering based methods are suffering from adopting single classifier and using fixed features for classifying all five types of heartbeats. This introduces difficulties in identification of the problematic heartbeats and limits the overall classification performance. The deep-learning based methods are usually not evaluated in a realistic manner and report overoptimistic results which may hide potential limitations of the models. Moreover, the lack of consideration of frequency patterns and the heart rhythms can also limit the model performance. To fill in the gaps, we propose a framework for arrhythmia detection from IoT-based ECGs. The framework consists of two modules: a data cleaning module and a heartbeat classification module. Specifically, we propose two solutions to the heartbeat classification task, namely *Dynamic Heartbeat Classification with Adjusted Features* (DHCAF) and *Multi-channel Heartbeat Convolution Neural Network* (MCHCNN). DHCAF is a feature-engineering based approach, in which we introduce *dynamic ensemble selection* (DES) technique and develop a

result regulator to improve classification performance. MCHCNN is deep-learning based solution that performs multi-channel convolutions to capture both temporal and frequency patterns from heartbeat to assist the classification. We evaluate the proposed framework with DHCAF and with MCHCNN on the well-known MIT-BIH-AR database, respectively. The results reported in this paper have proven the effectiveness of our framework.

5.2 Introduction

Cardiac arrhythmia is a type of cardiovascular diseases (CVDs) that threatens millions of people's lives around the world. The easiest way to identify arrhythmia is to perform a manual inspection on 24 to 72 hours electrocardiograms (ECG). Traditionally, to have such long-term ECG recordings, patients need to wear a *Holter Monitor* for a continuous time period, which is a very uncomfortable experience. The rapid growth of *Internet-of-Things* (IoT) techniques has spawned novel ways, like Fitbit, Apple Watch, or Android Wear, for heart status tracking [197]. In comparison to the *Holter Moniter*, the IoT-based devices are more human-friendly because they have fewer cords and smaller-sizes, and cause fewer disruptions to patient's daily routines. However, on the other hand, the prevalence of IoT-based devices has also resulted in a dramatic increase of ECG data, posing a great challenge to the ECG interpretation. Manual inspections become time-consuming and error-prone, which is no longer possible. An automated method is highly demanded to provide a cost-effective screening for arrhythmia and allow at-risk patients to receive timely treatments.

Heartbeat classification plays a crucial role in identification of arrhythmia. Basically, heartbeats can be classified into five classes: Normal(N), Supra-ventricular (S) ectopic, Ventricular (V) ectopic, Fusion (F) and Unknown (Q) beats [8]. Par-

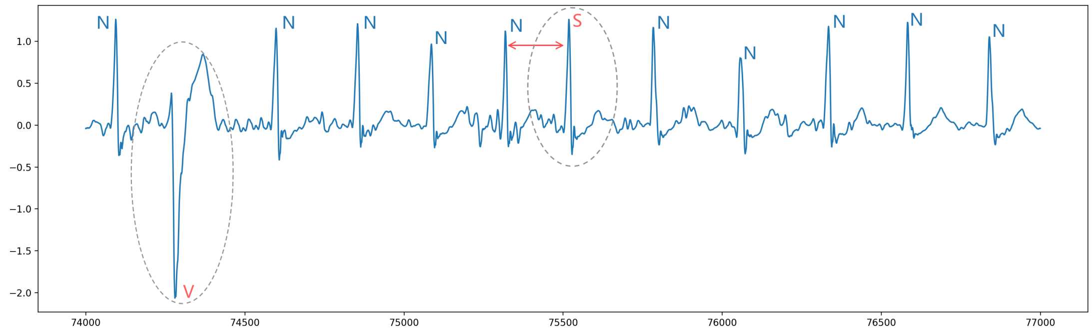


Figure 5.1: A sample ECG recording that contains N, S and V heartbeats. Note: RR-intervals denote the time distance between two successive R peaks.

ticularly, most arrhythmias are found in S and V beats. Fig.5.1 presents a sample ECG segment, where the problematic heartbeats are highlighted by circles. It can be seen that the S beat exhibits a great morphological similarity in temporal dimension to the normal heartbeats. Since ECG recordings are mostly dominated by normal heartbeats for the majority of patients [68], such similarity bring a great difficulty in distinguishing the S beats from the normal ones.

Many research attempts have been made to provide solutions to automated heartbeat classification. The existing methods are roughly divided as feature-engineering based and deep-learning based methods. However, none of these methods has achieved a clinical significance. Most feature-engineering methods are facing a bottleneck of applying a standalone classifier and using a static feature set to classify all heartbeat samples [23, 37, 38, 120, 204]. This has been shown to cause huge impacts on identification of the problematic heartbeats. The deep-learning based methods are commonly limited by learning temporal patterns from the raw ECG heartbeats only. The frequency patterns and the RR-intervals have not been well considered to assist the classification. Moreover, to supply sufficient training data for driving the deep neural networks, many works [2, 3, 94, 203, 207, 212] followed a biased evaluation procedure, in which they synthesized heartbeat samples from the whole dataset and then randomly split all heartbeats for model training, validation and test. Consequently, heartbeats from the same patient are likely to appear in both the training and test datasets, leading to an over estimation of the model performance. The overoptimistic results may hide potential limitations of the neural networks.

Besides, data quality also presents challenges for an IoT-based arrhythmia detection method. First, the IoT-based heart rate sensors may vary the rate of measurement for battery preservation [11]. Second, the collected ECG recordings are likely interrupted by background noises and baseline wonders (the effect that the base axis (X-axis) of individual heartbeats appears to move up or down rather than being straight all the time).

To solve these problems, we propose a framework for arrhythmia detection from IoT-based ECGs. The framework consists of a data cleaning module and a heartbeat classification module. Specifically, we provide two novel solutions to the heartbeat classification task. The first one is a feature-engineering based method, in which we introduce the *Dynamic Ensemble Selection* (DES) technique and specially design a result regulator to improve the problematic heartbeats detection. The other one is a deep neural network that performs multi-channel convolutions in parallel to manage both temporal and frequency patterns to assist the classification. To remedy the impact brought by the lack of consideration of heart rhythms, the proposed network accepts heart rhythms (RR-intervals) as part of the input. In order to reveal the performance of the proposed methods in real-world practices, we evaluate the models on the benchmark MIT-BIH arrhythmia database following the inter-patient evaluation paradigm proposed in [38]. The paradigm divides the benchmark database into a training and a test dataset at patient level, making the heartbeat classification a significantly more difficult task.

The rest of this chapter is structured as follows. Section 5.3 reviews current methods in heartbeat classification. Section 5.4 presents the proposed framework and the two embedded solutions for heartbeat classification. The experiment results and discussion are presented in Section 5.5. Section 5.6 concludes this chapter and discusses the future work.

5.3 Related Work

This section provides a comprehensive review of current methods for heartbeat classification. As mentioned before, the existing methods can be roughly allocated to either the feature-engineering based or the deep-learning based category. The differences between them are summarized in Table 5.1.

The feature-engineering based methods focus on signal feature extraction and classifier selection. Commonly used features includes RR-intervals [4,23,209], samples or segments of ECG curves [145], higher-order statistics [4,40], wavelet coefficients [37,60,154], and signal energy [204]. They are mostly extracted from cardiac rhythm, or time/frequency domains. Feature correlation and effectiveness are important concerns for this type of methods. To avoid negative impacts of noisy data, techniques, like the *floating sequential search* [115] and the *weighted LD model* [46], must be employed to reduce the feature space. Regarding the selection of classifiers, the *support vector machine* (SVM) is the most widely used for its robustness, good generalization and computationally efficiency [1,35]. Besides, the *nearest neighbors* (NN) and *artificial neural networks* (ANN) are also frequently found in the literature. The performances of current feature-engineering based methods are mainly limited by the application of single classifiers and the use of fixed features to classify all heartbeat types. On one hand, in consideration of the intra- and inter-subjects variations of the feature values, it is difficult for a

Table 5.1: Comparison between Feature-engineering based and Deep-learning based methods

	Feature Engineering	Deep Learning
Work flow	Feature extraction, selection and classifier determination	End-to-end processing
Commonly used features	RR-intervals, higher-order statistics, wavelet, signal energy coefficients, etc.	Learned by networks, including CNN, RNN, LSTM, etc.
Feature selection	PCA, floating sequential search, weighted LD model	N.A.
Commonly used classifiers	SVM, nearest neighbors, artificial neural networks, weighted linear discriminant, optimum-path forest	N.A.
Training data	Less	More
Parameters	Less	More
Explainability	High	Low
Current limitations	Use of fixed features for all heartbeat types classification; Limitation of static classifiers to handle both intra- and inter patients variations	Lack of considerations of frequency patterns and heart rhythms; A biased evaluation is followed.

single classifier to well handle a wide region of the feature space [210]. Although some ensemble methods, such as *random forest* [4] and *ensemble of support vector machine* [74], have been employed to remedy the disadvantages, the problem is still open because the diversity of the traditional ensembles is relatively low. On the other hand, using fixed features tends to make sporadically occurred S beats be wrongly classified as V beats because both heartbeats types exhibits anomalies in heart rhythms.

By contrast, the deep-learning based methods are more straightforward and integrated, in which features and classifiers are not concerns. They provide end-to-end solutions to the heartbeat classification task. The existing deep learning models are mainly extensions of convolution neural network (CNN) [2, 3, 94, 157] or combinations of CNN and recurrent neural network (RNN) [203, 207]. However, most of the CNN models are limited by the lack of consideration of frequency patterns and the heart rhythm to assist the classification. Moreover, in order to provide enough training data, many of them are evaluated in an ideal experimental setting where heartbeats from the same patient are allowed to appear in both training and test sets. The results can not reveal the true performances of the models in real-world practices and also may hide potential limitations of the methods. As compared to the feature-engineering based methods, both the results and the intermediate process of deep neural networks are less explainable. This is a potential impediment that prevents deep learning models from being widely applied in practices because explainability is important for clinicians to justify and rationalize the model outcome.

5.4 The Proposed Framework for Arrhythmia Detection

The proposed framework for arrhythmia detection from IoT-based ECGs is presented in this section. Fig 5.2 shows the framework architecture and the whole life-cycle of arrhythmia detection from IoT-based ECGs. The framework consists of a data cleaning module and a heartbeat classification module. It accepts raw ECG signals that are collected from different IoT devices as input and outputs predictions for individual heartbeats.

To reduce the impact of noisy data on the prediction accuracy, a series of preprocessing steps are applied to the input signals, such as frequency calibration, base-

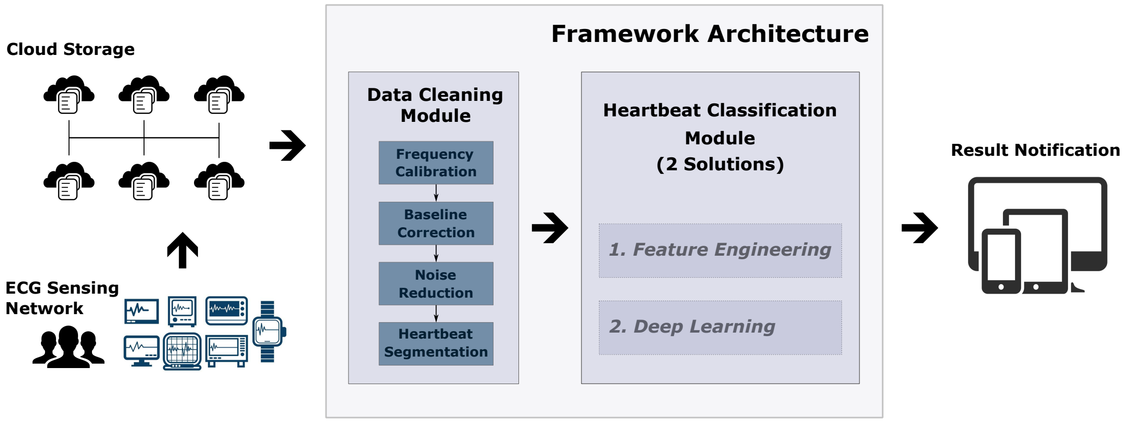


Figure 5.2: Architecture of the proposed framework. The whole life-cycle of arrhythmia detection from IoT-based ECGs includes 4 phases: data collection, storage, analysis and results notification. Specifically, ECG sensing network generates ECG recordings for patients and transmits the produced data to the IoT cloud, where fast access storage is conducted. The proposed framework is deployed in the IoT cloud to provide data analysis. Results from the framework will be pushed to patients' ends via Internet.

line correction, and noise reduction, before heartbeat classification. We propose two solutions, namely *Dynamic Heartbeat Classification with Adjusted Features* (DHCAF) and *Multi-channel Heartbeat Convolution Neural Network* (MCHCNN), for the heartbeat classification task. DHCAF is a feature-engineering based method, whereas MCHCNN is a deep-learning based method.

Details of the data cleaning module and two heartbeat classification solutions are presented below.

5.4.1 Data Cleaning Module

Frequency Calibration

To avoid the possible bias in sampling frequency caused by different ECG collectors, we develop a frequency calibration component to re-sample all incoming ECG recordings to 360 Hz at the input of the system.

Baseline Correction

To correct the baseline wanders, we process each ECG recording with a 200-ms width median filter followed by a 600-ms median filter to obtain the recording baseline, and then subtract the baseline from the raw ECG recording to get the baseline corrected data.

Noise Reduction

For noise reduction, we apply *discrete wavelet transform* [159] with *Daubechies-4* mother wavelet function to remove recordings' Gaussian white noise. The *Daubechies-4* function has short vanishing moment, which is ideal for analyzing signals like ECG with sudden changes. Concretely, in the noise reduction component, the baseline corrected recordings are decomposed to different frequency bands with various resolutions. The coefficients of detail information (cD_x) in each frequency band is then processed by a high-pass filter with a threshold value

$$T = \sqrt{2 * \log(n)},$$

where n indicates the length of the input recording. Coefficients that are blocked by the filter are set to zero. Finally, the clean recordings are obtained by employing inverse discrete wavelet transform on all the coefficients.

Heartbeat Segmentation

The clean signals are segmented to individual heartbeats by taking advantage of the R peak locations that are detected by the *Pan-Tompkins* algorithm [146]. For each R peak, 90 samples (250 ms) before R peak and 144 samples (400 ms) after R peak are taken to represent a heartbeat, which is long enough to catch samples to represent the re-polarization of ventricles and short enough to exclude the neighbor heartbeats [4].

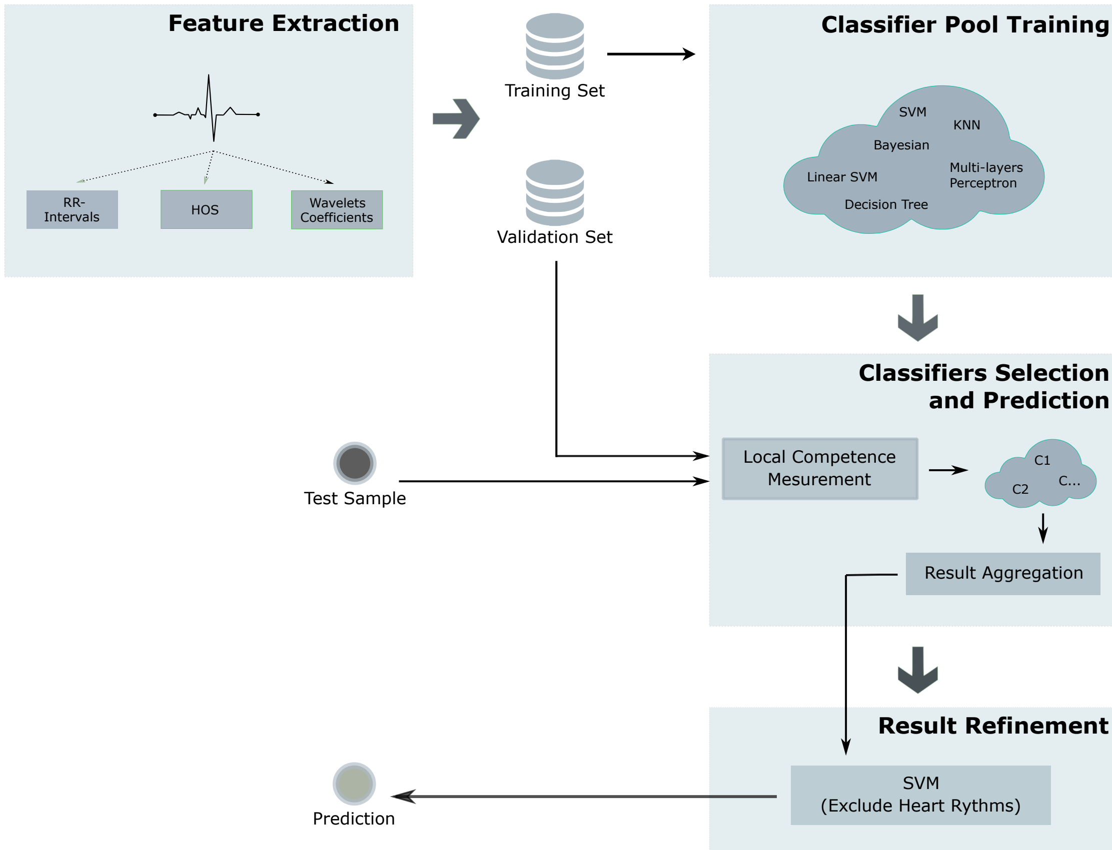


Figure 5.3: Architecture of the proposed DHCAF.

5.4.2 Dynamic Heartbeat Classification with Adjusted Features

Architecture of the proposed DHCAF is shown in Fig 5.3. The model contains 4 processing stages: Feature Extraction, Classifier Pool Training, Classifier Selection and Prediction, and Result Refinement.

Feature Extraction.

In this stage, three types of features are extracted to represent individual heartbeats: RR-intervals, higher order statistics and wavelet coefficients.

As experimentally proven in [209], the RR-interval is one of the most indispensable features for heartbeat classification and it has great capacity to tell both the S and V beats from the normal beats. In this work, four types of RR-intervals are extracted from ECG signals: pre_RR, $post\_RR$, $local\_RR$ and $global\_RR$ [119]. The RR-intervals can significantly vary with patients. To reduce the negative impact of the variation, we normalize the RR-intervals in the way below:

$$nomalized\_pre\_RR = \frac{pre\_RR}{mean(ds.pre\_RR)} \tag{5.1}$$

$$nomalized\_post\_RR = \frac{post\_RR}{mean(ds.post\_RR)} \tag{5.2}$$

$$nomalized\_local\_RR = \frac{local\_RR}{mean(ds.local\_RR)} \tag{5.3}$$

$$nomalized\_global\_RR = \frac{global\_RR}{mean(ds.global\_RR)} \tag{5.4}$$

where $ds.pre\_RR$ denotes the average of all pre_RRs in the ds that the heartbeat belongs to, and so on.

Regarding the *higher order statistics* (HOS), it is reported being useful in catching subtle changes in ECG data [125]. In this work, the *skewness* (3rd order statistics) and *kurtosis* (4th order statistics) are calculated for each heartbeat. They can be mathematically defined as follows, where $X_{1\ldots,N}$ denotes all the data samples in a signal, $\bar{X}$ is the mean and s is the standard deviation.

$$Skewness = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^3/N}{s^3} \tag{5.5}$$

$$Kurtosis = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^4/N}{s^4} - 3 \tag{5.6}$$

The wavelet coefficients provide both time and frequency domain information of a signal, which is claimed to be the best features of ECG signal [119]. The choice of the mother wavelet function used for coefficients extraction is crucial to the final classification performance. In this work, the *Haar* wavelet function is chosen because of its simplicity and that it has been demonstrated as the ideal wavelet for short time signal analysis [204].

Classifier Pool Training.

In this stage, a collection of classifiers, including *multi-layers perceptron*, *support vector machine* (SVM), *linear SVM*, *Bayesian model with Gaussian kernel*, *decision tree*, and *K-nearest neighbors model*, are trained using the extracted features, to create an accurate and diverse classifier pool.

Classifier Selection and Prediction.

This stage plays a core role in the model. The *Dynamic Ensemble Selection* (DES) [33] technique is introduced in this stage to select the most competent classifiers for making predictions of the test samples. It helps to solve both the intra- and inter-subjects variations of the feature values.

In DES, the competence of a classifier in the pool is measured by its performance over a local region of the feature space where the testing sample is located. Methods for defining a local region includes clustering [111], k-nearest neighbors [160], potential function model [194, 195] and decision space [16]. The criterion for measuring the performance of a base classifier can be divided as individual-based and group-based criterion. In the individual-based criterion, each base classifier is independently measured by evaluation metrics such as ranking, accuracy, probabilistic, behavior [16], meta-learning [32]. In the group-based criterion, the performance of a base classifier relates to its iterations with other classifiers in the pool. For example, diversity, data handling [196] and ambiguity [47] are widely used group-based performance metrics.

Once the candidates classifiers are selected, aggregation of results from these classifiers is then performed to give a united decision. There are three main strategies for results combination: static combiner, trained combiner and dynamic weighting. The majority voting scheme is a representative of static combiner, which is also commonly used in the traditional ensemble methods. In trainable combiners, the outputs of the selected based classifiers are used as the input features for another learning algorithm, such as [15, 129]. In dynamic weighting, higher weight value will be allocated to the most competent classifier and then the outputs of all the weighted classifiers are aggregated to give the united decision.

Result Refinement.

The aggregated result from the previous stage will be refined in this stage by our adjusted features strategy. Specifically, we train an SVM classifier with only HOS and wavelet coefficients (the RR-intervals are removed) to improve the results of S and V beats. The rationale of such a classification strategy is that the sensitivities to certain feature varies with heartbeat types [209] . For instance, the RR-intervals are indispensable for identifying disease heartbeats from the normal ones. However, the RR-intervals can also cause troubles to make a distinction between different kinds of disease heartbeats, such as S and V beats.

5.4.3 Multi-channels Heartbeat Convolution Neural Network

The architecture of the proposed Multi-channels Convolution Neural Network (MCHCNN) is presented in Fig 5.4. The network accepts two inputs: raw ECG heartbeat and heart rhythm (RR-intervals). As motivated by an electroencephalogram (EEG) processing network [170] which uses different sizes of convolution filters to capture temporal and frequency patterns from EEG signals, the proposed MCHCNN performs 3 channels of convolutions in parallel on the input ECG heartbeats to extract the temporal and frequency information. The convolution filter size varies with channels, where the smaller filter is used to capture temporal patterns and the larger filter is used to capture frequency patterns. We denote the convolution process as $Conv(x, y)$ in Fig 5.4, where x is the convolution filter size and y is the amount of the output feature maps. Each convolution operation is followed by a batch normalization and a ReLu activation. The batch normalization normalizes the output of the convolution by subtracting the batch mean and dividing by the batch standard deviation, which reduces the problem of *internal covariate shift* [76] and overfitting. The introduction of a ReLu activation is to allow the network to extract nonlinear features.

Every three stacked convolutions are wrapped into a building block and bypassed by a shortcut connection. The learned features are added to the shortcut at the end of each building block. Such a design helps to reduce the network degradation problem [69]. Each channel contains 3 building blocks. Learned features from the three channels are integrated by addition before a pooling layer. The pooling layer is used to reduce feature dimensions, after which the learned features are reduced to half-size. It helps to reduce the number of parameters in the following fully connected layer and lower the risk of overfitting.

A *Rhythm Integration* layer is specially designed to concatenate the learned features and the input heart rhythms. It reduces the impact brought by the lack of consideration of heart rhythms on identification of disease heartbeats in many existing network models.

Next, the dense layer is used to learn non-linear combinations of the learned features. The softmax layer gives the probabilities of each heartbeat type.

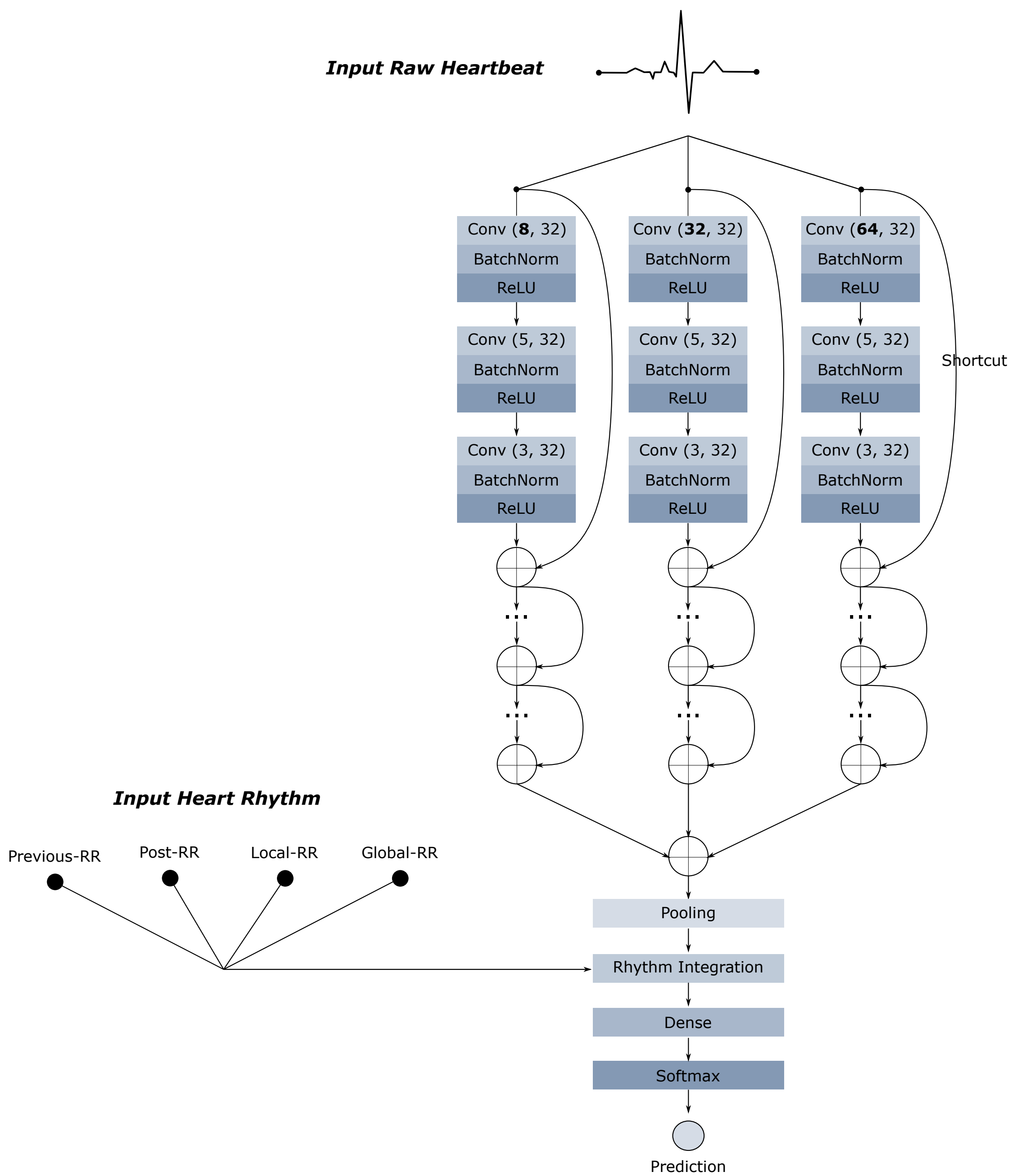


Figure 5.4: Architecture of the proposed Multi-channels Heartbeat Convolution Neural Network (MCHCNN).

5.5 Evaluation

In this section, we evaluate the proposed framework equipped with DHCAF and with MCHCNN, respectively. The MIT-BIH-AR database [137] is used as the benchmark database. It is the most representative database for arrhythmia detection and it has been used for most of the published research [38]. Details of the database are given below.

5.5.1 The MIT-BIH-AR Database

The MIT-BIH-AR database contains 48 two-leads ambulatory ECG records from 47 patients (22 females and 25 males). Each record has approximately 30 minutes in length. These recordings were digitized at $360Hz$. For most of them, the first lead is *modified limb lead II* (except for the recording 114). The second lead is a pericardial lead (usually $V1$, sometimes are $V2$, $V4$ or $V5$, depending on subjects).

In order to reveal the performance of the proposed framework, we follow the evaluation paradigm proposed in [38] to divide the database into a training and a test dataset. The paradigm avoids heartbeats of the same patient appearing in both training and test stages, ensuring a fair evaluation. Table 5.2 shows the division details, where DS1 is the training set and DS2 denotes the test set.

Noticing that DS1 is extremely imbalanced and dominated by N beats, we apply the SMOTEENN technique [6, 21, 191] on DS1 to over-sample the minority heartbeats (S and V) to the same amount of N beats.

5.5.2 Evaluation Metrics

Evaluation metrics used in this work are sensitivity (Se), positive predictive value ($+P$) and accuracy value (Acc). According to the AAMI standard [38], penal-

Table 5.2: Recording distributions and class proportions on DS1 and DS2.

Data set	N	S	V	F	Q	Recordings (Patient ID)[1]
DS1	45808	943	3786	414	8	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	44198	1836	3219	388	7	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

[1] Each recording is denoted by a 3-digits number and the numbers are originally discontinuous.

ties would not be applied for the misclassification of F and Q beats, as they are naturally unclassifiable.

5.5.3 Results of the proposed framework

Confusion matrixs of the proposed framework with DHCAF and with MCHCNN on DS2 are presented in Table 5.3. We summarize the results and compare our framework with multiple state-of-the-art methods in Table 5.4. All results reported in Table 5.4 are obtained under the same evaluation paradigm on DS2 of MIT-BIH-AR database.

It is clear that the proposed framework with DHCAF achieves the best sensitivity of both class S and V, and maintains a good performance in overall accuracy and classification of class N. Shan's model [23] obtains the highest accuracy and class N sensitivity. However, it fails in the detection of class S, with the sensitivity of class S being merely 29.5%, which limit the model's practical significance. The proposed framework with MCHCNN outperforms DHCAF in terms of the overall accuracy, sensitivity of N beats, and the positive predictive value of S beat, but its performance on sensitivity of S beat is less satisfactory. In fact, it can be found

Table 5.3: Confusion matrixs of **DHCAF** and **MCHCNN** on DS2.

		Predicted class (**DHCAF**)					Predicted class (**MCHCNN**)				
		N	S	V	F	Q	N	S	V	F	Q
True class	N	40698	2053	1199	0	119	42356	903	734	83	2
	S	125	1472	157	0	0	669	686	396	3	0
	V	102	103	3013	0	1	259	30	2924	6	0
	F	263	2	122	0	1	381	1	6	0	0
	Q	3	0	4	0	0	3	0	4	0	0

that the positive predictive values of S beats for most listed works in Table 5.4 are relatively low, as compared to other metrics. This is mainly caused by some N beats being misclassified as S beats. As we mentioned in the Introduction section, the similar QRS complex and the data imbalance problem have introduced a great difficulty in distinguishing the S from the N beats. We compare proposed framework with MCHCNN to another deep-learning based method by Sellami et al. [157], which reported model performance under the same unbiased evaluation. The results show that Sellami's work has achieved a promising performance on identification of both the problematic S and V beats, being close to that of the proposed framework with DHCAF. However, this is at the cost of the overall accuracy and sensitivity of normal beats. In real-world practices, the large amount of misclassification of normal heartbeats as the disease heartbeats will result in an unnecessary waste in medical resources.

From the above analysis, the proposed framework with DHCAF is believed to be a more appropriate choice than the other listed works for cardiac arrhythmia detection, because it achieves the best identification performance on disease heartbeats while maintaining a good overall accuracy and classification performance on the normal heartbeats.

Table 5.4: Arrhythmia detection results of the proposed framework and the stat-of-the-art methods on DS2

Method	Type	Acc	N		S		V	
			Se	+P	Se	+P	Se	+P
Proposed (DHCAF)	feature engineering	91.4	92.4	98.8	**84.0**	40.6	**93.6**	67.3
De Chazal [38]	feature engineering	81.9	86.9	99.2	75.9	38.5	77.7	81.9
Ye C [200]	feature engineering	86.4	88.5	97.5	60.8	52.3	81.5	63.1
Zhang Z [209]	feature engineering	86.7	88.9	99.0	79.1	36.0	85.5	92.8
Shan C [23]	feature engineering	93.1	98.4	95.4	29.5	38.4	70.8	85.1
Mariano L [115]	feature engineering	78.0	78.0	99.0	76.0	41.0	83.0	88.0
Proposed (MCHCNN)	deep learning	93.0	96.1	97.0	39.1	**42.3**	90.1	72.0
Sellami Ali [157]	deep learning	88.3	88.5	98.8	82.0	30.4	92.0	72.1

5.5.4 Ablative analysis

We perform ablative analysis for the proposed DHCAF and MCHCNN to demonstrate the effectiveness of the model architectures. The results are summarized in Table 5.5 and Table 5.6, respectively.

Two baselines are used in the ablative analysis for DHCAF. One is DHCAF with the result refinement stage removed. The other one is DHCAF with the dynamic ensemble selection classification of DHCAF replaced by ensemble of SVM classification. It is apparent that the result regulator has made unique contribu-

Table 5.5: Ablative analysis of DHCAF

Method	Acc	N		S		V	
		Se	+P	Se	+P	Se	+P
Proposed (DHCAF)	**91.4**	**92.4**	98.8	**84.0**	40.6	93.6	**67.3**
Proposed (DHCAF) without Result Refinement	90.8	92.4	98.8	63.6	36.1	95.0	60.8
Proposed (DHCAF) with SVM Ensemble	82.8	82.9	99.4	72.3	30.9	**96.7**	36.8

tions to DHCAF, with which the overall accuracy, sensitivity of class S, and positive predictive of class S and V are visibly increased. On the other hand, the poor classification performance of the SVM ensemble has demonstrated the importance of the introduction of *dynamic ensemble selection* to the proposed method.

As discussed in Section 5.2 and 5.3, many existing deep neural network models have not taken the heart rhythms into account for heartbeat classification, but this limitation is hidden by the over-optimistic results obtained in a biased evaluation paradigm. In the ablative test of MCHCNN, we want to know the actual impact of heart rhythms on model performances. Therefore, we construct a baseline MCHCNN which only takes raw ECG heartbeats as input. The results, as seen in Table 5.6, indicate that heart rhythms (RR-intervals) are necessary for identification of the disease heartbeats. Without consideration of heart rhythm, the baseline can hardly detect S beats. The detection on V beats is also affected. The outcome is in line with the medical fact. As we can see in Fig 5.1, most V beats present a huge morphological difference with other heartbeats. That is why the baseline can still maintain 73.8% sensitivity on V beats. However, for S beats, the heart rhythm is essential for distinguishing them from the normal heartbeats.

Although heartbeat rhythms have been part of the input to the proposed

Table 5.6: Ablative analysis of MCHCNN

Method	Acc	N		S		V	
		Se	+P	Se	+P	Se	+P
Proposed (MCHCNN)	**93.0**	96.1	**97.0**	**39.1**	**42.3**	**90.1**	72.0
Proposed (MCHCNN) without Heart Rhythms	92.1	97.5	95.0	10.3	20.1	73.8	86.0

MCHCNN, the S beats detection performance is still less satisfied. This indicates that the raw heartbeat rhythms provide limited assistance to our MCHCNN in identification of S beats. A possible explanation is that the heartbeat rhythms are not integrated well to the network and also easily affected by other learned features. A future study is needed to investigate this issue.

5.6 Chapter Conclusion

Millions of people around the world are suffering from cardiac arrhythmia. In this work, we propose a framework for automated arrhythmia detection from IoT-based ECGs. The framework consists of two modules: a data cleaning module to tackle the challenges presented by IoT-based ECGs, and a heartbeat classification module for identification the disease heartbeats. Specifically, we proposed two solutions, DHCAF and MCHCNN, for the heartbeat classification task. DHCAF is a feature-engineering based method which introduces the *dynamic ensemble selection* techniques and uses an adjust-feature strategy to assist disease heartbeats identification. By contrast, MCHCNN is an end-to-end solution that performs multi-channel convolutions to capture both the temporal and frequency information from the raw heartbeats to improve the classification performance. We evaluate the proposed framework on the MIT-BIH-AR database under the inter-patient evaluation paradigm. The results show that the proposed framework with DHCAF is a qualified candidate for automated arrhythmia detection from IoT-based ECGs. Besides, although the S beats detection performance of MCHCNN is less satisfied, the network still provide some insights to our future study.

This work is a first step to provide a solution for the automated arrhythmia detection in the era of *Internet-of-Things*. In our next study, we aim to investigate a more effective way for integration of the heart rhythms into a neural network.

CHAPTER 6

AN ADVANCED TWO-STEP DNN-BASED FRAMEWORK FOR ARRHYTHMIA DETECTION

6.1 Chapter Abstract

Heart arrhythmia is a severe heart problem that takes people's lives without warning at every corner of the world. Automated heartbeat classification provides a cost-effective screening for heart arrhythmia and allows at-risk patients to receive timely treatments, which is a highly demanded but challenging task. Deep neural networks (DNNs) have brought visible improvements to this area, but to identify the problematic *supraventricular ectopic* (S-type) heartbeats is still a bottleneck in most of the existing studies. This is mainly due to morphological similarity between the S-type heartbeats and the normal ones, the imbalanced heartbeat occurrence rate, and both the inter- and intra-patients variations in heart rhythms. This paper presents a two-step DNN-based classification framework to identify problematic heartbeats for arrhythmia detection. In the first step, a deep dual-channel convolutional neural network (DDCNN) is proposed to classify all heartbeat classes, except for the S-type and the normal heartbeats. In the second stage, a central-towards LSTM supportive model (CLSM) is specially designed to distinguish S-type heartbeats from the normal ones. By processing heart rhythms in central-towards directions, the proposed CLSM learns and abstracts hidden temporal information between a heartbeat and its neighbors to reveal the deep differences between the two heartbeat types. As an improvement, we also propose a rule-based data augmentation method to solve the training data imbalance problem. The proposed framework is evaluated over three real-world ECG databases. The results show that our method outperforms many baselines methods in most

evaluation metrics.

6.2 Introduction

Heart arrhythmia is known as abnormal heart rhythms, in which the heart beats too fast, too slowly or erratically. Arrhythmia threatens people's lives by preventing their hearts from pumping enough blood into the vital organs. It has been a major worldwide health problem for years, accounting for nearly 12% of global deaths every year [133]. Early detection and timely treatment are the keys to survival from arrhythmia. The electrocardiogram (ECG) plays a pivotal role in the diagnosis of arrhythmia since it captures heart rate, rhythm, and vital information regarding the electrical heart activities and related conditions. However, the manual interpretation of ECG recordings is time-consuming and error-prone, especially for the long-term ECG recording, which is essential for capturing the sporadically occurred arrhythmia [209]. Therefore, an automated method to assist clinicians in detecting arrhythmia heartbeats from ECG is highly demanded.

Heartbeat classification on ECG is a core step towards identifying arrhythmia. As reported by the Association for Advancement of Medical Instrumentation (AAMI) [8], heartbeats can be categorized into five super classes: Normal (N), Supraventricular (S) ectopic, Ventricular (V) ectopic, Fusion (F) and Unknown (Q). In particular, problematic arrhythmias are mostly found in ***S*-type** and ***V*-type** heartbeats [38]. Fig.6.1 shows several ECG samples of different heartbeat types. It can be observed that the V-type heartbeat exhibits a huge morphological difference against other heartbeats, while the *normal* (N-type) and the S-type heartbeats are similar in shape. It is less likely to provide accurate identification of the S-type heartbeats from the normal ones merely based on the morphology. In clinical practice, special rhythm information between two heartbeats, known as the

RR-interval, is needed to help identify the S-type heartbeats because the S-type heartbeats are premature heartbeats and they normally have shorter previous-RR-intervals than the normal heartbeats. However, the inter- and intra-patients variations existing in the heart rhythms still impose great challenges to the detection tasks. Besides, the sporadical occurrence of the S-type heartbeats (most commonly normal heartbeats for the majority of patients) can also be an issue that tends to bias an automated heartbeat classification method.

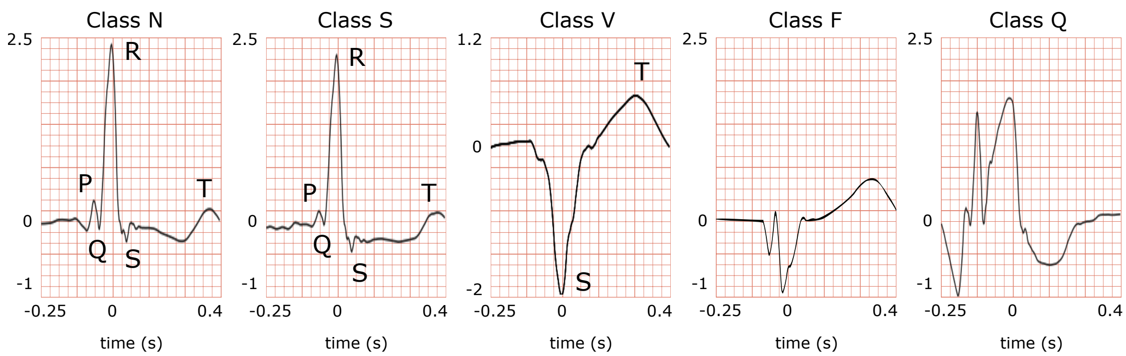


Figure 6.1: Examples of different types of heartbeats. Letters indicate the P-waves, R-peaks, QRS-complexes and T-waves, corresponding to their references in the medical literature. Time gap between two successive R peaks is known as RR-interval. Specifically, **previous-$\boldsymbol{RR}$-interval** denotes the interval between the current R peak and the previous R peak. In comparison to the normal heartbeat (class N), the S-type heartbeat has a less obvious P-wave which is due to *junctional premature beating*. The V-type heartbeat exhibits a deep and capacious S-wave caused by *left bundle branch block*. Class F is a fusion of paced and normal heartbeats. The unclassifiable beat is denoted as class Q.

Existing solutions to the heartbeat classification problem are mostly following a traditional pattern recognition paradigm [23, 38, 200, 209], in which the fluctuations of the raw ECG signals are modeled by a set of carefully extracted features, such as RR-intervals, wavelet coefficients, and morphological amplitudes. However, pattern-based classification models often experience difficulty in achieving satisfactory performance on abnormal heartbeat detection, especially when S-type arrhythmia heartbeats are involved. Besides, the effectiveness of extracted features, the mutual-influences among features, and the compatibility between the feature distribution and the classifiers [33] are three major factors that lead to a solid upper-bound on model performance.

Recent advances in heartbeat classification are largely driven by deep neural networks (DNNs). A DNN is a computational model consisting of multiple processing layers, which can automatically learn the high-level representations of the raw ECG recordings without extensive data preprocessing. In consideration of the sporadic occurrence of S-type heartbeats, which imposes a great challenge to DNN training, many DNN-based studies used synthetic heartbeats for model training and evaluation [3, 113, 114, 203, 207]. However, these efforts suffer from data leakage because, after augmentation, data is not partitioned patient-wise into training and test sets. So that beats from the same patient may appear in training and test, and the deep learning algorithms may learn patient-specific characteristics during training which then appear on test data. Additionally, the over-optimistic results obtained from data leakage have hidden a potential limitation of these DNN models in which only the ECG segmented heartbeats are accepted as inputs. The inter-heartbeat rhythm information is not well considered in these models. As mentioned above, the rhythm provides indispensable information to distinguish the S-type arrhythmia heartbeats. Without such information, a high misclassification rate is probably obtained on S-type heartbeats. The problem is still open.

Contributions: In this work, we propose a two-step deep neural network-based heartbeat classification framework. Due to the observed difficulty of detecting S-type heartbeats from N-type heartbeats, the proposed framework trains a deep dual-channel convolutional neural network (DDCNN) which accepts segmented heartbeats as input in the first step to classify V-type, F-type and Q-type heartbeats. At this stage, S-type and N-type heartbeats are not the targets, so they are put into one bundle to be studied in the next step. In the second step, a central-towards LSTM supportive model (CLSM) is specially designed to distinguish S-type heartbeats from N-type ones. The RR-intervals of a heartbeat and its neighbors are arranged in sequence form, serving as the input to CLSM. In particular, CLSM learns and extracts hidden temporal dependency between heartbeats by processing the input RR-interval sequence in central-towards directions. Instead of using raw individual RR-intervals, the abstractive, mutual-connected temporal information provides stronger and more stable support for identifying the problematic S-type heartbeats. Besides, as an improvement as well as a necessary driver for activating the CLSM, a rule-based data augmentation method is also proposed to supply high-quality synthetic samples for the under-represented S-type RR-interval sequences. To avoid data leakage, the benchmark dataset is split into training and test sets at patient level following the well-recognized inter-patient division paradigm proposed in [38]. The synthetic training samples are generated from the training set only.

The rest of the chapter is structured as follows: Section 6.3 reviews the related Deep Neural Network algorithms that are widely used in arrhythmia detection work; Section 6.4 shows the details of the proposed framework and the rule-based data augmentation method; Section 6.5 presents our experiments on real-world ECG data and discusses the results; finally, the entire chapter is summarized with our achievements in Section 6.6.

6.3 Related Work

The proposed DDCNN and CLSM are novel models that involve convolutional neural networks, residual network, and long short-term memory networks. In this section, all three related deep learning algorithms are reviewed to provide a basic understanding of the proposed models to be presented in the next section.

6.3.1 Convolutional Neural Networks

The convolutional neural network (CNN) is useful in learning representations of data. CNNs are commonly applied in image and video recognition, recommend systems, image classification, and natural language processing. Recently, CNNs have attracted more and more attention in the applications in ECG signal classification due to their capability of effectiveness in recognizing key patterns and learn useful features, such as P-waves and QRS-complexes of ECG heartbeats [65].

A convolutional neural network is normally made up of an input layer, an output layer, multiple convolutional layers, pooling layers, and dense Layers. The convolutional layer is the core building block, where most of the computational heavy lifting is done. Given input data, a convolution is known as a linear combination of each data point with its neighbors. It is usually followed by a ReLU activation to enable the network to learn non-linear patterns from the input data. The pooling layer is used to reduce the size of the learned representations to reduce the number of network parameters. The purpose of the dense layer is to provide an overall regulation of the previously learned representations.

6.3.2 Residual Networks

The residual network (ResNet) was first introduced in [69] to solve the network degradation problem: with a network getting deeper, its accuracy gets saturated and then degrades rapidly. Fig. 6.2 shows a sample ResNet block, which consists of two stacked layers and an identity mapping (the shortcut connection).

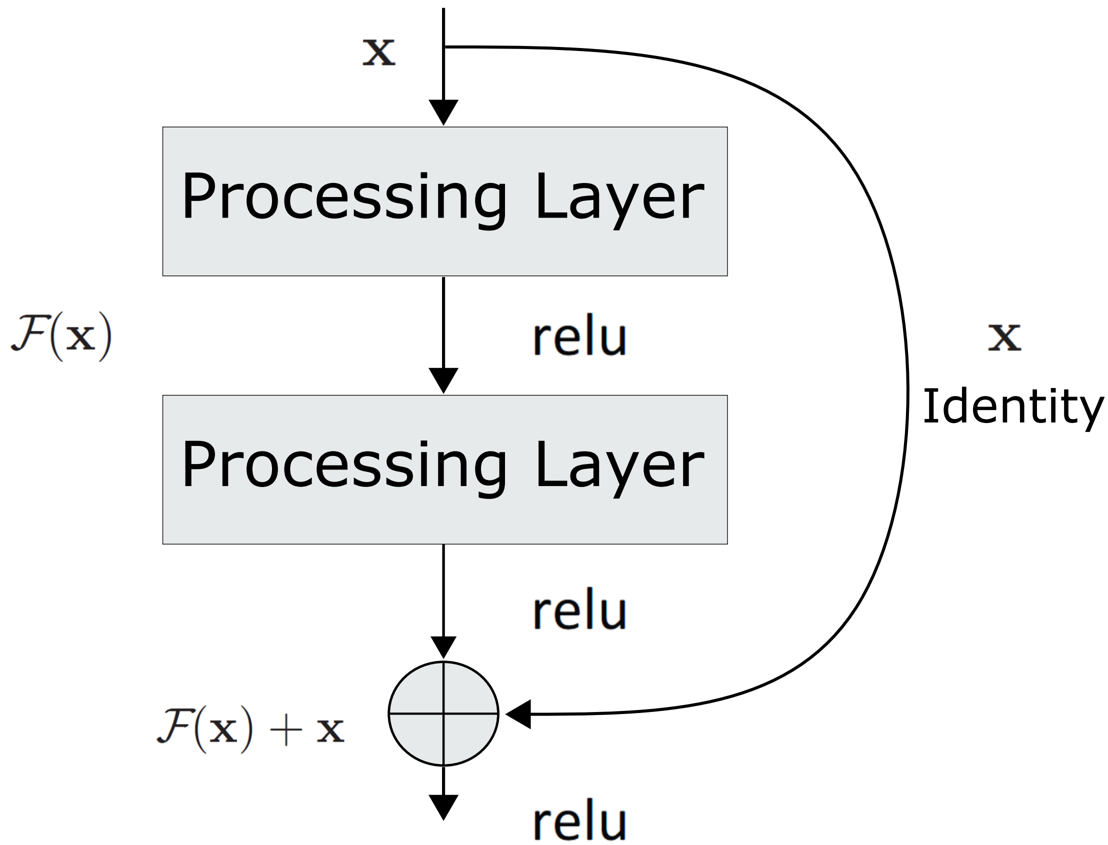


Figure 6.2: A Sample ResNet Block.

The core idea of ResNet is to use the stacked processing layers, such as convolution layer and dense layer, to fit the residual mapping. The rationale behind such a design is that if the input has already been optimal, it would be easier to push the stacked layers to zero than to make them an identity mapping. Hence, the network will not degrade with depth.

Let $\mathbf{x}$ denotes the input and $\mathcal{H}(\mathbf{x})$ denotes the desired underlying mapping. A residual mapping is expressed as :

$$\mathcal{F}(\mathbf{x}) := \mathcal{H}(\mathbf{x}) - \mathbf{x}. \tag{6.1}$$

Therefore, the desired underlying mapping $\mathcal{H}(\mathbf{x})$ can be recast into :

$$\mathcal{H}(\mathbf{x}) := \mathcal{F}(\mathbf{x}) + \mathbf{x} \tag{6.2}$$

and represented by the stacked layers plus with an identity mapping.

6.3.3 Long Short-Term Memory Networks

The long short-term memory network (LSTM) [73] is a variation of recurrent neural networks (RNN). It alleviates the vanishing gradient problem presented in the ordinary RNNs and it is able to learn temporal relationship across long periods of time.

A common LSTM unit consists of a *cell* c_t, an *input* gate i_t, a *forget* gate f_t and an *output* gate o_t, as shown in Fig.6.3. The *cell* remembers the time dependency between elements in the input sequence. Its memory can be effectively conveyed along the entire processing chain with just limited linear interactions. The *input* gate controls the new information to be stored in the *cell*. The *forget* gate decides the information to be thrown away from the *cell*. The unit output is managed by the *output* gate based on the current *cell*'s memory.

As an illustration, let W_n and U_n be the weights of inputs and recurrent connections respectively, and b_n be the bias. The subscript n can be the *forget* gate f, *input* gate i, *output* gate o or the *cell* c. Given the input x_t, the LSTM unit at time t is updated as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{6.3}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{6.4}$$

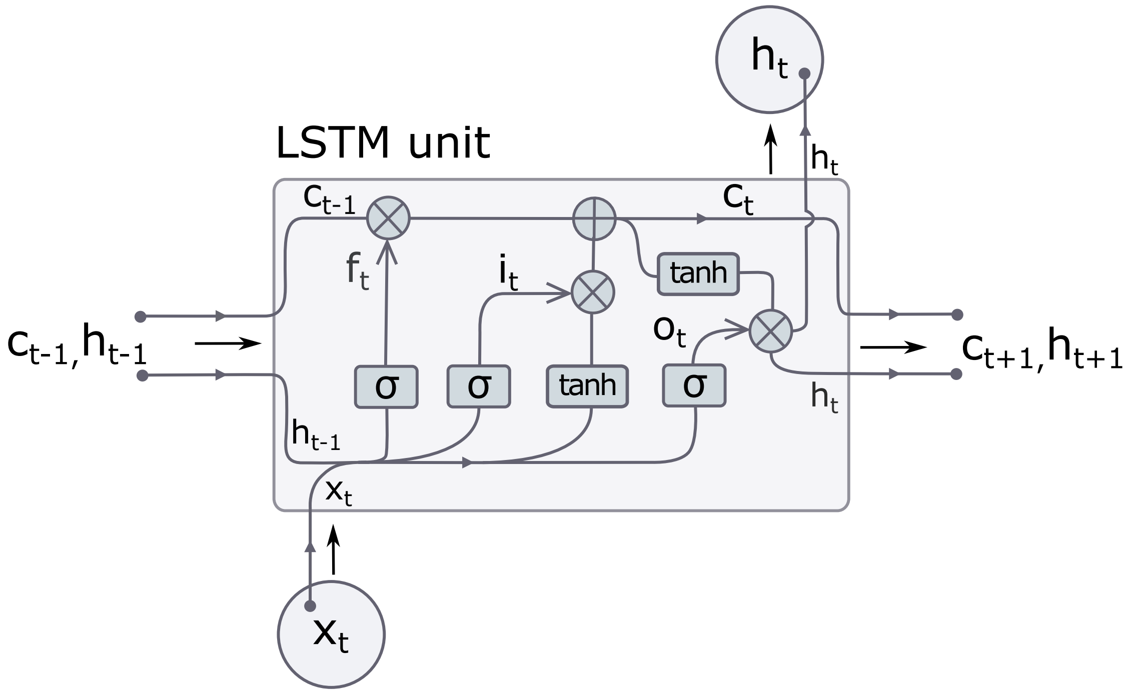


Figure 6.3: A LSTM Unit.

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{6.5}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{6.6}$$

$$h_t = o_t \circ tanh\left(c_t\right) \tag{6.7}$$

where σ represents the sigmoid function and the operator $\circ$ denotes the element-wise product.

6.4 The Proposed Framework for Arrhythmia Detection

The proposed framework consists of DDCNN and CLSM. DDCNN is used to capture both the morphological and frequent patterns of heartbeats, and CLSM is specially designed to handle the temporal information between heartbeats. Details of these two models and the proposed data augmentation for driving CLSM are stated below.

6.4.1 Deep Dual-channel Convolutional Neural Network

The architecture of DDCNN is presented in Fig.6.4. The network accepts segmented ECG heartbeats (modified limb lead II, sampled at 360Hz) as input, and outputs a prediction of probabilities of the $N\&S$-bundle, V, F and Q classes. Being inspired by the fact that signal processing experts capture the temporal and frequency patterns in electroencephalogram (EEG) signals with different sizes of convolutional filters [170], the proposed DDCNN is designed as a dual-channel convolutional neural network, with the small filter channel $Conv(8, 32)$ dealing with the temporal information and the larger filter channel $Conv(64, 32)$ handling the frequency information. The learned temporal and frequency information are added together before the pooling operation. The entire DDCNN contains 18 convolutional layers, a pooling layer, a concat layer, a dense layer, and a *softmax* output layer. Specifically, the **concat** layer is designed to concatenate rhythm information (RR-intervals) to assist heartbeat classification. Each convolution operation is followed by a batch normalization and a ReLU activation. Every three convolutional layers of each channel are packed into a residual block and bypassed by a shortcut connection. The stacked residual blocks design reduces the network degradation risk and accelerates the training process.

6.4.2 Central-towards LSTM Supportive Model

The proposed CLSM consists of two specially designed central-towards LSTM layers and one *softmax* output layer. The term '*central-towards*' means that information in an LSTM chain flows from the farthermost units in both sides towards the central unit, without crossing over with each other. A graphical representation of our model is provided in Fig. 6.5.

The proposed CLSM accepts heartbeats' previous-RR-interval sequences as

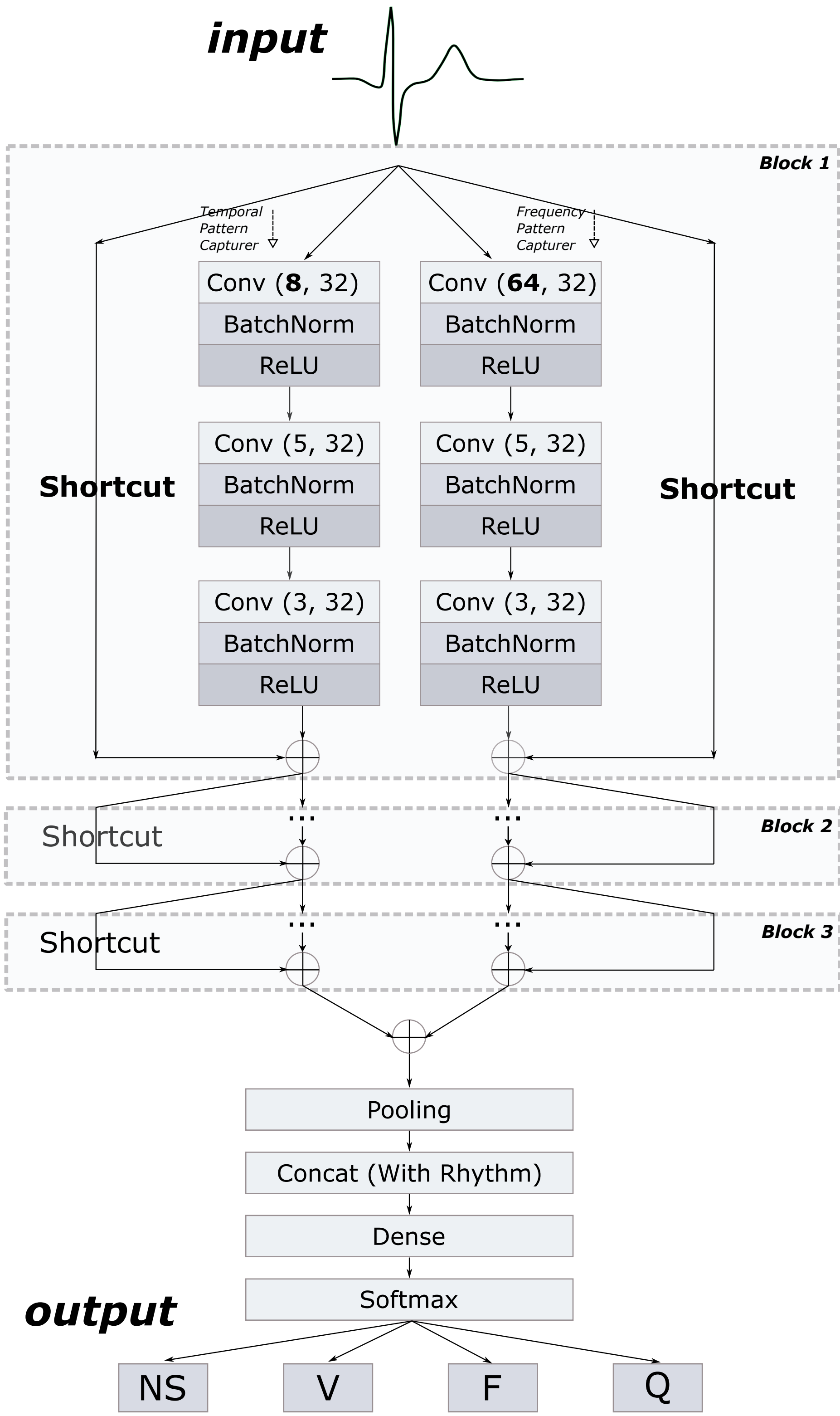


Figure 6.4: Architecture of DDCNN, where $Conv(x, y)$ denotes a convolutional layer with a kernel in size x and an output of y feature maps.

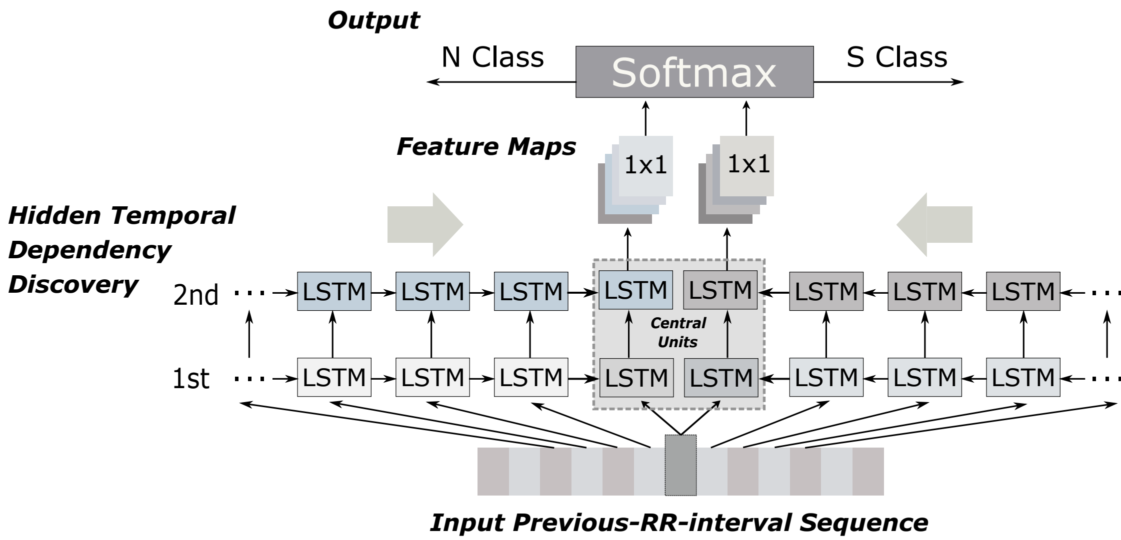


Figure 6.5: Central-towards LSTM Supportive Model architecture.

inputs. A previous-RR-interval sequence of the t^{th} heartbeat hb_t is defined as

$$S_t = [R_{t-NeRan}, ..., R_{t-1}, R_t, R_{t+1}, ..., R_{t+NeRan}] \tag{6.8}$$

where R_t denotes the RR-interval between the $t-1^{th}$ and t^{th} heartbeats, and $NeRan$ defines the range of a heartbeat's neighborhood. The default value of $NeRan$ is 25. A previous-RR-interval sequence S_t is labeled as the same label of the central heartbeat hb_t, which is N-type or S-type.

Each central-toward LSTM layer contains $2*(NeRan+1)$ common LSTM units. The outputs of the 1^{st} layer serve as the inputs to the 2^{rd} layer. Particularly, the central units receive and process the learned temporal dependencies from the previous and the next heartbeats, respectively. There is no pathway between the two central LSTM units in each layer, which is to avoid the mutual-interruption of the learned temporal dependencies from both directions.

Given an input S_t, update equations of a unit in the proposed central-towards

LSTM layer depend on the unit's position n at the layer, where $n \in [0, 2*NeRan+1]$. Let $g_{f,n}$, $g_{i,n}$, $g_{o,n}$, h_n denote the *forget* gate, *input* gate, *output* gate, and the output of the n^{th} unit, respectively.

$$g_{q,n} = \begin{cases} \sigma\left(W_q S_t[n] + U_q h_{n-1} + b_q\right), n < NeRan + 1 \\ \\ \sigma\left(W_q S_t[n] + U_q h_{n+1} + b_q\right), n \geq NeRan + 1 \end{cases}, q \in f, i, o \tag{6.9}$$

where W and U are the weight matrix of inputs and recurrent connections, respectively, and b denotes the bias. We define the change of the memory as:

$$\tilde{c}_n = \begin{cases} tanh\left(W_c S_t[n] + U_c h_{n-1} + b_c\right), n < NeRan + 1 \\ \\ tanh\left(W_c S_t[n] + U_c h_{n+1} + b_c\right), n \geq NeRan + 1 \end{cases} \tag{6.10}$$

Then the cell state is determined by the following equation:

$$c_n = \begin{cases} g_{f,n} \circ c_{n-1} + g_{i,n} \circ \tilde{c}_n, \quad n < NeRan + 1 \\ \\ g_{f,n} \circ c_{n+1} + g_{i,n} \circ \tilde{c}_n, \quad n \geq NeRan + 1 \end{cases} \tag{6.11}$$

The output of the unit depends on the cell state, which is given by:

$$h_n = g_{o,n} \circ tanh\left(c_n\right) \tag{6.12}$$

In the 2^{nd} central-towards LSTM layer, the central units output 32 feature maps in size 1×1. The feature maps are flattened before being processed by a *softmax* function for classification. The model outputs probabilities of the central heartbeat being normal and S-type.

6.4.3 Rule-based Data Augmentation

The sporadic occurrence of the S-type heartbeats has resulted in a serious class imbalance problem in the benchmark training heartbeat data, which puts an obstacle to the successful training of CLSM. To generate synthetic samples for the under-represented S-type heartbeats becomes necessary and critical. Many oversampling techniques, such as SMOTE [21], have been introduced for data augmentation purpose, but these techniques are mainly designed for data samples that are represented by extracted features. Elements in a heartbeat's previous-RR-interval sequence $S_t = [R_{t-NeRan}, ..., R_{t-1}, R_t, R_{t+1}, ..., R_{t+NeRan}]$ have evident linear correlations, which are different from the mutual-independent features. Applying the existing oversampling methods will introduce invalid samples and make the training even worse.

To solve the problem, we propose a rule-based data augmentation method to generate synthetic previous-RR-sequences of the S-type heartbeats. Basically, a valid synthetic previous-RR-interval sequence is subject to 3 medical facts:

a. S-type heartbeats normally have shorter previous-RR-intervals than the normal ones.

   ($question_1$: what is the valid range of distance between previous-RR-intervals of S-type and normal heartbeats?)

b. Heartbeats of the same type exhibit a limited variation in the previous-RR-intervals within a short period.

   ($question_2$: how much the variation is?)

c. Some normal heartbeats can be found within the neighborhood scope of a S-type beat.

   ($question_3$: how many normal heartbeats can be found?)

The above medical facts provide a qualitative overview of what a valid synthetic sample should be. To synthesize a new valid sample, we still need to explicitly answer the questions following each medical fact.

The proposed method seeks for the answers via performing a statistical analysis on the benchmark training set (DS1 of MIT-BIH-AR [137]). We define three variables, gap, $varPct$ and $nAmt$ for question 1, 2 and 3, respectively. Statistically, we have:

$$\begin{aligned} gap &\approx 0.1; \\ varPct &\approx 3\%; \\ given\ NeRan = 25,\ nAmt &= Range([0, 48]). \end{aligned}$$

Let $nVals$ and $sVals$ be the collections of previous-RR-intervals of the normal and the S-type heartbeats in the training set, respectively. The proposed rule-based data augmentation method is thoroughly illustrated in Algorithm 7.

By complying with the rules (line 3, 4 & 6 in Alg.7) and creating new combinations (line 2, 5 & 7 in Alg.7) from the existing data, our method is able to generate high-quality and diversified training samples for activating CLSM.

Algorithm 7 Rule-based Data Augmentation

Require: gap, $varPct$, $nAmt$, $nVals$ and $sVals$;

Ensure: $synSeq$;

1: $synSeq \leftarrow new(list)$;

2: $centralS \leftarrow$ a random pick from $sVals$;

3: $amount \leftarrow$ a random pick from $nAmt$;

4: $var \leftarrow$ a random float in $[1 - varPct,\ 1 + varPct]$;

5: $candidate \leftarrow$ a random pick from $nVals$;

6: **while** $candidate < centralS + gap$ **do**

7:     $candidate \leftarrow$ a random pick from $nVals$;

8: **end while**

9: **for** i in range($amount$) **do**

$synSeq.add(candidate * var)$;
end for
for i in range($2 * NeRan - amount - 2$) **do**
$synSeq.add(centralS * var)$;
end for
shuffle($synSeq$);
Insert $centralS$ into the central position of $synSeq$;
return $synSeq$;

6.5 Experimental Evaluation

Extensive experiments on three real-world ECG databases are implemented to evaluate the proposed framework and the rule-based data augmentation method. Experimental databases, settings, and results are discussed as follows.

6.5.1 Arrhythmia Datasets

The real-world ECG datasets used in this study are: (1) MIT-BIH Arrhythmia database (MIT-BIH-AR); (2) MIT-BIH Supraventricular Arrhythmia database (MIT-BIH-SUP) and (3) St.-Petersburg Institute of Cardiological Technics 12-lead Arrhythmia database (INCART). The databases are all publicly available in the Physiobank [137].

MIT-BIH-AR is the benchmark database for arrhythmia detection, which is used in most published research [119]. To fairly compare against existing methods, we train and test our framework in this database following the well-recognized inter-patient evaluation paradigm [38]. MIT-BIH-SUP and INCART are used to demonstrate the generalizability of the proposed framework to external data. Details of the databases are listed as follows.

- **MIT-BIH-AR** contains 48 two-lead ambulatory ECG recordings from 47 patients (including 22 females and 25 males). Each recording is denoted by a 3-digits number. The recordings were digitized at 360 Hz per second per channel with 11-bit resolution over a 10-mV range. For most of them, the first lead is modified limb lead II (except for the recording 114). The second lead is a pericardial lead (usually *V1*, sometimes are *V2*, *V4* or *V5*, depending on subjects). In this study, only the modified limb lead II is used.
- **MIT-BIH-SUP** consists of 78 two-leads recordings, with each of them approximately 30 minutes in length. The recordings are sampled at 128 Hz. They were chosen to supplement the examples of supraventricualr arrhythmias in the MIT-BIH-AR database. Heartbeats in the recordings are well annotated and the original labeling can be adapted to the AAMI recommendations.
- **INCART** consists of 75 ECG recordings sampled at 257 Hz. Each recording contains 12 standard leads. Similarly, only the modified limb lead II is used in this study. The annotations were first produced by an automatic algorithm and then corrected manually based on the standard PhysioBank beat annotation definitions. None of the recordings contains pacemakers, but most of them have ventricular ectopic heartbeats.

6.5.2 Experiment Setup

The experimental setup procedures are shown as follows.

- *Benchmark Training and Test Datasets.* We divide MIT-BIH-AR into a training and a test set at patient level following the well-recognized inter-patient evaluation scheme [38]. Table 6.1 presents the division in detail, where **DS1** is the training set and **DS2** in the test set. The division scheme balances the heartbeat distribution on both DS1 and DS2, and more importantly, it avoids the training and the test heartbeats coming from the same patient.

- *Heartbeats Segmentation.* We segment each recording to heartbeats based on the R peak locations in notations. For each R peak, 70 samples (200-ms) before R peak and 100 samples (280-ms) after R peak were taken to represent a heartbeat. This is long enough to catch the samples representing the re-polarization of ventricular and short enough to exclude the neighbor heartbeats. After segmentation, the heartbeat distributions of each dataset are shown in Table 6.2.

- *Previous-RR-Interval Sequence Generation.* For each segmented normal or S-type heartbeat hb_t, we extract the previous-RR-intervals of 25 heartbeats previous and next to hb_t, respectively. The previous-RR-interval sequence of hb_t is then given by $S_t = [R_{t-25}, ..., R_{t-1}, R_t, R_{t+1}, ..., R_{t+25}]$, where R_t denotes the previous-RR-interval of hb_t.

- *Data Augmentation.* We generate synthetic S-type previous-RR-interval sequences from the training set (DS1) using our rule-based data augmentation method. After data augmentation, the sequences for training CLSM are made up of 44738 normal and 45908 S-type sequences.

- *Training Specification.* Both the proposed DDCNN and CLSM are trained with a variant of the gradient decent algorithm named Adam [92]. The learning rate is set to 0.001 with no decay. The *Categorical Cross-Entropy* function is used to measure the loss.

Table 6.1: The inter-patient division paradigm (for MIT-BIH-AR)

dataset	Recordings[1]
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230
DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234

[1] Recordings 102, 104, 107 and 217 containing paced beats are excluded [8].

Table 6.2: Heartbeat distributions in MIT-BIH-AR, MIT-BIH-SUP and INCART

Database	N	S	V	F	Q
DS1	45808	943	3786	414	8
DS2	44198	1836	3219	388	7
MIT-BIH-SUP	158760	11976	9718	23	76
INCART	150210	1917	19621	218	6

- *Evaluation Metrics.* The evaluation metrics used in this study are accuracy (ACC), precision (PRE), recall (REC) and $f1$ score (F1).

6.5.3 $Experiment_1$: Overall heartbeat classification

In this section, we evaluate the heartbeat classification performance of the proposed framework on the benchmark database and compare the results against multiple baseline algorithms [3, 23, 38, 115, 200, 209] derived from literature. Table 6.3 summarizes the comparative results. The comparison focuses on the normal, S-type and V-type heartbeats because, according to the AAMI standard [38], the F-type and Q-type heartbeats are naturally unclassifiable and penalties should not be applied for the misclassification of these heartbeats. The proposed DDCNN + CLSM architecture performs significantly better than the baseline algorithms on the overall accuracy (95.1% vs 78.0% - 93.1%), F1 score of normal heartbeats (97.6% vs 87.3% - 96.9%), recall rate of S-type heartbeats (83.8% vs 29.5% -

76.0%), precision rate of S-type heartbeats (59.4% vs 36.0% - 52.3%), and F1 score of S-type heartbeats (69.5% vs 33.4% - 56.3%). The performance on V-type beats is above the average, ranking the 3^{rd} place of the listed works.

It is apparent from Table 6.3 that most of the listed works struggled in the detection of the S-type heartbeats. We re-implement and evaluate the DNN model proposed in [3] following the inter-patient paradigm. The result confirms that, without considerations of heart rhythm, a DNN is less likely to identify S-type heartbeats. Zhang et al. [209] and Mariano et al. [115] achieve close recall rates of the S-type heartbeats as our framework, but they sacrifice S-type heartbeats precision rates (36.0% and 41%, respectively) and normal heartbeats recall rates (88.9% and 78.0%, respectively). This implies that both these two works misclassify a large portion of normal heartbeats as S-type heartbeats. In clinical practice, the erroneous misclassification of normal heartbeats as disease heartbeats leads to unnecessary additional tests, unnecessary patient treatments, expensive costs, and risks for patients.

An ablative analysis is also performed. We remove CLSM from the proposed framework and use standalone DDCNN for overall classification of all five types of heartbeats. The result is shown as **DDCNN Only** in Table 6.3. To further investigate whether raw RR-intervals help to identify problematic heartbeats, we train a DDCNN without the concat layer for comparison. The result is denoted as **DDCNN Only (without Concat)**. It is clear that, without the proposed CLSM, both standalone DDCNNs can hardly detect the S-type heartbeats. The DDCNN with the concat layer performs better on both S-type and V-type heartbeats than the DDCNN without the concat layer. The outcome indicates that RR-intervals help to identify problematic heartbeats, especially for S-type heartbeats, but the assistance of raw RR-intervals is limited because they are likely to be influenced

Table 6.3: Performance comparison on DS2 of MIT-BIH-AR

Method	ACC	N			S			V		
		REC	PRE	F1	REC	PRE	F1	REC	PRE	F1
DDCNN + CLSM	**95.1**	97.5	97.6	**97.6**	**83.8**	**59.4**	**69.5**	80.4	90.2	85.0
DDCNN Only	93.4	97.9	95.7	96.7	13.2	20.7	16.1	87.2	87.7	87.5
DDCNN Only (without Concat)	85.9	90.2	95.9	93.0	3.9	3.5	3.7	82.4	46.3	59.2
Acharya, U.R [3]	71.3	73.3	95.0	82.6	6.3	2.3	3.4	90.8	28.2	43.5
De Chazal [38]	81.9	86.9	99.2	92.6	75.9	38.5	51.1	77.7	81.9	80.0
Ye C [200]	86.4	88.5	97.5	92.8	60.8	52.3	56.3	81.5	63.1	71.2
Zhang Z [209]	86.7	88.9	99.0	93.7	79.1	36.0	49.5	85.5	**92.8**	**89.0**
Shan C [23]	93.1	**98.4**	95.4	96.9	29.5	38.4	33.4	70.8	85.1	77.3
Mariano L [115]	78.0	78.0	**99.1**	87.3	76.0	41.0	53.3	83.0	88.0	85.4

* Results in this table are presented in percentage (%), which are obtained on DS2 of MIT-BIH-AR following the same evaluation procedures.

Table 6.4: Generalization performances (%) on MIT-BIH-SUP and INCART

Method	Dataset	ACC	N		S		V	
			REC	PRE	REC	PRE	REC	PRE
Proposed	MIT-BIH-SUP	88.2	90.6	97.8	72.6	53.5	70.0	43.0
Proposed	INCART	**91.6**	**92.0**	**99.6**	81.0	**14.4**	**91.0**	81.9
Mariano L [115]	INCART	91.0	92.0	99.0	85.0	11.0	82.0	88.0

by the intra- and inter-patients variations. Therefore, having a consideration of neighbor heartbeats and performing an abstraction of the temporal dependency from the raw RR-intervals is necessary.

6.5.4 $Experiment_2$: Generalization of the proposed framework

We apply the proposed framework (trained in DS1) to MIT-BIH-SUP and INCART to demonstrate its generalizability. To be fitted, ECG recordings in these two databases are re-sampled to 360 Hz. Table 6.4 summarizes the results.

To the best of our knowledge, this work is the first one to report heartbeat classification results on MIT-BIH-SUP. When being applied on MIT-BIH-SUP, the proposed framework experiences a slight performance drop on the V-type heartbeats detection. However, this is mainly due to the low-resolutions of the source ECG recordings which are originally sampled at 128 Hz.

We compare the proposed framework to Mariano's work [115] on INCART. Mariano's work is one of the few works that conduct model evaluation on both MIT-BIH-AR and INCART. The results show that both works achieve promising performances, where the proposed framework slightly outperform Mariano's

Table 6.5: The impact of data augmentation method on CLSM's performance

Method	ACC	N		S	
		REC	PRE	REC	PRE
CLSM + Rule-based Method	**97.7**	**98.2**	**99.4**	**85.6**	**65.7**
CLSM + SMOTE	94.7	97.7	96.8	19.6	25.5

work [115] in majority metrics. The commonly low precision rates of the S-type heartbeats are due to the extreme imbalance of the INCART database.

6.5.5 $Experiment_3$: Rule-based data augmentation versus SMOTE

We investigate the effectiveness of our rule-based data augmentation method in this section. The SMOTE algorithm [21] is used as a baseline. We train individual CLSMs with the rule-based augmented sequences and the SMOTE augmented sequences, respectively, and evaluate their classification performances using all the normal and S-type heartbeats in DS2. Table 6.5 summarizes the results.

Apparently, SMOTE failed to generate valid previous-RR-interval sequences for training the proposed CLSM. The CLSM trained with SOMTE-generated samples couldn't effectively identify the S-type heartbeats, with both the recall and precision rates being lower than 30%. The poor result is not surprising because the SMOTE method is designed for featurized data oversampling. Thus, data like previous-RR-interval sequences with internal connections between elements will disable the SMOTE method. By contrast, using the medical rules as a guide, the proposed rule-based data augmentation method can generate high-quality synthetic sequences that reflect the true distribution of the real-world data to support the CLSM.

6.5.6 Discussion

Experimental results achieved on the three real-world ECG databases have proven the effectiveness and the robustness of the proposed framework and indicated that the proposed framework has the potential to make a substantial clinical impact. In particular, the proposed CLSM structure distinguishes our framework from the others. It provides a promising solution for separating S-type heartbeats from normal heartbeats which is one of the most problematic tasks for existing arrhythmia detection methods.

While CLSM has provided a novel idea of how to incorporate heart rhythm to help individual heartbeat classification, we have implemented several experiments to demonstrate how CLSM outperforms traditional LSTMs, and to investigate the performance-influencing factors of CLSM. Moreover, we have also implemented an experiment to explore the supportive nature of CLSM.

Central-towards LSTM VS traditional LSTMs

We compare the proposed central-towards LSTM layer against two baseline LSTM layers, the ordinary LSTM and the bidirectional LSTM. For a fair comparison, we replace the central-towards LSTM layers in the proposed CLSM with ordinary LSTM and bidirectional LSTM layers, respectively, keeping other layers and training data unchanged and all hyperparameters in default. The results are shown in Table.6.6, which are obtained on all the normal and S-type heartbeats in DS2. Apparently, the proposed central-towards LSTM layers exhibits an overwhelming superiority against the ordinary and bidirectional LSTM layers in discovering useful temporal dependencies from a previous-RR-interval sequence for classification of the central heartbeat. One possible explanation is that, in ordinary LSTM and bidirectional LSTM networks, the learned temporal dependencies are gathered and

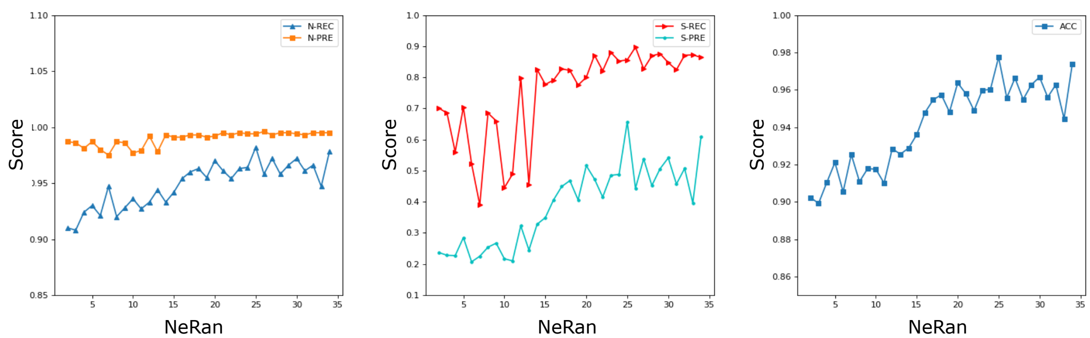
Table 6.6: Central-towards LSTM versus baselines

Method	ACC	N		S	
		REC	PRE	REC	PRE
central-towards LSTM	**97.7**	**98.2**	**99.4**	**85.6**	**65.7**
ordinary LSTM	88.5	91.0	96.9	27.2	10.6
bidirectional LSTM	84.1	86.1	97.0	32.9	8.6

outputted at the first or the last unit instead of the central unit, which leads to the outcome that the learned information is less relevant to the central target heartbeat.

Hyperparameters of CLSM

The proposed CLSM has two hyperparameters: $NeRan$ and $FeMaps$. The former denotes the neighborhood range of a heartbeat in a previous-RR-interval sequence. The latter decides the number of feature maps generated in the CLSM. As mentioned in Sec.6.4, the default values of $NeRan$ and $FeMaps$ are 25 and 32, respectively. Fig. 6.6 shows how model performance varies with different value of $NeRan$. The results are obtained on all the normal and S-type heartbeats in DS2. When $NeRan$ is smaller than 16, the model performances exhibit some fluctuations. The most representative one is the recall rate of the S-type heartbeats in Fig. 6.6 (a). The fluctuations imply that there is not enough information stored in the previous-RR-interval sequence for supporting the model to make decisions, which makes the model fall into a 'random guess' state for a portion of S-type heartbeats. In spite of the fluctuations, if we have a bird's-eye view of all three sub-figures, we can still see a growing trend of model performances with $NeRan$ increasing from 2 to 16. The model performances become stable and optimal when $NeRan$ is set to 20 or over. Therefore, we suggest that $NeRan > 20$ is necessary for CLSM to achieve near-optimal performance. The impact of $FeMaps$ on model performance is shown in Fig. 6.7, where $FeMaps$ is set to 8, 16, 32, 64, and 128 for comparison. The statistics curves exhibit a visible pattern: with the increase of $FeMaps$, the proposed CLSM presents a steady growth in performance. The optimal performance is achieved on $FeMaps = 32$. For $FeMaps > 32$, the model is kept on its optimal status.

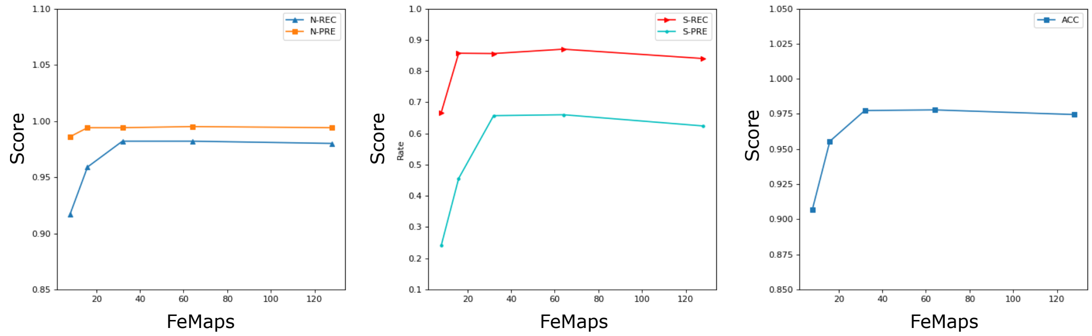


(a) recall and precision rate of class N (b) recall and precision rate of class S (c) overall classification accuracy

Figure 6.6: The impact of $NeRan$ on CLSM's performance. (a): recall rate (N-REC) and precision rate (N-PRE) of the normal heartbeats as a function of $NeRan$. (b): recall rate (S-REC) and precision rate (S-PRE) of the S-type heartbeats as a function of $NeRan$. (c): overall classification accuracy (ACC) as a function of $NeRan$.

Number of layers on CLSM

Next, we consider the impact of the number of central-towards LSTM layers on model performance. Fig. 6.8 (a) shows the overall classification accuracy, F1 scores for normal and S-type heartbeats as functions of the number of central-towards LSTM layers, in which results are obtained on all the normal and S-type heartbeats in DS2. The F1 score for the S-type heartbeats experience a slight decline when the number of layers increased from 2 to 6, whereas the F1 score for normal heartbeats and the overall classification accuracy are maintained at the same level with the increasing layers. The results indicate that the number of layers has a limited



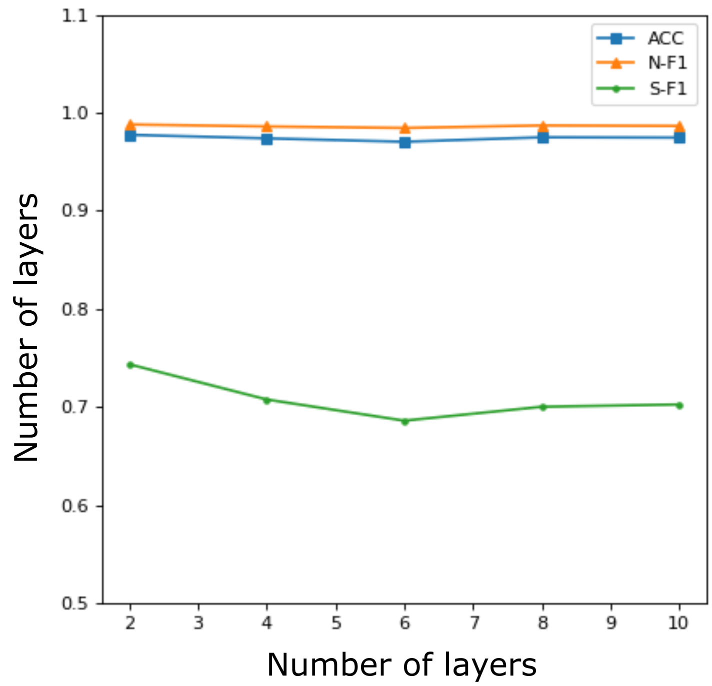
(a) recall and precision rate of class N (b) recall and precision rate of class S (c) overall classification accuracy

Figure 6.7: The impact of $FeMaps$ on CLSM's performance. (a): recall rate (N-REC) and precision rate (N-PRE) of the normal heartbeats as a function of $FeMaps$. (b): recall rate (S-REC) and precision rate (S-PRE) of the S-type heartbeats as a function of $FeMaps$. (c): overall classification accuracy (ACC) as a function of $FeMaps$.

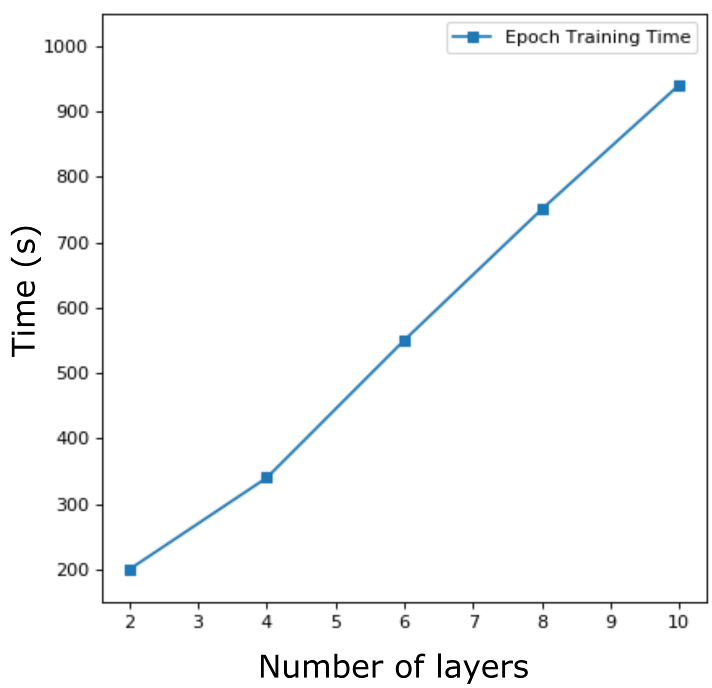
influence on CLSM's performance. On the other hand, however, with the increase of layer amount, the training time of an epoch steadily grows, as shown in Fig. 6.8 (b). Therefore, we suggest that the optimal layer number of the proposed CLSM is two.

The supportive nature of CLSM

Although CLSM is initially designed as the second-step structure in the proposed framework, it is a general and flexible binary classifier. We gather all normal and S-type heartbeats from the DS2, MIT-BIH-SUP and INCART databases for individual performance evaluations of the proposed CLSM which is trained on DS1. The results are shown in table 6.7. It can be seen that CLSM achieves decent performances on all three datasets, except for a relative low precision rate of the S-type heartbeats on the NS-INCART dataset where the normal heartbeats are nearly 80 times to the S-type heartbeats. In spite of the low precision rate, the CLSM makes a breakthrough on the S-type recall rate (94.7%) on the NS-INCART dataset. The results demonstrate the effectiveness of CLSM in identifying



(a) overall accuracy, F1 scores for class N and S



(b) Training time for an epoch

Figure 6.8: The impact of number of layers on CLSM's performance. (a): overall classification accuracy (ACC) and F1 scores for the normal (N-F1) S-type heartbeats (S-F1) as a function of the number of layers. (b): epoch training time of the proposed CLSM as a function of the number of layers.

Table 6.7: Individual performance evaluations of CLSM

Database	ACC	N		S	
		REC	PRE	REC	PRE
NS-DS2[1]	97.7	98.2	99.4	85.6	65.7
NS-MIT-BIH-SUP[2]	94.9	96.0	98.6	81.3	60.4
NS-INCART[3]	94.7	94.7	99.9	94.7	18.6

[1] A dataset consisting of all normal and S-type heartbeats of the MIT-BIH-AR database.
[2] A dataset consisting of all normal and S-type heartbeats of the MIT-BIH-SUP database.
[3] A dataset consisting of all normal and S-type heartbeats of the INCART database.

the S-type heartbeats from the normal ones. For those works suffering from the confusion of the S-type and the normal heartbeats, CLSM can be easily integrated as a complement without changing their original structures. This is why we define CLSM as a supportive model.

6.6 Conclusion

This work presents a two-step DNN-based classification framework to identify arrhythmia-related heartbeats from ECG recordings. The framework consists of a deep dual-channel convolutional neural network (DDCNN) and a central-towards LSTM supportive model (CLSM). In step-1, DDCNN incorporates both temporal and frequent patterns to identify the V, F and Q-type heartbeats. In step-2, CLSM distinguishes S-type heartbeats from normal ones by taking advantage of the central-towards LSTMs to learn and abstract hidden temporal information of each heartbeat. The experimental results obtained on three real-world databases show that the proposed framework has the potential to make a substantial clinical impact.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 Conclusion

Heart arrhythmia is a severe heart problem. Automated heartbeat classification provides a cost-effective screening for heart arrhythmia and allows at-risk patients to receive timely treatments, which is a highly demanded but challenging task. In this thesis, the author tries to tackle the automated heartbeat classification problem. The main contributions of this thesis are to mathematically represent the problem, to analyze the challenges, and to propose methods to solve the problem.

There are four practical heartbeat classification models proposed in this work. The first model is named D-ECG, which introduces the *dynamic ensemble selection* techniques and designs a result regulator to improve the detection performance of disease heartbeats. However, due to the dynamic nature, it is difficult for D-ECG to provide a real-time response to the streaming ECG signals. To tackle the heartbeat classification problem in real-time scenarios, the second model, a pyramid-like model, is proposed. This model separates the classification of normal and supraventricular ectopic beats from the overall heartbeat classification, and customizes an algorithm to take advantage of the neighbor-related information to assist identification of supraventricular ectopic bests. Compared to D-ECG, the pyramid-like model can provide more timely response to an unknown heartbeat while maintaining a good overall classification performance.

Since recent advances in heartbeat classification are brought by deep neural networks, the author examines these advances and proposes a DNN-based solution named Multi-channels Convolution Neural Network (MCHCNN) to solve the problems of current deep-learning based heartbeat classification models. As an improvement to other networks, MCHCNN considers heart rhythm to assist identification of disease heartbeats. Moreover, MCHCNN uses different sizes of convolution filters in parallel to capture temporal and frequency patterns from ECG signals. However, there is still a long way before MCHCNN can make practical impacts because its detection performance of S-type heartbeats is still relatively low. In fact, not just for MCHCNN, the identification of S-type heartbeats is one of the most problematic tasks for majority existing methods. To tackle this problem, the author investigates the potential causes and proposes an advanced two-step DNN-based classification framework. Specifically, in the first step, a deep dual-channel convolutional neural network (DDCNN) is proposed to classify all heartbeat classes, except for the normal and S-type heartbeats. In the second stage, a central-towards LSTM supportive model (CLSM) is specially designed to distinguish S-type heartbeats from the normal ones. By processing heart rhythms in central-towards directions, CLSM learns and abstracts hidden temporal information between a heartbeat and its neighbors to reveal the deep differences between the two heartbeat types.

The author has conducted extensive experiments to provide a comprehensive evaluation for each proposed model. The results prove that this thesis brings practical ideas and solutions to the automated heartbeat classification problem.

7.2 Future Work

Future work will focus on performing clinical deployment tests for the proposed methods, which includes generalization performance test and efficiency test. Based on the test results, further improvements and upgrades will be made to the proposed methods. Specifically, for D-ECG, the pyramid-like model, MCHCNN, and DHCAF which are designed for IoT applications, the author aims to improve these methods to provide a timely response, whereas for the DNN-based methods, the author will pay more attentions to improve the classification performance. Besides, as mentioned above, since the identification of S-type heartbeats is one of the most problematic tasks for majority existing methods, the author will continue to improve the proposed CLSM, making it easier to be incorporated into exiting methods to assist the identification of disease heartbeats.

Additionally, the information contained in ECG are not just useful in identification of heart disease. In the next study, the author plans to explore the relationship between ECG and sleep. The author aims to identify the sleep stages from ECG signals. Traditionally, this is achieved by analyzing the electroencephalography (EEG) data. However, as compared to EEG data, the acquisition of ECG data is easier. Therefore, building a connection between ECG and sleep stages will allow broader applications of a sleep stages classification model.

BIBLIOGRAPHY

[1] Jemal H Abawajy, Andrei V Kelarev, and M Chowdhury. Multistage approach for clustering and classification of ecg data. *Computer methods and programs in biomedicine*, 112(3):720–730, 2013.

[2] U Rajendra Acharya, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Muhammad Adam. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences*, 415:190–198, 2017.

[3] U Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, and Ru San Tan. A deep convolutional neural network model to classify heartbeats. *Computers in biology and medicine*, 89:389–396, 2017.

[4] Rashid Ghorbani Afkhami, Ghanbar Azarnia, and Mohammad Ali Tinati. Cardiac arrhythmia classification using statistical and mixture modeling features of ecg signals. *Pattern Recognition Letters*, 70:45–51, 2016.

[5] Foteini Agrafioti, Dimitris Hatzinakos, and Adam K Anderson. Ecg pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3(1):102–115, 2011.

[6] Roberto Alejo, José Martínez Sotoca, Rosa Maria Valdovinos, and P Toribio. Edited nearest neighbor rule for improving neural networks classifications. In *International Symposium on Neural Networks*, pages 303–310. Springer, 2010.

[7] Felipe Alonso-Atienza, Eduardo Morgado, Lorena Fernandez-Martinez, Arcadi García-Alberola, and José Luis Rojo-Alvarez. Detection of life-threatening arrhythmias using feature selection and support vector machines. *IEEE Transactions on Biomedical Engineering*, 61(3):832–840, 2014.

[8] ANSI/AAMI. Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms. *Association for the Advancement of Medical Instrumentation -AAMI*, ISO EC57, 1998-2008.

[9] Muhammad Arif et al. Robust electrocardiogram (ecg) beat classification using discrete wavelet transform. *Physiological measurement*, 29(5):555, 2008.

[10] Babak Mohammadzadeh Asl, Seyed Kamaledin Setarehdan, and Maryam Mohebbi. Support vector machine-based arrhythmia classification using

reduced features of heart rate variability signal. *Artificial intelligence in medicine*, 44(1):51–64, 2008.

[11] Brandon Ballinger, Johnson Hsieh, Avesh Singh, Nimit Sohoni, Jack Wang, Geoffrey H Tison, Gregory M Marcus, Jose M Sanchez, Carol Maguire, Jeffrey E Olgin, et al. Deepheart: semi-supervised sequence learning for cardiovascular risk prediction. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[12] Yakoub Bazi, Naif Alajlan, Haikel AlHichri, and Salim Malek. Domain adaptation methods for ecg classification. In *2013 international conference on computer medical applications (ICCMA)*, pages 1–4. IEEE, 2013.

[13] Iva Bogdanova, Francisco Rincón, and David Atienza. A multi-lead ecg classification based on random projection features. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 625–628. IEEE, 2012.

[14] G Bortolan, II Christov, and W Pedrycz. Hyperbox classifiers for ecg beat analysis. In *2007 Computers in Cardiology*, pages 145–148. IEEE, 2007.

[15] Alceu S Britto Jr, Robert Sabourin, and Luiz ES Oliveira. Dynamic selection of classifiers—a comprehensive review. *Pattern Recognition*, 47(11):3665–3680, 2014.

[16] Paulo R Cavalin, Robert Sabourin, and Ching Y Suen. Dynamic selection approaches for multiple classifier systems. *Neural Computing and Applications*, 22(3-4):673–688, 2013.

[17] Rahime Ceylan and Yüksel Özbay. Comparison of fcm, pca and wt techniques for classification ecg arrhythmias using artificial neural network. *Expert Systems with Applications*, 33(2):286–295, 2007.

[18] Rahime Ceylan, Yüksel Özbay, and Bekir Karlik. A novel approach for classification of ecg arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications*, 36(3):6721–6726, 2009.

[19] Sucheta Chauhan and Lovekesh Vig. Anomaly detection in ecg time signals via deep long short-term memory networks. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–7. IEEE, 2015.

[20] MPS Chawla. A comparative analysis of principal component and independent component techniques for electrocardiograms. *Neural Computing and Applications*, 18(6):539–556, 2009.

[21] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[22] Lili Chen, Xi Zhang, and Changyue Song. An automatic screening approach for obstructive sleep apnea diagnosis based on single-lead electrocardiogram. *IEEE Transactions on Automation Science and Engineering*, 12(1):106–115, 2014.

[23] Shanshan Chen, Wei Hua, Zhi Li, Jian Li, and Xingjiao Gao. Heartbeat classification using projected and dynamic features of ecg signal. *Biomedical Signal Processing and Control*, 31:165–173, 2017.

[24] Szi-Wen Chen, Peter M Clarkson, and Qi Fan. A robust sequential detection algorithm for cardiac arrhythmia classification. *IEEE Transactions on Biomedical Engineering*, 43(11):1120–1124, 1996.

[25] Ping Cheng and Xiaodai Dong. Life-threatening ventricular arrhythmia detection with personalized features. *IEEE Access*, 5:14195–14203, 2017.

[26] Ivalyo Christov and G Bortolan. Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks. *Physiological Measurement*, 25(5):1281, 2004.

[27] Kuang Chua Chua, Vinod Chandran, U Rajendra Acharya, and Choo Min Lim. Application of higher order statistics/spectra in biomedical signals—a review. *Medical engineering & physics*, 32(7):679–689, 2010.

[28] Teck Wee Chua and Woei Wan Tan. Non-singleton genetic fuzzy logic system for arrhythmias classification. *Engineering Applications of Artificial Intelligence*, 24(2):251–259, 2011.

[29] Kwok Tai Chui, Kim Fung Tsang, Hao Ran Chi, Bingo Wing Kuen Ling, and Chung Kit Wu. An accurate ecg-based transportation safety drowsiness detection scheme. *IEEE Transactions on Industrial Informatics*, 12(4):1438–1452, 2016.

[30] Douglas A Coast, Richard M Stern, Gerald G Cano, and Stanley A Briller.

An approach to cardiac arrhythmia analysis using hidden markov models. *IEEE Transactions on biomedical Engineering*, 37(9):826–836, 1990.

[31] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[32] Rafael MO Cruz, Robert Sabourin, and George DC Cavalcanti. Meta-des. oracle: Meta-learning and feature selection for dynamic ensemble selection. *Information fusion*, 38:84–103, 2017.

[33] Rafael MO Cruz, Robert Sabourin, and George DC Cavalcanti. Dynamic classifier selection: Recent advances and perspectives. *Information Fusion*, 41:195–216, 2018.

[34] Hugo Plácido da Silva, Carlos Carreiras, André Lourenço, Ana Fred, Rui César das Neves, and Rui Ferreira. Off-the-person electrocardiography: performance assessment and clinical correlation. *Health and Technology*, 4(4):309–318, 2015.

[35] Abdelhamid Daamouche, Latifa Hamami, Naif Alajlan, and Farid Melgani. A wavelet optimization approach for ecg signal classification. *Biomedical Signal Processing and Control*, 7(4):342–349, 2012.

[36] Ingrid Daubechies. *Ten lectures on wavelets*, volume 61. Siam, 1992.

[37] Victor Hugo C de Albuquerque, Thiago M Nunes, Danillo R Pereira, Eduardo José da S Luz, David Menotti, João P Papa, and João Manuel RS Tavares. Robust automated cardiac arrhythmia detection in ecg beat signals. *Neural Computing and Applications*, 29(3):679–693, 2018.

[38] Philip De Chazal, Maria O'Dwyer, and Richard B Reilly. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE transactions on biomedical engineering*, 51(7):1196–1206, 2004.

[39] Gael De Lannoy, Damien François, Jean Delbeke, and Michel Verleysen. Weighted svms and feature relevance assessment in supervised heart beat classification. In *International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 212–223. Springer, 2010.

[40] Gaël De Lannoy, Damien François, Jean Delbeke, and Michel Verleysen. Weighted conditional random fields for supervised interpatient heartbeat

classification. *IEEE Transactions on Biomedical Engineering*, 59(1):241–247, 2012.

[41] Selim Dilmac and Mehmet Korurek. Ecg heart beat classification method based on modified abc algorithm. *Applied Soft Computing*, 36:641–655, 2015.

[42] Sofia-Maria Dima, Christos Panagiotou, Evangelos B Mazomenos, James A Rosengarten, Koushik Maharatna, John V Gialelis, Nick Curzen, and John Morgan. On the detection of myocadial scar based on ecg/vcg analysis. *IEEE Transactions on Biomedical Engineering*, 60(12):3399–3409, 2013.

[43] Cüneyt Dirican. The impacts of robotics, artificial intelligence on business and economics. *Procedia-Social and Behavioral Sciences*, 195:564–573, 2015.

[44] Zümray Dokur and Tamer Ölmez. Ecg beat classification by a novel hybrid neural network. *Computer methods and programs in biomedicine*, 66(2-3):167–181, 2001.

[45] Xunde Dong, Cong Wang, and Wenjie Si. Ecg beat classification via deterministic learning. *Neurocomputing*, 240:1–12, 2017.

[46] Gauthier Doquire, Gael De Lannoy, Damien François, and Michel Verleysen. Feature selection for interpatient supervised heart beat classification. *Computational intelligence and neuroscience*, 2011:1, 2011.

[47] Eulanda M Dos Santos, Robert Sabourin, and Patrick Maupin. A dynamic overproduce-and-choose strategy for the selection of classifier ensembles. *Pattern Recognition*, 41(10):2993–3009, 2008.

[48] Saibal Dutta, Amitava Chatterjee, and Sugata Munshi. Correlation technique and least square support vector machine combine for frequency domain based ecg beat classification. *Medical engineering & physics*, 32(10):1161–1169, 2010.

[49] Reza Ebrahimpour, Naser Sadeghnejad, Atena Sajedin, and Nima Mohammadi. Electrocardiogram beat classification via coupled boosting by filtering and preloaded mixture of experts. *Neural Computing and Applications*, 23(3-4):1169–1178, 2013.

[50] Ataollah Ebrahimzadeh, B Shakiba, and Ali Khazaee. Detection of electrocardiogram signals using an efficient method. *Applied Soft Computing*, 22:108–117, 2014.

[51] Fatin A Elhaj, Naomie Salim, Arief R Harris, Tan Tian Swee, and Taqwa Ahmed. Arrhythmia recognition and classification using combined linear and nonlinear features of ecg signals. *Computer methods and programs in biomedicine*, 127:52–63, 2016.

[52] Mehmet Engin. Ecg beat classification using neuro-fuzzy network. *Pattern Recognition Letters*, 25(15):1715–1722, 2004.

[53] Miguel Angel Escalona-Morán, Miguel C Soriano, Ingo Fischer, and Claudio R Mirasso. Electrocardiogram classification using reservoir computing with logistic regression. *IEEE journal of biomedical and health informatics*, 19(3):892–898, 2015.

[54] Themis P Exarchos, Markos G Tsipouras, Costas P Exarchos, Costas Papaloukas, Dimitrios I Fotiadis, and Lampros K Michalis. A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree. *Artificial Intelligence in medicine*, 40(3):187–200, 2007.

[55] TP Exarchos, MG Tsipouras, D Nanou, C Bazios, Y Antoniou, and DI Fotiadis. A platform for wide scale integration and visual representation of medical intelligence in cardiology: the decision support framework. In *Computers in Cardiology, 2005*, pages 167–170. IEEE, 2005.

[56] Adam Gacek. Data structure-guided development of electrocardiographic signal characterization and classification. *Artificial intelligence in medicine*, 59(3):197–204, 2013.

[57] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[58] Pedro R Gomes, Filomena O Soares, JH Correia, and CS Lima. Ecg data-acquisition and classification system by using wavelet-domain hidden markov models. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 4670–4673. IEEE, 2010.

[59] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[60] İnan Güler and Elif Derya Übeylı. Ecg beat classifier designed by combined neural network model. *Pattern recognition*, 38(2):199–208, 2005.

[61] Salih Güneş, Kemal Polat, and Şebnem Yosunkaya. Multi-class f-score feature selection approach to classification of obstructive sleep apnea syndrome. *Expert systems with applications*, 37(2):998–1004, 2010.

[62] Sandeep Gutta and Qi Cheng. Joint feature extraction and classifier design for ecg-based biometric recognition. *IEEE journal of biomedical and health informatics*, 20(2):460–468, 2015.

[63] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[64] Fredric M Ham and Soowhan Han. Classification of cardiac arrhythmias using fuzzy artmap. *IEEE Transactions on Biomedical Engineering*, 43(4):425–429, 1996.

[65] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65, 2019.

[66] Hassan Hamsa Haseena, Abraham T Mathew, and Joseph K Paul. Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. *Journal of Medical Systems*, 35(2):179–188, 2011.

[67] Ahnaf Rashik Hassan and Md Aynal Haque. An expert system for automated identification of obstructive sleep apnea from single-lead ecg using random under sampling boosting. *Neurocomputing*, 235:122–130, 2017.

[68] Jinyuan He, Le Sun, Jia Rong, Hua Wang, and Yanchun Zhang. A pyramid-like model for heartbeat classification from ecg recordings. *PloS one*, 13(11):e0206593, 2018.

[69] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[70] Robert Hecht-Nielsen. Theory of the backpropagation neural network. In *Neural networks for perception*, pages 65–93. Elsevier, 1992.

[71] Frank L Hitchcock. The distribution of a product from several sources to numerous localities. *Journal of mathematics and physics*, 20(1-4):224–230, 1941.

[72] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.

[73] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[74] Huifang Huang, Jie Liu, Qiang Zhu, Ruiping Wang, and Guangshu Hu. A new hierarchical method for inter-patient heartbeat classification using random projections and rr intervals. *Biomedical engineering online*, 13(1):90, 2014.

[75] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.

[76] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[77] Mehrdad Javadi, Seyed Ali Asghar Abbaszadeh Arani, Atena Sajedin, and Reza Ebrahimpour. Classification of ecg arrhythmia by a modular neural network based on mixture of experts and negatively correlated learning. *Biomedical Signal Processing and Control*, 8(3):289–296, 2013.

[78] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4):230–243, 2017.

[79] Wei Jiang and Seong G Kong. Block-based neural networks for personalized ecg signal classification. *IEEE Transactions on Neural Networks*, 18(6):1750–1761, 2007.

[80] Md Enamul Kabir, Hua Wang, and Elisa Bertino. A role-involved purpose-based access control model. *Information Systems Frontiers*, 14(3):809–822, 2012.

[81] S Karimifard, A Ahmadian, Mohammad Khoshnevisan, and MS Nambakhsh.

Morphological heart arrhythmia detection using hermitian basis functions and knn classifier. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1367–1370. IEEE, 2006.

[82] Walter Karlen, Claudio Mattiussi, and Dario Floreano. Sleep and wake classification with ecg and respiratory effort signals. *IEEE Transactions on Biomedical Circuits and Systems*, 3(2):71–78, 2009.

[83] JS Karthika, Jan Mary Thomas, and Jubilant J Kizhakkethottam. Detection of life-threatening arrhythmias using temporal, spectral and wavelet features. In *Computational Intelligence and Computing Research (ICCIC), 2015 IEEE International Conference on*, pages 1–4. IEEE, 2015.

[84] Aya F Khalaf, Mohamed I Owis, and Inas A Yassine. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. *Expert Systems with Applications*, 42(21):8361–8368, 2015.

[85] Ahsan H Khandoker, Marimuthu Palaniswami, and Chandan K Karmakar. Support vector machines for automated recognition of obstructive sleep apnea syndrome from ecg recordings. *IEEE transactions on information technology in biomedicine*, 13(1):37–48, 2008.

[86] Ali Khazaee. Heart beat classification using particle swarm optimization. *International Journal of Intelligent Systems and Applications*, 5(6):25, 2013.

[87] Ali Khazaee and Ataollah Ebrahimzadeh. Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. *Biomedical Signal Processing and Control*, 5(4):252–263, 2010.

[88] Hamid Khorrami and Majid Moavenian. A comparative study of dwt, cwt and dct transformations in ecg arrhythmias classification. *Expert systems with Applications*, 37(8):5751–5757, 2010.

[89] Rami N Khushaba, Sarath Kodagoda, Sara Lal, and Gamini Dissanayake. Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE Transactions on Biomedical Engineering*, 58(1):121–131, 2010.

[90] Rami N Khushaba, Sarath Kodagoda, Sara Lal, and Gamini Dissanayake. Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition. *Fuzzy Sets and Systems*, 221:90–111, 2013.

[91] Jinkwon Kim, Hang Sik Shin, Kwangsoo Shin, and Myoungho Lee. Robust algorithm for arrhythmia classification in ecg using extreme learning machine. *Biomedical engineering online*, 8(1):31, 2009.

[92] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[93] Serkan Kiranyaz, Turker Ince, and Moncef Gabbouj. Real-time patient-specific ecg classification by 1-d convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 63(3):664–675, 2015.

[94] Serkan Kiranyaz, Turker Ince, and Moncef Gabbouj. Real-time patient-specific ecg classification by 1-d convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 63(3):664–675, 2016.

[95] Albert HR Ko, Robert Sabourin, and Alceu Souza Britto Jr. From dynamic classifier selection to dynamic ensemble selection. *Pattern Recognition*, 41(5):1718–1731, 2008.

[96] Mehmet Korürek and Berat Doğan. Ecg beat classification using particle swarm optimization and radial basis function neural network. *Expert systems with Applications*, 37(12):7563–7569, 2010.

[97] Mehmet Korürek and Ali Nizam. Clustering mit–bih arrhythmias with ant colony optimization using time domain and pca compressed wavelet coefficients. *Digital Signal Processing*, 20(4):1050–1060, 2010.

[98] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[99] R Ganesh Kumar and YS Kumaraswamy. Investigation and classification of ecg beat using input output additional weighted feed forward neural network. In *2013 International Conference on Signal Processing, Image Processing & Pattern Recognition*, pages 200–205. IEEE, 2013.

[100] Yakup Kutlu and Damla Kuntalp. A multi-stage automatic arrhythmia recognition and classification system. *Computers in Biology and Medicine*, 41(1):37–45, 2011.

[101] Yakup Kutlu and Damla Kuntalp. Feature extraction for ecg heartbeats

using higher order statistics of wpd coefficients. *Computer methods and programs in biomedicine*, 105(3):257–267, 2012.

[102] Martin Lagerholm, Carsten Peterson, Guido Braccini, Lars Edenbrandt, and Leif Sornmo. Clustering ecg complexes using hermite functions and self-organizing maps. *IEEE Transactions on Biomedical Engineering*, 47(7):838–848, 2000.

[103] Pablo Laguna, Raimon Jané, and Pere Caminal. Automatic detection of wave boundaries in multilead ecg signals: Validation with the cse database. *Computers and biomedical research*, 27(1):45–60, 1994.

[104] Antonio Lanata, Gaetano Valenza, C Mancuso, and Enzo Pasquale Scilingo. Robust multiple cardiac arrhythmia detection through bispectrum analysis. *Expert Systems with Applications*, 38(6):6798–6804, 2011.

[105] Hongqiang Li, Danyang Yuan, Xiangdong Ma, Dianyin Cui, and Lu Cao. Genetic algorithm for the optimization of features and neural networks in ecg signals classification. *Scientific reports*, 7:41011, 2017.

[106] Min Li, Xiaoxun Sun, Hua Wang, Yanchun Zhang, and Ji Zhang. Privacy-aware access control with trust management in web service. *World Wide Web*, 14(4):407–430, 2011.

[107] Ming Li, Viktor Rozgica, Gautam Thatte, Sangwon Lee, Adar Emken, Murali Annavaram, Urbashi Mitra, Donna Spruijt-Metz, and Shrikanth Narayanan. Multimodal physical activity recognition by fusing temporal and cepstral information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(4):369–380, 2010.

[108] Pengfei Li, Yu Wang, Jiangchun He, Lihua Wang, Yu Tian, Tian-shu Zhou, Tianchang Li, and Jing-song Li. High-performance personalized heartbeat classification model for long-term ecg signal. *IEEE Transactions on Biomedical Engineering*, 64(1):78–86, 2016.

[109] Qiao Li, Cadathur Rajagopalan, and Gari D Clifford. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Transactions on Biomedical Engineering*, 61(6):1607–1613, 2013.

[110] Konstantinos G Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. Machine learning in agriculture: A review. *Sensors*, 18(8):2674, 2018.

[111] Chen Lin, Wenqiang Chen, Cheng Qiu, Yunfeng Wu, Sridhar Krishnan, and Quan Zou. Libd3c: ensemble classifiers with a clustering and dynamic selection strategy. *Neurocomputing*, 123:424–435, 2014.

[112] Chia-Hung Lin, Yi-Chun Du, and Tainsong Chen. Adaptive wavelet network for multiple cardiac arrhythmias recognition. *Expert Systems with Applications*, 34(4):2601–2611, 2008.

[113] Fan Liu, Xingshe Zhou, Jinli Cao, Zhu Wang, Hua Wang, and Yanchun Zhang. Arrhythmias classification by integrating stacked bidirectional lstm and two-dimensional cnn. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 136–149. Springer, 2019.

[114] Fan Liu, Xingshe Zhou, Jinli Cao, Zhu Wang, Hua Wang, and Yanchun Zhang. A lstm and cnn based assemble neural network framework for arrhythmias classification. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1303–1307. IEEE, 2019.

[115] Mariano Llamedo and Juan Pablo Martínez. Heartbeat classification using feature selection driven by database generalization criteria. *IEEE Transactions on Biomedical Engineering*, 58(3):616–625, 2011.

[116] Zhengji Long, Guangyuan Liu, and Xuewu Dai. Extracting emotional features from ecg by using wavelet transform. In *2010 International Conference on Biomedical Engineering and Computer Science*, pages 1–4. IEEE, 2010.

[117] Rose Luckin, Wayne Holmes, Mark Griffiths, and Laurie B Forcier. Intelligence unleashed: An argument for ai in education. 2016.

[118] Eduardo José Da S Luz, Thiago M Nunes, Victor Hugo C De Albuquerque, Joao P Papa, and David Menotti. Ecg arrhythmia classification based on optimum-path forest. *Expert Systems with Applications*, 40(9):3561–3573, 2013.

[119] Eduardo José da S Luz, William Robson Schwartz, Guillermo Cámara-Chávez, and David Menotti. Ecg-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine*, 127:144–164, 2016.

[120] Jiangang Ma, Le Sun, Hua Wang, Yanchun Zhang, and Uwe Aickelin. Supervised anomaly detection in uncertain pseudoperiodic data streams. *ACM Transactions on Internet Technology (TOIT)*, 16(1):4, 2016.

[121] Venkatachalam Mahesh, Arumugam Kandaswamy, Chandran Vimal, and Balakrishnan Sathish. Ecg arrhythmia classification based on logistic model tree. *Journal of Biomedical Science and Engineering*, 2(06):405, 2009.

[122] M Sabarimalai Manikandan and Samarendra Dandapat. Wavelet-based electrocardiogram signal compression methods and their performances: a prospective review. *Biomedical Signal Processing and Control*, 14:73–107, 2014.

[123] Yishay Mansour. Pessimistic decision tree pruning based on tree size. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, pages 195–201. Citeseer, 1997.

[124] Tanis Mar, Sebastian Zaunseder, Juan Pablo Martínez, Mariano Llamedo, and Rüdiger Poll. Optimization of ecg classification by means of feature selection. *IEEE transactions on Biomedical Engineering*, 58(8):2168–2177, 2011.

[125] Roshan J Martis, U Rajendra Acharya, Ajoy K Ray, and Chandan Chakraborty. Application of higher order cumulants to ecg signals for the cardiac health diagnosis. In *Engineering in medicine and biology society, embc, 2011 annual international conference of the ieee*, pages 1697–1700. IEEE, 2011.

[126] Roshan Joy Martis, U Rajendra Acharya, Choo Min Lim, and Jasjit S Suri. Characterization of ecg beats from cardiac arrhythmia using discrete cosine transform in pca framework. *Knowledge-Based Systems*, 45:76–82, 2013.

[127] Roshan Joy Martis, U Rajendra Acharya, KM Mandana, Ajoy Kumar Ray, and Chandan Chakraborty. Cardiac decision making using higher order spectra. *Biomedical Signal Processing and Control*, 8(2):193–203, 2013.

[128] Roshan Joy Martis, U Rajendra Acharya, and Lim Choo Min. Ecg beat classification using pca, lda, ica and discrete wavelet transform. *Biomedical Signal Processing and Control*, 8(5):437–448, 2013.

[129] Saeed Masoudnia and Reza Ebrahimpour. Mixture of experts: a literature survey. *Artificial Intelligence Review*, 42(2):275–293, 2014.

[130] EB Mazomenos, Taihai Chen, Amit Acharyya, Arnab Bhattacharya, James Rosengarten, and Koushik Maharatna. A time-domain morphology and gradient based algorithm for ecg feature extraction. In *2012 IEEE International Conference on Industrial Technology*, pages 117–122. IEEE, 2012.

[131] Yeong Pong Meau, Fatimah Ibrahim, Selvanathan AL Narainasamy, and Razali Omar. Intelligent classification of electrocardiogram (ecg) signal using extended kalman filter (ekf) based neuro fuzzy system. *Computer Methods and Programs in Biomedicine*, 82(2):157–168, 2006.

[132] Hila Mehr, H Ash, and D Fellow. Artificial intelligence for citizen services and government. *Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch., no. August*, pages 1–12, 2017.

[133] Rahul Mehra. Global public health problem of sudden cardiac death. *Journal of electrocardiology*, 40(6):S118–S122, 2007.

[134] MO Mendez, J Corthout, Sabine Van Huffel, Matteo Matteucci, T Penzel, Sergio Cerutti, and AM Bianchi. Automatic screening of obstructive sleep apnea from the ecg based on empirical mode decomposition and wavelet analysis. *Physiological measurement*, 31(3):273, 2010.

[135] Amit K Mishra and Shantanu Raghav. Local fractal dimension based ecg arrhythmia classification. *Biomedical Signal Processing and Control*, 5(2):114–123, 2010.

[136] Majid Moavenian and Hamid Khorrami. A qualitative comparison of artificial neural networks and support vector machines in ecg arrhythmias classification. *Expert Systems with Applications*, 37(4):3088–3093, 2010.

[137] George B Moody and Roger G Mark. The impact of the mit-bih arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3):45–50, 2001.

[138] Joseph J Oresko, Zhanpeng Jin, Jun Cheng, Shimeng Huang, Yuwen Sun, Heather Duschl, and Allen C Cheng. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Transactions on Information Technology in Biomedicine*, 14(3):734–740, 2010.

[139] S Osowski, T Markiewicz, and L Tran Hoai. Recognition and classification system of arrhythmia using ensemble of neural networks. *Measurement*, 41(6):610–617, 2008.

[140] Stanislaw Osowski, Linh Tran Hoai, and Tomasz Markiewicz. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE transactions on biomedical engineering*, 51(4):582–589, 2004.

[141] Stanislaw Osowski and Tran Hoai Linh. Ecg beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering*, 48(11):1265–1271, 2001.

[142] MI Owis, A-BM Youssef, and YM Kadah. Characterisation of electrocardiogram signals based on blind source separation. *Medical and Biological Engineering and Computing*, 40(5):557–564, 2002.

[143] Mohamed I Owis, Ahmed H Abou-Zied, A-BM Youssef, and Yasser M Kadah. Study of features based on nonlinear dynamical modeling in ecg arrhythmia detection and classification. *IEEE transactions on Biomedical Engineering*, 49(7):733–736, 2002.

[144] Yüksel Özbay, Rahime Ceylan, and Bekir Karlik. A fuzzy clustering neural network architecture for classification of ecg arrhythmias. *Computers in Biology and Medicine*, 36(4):376–388, 2006.

[145] Yüksel Özbay and Gülay Tezel. A new method for classification of ecg arrhythmias using neural network with adaptive activation function. *Digital Signal Processing*, 20(4):1040–1049, 2010.

[146] Jiapu Pan and Willis J Tompkins. A real-time qrs detection algorithm. *IEEE Trans. Biomed. Eng*, 32(3):230–236, 1985.

[147] KS Park, BH Cho, DH Lee, SH Song, JS Lee, YJ Chee, IY Kim, and SI Kim. Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function. In *Computers in Cardiology, 2008*, pages 229–232. IEEE, 2008.

[148] Edoardo Pasolli and Farid Melgani. Active learning methods for electrocardiographic signal classification. *IEEE Transactions on Information Technology in Biomedicine*, 14(6):1405–1416, 2010.

[149] Pavel Pudil, Jana Novovičová, and Josef Kittler. Floating search methods in feature selection. *Pattern recognition letters*, 15(11):1119–1125, 1994.

[150] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

[151] Sandeep Raj and Kailash Chandra Ray. Ecg signal analysis using dct-based dost and pso optimized svm. *IEEE Transactions on instrumentation and measurement*, 66(3):470–478, 2017.

[152] Jimena Rodriguez, Alfredo Goni, and Arantza Illarramendi. Real-time classification of ecgs on a pda. *IEEE Transactions on Information Technology in Biomedicine*, 9(1):23–34, 2005.

[153] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. A metric for distributions with applications to image databases. In *Computer Vision, 1998. Sixth International Conference on*, pages 59–66. IEEE, 1998.

[154] Santanu Sahoo, Bhupen Kanungo, Suresh Behera, and Sukanta Sabut. Multiresolution wavelet transform based feature extraction and ecg classification to detect cardiac abnormalities. *Measurement*, 108:55–66, 2017.

[155] Mohammad Sarfraz, Ateeq Ahmed Khan, and Francis F Li. Using independent component analysis to obtain feature space for reliable ecg arrhythmia classification. In *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*, pages 62–67. IEEE, 2014.

[156] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.

[157] Ali Sellami and Heasoo Hwang. A robust deep convolutional neural network with batch-weighted loss for heartbeat classification. *Expert Systems with Applications*, 122:75–84, 2019.

[158] Hemant Sharma and KK Sharma. An algorithm for sleep apnea detection from single-lead ecg using hermite basis functions. *Computers in biology and medicine*, 77:116–124, 2016.

[159] Mark J Shensa. The discrete wavelet transform: wedding the a trous and mallat algorithms. *IEEE Transactions on signal processing*, 40(10):2464–2482, 1992.

[160] Basilio Sierra, Elena Lazkano, Itziar Irigoien, Ekaitz Jauregi, and Iñigo Mendialdua. K nearest neighbor equality: giving equal chance to all existing classes. *Information Sciences*, 181(23):5158–5168, 2011.

[161] Tom Simonite. For superpowers, artificial intelligence fuels new global arms race. *Wired Magazine*, page 9, 2017.

[162] Rodrigo GF Soares, Alixandre Santana, Anne MP Canuto, and Marcílio Carlos Pereira de Souto. Using accuracy and diversity to select classifiers to

build ensembles. In *Neural Networks, 2006. IJCNN'06. International Joint Conference on*, pages 1310–1316. IEEE, 2006.

[163] Changyue Song, Kaibo Liu, Xi Zhang, Lili Chen, and Xiaochen Xian. An obstructive sleep apnea detection approach using a discriminative hidden markov model from ecg signals. *IEEE Transactions on Biomedical Engineering*, 63(7):1532–1542, 2015.

[164] Mi Hye Song, Jeon Lee, Sung Pil Cho, Kyoung Joung Lee, and Sun Kook Yoo. Support vector machine based arrhythmia classification using reduced features. 2005.

[165] M Llamedo Soria and JP Martínez. Analysis of multidomain features for ecg classification. In *2009 36th Annual Computers in Cardiology Conference (CinC)*, pages 561–564. IEEE, 2009.

[166] Donald F Specht. Probabilistic neural networks. *Neural networks*, 3(1):109–118, 1990.

[167] Lili Sun, Hua Wang, Jianming Yong, and Guoxin Wu. Semantic access control for cloud computing based on e-healthcare. In *Computer Supported Cooperative Work in Design (CSCWD), 2012 IEEE 16th International Conference on*, pages 512–518. IEEE, 2012.

[168] Xiaoxun Sun, Min Li, Hua Wang, and Ashley Plank. An efficient hash-based algorithm for minimal k-anonymity. In *Proceedings of the thirty-first Australasian conference on Computer science-Volume 74*, pages 101–107. Australian Computer Society, Inc., 2008.

[169] Xiaoxun Sun, Hua Wang, Jiuyong Li, and Yanchun Zhang. Injecting purpose and trust into data anonymisation. *computers & security*, 30(5):332–345, 2011.

[170] Akara Supratak, Hao Dong, Chao Wu, and Yike Guo. Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(11):1998–2008, 2017.

[171] Araz Taeihagh and Hazel Si Min Lim. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport reviews*, 39(1):103–128, 2019.

[172] Manal M Tantawi, Kenneth Revett, Abdel-Badeeh Salem, and Mohamed F Tolba. A wavelet feature extraction method for electrocardiogram (ecg)-based biometric recognition. *Signal, Image and Video Processing*, 9(6):1271–1280, 2015.

[173] Vahid Tavakoli, Nima Sahba, and Nima Hajebi. A fast and accurate method for arrhythmia detection. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 1897–1900. IEEE, 2009.

[174] Tomas Teijeiro, Paulo Felix, Jesus Presedo, and Daniel Castro. Heartbeat classification using abstract features from the abductive interpretation of the ecg. *IEEE journal of biomedical and health informatics*, 2016.

[175] Manu Thomas, Manab Kr Das, and Samit Ari. Automatic ecg arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69(4):715–721, 2015.

[176] Markos G Tsipouras, Dirnitrois I Fotiadis, and D Sideris. Arrhythmia classification using the rr-interval duration signal. In *Computers in cardiology*, pages 485–488. IEEE, 2002.

[177] Elif Derya Übeyli. Adaptive neuro-fuzzy inference system for classification of ecg signals using lyapunov exponents. *Computer methods and programs in biomedicine*, 93(3):313–321, 2009.

[178] Elif Derya Übeyli. Recurrent neural networks employing lyapunov exponents for analysis of ecg signals. *Expert systems with applications*, 37(2):1192–1199, 2010.

[179] MH Vafaie, M Ataei, and Hamid R Koofigar. Heart diseases prediction based on ecg signals' classification using a genetic-fuzzy system and dynamical model of ecg signals. *Biomedical Signal Processing and Control*, 14:291–296, 2014.

[180] Pasupathy Vimalachandran, Hua Wang, Yanchun Zhang, Ben Heyward, and Yueai Zhao. Preserving patient-centred controls in electronic health record systems: A reliance-based model implication. In *Orange Technologies (ICOT), 2017 International Conference on*, pages 37–44. IEEE, 2017.

[181] Pasupathy Vimalachandran, Hua Wang, Yanchun Zhang, Guangping Zhuo, and Hongbo Kuang. Cryptographic access control in electronic health record

systems: A security implication. In *International Conference on Web Information Systems Engineering*, pages 540–549. Springer, 2017.

[182] Saeid Wahabi, Shahrzad Pouryayevali, Siddarth Hari, and Dimitrios Hatzinakos. On evaluating ecg biometric systems: session-dependence and body posture. *IEEE Transactions on Information Forensics and Security*, 9(11):2002–2013, 2014.

[183] Hua Wang, Jinli Cao, and Yanchun Zhang. A flexible payment scheme and its role-based access control. *IEEE Transactions on knowledge and Data Engineering*, 17(3):425–436, 2005.

[184] Hua Wang, Zonghua Zhang, and Tarek Taleb. Special issue on security and privacy of iot. *World Wide Web*, 21(1):1–6, 2018.

[185] Jeen-Shing Wang, Wei-Chun Chiang, Yu-Liang Hsu, and Ya-Ting C Yang. Ecg arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing*, 116:38–45, 2013.

[186] Ye Wang, Hu Li, Hua Wang, Bin Zhou, and Yanchun Zhang. Multi-window based ensemble learning for classification of imbalanced streaming data. In *International Conference on Web Information Systems Engineering*, pages 78–92. Springer, 2015.

[187] Cheng Wen, Teng-Chiao Lin, Kuang-Chiung Chang, and Chih-Hung Huang. Classification of ecg complexes using self-organizing cmac. *Measurement*, 42(3):399–407, 2009.

[188] Jenna Wiens and John V Guttag. Active learning applied to patient-adaptive heartbeat classification. In *Advances in neural information processing systems*, pages 2442–2450, 2010.

[189] Wikipedia. Electrocardiography, 2019.

[190] Wikipedia contributors. Heart arrhythmia, 2018.

[191] Dennis L Wilson. Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*, 3:408–421, 1972.

[192] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[193] Tomasz Woloszynski and Marek Kurzynski. A measure of competence based on randomized reference classifier for dynamic ensemble selection. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 4194–4197. IEEE, 2010.

[194] Tomasz Woloszynski and Marek Kurzynski. A probabilistic model of classifier competence for dynamic ensemble selection. *Pattern Recognition*, 44(10-11):2656–2668, 2011.

[195] Tomasz Woloszynski, Marek Kurzynski, Pawel Podsiadlo, and Gwidon W Stachowiak. A measure of competence based on random classification for dynamic ensemble selection. *Information Fusion*, 13(3):207–213, 2012.

[196] Jin Xiao, Ling Xie, Changzheng He, and Xiaoyi Jiang. Dynamic classifier ensemble model for customer classification with imbalanced class distribution. *Expert Systems with Applications*, 39(3):3668–3675, 2012.

[197] Zhe Yang, Qihao Zhou, Lei Lei, Kan Zheng, and Wei Xiang. An iot-cloud based wearable ecg monitoring system for smart healthcare. *Journal of medical systems*, 40(12):286, 2016.

[198] Can Ye, Miguel Tavares Coimbra, and BVK Vijaya Kumar. Arrhythmia detection and classification using morphological and dynamic features of ecg signals. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 1918–1921. IEEE, 2010.

[199] Can Ye, BVK Vijaya Kumar, and Miguel Tavares Coimbra. Combining general multi-class and specific two-class classifiers for improved customized ecg heartbeat classification. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pages 2428–2431. IEEE, 2012.

[200] Can Ye, BVK Vijaya Kumar, and Miguel Tavares Coimbra. Heartbeat classification using morphological and dynamic features of ecg signals. *IEEE Transactions on Biomedical Engineering*, 59(10):2930–2941, 2012.

[201] Yun-Chi Yeh, Wen-June Wang, and Che Wun Chiou. Cardiac arrhythmia diagnosis method using linear discriminant analysis on ecg signals. *Measurement*, 42(5):778–789, 2009.

[202] Yun-Chi Yeh, Wen-June Wang, and Che Wun Chiou. A novel fuzzy c-means method for classifying heartbeat cases from ecg signals. *Measurement*, 43(10):1542–1555, 2010.

[203] Özal Yildirim. A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification. *Computers in biology and medicine*, 96:189–202, 2018.

[204] Sung-Nien Yu and Ying-Hsiang Chen. Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. *Pattern Recognition Letters*, 28(10):1142–1150, 2007.

[205] Sung-Nien Yu and Kuan-To Chou. Integration of independent component analysis and neural networks for ecg beat classification. *Expert Systems with Applications*, 34(4):2841–2846, 2008.

[206] Sung-Nien Yu and Kuan-To Chou. Selection of significant independent components for ecg beat classification. *Expert Systems with Applications*, 36(2):2088–2096, 2009.

[207] Chenshuang Zhang, Guijin Wang, Jingwei Zhao, Pengfei Gao, Jianping Lin, and Huazhong Yang. Patient-specific ecg classification based on recurrent neural networks and clustering technique. In *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, pages 63–67. IEEE, 2017.

[208] Ji Zhang, Hongzhou Li, Xuemei Liu, Yonglong Luo, Fulong Chen, Hua Wang, and Liang Chang. On efficient and robust anonymization for privacy protection on massive streaming categorical information. *IEEE Transactions on Dependable and Secure Computing*, 14(5):507–520, 2017.

[209] Zhancheng Zhang, Jun Dong, Xiaoqing Luo, Kup-Sze Choi, and Xiaojun Wu. Heartbeat classification using disease-specific feature selection. *Computers in biology and medicine*, 46:79–89, 2014.

[210] Xingquan Zhu, Xindong Wu, and Ying Yang. Dynamic classifier selection for effective mining from noisy data streams. In *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on*, pages 305–312. IEEE, 2004.

[211] Zahia Zidelmal, Ahmed Amirou, Djaffar Ould-Abdeslam, and Jean Merckle. Ecg beat classification using a cost sensitive classifier. *Computer methods and programs in biomedicine*, 111(3):570–577, 2013.

[212] Muhammad Zubair, Jinsul Kim, and Changwoo Yoon. An automated ecg beat classification system using convolutional neural networks. In *2016 6th international conference on IT convergence and security (ICITCS)*, pages 1–5. IEEE, 2016.