

This is an author-created, un-copyedited version of an article accepted for publication/published in Physiological Measurement. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at <https://doi.org/10.1088/1361-6579/ac6f40>

Citation:

Apoorva Srivastava et al 2022 Physiol. Meas. in press <https://doi.org/10.1088/1361-6579/ac6f40>

DOI:

<https://doi.org/10.1088/1361-6579/ac6f40>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

ACCEPTED MANUSCRIPT

A deep residual inception network with channel attention modules for multi-label cardiac abnormality detection from reduced-lead ECG

To cite this article before publication: Apoorva Srivastava *et al* 2022 *Physiol. Meas.* in press <https://doi.org/10.1088/1361-6579/ac6f40>

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is © 2022 Institute of Physics and Engineering in Medicine.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by-nc-nd/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

A Deep Residual Inception Network with Channel Attention Modules for Multi-label Cardiac Abnormality Detection from Reduced-Lead ECG

Apoorva Srivastava^{1,*}, Sawon Pratiher¹, Sazedul Alam², Ajith Hari¹, Nilanjan Banerjee², Nirmalya Ghosh¹, Amit Patra¹

¹Department of Electrical Engineering, Indian Institute of Technology Kharagpur, West Bengal, India.

²Computer Science and Electrical Engineering, University of Maryland-Baltimore County, Maryland, USA.

E-mail: apoorva.s.2311@gmail.com (*Corresponding Author)

Keywords: Electrocardiogram; cardiac arrhythmia; multi-label classification; deep neural network; inception network; residual learning; channel attention.

Abstract. *Objective.* Most arrhythmias due to cardiovascular diseases alter the heart’s electrical activity, resulting in morphological alterations in electrocardiogram (ECG) recordings. ECG acquisition is a low-cost, non-invasive process and is commonly used for continuous monitoring as a diagnostic tool for cardiac abnormality identification. Our objective is to diagnose twenty-nine cardiac abnormalities and sinus rhythm using varied lead ECG signals. *Approach.* This work proposes a deep residual inception network with channel attention mechanism (**RINCA**) for twenty-nine cardiac arrhythmia classification (CAC) along with normal ECG from multi-label ECG signal with different lead combinations. The **RINCA** architecture employing the Inception-based convolutional neural network backbone uses residual skip connections with the channel attention mechanism. The Inception model facilitates efficient computation and prevents overfitting while exploring deeper networks through dimensionality reduction and stacked 1-dimensional convolutions. The residual skip connections alleviate the vanishing gradient problem. The attention modules selectively leverage the temporally significant segments in a sequence and predominant channels for multi-lead ECG signals, contributing to the decision-making. *Main results.* Exhaustive experimental evaluation on the large-scale ‘PhysioNet/Computing in Cardiology Challenge (2021)’ dataset demonstrates **RINCA**’s efficacy. On the hidden test data set, **RINCA** achieves the challenge metric score of 0.55, 0.51, 0.53, 0.51, and 0.53 (ranked 2nd, 5th, 4th, 5th and 4th) for the twelve-lead, six-lead, four-lead, three-lead, and two-lead combination cases, respectively. *Significance.* The proposed **RINCA** model is more robust against varied sampling frequency, recording time, and data with heterogeneous demographics than the existing art. The explainability analysis shows **RINCA**’s potential in clinical interpretations.

1. Introduction

1.1. Background

Burgeoning cardiovascular diseases (CVDs) have become one of the world's most critical public healthcare issues, taking an estimated 17.9 million lives each year (WHO 2020). Strokes and heart attacks account for more than four out of every five CVD fatalities, with one-third of these deaths occurring under 70 years of age (Virani et al. 2021). CVDs comprise heart and blood vessel-related abnormalities in the coronary artery and cerebrovascular, cardiomyopathy, rheumatic heart disease, etc. Most cardiac arrhythmias due to CVDs alter the heart's electrical activity and may be detected by an electrocardiogram (ECG) (Berkaya et al. 2018). In general, the 12-lead ECG signal acquisition is employed for early-stage multiple cardiac abnormality detection (Roth, Johnson, et al. 2018). However, manual ECG interpretation is time-consuming, tedious, and requires trained cardiologists to investigate and distinguish abnormal inter-beat and intra-beat patterns meticulously. Moreover, manual ECG waveform analysis is vulnerable to inter-observer and intra-observer variabilities. As such, reliable CVD diagnosis methods for automated ECG pattern classification are essential to aid cardiologists with timely interventions for prompt therapies and improve clinical outcomes (Roth, Mensah, et al. 2020).

Conventional CVD identification approaches employ feature-based machine learning (ML) techniques to extract disease-specific handcrafted features from the ECG signals. These include empirical mode decomposition (Abdelazez, Rajan, and Chan 2020), S-transform, wavelet transform, and Hermite interpolation (Ince, Kiranyaz, and Gabbouj 2009, Ye, Kumar, and Coimbra 2012), with ML classifiers-based hybrid expert systems with genetic algorithms (Moavenian and Khorrami 2010), neural networks (NN) (Hammad et al. 2020), support vector machines, K-nearest neighbor algorithm, decision trees (Qin et al. 2021), and random forests (Manibardo et al. 2019). However, these shallow ECG descriptors poorly comprehend the ECG signal's intricate properties pertinent to different CVDs with many non-linear components. Moreover, these ML models often fail to decipher the wide inter- and narrow intra-record variability in ECG records with high precision and pose challenges for end-to-end CVD detection. Further, ECG feature engineering requires time and domain expertise to check the feasibility of designing disease-specific morphological attributes from the ECG waveforms, which limits its practical application (Hong et al. 2020).

Recently deep learning (DL) methods have shown efficacy in achieving promising results and democratized the CVD recognition task with improved performance through supervised or unsupervised learning (Hannun et al. 2019, A. H. Ribeiro et al. 2020, Strodthoff et al. 2020). DL architectures strongly encapsulate the underlying non-linear and complex relationships in ECG signals with significantly improved performance (Mousavi and Afghah 2019, Hong et al. 2020). DL models envisage automated ECG feature representations for highly efficient CVD recognition in an end-to-end manner and significantly lessen the data processing complexity (Natarajan et al. 2020, Kirodiwal et al. 2020, Bos et al. 2021). A broad comparative evaluation with the prior work spanning ML and DL methods for CVD diagnosis is summarized in this work.

1.2. *Research Motivation and Problem Statement*

Most of the above-mentioned handcrafted feature-based ML and DL methods focus on single-label diagnosis, not multi-label cardiac arrhythmia classification (CAC) (Smith et al. 2019, Kirodiwal et al. 2020, Z. Xiong, Stiles, and J. Zhao 2017). The ECG data for method validation generally comes from Holter monitor or cardiac event monitor-based acquisition and has a duration spanning between a few seconds to a few hours from the same demographic region (Hannun et al. 2019, Clifford et al. 2017, Luz et al. 2016). Further, there are few publicly available databases aimed at multi-label CAC (Clifford et al. 2017, Luz et al. 2016). The MIT-BIH Malignant Ventricular Ectopy and MIT-BIH atrial fibrillation database have two-channel ECG recordings of 22 and 25 subjects, respectively, and it contains three CVD classes (Moody 1983, Greenwald 1986). The Creighton University (CU) ventricular tachyarrhythmia database contains three CVD classes with one channel recording of 35 subjects (Nolle et al. 1986). The prior art abounds in ML and DL algorithms validated on these publicly available databases for CAC, comprised of training data ranging from 20 to 50,000 collected from a very few subjects and thus limits its generalization potency (Oh et al. 2018).

The present study focuses on normal sinus rhythm and twenty-nine multi-label CAC from varied ECG lead combinations. The proposed framework is experimentally evaluated on the PhysioNet/Computing in Cardiology Challenge (2021) database having more than 1,11,000 ECG records with diverse demographics (Alday et al. 2020, Reyna et al. 2021). A detailed description of the studied cardiac abnormalities and their class-wise subject count is described in https://github.com/physionetchallenges/evaluation-2021/blob/main/dx_mapping_scored.csv. The training data comprises twelve-lead ECGs. The validation and test data encompass twelve-lead, six-lead, four-lead, three-lead, and two-lead ECGs, with various lead combinations enumerated below:

- **12-lead:** I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6
- **6-lead:** I, II, III, aVR, aVL, aVF
- **4-lead:** I, II, III, V2
- **3-lead:** I, II, V2
- **2-lead:** I, II

1.3. *Main Contribution*

The present work is in response to the invitation to submit an extended article to the Physiological Measurement Focus Issue (refer https://iopscience.iop.org/journal/0967-3334/page/Classification_Multilead_ECGs). A shorter version of this article was presented at the IEEE 2021 Computing in Cardiology (CinC) Conference for the PhysioNet/Computing in Cardiology Challenge (2021) (Srivastava et al. 2021). Our team participated in PhysioNet/Computing in Cardiology Challenge (2021) under the name "cardiochallenger". In this treatise, technical intricacies of the proposed deep residual inception network with channel attention mechanism (**RINCA**) framework, its different variants, hyperparameters tuning, exhaustive experimental validation results, and explainability of **RINCA**'s underlying predictions for clinical interpretations are expounded in detail. The **RINCA** architecture is inspired from the recently proposed

squeeze and excitation (SE) network (Hu, Shen, and G. Sun 2018, Zhu, H. Wang, et al. 2020). The design and evaluation of our **RINCA** framework on a large-scale database of more than 1,11,000 records for twenty-nine multi-label CAC provide the following contributions:

- To the best of authors' knowledge, **RINCA** framework is the first to employ an Inception architecture with residual skip connections and attention modules for multi-label CAC from varied length ECG signals, with **RINCA** outperforming the prior art (refer <https://physionetchallenges.org/2021/leaderboard/>).
- The Inception backbone and the residual skip connections in the **RINCA** architecture improve the multi-label CAC performance for twenty-nine classes without increasing the model's complexity, as evident by 6.85% less trainable parameters than the SE network (Hu, Shen, and G. Sun 2018).
- **RINCA**'s multi-head channel attention mechanism leverage significant temporal segments among the important leads, imparting robustness against variable sampling frequency, recording duration, and varied lead combinations, an essential aspect for clinical deployment in resource-constrained challenging environments.

Ongoing, Section 2 outlined the data pre-processing and the technical intricacies of the **RINCA** architecture (refer to Figure 1, 2, 3). Experimental results are discussed in Section 3 and Section 4 concludes the paper.

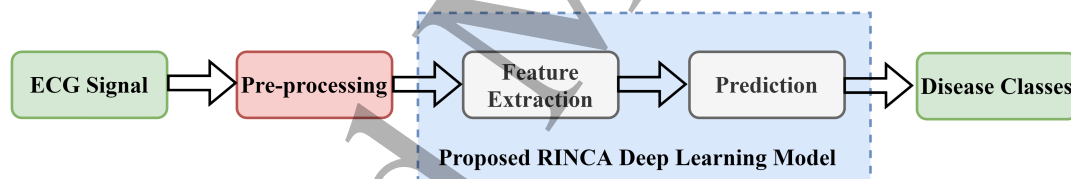


Figure 1: Pipeline showing the proposed **RINCA** framework and component modules.

2. Methodology

2.1. Pipeline for multi-label CAC from Pre-processed ECG

Most of the prior DL-based CAC research followed a similar workflow (Figure 1): Initial ECG pre-processing, ECG segmentation into equal-length segments, and then fed into DL frameworks. The ECG records from the PhysioNet/Computing in Cardiology Challenge (2021) (Reyna et al. 2021) database have variable sampling frequency and varying signal duration ranging between 6 seconds to 30 minutes. In the pre-processing stage, all the ECG records are down-sampled to 125 Hz (Ellis et al. 2015) by selecting every fourth ECG sample. All the model layers are modified to accept variable-length input and propagate the corresponding masks. ECG records longer than two minutes are condensed to two minutes (Only 74 records are affected, i.e., < 0.01 % of the total samples). The pre-processed ECG segments are fed as input to the **RINCA** model with the multi-label output array having an entry value equal to one when a particular abnormality is present, else set to zero.

2.2. Proposed RINCA Architecture

RINCA's backbone architecture consists of the SE-based Inception blocks (Hu, Shen, and G. Sun 2018) and residual skip connections. Multi-head attention and channel attention modules are fused to focus more on the salient temporal segments and channel-wise spatial feature representations. Source code with detailed technical specifications will be released (https://github.com/apoorvasrivastava23/ECG_RINCA). A modular level schematic of RINCA and its variants is shown in Figure 2 & 3, and brief functionalities of the different modules is explained below.

2.2.1. Inception Modules and Residual Connections: The fully connected layers in the convolutional neural network (CNN) are replaced with sparsely connected Inception network (Szegedy et al. 2015), which reduces the computational complexity while exploring a deeper multi-scale spatial feature representation. The introduction of residual skip connections alleviates the vanishing gradient problem, results in fewer extra parameters with increasing network depth, and accelerates training convergence (He et al. 2016). Mathematically, the residual block's (Resblock) output can be depicted as $y = \mathcal{R}(h_j) + \mathcal{D}(x_j)$, where $\mathcal{R}(\cdot)$ is the residual mapping learned during the the embedding phase. The convolutional operation is depicted as $\mathcal{D}(\cdot)$. Figure 2(b) demonstrate the Resblock's connection in the RINCA architecture, which is followed by max-pooling layer for feature map condensation (Scherer, Müller, and Behnke 2010).

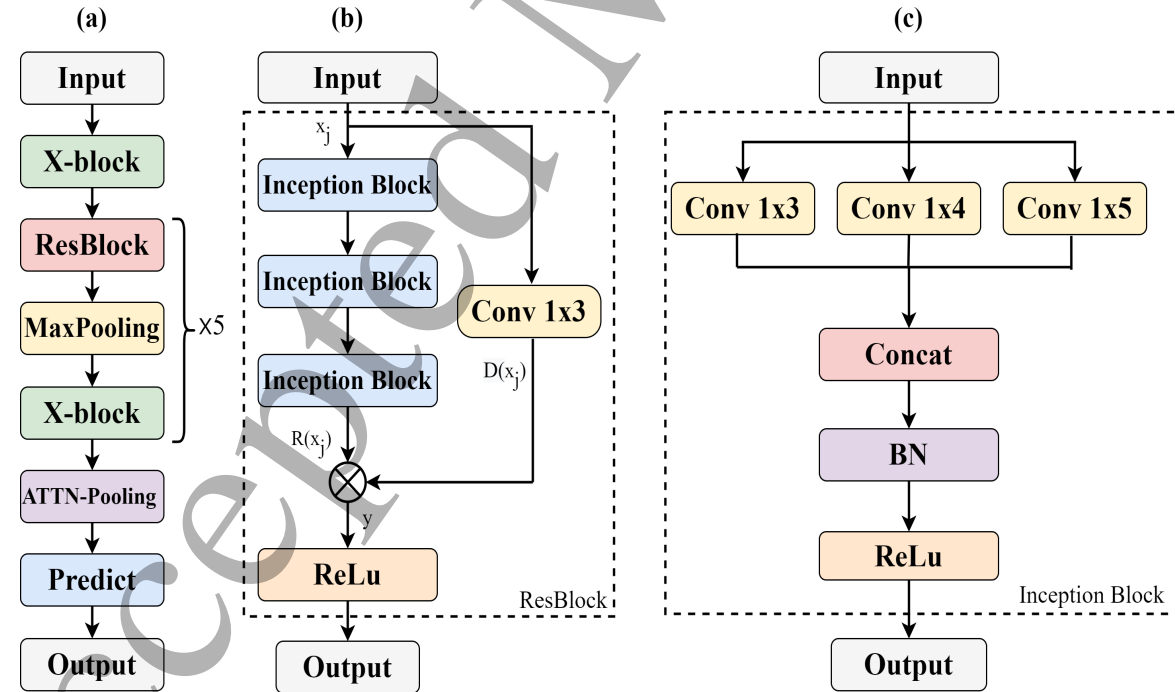


Figure 2: RINCA framework and its architectural variants. The grey output block of the model is one-hot encoded cardiac arrhythmias with the normal class. The Figure exemplify: (a) proposed RINCA model; (b) Residual block; (c) Inception block.

2.2.2. Attention Mechanism in RINCA: Attention in a DL framework boosts the important input channels and salient segments in input for improved model performance

and aids in prediction explainability. In the proposed **RINCA** model, two kinds of attention mechanisms are explored: channel attention and attention pooling.

2.2.3. X-block (Channel Attention): The CNN filters extract the feature maps, whose feature vector length is the number of neuronal units of the last residual block (Albawi, Mohammed, and Al-Zawi 2017). The spatial information in each input channel is extracted from the local receptive fields of the CNN layers. As depicted in Figure 3, the input to the **X-block** is of dimension $X \in \mathbb{R}^{T \times C}$, where T are the extracted spatial features and C is number of input channels from which N features are extracted in the feature extraction block. The various channel attention strategies (**X-block**) proposed for abstracting the relative importance of the input channels are explained below.

2.2.4. Squeeze and Excitation Network (SE-Net): The existing CNN architectures consider all the input channels of the same importance while computing the output feature maps. However, recent studies (Ilse, Tomczak, and Welling 2018, Tao et al. 2018, Nejedly et al. 2021) have highlighted the significance of channel attention in making the final prediction. The **SE-Net** architecture pioneered the incorporation of spatial channel-wise information within the local receptive fields at each layer (Hu, Shen, and G. Sun 2018). This work employs the **SE-Net**'s adaptive channel weighting strategy for multi-label CAC. Here, global average pooling is explored for a channel's importance comprehension (refer Figure 3 (a)).

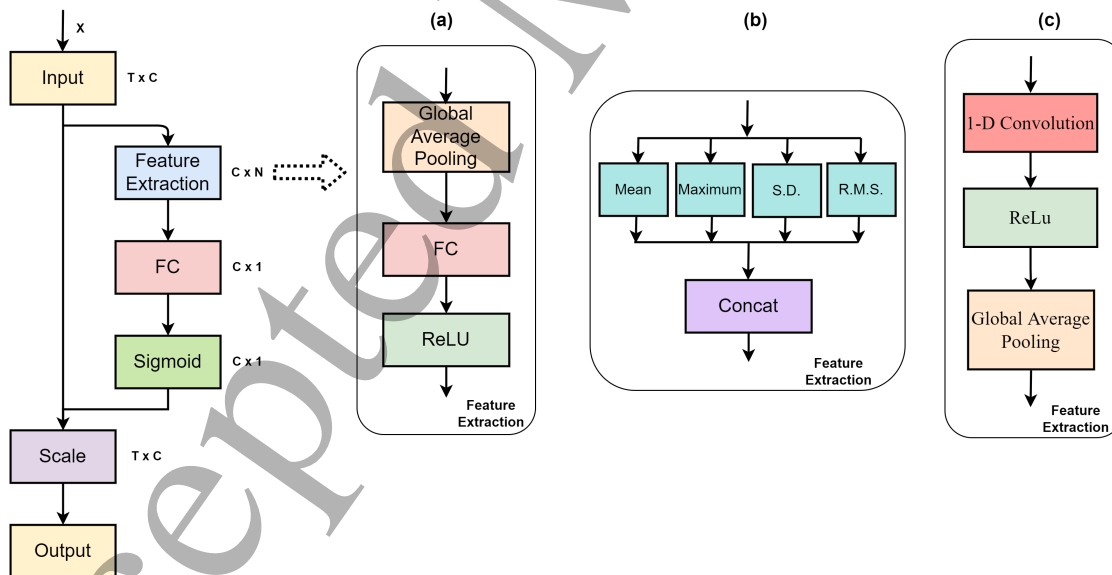


Figure 3: Technical intricacies of the X-block in the **RINCA** architecture for different variants. The overall X-block module is shown in the left. The variants of feature extraction modules are illustrated in (a) SE network; (b) Statistical feature-based RINCA architecture; (c) Deep feature-based RINCA architecture.

2.2.5. Statistical feature based RINCA (s-RINCA): The design of **s-RINCA** is inspired from SE-Net (Hu, Shen, and G. Sun 2018). **s-RINCA** uses multiple channel features like mean, maximum, root mean square (R.M.S.), and standard deviation (S.D.) instead of a single channel-wise feature for channel attention (refer Figure 3 (b)).

2.2.6. Deep feature-based RINCA (d-RINCA): In d-RINCA deep features are explored to comprehend the interdependence among the different ECG channels and use the SE-Net as the backbone architecture. Here, the deep features are extracted from the output of the **Resblock**'s after passing it through a one-dimension CNN block followed by ReLu activation (refer Figure 3 (c)).

2.2.7. Attention Pooling layer: The attention pooling is based on adaptive multi-instance pooling methodology (Ilse, Tomczak, and Welling 2018). Attention pooling helps to concentrate on informative and important relative features exacted from ECG segments. Likewise, multi-head attention mechanism incorporates an ensemble of multi-label class-specific feature importance depending on the number of output classes (Tao et al. 2018). The output of the multi-head attention block in **RINCA** is fed into the Softmax activation function. For $V = \{v_1, v_2, \dots, v_K\}$ being a bag of K feature vectors, the attention pooling operation is defined as: $p = \sum_{k=1}^K a_k v_k$, $p \in \mathbb{R}^{N \times L}$ is feature vector corresponding to N disease classes with L features. The attention weights (a_k) are given by $a_k = \frac{\exp\{W^T \tanh(Uv_k^T)\}}{\sum_{j=1}^K \exp\{W^T \tanh(Uv_j^T)\}}$, where, U and W are trainable. In the multi-head attention module, thirty attention units correspond to each one of the twenty-nine cardiac abnormalities and normal sinus rhythm under study (refer Figure 4). The number of attention units is motivated by the number of desired classes.

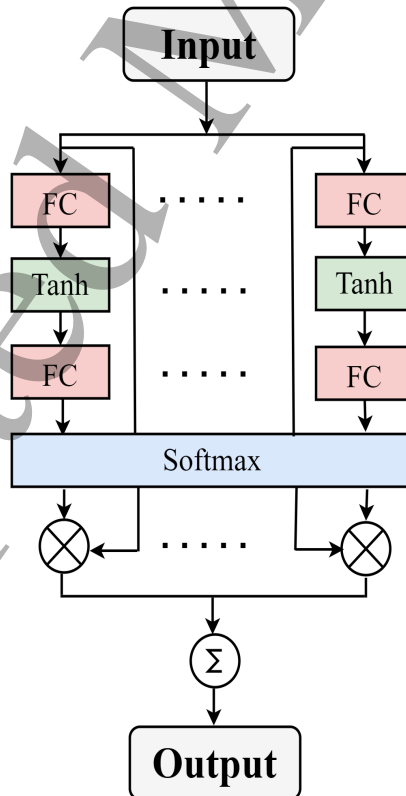


Figure 4: Technical intricacies of the attention pooling module

2.3. Threshold Optimisation

The output of the attention pooling layer goes into the prediction layer. The sigmoid activation in the prediction layer will generate the probabilities of occurrence of twenty-nine cardiac abnormalities and sinus rhythm. Converting these probabilities in binary format by applying a threshold to these prediction values to evaluate the challenge metric (CM) score is required. A value of one is assigned to the element of the multi-label output array corresponding to a particular cardiac abnormality class is present else, it is set to zero. Optimal thresholds values are calculated by maximizing the CM score on the validation dataset by using a genetic algorithm (Whitley 1994).

3. Experiment Results and Discussion

3.1. Experimental setup and Challenge Evaluation Metric

Although the dataset for the method validation comprises 133 cardiac arrhythmias, only twenty-nine cardiac abnormalities and the normal sinus rhythm are classified in the challenge. The Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) codes of the CAC classes are used as ECG record labels (Alday et al. 2020). For the model the loss function, $loss(x, y) = \alpha \cdot BCE(x, y) - S_{normalised}$ is optimized during the training phase. Here BCE is the binary cross-entropy, $S_{normalised}$ is the normalized CM score, α is the scaling factor, x and y are the true and predicted arrays, respectively. If the divergence between the ground truth and predicted label is less than 0.3, the BCE loss is scaled by a factor of 0.1, else scaled by a factor of 1, and $S_{normalised}$ is evaluated as: $S_{normalised} = \frac{S_{observed} - S_{inactive}}{S_{true} - S_{inactive}}$. Here, $S_{(x,y)} = X^T \times (\frac{y}{norm}) \cdot W_{reward}$, $norm = \max(x + y - xy, 1)$, W_{reward} is the reward matrix (Alday et al. 2020). The reward matrix is designed to give partial credit for wrong predictions and full credit for correct prediction, considering the symptoms and medication of the diseases.

In addition to standard performance evaluation metrics (refer Table 4), the official scoring methodology (refer Table 3) given by the PhysioNet/CinC Challenge 2021 (Alday et al. 2020) has been used for the proposed method's assessment. The challenge metric conforms with real-world clinical practice (Reyna et al. 2021). Some misdiagnoses are less detrimental than others, and partial credit is given to the misdiagnoses that result in comparable results as the ground truth. Further, the CM score encapsulates the clinical fact that certain misdiagnoses are more dangerous than others and should be graded appropriately (Zhu, Lan, et al. 2021). Comprehensive technical details can be found in <https://github.com/physionetchallenges/evaluation-2021>.

3.2. Training Details

The **RINCA** framework and its variants are trained using the Adam optimizer (Kingma and Ba 2014) with a learning rate of 0.001 for 100 epochs and a batch size of 32. The **d-RINCA** architecture has "715,512 (≈ 0.7 million)" trainable parameters and the model hyper-parameters are initialized using the Xavier uniform initializer (Glorot and Bengio 2010). Early stopping is used to evade model overfitting. The model is implemented in Python using Tensorflow and Keras libraries in Ubuntu 20.04 LTS x64 OS with the

system hardware consisting of 1xTesla K80, compute 3.7, having 2496 CUDA cores GPU, 12GB GDDR5 VRAM, and single-core hyperthreaded Xeon Processors @2.3Ghz CPU.

3.3. *Dataset Description*

We used annotated ECG data collected from 7 sources as provided by the organizers of the PhysioNet challenge, which spans four countries across three continents. There are more than 1,11,000 recordings of 12-lead ECG signals. Out of those recordings, the organizers offered 88,000 for training, 6,630 for validation, and more than 36,000 for testing purposes. Table 2 summarize the databases. The six-lead, four-lead, three-lead, and two-lead versions of the recordings are created from the twelve-lead recordings. A total of twenty-nine cardiac arrhythmia are studied listed in Table 1

3.4. *Comparative Analysis of the RINCA Framework and its Variants*

The averaged 3-fold cross validated CM scores on the in-house validation dataset for a variable number of ECG leads are tabulated in Table 3. The in-house training and validation comprise the publicly available ECG records from the PhysioNet/CinC challenge (<https://physionetchallenges.org/>). The in-house non-overlapping train and validation split for each fold is 70% and 30%.

Table 3 tabulate the CM scores on the hidden 6,630 and 16,630 ECG recordings for validation and test, respectively. The model's performance for varying lead combinations remains similar to the 12-lead case. There are 0.04% and 0.02% variances in the CM scores on the hidden validation and test datasets over the various lead sets. The model reduces the necessity of all the 12-lead ECG signal acquisition as per the performance.

For **s-RINCA** statistical features like minimum, skewness, kurtosis, crest factor, etc., are explored, but they did not improve the CAC results. The efficacy of **d-RINCA** compared to **s-RINCA**, and the SE-Net models may be attributed to the competency of non-linearity abstraction by the CNN layers. In SE-Net, the global average pooling outputs the single feature per channel, indicating averaged magnitude of signal. Instead of outputting single feature, **s-RINCA** computed four different statistical features for each channel. **d-RINCA** further extrapolates the idea of extracting multiple features. With CNN layer in **d-RINCA** 36 deep non-linear features per channel is extracted. The number of trainable parameters in **s-RINCA** and **d-RINCA** models are less than the SE-Net architecture.

3.5. *Detailed Performance Evaluation Scores Across Different Datasets*

Different evaluation metrics such as the area under piecewise linear function with sensitivity and specificity (AUROC), the area under piecewise linear function with recall and precision (AUPRC), CM score, etc., are used for **d-RINCA**'s efficacy evaluation in multi-label CAC performance. Table 4 highlights the different performance metrics for the hidden validation and test datasets with the **d-RINCA** model taking 2497 minutes for training. Here, 78% and 50% of CPSC and G12EC datasets, respectively, are taken for the **d-RINCA**'s training, and correspondingly, 11% and 25% of CPSC and G12EC

Table 1: Disease annotated in the ECG signals along with its abbreviation

Abbreviation	Disease Name
IAVB	1st Degree of AV Block
AF	Atrial fibrillation
AFL	Atrial flutter
BBB	Bundle branch block
Brady	Bradycardia
CLBBB	Complete left bundle branch block
CRBBB	Complete right bundle branch block
IRBBB	Incomplete right bundle branch block
LAnFB	Left anterior fascicular block
LAD	Left axis deviation
LBBB	Left bundle branch block
LQRSV	Low QRS voltage
NSIVCB	Nonspecific intraventricular conduction disorder
PR	Pacing rhythm
PRWP	Poor R wave progression
PAC	Premature atrial contraction
PVC	Premature ventricular contraction
LPR	Prolonged PR interval
LQT	Prolonged QT interval
QAb	Q wave abnormal
RAD	Right axis deviation
RBBB	Right bundle branch block
SA	Sinus arrhythmia
SB	Sinus bradycardia
ST	Sinus tachycardia
SVPB	Supraventricular premature beats
TAb	T wave abnormal
TInv	T wave inversion
VPB	Ventricular premature beats

Table 2: Dataset Description (Reyna et al. 2021).

Name	Origin	Number of Samples	Duration	Sampling frequency
CPSC and CPSC-Extra	China	13256	6-144 seconds	500 Hz
INCART	St Petersburg, Russia	74	30 minutes	257 Hz
PTB and PTB-XL	Germany	22353	10-120 seconds	500 or 1000 Hz
G12EC	Georgia	20672	5-10 seconds	500 Hz
Augmented Undisclosed	USA	10000	-	-
Chapman-Shaoxing and Ningbo	Shaoxing, Ningbo	45152	10 seconds	500 Hz
UMich	Michigan	19642	10 seconds	250 or 500 Hz

Table 3: Overall performance of **RINCA** model variants and **d-RINCA**'s ranking.

Lead	Training set			d-RINCA's CM score set		
	SE-Net	s-RINCA	d-RINCA	Validation Set	Test Set	Rank
12-Lead	0.71	0.77	0.81	0.64	0.55	2 nd
6-Lead	0.69	0.74	0.79	0.64	0.51	5 th
4-Lead	0.69	0.75	0.77	0.64	0.53	4 th
3-Lead	0.67	0.70	0.76	0.63	0.51	5 th
2-Lead	0.66	0.68	0.74	0.63	0.53	4 th

BOLD signifies in-house better performance.

datasets, respectively, are taken for both model validation and testing on the hidden dataset.

For the hidden validation dataset, the CM score on the CPSC and the G12EC dataset is higher than the hidden UMich and undisclosed test dataset (Table 4). On the hidden test dataset, the **d-RINCA** performance was superior with 12-lead ECG signals as compared to other lead combinations. The CM score degrades by 4.0% from using 12-lead, as we lower the number of leads. It may be due to the loss of clinically relevant information with the left-out ECG leads. There is an 8.87% increment and a 3.58% decrement in CM score for the CPSC and the G12EC test data, respectively, compared to the hidden validation data, which may be attributed to more CAC classes in the G12EC database. The CM score on the UMich and the undisclosed test data is reduced due to fewer subjects from North America in the training dataset.

3.6. Comparison with Other Challenge Approaches

Table 5 tabulate the prevalent other challenges approaches for CAC using ML and DL-based techniques and compares them with respect to the CM scores in the hidden validation and test datasets. From Table 5, the potency of DL models are apparent for CAC from ECG signals (Al-Nashash 2000, Y. Wang et al. 2001, Huang et al. 2019, J.-S. Wang et al. 2013, Jadhav, Nalbalwar, and Ghatol 2010). The top-scoring methods in Table 5 particularly uses different variants of deep neural network (DNN) architectures.

Proposed **d-RINCA**'s performance in the PhysioNet/CinC Challenge 2021 placed our team in the top five and outperformed the other challenge techniques (refer Table 5). Table 5 highlights a wide difference between the CM scores on the confidential validation data and hidden test data for a majority of the top-scoring teams in the PhysioNet/CinC Challenge 2021. For some of these methods, the CM score dropped by more than 25% on the confidential test data, which shows the issue of overfitting in the top-scoring methods (refer to Table 5). In contrast, proposed **d-RINCA** showed only a 14% drop in the CM score from the confidential validation data to the confidential test data, which demonstrates better generalization capability of **d-RINCA**.

Table 4: Detailed performance evaluation of **d-RINCA** across different datasets.

Dataset	Metric	Number of Leads				
		12	6	4	3	2
Training set	Training Run Time (minutes)	2497	2497	2497	2497	2497
Validation	AUROC	0.924	0.909	0.916	0.902	0.896
	AUPRC	0.474	0.477	0.469	0.454	0.452
	Accuracy	0.321	0.357	0.32	0.308	0.31
	F-measure	0.412	0.421	0.41	0.399	0.394
	Challenge Metric	0.642	0.641	0.638	0.626	0.625
	Test Run Time (minutes)	6	7	6	7	6
CPSC Test	AUROC	0.95	0.96	0.956	0.946	0.953
	AUPRC	0.805	0.84	0.829	0.796	0.825
	Accuracy	0.405	0.526	0.453	0.388	0.411
	F-measure	0.177	0.192	0.18	0.172	0.179
	Challenge Metric	0.699	0.784	0.738	0.675	0.706
	Test Run Time (minutes)	3	3	3	3	3
G12EC Test	AUROC	0.906	0.881	0.898	0.894	0.884
	AUPRC	0.48	0.474	0.468	0.472	0.455
	Accuracy	0.268	0.292	0.271	0.27	0.266
	F-measure	0.412	0.407	0.401	0.395	0.39
	Challenge Metric	0.619	0.61	0.614	0.613	0.603
	Test Run Time (minutes)	3	3	3	3	3
Undisclosed Test	AUROC	0.866	0.85	0.884	0.876	0.852
	AUPRC	0.472	0.464	0.487	0.492	0.476
	Accuracy	0.253	0.196	0.268	0.196	0.199
	F-measure	0.309	0.275	0.298	0.305	0.287
	Challenge Metric	0.434	0.407	0.4	0.358	0.433
	Test Run Time (minutes)	6	6	6	6	6
UMich Test	AUROC	0.909	0.892	0.9	0.9	0.891
	AUPRC	0.485	0.48	0.485	0.48	0.466
	Accuracy	0.297	0.306	0.309	0.303	0.309
	F-measure	0.418	0.402	0.411	0.403	0.396
	Challenge Metric	0.568	0.513	0.554	0.548	0.551
	Test Run Time (minutes)	12	12	12	12	13

3.7. Attention Weight Visualization for Clinical Interpretation

The explainability of **d-RINCA**'s predictions with physiological interpretations is visually depicted in Figure 5. Experimental validation of **d-RINCA** shows state-of-the-art superior performance. However, the clinical acceptance of **d-RINCA** requires explainable artificial intelligence (XAI) to analyze **d-RINCA**'s transparency and the social right to explain **d-RINCA**'s inferences (Došilović, Brčić, and Hlupić 2018). Here attention weight visualization is encompassed for a better understanding of **d-RINCA**'s interpretations. A higher attention weight is associated with an ECG segment

Table 5: Other challenge approach comparison on multi-label CAC from 12-lead ECG signals.

Reference	Method	Validation score	Test score
Cai et al. 2021	Deep residual convolutional neural networks	0.56	0.52
Bugata et al. 2021.	1D variant of the ResNet50 network	0.70	0.52
Vazquez et al. 2021	CNN and hand-crafted features	0.57	0.52
Ren, M. Xiong, and Hooi 2021	Deep Residual Network	0.64	0.51
Li et al. 2021	SE-ResNet, and ensemble models	0.67	0.49
Seki et al. 2021	DNN with DivideMix and stochastic weight averaging	0.62	0.49
This Work	d-RINCA	0.64	0.55

ResNet = Residual Neural Network.

being more salient and informative towards making the final prediction (Adadi and Berrada 2018). The multi-head attention module generates an ensemble of twenty-nine weights from one ECG recording. These twenty-nine weight vector entries represent the probabilities of a particular distinct cardiac arrhythmia being present. Figure 5 shows the efficacy of attention modules in improved multi-label CAC performance with clinical saliency corresponding to three commonly occurring cardiac arrhythmias, namely Atrial Fibrillation (AF), Premature Atrial Contraction/ Supraventricular Premature Beats (PAC/SVPB), and Premature Ventricular Contractions/ Ventricular Premature Beats (PVC/VPB). Due to similar symptoms of cardiac arrhythmia, PAC and SVPB, PVC, and VPB are considered the same class. In Figure 5 shows the ECG signals with clinical demarcated abnormal regions (green), and highlights the attention weights (red; salient ECG segments which the DL model gives more weightage for a making a prediction). The attention weights are higher in the abnormal ECG segments, and they can be used to highlight abnormal ECG patterns, as seen in Figure 5.

4. Conclusion

The present work proposes a DL framework based on a residual-Inception architecture with channel attention blocks for automated multi-label CAC from different ECG lead combinations with varying durations. Compared to the prevalent art, the proposed **RINCA** model is more robust against variable sampling frequency, recording duration, and data having diverse demographics. Out of all the **RINCA** variants, the deep feature-based **RINCA** performs superior to the statistical-based **RINCA** and SE-Net models (refer to Table 3). The channel attention module in **RINCA** perceives the spatial interdependence among the channels, and the attention pooling identifies the salient temporal ECG segments having decision-making implications. The potency of the attention mechanisms as mentioned above is evident from the improved CM scores with scaled BCE loss. The CM score is highest for the **d-RINCA** architecture (refer Figure 3). In the future, we plan to explore techniques to mitigate the class imbalance problem and explore DL frameworks aiming to improve the CM score for the multi-label CAC task with different combinations of ECG leads.

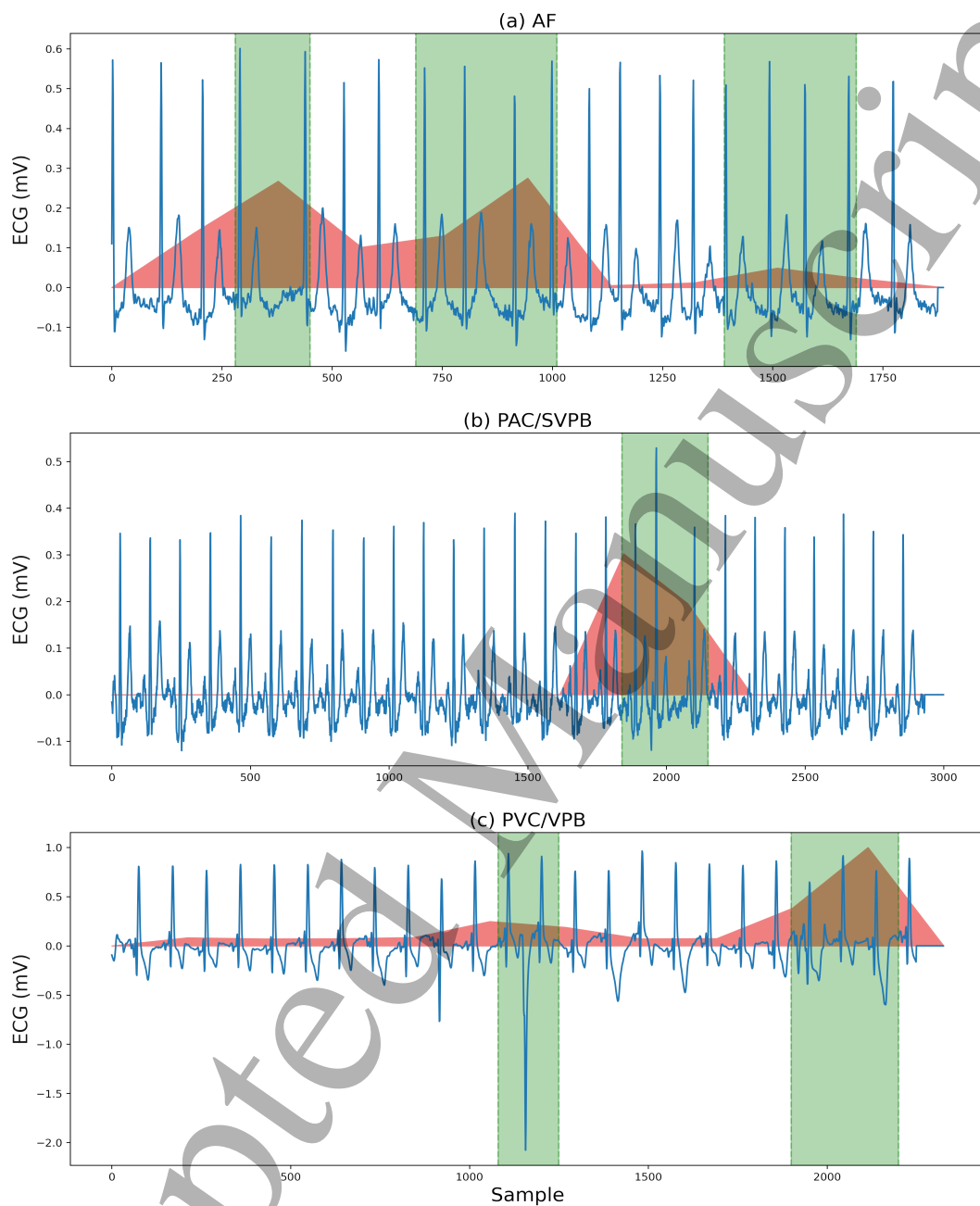


Figure 5: **d-RINCA**'s explainability visualization. (a) ECG recordings (blue), (b) overlaid clinically demarcated abnormal regions (green), and (c) overlaid attention weights (red), showing the ECG segments' saliency toward the final prediction.

Acknowledgments

The authors would like to thank Kharagpur Learning, Imaging & Visualization (KLIV) research group for providing GPU access to test our frameworks.

REFERENCES

15

References

Abdelazez, M., Rajan, S., and Chan, A. D. (2020). "Detection of Atrial Fibrillation in Compressively Sensed Electrocardiogram Measurements". In: *IEEE Transactions on Instrumentation and Measurement* 70, pp. 1–9.

Adadi, A. and Berrada, M. (2018). "Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)". In: *IEEE Access* 6, pp. 52138–52160.

Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). "Understanding of A Convolutional Neural Network". In: *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6.

Alday, E. A. P., Gu, A., Shah, A. J., Robichaux, C., Wong, A.-K. I., Liu, C., Liu, F., Rad, A. B., Elola, A., Seyed, S., Li, Q., Sharma, A., Clifford, G. D., and Reyna, M. A. (2020). "Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020". In: *Physiological Measurement* 41.12, p. 124003.

Berkaya, S. K., Uysal, A. K., Gunal, E. S., Ergin, S., Gunal, S., and Gulmezoglu, M. B. (2018). "A Survey on ECG Analysis". In: *Biomedical Signal Processing and Control* 43, pp. 216–235.

Bos, J. M., Attia, Z. I., Albert, D. E., Noseworthy, P. A., Friedman, P. A., and Ackerman, M. J. (2021). "Use of Artificial Intelligence and Deep Neural Networks in Evaluation of Patients with Electrocardiographically Concealed Long QT Syndrome from The Surface 12-lead Electrocardiogram". In: *JAMA Cardiology* 6.5, pp. 532–538.

Bugata, P., Jr., P. B., Kmecova1, V., Stankova, M., Gajdos, D., Hudak, D., Stana, R., Horvat, S., Antoni, L., Vozarikova, G., Bruoth, E., and Szabari, A. (2021). "A Two-Phase Multilabel ECG Classification Using One-Dimensional Convolutional Neural Network and Modified Labels". In: URL: <https://www.cinc.org/2021/Program/accepted/78.html>.

Cai, W., Liu, F., Wang, X., Xu, B., and Wang, Y. (2021). "Classifying Different Dimensional ECGs using Deep Residual Convolutional Neural Networks". In: URL: <https://www.cinc.org/2021/Program/accepted/105.html>.

Clifford, G. D., Liu, C., Moody, B., Li-wei, H. L., Silva, I., Li, Q., Johnson, A., and Mark, R. G. (2017). "AF Classification from A Short Single Lead ECG Recording: The PhysioNet/Computing in Cardiology Challenge 2017". In: *2017 Computing in Cardiology (CinC)*. IEEE, pp. 1–4.

Došilović, F. K., Brčić, M., and Hlupić, N. (2018). "Explainable Artificial Intelligence: A Survey". In: *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, pp. 0210–0215.

Ellis, R. J., Zhu, B., Koenig, J., Thayer, J. F., and Wang, Y. (2015). "A Careful look at ECG Sampling Frequency and R-peak Interpolation on Short-term Measures of Heart Rate Variability". In: *Physiological measurement* 36.9, p. 1827.

Glorot, X. and Bengio, Y. (2010). "Understanding the Difficulty of Training Deep Feedforward Neural Networks". In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256.

Greenwald, S. D. (1986). "The Development and Analysis of A Ventricular Fibrillation Detector". PhD thesis. Massachusetts Institute of Technology.

REFERENCES

16

- Hammad, M., Ilyasu, A. M., Subasi, A., Ho, E. S., and Abd El-Latif, A. A. (2020). “A Multitier Deep Learning Model for Arrhythmia Detection”. In: *IEEE Transactions on Instrumentation and Measurement* 70, pp. 1–9.
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., and Ng, A. Y. (2019). “Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms using a Deep Neural Network”. In: *Nature Medicine* 25.1, pp. 65–69.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep Residual Learning for Image Recognition”. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
- Hong, S., Zhou, Y., Shang, J., Xiao, C., and Sun, J. (2020). “Opportunities and Challenges of Deep Learning Methods for Electrocardiogram Data: A Systematic Review”. In: *Computers in Biology and Medicine*, p. 103801.
- Hu, J., Shen, L., and Sun, G. (2018). “Squeeze-and-Excitation Networks”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141.
- Huang, J., Chen, B., Yao, B., and He, W. (2019). “ECG Arrhythmia Classification using STFT-based Spectrogram and Convolutional Neural Network”. In: *IEEE Access* 7, pp. 92871–92880.
- Ilse, M., Tomczak, J., and Welling, M. (2018). “Attention-based Deep Multiple Instance Learning”. In: *International Conference on Machine Learning*, pp. 2127–2136.
- Ince, T., Kiranyaz, S., and Gabbouj, M. (2009). “A Generic and Robust Ssystem for Automated Patient-Specific Classification of ECG Signals”. In: *IEEE Transactions on Biomedical Engineering* 56.5, pp. 1415–1426.
- Jadhav, S. M., Nalbalwar, S., and Ghatol, A. (2010). “Artificial Neural Network based Cardiac Arrhythmia Classification using ECG Signal Data”. In: *2010 International Conference on Electronics and Information Engineering*. Vol. 1. IEEE, pp. V1–228.
- Kingma, D. P. and Ba, J. (2014). “Adam: A Method for Stochastic Optimization”. In: *arXiv preprint arXiv:1412.6980*.
- Kirodiwal, A., Srivastava, A., Dash, A., Saha, A., Penaganti, G. V., Pratiher, S., Alam, S., Patra, A., Ghosh, N., and Banerjee, N. (2020). “A Bio-toolkit for Multi-Cardiac Abnormality Diagnosis Using ECG Signal and Deep Learning”. In: *2020 Computing in Cardiology*, pp. 1–4. DOI: [10.22489/CinC.2020.225](https://doi.org/10.22489/CinC.2020.225).
- Li, X., Li, X., Li, C., Xu, X., Wei, Y., Wei, J., Sun, Y., Qian, B., and Xu, X. X. (2021). “Towards Generalization of Cardiac Abnormality Classification Using ECG Signal”. In: URL: <https://www.cinc.org/2021/Program/accepted/212.html>.
- Luz, E. J. d. S., Schwartz, W. R., Cámara-Chávez, G., and Menotti, D. (2016). “ECG-based Heartbeat Classification for Arrhythmia Detection: A Survey”. In: *Computer Methods and Programs in Biomedicine* 127, pp. 144–164.
- Manibardo, E., Irusta, U., Del Ser, J., Aramendi, E., Isasi, I., Olabarria, M., Corcuera, C., Veintemillas, J., and Larrea, A. (2019). “ECG-based Random Forest Classifier for Cardiac Arrest Rhythms”. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 1504–1508.

REFERENCES

17

Moavenian, M. and Khorrami, H. (2010). "A Qualitative Comparison of Artificial Neural Networks and Support Vector Machines in ECG Arrhythmias Classification". In: *Expert Systems with Applications* 37.4, pp. 3088–3093.

Moody, G. (1983). "A New Method for Detecting Atrial Fibrillation using RR Intervals". In: *Computers in Cardiology*, pp. 227–230.

Mousavi, S. and Afghah, F. (2019). "Inter-and Intra-Patient ECG Heartbeat Classification for Arrhythmia Detection: A Sequence to Sequence Deep Learning Approach". In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 1308–1312.

Al-Nashash, H. (2000). "Cardiac Arrhythmia Classification using Neural Networks". In: *Technology and Health care* 8.6, pp. 363–372.

Natarajan, A., Chang, Y., Mariani, S., Rahman, A., Boverman, G., Vij, S., and Rubin, J. (2020). "A Wide and Deep Transformer Neural Network for 12-Lead ECG Classification". In: *Computing in Cardiology*, pp. 1–4.

Nejedly, P., Ivora, A., Viscor, I., Loscova, Z., Smisek, R., Jurak, P., and Plesinger, F. (2021). "Classification of ECG using Ensemble of Residual CNNs with Attention Mechanism". In: URL: <https://www.cinc.org/2021/Program/accepted/14.html>.

Nolle, F., Badura, F., Catlett, J., Bowser, R., and Sketch, M. (1986). "CREI-GARD, A New Concept in Computerized Arrhythmia Monitoring Systems". In: *Computers in Cardiology* 13, pp. 515–518.

Oh, S. L., Ng, E. Y., Tan, R. S., and Acharya, U. R. (2018). "Automated Diagnosis of Arrhythmia using Combination of CNN and LSTM Techniques with Variable Length Heart Beats". In: *Computers in Biology and Medicine* 102, pp. 278–287. ISSN: 0010-4825.

Qin, L., Xie, Y., Liu, X., Yuan, X., and Wang, H. (2021). "An End-to-End 12-Leading Electrocardiogram Diagnosis System Based on Deformable Convolutional Neural Network With Good Antinoise Ability". In: *IEEE Transactions on Instrumentation and Measurement* 70, pp. 1–13.

Ren, H., Xiong, M., and Hooi, B. (2021). "Robust and Task-Aware Training of Deep Residual Networks for Varying-Lead ECG Classification". In: URL: <https://www.cinc.org/2021/Program/accepted/122.html>.

Reyna, M. A., Sadr, N., Alday, E. A. P., Gu, A., Shah, A. J., Robichaux, C., Rad, A. B., Elola, A., Seyed, S., Ansari, S., Ghanbari, H., Li, Q., Sharma, A., and Clifford, G. D. (2021). "Will Two Do? Varying Dimensions in Electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021". In: *Computing in Cardiology* 48, pp. 1–4.

Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., Ferreira, M. P., Andersson, C. R., Macfarlane, P. W., Meira Jr, W., et al. (2020). "Automatic Diagnosis of the 12-lead ECG using a Deep Neural Network". In: *Nature Communications* 11.1, pp. 1–9.

Roth, G. A., Johnson, C. O., Abate, K. H., Abd-Allah, F., Ahmed, M., Alam, K., Alam, T., Alvis-Guzman, N., Ansari, H., Ärnlov, J., et al. (2018). "The Burden of Cardiovascular Diseases Among US States, 1990-2016". In: *JAMA cardiology* 3.5, pp. 375–389.

REFERENCES

18

- Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., et al. (2020). “Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study”. In: *Journal of the American College of Cardiology* 76.25, pp. 2982–3021.
- Scherer, D., Müller, A., and Behnke, S. (2010). “Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition”. In: *International conference on artificial neural networks*. Springer, pp. 92–101.
- Seki, H., Nakano, T., Ikeda, K., Hirooka, S., Kawasaki, T., Yamada, M., Saito, S., Yamakawa, T., and Ogawa, S. (2021). “Reduced-Lead ECG Classifier Model Trained with DivideMix and Model Ensemble”. In: URL: <https://www.cinc.org/2021/Program/accepted/79.html>.
- Smith, S. W., Walsh, B., Grauer, K., Wang, K., Rapin, J., Li, J., Fennell, W., and Taboulet, P. (2019). “A Deep Neural Network Learning Algorithm Outperforms a Conventional Algorithm for Emergency Department Electrocardiogram Interpretation”. In: *Journal of Electrocardiology* 52, pp. 88–95. ISSN: 0022-0736.
- Srivastava, A., Hari, A., Pratiher, S., Alam, S., Ghosh, N., Banerjee, N., and Patra, A. (2021). “Channel Self-Attention Deep Learning Framework for Multi-Cardiac Abnormality Diagnosis from Varied-Lead ECG Signals”. In: *2021 Computing in Cardiology (CinC)*. Vol. 48, pp. 1–4. DOI: [10.23919/CinC53138.2021.9662886](https://doi.org/10.23919/CinC53138.2021.9662886).
- Strodthoff, N., Wagner, P., Schaeffter, T., and Samek, W. (2020). “Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL”. In: *IEEE Journal of Biomedical and Health Informatics* 25.5, pp. 1519–1528.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). “Going Deeper with Convolutions”. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9.
- Tao, C., Gao, S., Shang, M., Wu, W., Zhao, D., and Yan, R. (2018). “Get The Point of My Utterance! Learning Towards Effective Responses with Multi-Head Attention Mechanism.” In: *International Joint Conference on Artificial Intelligence*, pp. 4418–4424.
- Vazquez, C. G., Breuss, A., Gnarra, O., Portmann, J., and Poian, G. D. (2021). “Two will do: Convolutional Neural Network with Asymmetric Loss and Self-Learning Label Correction for Imbalanced Multi-Label ECG Data Classification”. In: URL: <https://www.cinc.org/2021/Program/accepted/24.html>.
- Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Cheng, S., Delling, F. N., et al. (2021). “Heart Disease and Stroke Statistics—2021 Update: A Report from The American Heart Association”. In: *Circulation* 143.8, e254–e743.
- Wang, J.-S., Chiang, W.-C., Hsu, Y.-L., and Yang, Y.-T. C. (2013). “ECG Arrhythmia Classification using A Probabilistic Neural Network with A Feature Reduction Method”. In: *Neurocomputing* 116, pp. 38–45.
- Wang, Y., Zhu, Y.-S., Thakor, N. V., and Xu, Y.-H. (2001). “A Short-Time Multifractal Approach for Arrhythmia Detection based on Fuzzy Neural Network”. In: *IEEE Transactions on Biomedical Engineering* 48.9, pp. 989–995.

REFERENCES

19

Whitley, D. (1994). “A Genetic Algorithm Tutorial”. In: *Statistics and computing* 4.2, pp. 65–85.

WHO (2020). *Cardiovascular Diseases*. Accessed 03 January 2021. URL: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1.

Xiong, Z., Stiles, M. K., and Zhao, J. (2017). “Robust ECG Signal Classification for Detection of Atrial Fibrillation using A Novel Neural Network”. In: *2017 Computing in Cardiology (CinC)*, pp. 1–4.

Ye, C., Kumar, B. V., and Coimbra, M. T. (2012). “Heartbeat Classification using Morphological and Dynamic Features of ECG Signals”. In: *IEEE Transactions on Biomedical Engineering* 59.10, pp. 2930–2941.

Zhu, Z., Lan, X., Zhao, T., Guo, Y., Kojodjojo, P., Xu, Z., Liu, Z., Liu, S., Wang, H., Sun, X., et al. (2021). “Identification of 27 Abnormalities from Multi-Lead ECG Signals: An Ensembled SE-ResNet Framework with Sign Loss Function”. In: *Physiological Measurement*.

Zhu, Z., Wang, H., Zhao, T., Guo, Y., Xu, Z., Liu, Z., Liu, S., Lan, X., Sun, X., and Feng, M. (2020). “Classification of Cardiac Abnormalities From ECG Signals Using SE-ResNet”. In: *Computing in Cardiology*, pp. 1–4.