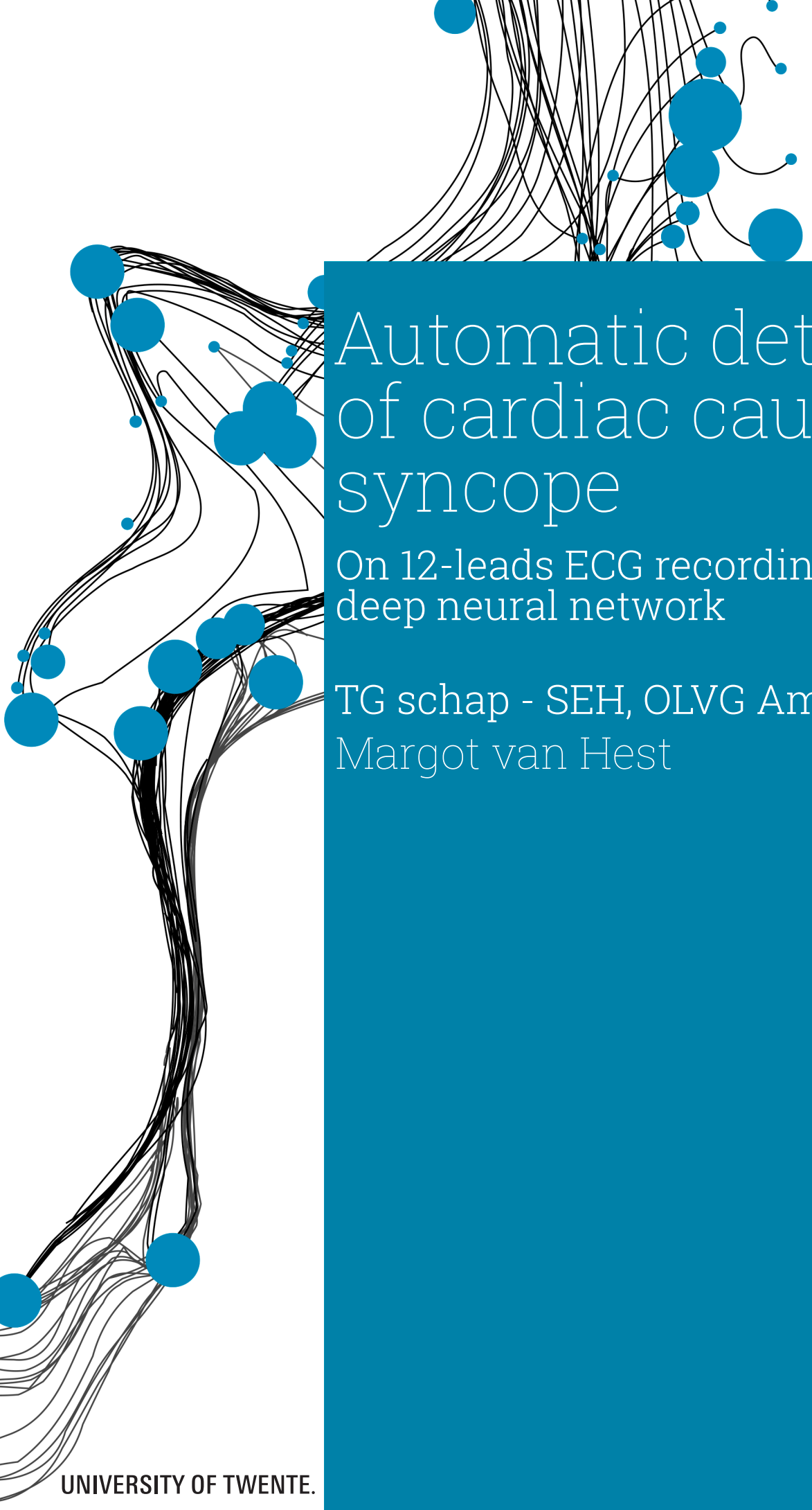


Automatic detection of cardiac causes of syncope

On 12-leads ECG recordings using a deep neural network

TG schap - SEH, OLVG Amsterdam
Margot van Hest

UNIVERSITY OF TWENTE.

Automatic detection of cardiac causes of syncope

On 12-leads ECG recordings using a deep neural network

by

Margot van Hest

Studentnumber: s2280299

Medical Supervisor (OLVG Amsterdam): Dr. Ir. J.M.M. van Breugel-Smits
Medical Supervisor (OLVG Amsterdam): Drs. M.M.S. Zwartsenburg
Technical Supervisor (OLVG Amsterdam): MSc. M. van den Boorn
Technical Supervisor (Universiteit Twente): dr. A. John

December 2023 - February 2024
Master Technical Medicine, track: Medical Sensing and Stimulation
Faculty: Faculty of Science and Technology, Enschede

UNIVERSITY OF TWENTE.

Abstract

The detection of cardiac causes for Syncope is crucial as cardiac-related syncopal events are associated with an increased one-year mortality. For this reason, this study utilized the Chapman Ningbo dataset, comprising 45,152 ECG recordings, for the development of two residual neural networks. A ResNet models with and without an Efficient Channel Attention (ECA) module were trained and evaluated on this dataset. The original ResNet model obtained an F1 score of 0.72 and a recall score 0.67, while the model with ECA module obtained an F1 score of 0.70 and recall score of 0.63. Furthermore, the models were only able to classify two diagnoses in the dataset related to the WOBBLER acronym for ECG features of cardiac syncope with reasonable recall scores. For this reason, the ResNet models described in the current study are not suitable for deployment in clinical practice due to their inadequate recall and F1 scores. As it stands, deploying these models would not contribute to the goal of automatically detecting cardiac causes of syncope from ECG signals, nor would it enhance the care and prognoses of these patients.

Click here for the corresponding Github page containing the used files

Contents

1

Introduction

Syncope, a common presentation in emergency departments, poses diagnostic challenges, with up to a third of cases lacking a definitive etiology.[3, 9] This is particularly worrisome, especially given the association of cardiac-related syncopal events with increased one-year mortality rates.[9]

Among the various causes of cardiac syncope, arrhythmias occur the most frequently.[9, 2] Detecting these arrhythmias in syncopal patients however, poses a unique challenge as they may not manifest continuously. Additionally, the fast-paced environment of emergency departments (ED) often hinders continuous electrocardiogram (ECG) monitoring, potentially leading to the oversight of underlying arrhythmias.

Recent advancements in deep learning (DL) approaches for classifying arrhythmias from ECG recordings show promising outcomes. For instance, Ribeiro et al. developed a deep neural network (DNN) model that outperformed a 4th-year cardiology resident in identifying six types of abnormalities in 12-lead ECG recordings.[14] This underscores the potential of DL in providing accurate ECG classification and the potential for continuous monitoring.

The primary aim here is to craft a robust DL model capable of independently recognizing and flagging arrhythmias leading to syncope in patients. This model could be integrated into a system where flagged segments of ECG recordings trigger an automatic saving and upload into the electronic patient dossier. Therefore, allowing these recordings to undergo evaluation by a physician and preventing an arrhythmia as the cause of a syncopic event from being missed. This could enhance patient care outcomes as the underlying can be treated.

2 Background

2.1. Definition

This report adopts the definition of syncope as outlined in the European Society of Cardiology (ESC) guideline for the diagnosis and management of syncope. Syncope is described as a transient loss of consciousness (TLOC) resulting from cerebral hypoperfusion. It manifests with a sudden onset, brief duration, and spontaneous complete recovery.[2]

2.2. Pathofysiology

Syncope encompasses three primary categories distinguished by their shared pathophysiological mechanisms: reflex (neurally mediated) syncope, orthostatic hypotension syncope, and cardiac syncope.[2] These categories share the same general pathological principle which is that a decline in systemic blood pressure (BP) leads to inadequate cerebral blood flow thereby causing the syncopal event. Blood pressure is the product of cardiac output (CO) and peripheral resistance. The underlying etiology of syncope typically involves a reduction in one or both of these components.

2.2.1. Reflex (neurally mediated) syncope

Reflex syncope is any kind of syncope that is caused by failure of autoregulation to maintain blood pressure, leading to decreased cerebral perfusion pressure which results in the transient loss of consciousness. This decline in BP in reflex syncope can be attributed to decreased reflex activity, causing vasodilation or a cardioinhibitory reflex.[2, 22]

Reflex syncope can be further divided into the following subtypes:

- vasovagal syncope: This subtype can be orthostatic (occurring while lying or standing) or emotionally triggered (by stimuli as fear, pain (somatic or visceral), instrumentation, blood phobia)
- situational syncope: This subtype is triggered by specific situations like micturition, gastrointestinal stimulation, coughing, or post-exercise.
- carotid sinus syndrome or non-classical forms[2, 22]

2.2.2. Orthostatic hypotension syncope

Orthostatic hypotension is defined as a drop in systolic BP of $>$20 mmHg. This decrease in systolic BP leads to a transient fall in cerebral blood flow leading to the syncopal event.[2] Orthostatic hypotension is most often drug-induced, but can also be caused by volume depletion (due to hemorrhage, diarrhea, vomiting, etc.), primary autonomic failure, and secondary autonomic failure.[2]

2.2.3. Cardiac syncope

Cardiac syncope involves a decrease in blood pressure resulting from a drop in cardiac output caused by arrhythmias (electrical abnormalities), structural abnormalities, or pathologies in the cardiopulmonary circulation and major vessels. Arrhythmias are the primary cause of cardiac syncope.[2] These arrhythmias can be divided into bradycardiac arrhythmias and tachycardic arrhythmias which are further elaborated below.

Bradycardia

Bradycardia can lead to syncope because the heart beats too slow to generate sufficient CO to maintain adequate BP.

Sinus node dysfunction This term describes the inability of the sinus node to generate an adequate heart rate that meets the physiological needs of a patient.[6, 4] The term encompasses various sinus node abnormalities, including:

- Sinus bradycardia: Electrical activity of the sinus node drops below <40 beats per minute.[6, 4]
- Sinus pauses or sinus arrest: Intermittent periods of more than 3 seconds without sinus activity.[6]
- Sinoatrial exit block: impaired transmission of electrical pulses from the sinus node to adjacent atrial tissue.[4]
- tachy-brady syndrome: Bradycardia which is alternated with atrial fibrillation.[4]

Atrioventricular conduction system disease Abnormalities in the conduction system can cause irregularities in the transmission of electrical impulses leading to not all electrical impulses from the SN being followed by ventricular contraction.

- First-degree AV block: the P waves are always followed by a QRS complex, but the PR interval is prolonged (>200 ms). The prolonged PR interval is characteristically seen on the ECG.[7]
- Second-degree AV block Morbitz type I (Wenckebach): Progressive prolongation of the PR-interval, which eventually leads to a dropped beat due to an electrical impulse not being conducted by the AV-node.[7]
- second-degree AV block Morbitz type II: Intermittent, randomly non-conducted P waves, with constant P-P intervals on the ECG.[7]
- third-degree AV block: Complete blockage of the AV node. The QRS complexes on the ECG are not related to the P waves. Ventricular activity is maintained by the junctional or ventricular escape rhythm.[7]
- bifascicular block: This is a conduction block of the right bundle branch (RBBB) and the anterior or posterior fascicle. A complete left bundle branch also constitutes as a bifascicular block as both the anterior and posterior fascicles are involved.[16]

Tachycardia

Tachycardia can provoke syncope due to the heart beating excessively fast, hindering adequate ventricular filling during the diastolic phase. This leads to failure to maintain sufficient cardiac output and therefore cerebral perfusion. Tachycardia can be further divided into supraventricular tachycardia, ventricular tachycardia, and accessory conduction pathways.

Supraventricular tachycardia This term encompasses tachycardia (atrial and/or ventricular rates >100 beats per minute) originating at or above the AV node. Key characteristics of the ECG include a narrow QRS complex of less than 120 milliseconds and for most of these arrhythmias, the p-wave is not easily discernible.[11]

- Atrial fibrillation: Occurs when atrial tissue depolarizes spontaneously outside the sinoatrial node, leading to a fast and irregular heartbeat due to random pulse transmission by the AV node. This arrhythmia is characterized by an irregular pattern of QRS complexes on the ECG.[20]
- Atrial flutter: Characterized by a fast atrial rate with a fixed or variable ventricular rate which is slower than the atrial rate as not all impulses are conducted by the AV node. Contrary to atrial fibrillation the conduction of pulses by the AV-node by a fixed proportion. This arrhythmia is characterized by a “sawtooth” pattern on the ECG.[20]
- AV re-entrant tachycardia (AVRT): Involves an accessory pathway that exists between the atrium and the ventricle. This accessory pathway can cause pre-excitation of the ventricle and causes the characteristic short PR interval (<120 ms) and broadening of the QRS complex on the ECG.[11, 20]
- AV nodal re-entrant tachycardia (AVNRT): Occurs due to the presence of two functionally distinct electrophysiological tracts: one fast and one slow. Due to the existence of these functionally different pathways in the AV node a premature beat can initiate a repetitive re-entry loop. On the ECG this is recognizable as narrow QRS tachycardia and the P-waves are not easily distinguishable.[11, 20]

Ventricular Tachycardia Non-sustained ventricular tachycardia (VT) is a wide complex arrhythmia originating in the ventricle, with a rate exceeding 100 beats per minute and lasting less than 30 seconds.[5] A distinction can be made based on the QRS morphology into monomorphic or polymorphic VT. This type of tachycardia often arises secondary to mechanical or structural heart diseases, or channelopathies such as long QT syndrome and Brugada Syndrome.

- Long QT syndrome: This condition can be recognized by a prolonged QT interval on the ECG.[10]
- Brugada Syndrome: Characterized by ST-segment elevation in the right precordial leads (V1-V3) and a J-point elevation on the ECG.[10]

Accesory conduction pathways An example of an accessory conduction pathway associated with syncope is Wolff-Parkinson-White (WPW) syndrome. WPW syndrome is a congenital disease where an accessory electrical pathway is present between the atria and ventricle.[12] Due to this accessory pathway electrical impulses can bypass the AV node and are not delayed which can cause tachyarrhythmias. Characteristic findings on the ECG of WPW include a short PR interval and a prolonged QRS complex with 'delta' wave.

2.3. Diagnostic evaluation of syncope

For the diagnosis of syncope, a careful anamnesis of the current attack with clinical history is very important.[2, 15] Additionally, hetero-anamnesis of an eyewitness can also be very helpful for distinguishing between syncope or non-syncopal loss of consciousness and the specific type of syncope. Alongside the anamnesis, physical examination including supine and standing BP measurements and ECG registration are also important. Reflex syncope can often be inferred or ruled out based on the patient's history. Orthostatic hypotension syncope can also be inferred from historical clues and confirmed through physical examination revealing orthostatic hypotension. Cardiac syncope becomes more plausible if there's a family history or underlying cardiac pathology. Confirmation of cardiac syncope can be made based on ECG findings indicative of arrhythmias or structural abnormalities. If initial evaluations through the patient history, physical examination, and ECG recordings are not enough to make a conclusive diagnosis further diagnostic tests can be deployed. Further examinations include tests such as a cardiac echo, table tilt test, carotid sinus massage, and blood tests.[2]

2.4. Management of cardiac syncope

The treatment of cardiac syncope depends upon identifying the underlying etiology.[2] For bradycardia-related syncope, the necessity for a pacemaker may arise. In cases involving tachycardia, an implantable cardioverter-defibrillator (ICD) might be indicated as part of the treatment strategy.

2.5. Deep learning

Deep learning (DL) is a subfield of machine learning that has gained significant attention due to its ability to learn from data and make classifications in various fields. For this reason, DL has also become of great interest in medical research. DL models are being applied to medical image analysis, surgical navigation strategies, and disease diagnosis.[1, 13] Deep neural networks are computational models inspired by the structure and function of human brains.[18] It is a network of multiple layers of interconnected neurons. Neurons are the basic units of the network and are inspired by the neurons of the brain. A neural network consisting of more than one hidden layer is called a deep neural network.

Contrary to machine learning approaches, DL does not require explicit feature extraction as a preprocessing step. Instead, large datasets are fed into the neural network directly and the weights of the network are updated during training via an optimization algorithm such that the loss function is minimized. The main advantage of this approach is that the performance of the model is not dependent on the quality of the input features. Because, as intuitively can be understood, if the input features are not relevant to the task, the performance of an Ml model will be poorer.

2.5.1. Residual neural network

The residual neural network (ResNet) is a type of DNN that overcomes one of the problems encountered with deeper neural networks, namely the gradient decrease with increasing depth of the network. This decrease in gradient during backpropagation causes the learning speed to decrease and thereby

slows down training.[18] A ResNet overcomes this problem by implementing a skip connection which allows the gradient to be directly backpropagated skipping over a convolution weight layer. During training, these skip connections initially allow the model to circumvent convolutional layers. As training progresses and the loss function is minimized, the model gradually learns to utilize the forward connection through these layers. This mechanism not only aids in capturing more meaningful features from the data but also facilitates a smoother flow of gradients, leading to more efficient training.

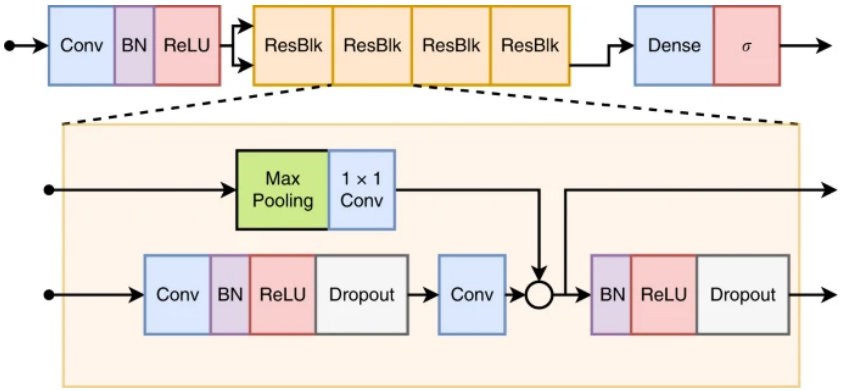


Figure 2.1: The unidimensional residual neural network architecture used for ECG classification. Ribeiro et al. (2020)

The current study adopts the ResNet model described in the paper of Ribeiro et al (2020) as a starting point as it was shown that this ResNet model could outperform cardiology resident medical doctors in recognizing 6 types of abnormalities in 12-lead ECG recordings, with F1 scores described above 80 %.[14] The network consists of a 1D convolution layer followed by batch normalization and ReLu activation. This is followed by four residual blocks with each residual block containing two convolution layers per block. Each convolution layer is followed by batch normalization for rescaling and ReLu activation. Dropout is performed after the ReLu activations in the residual block. The output of the last residual block is fed into a dense layer (consisting of fully connected neurons) with a sigmoid activation function. The sigmoid activation function was used for this multiclass problem as it is also a multilabel problem meaning that one instance can belong to multiple classes.

The model from the original paper was trained on 2,322,513 ECG records from 1,676,384 different patients in Brazil (CODE dataset).[14] The ResNet model could classify a first-degree AV block, right bundle branch block, left bundle branch block, sinus bradycardia, atrial fibrillation, and sinus tachycardia.

2.5.2. Efficient Channel Attention Module

The Efficient Channel Attention (ECA) module is a module that aims to improve the performance of a model by efficiently capturing relevant patterns by channel attention, while minimizing computational cost compared to the squeeze and excite module.[19] This enhances the ability of the model to make accurate predictions. For this purpose, the ECA module incorporates several key components. First,

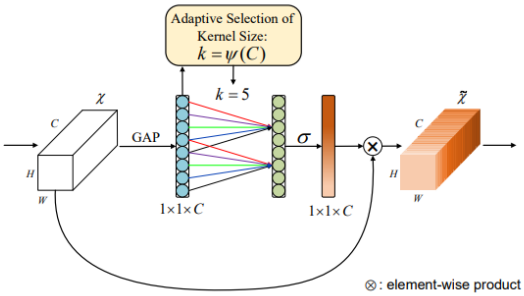


Figure 2.2: Efficient Channel Attention module. Wang et al. (20200

it applies global average pooling per channel without dimensionality reduction, allowing it to extract im-

portant information from each channel independently. Then local cross-channel interaction is captured by considering every channel and its k neighbors (by performing 1D convolution of size k). Kernel size k is proportional to channel dimension. This step enables the module to analyze the relationships between channels and identify relevant patterns. This is followed by a Sigmoid function which modulates the attention scores thereby emphasizing informative channels while suppressing irrelevant ones. This module adds very few additional parameters to the model and negligible computations, while it has been shown to significantly improve performance.[19]

2.5.3. Training

The concept of training for DNN is the process of determining the value of the weights in the network. For the purpose, of the current study a classification task: the input of the network will be ECG recording fragments and the output will be a vector of class probabilities. The goal during training is to determine the weights that maximize the probability of the correct classes and minimize the probabilities of incorrect classes.[18] The difference between the correct probability and the predicted probability by the model is referred to as the loss. During training the average loss is minimized using an optimization algorithm that updates the weights of the model to minimize the loss.[18]

There are various techniques to improve the efficiency and robustness of the training process and therefore improve the performance of the model. This is done by tuning the hyperparameters of the model. The hyperparameters which were tuned in the current study will briefly be discussed below:

- **Batch size**: Refers to the number of ECG recordings processed by the model before the weights are updated [18, 21]. A larger batch size can increase the speed of training.
- **Learning rate**: Determines the size of steps taken to update the weights of the model. A large learning rate means big steps and can result in quicker learning.[18, 21]
- **Dropout rate**: A regularization technique applied to prevent overfitting and thereby increase the generalization of the model. It controls the probability of neurons being randomly dropped during training.[18, 21]
- **Activation functions**: Introduce non-linearity into the neural network, allowing the model to learn complex patterns in data. Different activation functions have unique characteristics that can impact the model's ability to learn intricate patterns.[18, 21]
- **Number of epochs**: Refers to the number of times the model passes through the complete training dataset during training. During each epoch, the model's weights are updated after processing one batch.[18, 21]
- **Kernel size**: Refers to the dimension of the convolutional filter. A larger kernel size results in a larger filter, allowing for the extraction of more global features, while smaller filters can capture more detailed features.[18, 21]
- **Kernel initializer**: Refers to the technique used to initialize the values of the kernels before training starts. This can speed up training and influence the performance of the model.[18, 21]

2.5.4. Evaluation Metrics

The evaluation metrics deployed in this study will be shortly introduced. The F1 score, which represents the harmonic mean of precision and recall, provides a comprehensive measure of a model's accuracy in classification tasks. This performance score can also give an adequate evaluation of the model's performance when dealing with imbalanced datasets, such as in medical diagnosis. Recall (or Sensitivity) represents the proportion of true positive classifications (as it is calculated by the true positives divided by the true positives plus false negatives).[18, 21] This is an important evaluation measure in this study as it is crucial that cardiac causes of Syncope are not missed at the ED as this is associated with a poor prognosis. Precision, on the other hand, displays the model's ability to correctly identify positive cases out of all cases predicted as positive (true positives and false positives).[18, 21] It is important that this score does not drop too significantly in the context of this study as this would lead to many false positives and therefore unnecessary alarms. As there is always a trade-off between recall and precision for the goal of this study we will prioritize recall. Additionally, the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve assesses the model's ability to distinguish between classes also the ROC gives a great visualisation representation of the model performance.[18, 21]

3
Method

3.1. Dataset
For the current study, the 12-lead electrocardiogram (ECG) database, developed collaboratively by Chapman University, Shaoxing People's Hospital (affiliated with Zhejiang University School of Medicine), and Ningbo First Hospital, was utilized (Zheng et al., 2020). This database is publicly accessible on PhysioNet and comprises 12-lead ECG signals collected from 45,152 patients.[23] The signals were recorded over a duration of 10 seconds at a sampling rate of 500Hz. The amplitude unit of the ECG signal is microvolt, with each A/D bit corresponding to 4.88 volts and a 32-bit resolution for the A/D converter. Each ECG signal in the dataset was annotated for rhythm and cardiac conditions by two independent licensed physicians. In cases where discrepancies arose between the annotations provided by the physicians, a third senior physician intervened to determine the final label.[23] The labels assigned to each ECG recording were represented using SNOMED CT codes, along with a corresponding file containing the mapping between the original labels and the SNOMED CT code system.

3.2. Study population
The PhysioNet repository for the Chapman Ningbo dataset does not include patient characteristics. However, the original paper introducing the Chapman Ningbo dataset, reports 40,258 12-lead ECG records collected from 22,599 males and 17,659 females.[23] The current dataset contains an additional 4,894 ECG records of which it is assumed that the gender and age distribution is similar as reported in the original paper. These participants were randomly selected from individuals who visited Shaoxing People's Hospital and Ningbo First Hospital of Zhejiang University between 2013 and 2018. Regarding age distribution, the dataset predominantly includes individuals aged 51-60, 61-70, and 71-80 years, accounting for 19.8%, 61-70% and 71-80% of the dataset, respectively.

3.3. Preprocessing
Since all ECG recordings in the Chapman Ningbo dataset have consistent length and sampling frequency, no preprocessing steps were undertaken to modify these recordings. The objective of this study was to classify arrhythmias causing syncopic events and other abnormalities from ECG recordings, focusing on 45 out of the 110 diagnoses present in the dataset (Appendix A 1). To facilitate classification, the original SNOMED-CT coded labels corresponding to these 45 diagnoses were One-Hot encoded. This encoding method represents each diagnosis as a bit in a 45-bit long array.

3.4. Model Architecture
The ResNet model described by Ribeiro et al. (2020) was adopted as the foundation for the current study (see section 2.5.1 Residual neural network). This decision was motivated by the model's outstanding performance, surpassing that of cardiology resident physicians with F1 scores exceeding 80%. The original ResNet architecture was implemented without any modifications, aside from adjusting the input shape to match the dataset used and the output shape to accommodate the additional classes.

In an effort to enhance the ResNet model, an efficient channel attention module was introduced. This attention module was integrated into the original ResNet architecture, positioned after the last residual block. The rationale behind incorporating the channel attention module was to leverage its capabilities in selectively highlighting informative channels, thereby enhancing the model's ability to capture relevant features from the ECG signals.

3.5. Model development

The Chapman Ningbo dataset served as the foundation for model development. This dataset was split into a training set containing 38,379 recordings (85%) and a test set containing 6,773 recordings (15%). The split was achieved using the MultiLabelStratifiedShuffleSplit method from the iterative stratification library (Release 0.1.7) with a random state of 33. Subsequently, the training set was further divided into 10 folds using stratified K-split (StratifiedKFold from the Scikit Learn library, Version 1.4.1) with a random seed of 33. This stratification ensured that the distribution of labels across the subsets remained comparable. To address the class imbalance inherent in the dataset, class weights were computed using the balanced parameter (Scikit Learn library, Version 1.4.1). This approach ensured that each class contributed proportionately to the model's training process, mitigating the impact of class imbalance on model performance. Next, evaluation metrics were defined as binary accuracy, recall, precision, and auc for training the model. Additionally, two callbacks were implemented to optimize model training. The first callback reduced the learning rate with a factor of 0.1 if the validation loss did not decrease during 2 epochs. The second callback stopped training if the validation loss did not decrease by 0.0001 over 5 epochs. These callbacks were designed to prevent overfitting and terminate training if the model's performance did not improve any further. Next, parameter values were optimized using the training dataset with the Adam optimizer algorithm and the binary cross-entropy loss function. The optimal hyperparameter values were determined for batch size (16, 32, 64, 128), learning rate (0.05, 0.01, 0.005, 0.001), kernel size (8, 16, 24, 32), kernel initializer ('he normal', 'truncated normal', 'orthogonal'), dropout keep probability (0.5, 0.6, 0.7, 0.8, 0.9), and activation function ('ReLU', 'Sigmoid', 'Tanh'). The Bayesian Optimization method of Keras Tuner was employed for hyperparameter tuning, limiting the number of iterations to a maximum of 25 epochs per model. A total of 20 models with different hyperparameter configurations were evaluated during the tuning process.

3.6. Performance evaluation

The final ResNet model and the ResNet model with the ECA module were assessed on the test dataset. The primary outcome measures was the Recall and F1 score, calculated based on the model's performance on the test dataset. Secondary outcome measures included precision, and the area under the curve of the receiver operating characteristic (ROC) curve, all calculated based on the test set. These metrics were utilized to comprehensively evaluate the models' performance and assess their effectiveness in accurately classifying arrhythmias and other abnormalities in ECG recordings.

4

Method

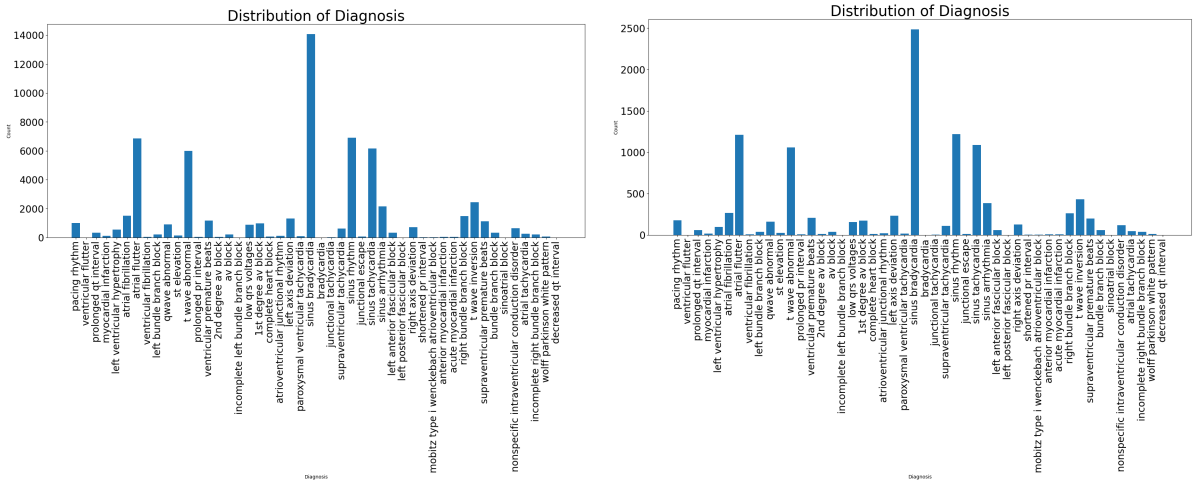
4.1. Chapman Ningbo dataset

Table 4.1: Demographic characteristics

Male (%	88.0%
Age (Years)	58 ± 19.8

*Mean ± standard deviation

The dataset characteristics information available is summarized in Table 4.1. A description of the dataset acquisition characteristics and annotation procedure is described in 3. The dataset was split into a training set containing 38,379 recordings (85%) and a test set containing 6,773 recordings (15%. The distribution of the labels in the dataset is shown in figure 4.1.



(a) Train dataset

(b) Test dataset

Figure 4.1: Distribution of labels over the dataset

4.2. Hyperparameter tuning

The results of the hyperparameter tuning are summarized in Table 4.2. From this table, it can be seen that for both models the found best hyperparameters were identical.

Table 4.2: Best hyperparameter values found using Keras tuner

Hyperparameter (%	ResNet	ResNet with ECA module
Kernel size	24	24
Kernel initializer	he normal	he normal
dropout keep probability	0.8	0.8
activation function	ReLu	ReLu
learning rate	0.001	0.001
batch size	32	32

*Mean $\pm$ standard deviation

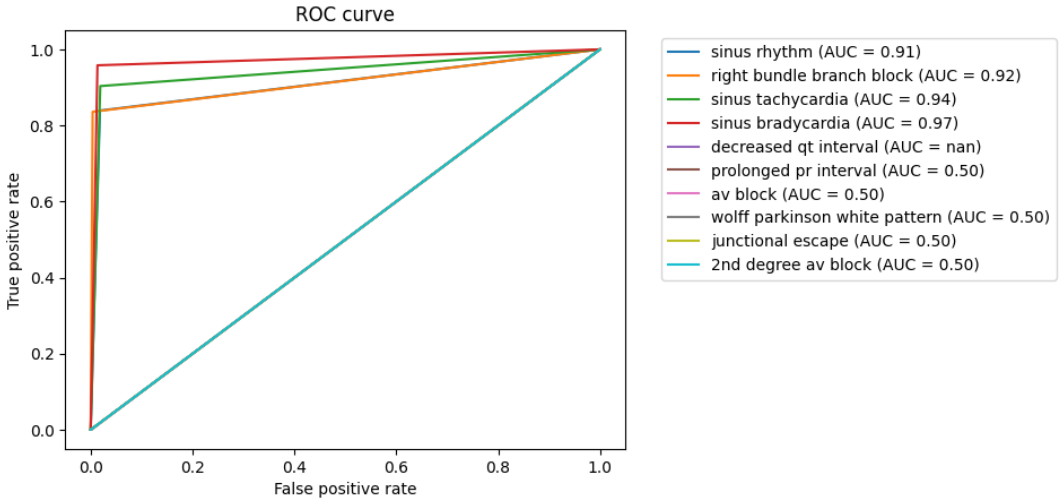


Figure 4.2: Best and worst ROC curves per label as classified by the model for ResNet model

4.3. ResNet model

Figure 4.2 displays the ROC curves for the labels with the highest and lowest AUC scores obtained by the ResNet model on the test dataset. From this figure, it can be seen that the model obtained high AUC curves for the labels sinus rhythm, right bundle branch block, sinus tachycardia, and sinus bradycardia. Additionally, the figure shows that the lowest AUC curves were obtained for the labels decreased qt interval, prolonged pr interval, AV block, Wolf Parkinson White pattern, junctional escape, and second-degree AV block. A figure containing the ROC curves of all labels can be found in Appendix 1a. Table 4.3 presents the classification scores of the ResNet model on the test dataset, including precision, recall, and F1-score per label. Additionally, the table provides the average score for each metric. Upon examination, it is evident that the average micro recall score is relatively low at 0.67, with a mediocre F1-score of 0.72. Focusing on individual recall scores for ECG diagnoses associated with the WOBBLER acronym for syncope, the following scores are observed: Wolf Parkinson White pattern 0.18, second-degree AV block 0.30, AV block 0.05, first-degree AV block 0.35, Mobitz type I Wenckebach atrioventricular block 0.00, left bundle branch block 0.44, left anterior fascicular block 0.35, left posterior fascicular block 0.00, right bundle branch block 0.84, left ventricular hypertrophy 0.75, prolonged QT interval 0.39, and decreased QT interval 0.00.

4.4. ResNet model with ECA attention module

In figure 4.4 the ROC curves for the labels with the highest and lowest AUC scores obtained by the ResNet model with the ECA module on the test dataset are displayed. From this figure, it can be seen that the model obtained high AUC curves for the labels sinus rhythm, right bundle branch block, sinus tachycardia, and sinus bradycardia. Additionally, the figure shows that the lowest AUC curves were obtained for the labels decreased qt interval, prolonged pr interval, AV block, Wolf Parkinson White pattern, junctional escape, and second-degree AV block. A figure containing the ROC curves of all

	precision	recall	f1-score	support
pacing rhythm	0.91	0.82	0.86	177
ventricular flutter	1.00	1.00	1.00	1
prolonged qt interval	0.33	0.39	0.36	59
myocardial infarction	0.68	0.72	0.70	18
left ventricular hypertrophy	0.48	0.75	0.59	97
atrial fibrillation	0.27	0.42	0.33	267
atrial flutter	0.79	0.48	0.59	1209
ventricular fibrillation	1.00	1.00	1.00	9
left bundle branch block	0.53	0.44	0.48	36
qwave abnormal	0.45	0.19	0.27	159
st elevation	0.19	0.38	0.26	26
t wave abnormal	0.61	0.42	0.50	1056
prolonged pr interval	0.60	0.75	0.67	8
ventricular premature beats	0.80	0.74	0.77	208
2nd degree av block	0.75	0.30	0.43	10
av block	1.00	0.05	0.10	37
incomplete left bundle branch block	0.00	0.00	0.00	1
low qrs voltages	0.48	0.54	0.51	156
1st degree av block	0.39	0.35	0.37	171
complete heart block	1.00	0.27	0.43	11
atrioventricular junctional rhythm	0.39	0.33	0.36	21
left axis deviation	0.56	0.50	0.52	232
paroxysmal ventricular tachycardia	0.65	0.69	0.67	16
sinus bradycardia	0.98	0.95	0.97	2484
bradycardia	0.00	0.00	0.00	1
junctional tachycardia	1.00	0.75	0.86	4
supraventricular tachycardia	0.58	0.84	0.69	109
sinus rhythm	0.92	0.85	0.88	1219
junctional escape	0.75	0.27	0.40	11
sinus tachycardia	0.92	0.92	0.92	1088
sinus arrhythmia	0.74	0.53	0.62	383
left anterior fascicular block	0.83	0.35	0.49	57
left posterior fascicular block	0.00	0.00	0.00	1
right axis deviation	0.46	0.18	0.26	128
shortened pr interval	1.00	0.33	0.50	3
mobitz type i wenckebach atrioventricular block	0.00	0.00	0.00	5
anterior myocardial infarction	0.67	0.22	0.33	9
acute myocardial infarction	1.00	0.43	0.60	7
right bundle branch block	0.94	0.84	0.89	262
t wave inversion	0.58	0.36	0.44	432
supraventricular premature beats	0.44	0.08	0.13	198
bundle branch block	0.89	0.14	0.24	58
sinoatrial block	0.00	0.00	0.00	1
nonspecific intraventricular conduction disorder	0.68	0.22	0.34	116
atrial tachycardia	0.39	0.16	0.22	45
incomplete right bundle branch block	1.00	0.05	0.10	37
wolff parkinson white pattern	1.00	0.18	0.31	11
decreased qt interval	0.00	0.00	0.00	0
micro avg	0.79	0.67	0.72	10654
macro avg	0.62	0.42	0.46	10654
weighted avg	0.78	0.67	0.70	10654
samples avg	0.78	0.74	0.73	10654

Figure 4.3: Classification report of ResNet model

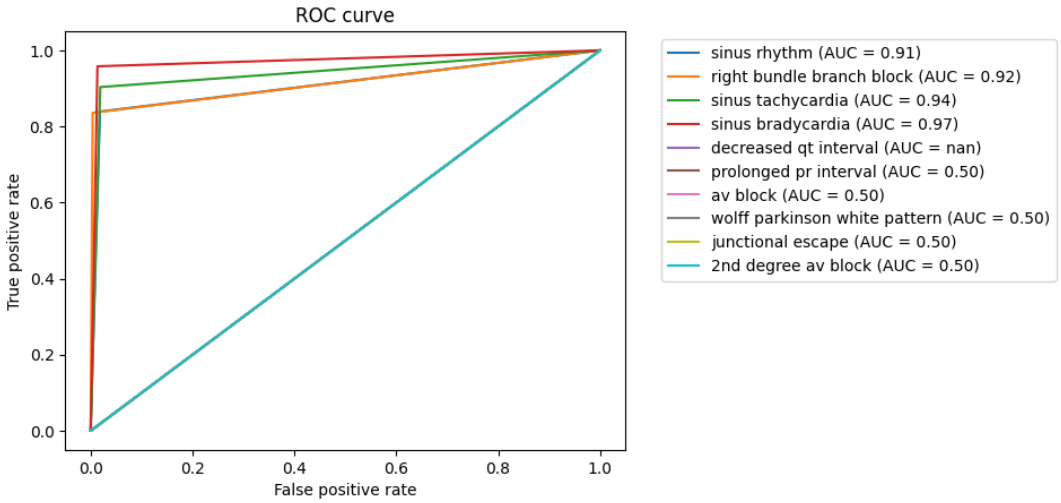


Figure 4.4: Best and worst ROC curves per label as classified by the model for ResNet model with ECA module

labels can be found in Appendix 1b. Table 4.3 presents the classification scores of the ResNet model with the ECA module on the test dataset, including precision, recall, and F1-score per label. Additionally, the table provides the average score for each metric. Upon examination, it is evident that the average micro recall score is relatively low at 0.63, with a mediocre F1-score of 0.70. Focusing on individual recall scores for ECG diagnoses associated with the WOBBLER acronym for syncope, the following scores are observed: Wolf Parkinson White pattern 0.00, second-degree AV block 0.30, AV block 0.03, first-degree AV block 0.06, Mobitz type I Wenckebach atrioventricular block 0.00, left bundle branch block 0.36, left anterior fascicular block 0.16, left posterior fascicular block 0.00, right bundle branch block 0.79, left ventricular hypertrophy 0.71, prolonged QT interval 0.20, and decreased QT interval 0.00.

	precision	recall	f1-score	support
pacing rhythm	0.89	0.77	0.82	177
ventricular flutter	1.00	1.00	1.00	1
prolonged qt interval	0.33	0.20	0.25	59
myocardial infarction	0.54	0.83	0.65	18
left ventricular hypertrophy	0.45	0.71	0.55	97
atrial fibrillation	0.26	0.35	0.30	267
atrial flutter	0.77	0.49	0.60	1209
ventricular fibrillation	1.00	0.89	0.94	9
left bundle branch block	0.33	0.36	0.35	36
qwave abnormal	0.51	0.25	0.33	159
st elevation	0.16	0.27	0.20	26
t wave abnormal	0.60	0.34	0.44	1056
prolonged pr interval	0.25	0.75	0.38	8
ventricular premature beats	0.78	0.68	0.73	208
2nd degree av block	0.75	0.30	0.43	10
av block	1.00	0.03	0.05	37
incomplete left bundle branch block	0.00	0.00	0.00	1
low qrs voltages	0.44	0.54	0.48	156
1st degree av block	0.17	0.06	0.09	171
complete heart block	0.43	0.27	0.33	11
atrioventricular junctional rhythm	0.15	0.24	0.18	21
left axis deviation	0.55	0.41	0.47	232
paroxysmal ventricular tachycardia	0.53	0.56	0.55	16
sinus bradycardia	0.97	0.95	0.96	2484
bradycardia	0.00	0.00	0.00	1
junctional tachycardia	0.43	0.75	0.55	4
supraventricular tachycardia	0.62	0.83	0.71	109
sinus rhythm	0.92	0.81	0.86	1219
junctional escape	0.43	0.27	0.33	11
sinus tachycardia	0.93	0.91	0.92	1088
sinus arrhythmia	0.71	0.41	0.52	383
left anterior fascicular block	0.69	0.16	0.26	57
left posterior fascicular block	0.00	0.00	0.00	1
right axis deviation	0.46	0.09	0.16	128
shortened pr interval	0.00	0.00	0.00	3
mobitz type i wenckebach atrioventricular block	0.00	0.00	0.00	5
anterior myocardial infarction	0.67	0.22	0.33	9
acute myocardial infarction	1.00	0.29	0.44	7
right bundle branch block	0.89	0.79	0.84	262
t wave inversion	0.59	0.32	0.42	432
supraventricular premature beats	0.27	0.06	0.10	198
bundle branch block	1.00	0.02	0.03	58
sinoatrial block	0.00	0.00	0.00	1
nonspecific intraventricular conduction disorder	0.42	0.07	0.12	116
atrial tachycardia	0.36	0.11	0.17	45
incomplete right bundle branch block	0.00	0.00	0.00	37
wolff parkinson white pattern	0.00	0.00	0.00	11
decreased qt interval	0.00	0.00	0.00	0
micro avg	0.79	0.63	0.70	10654
macro avg	0.48	0.36	0.37	10654
weighted avg	0.76	0.63	0.67	10654
samples avg	0.76	0.71	0.71	10654

Figure 4.5: Classification report of ResNet model with ECA module

5 Discussion

In the current study it was found that out of the 12 diagnoses present in the Chapman Ningbo dataset that are related to the WOBBLER acronym for ECG features of cardiac syncope, only two were classified by the described ResNet models with reasonable recall scores. Specifically, the right bundle branch block achieved a recall score of 0.84 in the ResNet model and 0.79 in the RNN model with an ECA module, while left ventricular hypertrophy attained scores of 0.75 and 0.71, respectively. However, for the remaining diagnoses, the performance was suboptimal. Thereby, it should be concluded that the ResNet models described in the current study are not suitable for deployment in clinical practice for detecting cardiac causes of Syncope from a 12-leads ECG. Additionally, it's noteworthy that the original ResNet model outperformed its counterpart with an ECA module across evaluation metrics with an average F1-score of 0.72 compared to 0.70 and a recall score of 0.67 compared to 0.63.

5.1. Comparison to current literature

This study has been unable to demonstrate the similar high evaluation scores obtained by Ribeiro et al (2020) using the same ResNet model.[14] Ribeiro's paper reported F1 scores per class exceeding 0.87 and recall scores per class surpassing 0.77. Several factors may have contributed to the significantly lower evaluation scores obtained in the models described in this study. The first and probably most important factor is that the Chapman Ningbo dataset utilized in this study comprised of just 45,152 ECG recordings, whereas Ribeiro et al. deployed a dataset of roughly 2,3 million recordings. This difference in dataset size could be the reason the models described in this study underperformed compared to the model described by Rebeiro which had the same model architecture, because the model did not have as many recordings to train upon. Additionally, while Ribeiro et al. attempted to classify only six different diagnoses from the ECG recordings, our study aimed to classify 48 different diagnoses. This larger number of classes poses a greater challenge for the model to learn discriminative features for each class, particularly if the classes share commonalities. This is also not helped by the fact that this model was trained on a smaller dataset. This complexity could have contributed to the diminished performance observed in this study.

Contrary to findings from previous studies, this study found an unexpected result regarding the performance of the ResNet model incorporating an ECA module.[8] In contrast to the anticipated enhancement from channel attention mechanisms, our study showed that the ResNet model with the ECA module did not outperform the original ResNet model. This finding is surprising, given that the incorporation of channel attention typically improves the performance of neural networks by selectively weighting feature maps based on context and thereby focussing the attention of the network to relevant features for classification. One possible explanation for this discrepancy could be the application of the ECA module solely after the four Residual blocks, rather than after each Residual block, as described in some studies.[8]

5.1.1. Limitations

The primary objective of this study was to detect cardiac causes of syncope from 12-lead ECG recordings. With this goal in mind, the main focus in enhancing the performance of the ResNet model was

improving the recall score As improving the recall score signifies a reduction in false negative classifications, This reduction implies that if the model were deployed in clinical practice, fewer instances of cardiac causes of syncope would be missed. This is desired because of the bad prognosis of people with a cardiac cause of syncope when this is missed at the ED.[9] One potential approach to enhancing the recall score involves lowering the classification threshold, allowing the model to classify more instances as true positives. This strategy was considered but as the other evaluation scores such as the F1 score and precision in the current ResNet models are not acceptable, improving the recall by lowering the threshold would not mean the the model suddenly performs well enough to be deployed in clinical practice.

5.2. Future research and clinical perspective

A few options that could be explored to improve the performance of the model in the future will be discussed next. Firstly, increasing the size of the dataset on which the model is trained is a promising approach. It is widely recognized that neural networks tend to perform better and exhibit greater generalizability when trained on larger datasets.[18, 1] A larger dataset would provide the model with more diverse recordings, including arrhythmias that are less commonly encountered. Consequently, the model would have a better opportunity to learn features unique to these arrhythmias as it has more instances to learn from, thereby improving its overall performance. Furthermore, some labels were underrepresented in the dataset used, and certain ECG findings associated with cardiac causes of syncope, such as the epsilon wave and Brugada syndrome, were absent. As a result, the network did not have the opportunity to learn these features, highlighting the importance of a more comprehensive dataset for training. Additionally, it is crucial to use clear and standardized guidelines and terminology for labeling ECG data when developing new 12-lead ECG datasets so that one can deploy multiple datasets without having to make multiple adjustments.

Another method worth exploring for improving the performance of the model is increasing the depth of the ResNet model by adding more Residual blocks to the network. Research has demonstrated that increasing the depth of neural networks enhances model performance, as it allows the model to learn more intricate features.[18]

Additionally, adapting the workflow to a two-step process instead of the single-step process deployed in this study should be considered. In the two-step approach, the first step involves distinguishing between normal and abnormal ECG recordings, while the second step focuses on classifying abnormal ECG recordings into specific diagnoses. This method naturally helps reduce class imbalance, as the largest class comprises normal ECG recordings. Therefore, if the model can effectively distinguish between normal and abnormal recordings in the first step, the subsequent model can learn from a more balanced dataset. This could improve the overall classification performance as the bias towards the majority class of a model trained on an imbalanced dataset will be smaller in a dataset that is at least slightly more balanced.

Finally, by analyzing the nature of misclassifications, targeted strategies can be implemented. For instance, it may be discovered that the model struggles to differentiate between various types of AV blocks, leading to misclassification errors such as classifying a first-degree AV block as a second-degree AV block. If a majority of misclassifications occur within the same group of arrhythmias, it may be beneficial to consider grouping these together or potentially transitioning to a two- or even three-step approach in which the arrhythmias are first classified as belonging to a specific group and afterward a distinction within the group is made.

In addition to the methods mentioned for improving performance, several considerations should be taken into account before deploying such a model in clinical practice. One noteworthy aspect is the incorporation of explainability of the classifications made by the model. Enhancing model explainability enhances the model's deployability within clinical practice and could potentially improve model performance.[13] By delving into the rationale behind the model's decisions and understanding which features are utilized, clinicians can assess the logical coherence of the model's classifications. This helps physicians build trust in the model and thereby acceptance in clinical practice. Additionally, it can help identify errors made by the model for which targeted corrections and improvements can be made which could lead to improved performance. It's important to exercise caution before implementing such a model in clinical practice. Research suggests that faulty automated interpretations can result in inaccurate physician overreads, whereas accurate interpretations are correlated with dependable

physician assessments.[17] A similar effect could be observed when the ResNet model is deployed in clinical practice and classifies an ECG recording as a certain arrhythmia. Therefore, it is crucial when implementing the model in clinical settings to ensure that medical professionals understand its limitations and are aware of this phenomenon. Before deploying such a model for detecting cardiac causes of syncope from 12-lead ECG recordings in clinical practice, an additional complication that needs to be addressed is the continuous extraction of ECG signals from Philips monitors. These monitors are secured, and extracting the signal requires a software license from ixTrend. Future work should look into how to incorporate this program or ensure its compatibility with the applications deploying a model for assessing the ECG signal that needs to be developed. This is crucial for the deployment of an ResNet model for detection of cardiac causes of Syncope in clinical practice.

Conclusion

The ResNet models described in the current study are not suitable for deployment in clinical practice due to their inadequate recall and F1 scores. As it stands, deploying these models would not contribute to the goal of automatically detecting cardiac causes of syncope from ECG signals, nor would it enhance the care and prognoses of these patients. In theory, such models could potentially support emergency departments (EDs) in monitoring for cardiac causes of syncope. However, before deployment in clinical settings, it is imperative to address the limitations associated with the model's performance. This includes improving recall and F1 scores to ensure reliable and accurate detection of cardiac-related syncope from ECG signals.

References

[1] Laith Alzubaidi et al. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. In: *Journal of Big Data* 8.1 (Dec. 2021), p. 53. ISSN: 21961115. DOI: 10.1186/S40537-021-00444-8. URL: /pmc/articles/PMC8010506/%20/pmc/articles/PMC8010506/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8010506/.

[2] Michele Brignole et al. “2018 ESC Guidelines for the diagnosis and management of syncope”. In: *European Heart Journal* 39.21 (June 2018), pp. 1883–1948. ISSN: 0195-668X. DOI: 10.1093/eurheartj/ehy037. URL: https://doi.org/10.1093/eurheartj/ehy037.

[3] Fabrizio D’Ascenzo et al. “Incidence, etiology and predictors of adverse outcomes in 43,315 patients presenting to the Emergency Department with syncope: An international meta-analysis”. In: *International Journal of Cardiology* 167.1 (July 2013), pp. 57–62. ISSN: 01675273. DOI: 10.1016/j.ijcard.2011.11.083. URL: http://www.internationaljournalofcardiology.com/article/S0167527311021401/fulltext%20http://www.internationaljournalofcardiology.com/article/S0167527311021401/abstract%20https://www.internationaljournalofcardiology.com/article/S0167-5273(11)02140-1/abstract.

[4] Wael Dakkak and Rami Doukky. “Sick Sinus Syndrome”. In: *StatPearls* (July 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK470599/.

[5] Christopher Foth et al. “Ventricular Tachycardia”. In: *StatPearls* (July 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK532954/.

[6] Fouad Jabbour and Arun Kanmanthareddy. “Sinus Node Dysfunction”. In: *StatPearls* (July 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK544253/%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC483009.

[7] Anthony H. Kashou et al. “Atrioventricular Block”. In: *StatPearls* (Feb. 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK459147/.

[8] Sander R Klomp et al. “Performance-Efficiency Comparisons of Channel Attention Modules for ResNets”. In: *Neural Processing Letters* 55 (2023), pp. 6797–6813. DOI: 10.1007/s11063-023-11161-z. URL: https://doi.org/10.1007/s11063-023-11161-z.

[9] Esther M. Mizrachi and Kranthi K. Sitammagari. “Cardiac Syncope”. In: *StatPearls* (Apr. 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK526027/%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5866177.

[10] Yukiko Nakano and Shimizu Wataru. “Syncope in patients with inherited arrhythmias”. In: *Journal of Arrhythmia* 33.6 (Dec. 2017), pp. 572–578. ISSN: 18832148. DOI: 10.1016/J.JOA.2017.07.007.

[11] Laryssa Patti and William G. Gossman. “Supraventricular Tachycardia”. In: *StatPearls* (Aug. 2023). URL: https://www.ncbi.nlm.nih.gov/books/NBK441972/.

[12] Tam Dan N. Pham and Mark E. Alexander. “Wolff-Parkinson-White Syndrome”. In: *Exercise Physiology for the Pediatric and Congenital Cardiologist* (Aug. 2023), pp. 227–234. DOI: 10.1007/978-3-030-16818-6{_}31. URL: https://www.ncbi.nlm.nih.gov/books/NBK554437/.

[13] P Rajpurkar et al. “AI in health and medicine”. In: *nature.comP Rajpurkar, E Chen, O Banerjee, EJ TopolNature medicine, 2022•nature.com* (). URL: https://www.nature.com/articles/s41591-021-01614-0.

[14] Antônio H. Ribeiro et al. “Automatic diagnosis of the 12-lead ECG using a deep neural network”. In: *Nature Communications 2020 11:1* 11.1 (Apr. 2020), pp. 1–9. ISSN: 2041-1723. DOI: 10.1038/s41467-020-15432-4. URL: https://www.nature.com/articles/s41467-020-15432-4.

[15] LLOYD A. RUNSER, ROBERT L. GAUER, and ALEX HOUSER. “Syncope: Evaluation and Differential Diagnosis”. In: *American Family Physician* 95.5 (Mar. 2017), pp. 303–312. URL: https://www.aafp.org/pubs/afp/issues/2017/0301/p303.html.

[16] Muhammad Asim Shabbir et al. “Bifascicular block in unexplained syncope is underrecognized and under-evaluated: A single-center audit of ESC guidelines adherence”. In: *PLoS ONE* 17.2 (Feb. 2022). ISSN: 19326203. DOI: 10.1371/JOURNAL.PONE.0263727. URL: /pmc/articles/PMC8884493/%20/pmc/articles/PMC8884493/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8884493/.

[17] Stephen W Smith et al. “A deep neural network learning algorithm outperforms a conventional algorithm for emergency department electrocardiogram interpretation”. In: (2018). DOI: 10.1016/j.jelectrocard.2018.11.013. URL: https://doi.org/10.1016/j.jelectrocard.2018.11.013.

[18] Vivienne Sze et al. “Efficient Processing of Deep Neural Networks: A Tutorial and Survey”. In: ().

[19] Qilong Wang et al. “ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Oct. 2019), pp. 11531–11539. ISSN: 10636919. DOI: 10.1109/CVPR42600.2020.01155. URL: https://arxiv.org/abs/1910.03151v4.

[20] Steven E. Williams, Mark O’Neill, and Irum D. Kotadia. “Supraventricular tachycardia: An overview of diagnosis and management”. In: *Clinical Medicine* 20.1 (2020), p. 43. ISSN: 14734893. DOI: 10.7861/CLINMED.CME.20.1.3. URL: /pmc/articles/PMC6964177/%20/pmc/articles/PMC6964177/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6964177/.

[21] Tong Yu and Hong Zhu. “Hyper-Parameter Optimization: A Review of Algorithms and Applications”. In: (Mar. 2020). URL: https://arxiv.org/abs/2003.05689v1.

[22] Munir Zaqqa and Ali Massumi. “Neurally Mediated Syncope”. In: *Texas Heart Institute Journal* 27.3 (2000), p. 268. ISSN: 07302347. URL: /pmc/articles/PMC101078/%20/pmc/articles/PMC101078/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC101078/.

[23] Jianwei Zheng et al. “A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients”. In: *Scientific Data* 7.1 (Dec. 2020). ISSN: 20524463. DOI: 10.1038/S41597-020-0386-X.

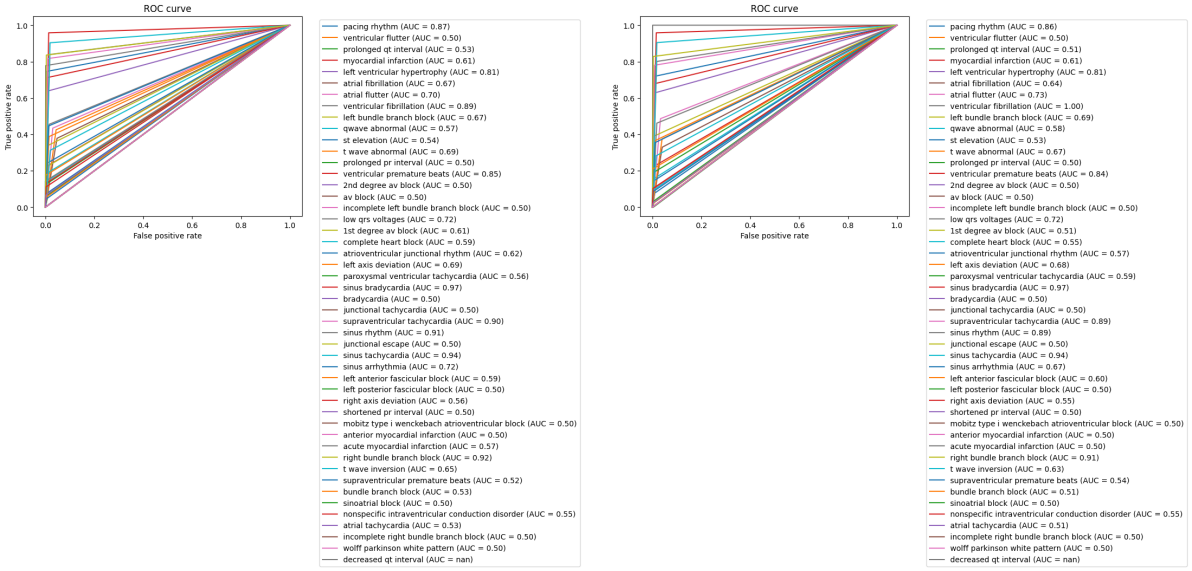
Appendix A

Dx	SNOMED CT Code	Abbreviation
1st degree av block	270492004	IAVB
atrial fibrillation	164889003	AF
atrial flutter	164890007	AFL
bradycardia	426627000	Brady
complete right bundle branch block	713427006	CRBBB
incomplete right bundle branch block	713426002	IRBBB
left anterior fascicular block	445118002	LAnFB
left axis deviation	39732003	LAD
left bundle branch block	164909002	LBBB
low qrs voltages	251146004	LQRSV
nonspecific intraventricular conduction disorder	698252002	NSIVCB
pacing rhythm	10370003	PR
premature atrial contraction	284470004	PAC
premature ventricular contractions	427172004	PVC
prolonged pr interval	164947007	LPR
prolonged qt interval	111975006	LQT
qwave abnormal	164917005	QAb
right axis deviation	47665007	RAD
right bundle branch block	59118001	RBBB
sinus arrhythmia	427393009	SA
sinus bradycardia	426177001	SB
sinus rhythm	426783006	SNR
sinus tachycardia	427084000	STach
supraventricular premature beats	63593006	SVPB
t wave abnormal	164934002	TAb
t wave inversion	59931005	TInv
ventricular premature beats	17338001	VPB
2nd degree av block	195042002	IIAVB
acute myocardial infarction	57054005	AMI
acute myocardial ischemia	413444003	AMIs
anterior ischemia	426434006	AnMIs
anterior myocardial infarction	54329005	AnMI
atrial fibrillation and flutter	195080001	AFAFL
atrial tachycardia	713422000	ATach
atrioventricular junctional rhythm	29320008	AVJR
av block	233917008	AVB
brady tachy syndrome	74615001	BTS
bundle branch block	6374002	BBB
cardiac dysrhythmia	698247007	CD
chronic atrial fibrillation	426749004	CAF
chronic myocardial ischemia	413844008	CMI
complete heart block	27885002	CHB
congenital incomplete atrioventricular heart block	204384007	CIAHB
diffuse intraventricular block	82226007	DIB
heart failure	84114007	HF
heart valve disorder	368009	HVD
incomplete left bundle branch block	251120003	ILBBB

inferior ischaemia	425419005	IIs
junctional escape	426995002	JE
junctional tachycardia	426648003	JTach
lateral ischaemia	425623009	LIs
left posterior fascicular block	445211001	LPFB
mobitz type i wenckebach atrioventricular block	54016002	MoI
myocardial infarction	164865005	MI
myocardial ischemia	164861001	MIs
paroxysmal atrial fibrillation	282825002	PAF
paroxysmal supraventricular tachycardia	67198005	PSVT
paroxysmal ventricular tachycardia	425856008	PVT
rapid atrial fibrillation	314208002	RAF
sinoatrial block	65778007	SAB
sinus node dysfunction	60423000	SND
st elevation	164931005	STE
supraventricular tachycardia	426761007	SVT
transient ischemic attack	266257000	TIA
ventricular fibrillation	164896001	VF
ventricular flutter	111288001	VFL
ventricular tachycardia	164895002	VTach
wolff parkinson white pattern	74390002	WPW
Brugada Syndrom	1204168008	Brug
shortened pr interval	49578007	SPRI
left ventricular hypertrophy	164873001	LVH
decreased qt interval	77867006	SQT

Table 1: Classification Labels SNOMED-CT mappings

Appendix B



(a) ResNet model

(b) ResNet model with ECA module

Figure 1: ROC curves per label obtained on test dataset