COM3025 Deep Learning and Advanced AI – Detecting Abnormal Heartbeats

[redacted], Rui Santos and [redacted] Computer Science (BSc) Year 3 - 2020

Introduction

More than 2 million people in the UK experience heart rhythm problems at some point in their lifetime.[nhs,2018] An ECG (Electrocardiogram) records the electrical activity of a person's heart and can show rhythm irregularities. As it is reviewed by doctors, in some rare cases there are chances for human error to occur and abnormalities go unnoticed. An ECG can also only show a person's condition at that moment in time and it is very difficult, if not impossible for doctors to predict the changes in a heart over time. It is therefore possible for someone with an irregular heartbeat to be categorised as normal and then later suffer serious complications

Problem Background – The Heart and ECG

Having a heart arrhythmia can put you at greater risk of developing serious and life threatening conditions. For example, someone living with atrial fibrillation is five times more likely to have a stroke than someone with a normal heart rhythm.[nhs,2018] There are also certain types of arrhythmia that can cause sudden cardiac death. In the UK alone 100,000 people die each year from this, some of which could have been prevented had the arrhythmia been diagnosed sooner.[nhs,2018]

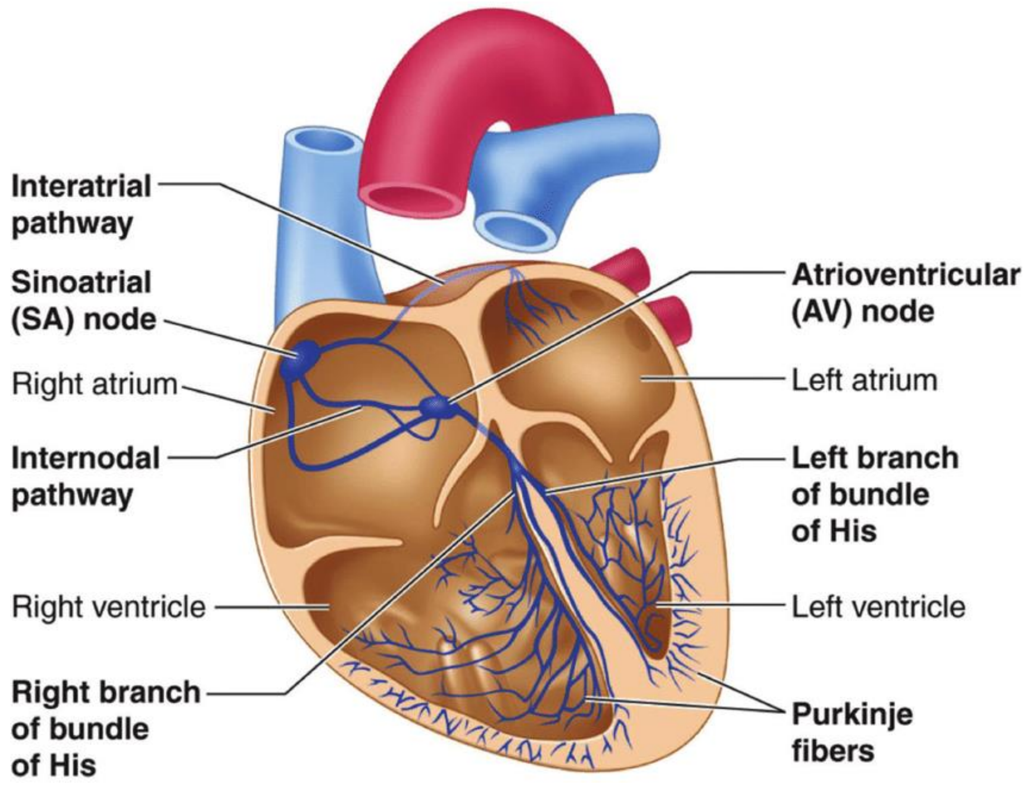


Fig 1. Electrical conduction system of the heart [Ganesan et al, 2016]

An Introduction to the Electrical Function of the Heart

The Heart's electrical function is responsible for regulating the rhythm and rate of your heartbeat. Your heart rate is the number of beats in one minute and the rhythm is the synchronised pumping of your four heart chambers. ["Electrical System of the Heart", n.d].

The Process

The first step to your heartbeat starts in a group of cells called the sinoatrial node (SA node aka 'The Pacemaker of the Heart). These cells generate and send a signal through your heart to make the chambers of your heart contract in turn. The top two chambers (Right and Left Atrium) contract first. The signal then reaches the atrioventricular node (AV node) which slows the signal. This causes a brief delay to allow blood to pump from the atria to the ventricles. The signal is then sent to the ventricles which contract in the same way as the atrium. This allows blood to be pumped out of the heart and to the rest of the body. The right ventricle pumps blood to your lungs and the left pumps blood to the rest of your body. After all of this has happened, each part of the heart performs a 'reset' ready for the cycle to happen again. This is one heartbeat.

Analysing a Heartbeat Cycle

- **P-Wave –** Represents the contraction of the two upper chambers of the heart (Atria).
- **PR Interval –** This represents the time taken from the signal to travel from the SA node through the AV node. The flat line in between the P-wave and the QRS complex is the AV node slowing the signal causing a brief delay. This is normal as described above.
- **QRS Complex –** This represents the contraction of the left and right ventricles. As they have a larger muscle mass than the atria, there is a larger amplitude than the P-wave.
- **T-Wave –** The T-wave is formed at the end of the cycle and is the process where the ventricles return to their resting state in order for the process to begin again.
- **QT and ST –** The ST segment represents the period of time when the ventricles are contracted. The QT interval is the time from the beginning of QRS until the end of the T-Wave. ("Electrical conduction system of the heart", 2020)

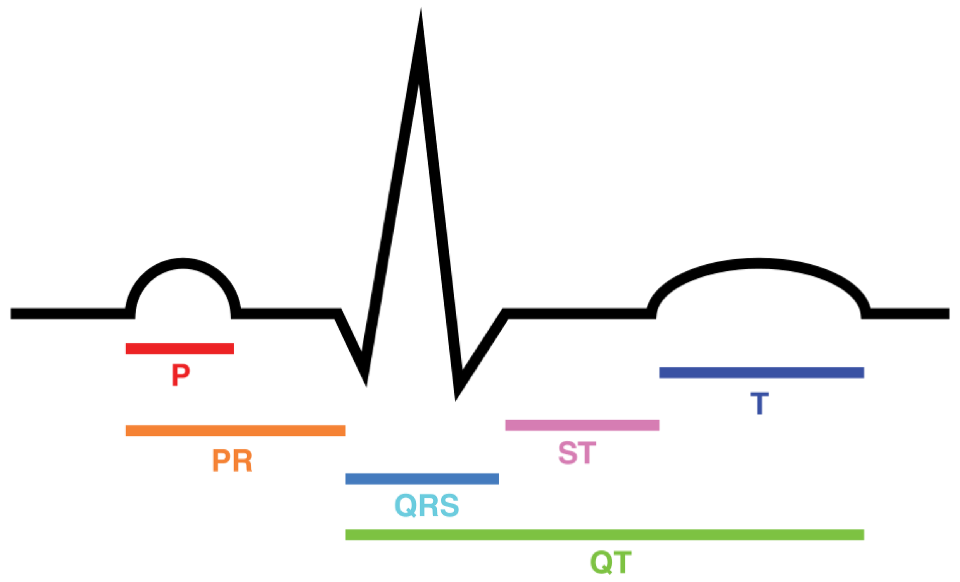


Fig 2. Electrical conduction system of the heart [Drricksanchez, 2020]

What is an ECG?

An Electrocardiogram (ECG) is a graph produced during an ECG test that shows the electrical activity of the heart. It measures voltage levels over time and produces a waveform similar to that of fig. 2. The test is conducted by sticking small sensors (electrodes) on the chest, arms and legs of the patient. These are attached to a machine that records the data and produces the graph. The graph is then typically reviewed by a doctor to see if there are any abnormalities. The components described above are all found in a typical ECG.

Problem Background – Arrhythmias

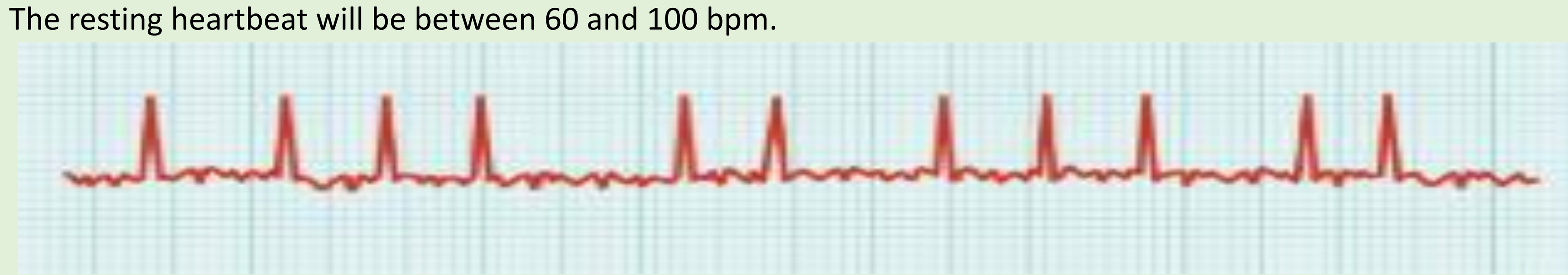
Arrhythmia is where the heart beats with an irregular rhythm. When this occurs it will produce an abnormal ECG. The main types of arrhythmia are:

- **Atrial Fibrillation -** This is where the heart beats abnormally fast. This occurs when the atria contract randomly and so fast that they do not properly relax before the next contraction. This reduces the efficiency of the heart. The random contractions are often caused by abnormal electrical impulses which override the SA node. The SA node can then no longer control the rhythm of the heart. The cause of this is not yet fully understood.
- **Supraventricular tachycardia (SVT)** – Similar to Atrial Fibrillation, the main symptom of SVT is an abnormally rapid heartbeat. However, in this instance it occurs in episodes and can slow down extremely abruptly after each episode. The causes for this differ, but the main causes all relate to the electrical system of the heart not functioning normally. This normally relates to having extra/ missing electrical pathways in between the atria and the ventricles. If treated, SVT is not normally serious.
- **Bradycardia –** This is where the heart beats too slowly. Normally, anything lower than 60bpm qualifies as bradycardia. However, there are some exceptions such as deep sleep and athletes. The main causes are: problems with the SA node, problems in the electrical pathways, metabolic problems (such as low thyroid syndrome), Damage from heart disease or a heart attack or certain heart medications that can cause a slow heartbeat as a side effect.
- **Ventricular Fibrillation –** This is the most serious type of arrhythmia and causes rapid, unpredictable heart beats. It often comes at extremely short notice and can cause sudden death if not treated immediately. The chambers in your heart beat so fast that they simply vibrate instead of pumping blood to your body. This in turn starves all of your organs of blood and makes your blood pressure plummet. The person will collapse in seconds of ventricular fibrillation starting and they will require cardiopulmonary resuscitation (CPR) and shocks to the heart with an automated external defibrillator (AED). This treatment will need to be implemented immediately for any chance of survival. The cause isn't always known, however it often comes as a result of a previous heart attack or an undiagnosed heart condition.

Some Key ECG patterns

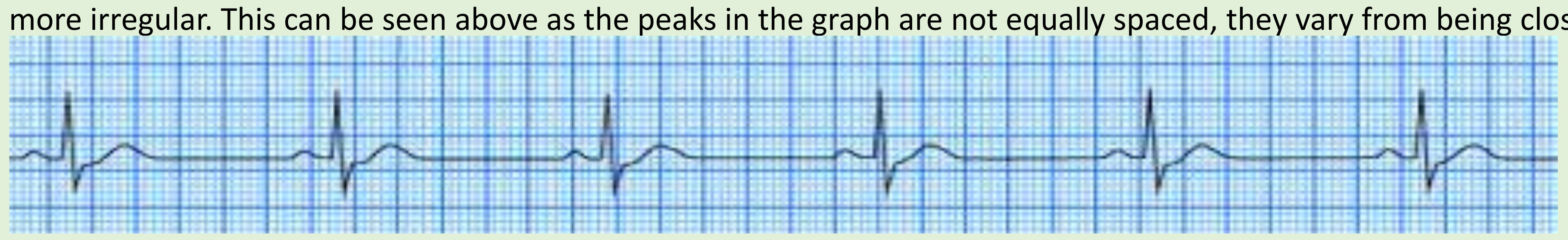
Sinus Rhythm

As seen above, a sinus rhythm will produce a clear regular waveform where all of the components are clearly seen. The resting heartbeat will be between 60 and 100 bpm.



Atrial Fibrilation (AF)

In this instance there will be a much higher heart rate and the wavelength (distance between beats) will be much more irregular. This can be seen above as the peaks in the graph are not equally spaced, they vary from being close



Bradycardia

The waveform produced will be similar to sinus rhythm, however the heart rate will be much slower. The distance between peaks will therefore be greater as the time between each heartbeat increases.

Ventricular Fibrillation (VF)

VF produces an extremely erratic graph where there are no clear components of the heartbeat displayed. The distance between each beat is minimal as the heart is beating so fast. The amplitude is also quite small.

Project Outline

The overall aim for the project is to develop a system that can analyse a given ECG and determine any abnormalities. We hope to be able to implement multiple methods and determine which is best. At first we aim to be able to be able to build a simple classifier that will determine whether the ECG is normal or abnormal. We will then build on this to detect all of the arrhythmias specified previously. We will use multiple different models using both plotted ECG images and time series data. We hope that this type of system would be of benefit as it could significantly reduce the time taken to review the ECG and potentially discover something that a human may miss, therefore eliminating any chance of human error.

Objectives

1. Build Binary Classifier to determine whether an ECG is normal or not.
2. Build a Multi-Class Classifier to determine which class the ECG falls into.
3. Build upon the models by implementing measures to improve the accuracy. This includes (but not limited to):
   - Experimenting with the data to make it better to train with.
   - Using more advanced techniques such as Transfer Learning.

Deliverables

1. Binary Classifier Model
2. 1DCNN Multi-Class Classifier
3. Transfer Learning model (Using Binary to train, utilised on Multi-Class)
4. Poster

COM3025 Deep Learning and Advanced AI – Detecting Abnormal Heartbeats

[redacted], Rui Santos and [redacted] Computer Science (BSc) Year 3 - 2020

Literature Review / Related Work

What is CNN: (Rawat and Wang, 2017)
A Convolutional Neural network can be described as a "Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other." This means that "with enough training, CNNs have the ability to learn these filters/characteristics." The architecture resembles the "connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area."
What is Binary & Multiclass Classification: (Binary Classification - an overview | ScienceDirect Topics, 2020)
Binary classification involves "assigning an individual to one of two categories, by measuring a series of attributes. An example is medical diagnosis for a single medical condition (say disease vs. no disease) based on a battery of tests." Whereas, (1.12. Multiclass and multilabel algorithms — scikit-learn 0.23.1 documentation, 2020)
Multiclass classification involves assigning individuals to a class when there are more than two classes available. "Each sample can only be labelled as one class."
What is Transfer Learning: (Brownlee, 2020)
"Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.
It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems."
Other work in the field:
Paper 1 (Sherin et al. 2018) proposes an "accurate and robust approach using deep learning for single led ECG classification." The research provided in the paper describes that "the most successful type of deep learning models are restricted Boltzmann machines, stacked autoencoder, Convolutional Neural Networks and Deep Belief Networks." Reference 43 in the paper further shows this by outlining "a CNN based classification system was developed which automatically learns a suitable feature representation from two lead ECG data, thereby negating the need of hand-crafted features." Furthermore, in 42, "segmented ECGs are processed by an eleven-layer convolutional neural network resulting in maximum accuracy of 93.18% using short duration ECG data." Further CNN architectures are suggested on reference 44, by employing a "1-D convolutional 7 neural networks (CNNs) that fused the feature extraction and classification for two-lead ECG classification. For each patient, an individual and simple CNN was trained by using relatively small common and patient-specific training data, and such patient-specific feature extraction ability did provide an improvement in the classification performance."

Paper 2 (Sannino et al. 2018) presents "present a novel deep learning approach for ECG beat classification" using Tensor Flow framework and the deep learning library from Google, in Python.
The research on that paper suggests that heartbeat categorization can be categorized into "intra-patient" and "inter-patient". "The intra-patient paradigm partitions the whole dataset into training and testing subsets based only on the beat label, and therefore an ECG recording may partly appear in both the data subsets." The research argues that "with this scheme, the classifiers usually produce over-optimistic results."
To provide a practical solution from this, reference 15 proposes "inter-patient paradigm where the training and testing subsets were constructed from different ECG recordings so that the inter-individual variation would be taken into account and the classifier would exhibit a better generalization ability." "In [15] the best-performing single-lead classifier obtained a maximum accuracy of 83.0%. Instead, the sensitivity and specificity results reveal that, while it achieved a high specificity (88.1%), the resulting sensitivities were very low." Reference 22 suggests instead "a hybrid paradigm called 'patient-specific'". In this proposed solution, "a global classifier was trained and then a local classifier was employed to tune the global classifier. The gross performance of the system presented in [22] was accuracy 97.4%, sensitivity 94.4%, and specificity 98.4%."

Paper 3 (Pyakillya *et al* 2017 *J. Phys.: Conf. Ser.* 913 012004)proposes using "deep learning (DL) architectures where first layers of convolutional neurons behave as feature extractors and in the end some fully-connected (FCN) layers are used for making final decision about ECG classes." In this proposed solution, "the deep learning architecture with 1D convolutional layers and FCN layers for ECG classification is presented and some classification results are showed."
In the research section of this paper, reference 2 shows that a "34-layer convolutional neural network and exceed the average cardiologist performance in both recall and precision [2]." In this research however, part of the crucial property to enable this to work was the "use of a dataset with more than 500 times the number of unique patients than other well-known corpora." In comparison, "other works present two-class ECG classification algorithms with simple neural network architecture (ischemic heart disease or normal sinus rhythm), which works well but in case of low amount of data and special problem statement [3–5]." However due to the lack of access to use that large of a dataset, a solution that is discussed to solve this issue is "to use of data augmentation technics, which are well fitted to time series classification problems [6–9]."

Proposed Approach

Our approach to meet our original objectives consists of three main parts. Firstly, we are going to build a binary classifier to distinguish between abnormal and normal ECG's. After we have built the model, we will train and test it to achieve a 'base accuracy'. We will then try and improve this figure by looking at how we can modify the dataset or the model. Once we achieve an accuracy that we are happy with, we will begin the next model.
Once our binary classifier works, we will go further and try and build a multi class classifier in the form of a 1 Dimensional Convolutional Neural Network. We hope to be able to classify between all of the classes in the dataset with a high accuracy. We will again get a starting accuracy with just the data. We will then try and improve it and track our changes with the accuracy metric.
The third thing we hope to achieve in this project is the use of Transfer Learning. We aim to implement this by building and training the model with the binary classifier. We will then use the model on the multi class problem.
In all instances, we will be tracking and recording our changes to see how our accuracy changes to hopefully build the most accurate system possible. At the end of the project, it will also allow us to compare all of our results and learn how to better approach a similar project in the future.

The Dataset

The dataset we are using is the 'ECG Heartbeat Categorization Dataset' from Kaggle. The dataset can be found at https://www.kaggle.com/shayanfazeli/heartbeat.
The dataset is a collection of csv files all of which contain the raw data from a series of ECG's. Each column is a voltage level which would produce the plot on a ECG graph. Each column could be considered as a time interval. The last column in each row is the class label.

Important Information about the dataset
Number of Samples: 109446
Number of Categories: 5
Sampling Frequency: 125Hz
Data Source: Physionet's MIT-BIH Arrhythmia Dataset
Classes: ['N': 0, 'S': 1, 'V': 2, 'F': 3, 'Q': 4]

Classes
N: Normal beat
S: Supraventricular premature beat
V: Premature ventricular contraction
F: Fusion of ventricular and normal beat
Q: Unclassifiable beat

The classes in the dataset do not have an equal amount of samples each. The number of samples in each class is shown below:
['N': 9589, 'S': 2779, 'V': 7236, 'F': 803, 'Q': 8033]

Training data is 80% of all samples. Test data is 20% of all samples
NB: For validation, some samples will be removed prior to the split.

Experiments

1. 1DCNN (Binary Classifier) with raw data.
2. 1DCNN (Binary Classifier) with noise reduction
3. 1DCNN (Binary Classifier) Interpolated
4. 1DCNN (Binary Classifier) Interpolated with noise reduction

5. 1DCNN (Multi Class Classifier) with raw data.
6. 1DCNN (Multi Class Classifier) with noise reduction
7. 1DCNN (Multi Class Classifier) Interpolated
8. 1DCNN (Multi Class Classifier) Interpolated with noise reduction

9. 1DCNN (Transfer Learning) with raw data.
10. 1DCNN (Transfer Learning) with noise reduction
11. 1DCNN (Transfer Learning) Interpolated
12. 1DCNN (Transfer Learning) Interpolated with noise reduction

Possible Challenges

Challenge	Solution
There are not equal amounts of each class, training may become biased / overfitted with classes where there are more samples	To make the system fair we have implemented up sampling and down sampling to balance the number of samples in each scenario. The list of samples is below: **Data** **Action** 1DCNN MC Training Up sample class 1-3 to 20k each 1DCNN MC Testing original: [8466, 265, 692, 67, 758] down sample: [2500, 265, 692, 67, 758] 1DCNN BC Training original [75609, 23589] up sample:[75609, 50000] 1DCNN BC Testing original [19026, 5774] down sample [5000, 5774] ***MC** – Multi Class Classifier, **BC** – Binary Classifier
Samples are different lengths. Important information (that could determine class) may be missed if all samples are truncated to shortest sample.	The interpolation of the data was done to make each sample of equal length and hopefully normalize the data for the model. This was done by stretching the dataset to the length of the longest sample and then using an interpolation function to fill in the data-points that were missing after the stretch. A polynomial interpolation was used as it best fitted the shape of the ECG.
Data is noisy. There are lots of insignificant changes in voltage (that are normal) which create noise. The system may detect these as significant.	To reduce noise we first split the ECG into sections. This will help to identify the unique features(P- wave, QRS etc). We take a rolling average across the section to smooth the data(eg. An increase of +0.1 followed by a decrease of -0.1 would result in an average change of 0 and the data would change such that the small increase and decrease were removed. So that we do not miss important features in the ECG we do not change any maximum or minimum points as this would change the overall ECG pattern(eg. A QRS complex that increases by +0.6 and then decreases -0.6 would be removed by the noise cancelling process).

The Models

Binary Classifier

Layer	Output Shape	Parameters
Reshape	(N, 187, 1)	0
Convolutional	(N, 185, 64)	256
Convolutional	(N, 183, 64)	12352
Dropout	(N, 183, 64)	0
Max Pooling	(N, 30, 64)	0
Flatten	(N, 1920)	0
Dense	(N, 100)	192100
Dense	(N, 1)	101

Multi Class Classifier

Layer	Output Shape	Parameters
SAME AS BINARY CLASSIFIER APART FROM LAST DENSE LAYER		
Dense	(N, 1)	505

Transfer Learning

Layer	Output Shape	Parameters
SAME AS BINARY UP UNTIL MAX POOLING LAYER		
Convolutional	(N, 181, 64)	12352
Convolutional	(N, 179, 64)	12352
Max Pooling	(N, 29, 64)	0
Flatten	(N, 1856)	0
Dense	(N, 100)	185700
Dense	(N, 5)	505

RESEARCH, PLANNING AND IMPLEMENTATION

COM3025 Deep Learning and Advanced AI – Detecting Abnormal Heartbeats

[redacted], Rui Santos and [redacted] Computer Science (BSc) Year 3 - 2020

Binary Classifier

The binary classifier performed well with our highest accuracy being 97%. However, this was with the raw data which is unexpected. We believe this may be to slight overfitting as there were still a significantly larger amount of samples in the normal class (even with up sampling). If this is taken into account we believe our measures did improve the system as we had a slight increase in accuracy with our measures in place.

The results are shown below.

Classifier		Accuracy	Val
Binary Classifier	**Raw**	97	96.9
	NR	96.5	95.5
	IN	96.6	93.1
	NR+IN	96.6	95.5

NR – Noise Reduction, IN – Interpolated

Multi Class Classifier

There was also a very pleasing result with our CNN model to detect multiple classes. Our highest accuracy was 93 with the interpolated data. We did see a slight increase from the raw data in both accuracy and validation accuracy. However we believe the model may have been slightly overfitted again. This is again due to the large number of samples in one class and the high amount of up sampling required in the training data.

The results are shown below

Classifier		Accuracy	Val
Multi-Class Classifier	**Raw**	92.6	91.7
	NR	92.6	90.8
	IN	93	91.9
	NR+IN	92.3	92.5

NR – Noise Reduction, IN – Interpolated

Transfer Learning

The transfer learning model produce some impressive results given the small training set available. On top of a high performance for all input types, Noise reduction seemed to be the best modification to the training data. However the differences were very small between the two top performing models. Some overfitting seemed to have occurred in IN as validation accuracy reduced while training accuracy kept increasing.

The results are shown below

Classifier		Accuracy	Val
Transfer Learning	**Raw**	96.1	96.4
	NR	96.7	96.4
	IN	96.5	94.7
	NR+IN	96.4	96.6

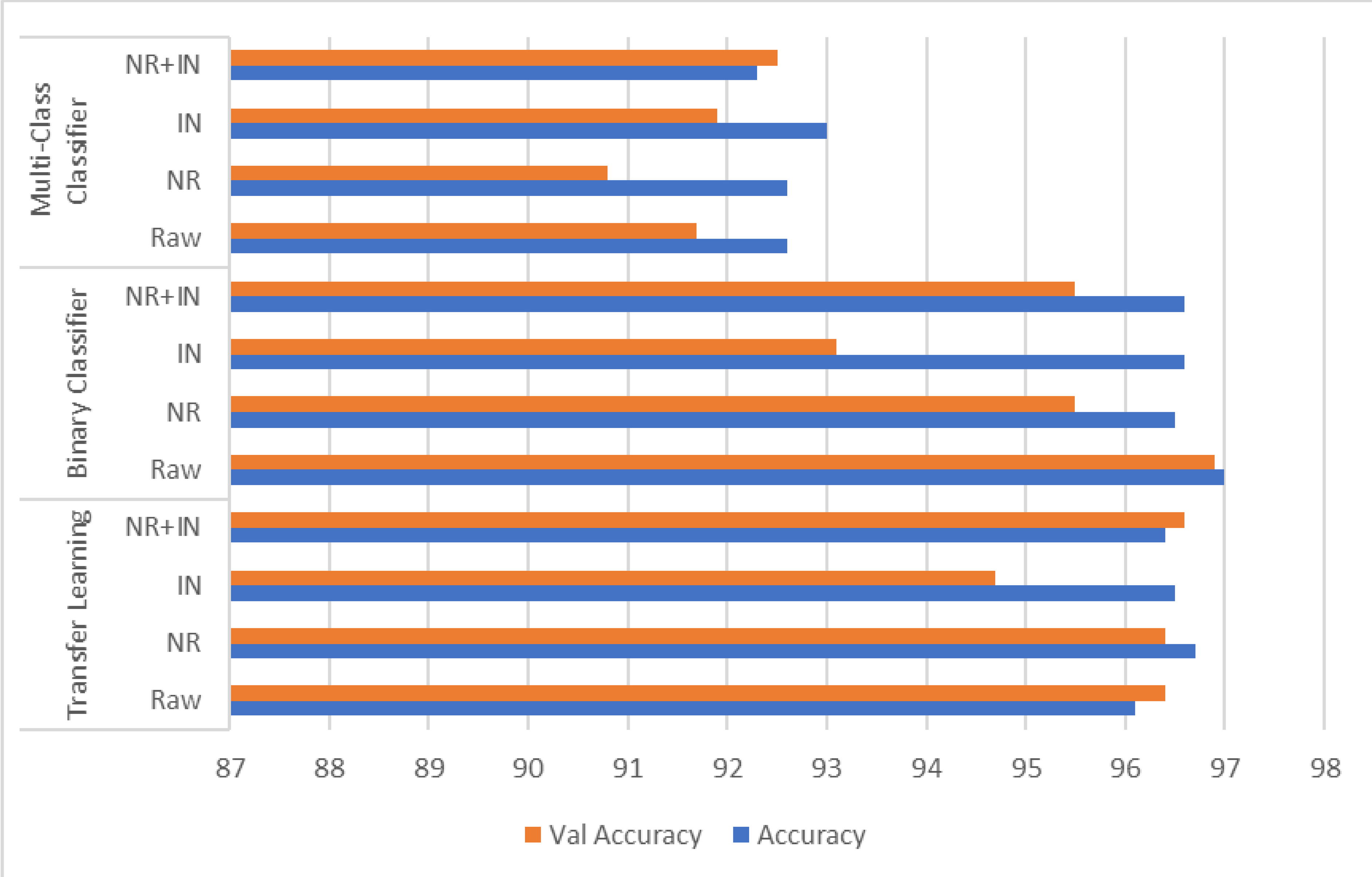
NR – Noise Reduction, IN – Interpolated

Results

The Binary Classifier(BC) performed very well, and this can be attributed to solving an easier classification problem and have more unique samples to train on. The attempts to normalize the data did not provide any benefit to validation accuracy over training on the raw samples.

The Multi-class Classifier(MC) struggled to reach accuracies above 95% which lead to a disappointing result. This could be because of the large discrepencies in the spread of samples of each class. In this case normalisation of the data did provide an advantage to validation accuracy. However it is unclear why this effect was not seen in the binary classifier.

The Transfer Learning Classifier(TLC) produced the most promising results and showed an effectiveness of transfer learning for training on small datasets. The transfer learning model training on the same dataset as the multi-class classifier and managed to outperform it very well. This model was deeper which may have contributed to a better understanding on the input data by the model however utilizing the weights created in the binary classifier prevented some overfitting issues the MC experienced. Combining the two models allowed the multi-class problem to be solved with the accuracy of a binary class problem which is an impressive result.



Evaluation of Objectives

- **Build Binary Classifier to determine whether an ECG is normal or not.**

Overall the outcome for this task was a success. We were able to detect normal and abnormal ECG's with a high accuracy of 97%. With over 125,000 samples tested, in a real world scenario this could save a lot of time and significantly benefit the industry. However, in practice the accuracy would probably still need to improve. In future, we would try and make the dataset even fairer to eradicate any overfitting.

- **Build a Multi-Class Classifier to determine which class the ECG falls into.**

Again, we are happy with the outcome of this task as even with the more complicated job of distinguishing between 5 classes our model still performed well, with a top accuracy of 93%. There is a larger area for improvement here and one aspect to possibly improve in future, would be to implement a "unrecognised" class so that if there is a doubt in which class it should be, it could be reviewed.

- **Build upon the models by implementing measures to improve the accuracy. This includes (but not limited to):**

- Experimenting with the data to make it better to train with.

Although our methods did not greatly improve the results of our system, we still believe they made a positive impact on the system. The skills learnt from this task will still be useful in the future.

- Using more advanced techniques such as Transfer Learning.

The outcome of this experiment was very successful and allowed for a small data set with uneven sample numbers to solve the problem with a very high accuracy. Considering the model was training with one class covering 90% of unique samples, the transfer learning approach allow us to create an efficient model despite a less than optimal training set.

Further Research

There are a few areas that could be pursued in future with the aim to improve the project.

1. Implement a "unrecognised" class so that data samples that cannot be categorised are sent for review. This would increase the accuracy of the system as the model would not simply "guess" the right class.
2. Experiment with more data manipulation to make the original dataset fairer. The biggest problem we had was the heavily biased nature of the original dataset. Spending more time making this fair would be a good area for improvement.
3. Experiment and compare with different model architectures. We only implemented CNNs in our project. A good area of research would be to find out if other architectures have a better result and find out whether they are useful for this purpose.

Concluding Thoughts

Overall the result of this experiment was thought to be a success. The main difficulty was managing to utilize an ill fitted dataset, and this was overcome quite well. It would be interesting to gather more datasets to train MC model and compare its result with the TLC model. With the current dataset , the best model was the TLC model and if we would suggest that this kind of model be used to diagnose ECGs. Time is a factor that was not assessed in this research project however, it does have implications on the success of a real-world application. For this reason, comparison between prediction speeds would be a reasonable enquiry for the future. The result of this experiment was a success and we would suggest the TLC technique for training an ECG Classification model, however there are some real-world factors that would need to be address before this model could be useful in real-time ECG classification.

RESULTS AND EVALUATION

References

- Ganesan, Prasanth & Sterling, Mark & Ladavich, Steven & Ghoraani, Behnaz. (2016). Computer-Aided Clinical Decision Support Systems for Atrial Fibrillation. 10.5772/65620.
- Electrical System of the Heart. (n.d.). Retrieved from https://myhealth.alberta.ca/Health/pages/conditions.aspx?hwid=te7147abc
- Electrical conduction system of the heart. (2020, March 6). Retrieved from https://en.wikipedia.org/wiki/Electrical_conduction_system_of_the_heart
- Rawat, W. and Wang, Z., 2017. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, 29(9), pp.2352-2449.
- Sciencedirect.com. 2020. *Binary Classification - An Overview | Sciencedirect Topics*. [online] Available at: <https://www.sciencedirect.com/topics/computer-science/binary-classification> [Accessed 26 May 2020].
- Scikit-learn.org. 2020. *1.12. Multiclass And Multilabel Algorithms — Scikit-Learn 0.23.1 Documentation*. [online] Available at: <https://scikit-learn.org/stable/modules/multiclass.html> [Accessed 26 May 2020].
- Brownlee, J., 2020. *A Gentle Introduction To Transfer Learning For Deep Learning*. [online] Machine Learning Mastery. Available at: <https://machinelearningmastery.com/transfer-learning-for-deep-learning/> [Accessed 26 May 2020].
- Sherin M. Mathews, Chandra Kambhamettu, Kenneth E. Barner,A novel application of deep learning for single-lead ECG classification,Computers in Biology and Medicine,Volume 99,2018,https://doi.org/10.1016/j.compbiomed.2018.05.013.
- G. Sannino, G. De Pietro,A deep learning approach for ECG-based heartbeat classification for arrhythmia detection,Future Generation Computer Systems,Volume 86,2018,https://doi.org/10.1016/j.future.2018.03.057.
- B Pyakillya, N Kazachenko, & N Mikhailovsky (2017). Deep Learning for ECG Classification*Journal of Physics: Conference Series, 913*, 012004.