Cardiac rhythm interpretation

Mari Amdalsrød Hognestad[1], Sander A. Søndeland[1]

[1] Faculty of Science and Technology, University of Stavanger, Stavanger, Norway

Abstract

Sometimes a person may have a sudden cardiac arrest without any medical personnel present. If this were to happen, it would have been useful to have an automated external defibrillator nearby to tell you when to shock the person.

In this study, a dataset containing ECG recordings from Out-of-hospital-cardiac-arrest will be trained with a convolutional neural network. The dataset is classified into five different heart rhythms: asystole, pulseless electrical activity, pulse-generating rhythm, ventricular fibrillation and ventricular tachycardia. The aim is to create a model that provides the best accuracy. Several trials were therefore carried out, with each trial receiving a new adjustment of parameters. In the end, the best achievable result was 85% accuracy.

1. Introduction

During cardiopulmonary resuscitation, the patient is treated with heart compressions, ventilations and electroshock. Normally, a cardiac arrest will be treated in the hospital, but in some cases it happens before you are able to get there. This scenario is called Out-of-hospital-cardiac-arrest, also known as OHCA [1]. Sudden and unexpected cardiac arrest that occurs outside the hospital affects around 3,000 people in Norway each year [3].

OHCA can happen anytime, anywhere, and with anyone without warning. When a person collapses with sudden cardiac arrest, every minute counts. In some cases where OHCA occurs, healthcare personnel will not be present, but an automated external defibrillator (AED) will be. An AED is used when a person goes into sudden cardiac arrest. It uses advanced technology to automatically evaluate and deliver a shock that will restart the heart. The AED is programmed to recognize and deliver shocks to two types of abnormal heart rhythms. These two types are called ventricular fibrillation (VF) and ventricular techycardia (VT). Due to this programming, the AED will not deliver a shock to someone with a normal heart rhythm. The shock from an AED is used to correct abnormal electrical disturbance caused by irregular heart rhythm which will in turn establish normal heart function[11]. Not only will an AED be easily accessible to individuals, it also imporove the chance of survival up to 74% percent [10].

In this study, a model for categorizing heart rhythms will be implemented. Various experiments will be conducted to find the most suitable model that can be used during OHCA. The report consists of the method used to find the model followed by experiments. Finally, the result is presented together with a discussion.

2. Dataset

In order to build a highly accurate model, a lot of training data is required. The more data, the better the algorithm until overfitting.

The data is taken from a study on quality in cardiopulmonary resuscitation with real-time automated feedback. The dataset used in this project is a subset of the data collected from the study. ECG signals, time signal events, accelerometer signals and trans-thoracic impedance are included in the recordings. The sampling frequency is 500 Hz with 16-bit resolution and the data is annotated by experts. In this case, the data is downsampled to 250 Hz and into 4 second segments from 100 different patients. In total there are 2833 segments containing 423 AS, 912 PEA, 689 PR, 643 VF and 166 VT [7]. An explanation of the various classes can be found in subchapter 3.

3. Classification of heart rhythm

The heart rhythms are classified into two different main classes, shockable or non-shockable rhythm. The shockable rhythm is again divided into two classes called coarse ventricular fibrillation (VF) and rapid ventricular tachycardia (VT). If the amplitude of a coarse ventricular fibrillation signal is greater than 200 μV, the rhythm is shockable. Rapid ventricular tachycardia (VT) is shockable at a rate greater than 150 bpm.

The non-shockable rhythms contain three classes, asystole (AS), pulseless electrical activity (PEA) and pulse generatic rhythm (PR). When there is no heartbeat or electrical activity, i.e. the rhythm is flat, it is classified as asystole (AS). When there is a flatline rhythm, it means that the heart and defibrillation are not working. Less than 2% of people with Asystole survives.

Pulseless electrical activity(PEA) is a cardiac rhythm that is life-threatening and unshockable. Even though there is presence of coordinated cardiac electrical activity there is no perceptible pulse that is shockable. Pulse generatic rhythm (PR) are rhythms that is not associated with cardiac arrest and therefore can not be shockable rhythms. Normal sinus rhythm (NSR) and supraventricular tachycardias are some examples in this class. [2]

4. Method

The data set that was available had to be fitted to be processed by the neural network. A data frame was established which contained the patient's ID, ECG signal, shockable label and rhythm categories respectively. The shockable label divides the rhythm categories into what is shockable (VF and VT) or non-shockable (AS, PEA and PR) [1]. Furthermore, the number of signals in each group was shortened to equal length to avoid one class being overfitted.

Finally, the data frame was divided into training, test and validation sets using stratified 10-fold cross-validator.The purpose of a cross-validator is to ensure that recordings from a person's heartbeat would only occur in one data set. Stratified cross validation will ensure that the five different classes (AS, PEA, PR, VT and VF) are represented in each dataset. This type of train-test-split will lead to good accuracy [1].

4.1. Convolutional neural network

In this study, a convolutional neural network (CNN) is the deep learning model used to classify the category of a heart rhythm. Although CNNs are considered the most favorable for image classification, a CNN will also be able to provide good accuracy for one-dimensional signals.

The layers in the deep learning model are considered the model's architecture. There are different types of layers that can be used in the model where all the layers have their own meaning based on their functions [13]. A one-dimensional CNN will have input and output data that are two-dimensional, while the kernel will slide along one dimension. A kernel is a vector of weights that is convolved with a vector of input data. This linear operation will result in the input vector being longer than the output. To produce equal size vectors, zero padding is added to the beginning and end of the input vector. The padding that is added has a zero value so it will therefore have no effect when multiplied by the kernel. Because of this, paddig will improve performance by keeping the information at the borders [4].

A problem with the output feature maps is that they are sensitive to the location of the the features in the input. One way to adress this sensitivity is to downsample the feature map. The resulting downsampeled feature maps is more robust to changes in the position. For this reason, a pooling layer is requried when using padding [5].

The pooling layer reduces the length of the signal by summing the presence of features in patches. There are two main methods of pooling layers - average and max pooling. The average pooling layer, as the name suggests, takes the average value for each patch on the feature map. This causes it to smoothly extract features and will therefore not extrat the salient features. Max pooling, on the other hand, calculates the maximum value for each patch, which in turn causes it to extract the most prominent features of the data. It must also be taken into account that downsampling will limit the identification of features, which in turn results in a loss of detail. Too many pooling layers will reduce the model size to such a large extent that it leads to loss of information [8].

It is also possible to replace the pooling layer by increasing the strides on the CNN layer, but it is not seen as an equally robust approach.

After the model has learned the features, it is time to fit the last feature map to a neural network for classification purpose. However, since the fully connected layer only takes in a one dimentional array, the model needs to be flattened first. Therefore a flattened layer will be needed to convert the data into a one dimentional array. After flattening a dense layer is added to perform classification.

5. Experiments

In order to produce the best model for the classification of heart rhythms, several experiments were carried out. The aim of the experiments was to compare the performance of different structures in a CNN model. Yoshua Bengio proposed a simple rule on how to achieve the best possible model. *"Very simple. Just keep adding layers until the test error does not improve anymore"* [12]. The experiments conducted are based on this simple remark.

The first experiment was to test the effect of zero padding. A simple model with four CNN layers was therefore trained with and without padding. Although padding will ensure that the information on the borders is also included, there was no difference in the result.

Both models gave poor accuracy and had a long processing time.

The reason for the poor accuracy comes from the fact that the output feature map is sensitive to the location of the features in the input. Therefore, it is necessary to reduce the spatial dimension in order to improve the prediction. One way to achieve this is by increasing the stride size, i.e. moving the kernel by several steps. Default stride size is one, and by setting this to two, the length of the vector will be reduced by half. By increasing the stride size, the processing time was much shorter but the accuracy improved marginally.

A more common and robust way to downsample the length of the vector is by using pooling layers. In this study, the aim is to distinguish different signals from each other. Therefore, maxpooling was used as the method focuses on the important features of the feature map.

To show that reducing the model size gave better accuracy, a test was carried out where a pooling layer was added after each CNN layer and where it was added after all the CNN layers. Not surprisingly, the model with the most pooling layers came out best. However, downsampling only gives better accuracy up to a certain point as information is also lost.

Until now, the experiments have been carried out with a four-layer CNN. Just by adding one extra CNN layer, the result, in accordance with Yoshua Bengio's simple rule, will improve. The five-layer CNN gives good results, but this too can be improved. For one dimentional CNN models, the choice of filter size is essential in finding the most important parts of the signal. The process of choosing the size of filter can be a time-consuming process. By only defining one filter, the neural network will only learn one simple feature in the first layer. Therefore, each layer is defined with several filters. After pooling layers, the CNN model will be able to learn a higher level of features, so therefore the filter size is increased for each layer. When it comes to kernel size, these are associated with the weights that are convolved with the input data. The most common kernel sizes are in odd numbers.

6. Results

The results from the experiments from subchapter 5 are shown in table 1 and 2.

The tables show the sensitivity and the positive predicted value (PPV) of each heart rhythm in the different experiments. In addition, the accuracy of each trial is also presented.

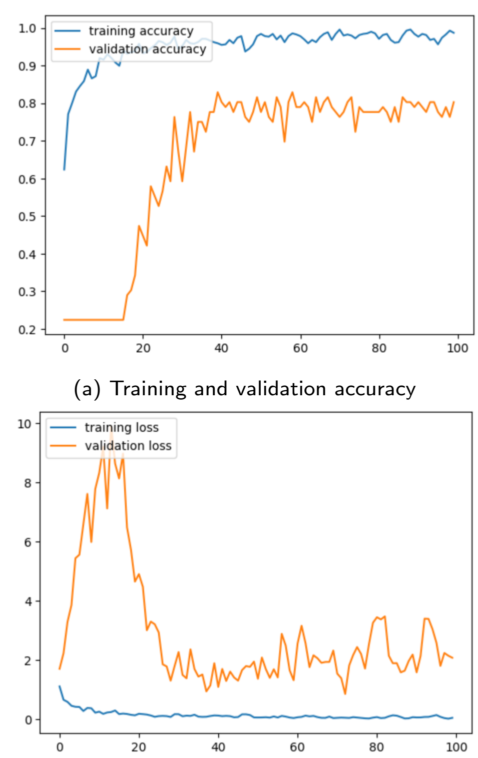
The sensitivity and PPV show the characteristic of the tests and can be found by comparing true positive (TP), false positive (FP) and false negative (FN) results [9].

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$PPV = \frac{TP}{TP + FP} \quad (2)$$

The sensitivity measures how suitable the model is to classify each heart rhythm and a high PPV value shows how accurate the prediction of the class is.

The model that gave the best results was a five-layer CNN with increasing numbers of filterst. Figure 1 shows the validation accuracy and loss for this model.



(a) Training and validation accuracy

(b) Training and validation loss

Figure 1: Training and validation data for the five-layer CNN

Ideally, the validation accuracy should converge towards one, while the validation loss should be as low as possible, close to zero.

		AY		PEA		PR		VF		VT	
Experiment	ACC	Sens	PPV	Sens	PPV	Sens	PPV	Sens	PPV	Sens	PPV
Without padding	0.20	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
With padding	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.20
Padding and stride	0.20	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Padding and mult.pooling	0.64	0.80	1.00	0.43	0.68	0.77	0.57	0.80	0.58	0.43	0.52
Padding and one pooling	0.23	0.00	0.00	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00

Table 1: Result with a four-layer CNN

		AY		PEA		PR		VF		VT	
Experiment	ACC	Sens	PPV	Sens	PPV	Sens	PPV	Sens	PPV	Sens	PPV
5 layer	0.73	0.87	1.00	0.56	0.50	1.00	0.66	0.87	0.86	0.36	0.73
Increasing filters	0.85	1.00	0.87	0.86	0.68	0.6	0.73	1.00	1.00	0.78	1.00

Table 2: Result with a five-layer CNN

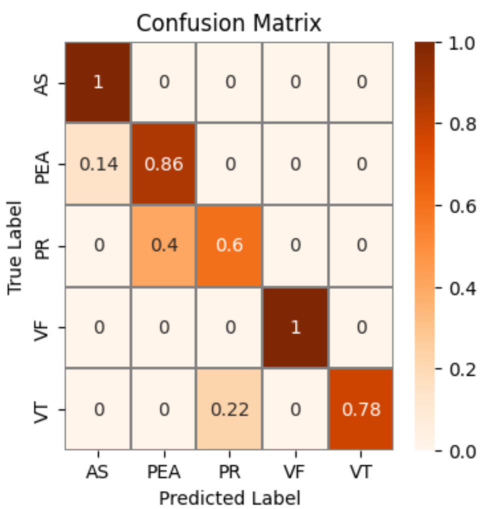


Figure 2: Confusion matrix for the five-layer CNN

The confusion matrix in figure 2 is also a well known way of presenting the result. A good matrix will have higher values along the diagonal and lower values, preferably zero, otherwise.

7. Discussion

We have now gone through various approaches to find the best model for classifying heart rhythms and as a result we obtained 85% accuracy.

The experiments from table 1 clearly shows that if you want to improve the model, you must add zero padding to the input vector followed by multiple pooling layers. Without the pooling layers, the feature maps will be too sensitive to the location of the features in the input and the result is that the model can only classify one type of heart rhythm. We can also observe that only one pooling layer has no effect and the result is the same as with no pooling layer.

Furthermore, one can observe that by only adding an extra layer, the accuracy increases considerably from 64% to 73%, see table 2. As the pooling layer sums together features in patches, this means that the CNN can learn a higher level of features. For this reason, we increased the number of filters for each layer.

This led to an accuracy of 85% which is a good prediction. From table 2 it can be seen that the sensitivity and PPV are relatively high compared to table 1. The model with an increasing number of filters manages to classify each of the heart rhythms AS, VF and VT, in addition, the sensitivity on PEA and PR is also high. If you look at the PPV values, these show high accuracy in predicting each class. Again, it is AS, VF and VT that gives best result, while PEA and PR give slightly worse.

One thing that was discovered a little later in the experiments was that training with the same model and parameters could produce different results. That's because the weights are randomly generated, which means that each training will be different. This can be fixed by instead of having random weights you can use a fixed seed that will not change. The seed will make sure to deliver predictable and repeatable results every time. Unfortunately, this was something we did not get to test in this project so the results for future tests may differ from the results in table 1 and 2.

As a conclusion we can say that that adding more layers improved the accuracy of prediction. With this being said, adding more layers will help you extract more features but we can only do that up to a certain extent. There is a limit, and if this limit is crossed, the data tend to be overfitted which can lead to errors like false positives. How many layers that is needed depends on how large and complex the data set is. I.e. Yoshua Bengio's simple rule is a good place to start and it got us to a 85% accuracy.

References

[1] Christopher Andreassen, Jon Tveit. Cardiac rhythm interpretation. https://nextcloud.ux.uis.no/s/ECfWQJWskTGjt4j. (accessed: 21.10.2022)

[2] Dahal, K., Ali, M.H., (2022). Overview of Machine Learning and Deep Learning Approaches for Detecting Shockable Rhythms in AED in the Absence or Presence of CPR. https://nextcloud.ux.uis.no/s/g3iFYCE2ekDK5C9. (accessed: 24.11.2022)

[3] HelseNorge. Hjertestans. https://www.helsenorge.no/sykdom/hjerte-og-kar/hjertestans/. (accessed: 15.11.2022)

[4] How Convolutional Layers Work in Deep Learning Neural Networks?. https://jinglescode.github.io/2020/11/01/how-convolutional-layers-work-deep-learning-neural-networks/. (accessed: 14.11.2022)

[5] Jason Brownlee. A Gentle Introduction to Pooling Layers for Convolutional Neural Networks. https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/. (accessed: 14.11.2022)

[6] Krasteva, V., Ménétré, S., Didon, J.P., Jekova, I. (2020). Fully Convoltuional Deep Neural Networks with Optimized Hyperparameters for Detection of Shockable and Non-shockable Rhythms. Sensors.

[7] Lasa, H., Eftestol, T., Aramendi, E. Rad, A.B., Kramer-Johansen, J., Wik, L. Multimodal Biosignal Analysis Algorithm for the Classification of Cardiac Rhythms During Resuscitation. https://www.cinc.org/archives/2020/pdf/CinC2020-347.pdf. (accessed: 12.11.2022)

[8] Priyanshi Sharma. MaxPool vs AvgPool. https://iq.opengenus.org/maxpool-vs-avg/. (accessed: 14.11.2022)

[9] Robert Trevethan. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. https://www.frontiersin.org/articles/10.3389/fpubh.2017.00307/full. (accessed: 27.11.2022)

[10] Røde Kors. Hva er fordelene med en helautomatisk hjertestarter?. https://www.rodekorsforstehjelp.no/rad-og-informasjon/produktinformasjon/hjertestarter/. (accessed: 15.11.2022)

[11] XABIER JAUREGUIBEITIA. Shock Decision Algorithms for Automated External Defibrillators Based on Convolutional Networks. https://nextcloud.ux.uis.no/s/Hwbe9dX596dKzT5. (accessed: 21.10.2022)

[12] Yoshua Bengio. Artificial Neural Networks: How can I estimate the number of neurons and layers?. https://www.quora.com/Artificial-Neural-Networks/Artificial-Neural-Networks-How-can-I-estimate-the-number-of-neurons-and-layers/answer/Yoshua-Bengio?share=7b58dc3b&srid=hqJd. (accessed: 17.11.2022)

[13] Yugesh Verma. A Complete Understanding of Dense Layers in Neural Networks. https://analyticsindiamag.com/a-complete-understanding-of-dense-layers-in-neural- (accessed: 11.11.2022)