Multi-channel ECG Classification with Deep Convolutional Neural Networks

Ali Gholami, Hamid Soltani*, Emad Soudani*
Dept. of Computer Engineering, Dept. of Bio-medical Engineering
Amirkabir University of Technology
{aligholamee,hamidsoltani,emad}@aut.ac.ir

Abstract

ECG or Electrocardiogram is an imperative diagnosis tool for measuring the electrical activity of the heartbeat. ECG Classification is important because of the sensitive nature of heart and pulses it propagates through body. Moreover, chronic heart diseases are reflected through these pulses and identifying such abnormalities is known to be a difficult problem even for human experts. Recently, the use of deep neural networks has become popular in the signal processing community. In this project, we explore the use of convolutional neural networks in detecting such abnormalities. We also compare the results with multi-layer perceptrons in the same task.

1 Introduction

An important diagnostic tool in identifying chronic and acute heart rhythm irregularities (cardiac arrhythmias) is the electrocardiogram (ECG), which measures electrical heart activity via electrodes placed on a patients skin. ECGs are ubiquitous in intensive care units (ICUs), where clinicians must be able to make critical care decisions quickly and accurately. The ability to correctly distinguish various arrhythmias from each other is crucial for patient well-being; in many cases, the wave morphologies of benign and lethal arrhythmias can be difficult to distinguish. Existing monitoring systems for ECGs record a myriad of vital signs and also utilize algorithms to determine changes in cardiac rhythm. However, accurate identification of arrhythmias is known to be challenging even for medical professionals, and requires considerable medical expertise. A study investigating diagnostic accuracy for licensed general practitioners showed a specificity of 92% and sensitivity of only 80% in distinguishing atrial fibrillation from healthy sinus rhythms [1]. Waveforms often show variation given an individuals unique biological characteristics, even for arrhythmias whose identifying patterns are known and well-documented. We incorporate the multi-channel data of Electrocardiogram signal, to increase the classification accuracy. Individual channels record cardiac electrical activity from multiple various spatial angles. Using a multi-channel signal, we can identify deeper and more complex underlying patterns that separate different Electrocardiogram signals.

2 Background

2.1 Cardiac Arrhythmias

The rhythm of a human heart is regulated by electrical signals produced by two nodes within the heart and conducted through a series of specialized cardiac cells. During healthy, normal operation, this occurs at regular intervals and the electrical signal, which causes the heart muscles to contract, propagates via the cardiac electrical conduction system along the correct path through the atria and ventricles. Cardiac arrhythmia occurs when the heartbeat is too fast (tachycardia), too slow (bradycardia), or altogether abnormal. Both atrial and ventricular arrhythmias can have any number of causes, including scar tissue from previous trauma (such as myocardial infarction) and coronary disease, and can even occur in healthy hearts. Though many arrhythmias are asymptomatic, those that are not can cause symptoms as mild as occasional palpitations or as severe as stroke and sudden cardiac death. As arrhythmias are caused by disorders of the electrical conduction system, they are reflected in ECG readings as abnormal waveforms. The complexity of this system often necessitates that clinicians use anywhere from six to twelve ECG leads to capture electrical activity across multiple spatial planes. A single lead only provides a projection of this activity across one specific plane and may not provide enough information to make accurate diagnoses of underlying pathologies. The rhythms we aim to classify are normal (**N**) rhythm, Supraventricular Ectopic Beat (**S**), Fusion Beat (**F**) and Unknown Beat (**Q**).

Category	**Annotation**
N	Normal, Atrial Escape, Nodal Escape
S	Aberrant Atrial Pre-mature, Nodal Pre-mature
V	Pre-mature Ventricular Contraction, Ventricular Escape
F	Fusion of Ventricular and Normal
Q	Paced, Fusion of Paced and Normal, Unclassifiable

Figure 2.1: Summary of mappings between beat annotations and PhysioNet MIT-BIH categories.

2.2 Multi-layer Perceptron

A multilayer perceptron (MLP) is a class of feedforward artificial neural network. An MLP consists of, at least, three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable [2]. Here is a formalization of what a multilayer perceptron with one hidden layer represents:

$$y = \sigma(Wx + b) \tag{2.1}$$

Multilayer perceptrons then leverage gradient descent and back-propagation to optimize their estimation with respect to a loss function [3].

2.3 Convolutinal Neural Networks

CNNs are a derivative of standard Multilayer Perceptron (MLP) neural networks optimized for 1-dimensional, 2-dimensional and 3-dimensional pattern recognition problems such as EEG or ECG Signal Classification. Instead of using fully connected hidden layers as described in the preceding section, the CNN introduces a special network structure, which consists of alternating named convolution and subsamlping layers. Feature maps generated by convolution layers, contain neurons that take their synaptic inputs from a local receptive field. The weights of neurons within the same feature map are shared. This represents ones of the characteristics of convolutional neural networks. It allows to have replicated units sharing the same configuration, thereby features can be detected regardless of their position in the visual field. Moreover, the fact that weights are shared increases learning efficiency by reducing the number of parameters being learnt. In order to have a data reduction, a sub-sampling operation called pooling is performed. This data reduction operation is applied to the predecessor convolution result by a local averaging over a predefined window. It partitions the input image into a set of non-overlapping windows and then for each sub-region outputs the maximum value. This step is important because it helps to eliminate non-maximal values and to provide a form an invariant translation. The output layers ensures the classification of the input character. In these layers all neurons are fully connected and have a unique set of weights so they can detect complex features and perform classification. Figure 2.3 represents one of the first convolutional neural network architectures used for digit recognition [4].

3 Implementation

In this section, we'll describe the pipeline used for the implementation of this system. We use Keras, a deep learning framework for fast implementation of various models of neural networks. We trained the models described below on an Nvidia 850m GTX with 640 CUDA Cores.

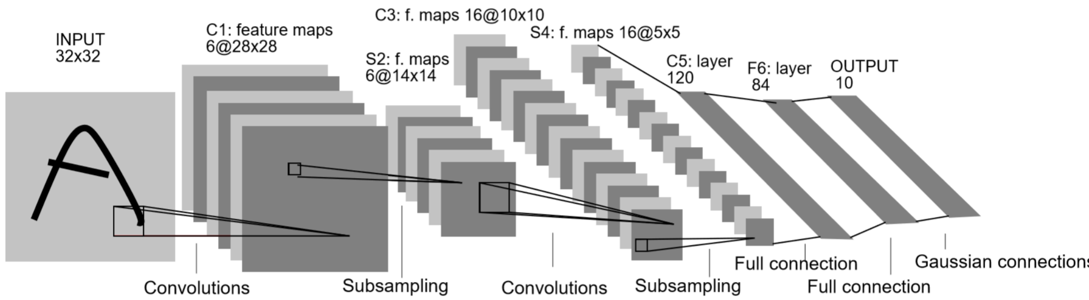

Figure 2.2: Architecture of LeNet-5, a Convolutional Neural Network for digits recognition. Each plane is a feature map, i.e., a set of units whose weights are constrained to be identical.

3.1 Method

3.1.1 Three Layer MLP

Our implementations consist of two major architectures. One is a multi-layer perceptron and the other one is a convolutional neural network. We analyzed various hyper-parameter settings to improve the accuracy of predictions for all classes. Figure 3.1.2 illustrates this model.

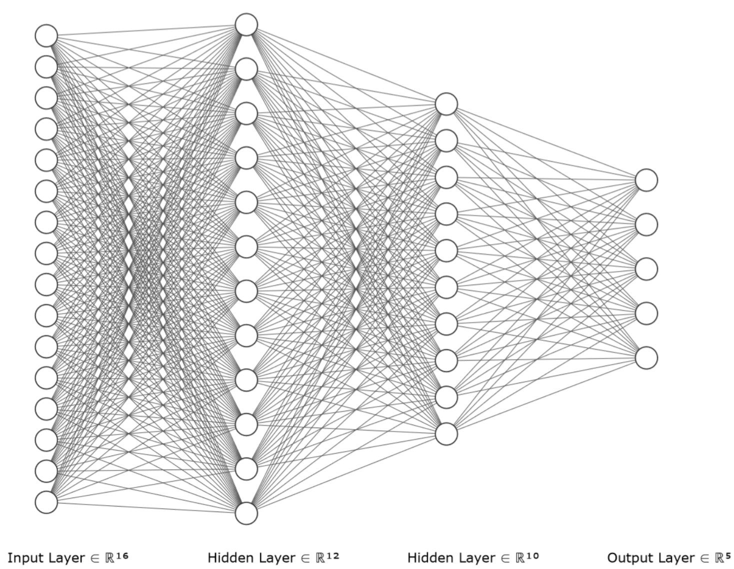

Figure 3.1: Architecture of our three layer perceptron.

This model leverages the complexity that hidden layers of a neural network provide. In the initial setup, we specify 128 hidden units for the first hidden layer and 64 hidden units for the second hidden layer. Outputs are passed through a Softmax to achieve probabilities for each class of prediction.

3.1.2 Five Layer MLP

This is our second implementation of multi-layer perceptron. This model employs five hidden layers with 256, 128, 64, 32 and five units in each hidden layer respectively.

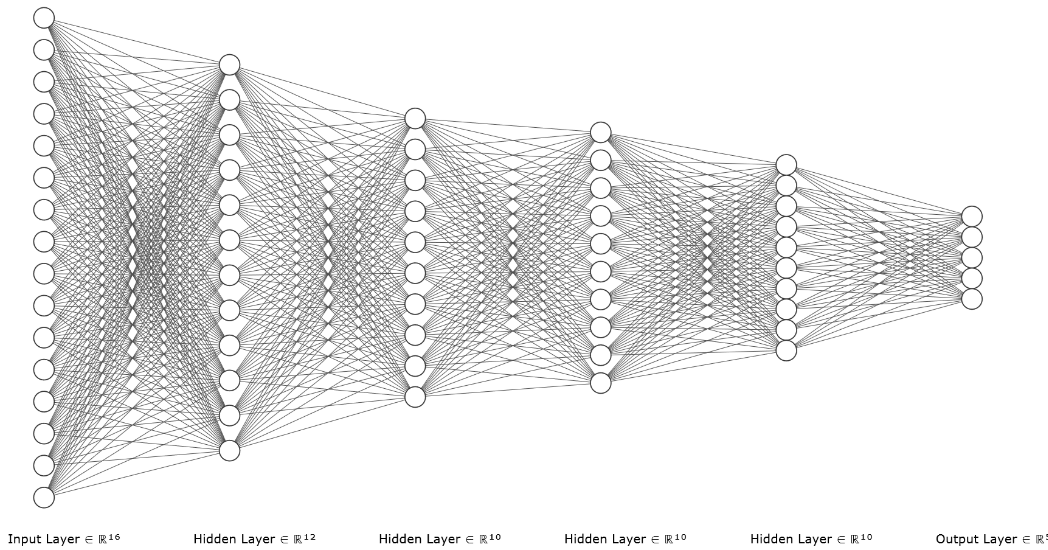

Figure 3.2: Architecture of our five layer perceptron.

3.1.3 Convolutional Neural Network Model

In this algorithm, we leveraged the power of convolutional neural networks for dimensionality reduction on signals. CNNs are also very good at obtaining the dependencies between various timesteps of the structured input data. Pooling operations are employed in the network for translation invariance and robustness of the model. Extracted feature maps are the passed through a classifier, a simple multi-layer neural network in this case and applied a Softmax for the final predictions.

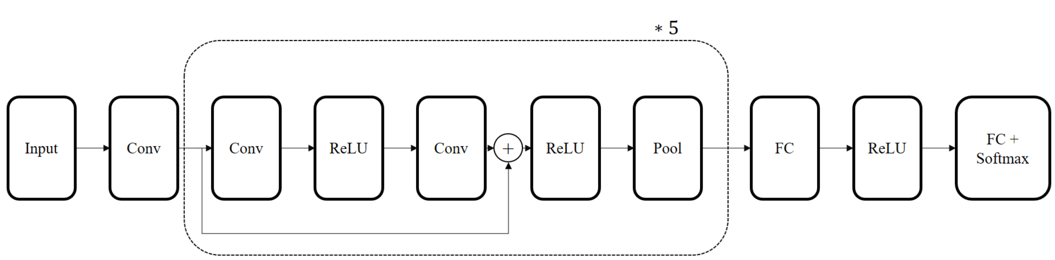

Figure 3.3: Architecture of the CNN we used for classification [5].

3.2 Dataset

The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were chosen at random from a set of 4000 24-hour ambulatory ECG recordings collected from a mixed population of inpatients (about 60%) and outpatients (about 40%) at Boston's Beth Hospital. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a ten mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database. To feed our neural networks, these signals are preprocessed and segmented, with each segment corresponding to a heartbeat. In the following section, we will describe the steps in which signals are preprocessed.

3.3 Preprocessing

As ECG beats are inputs of our model, we use an effective method presented in [5] for preprocessing ECG signals to extract heartbeats. The steps for achieving an extracted beat is as follows:

1. Splitting the continuous ECG signal to 10s windows and select a 10s window from an ECG signal.

2. Normalizing the amplitude values to the range of between zero and one.

3. Finding the set of all local maximums based on zero crossings of the first derivative.

4. Finding the set of ECG R-peak candidates by applying a threshold of 0.9 on the normalized value of the local maximums.

5. Finding the median of R-R time intervals as the nominal heartbeat period of that window (T).

6. For each R-peak, selecting a signal part with the length equal to 1.2T.

7. Padding each selected part with zeros to make its length equal to a predefined fixed length.

4 Experiments and Results

In this section, we will review a number of experiments we have done for classification of ECG signals. It is important to mention the class imbalance inside the dataset which reduced the accuracy of the model on the initial setup. We generated multiple samples extending our dataset for a balanced classification. This method is implemented in Keras as well and is present with the files attached with this paper. Moreover, we applied Scikit-learn confusion matrix to obtain a clear sense of how our model is performing on each of the classes. For example, the model risk of predicting normal wave compared to predicting an abnormality in the ECG signal is not the same. In this case, analyzing the confusion matrices of the model performance can be helpful. For instance, figure 4 represents the classification results using our convolutional neural network.

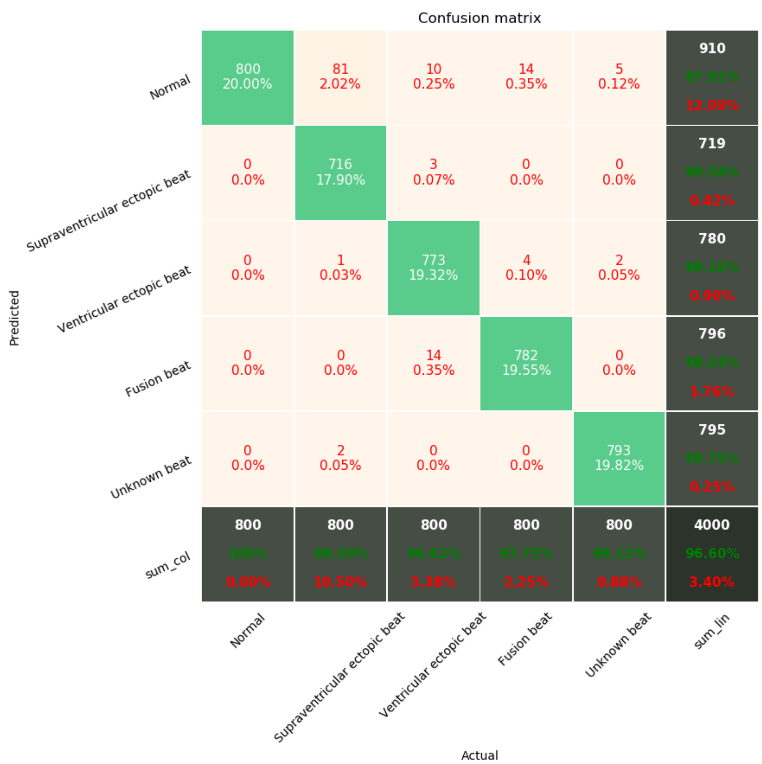

Figure 4.1: Confusion matrix for the classification results using the convolutional neural network.

Model	Precision	Recall	F1-Score
Three layer MLP	0.92	0.89	0.86
Five layer MLP	0.93	0.90	0.90
Convolutional Neural Network	0.97	0.96	0.94

Figure 4.2: Summary of performances for each of the models presented in section 3.1.

5 Conclusion

We showed that using known deep network algorithms for classifying time-series data allows for accurate classification of normal, benign, and critical arrhythmias as well as distinguishing artifacts and noise from multi-channel ECG recordings. A convolutional neural network can achieve relatively high accuracy and precision without the use of feature engineering or extraction of previously known waveform patterns. We also show that ECG classification greatly benefits from the use of multi-channel data, with nearly all classes and models showing markedly decreased accuracy when only one channel is used.

References

[1] Mant, J., Fitzmaurice, D. A., Hobbs, F. R., Jowett, S., Murray, E. T., Holder, R., ... & Lip, G. Y. (2007). The accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. Bmj, 335(7616), 380.

[2] Rosenblatt, F. (1961). Principles of neurodynamics. Perceptrons and the theory of brain mechanisms (No. VG-1196-G-8). CORNELL AERONAUTICAL LAB INC BUFFALO NY.

[3] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation (No. ICS-8506). California Univ San Diego La Jolla Inst for Cognitive Science.

[4] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[5] Kachuee, M., Fazeli, S., & Sarrafzadeh, M. (2018). ECG Heartbeat Classification: A Deep Transferable Representation. arXiv preprint arXiv:1805.00794.