
Classifying Arrhythmia using Artificial Neural Networks

Siddarth Sampangi

College of Information and Computer Sciences
University of Massachusetts, Amherst
Amherst, MA 01003
ssampangi@cs.umass.edu

Abstract

Cardiac arrhythmia refers to any abnormality in the sequence of electrical signals that govern heartbeats. There are many types of cardiac arrhythmia ranging in severity, including premature beats, atrial fibrillation, and ventricular fibrillation. While arrhythmia classification has been well researched, this study focuses on the use of recent techniques in deep learning to classify arrhythmia with minimal possible data pre-processing.

1 Introduction

2 Related work

3 Network architectures

In this section, we shall provide a brief introduction to convolutional neural networks (CNN) and Long Short-Term Memory (LSTM) networks. These two types of networks are themselves variants of a general class of discriminative classifier called artificial neural networks (ANN). The atomic unit of the ANN is called a neuron. A neuron can receive inputs from any number of real-valued signals, and it produces a real-valued output by taking the inner product of its inputs with a weight matrix W , adding a bias b , and then feeding the result to a nonlinear function. Multiple neurons that each receive identical copies of the input form a layer, and a basic ANN is formed by stacking layers such that the output of the i^{th} layer is the input of the $(i + 1)^{th}$ layer.

An ANN is trained via stochastic gradient descent (SGD) over a loss function (usually some form of $truelabel - prediction$) with respect to the W matrix and bias b of each neuron. To accomplish this, we use the backpropagation algorithm to localize dependencies so that, given a loss function L , we can easily calculate the gradient $\frac{\partial L}{\partial W_i}$ of the weights of a neuron i given the gradients $\frac{\partial L}{\partial W_j}$ of the weights of all neurons j that i outputs to. Thus, the backpropagation algorithm calculates the gradients with respect to the final layer of the network, and then propagates the gradient to each preceding layer. For this to work, it is very important for the nonlinearity used by each neuron to be differentiable.

3.1 Convolutional neural networks

Convolutional networks are designed to exploit the spatial dependencies between units of an input. For example, an image of a flower is recognized as an image of a flower not only because of the RGB values of each pixel, but also because of the position of each pixel in relation to the surrounding pixels. Randomly scrambling the pixel locations in an image of a flower would result in an image that can no longer be recognized. As such, valuable information could exist in the order that an input

is presented to the network. Currently, a vanilla ANN would output the same value regardless of the order of its inputs ($W_1x + W_2y = W_2y + W_1x$).

A one dimensional convolution operation involves a filter of size F , an input of size N , a stride of size S , and a padding of size P . A convolution is performed by mirroring the filter across its axis, and using it as a sliding window over the input, at each step taking the inner product of the filter with the units of the input that the filter’s window resides over at that step. The stride S dictates how many input units to skip at each step, and the padding P dictates the number of zero-padding units that are appended to both ends of the input. Since convolving a filter of size x with an input of size y results in an output of size $y - x + 1$, padding is used to ensure a fixed output size. Thus, the overall formula for calculating output size is $\frac{N-F+2P}{S} + 1$. 1D convolution can be trivially extended to 2D.

Intuitively, a convolution of a filter over an image can be understood as an operation where the filter checks for the presence or response of a specific signal G in the image. This response will be maximal when the filter’s window lies directly over the signal it is trained to respond to. Thus, the output of a convolution can be seen as a sort of heatmap of the likelihood of the presence of G at each point in the image.

A convolutional layer is specified by the number of filters it contains, and the size, stride, and padding (usually common across all filters in a layer) of each filter. The input to the layer is supplied to every filter, and the output of each filter is stacked and outputted to the next layer. A CNN is comprised of one or more convolutional layers, followed by one or more dense layers. A dense layer contains neurons that each receive the entire input of the layer.

3.2 LSTM networks

4 Data preprocessing

We perform the minimal necessary preprocessing to the data. The MIT-BIH Arrhythmia Database [1], found on PhysioNet [2], consists of 48 two-channel ambulatory ECG recordings from patients, each a little over 30 minutes long, and digitized at 360 Hz per channel with 11-bit resolution over a 10 mV range.

For our purposes, the most general purpose algorithm would be able to process and classify the heartbeats in the original format they were recorded in (as 48 separate sequences). While this is ideal, we want to avoid enforcing dependencies between the labels of sequential heartbeats. Since all of the 48 sequences contain largely the same class of heartbeat (either a majority of normal heartbeats, or a majority of abnormal heartbeats), enforcing a heartbeat-to-heartbeat independence would be impossible if the networks were trained on the original (non-randomized) sequences. We could also have trained the networks on randomized versions of the sequences, where each heartbeat of a training sequence is sampled uniformly from all heartbeats across all of the original sequences, but that introduces another set of problems, such as the need to normalize all heartbeats to have the same baseline. Thus, we decided to train our networks on each individual heartbeat.

To define and extract a single heartbeat from a continuous ECG sequence, we decided to extract a series of samples centered around the R peak of the heartbeat, since that would likely contain all relevant features. Since each heartbeat is annotated at its R peak, the only parameter we needed to decide on was the size of the window that we would extract. We determined empirically that the average distance between two R peaks was approximately 0.75 seconds, and so we decided on a window size of 1.0 seconds to allow for a margin of error. Thus, given our dataset’s constraints, we performed the minimal preprocessing.

A window size of 1 second led to nearly all examples containing 361 samples. 176 examples contained 360 samples, and to simplify their processing in the networks, we simply duplicated the last sample to allow a uniform sample size for each heartbeat. Exactly 100 annotated heartbeats could not be used as examples because their R peaks are too close (within 0.5 seconds) to either the beginning of their respective sequence or the end.

Starting with 110631 examples, we removed the 100 examples that did not contain the full 361 samples. We also removed the examples for which the annotation was neither ‘normal’ or some type of arrhythmia. Examples include paced heartbeats, ventricular flutter waves, unclassifiable

heartbeats, and heartbeats that are a fusion of normal and paced or normal and ventricular. After removing these irrelevant examples, we had a set of 93521 examples, of which 18489 (19.7%) are abnormal.

5 Results

6 Future work

7 Conclusion

more history on arrhythmia talk about recent techniques in deep learning briefly mention MIT-BIH and minimal preprocessing and architecture specification

7.1 Citations within the text

Citations within the text should be numbered consecutively. The corresponding number is to appear enclosed in square brackets, such as [1] or [2]-[5]. The corresponding references are to be listed in the same order at the end of the paper, in the **References** section. (Note: the standard `BIBTEX` style `unsrt` produces this.) As to the format of the references themselves, any style is acceptable as long as it is used consistently.

As submission is double blind, refer to your own published work in the third person. That is, use “In the previous work of Jones et al. [4]”, not “In our previous work [4]”. If you cite your other papers that are not widely available (e.g. a journal paper under review), use anonymous author names in the citation, e.g. an author of the form “A. Anonymous”.

7.2 Footnotes

Indicate footnotes with a number¹ in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).²

7.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction; art work should not be hand-drawn. The figure number and caption always appear after the figure. Place one line space before the figure caption, and one line space after the figure. The figure caption is lower case (except for first word and proper nouns); figures are numbered consecutively.

Make sure the figure caption does not get separated from the figure. Leave sufficient space to avoid splitting the figure and figure caption.

You may use color figures. However, it is best for the figure captions and the paper body to make sense if the paper is printed either in black/white or in color.

7.4 Tables

All tables must be centered, neat, clean and legible. Do not use hand-drawn tables. The table number and title always appear before the table. See Table 1.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

¹Sample of the first footnote

²Sample of the second footnote

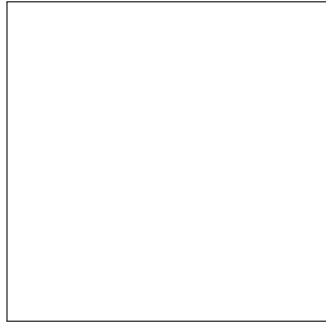


Figure 1: Sample figure caption.

Table 1: Sample table title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

References

- [1] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol* 20(3):45-50 (May-June 2001). (PMID: 11446209)
- [2] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/cgi/content/full/101/23/e215>]; 2000 (June 13).