Reduced-Lead Electrocardiogram Classification using Wavelet Analysis and Deep Learning

Adrian K Cornely, Alondra Carrillo, Grace M Mirsky

Benedictine University, Lisle, IL, USA

Abstract

For the 2021 PhysioNet/Computing in Cardiology Challenge, we developed a deep learning model for automated ECG classification for both the standard 12-lead as well as various reduced lead sets. SqueezeNet was leveraged in order to construct this deep learning model. Training inputs consisted of scalograms, which are time-frequency representations of the absolute value of the continuous wavelet transform coefficients. Multiple preprocessing steps were necessary in order to accommodate the timing and resource constraints of the Challenge. These steps included limiting the training samples to only single diagnosis samples, truncating the signal to only the first five seconds, and only using leads I, II and V2. In the Official Phase, Team Eagles received scores of 0.364, 0.372, 0.372, and 0.372 on 12-lead, 6-lead, 4-lead, 3-lead, 2-lead, respectively. Our preliminary rankings included 40th for the 2-lead, 42nd for the 6-lead, and 45th for the 3-, 4-, and 12-lead ECGs. The main finding from this work is that though wavelet analysis and deep learning can produce high accuracy in a variety of applications, the method applied to this classification task is too computationally intensive and requires substantial improvement to make it clinically viable.

1. Introduction

The standard twelve-lead electrocardiogram (ECG) is a well-established diagnostic tool for detecting cardiac arrhythmias and abnormalities [1]; however, not all hospitals and clinics worldwide have access to this equipment. The 2021 PhysioNet/Computing in Cardiology Challenge focuses on developing automated classification algorithms for a variety of arrhythmias using twelve-lead, six-lead, four-lead, three-lead and two-lead ECGs on a large, diverse dataset [2,3]. Developing algorithms for these different lead configurations will help to determine if robust, accurate classification is possible with reduced-lead ECGs.

Prior work has involved various algorithms to reconstruct the 12-lead ECG from a reduced lead set, with varying degrees of success [4]. Several studies have demonstrated that patient-specific lead derivation was superior to generalized derivations [5], except in the presence of ischemic events [6]. Techniques such as Independent Component Analysis (ICA) have shown to be effective at reconstructing precordial leads [7]. Other studies have cast doubt on the ability of reduced-lead sets to be able to accurately reconstruct precordial leads, particularly when rhythm and morphological abnormalities are present [8]. Though there is some disagreement, the literature as a whole demonstrates that some leads may be used to reconstruct others. This finding implies that the information in the different leads overlaps, at least somewhat, which helps to encourage the current study to determine if accurate classification is possible with a reduced set of leads.

2. Methods

2.1. Preprocessing

The Challenge requires the accurate classification of 30 diagnoses; however, the dataset contains a total of 133 different possible diagnoses. Therefore, there are 103 unscored diagnoses that are present in the training set. In addition, it is important to note that there is not only a single diagnosis per patient; rather, patients have varying numbers of diagnoses, which further complicates the classification. In particular, allowing for a range of between one to ten diagnoses per patient, the number of possible combinations for 30 diagnoses is over 53 million. If we include the unscored diagnoses, there are over 360 trillion possible combinations. That being said, these are the mathematical maxima; in practice, the number of combinations observed clinically is smaller than these maximum values. As a result, due to limitations in run-time imposed by the Challenge, we restricted our training set to only include patients with a single diagnosis from the list of 30 scored diagnoses. For each of these diagnoses, we used normal sinus rhythm as the negative case and the specified diagnosis as the positive case. The only exception was normal sinus rhythm, in which we used various other single diagnosis patient records as the negative cases.
In addition to limiting the number of patient records used, we truncated the signals to a uniform length of five seconds, adjusting for sampling rate. Five seconds was chosen in order to be no longer than the shortest signal length in the training dataset. However, this does require eliminating a substantial amount of data in some patients, since signal length ranged from five seconds to thirty minutes in the training set.

The most significant limitation imposed on the dataset during the preprocessing steps occurred due to the excessively long time necessary to train the models using our algorithm. In particular, our submissions consistently ran out of the allocated time on the scoring system when we attempted to use all 12 leads. Our compromise was to only use training data from leads I, II and V2, which allowed our code to complete training within the prescribed time limits, but obviously put us at a disadvantage since we were only using a fraction of the provided data. The leads used for each of the subsets are shown in Table 1.

Table 1. Leads available for training for each of the reduced lead subsets compared with the actual leads used in this method.

<table>
<thead>
<tr>
<th>#Leads</th>
<th>Available leads</th>
<th>Leads used</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>I, II</td>
<td>I, II</td>
</tr>
<tr>
<td>3</td>
<td>I, II, V2</td>
<td>I, II, V2</td>
</tr>
<tr>
<td>4</td>
<td>I, II, III, V2</td>
<td>I, II, V2</td>
</tr>
<tr>
<td>6</td>
<td>I, II, aVR, aVL, aVF</td>
<td>I, II</td>
</tr>
<tr>
<td>12</td>
<td>I, II, II, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6</td>
<td>I, II, V2</td>
</tr>
</tbody>
</table>

2.2. Models

Our approach utilizes wavelet analysis and transfer learning to create an individual deep learning model for each arrhythmia for each lead, leveraging the method described in [9] for the 2020 Challenge, which was based on the tutorial provided in [10]. We convert the ECG signals to scalograms, which are time-frequency representations of the absolute value of the continuous wavelet transform coefficients plotted over time and frequency.

It is necessary to convert the signals to two-dimensional representations since the deep learning network we use for transfer learning, SqueezeNet, is designed for image classification. These two-dimensional representations are called scalograms; several examples are shown in Figure 1. SqueezeNet is a well-known convolutional neural network, which has comparable accuracy to AlexNet [11] when evaluated on ImageNet data; however, SqueezeNet has several advantages over AlexNet, including having fewer parameters, being a smaller size model, having greater possible platform options and requiring less bandwidth to export the model [12].

MATLAB provides a built-in implementation of SqueezeNet in the Deep Learning Toolbox, which is available in the base R2020a version of the toolbox, further simplifying its use since support packages do not need to be installed in order to use SqueezeNet in MATLAB [13].

For the solver, we used the stochastic gradient descent with momentum (SGDM) optimizer, due to its fast, robust
convergence [14,15]. The specific parameters used to construct the model are described in detail in [9].

2.3. Voting

To assign a particular diagnosis, we require a minimum of one-third of the vote of the available leads. This design allows for the assignment of multiple concurrent diagnoses and also easily allows for additional arrhythmias to be included at a later time without any modification required to the existing classification system. However, it should be noted that this voting scheme was not practically used in our accepted submission since the maximum number of leads used was three. Nevertheless, we include it here since the confidence in the classification should be increased by using a voting scheme [16], so if the described method were applied using all twelve leads, we expect that the voting scheme would help to increase confidence in our results.

3. Results

A detailed explanation of the scoring algorithms used in the Challenge can be found in [2,3]. The scores for the best-performing entry for both the unofficial and official phases are shown in Table 2. The ranking for each event in the official phase is shown below in Table 3.

Table 2. Best scores for Team Eagles per lead subset in both phases of the Challenge.

<table>
<thead>
<tr>
<th># of Leads</th>
<th>Unofficial Phase</th>
<th>Official Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.351</td>
<td>0.372</td>
</tr>
<tr>
<td>3</td>
<td>0.352</td>
<td>0.372</td>
</tr>
<tr>
<td>4</td>
<td>n/a</td>
<td>0.372</td>
</tr>
<tr>
<td>6</td>
<td>0.352</td>
<td>0.372</td>
</tr>
<tr>
<td>12</td>
<td>0.344</td>
<td>0.364</td>
</tr>
</tbody>
</table>

Table 3. Ranking for Team Eagles out of the 58 teams in the Official Phase of the 2021 PhysioNet/Computing in Cardiology Challenge.

<table>
<thead>
<tr>
<th># of Leads</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>40th</td>
</tr>
<tr>
<td>3</td>
<td>45th</td>
</tr>
<tr>
<td>4</td>
<td>45th</td>
</tr>
<tr>
<td>6</td>
<td>42nd</td>
</tr>
<tr>
<td>12</td>
<td>45th</td>
</tr>
</tbody>
</table>

4. Discussion and Conclusions

The most obvious limitation of our approach is the use of only three leads for classification. Our method was too computationally intensive relative to the Challenge design to allow us to run the algorithm over all twelve leads. In addition, by limiting the signal length to the first five seconds, we are potentially creating signals for the training dataset without the specified arrhythmias present, which would occur if the arrhythmias are only present later in the signal.

In order to improve the accuracy of this method, the efficiency must be dramatically improved in order to be able to include models from all twelve leads, rather than only three select leads. Training the model offline would make this possible, but was unfortunately not allowed according to the Challenge rules. The voting methodology would likely provide more benefit had we been able to create models for all twelve leads, which would allow for a more reliable consensus decision.

Benefits of this approach include the ability to leverage transfer learning and the small size of SqueezeNet. Nevertheless, while our results show some promise, even given the substantial limitations described in the methods section above, there is clearly significant room for improvement in our team’s approach, particularly with respect to efficiency.

Acknowledgments

The authors would like to express their sincere thanks to the 2021 PhysioNet/Computing in Cardiology Challenge organizers for putting together a rigorous, valuable competition.

References


Address for correspondence:

Dr. Grace Mirsky
Benedictine University
5700 College Road
Lisle, IL 60532, USA
gmirsky@ben.edu