Semi-Supervised Learning for ECG Classification

Rui Rodrigues[1,2], Paula Couto[1,2]

[1] Center for Mathematics and Applications (CMA), FCT NOVA, Caparica, Portugal
[2]Department of Mathematics, Caparica, FCT NOVA, Portugal

Abstract

We present an approach for automatic cardiac abnormality detection using two leads ECG. This approach was developed in the context of the Physionet/Computing in Cardiology Challenge 2021.

Our model is decomposed into an Encoder and a Decoder. It is a huge neural network model with more than 36 million parameters. Although the Challenge training dataset consists of more than 88 thousand annotated ECGs, our model is extremely prone to overfitting to the training data.

The encoder is a convolution neural network followed by three transformer encoder blocks. The decoder is a transformer encoder block followed by a feedforward neural network.

To reduce the overfitting, we pretrain the Encoder in a semi-supervised way on three tasks. Given an ECG segment, L_1, the first task is to detect the QRS on L_1; the second task is to predict the ECG shape on an ECG segment, L_2 following L_1, given the QRS location on L_2; the third task is to predict the number of samples, after L_1, before the next QRS. The Decoder weights were firstly estimated with the frozen Endoder pre-trained parameters and then the whole model parameters were fine-tunned.

Our team, named matFCT, received a challenge score of 0.43 on the official test dataset. However, we were unable to qualify for ranking because we weren't able to submit the preprint to the Computing in Cardiology Conference before the deadline.

1. Introduction

Cardiovascular diseases are the leading cause of premature deaths and disabilities. The PhysioNet/Computing in Cardiology Challenge 2021, focused on automatic, open-source methods to detect cardiac abnormalities from ECGs using a reduced set of leads[1–3]. In that context, we present an approach for automatic cardiac abnormality classification from 2-leads ECG.

Complex tasks like the challenge focus are the natural target of large and complex machine learning models. However, models with many parameters tend to overfit to the training data. We present a way to fit the parameters of a huge neural network classification model to the challenge training data [4–9].

2. Methods

2.1. Pre-processing

We used only two ECG leads (I and II) and the signals were downsampled to 100 Hz. Otherwise, the input to our model is the raw signal. All the features used for classification are extracted and learned by our model.

2.2. Our Classification Model

Our Classification model consists of an encoder followed by a decoder, see fig. 1. It has 36,553,114 trainable

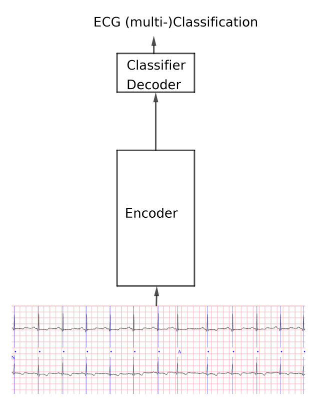

Figure 1. The Classification Model

parameters. Although the challenge training data consists of more than 88 thousand labeled examples, our model is extremely prone to overfitting.

Both encoder and decoder use the Transformer Encoder block [10], see fig. 2. Embedding dimension and feedforward neural network hidden layer number of units of the

Transformer Encoder block are always equal to 512 and it has 8 attention heads.

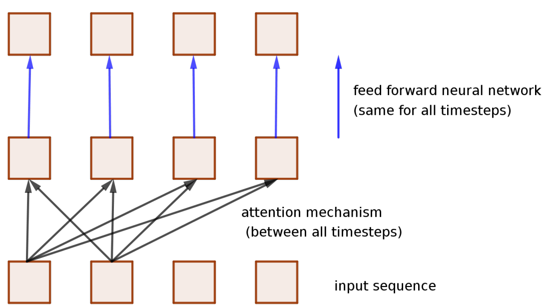

Figure 2. The Transformer Encoder Block

The encoder is a Convolution Neural Network followed by a Sequence of three Transformer Encoder blocks, see fig. 3.

The Convolution Neural Network has three layers with 128, 256 and 512 filters. In each convolution layer we applied batch normalization followed by 1D-maxpooling.

The encoder has 27,348,352 trainable parameters.

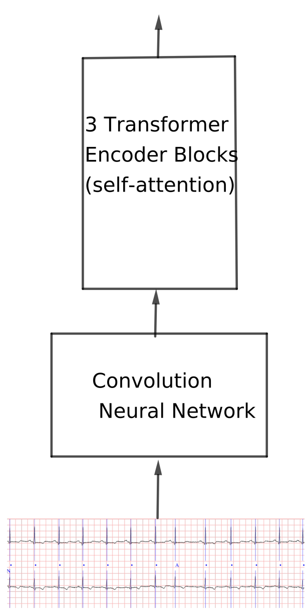

Figure 3. The Encoder

The decoder first component is a transformer encoder block. The first Transformer Encoder Block output timestep is used as input to a single hidden layer feed forward neural network. The decoder output is the ECG classification. See fig. 4.

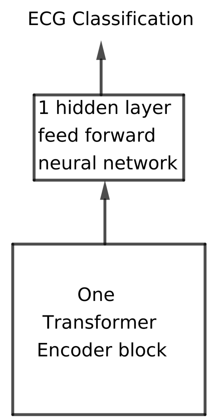

Figure 4. The Decoder

2.3. Training the model

2.3.1. Semi-Supervised Pre-Training of Encoder Parameters

To reduce overfitting to the training dataset we *pre-train* the encoder parameters in a *semi-supervised* way [11].

Three tasks were used in the pre-training process and, for each task, a decoder, like the Classification Model decoder, is added to the encoder.

For each ECG example, we use an initial segment, L_1, starting at the beginning of the ECG and ending in a moment randomly chosen between $\frac{2}{5}$ and $\frac{6}{7}$ of the ECG signal length. This initial segment is the input to the Encoder. On two of the three tasks we use a segment, L_2, 300 *ms* long, starting after L_1. The three encoder pre-training tasks are:

- *Task 1*: **Detect the QRS** in the first 3 seconds of the L_1 segment.
- *Task 2*: **predict the ECG shape in the segment** L_2. For this task, together with the encoder output, the QRS location in the segment L_2, or the information about its absence, is given to this task decoder.
- *Task 3*: **predict the number of samples until the next QRS** after L_1.

Figure 5 represents the model used in the encoder pre-training.

The three tasks used to pre-train the encoder impose a large number of constraints on the encoder weights that help to reduce overfitting. They also create an ECG representation on the encoder top layer that contains enough information to realize the three tasks by the decoder like models.

To train the model's weights we used the Mean Square Error loss on the three tasks with the Adam optimizer.

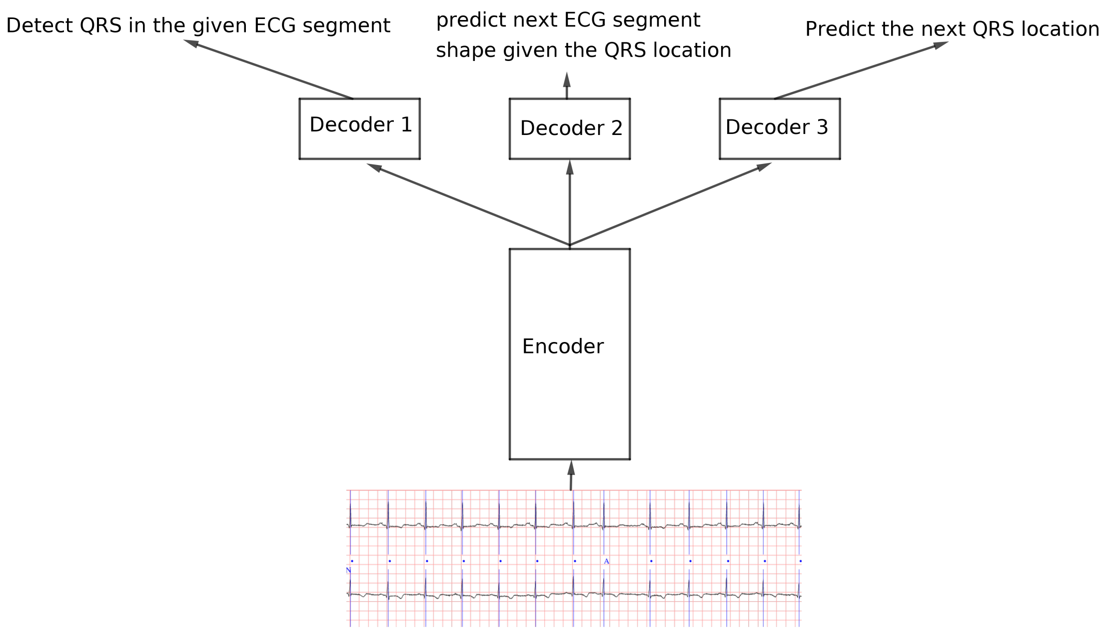

Figure 5. Model used in the encoder pre-training process

2.3.2. Pre-Training Data

As pre-training data we used the ECGs from the Challenge training dataset, ignoring the labels.

QRS locations for the three tasks were obtained using George Moody's *gqrs* detector.

When predicting the ECG shape on the above defined segment L_2 we attenuated the ECG baseline wander in L_2 subtracting from each ECG channel the result of a moving average filter one second long applied to the same channel.

2.3.3. The Classification Model Training Process

- To adjust the Classification Model's parameters to the Challenge Training Dataset we used as loss function a sotf(continuous) version of the unnormalized Challenge scoring metric [2, 12]. We chose the Adam optimizer.
- Starting with frozen Encoder parameters we train the decoder parameters in the Challenge Training Dataset (using the labels).
- Next, we train again the all model, with unfrozen encoder weights. To reduce overfitting, after three warmup epochs (learning rate: 10^{-8}, 10^{-7} and 10^{-6}), we trained 5 epochs with learning rate equal to 3×10^{-6}.

3. Results

F1-score for each class, on the train data 5-fold cross validation, are displayed on the table 1.

Table 2 contains our Challenge scores. We only report results on two leads because we did not use more ECG leads: the results on other lead combinations are the same.

We were unable to qualify for ranking because we weren't able to submit the preprint to the Computing in Cardiology Conference before the deadline.

4. Discussion

The use of semi-supervised training of ECG classifiers is a promising approach as it allows for the use of large complex and powerfull models. In the future, other tasks than the three we used should also be applied to pre-train the encoder. In particular, these new tasks should target a better result on classes that have a worse F1-score.

We limitated our approach to two lead ECG, it should be extended to other ECG lead subsets.

Acknowledgments

This work is funded by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the project UIDB/00297/2020 (Center for Mathematics and Applications

classe	AF	AFL	BBB	Brady	LBBB—CLBBB	RBBB—CRBBB
F1-score	0.437	0.609	0.092	0.000	0.397	0.448
classe	IAVB	IRBBB	LAD	LAnFB	LPR	LQRSV
F1-score	0.377	0.059	0.495	0.356	0.000	0.174
classe	LQT	NSIVCB	NSR	PAC—SVPB	PR	PRWP
F1-score	0.128	0.083	0.649	0.109	0.311	0.046
classe	VPB—PVC	QAb	RAD	SA	SB	STach
F1-score	0.133	0.000	0.177	0.000	0.815	0.650
classe	TAb	TInv				
F1-score	0.418	0.262				

Table 1. F1-score for each class, on the train data 5-fold cross validation (two leads only).

dataset	challenge score
5-fold cross validation on train data	0.55 ± 0.02
validation set	0.43
test set	0.43

Table 2. Challenge scores for our final entry (team matFCT), same score for all lead combination (we only used two leads).

References

[1] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000;101(23):e215–e220.

[2] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, Liu F, Rad BA, Elola A, Seyedi S, Li Q, Sharma A, Clifford GD, Reyna MA. Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. Physiological Measurement 2020;41.

[3] Reyna MA, Sadr N, Perez Alday EA, Gu A, Shah A, Robichaux C, Rad BA, Elola A, Seyedi S, Ansari S, Ghanbari H, Li Q, Sharma A, Clifford GD. Will Two Do? Varying Dimensions in Electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021. Computing in Cardiology 2021;48:1–4.

[4] Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu X, Liu Y, Ma C, Wei S, He Z, Li J, Kwee ENY. An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. Journal of Medical Imaging and Health Informatics 2018; 8(7):1368—1373.

[5] Tihonenko V, Khaustov A, Ivanov S, Rivin A, Yakushenko E. St Petersburg INCART 12-lead Arrhythmia Database. PhysioBank PhysioToolkit and PhysioNet 2008;Doi: 10.13026/C2V88N.

[6] Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomedizinische Technik 1995;40(S1):317–318.

[7] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, Schaeffter T. PTB-XL, a Large Publicly Available Electrocardiography Dataset. Scientific Data 2020;7(1):1–15.

[8] Zheng J, Zhang J, Danioko S, Yao H, Guo H, Rakovski C. A 12-lead Electrocardiogram Database for Arrhythmia Research Covering More Than 10,000 Patients. Scientific Data 2020;7(48):1–8.

[9] Zheng J, Cui H, Struppa D, Zhang J, Yacoub SM, El-Askary H, Chang A, Ehwerhemuepha L, Abudayyeh I, Barrett A, Fu G, Yao H, Li D, Guo H, Rakovski C. Optimal Multi-Stage Arrhythmia Classification Approach. Scientific Data 2020;10(2898):1–17.

[10] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Lu, Polosukhin I. Attention is all you need. In Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds.), Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc., 2017; .

[11] Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In Burstein J, Doran C, Solorio T (eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers). Association for Computational Linguistics, 2019; 4171–4186. URL https://doi.org/10.18653/v1/n19-1423.

[12] Reyna MA, Alday EAP, Gu A, Liu C, Seyedi S, Rad AB, Elola A, Li Q, Sharma A, Clifford GD. Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. In 2020 Computing in Cardiology (CinC), volume 47. IEEE, 2020; 1–4.

Address for correspondence:

Rui Rodrigues

Departamento de Matemática, Faculdade de Ciências e Tecnologia

2829-516 Caparica, Portugal

rapr@fct.unl.pt