

Wettbewerb künstliche Intelligenz in der Medizin

ECG-DualNet



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Attention Is All You Need

Christoph Reich & Vladislav Kruk

Supervisors

Prof. Dr.-Ing. Christoph Hoog Antink

Maurice Rohr, M. Sc.



1. Introduction

- Problem Setting
- Motivation
- Related Work

2. Method

- ECG-DualNet
- Augmentation Pipeline
- Training & Validation

3. Experiments

- Datasets
- ECG-DualNet Results
- Ablation Study

4. Conclusion & Discussion

1. Introduction

- Problem Setting
- Motivation
- Related Work

2. Method

- ECG-DualNet
- Augmentation Pipeline
- Training & Validation

3. Experiments

- Datasets
- ECG-DualNet Results
- Ablation Study

4. Conclusion & Discussion

Classify single-lead ECG signals with variable length.

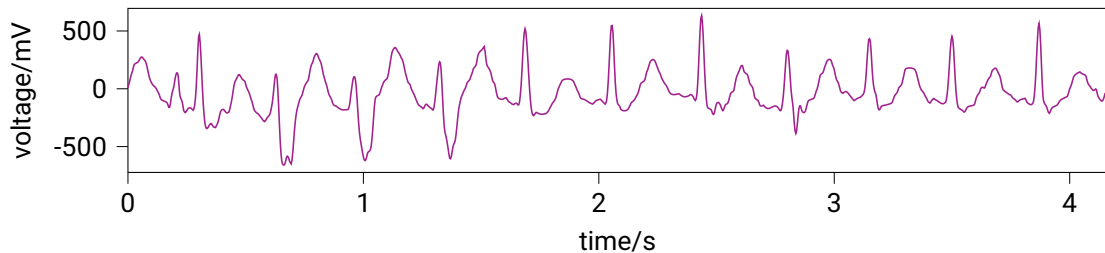


Figure: ECG signal of the 2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017] labeled as AF.



- Atrial fibrillation (AF) dangerous and often undetected
- AF one of the most common heart arrhythmia's
- AF can lead to strokes, dementia, and heart failure
- Increasing amount of single-lead ECG edge devices available
 - No expert knowledge typically available
 - Need for automated AF

[Becker, 2006]

[Herold, 2019]

- Atrial fibrillation (AF) dangerous and often undetected
- AF one of the most common heart arrhythmia's
- AF can lead to strokes, dementia, and heart failure
- Increasing amount of single-lead ECG edge devices available
 - No expert knowledge typically available
 - Need for automated AF

[Becker, 2006]

[Herold, 2019]

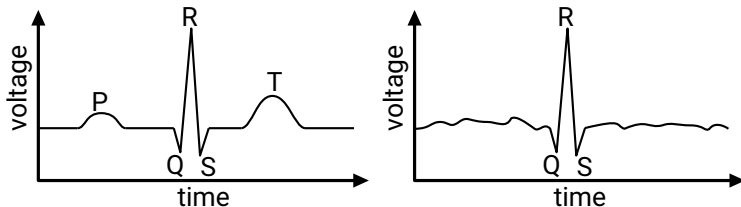


Figure: Regular heart beat left and atrial fibrillation on the right.

Introduction

Related Work

Transitional ML approaches

Deep learning approaches



Transitional ML approaches

- Preprocessing & Feature extraction
 - Data augmentation
 - ECG timing features
 - Robust interval features
 - Waveform features
- Learnable classifier
 - Random forest
 - Support vector machines
 - XGBoost
- [Hoog Antink et al., 2017, Smíšek et al., 2017]

Deep learning approaches

Transitional ML approaches

- Preprocessing & Feature extraction
 - Data augmentation
 - ECG timing features
 - Robust interval features
 - Waveform features
- Learnable classifier
 - Random forest
 - Support vector machines
 - XGBoost
- [Hoog Antink et al., 2017, Smíšek et al., 2017]

Deep learning approaches

- Preprocessing
 - Data augmentation
 - Data conversion (Spectrogram)
- Deep learning classifier
- [Zihlmann et al., 2017, Mousavi et al., 2019, Mashrur et al., 2019, Khriji et al., 2020, Nonaka and Seita, 2020]

1. Introduction

- Problem Setting
- Motivation
- Related Work

2. Method

- ECG-DualNet
- Augmentation Pipeline
- Training & Validation

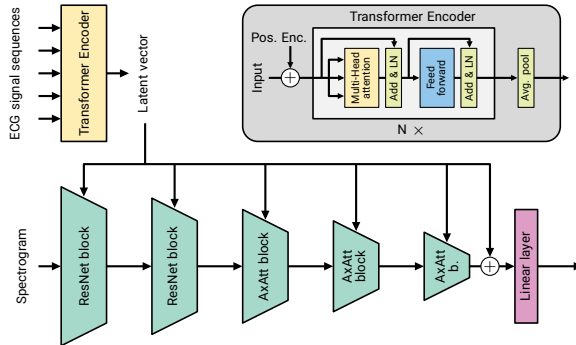
3. Experiments

- Datasets
- ECG-DualNet Results
- Ablation Study

4. Conclusion & Discussion

Method

ECG-DualNet(++)



[Vaswani et al., 2017]
[Wang et al., 2020]
[Dosovitskiy et al., 2020]
[de Vries et al., 2017]

Figure: ECG-DualNet++ architecture with signal and spectrogram encoder

Method

Augmentation Pipeline

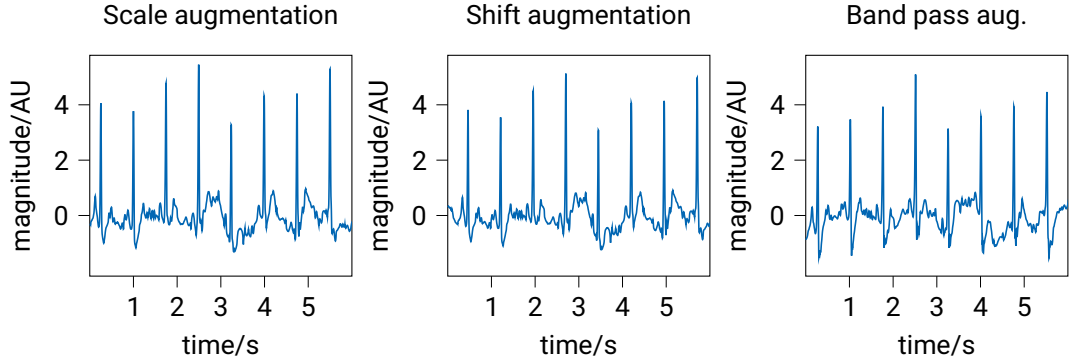


Figure: Different augmentations applied to ECG signal of the 2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017].

Method

Augmentation Pipeline

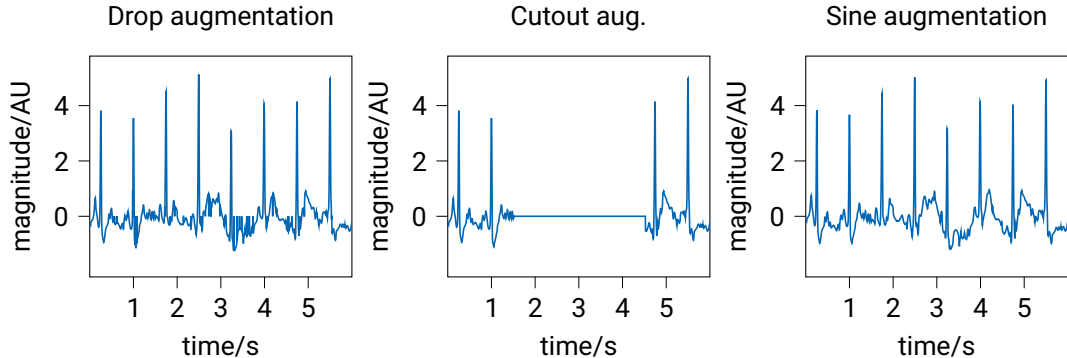


Figure: Different augmentations applied to ECG signal of the 2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017].

Method

Augmentation Pipeline

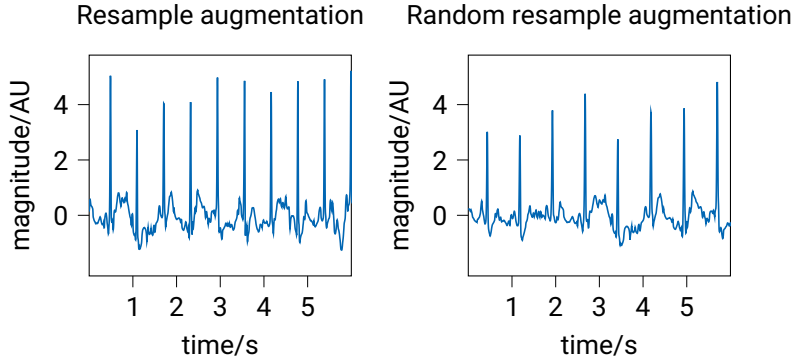


Figure: Different augmentations applied to ECG signal of the 2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017]. Random resample aug. inspired by 2D random elastic aug. [Simard et al., 2003].



Training loss (weighted cross entropy loss)

$$\mathcal{L} = -\frac{1}{N} \sum_{j=1}^N \sum_{i=1}^4 \alpha_i y_{ji} \log(\hat{y}_{ji})$$

Training loss (weighted cross entropy loss)

$$\mathcal{L} = -\frac{1}{N} \sum_{j=1}^N \sum_{i=1}^4 \alpha_i y_{ji} \log(\hat{y}_{ji})$$

Validation metrics (accuracy & F1)

$$\text{ACC} = \frac{1}{n} \sum_{j=1}^n \delta(\arg \max(\mathbf{y}_j), \arg \max(\hat{\mathbf{y}}_j)), \quad \text{F1} = \frac{1}{4} \sum_{i=1}^4 \frac{2\text{TP}_i}{2\text{TP}_i + \text{FP}_i + \text{FN}_i}$$

1. Introduction

- Problem Setting
- Motivation
- Related Work

2. Method

- ECG-DualNet
- Augmentation Pipeline
- Training & Validation

3. Experiments

- Datasets
- ECG-DualNet Results
- Ablation Study

4. Conclusion & Discussion



2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017]

- Single-lead ECG signals with variable length (2714 - 18286 samples)
- 8529 publically available data samples (7000 train & 1528 val.)
- Labels include four classes (normal, AF, other & noisy)

2017 PhysioNet/CinC Challenge dataset [Clifford et al., 2017]

- Single-lead ECG signals with variable length (2714 - 18286 samples)
- 8529 publically available data samples (7000 train & 1528 val.)
- Labels include four classes (normal, AF, other & noisy)

Icentia11k [Tan et al., 2019]

- Single-lead ECG signals of 11k patients.
- 550k data samples with a length of 1h
- Sparse six class rhythm labels including AF
- Cropped and resampled to match target dataset

Results

ECG-DualNet Results

Table: Classification results of our proposed approaches and baselines on the 2017 PhysioNet validation set.

| Model | ACC \uparrow | F1 \uparrow | # Parameters |
|--|----------------|--------------------|--------------|
| CNN baseline* [Zihlmann et al., 2017] | 0.812 | 0.790 [†] | ~ 3.5M |
| CRNN baseline* [Zihlmann et al., 2017] | 0.823 | 0.792 [†] | ~ 3.5M |
| ECG-DualNet S | 0.8527 | 0.8049 | 1.8M |
| ECG-DualNet M | 0.8560 | 0.7938 | 4.3M |
| ECG-DualNet L | 0.8514 | 0.8038 | 6.2M |
| ECG-DualNet XL | 0.8612 | 0.8164 | 20.7M |
| ECG-DualNet++ S | 0.8174 | 0.7291 | 1.8M |
| ECG-DualNet++ M | 0.8259 | 0.7730 | 2.6M |
| ECG-DualNet++ L | 0.8449 | 0.7859 | 3.7M |
| ECG-DualNet++ XL | 0.8593 | 0.8051 | 8.2M |
| ECG-DualNet++ 130M | 0.8534 | 0.7963 | 128M |

* Reported literature values.

[†] F1 score computed over three classes, thus not directly comparable.

Results

ECG-DualNet Results Pre-Training

Table: Classification results of our proposed approaches on the lcentia11k validation set. Only a single training for each model run was conducted.

| Model | ACC \uparrow | F1 \uparrow |
|------------------|----------------|---------------|
| ECG-DualNet XL | 0.8989 | 0.4564 |
| ECG-DualNet++ XL | 0.8899 | 0.4970 |

Results

ECG-DualNet Results Pre-Training

Table: Classification results of our proposed approaches on the Icentia11k validation set. Only a single training for each model run was conducted.

| Model | ACC \uparrow | F1 \uparrow |
|------------------|----------------|---------------|
| ECG-DualNet XL | 0.8989 | 0.4564 |
| ECG-DualNet++ XL | 0.8899 | 0.4970 |

Table: Classification results of our proposed approaches on the 2017 PhysioNet validation set and pre-trained on the Icentia11k dataset.

| Model | ACC \uparrow | F1 \uparrow |
|------------------|----------------|---------------|
| ECG-DualNet XL | 0.8534 | 0.8024 |
| ECG-DualNet++ XL | 0.8534 | 0.8004 |

Results

Ablation Study

Table: Classification results on the 2017 PhysioNet validation for different ablations. ECG-DualNet L configuration utilized.

| Data aug. & dropout | Signal encoder | Spectrogram encoder | ACC \uparrow | F1 \uparrow |
|---------------------|----------------|---------------------|----------------|---------------|
| \times | \checkmark | \checkmark | 0.8272 | 0.7493 |
| \checkmark | \times | \checkmark | 0.8440 | 0.7855 |
| \checkmark | \checkmark | \times | 0.7264 | 0.5813 |
| \checkmark | \checkmark | \checkmark | 0.8560 | 0.7938 |

1. Introduction

- Problem Setting
- Motivation
- Related Work

2. Method

- ECG-DualNet
- Augmentation Pipeline
- Training & Validation

3. Experiments

- Datasets
- ECG-DualNet Results
- Ablation Study

4. Conclusion & Discussion

Achievements

- Presented the novel ECG-DualNet for ECG classification
- Proposed an advanced augmentation pipeline
- Performed extensive testing (against various beelines) including pre-training

Achievements

- Presented the novel ECG-DualNet for ECG classification
- Proposed an advanced augmentation pipeline
- Performed extensive testing (against various beelines) including pre-training

Observations

- Highly overparameterized models (ECG-DualNet++ 130M) does not overfit in the classical sense → Deep Double Descent [[Nakkiran et al., 2020](#)] ?
- Spectrogram encoder is the most crucial part of ECG-DualNet
- Extensive pre-training on the lcentia11k does not lead to performance benefits on the target 2017 PhysioNet dataset



Becker, D. E. (2006).
Fundamentals of electrocardiography interpretation.
Anesthesia progress, 53(2):53–64.



Clifford, G. D., Liu, C., Moody, B., Li-wei, H. L., Silva, I., Li, Q., Johnson, A., and Mark, R. G. (2017).
AF Classification from a Short Single Lead ECG Recording: the PhysioNet/Computing in
Cardiology Challenge 2017.
In 2017 Computing in Cardiology (CinC), pages 1–4. IEEE.



de Vries, H., Strub, F., Mary, J., Larochelle, H., Pietquin, O., and Courville, A. (2017).
Modulating early visual processing by language.
In Advances in Neural Information Processing Systems, pages 6597–6607.



Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. (2020).

An image is worth 16x16 words: Transformers for image recognition at scale.
preprint arXiv:2010.11929.



Herold, G. (2019).

Innere Medizin.

Walter de Gruyter GmbH & Co KG.



Hoog Antink, C., Leonhardt, S., and Walter, M. (2017).

Fusing QRS Detection and Robust Interval Estimation with a Random Forest to Classify Atrial Fibrillation.

In 2017 Computing in Cardiology (CinC), pages 1–4. IEEE.



Khriji, L., Fradi, M., Machhout, M., and Hossen, A. (2020).
Deep Learning-based Approach for Atrial Fibrillation Detection.
In International Conference on Smart Homes and Health Telematics, pages 100–113. Springer.



Mashrur, F. R., Roy, A. D., and Saha, D. K. (2019).
Automatic identification of arrhythmia from ecg using alexnet convolutional neural network.
In 2019 4th International Conference on Electrical Information and Communication Technology (EICT), pages 1–5. IEEE.



Mousavi, S., Afghah, F., Razi, A., and Acharya, U. R. (2019).
ECGNET: Learning Where to Attend for Detection of Atrial Fibrillation with Deep Visual Attention.
In 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pages 1–4. IEEE.



Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., and Sutskever, I. (2020).
Deep double descent: Where bigger models and more data hurt.
In International Conference on Learning Representations.



Nonaka, N. and Seita, J. (2020).
Data Augmentation for Electrocardiogram Classification with Deep Neural Network.
preprint arXiv:2009.04398.



Simard, P. Y., Steinkraus, D., Platt, J. C., et al. (2003).

Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis.
In *Icdar*, volume 3.



Smíšek, R., Hejč, J., Ronzhina, M., Němcová, A., Maršánová, L., Chmelík, J., Kolářová, J.,
Provazník, I., Smital, L., and Vítek, M. (2017).

Svm based ecg classification using rhythm and morphology features, cluster analysis and
multilevel noise estimation.

In *2017 Computing in Cardiology (CinC)*, pages 1–4.



Tan, S., Androz, G., Chamseddine, A., Fecteau, P., Courville, A., Bengio, Y., and Cohen, J. P. (2019).

Icentia11k: An unsupervised representation learning dataset for arrhythmia subtype discovery.
preprint arXiv:1910.09570.



Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., and Polosukhin, I. (2017).

Attention is all you need.

In Advances in Neural Information Processing Systems, volume 30.



Wang, H., Zhu, Y., Green, B., Adam, H., Yuille, A., and Chen, L.-C. (2020).

Axial-DeepLab: Stand-Alone Axial-Attention for Panoptic Segmentation.
In European Conference on Computer Vision, pages 108–126.



Zihlmann, M., Perekrestenko, D., and Tschannen, M. (2017).
Convolutional Recurrent Neural Networks for Electrocardiogram Classification.
In *2017 Computing in Cardiology (CinC)*, pages 1–4. IEEE.

Code, weights, and paper are available at:
github.com/ChristophReich1996/ECG_Classification



Atrial Fibrillation Classification in Electrocardiography using Deep Learning

Christoph Reich¹, Vladislav Kruk
Department of Electrical Engineering and Information Technology,
Technische Universität Darmstadt
¹christoph.reich@inf.tu-darmstadt.de

Abstract—The evaluation of electrocardiogram recordings is the most common approach to diagnose and monitor cardiac arrhythmias such as atrial fibrillation. The electrocardiogram evaluation typically requires expert knowledge, which is not always available. We present a novel approach for the automated classification of atrial fibrillation in electrocardiogram recordings with variable length. Our deep learning approach utilizes both input data in the time and frequency domain. The performance of the proposed dual network for ECG classification (ECG-DualNet) is showcased on the 2017 PhysioNet/CinC Challenge dataset. ECG-DualNet outperforms recent Convolutional Neural Network approaches in terms of classification accuracy. Code and trained models are available at https://github.com/ChristophReich1996/ECG_Classification.

Index Terms—deep learning, attention, arrhythmia classification, atrial fibrillation classification, electrocardiography.

I. INTRODUCTION

Electrocardiography (ECG) is the most important tool for the diagnosis and the monitoring of cardiac arrhythmias [1]–[3]. The first recorded human heart beat dates back to the late 19th century [3]. Today 12-lead ECGs are the common standard [3]. The analysis of ECG recordings, especially the detection of cardiac arrhythmias, requires expert knowledge [1]. This expert knowledge is sometimes not available. Since a fast and accurate diagnosis of cardiac arrhythmias can highly affect the

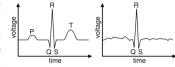


Fig. 1: Regular heart beat with QRS complex, P wave, and T wave on the left and atrial fibrillation on the right. AF can be detected by the noisy ECG at the P and T wave regions. Please note that this is only a theoretical illustration. Real ECGs of AF can be seen in the appendix.

II. RELATED WORK

Recent approaches for the task of AF classification in ECG recordings can be clustered in two groups. First, classical machine learning approaches [4], [9], and second, deep learning approaches [5], [10]–[13]. In general, deep learning approaches achieve better classification accuracy, however, sacrifice explainability [5], [10], [14], [15].

Classical machine learning approaches typically extract features, first and classify them in a second learnable step.

Challenge Submission

Backup

- ECG-DualNet XL utilized
- Pre-training on Icentia11k
- Optimized training (8000 samples) and validation (528 samples) split used

Table: Classification results of ECG-DualNet XL pre-trained on the Icentia11k dataset and fine-tuned on the PhysioNet dataset with optimized submission split. Metric computed on the small validation set.

| Model | ACC ↑ | F1 ↑ |
|----------------|--------|--------|
| ECG-DualNet XL | 0.8840 | 0.8549 |

Deep Double Decent

Backup

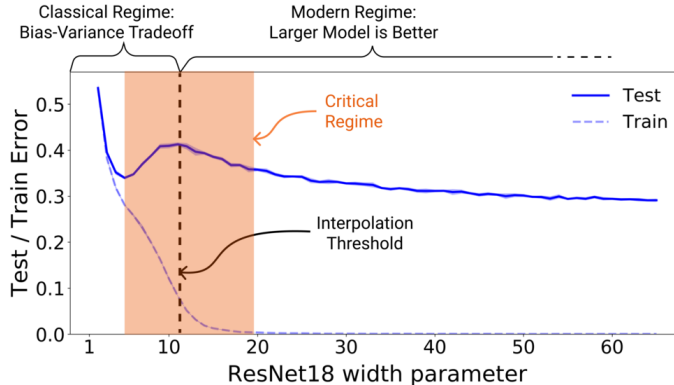


Figure: Illustration of the deep double decent phenomenon in image classification (CIFAR-10 & 15% label noise). Image taken from [Nakkiran et al., 2020].