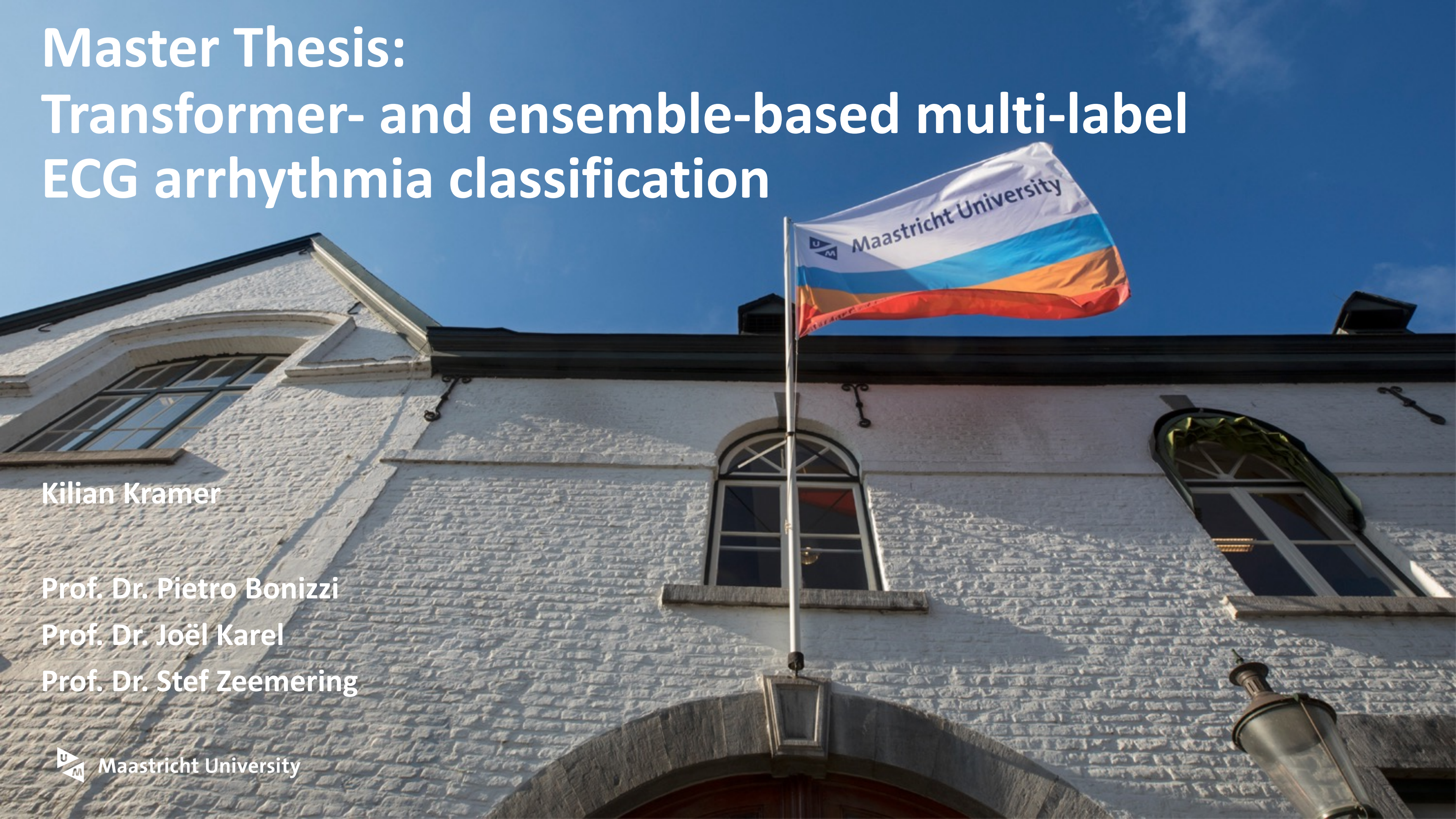
Master Thesis:
Transformer- and ensemble-based multi-label ECG arrhythmia classification



Kilian Kramer

Prof. Dr. Pietro Bonizzi
Prof. Dr. Joël Karel
Prof. Dr. Stef Zeemering

Maastricht University

Index

- Problem Introduction
- Related Work
- Research Questions
- Methodology
- Experiments and Evaluation
- Conclusion

Problem Introduction

- Cardiovascular diseases, such as Atrial Fibrillation, are among the leading causes of death in the population, resulting in an increased demand for cardiac assessment

- Deep learning can be used to develop multi-classification models for arrhythmia detection and improve medical monitoring

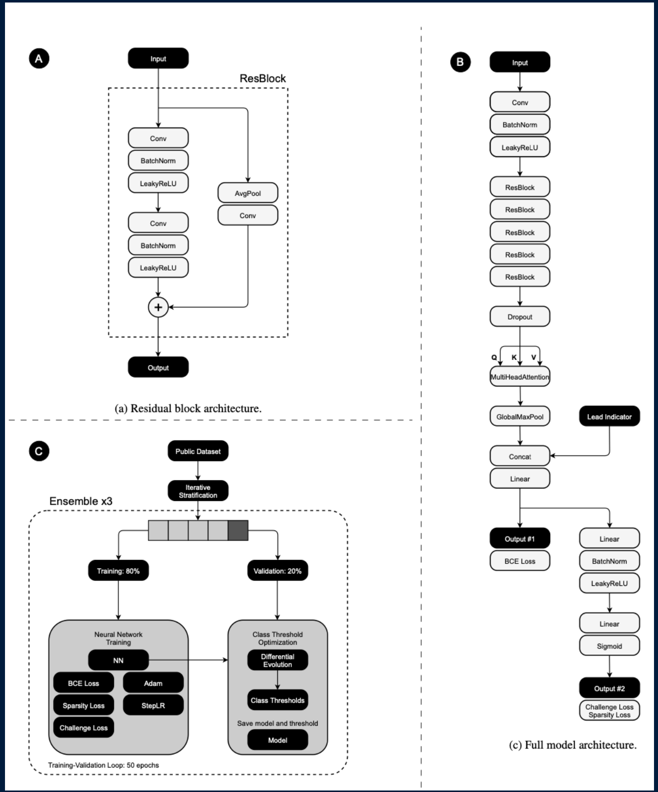
- Today, much research and accurate models are available for simpler arrhythmia classification tasks, e.g. AFIB classification (binary) or common arrhythmia types grouped (AAMI standard)

Problem Introduction

- Problem: Professional treatment requires individual and detailed ECG assessment
- In recent years, Transformer models have gained considerable popularity due to the self-attention mechanism and research papers that apply Transformer models on less comprehensive ECG arrhythmia classification tasks show good results
- However, research and accurate models are limited on comprehensive arrhythmia detection tasks including Transformer models and rare subdiseases, such as Atrial Flutter, Premature Ventricular Contractions, Prolonged QT interval etc.

Related Work: Classification of ECG Using Ensemble of Residual CNNs with Attention Mechanism

- winning paper of the Physionet 2021 challenge
- prop. by Nejedly et. Al
- Uses several residual CNN block and an encoder-block
- Specific designed loss function that fits to class weights in challenge evaluation metrics
- 58% accuracy on all subtasks (2-, 3-, 4-, 6- and 12-leads)
- Follow up study by authors shows that encoder-block does not improve performance (59% accuracy without)



Related Work: ECGBERT: Understanding Hidden Language of ECGs with Self-Supervised Representation Learning

- prop. by Choi et. al based on BERT methodology
- uses wave segment vocabulary that is created from 70 clusters using K-mean and Dynamic Time Warping (4 classifiers: P, QRS, T and background waves, such as PR or ST intervals) -> a heartbeat is composed of several wave segments
- ECG signal splitted by segments and each segment is assigned to a clusters and then a pre-defined token. This procedure creates training data that form sentences of repetitive wave tokens (max 8 consecutive heartbeats)
- Additionally!, CNN features from raw ECG are fed into the Transformer encoder using an U-Net
- Model is pre-trained with masking (similar to Masked Language Modeling in LLMs)
- Model is fine-tuned on several down-stream tasks by adapting/fine-tuning the output layers:
- Afib classification (binary), heartbeat classification (AAMI standard), ECG patient verification and sleep apnea detection

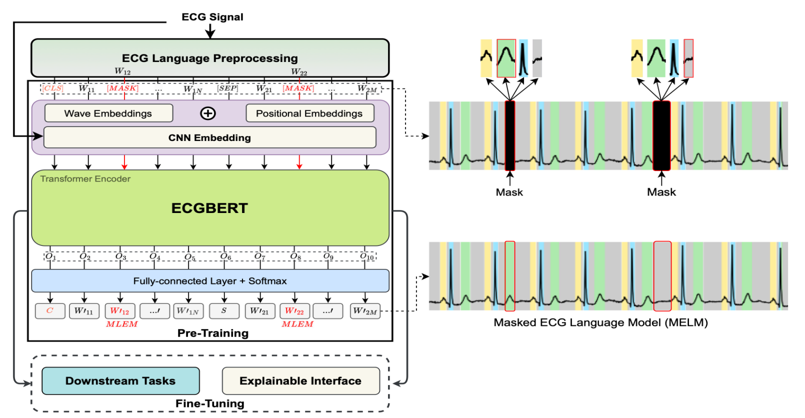


Figure 2: Illustration of the proposed ECGBERT architecture

Maastricht University

Research Questions

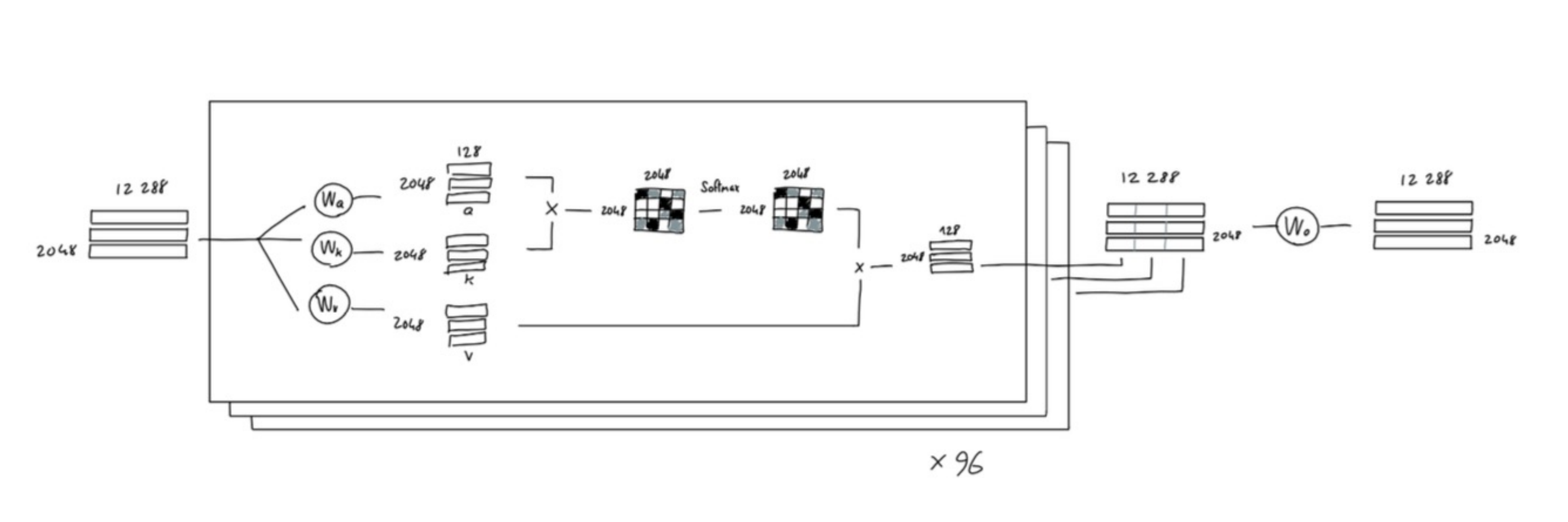
1. How well does a vanilla Transformer-based model perform on the Physionet 2021 challenge data compared to a feature-based model or a Convolutional Network?

2. Can an ensemble Transformer model and a Convolutional Network extracting wavelet features effectively capture spatio-temporal information and improve accuracy?

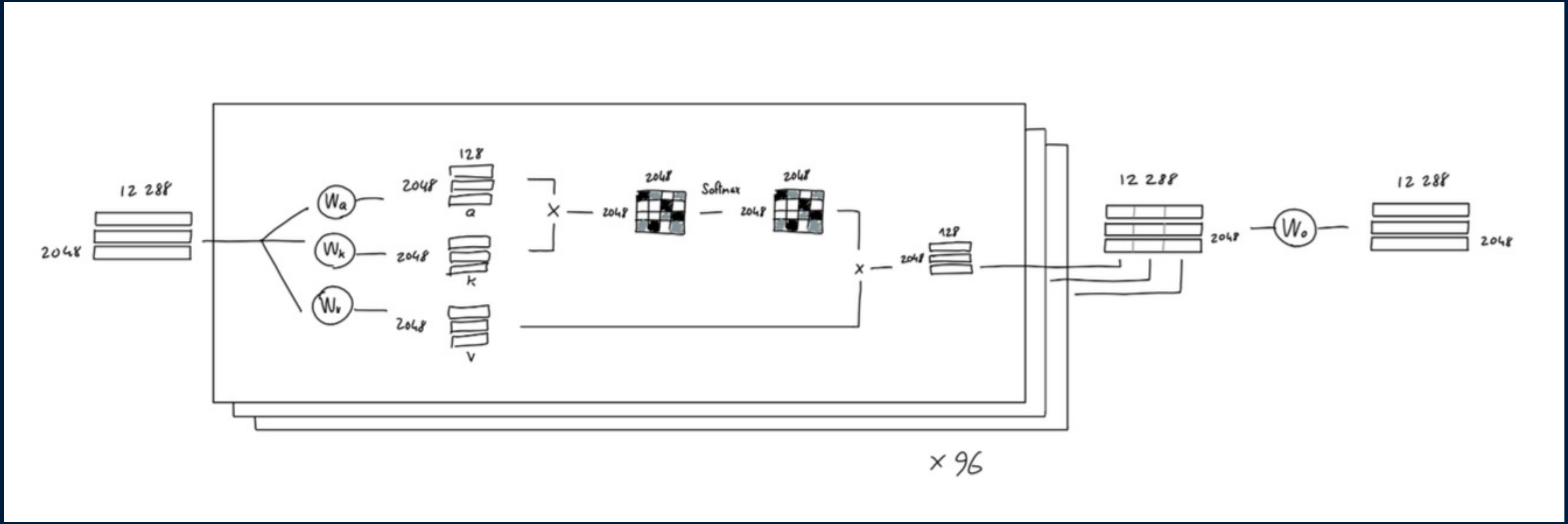
3. Which model performs best at discriminating SR, AF, AFL, PAC and PVC on both datasets?

4. What are the challenges in transferring the pre-trained models from the Physionet 2021 challenge data to the MyDiagnostick database? Do the models generalise well, even though different ECG devices were used?

Methodology 1: Transformer with equal-sized segments as input

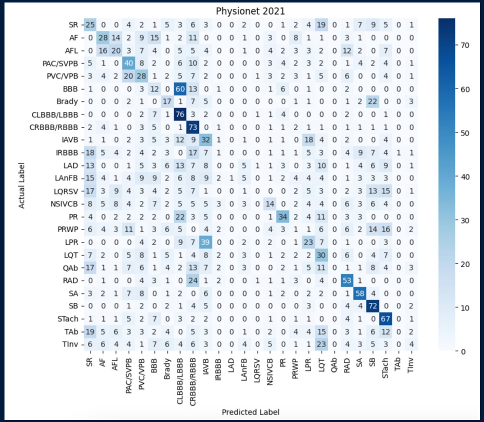


Methodology 1: Transformer with trainable embedding matrix



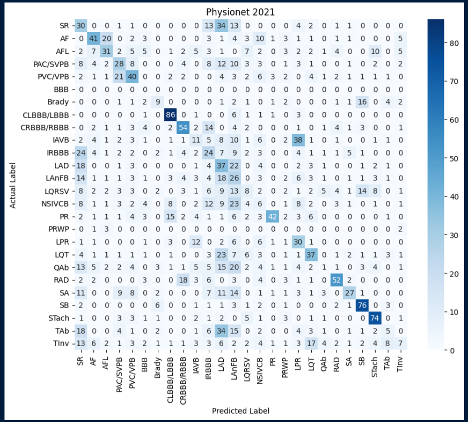
Methodology 2: Ensemble Model

Model 1



+

Model 2



+ …

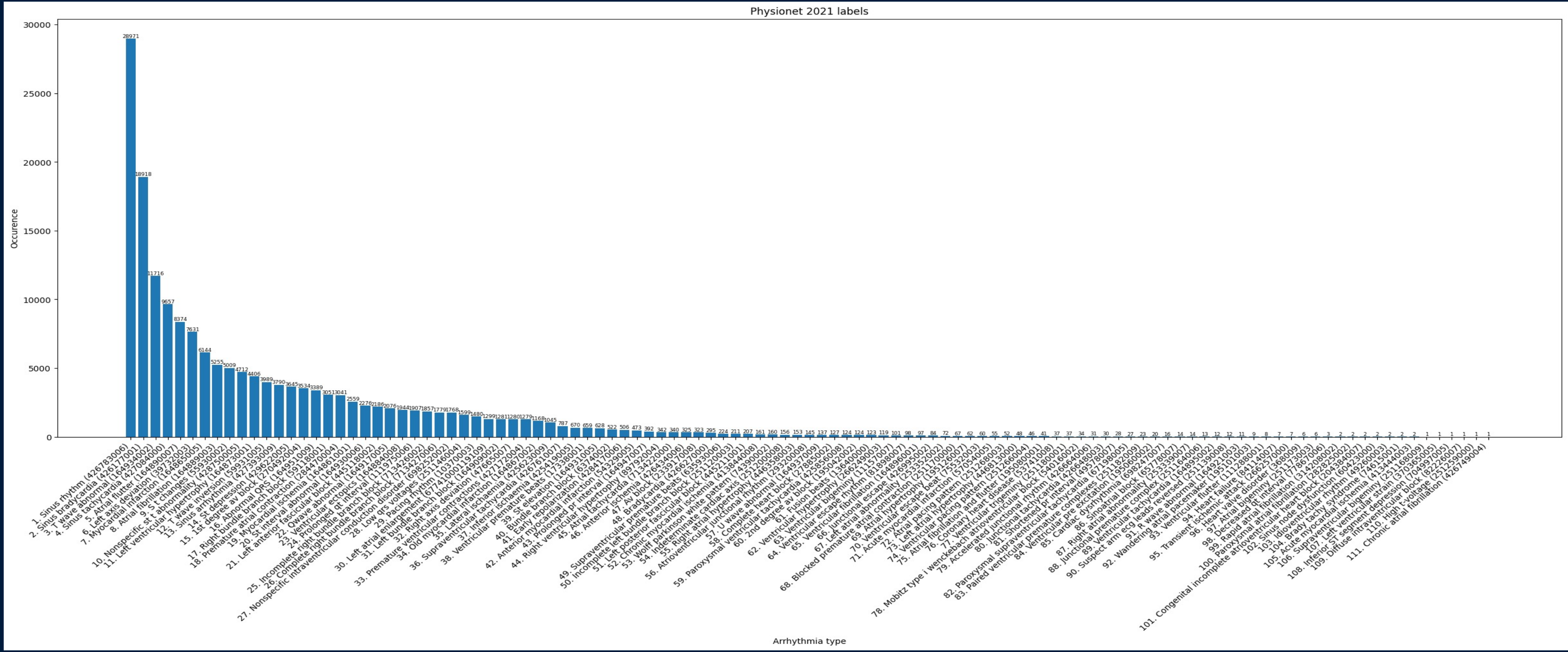
Experiments and Evaluation

Data: Physionet 2021 Challenge database

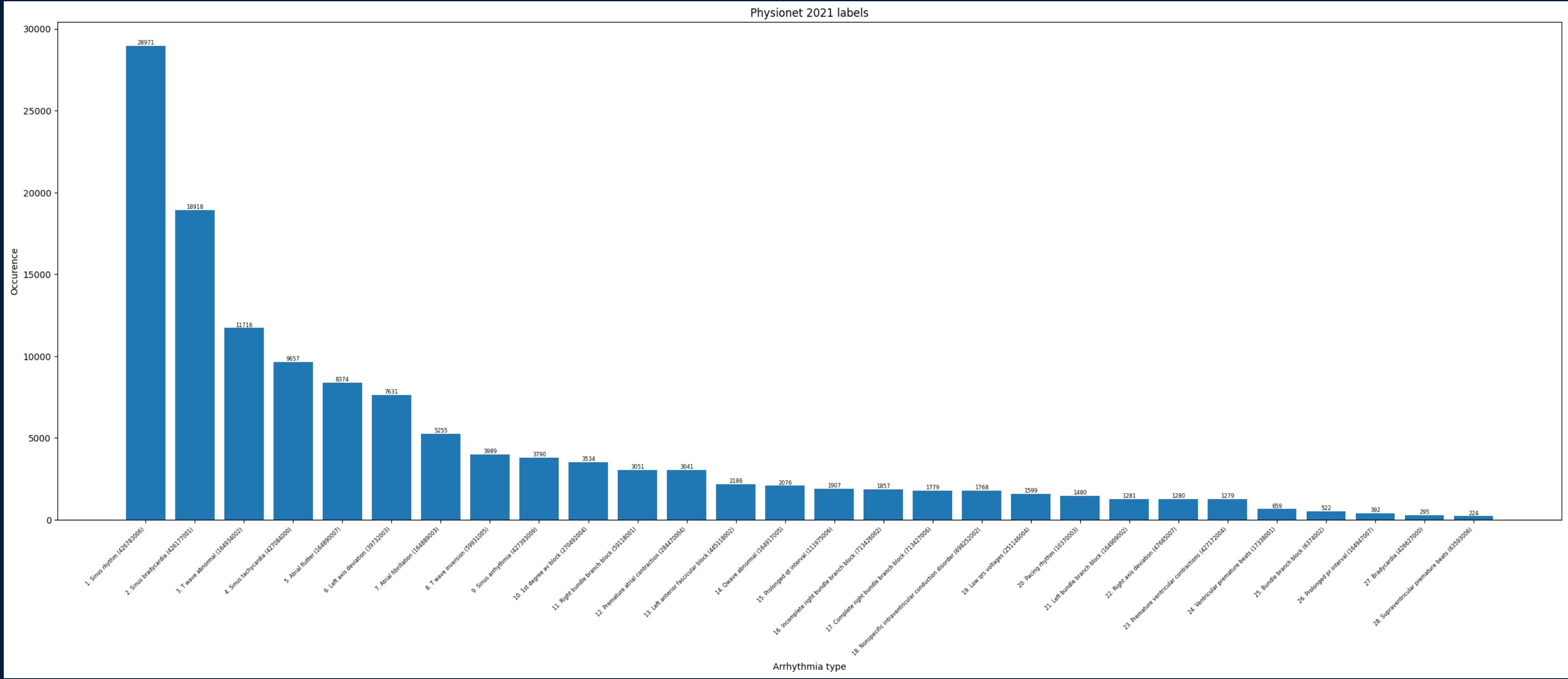
Dataset source	Average ECG length (seconds)	Data samples
Ningbo database	10s	34,905
PTB-XL database	10s	21,837
Chapman-Shaoxing database	10s	10,247
Georgia 12-lead challenge data	9s	10,344
CPSC database	15s	6. 877
CPSC-extra database	15s	3,453
PTB database	110s	516
INCART database	1800s	74

about 89.000 12-lead ECGS

Physionet 2021 data distribution



Physionet 2021 scored challenge data distribution (subset)



Maastricht University

Evaluation: Physionet 2021

Evaluation: Physionet 2021 metrics

Master's Thesis evaluation: Application of Transformer models on the Physionet 2021 challenge database

KK Kramer, Kilian (Stud. DACS / Alumni DACS)
An: contact@physionet.org
Mo, 24.06.2024 11:42

Diese Nachricht befindet sich in Englisch | Übersetzen nach Deutsch | Nie aus Englisch übersetzen

Dear organisers of the CinC Physionet 2021 Challenge,

first of all, a big thank you for the provided Physionet 2021 Challenge database, which is an important part of my Master's thesis. Regarding my current Master's thesis on the application of Transformer and other models (i.e. ensemble of CNN, Transformer and feature-based model) to multi-label arrhythmia classification, I need your help to evaluate my models. I have used the Physionet 2021 Challenge database to train my models and would like to compare my results with the official Challenge metrics and participant team scores (https://moody-challenge.physionet.org/2021/results/), but the official test data is private. Could you please run my best validated model on the test data so that I can successfully draw a proven conclusion from my evaluation?

Yours sincerely,

Kilian Kramer from Maastricht University, Netherlands

Evaluation: Physionet 2021 metrics

Conclusion