

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
- 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 subtypes, such as Atrial Flutter, Premature Ventricular Contractions, Prolonged QT interval etc.

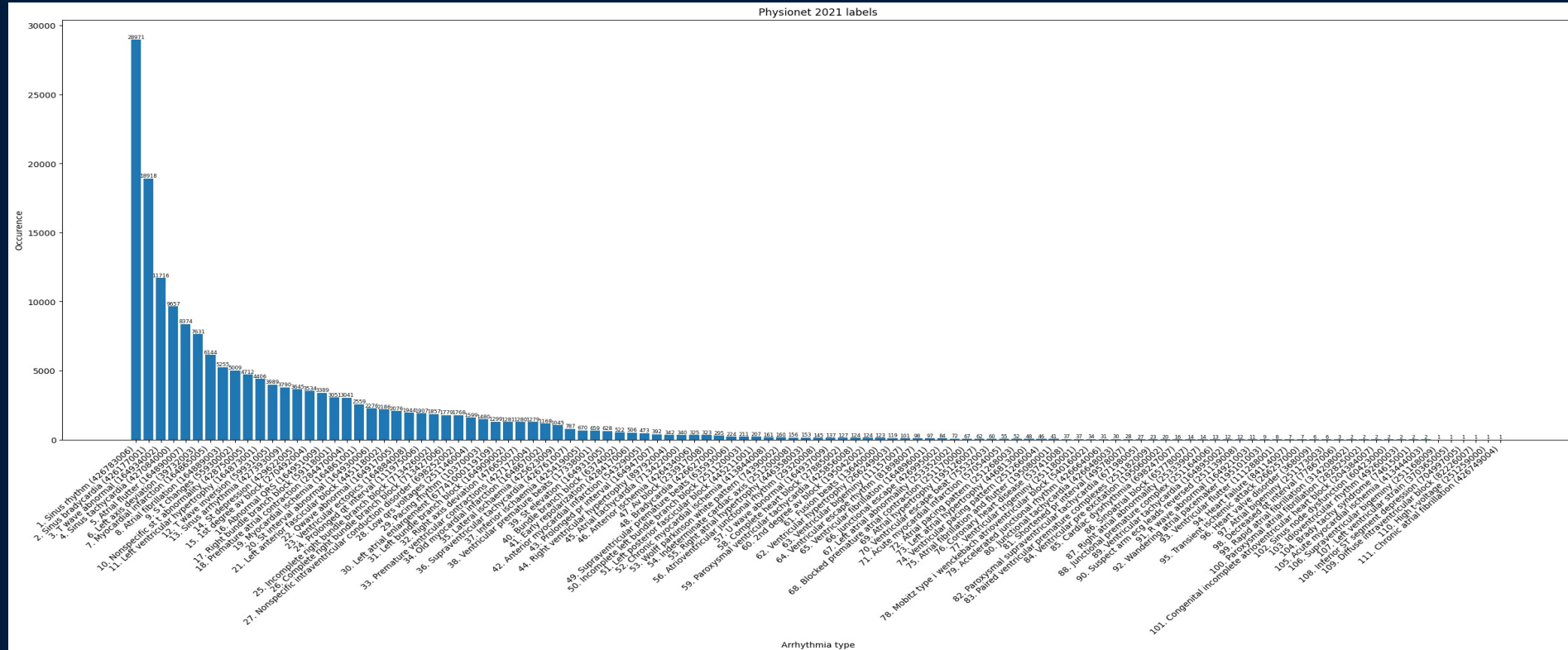
Related Work

Data: Physionet 2021 Challenge database

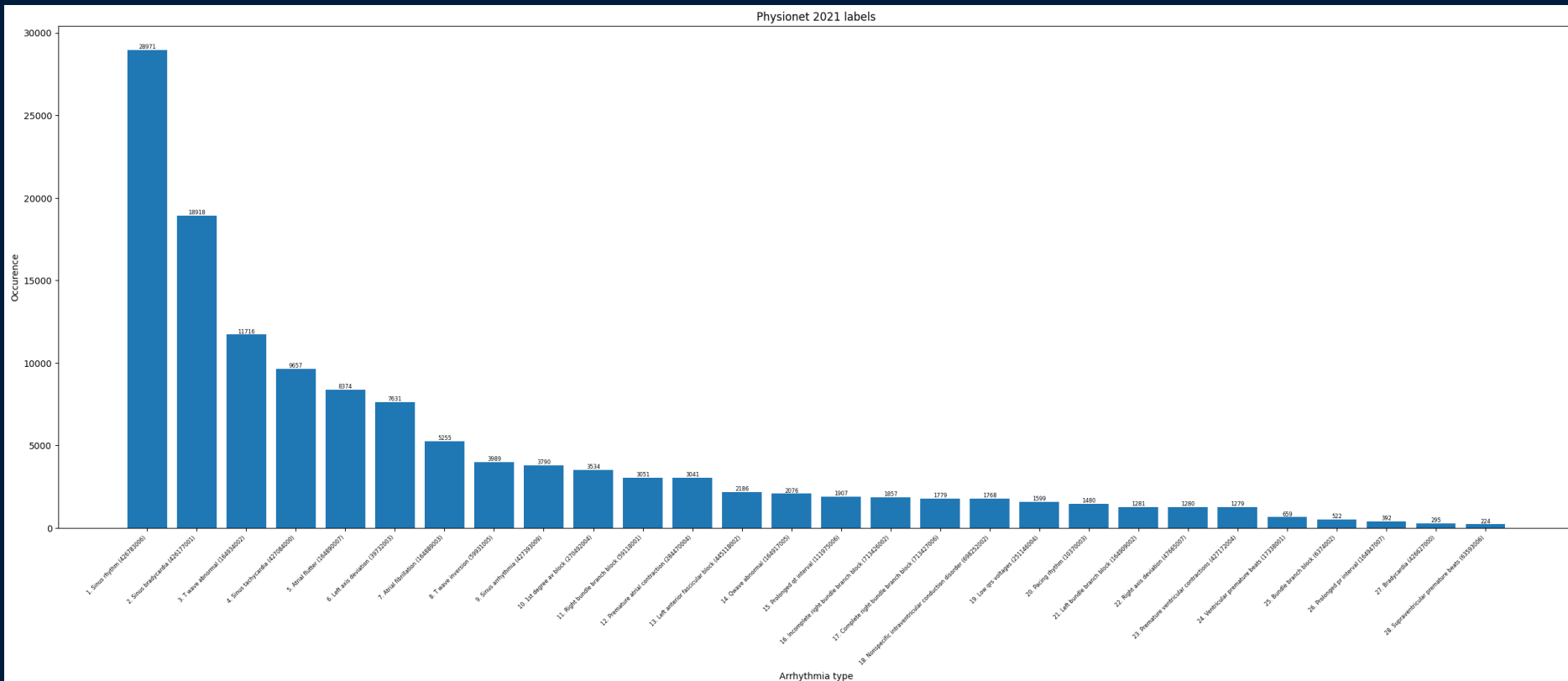
Dataset source	Average ECG length (seconds)	Data samples
Ningbo database	10s	34,905
PTB-XL database	10s	21,837
Chapman-Shaoxing database	10s	10,247
Georgia 12-lead challenge data	9s	10,344
CPSC database	15s	6. 877
CPSC-extra database	15s	3,453
PTB database	110s	516
INCART database	1800s	74

about 89.000 12-lead ECGS

Physionet 2021 data distribution



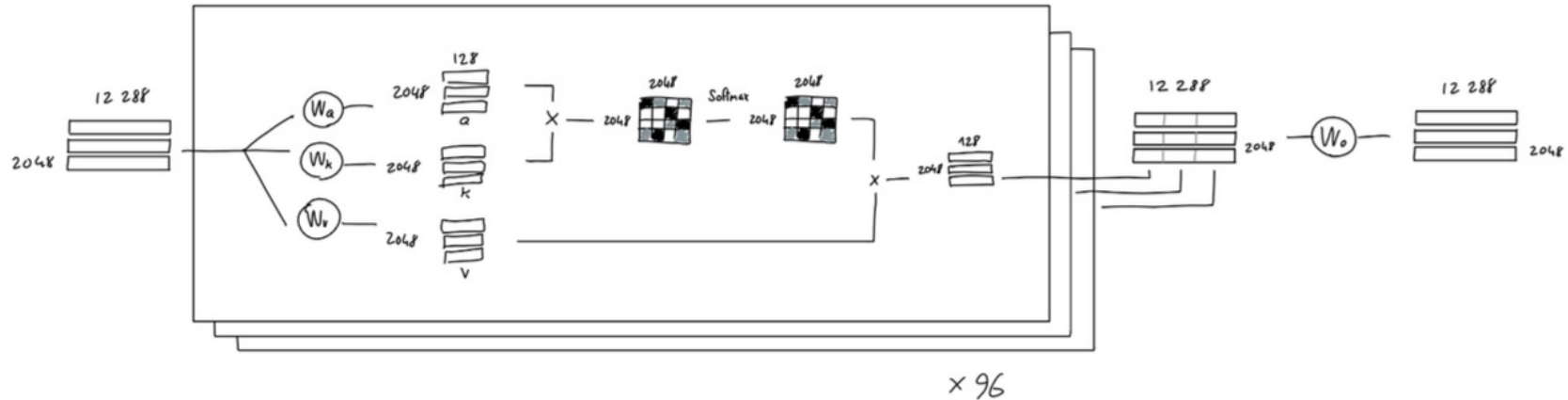
Physionet 2021 scored challenge data distribution (subset)



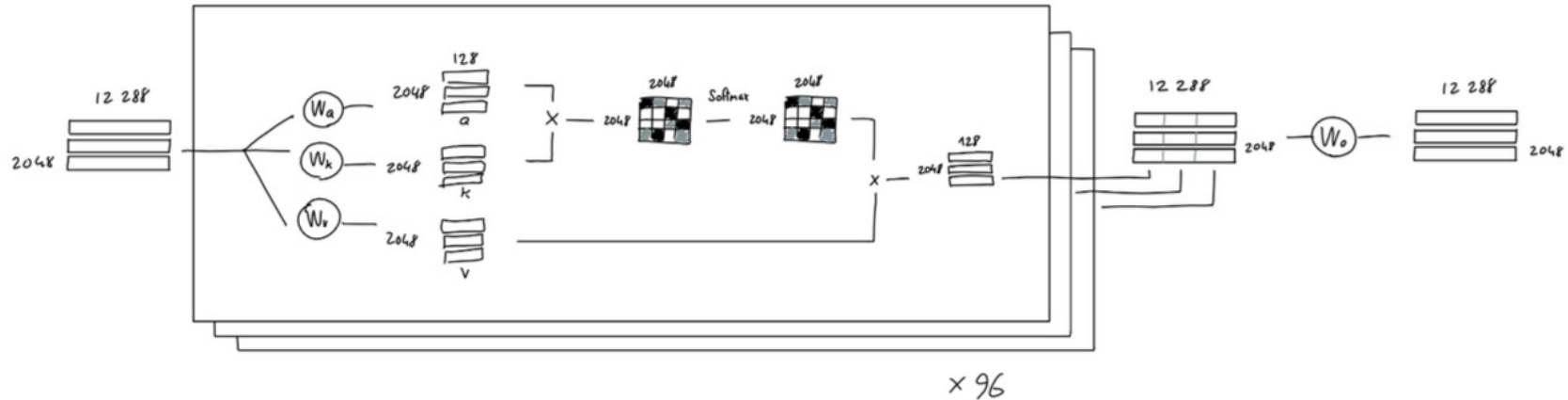
Research Questions

1. How well does a 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 Convolutional Network 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

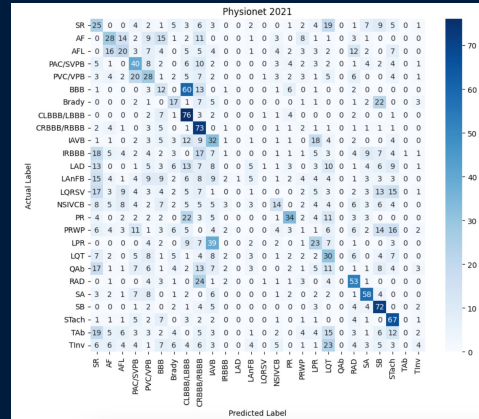


Methodology 1: Transformer with trainable embedding matrix



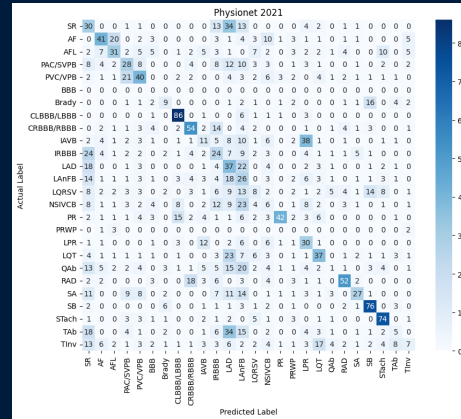
Methodology 2: Ensemble Model

Model 1



+

Model 2



+ ...

Experiments and Evaluation

Evaluation: Physionet 2021

Evaluation: Physionet 2021 metrics



Evaluation: Physionet 2021 metrics

Conclusion