ResNet-based ECG Diagnosis of Myocardial Infarction in the Emergency Department

UPPSALA UNIVERSITET

Daniel Gedon[1,*], Stefan Gustafsson[1,2,*], Erik Lampa[1], Antônio H. Ribeiro[1], Martin J. Holzmann[3], Thomas B. Schön[1], Johan Sundström[1,4]

[1]Uppsala University, Sweden;

[2]Sence Research AB, Sweden;

[3]Karolinska University Hospital and Karolinska Institutet, Sweden;

[4]University of New South Wales, Australia;

[*]equal contribution.

NeurIPS 2021 workshop

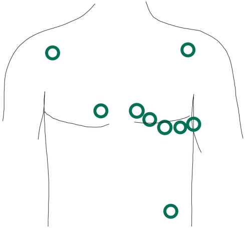
Machine learning from ground truth: New medical imaging datasets for unsolved medical problems.

Online, December 14, 2021

Background

- Emergency Department (ED)
- Myocardial Infarctions (MIs):
  - 9M deaths/year, 200M disability-adjusted life years/year, and rising.
  - False negatives: 10-50,000 missed cases/year at EDs in the United States.
  - False positives: Less than half of those hospitalized for a suspected MI are diagnosed.

  $\rightarrow$ High burden on public health.
- Electrocardiogram (ECG):
  - ST-elevation MI (STEMI) $\rightarrow$ detect in ECG
  - non-ST-elevation-MI (NSTEMI) $\rightarrow$ require blood testing



Background

Baselines:

- Human baseline (cardiologists): 75% acc. for STEMI[1]; much lower for NSTEMI.
- Deep learning models reach super-human performance but:
  - only classify STEMIs[2]
  - use managed data sets[2,3]

Goal: Provide well-calibrated prob. for STEMI/NSTEMI from ECGs at the ED.

Our contribution:

1. Extract a novel data set resembling the real-world setup.
2. Deep learning based model for diagnosing MIs in the ED.

[1] McCabe et al., "Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms".

[2] Cho et al., "Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography".

[3] Liu et al., "A Deep-Learning Algorithm for Detecting Acute Myocardial Infarction".

Data Set

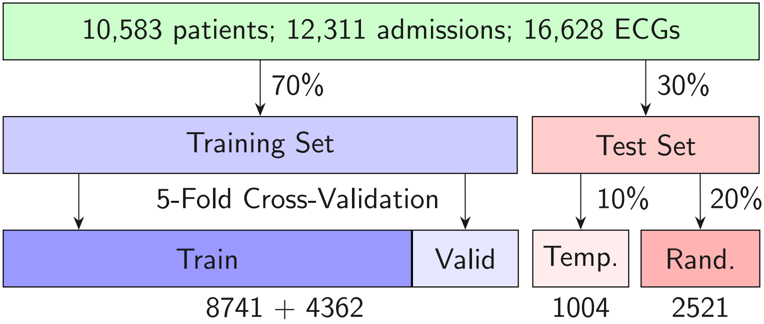
- Standard 10 seconds 12-lead ECGs.
- Adult patients at local ED visits in Stockholm region between 2007 and 2016.
- High risk patients $\rightarrow$ admitted to coronary care unit (CCU).
- Labels:
  - From SWEDEHEART registry[4]
  - By discharging physician that followed entire patient journey during hospitalisation.
- Filter to ensure:
  - inclusion of at event before-treatment ECGs
  - availability of outcome label

$\Rightarrow$ real-world scenario for unsolved problem

[4] `https://www.ucr.uu.se/swedeheart/dokument-sh/variabellista`

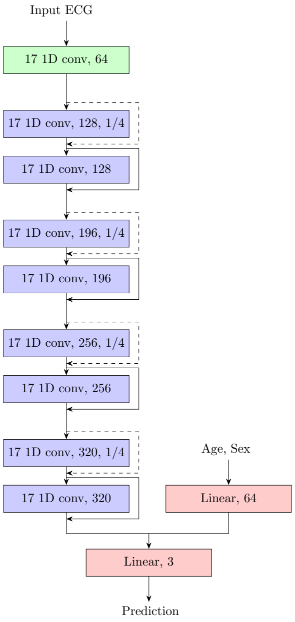
Data Set

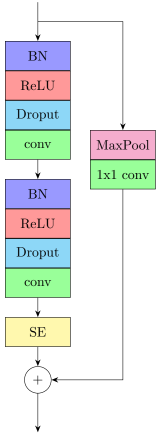
Splitting of the data set:



- Use repeated recordings during training as a form of data augmentation.
- Records from the same patient in the same split.
- Preprocess: Remove low frequency baseline, re-sample, zero-pad.

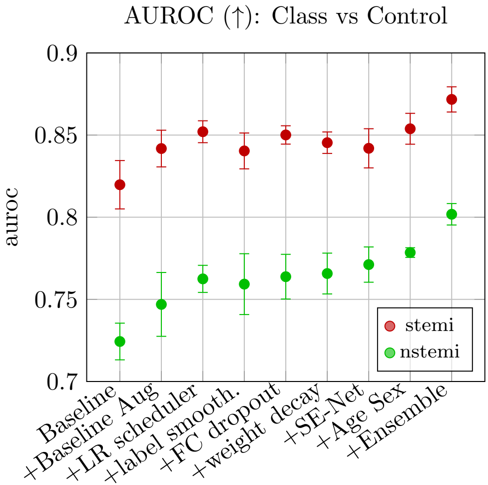
Model Architecture





Model Training

- Baseline from Ribeiro et al.[5]
- Results for 5-fold cross-validation:



[5]Ribeiro et al., "Automatic Diagnosis of the 12-Lead ECG Using a Deep Neural Network".

Results - Main

- Novel data set for unsolved problem $\rightarrow$ no direct baseline available.
- Results over 10 model seeds:

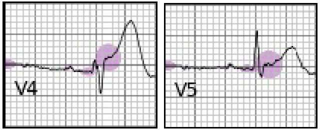
		Random	Temporal
AUROC (↑)	NSTEMI	0.76 (0.003)	0.74 (0.003)
	STEMI	0.85 (0.002)	0.82 (0.003)
AUPR (↑)	NSTEMI	0.69 (0.003)	0.64 (0.005)
	STEMI	0.76 (0.005)	0.64 (0.006)
Brier (↓)	NSTEMI	0.19 (0.001)	0.27 (0.002)
	STEMI	0.10 (0.001)	0.13 (0.001)
ECE (↓)	Multiclass	0.25 (0.004)	0.11 (0.012)

$\Rightarrow$ Super-human-level performance

Results - Model Analysis

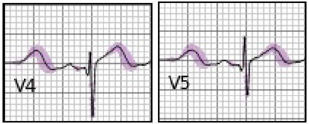
Grad-CAM plots → identify patterns of the model

STEMI



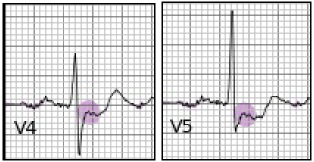
- ST-segment elevation
- typical for humans

STEMI



- Down-sloping T-wave
- untypical for humans

NSTEMI



- ST-segment depression
- humans would not suspect a MI

Extension - Larger Patient Population

- More general patient population: All patients at ED
- Data set statistics:

		Current CCU patients	*Extension* all ED patients
ECGS	Train	8,741+4,362	307,549+67,173
	Test Temp.	1,004	27,937
	Test Rand.	2,521	89,782
	Total	16,628	492,441
Patients		10,583	214,431
	Control	55.9%	98.5%
	STEMI	11.4%	0.4%
	NSTEMI	33.1%	1.1%

- Initial results: similar performance as in current study

Contact

UPPSALA UNIVERSITET

Daniel Gedon, Uppsala University

E-mail: `daniel.gedon@it.uu.se`
GitHub: `https://github.com/dgedon`
Twitter: `@danigedon`

Supported by the Kjell and Märta Beijer Foundation, Anders Wiklöf, the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by Knut and Alice Wallenberg Foundation, and Uppsala University via AI4Research.

APPENDIX

Appendix: Data Set

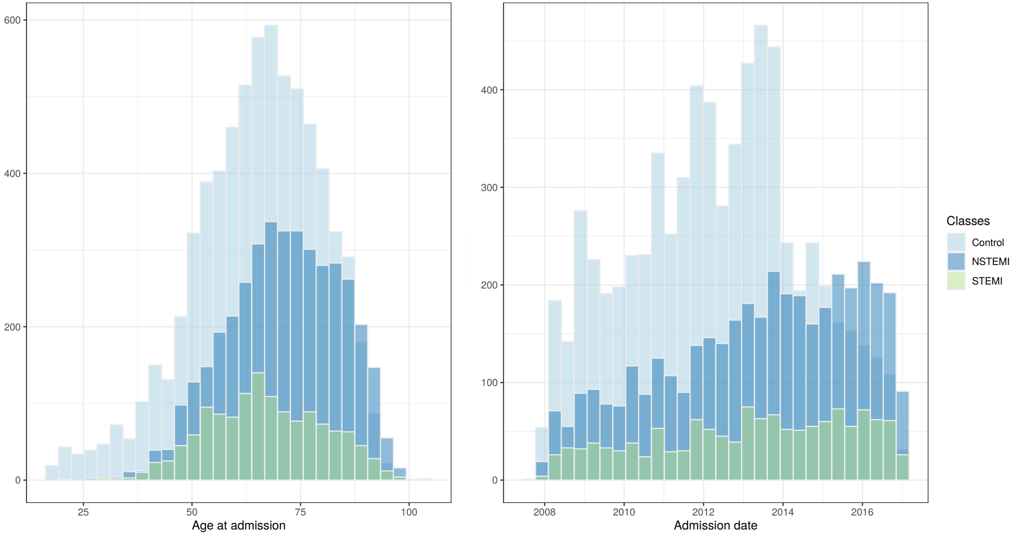
Similar data set: CODE[6,7]

- 811 counties in the state of Minas Gerais, Brazil.
- Collected between 2010 and 2016.
- 2.3M ECGs with 6 heart related labels; more available upon request.
- Separate high quality test data set[8].
- Example usage for anomaly classification[9], ecg-age[10], explaining ECG diagnosis[11] and many more.
- Available for research upon request.

---

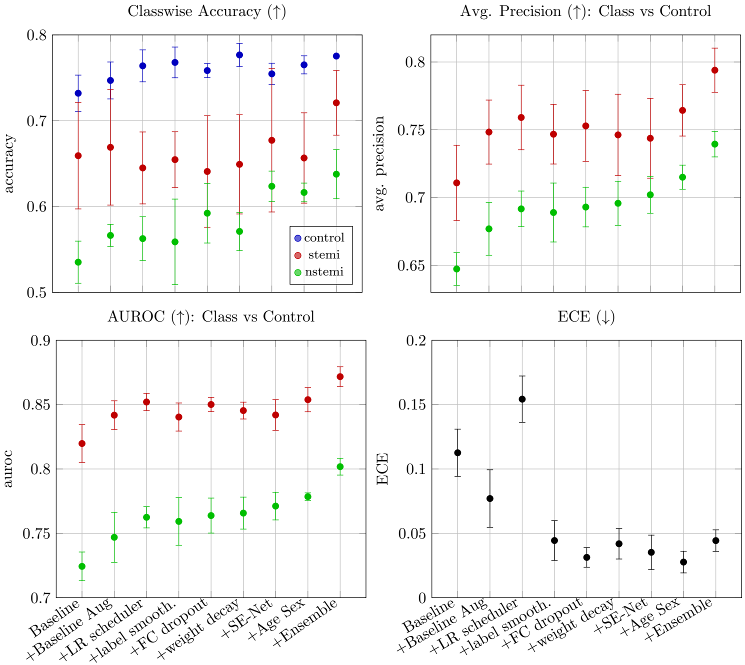
[6] Alkmim et al., "Improving patient access to specialized health care: the Telehealth Network of Minas Gerais, Brazil".
[7] Ribeiro et al., "Tele-electrocardiography and bigdata: The CODE (Clinical Outcomes in Digital Electrocardiography) study".
[8] doi.org/10.5281/zenodo.3765780
[9] Ribeiro et al., "Automatic Diagnosis of the 12-Lead ECG Using a Deep Neural Network".
[10] Lima et al., "Deep neural network-estimated electrocardiographic age as a mortality predictor".
[11] Oliveira et al., "Explaining End-to-End ECG Automated Diagnosis Using Contextual Features".

Appendix: CCU Admission Statistics

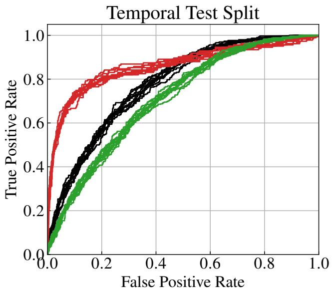


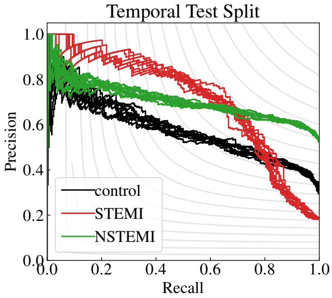
Appendix: Hyperparameter Selection

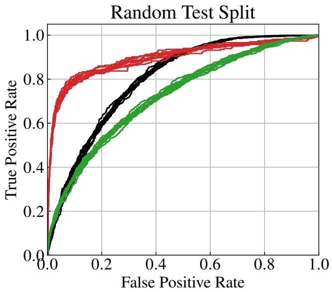
UPPSALA UNIVERSITET

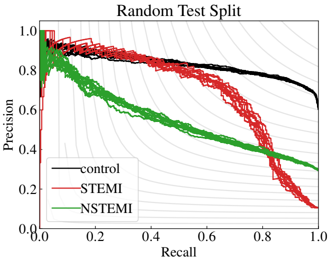


Appendix: Results ROC and PR curves

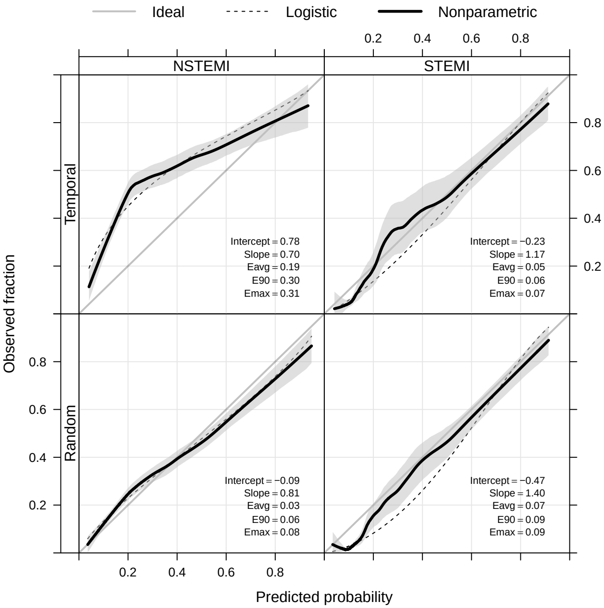








Results - Calibration Plot



Appendix: Discussion

Misclassifications:

- Follows known clinical / machine learning patterns.
- Myocarditis as imposter.

Limitations:

- Lack of external validation.
  $\rightarrow$ Temporal test set simulates new data distributions.
- Selected ECG trace:
  - Label was decided upon discharge from CCU.
  - We cannot ensure that the ECG we use is the one which guided the final diagnosis.

  $\rightarrow$ We mitigate partially by using repeated recordings during training if available.

Appendix: Extensions - (2) Blocked Artery

- Current medical classification: no MI, STEMI, NSTEMI
- Proposal of new classification: identify exact artery which is blocked

  → more fine grain classification

  → direct use for practicing physicians

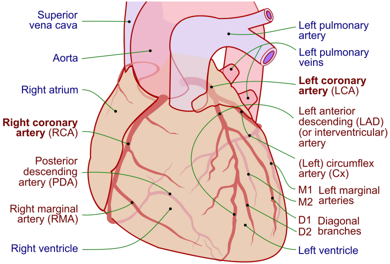


Figure: Coronary Arteries[12]

[12] en.wikipedia.org/wiki/Coronary_circulation

Appendix: Comparison with other Models

UPPSALA UNIVERSITET

Main comparison with Liu et al., “A Deep-Learning Algorithm for Detecting Acute Myocardial Infarction”.

Similarities:

- Residual blocks but not standardized ResNet structure.

Differences:

- Lead-wise residual network instead of combining all leads.
- No SE-net block but attention mechanism for each lead.

$\rightarrow$ Ability to use Grad-CAM plots with different highlight for different leads.
$\rightarrow$ Assume that leads are independent but they are highly correlated.