AI4Imaging-Hackathon-2024

CLASSIFICATION OF HEART DISEASE BASED ON CINE MRI SCAN

- Marvin Leonard Simak -

Course project
Intermediate Machine Learning
Opencampus.sh

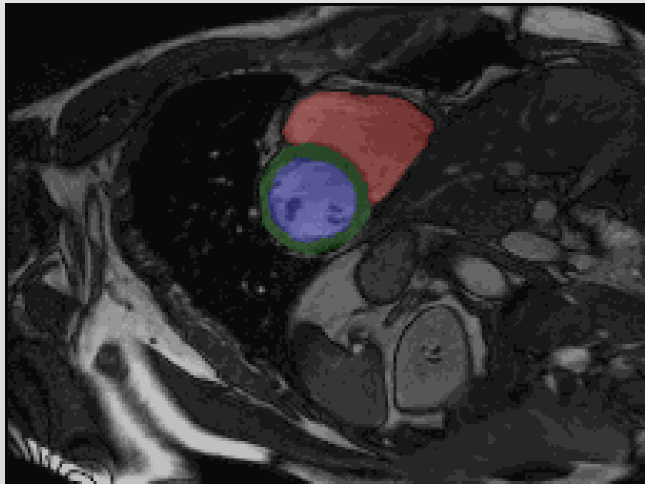
Introduction

- **Task**
  Multi-classification deep learning or radiomics model to predict the category of heart disease using two frames of cardiac MRI

- **Dataset**
  - Shuffled version of the Automated Cardiac Diagnosis Challenge
  - 150 MRI exams from patients (100 train + 50 test)
  - MRI data (3D-voxel intensity) for two cardiac phases: diastolic and systolic
  - Segmentation masks for region of interest (ROI)
  - 5 different classes

- **Evaluation**
  - Mean F1-Score

$$F1 = \frac{2 \ x \ Precision \ x \ Recall}{Precision + Recall}$$



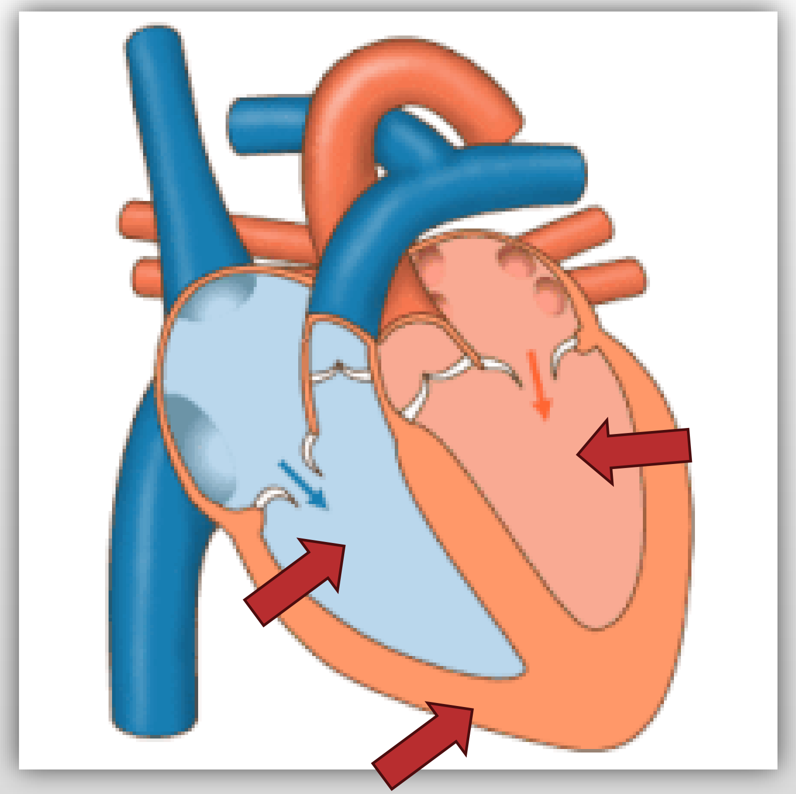
Heart structure & physiology

- The heart functions as a pump, delivering oxygenated blood to the body and deoxygenated blood to the lungs
- **Myocardium:** Cardiac / heart muscle

  Thick middle layer between the the pericardium and the endocardium
- **2 hearts / 4 chambers**:
  Right heart: Body
  right atrium, right ventricle, left atrium & left ventricle
- **Cardiac cycle**: Alternates between

  Systole (contraction) and

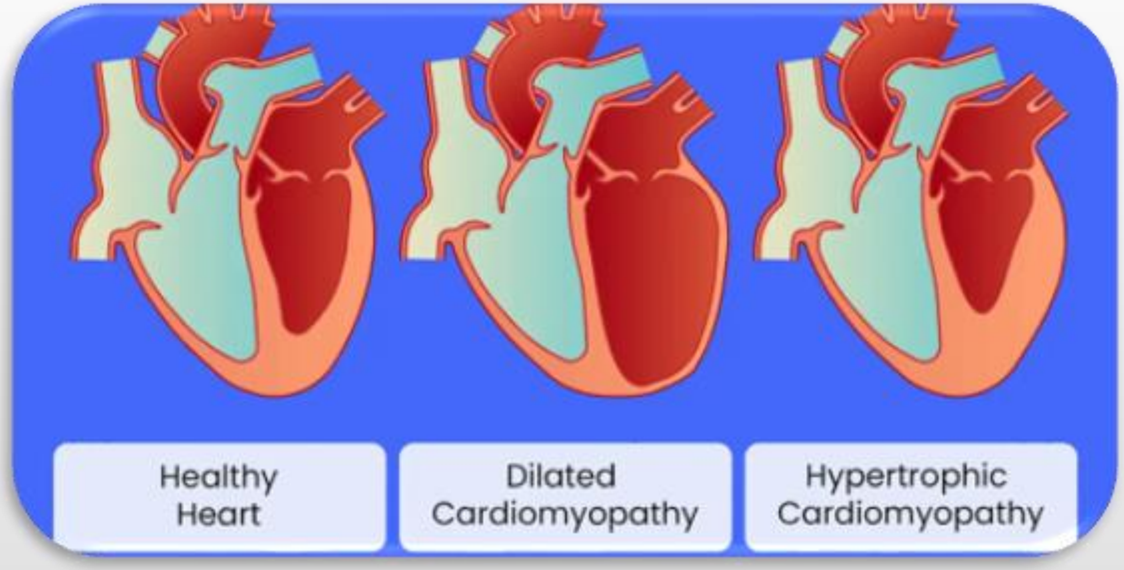
  Diastole (relaxation)



Cardiac cycle:
Body → Right Atrium → Right Ventricle (blue) → Lung → Left Atrium → Left Ventricle (red) → Body

Structural heart diseases

- **Myocardial infarction (MINF):**
  - Blood flow decreases or stops in one of the coronary arteries of the heart, causing infarction (tissue death) to the heart muscle
- **Dilated cardiomyopathy (DCM):**
  - Heart muscle disease that causes the heart chambers (ventricles) to thin and stretch, growing larger. It typically starts in the heart's main pumping chamber (left ventricle)



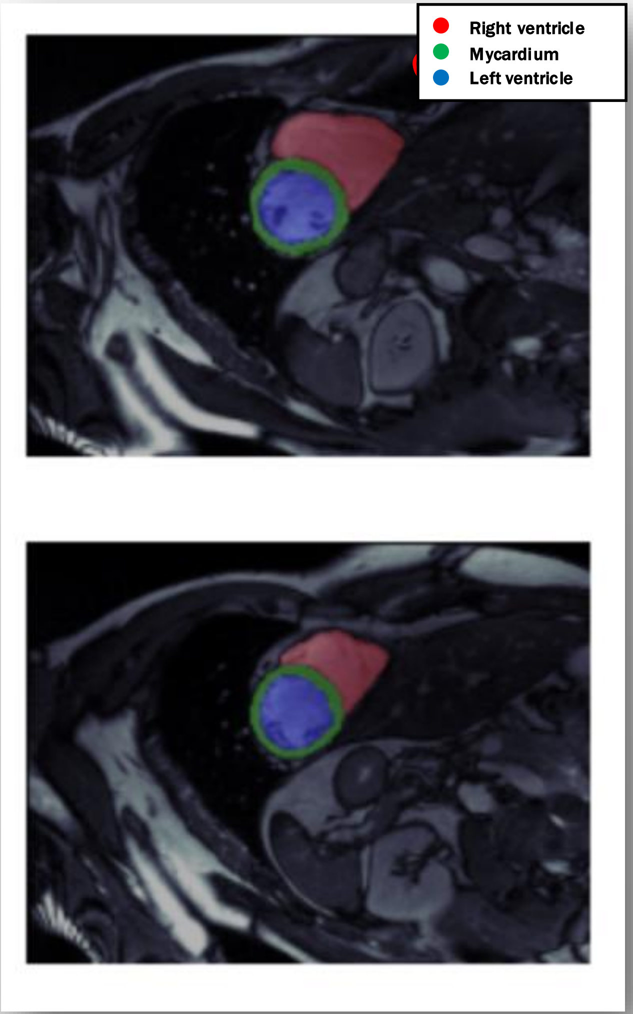
- **Hypertrophic cardiomyopathy (HCM):**
  - A condition affecting the left ventricle, the main pumping chamber of the heart. The walls of the left ventricle become thick and stiff
- **Abnormal right ventricle (RV):**
  - Abnormal enlargement of the cardiac muscle surrounding the right ventricle

Dataset

- 3D cardiac MRI images and corresponding segmentation masks for diastolic and systolic phase each
- 100 patients for training (and 50 patients for testing)
- Multi-label segmentation mask:

1 = right ventricle   3 = left ventricle   2 = myocardium

- **5 classes** (with 20 samples each)
  - **NOR**: Normal subjects
  - **MINF**: Myocardial infarction
  - **DCM**: Dilated cardiomyopathy
  - **HCM**: Hypertrophic cardiomyopathy
  - **RV**: Abnormal right ventricle



Baseline: Calculate volumes & use a simple model

	ESV_R	EDV_R	ESV_L	EDV_L	MV_{Sys}	MV_{Dia}
NOR	147.7	69.2	94.1	109.0	130.8	51.0
MINF	126.2	58.3	129.2	143.4	189.3	131.8
DCM	**186.1**	**128.6**	**158.1**	168.6	**275.6**	**224.9**
HCM	121.2	47.7	**168.0**	**192.4**	128.0	42.0
RV	**220.6**	**139.6**	81.63	91.02	125.5	54.0

- Voxel segmentation masks → Volume!
- Easily done with nibabel
- Calculate volumes of both ventricles and the myocardium for both phases

Build a simple ML model:

- Sklearn pipeline
- StratifiedKFold & GridSearchCV
- SVM, RandomForest, GradientBoosting & MLPClassifier

→ SVM (C= 10, kernel = rbf, scaling='passthrough')

```
Best classifier: SVM
              precision    recall  f1-score   support

         DCM       0.60      0.75      0.67         4
         HCM       1.00      0.75      0.86         4
        MINF       0.67      0.50      0.57         4
         NOR       0.67      1.00      0.80         4
          RV       1.00      0.75      0.86         4

    accuracy                           0.75        20
   macro avg       0.79      0.75      0.75        20
weighted avg       0.79      0.75      0.75        20
```

From the ACDC website:

Classification rules

Each group was clearly defined according to physiological parameter, such as the left or right diastolic volume or ejection fraction, the local contraction of the LV, the LV mass and the maximum thickness of the myocardium.

Different possibilities of ambiguous cases are detailed here:

- Patients with hypertrophic cardiomyopathy have a left ventricular ejection fraction higher than 55 %. Otherwise, patient with a left ventricular ejection fraction less than 40 % and a local increase of the myocardial thickness (as an adaptation of the myocardium to the disease) must be classified as patients with previous myocardial infarction.
- Patients with abnormal high left ventricular diastolic volume, low left ventricular ejection fraction and only several myocardial segments with abnormal contraction must be classified as patients with previous myocardial infarction. Indeed, the increase of the volume of the left ventricle is an adaptation of the left ventricle due to a myocardial infarction.
- A patient with dilated left and right ventricles (with or without abnormal function of the right ventricle) must be classified as patients with dilated cardiomyopathy. Indeed, dilated cardiomyopathy of left ventricle could have impact on the right ventricle.
- Patients with borderline values should not be included in one particular class. For example, an ejection fraction of the right ventricle greater than 45 % is considered as normal (Mac Kenna criteria) but an ejection fraction of the right ventricle between 40% and 45 % do not allow to classify a case as patient with abnormal right ventricle.

Baseline: A few more features

	ESV_R	EDV_R	ESV_L	EDV_L	MV_{Sys}	MV_{Dia}	$MT_{Sys,\ mean}$	$MT_{Dia,\ mean}$	$MT_{Sys,\ max}$	$MT_{Dia,\ max}$	SV	EF [%]
DCM	186.1	128.6	275.6	224.9	158.0	168.5	5.7	7.1	11.0	13.2	50.7	0.2
HCM	121.2	47.7	127.9	42.0	167.9	192.4	9.3	15.9	19.8	24.3	86.0	0.7
MINF	126.2	58.3	189.3	131.8	129.2	143.4	6.0	8.3	12.2	16.8	57.5	0.3
NOR	147.7	69.2	130.8	51.0	94.1	109.0	5.1	9.6	10.1	15.3	79.9	0.6
RV	220.6	139.6	125.5	54.0	81.6	91.0	4.5	8.0	9.7	14	71.5	0.6

Calculate more features:

- Volumes of RV & LV
- Myocard thickness (mean / max)
- Stroke Volume SV = LV_{sys}- LV_{DIA}
- Ejection fraction EF = SV / Lv_{sys}
- Slight improvement 0.77

```
Best classifier: SVM
              precision    recall  f1-score   support

         DCM       0.57      1.00      0.73         4
         HCM       1.00      0.75      0.86         4
        MINF       1.00      0.25      0.40         4
         NOR       0.80      1.00      0.89         4
          RV       1.00      1.00      1.00         4

    accuracy                           0.80        20
   macro avg       0.87      0.80      0.77        20
weighted avg       0.87      0.80      0.77        20
```

Published results for automatic cardiac diagnosis

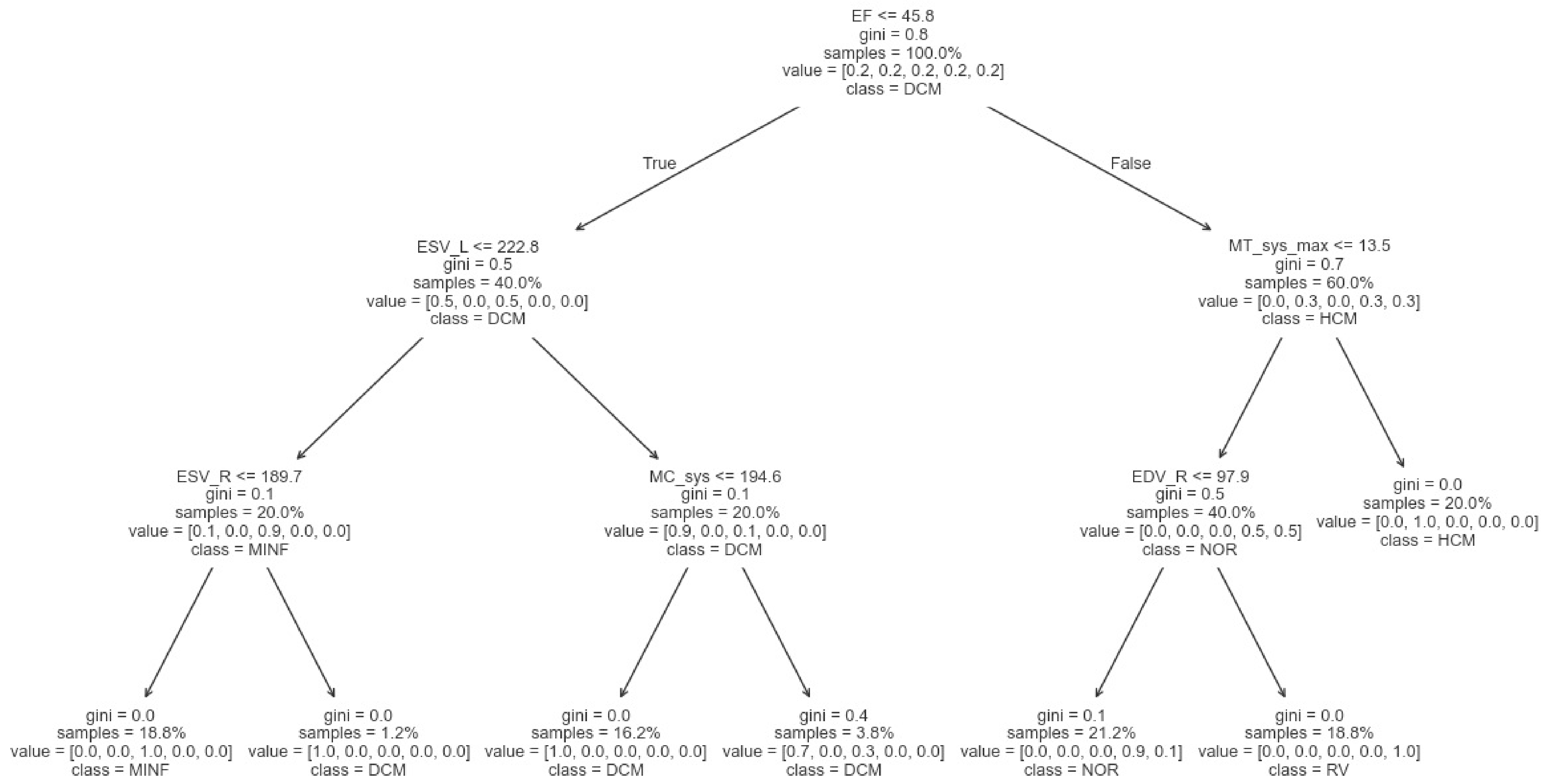
- Khened et al. [46] used 11 features, 9 derived from their segmentation map in addition to the patient weight and height → Acc: 96%
- Isensee et al. [44] extracted a series of instants and dynamic features from the segmentation maps →Acc: 92%

Similar approach but more features based on dynamics and anthropometrics (e.g. age, body mass, sex) – which are not available in the Kaggle dataset

➔ Calculating physiological features is close to medical practice and diagnostics

➔ Explainable / No black-box

Explainability



Baseline: Leaderboard

Prizes & Awards

Kudos

Does not award Points or Medals

Participation

8 Entrants

7 Participants

4 Teams

9 Submissions

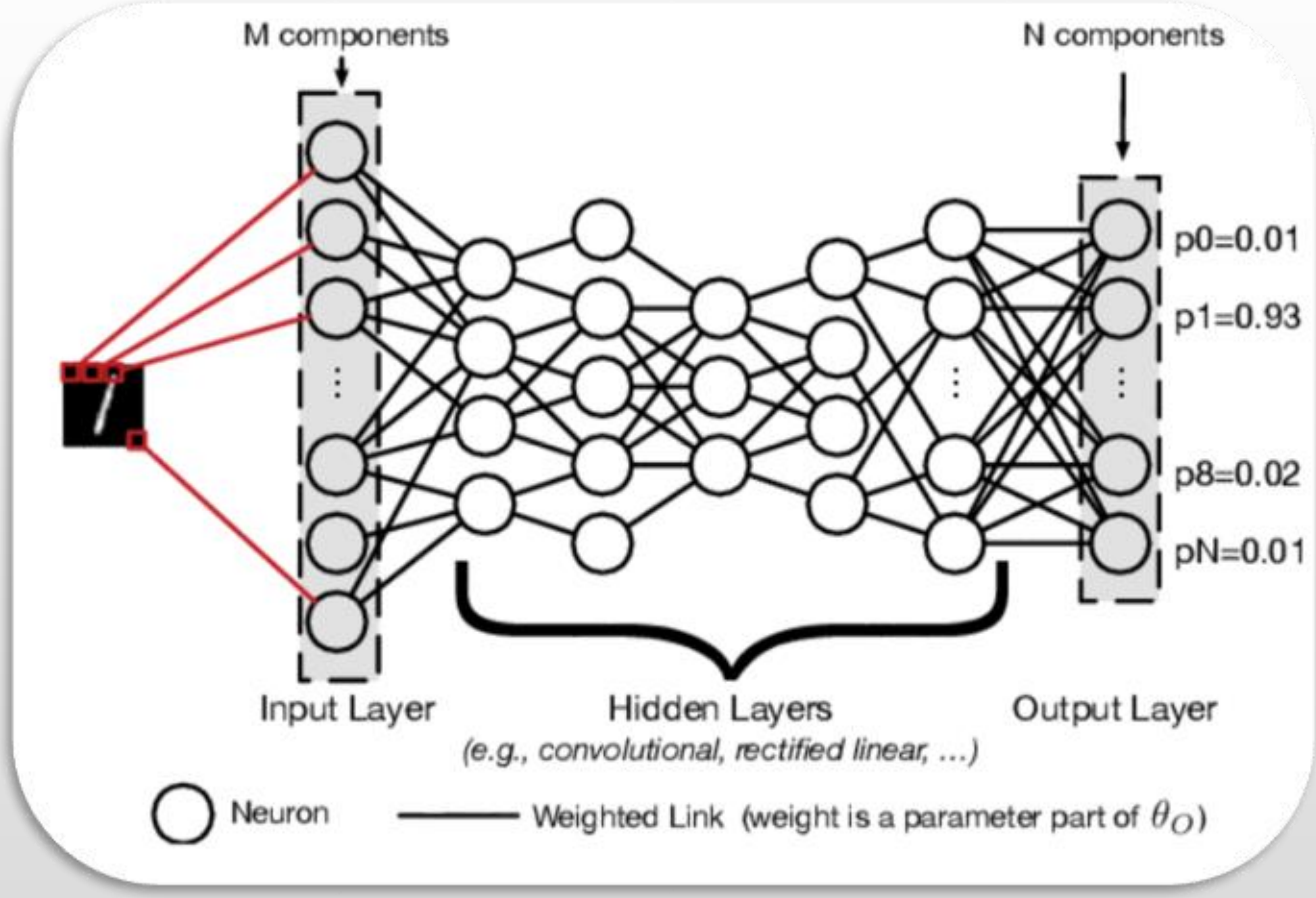
#	Team	Members	Score	Entries	Last	Solution
1	PixelPioneersss		0.64772	1	7mo	
2	PixelPioneers		0.53646	5	7mo	
3	PixelPioneersv1		0.26727	2	7mo	
4	CardioCoderz		0.16000	1	7mo	

- Not the most popular challenge – lets see if our baseline can beat that…

submission.csv **0.97994**

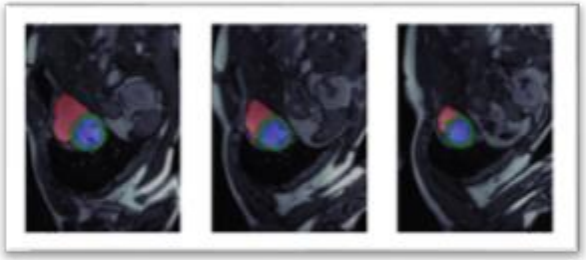
Complete (after deadline) · 2d ago · SVM & feature engineering

- That will be a hard to beat baseline ^^’

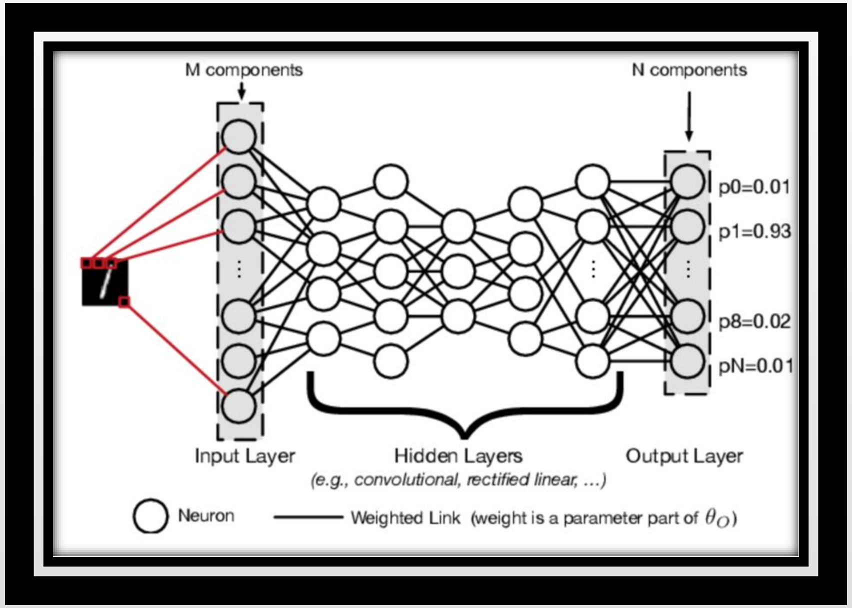


Deep Learning is all your need… Or is it?

Input Tensor
[B x C x D x W x H]

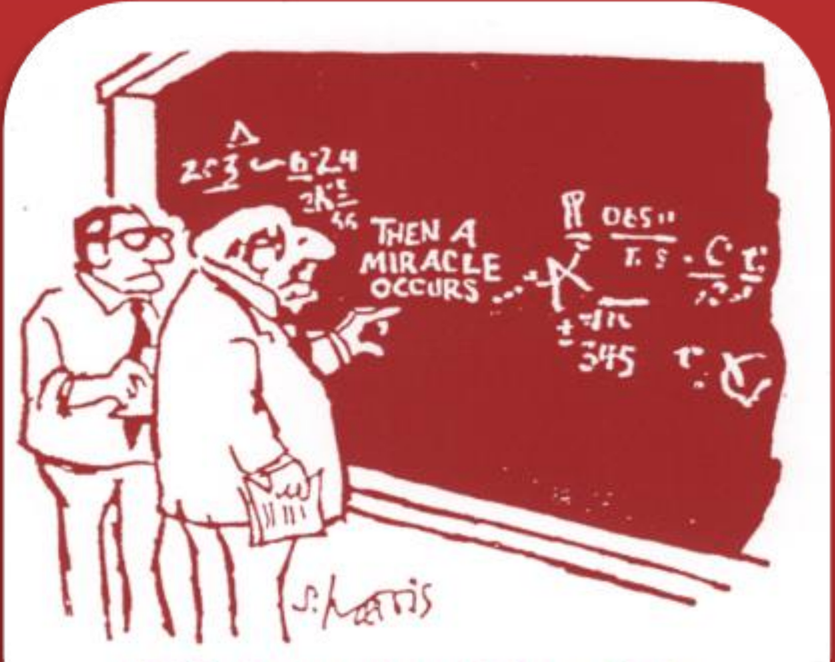


Magical black box
(aka DL model)



Success! eh..
Class prediction

NOR | MINF | DCM | HCM | RV



I think you should be a little more specific, here in Step 2

End-to-End Training possible?

Frameworks & Helpers

pandas

NiBabel

PyTorch

PyTorch Lightning

You do the research.
Lightning will do everything else.

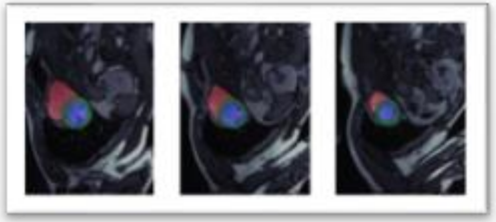
MONAI

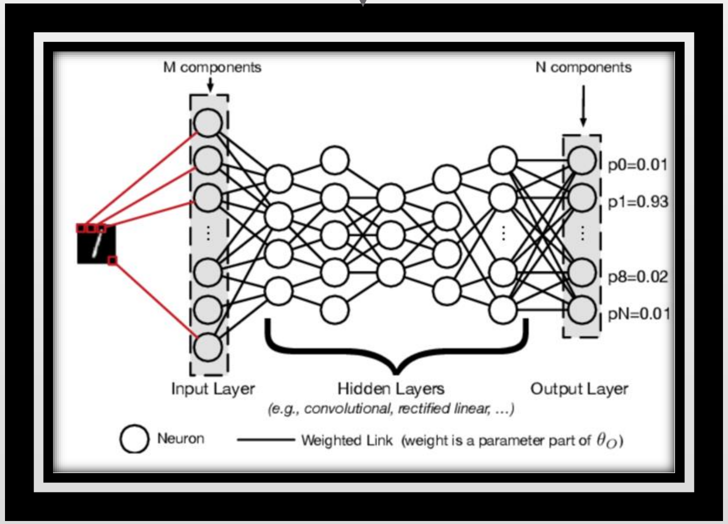
Open-source framework built for accelerating research and clinical collaboration in Medical Imaging

- Open Source Design
- Standardized
- User Friendly
- Reproducible
- Easy Integration
- High Quality

Easily done! Or is it…

- Just load the data, scale it and dump it into the model…
- Training:
  - PyTorch Lightning
  - Stratified 80/20 train-val-split
  - CNN (3D-ResNet18)
  - CrossEntropy-Loss
  - AdamW Optimizer
  - Batch Size 8
- Results:
  - F1-Score: 0.4 😭





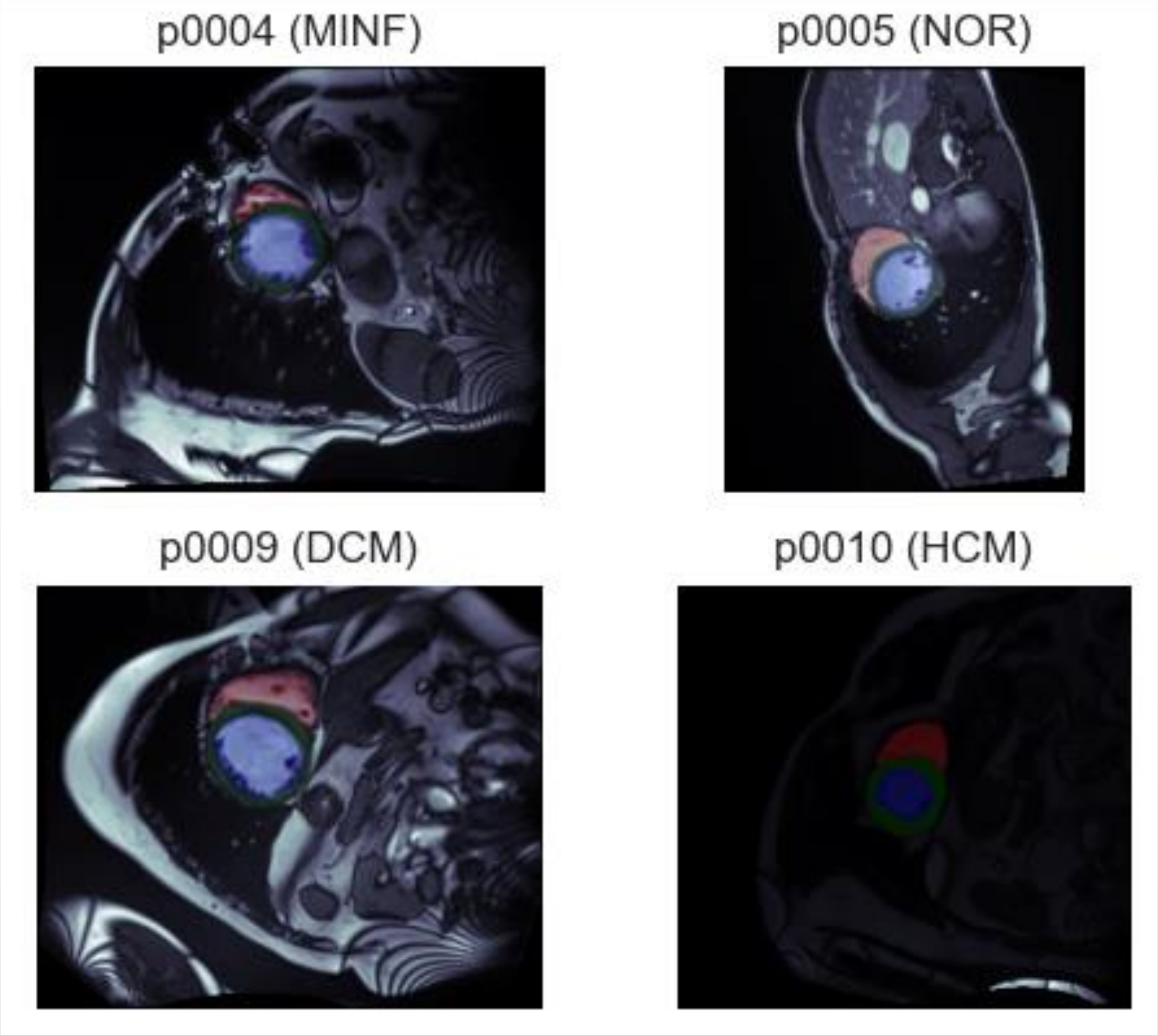
NOR | MINF | DCM | HCM | RV

Let’s have another look at the data

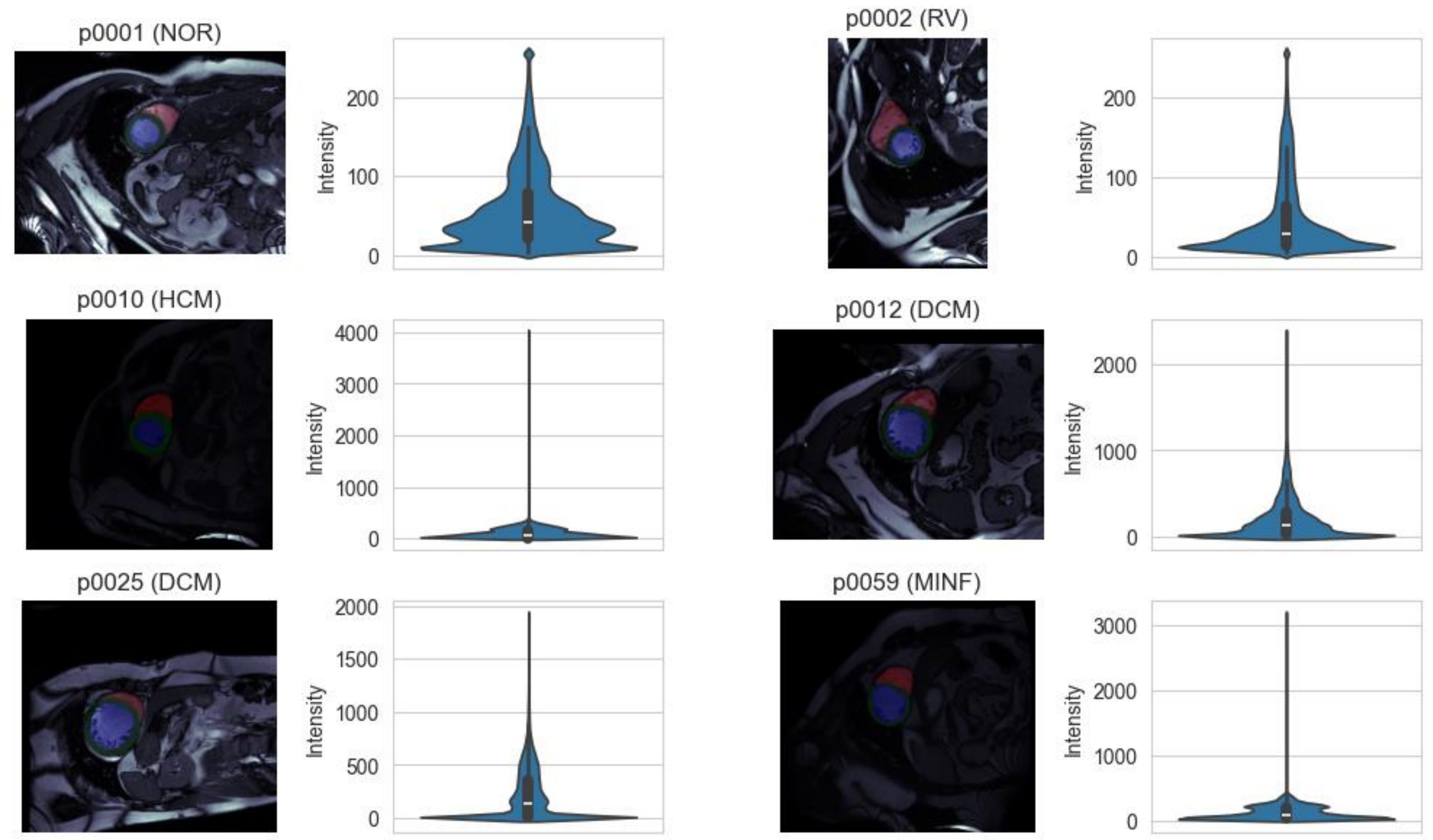
Problems:

- Different intensities
- Different orientation (wrong meta)
- Heart is small part of the picture

➔ Preprocessing is mandatory!



Intensity normalization

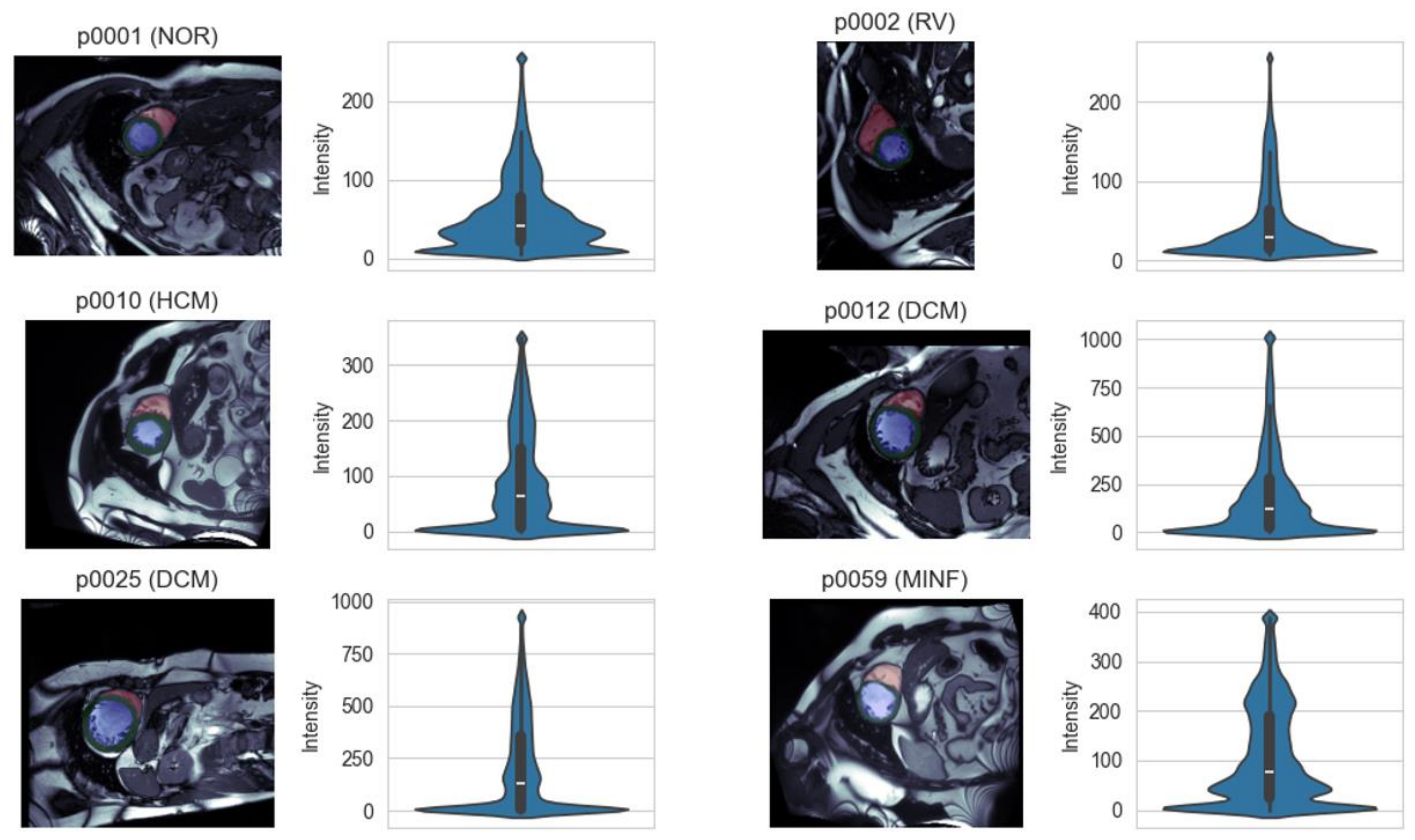


Problems:

- Completely different scales between patients
- Artefacts / Reflections

→ Clip outliers and scale intensity linearly

Intensity normalization



Result:

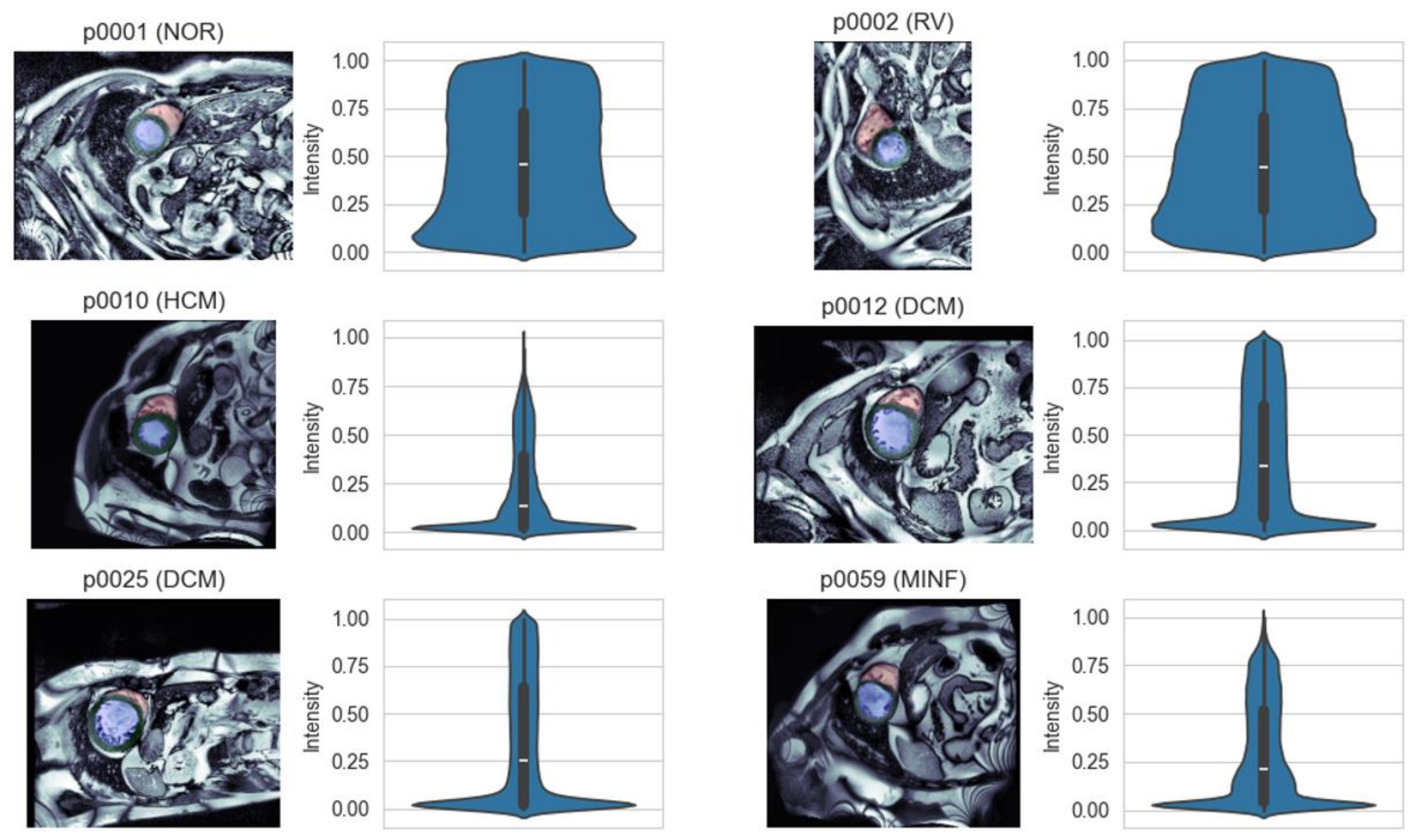
- Much better!

Further problems?

- Different contrast

→ Histogram normalization

Intensity normalization

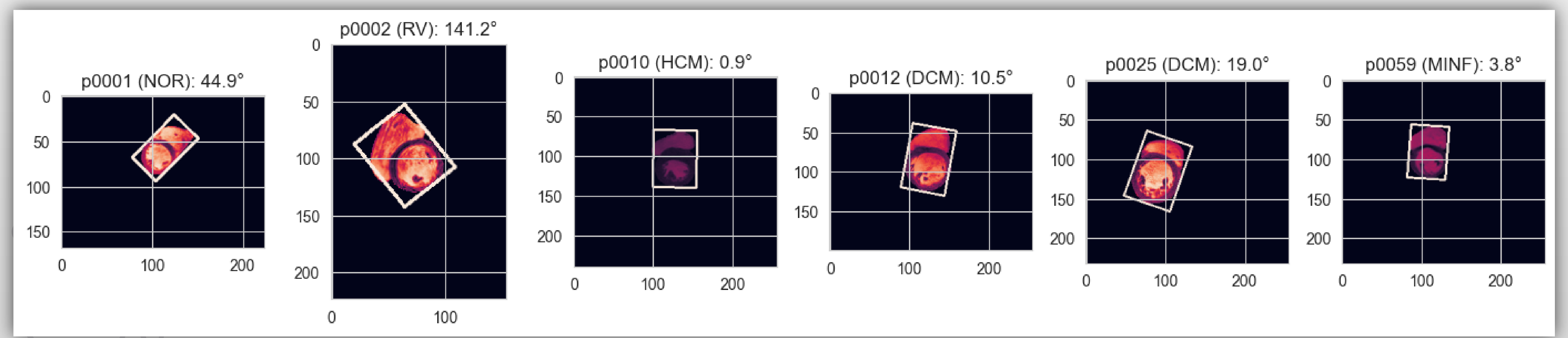


- Adaptive histogram normalization (CLAHE algorithm)
- Better contrast
- Less difference between subjects

Different Orientation

- Normally orientation is given in meta data
- Wrong meta data in some samples
- Ratio is unambiguous

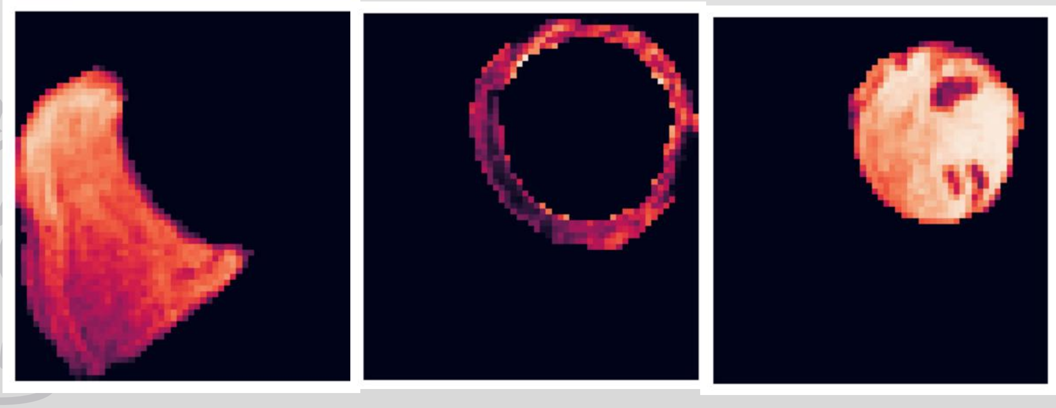
→ Solution use OpenCV to detect rotation of the heart (using the segmentation mask)

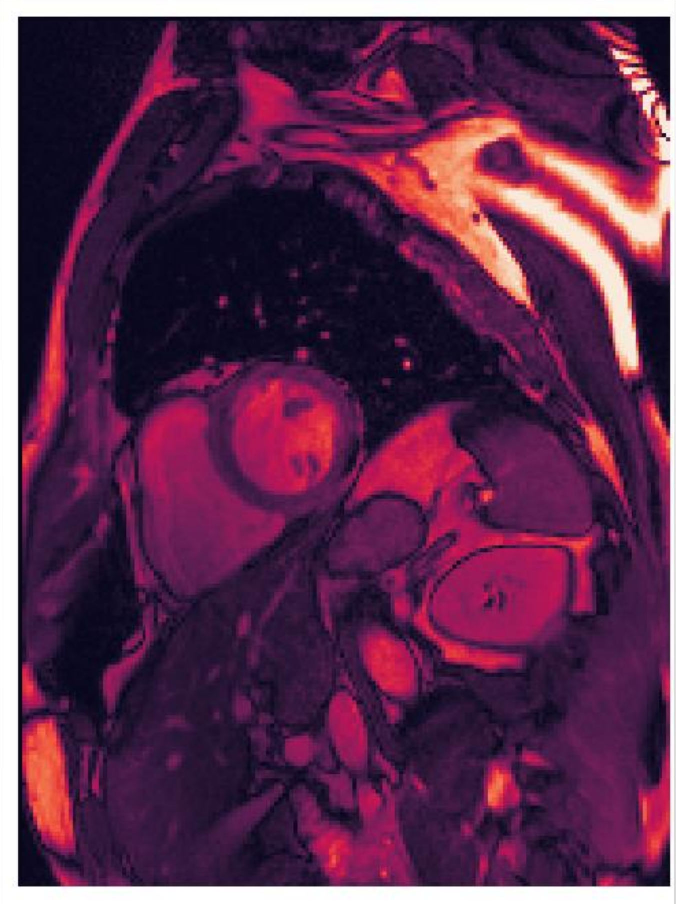


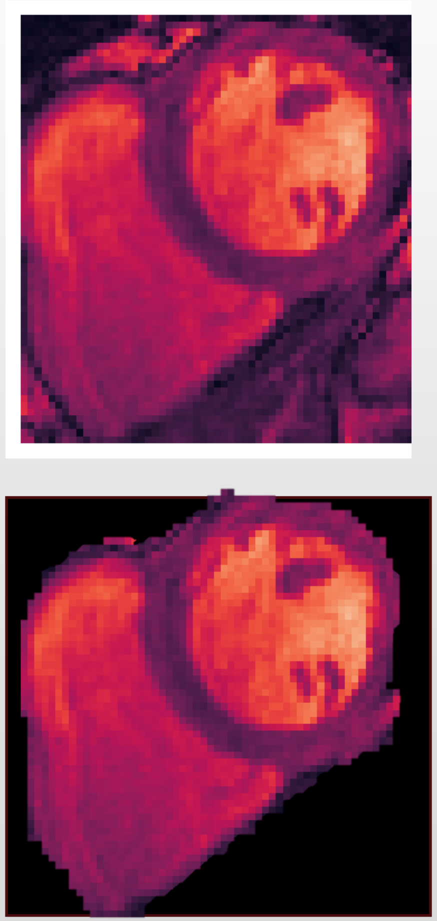
Cropping and masking

Options:

- Crop images to ROI area (bounding box around mask)
- Use the segmentations to mask the intensity
- Use the separate labels to mask different parts of the input







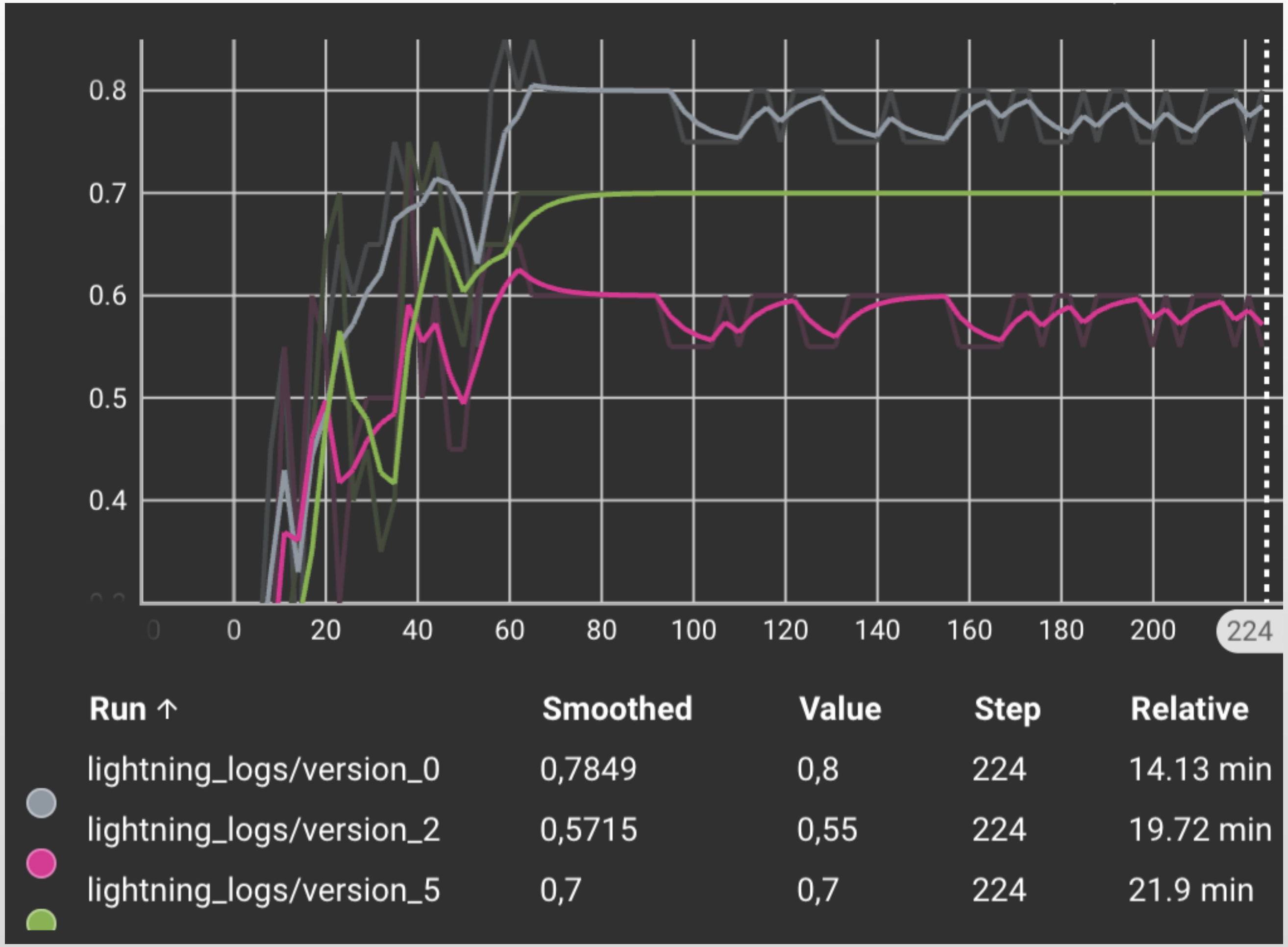
First results

- CNN trained with CrossEntropy using PyTorch Lightning
- 3D-ResNet18

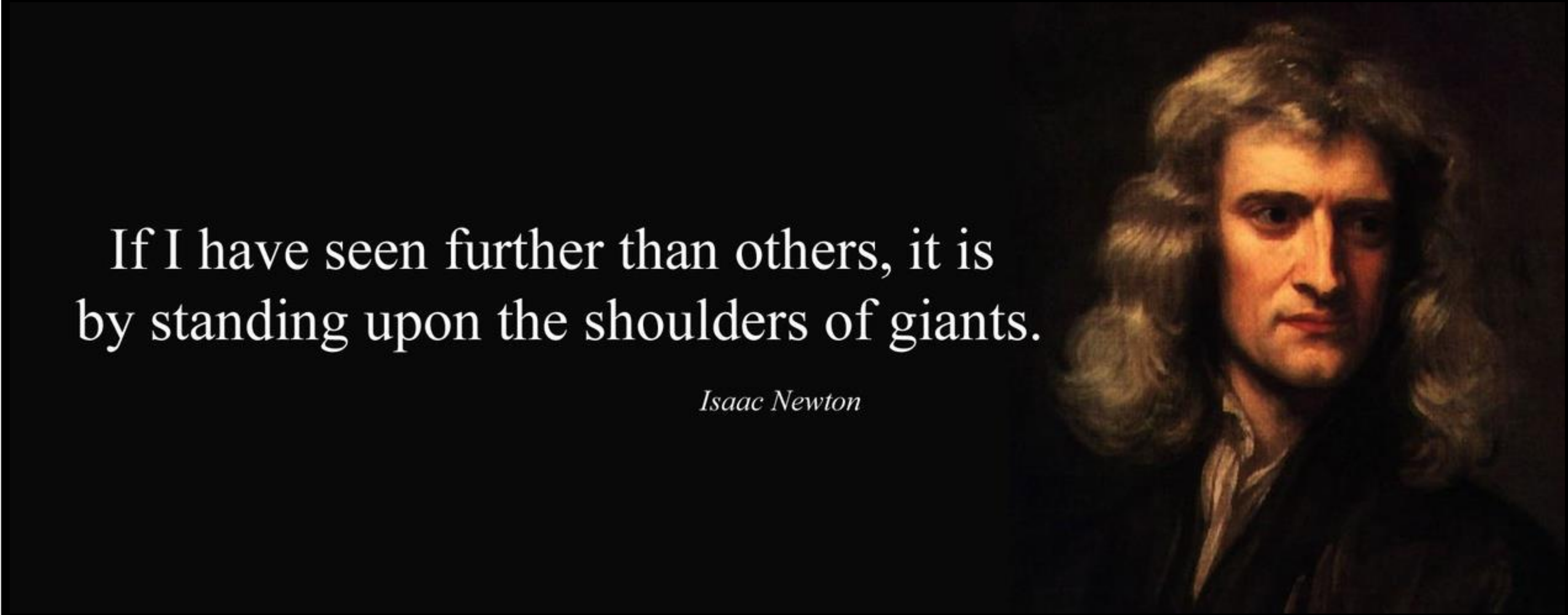
➔F1-Score: 0.7 - 0.8 (on valid)

What seemed to help:

- Intensity normalization
- Cropping
- Masking intensity
- Bigger batches / Gradient accumulation



Frameworks and other resources



Pretrained models?

- Med3D: Transfer Learning for 3D Medical Image Analysis
  - Chen et al. (2009)
  - Dataset with diverse modalities, target organs, and pathologies (MRI & CT)
  - A series of 3D-ResNet pre-trained models

- First results were not that great… F1: ~0.5-0.6
  - Potential problems:
    - Different normalization
    - Domain shift
    - Finetuning settings (frozen layers, LR, etc.)

Potential outlook

- Systematic hyperparameter Testing with optuna
- Re-check normalization used in MedicalNet
- Train the model to approximate the features as well as the class
- (Pre-) Train the model on segmentation task first
- Test different input augmentations
- Check sensitivity map / gradCAM for plausibility

Summary

- KISS – Keep it stupid simple!
- Really hard to beat the feature-based baseline
- Performance & Interpretability
- Why actually?
    - Except for learning ;-)
    - If you have domain knowledge, use it! (especially with small datasets)

Literature

- Bernard, O., Lalande, A., Zotti, C., et al. (2018). Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging*, *37*(11), 2514–2525. https://doi.org/10.1109/tmi.2018.2837502

- M. Khened, M., Alex, V., and Krishnamurthi, G. (2017). Densely connected fully convolutional network for short-axis cardiac cine mr image segmentation and heart diagnosis using random forest. Proc. STACOM-MICCAI, LNCS, 10663, 140–151.

- Isensee, F. , Jaeger, P., Full, P., et al. (2017), Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. Proc. STACOM-MICCAI, LNCS, 10663, 120–129.