Optimizing Cardiac MRI Segmentation: An Ensemble Approach with U-Net Variants

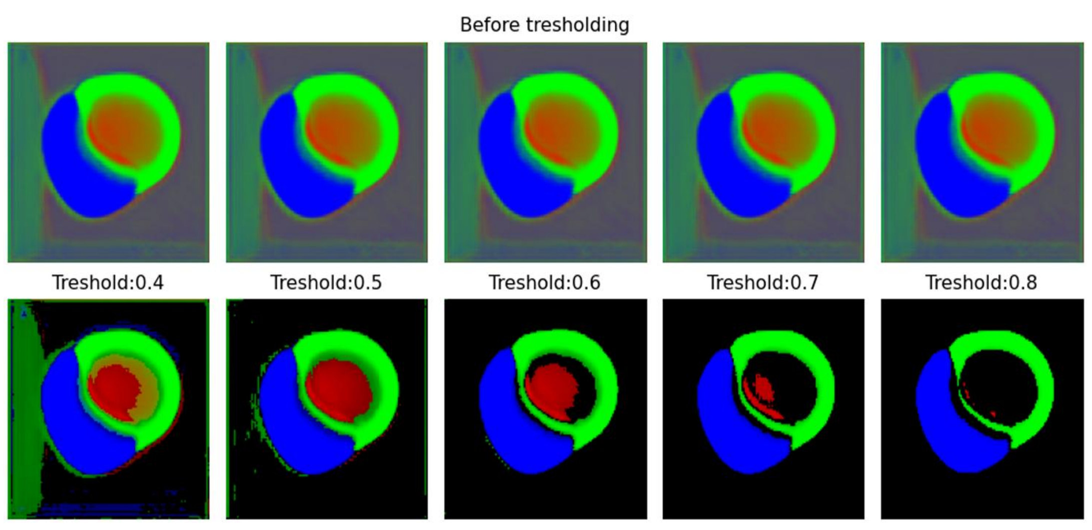
Bárdos-Deák Botond, Bodai Adrián Tibor, Mohammed Salah Al-Radhi

Introduction

- This study aims to improve performance of the segmentation of cardiac images to the left and right ventricles and the myocardium.
- To improve the performance, we used an ensemble approach with variants of the U-Net architecture.
- Multiple segmentation models were trained and optimized, and their outputs were combined using threshold-based binary conversion. And evaluated with ensemble strategies.

Models

- Four main model architectures were trained, all of them has advantages and disadvantages :
  - ➢ U-Net
  - ➢ DeepLabV3
  - ➢ MA-Net
  - ➢ LinkNet
- All models trained with different parameters and the best one from all has been chosen
- Each model independently generated segmentation masks for the left ventricle , right ventricle , and myocardium. To unify these predictions, a threshold based approach was applied to convert the continuous outputs of each model into binary masks (foreground vs. background).
- Effect of different threshold values on segmentation masks :



Evaluation Metrics

- We choose metrics to provide insights into various aspects of segmentation quality, such as overlap, accuracy, and handling of class imbalance.
- **F1-score** : The F1 score is the harmonic mean of precision and recall. Precision evaluates how many of the predicted positive pixels are correct, while recall measures how many of the true positive pixels were identified.
- **Pixel accuracy** : Measures the proportion of correctly classified pixels across the entire image
- **Intersection over Union (IoU)** : Evaluates the overlap between the predicted and ground truth masks.

Ensemble Method

- The ensemble was constructed by combining the outputs of four segmentation models: DeepLabV3, LinkNet, U-Net, and MA-Net.
- The outputs of the thresholded segmentation masks were aggregated using two distinct ensemble strategies: an **Averaging Method**, and a **Voting Method**.
- **Averaging Method :** For each pixel location, the mean value was calculated across all models:

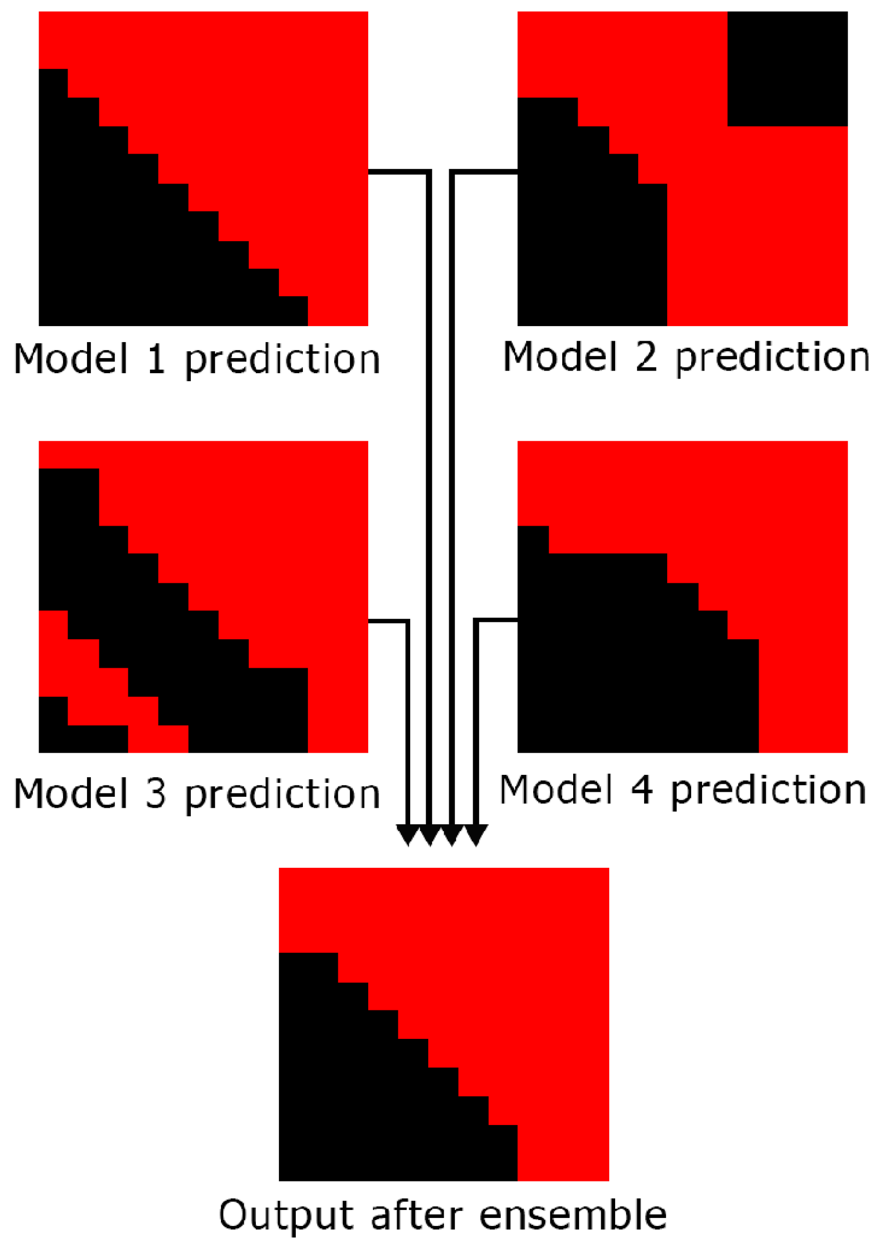
$$M_{\text{avg}}(x,y) = \frac{1}{N}\sum_{i=1}^{N} M_i(x,y)$$

- Where Mi(x,y) represents the binary mask of the i th model, and N is the total number of models in the ensemble. The resulting averaged mask smooths out in consistencies between model predictions and provides a more consistent output.
- **Voting Method :** The voting method used a majority voting scheme to determine the final classification for each pixel.

$$M_{\text{vote}}(x,y) = \begin{cases} 1 & \text{if } \sum_{i=1}^{N} M_i(x,y) > \frac{N}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

- This strategy is particularly robust to errors made by individual models, as it relies on consensus among the majority. It is especially useful in reducing the impact of outliers and noise in the predictions.
- The combination of these methods ensures that the ensemble achieves a robust solution for cardiac MRI segmentation tasks. er computational capacity.

Example for voting method



Results

Model	F1Score	Pixel Accuracy	IoU
UNet	0.33	0.98	0.31
DeepLabV3	0.31	0.98	0.29
LinkNet	0.32	0.98	0.30
MAnet	0.32	0.98	0.30
Ensemble	0.34	0.98	0.32

- The ensemble model slightly outperformed the individual models in all three metrics, achieving the highest F1 Score (0.34) and IoU (0.32), while maintaining the same high Pixel Accuracy (0.98). This demonstrates the effectiveness of the ensemble approach in improving segmentation performance

Budapest University of Technology and Economics
Department of Telecommunications and Artificial Intelligence
Budapest, Hungary