AortaSeg: Technical Description (lWM)

Marek Wodzinski[1,2] and Henning Müller[2,3]

[1]AGH University of Krakow, Department of Measurement and Electronics
Krakow, Poland
[2]University of Applied Sciences Western Switzerland (HES-SO Valais-Wallis)
Information Systems Institute, Sierre, Switzerland
[3] University of Geneva, Medical Faculty, Geneva, Switzerland
wodzinski@agh.edu.pl

Abstract. This document presents the short technical description of the method proposed for the AortaSeg challenge organized during the MICCAI 2024 conference. The goal of the work was to accurately perform multiclass aorta segmentation to partition it into meaningful and anatomically defined segments. We applied the data-centric approach and focused more on a proper data preparation, pre- and post-processing instead of the recent advances in segmentation architectures. We verified that the classical RUNet-based architectures outperform the more recent advances, at least in low data regimes.

Keywords: AortaSeg · Image Segmentation · Aorta · MICCAI · Challenge

1 Technical Description

1.1 Overview

The proposed approach is a two-step segmentation pipeline that starts with segmenting the whole aorta (binary segmentation). The output of the first step is then concatenated with the input image and passed to the second step that calculates the multiclass segmentation. The process for both binary and multiclass segmentation is multifold and patch-based. The idea behind the two-step approach was to ensure the connectivity between the classes that was naturally enforced by the binary segmentation step.

1.2 Models

During the algorithm development phase we evaluated several segmentation architectures: ResUNet, SwinUNETR, UNETR, AttentionUNet, SegMamba. During the internal evaluation we decided to use ResUNet which outperformed all other architectures by a considerable margin – probably due to limited dataset size (no external data for self-supervised pretraining was allowed). The ResUnet was applied both for the binary and multiclass segmentation.

1.3 Preprocessing & Postprocessing & Inference

The inference starts with splitting the image into patches (256x256x256) for the binary segmentation. Then, the patches are propagated through the first model. The binary segmentation (activation map, without thresholding) is concatenated with the input images and again divided into patches (192x192x192). The patches are propagated through the second model and the label map is produced by applying the argmax operator to the network output. The patches for both segmentation steps were propagated with 32x32x32 overlap.

During inference we applied two folds for both binary and multiclass segmentation. The reason to use only two folds was motivated by the 5 minutes inference limit that accounts also the Docker loading time. Therefore, to reduce the Docker size (to reduce the loading time accounting for 70of inference time) we reduced the number of folds. The real inference time (excluding Docker loading) was on average 50 seconds.

1.4 Training

The training was performed using the original resolution images, without resampling. The images were divided into patches - 256x256x256 and 192x192x192 for the binary and multiclass segmentation, respectively. The training was performed using TorchIO Queue interface using a single A100 GPU with 40 GB of VRAM, batch size equal to 1, optimizer update interval equal to 16 batches. The AdamW was used as the optimizer together with exponentially decaying learning rate scheduler. All folds were trained until convergence. The training was performed using the PLGRID supercomputing infrastructure. The final objective function was a combination of Generalized Dice Loss and Focal Loss, both available in the MONAI library.

2 Source Code & Instructions

Source code and details on how to reproduce the results are available at: `https://github.com/MWod/AortaSeg_2024`

3 Team Details

The contribution was prepared by the following team:
Grand-Challenge Username: **lWM**
Team name: **AGHSSO**

– Participant 1:
  - Marek Wodzinski
  - Email: *wodzinski@agh.edu.pl*
  - Affiliation 1: Department of Measurement and Electronics, AGH University of Kraków, Krakow, Poland

  - Affiliation 2: Information Systems Institute, University of Applied Sciences Western Switzerland (HES-SO Valais), Sierre, Switzerland
- Participant 2:
  - Henning Müller
  - Email: *henning.mueller@hevs.ch*
  - Affiliation 1: Information Systems Institute, University of Applied Sciences Western Switzerland (HES-SO Valais), Sierre, Switzerland
  - Affiliation 2: Medical Faculty, University of Geneva, Geneva, Switzerland

Acknowledgements We gratefully acknowledge Polish HPC infrastructure PLGrid support within computational grant no. PLG/2024/017079. The research was partially supported by the program "Excellence Initiative - Research University" for AGH University.

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.