A. MODEL SUMMARY

A1. Background on you/your team

- Competition Name: SenNet + HOA - Hacking the Human Vasculature in 3D
- Team Name: Ivan Panshin
- Private Leaderboard Score: 0.691797
- Private Leaderboard Place: 5th place

- Name: Ivan Panshin
- Location: Haifa, Israel
- Email: ivan.panshin@protonmail.com

A2. Background on you/your team

Ivan Panshin

- What is your academic/professional background?

I am a Team Lead of a Deep Learning team consisting of 6 people at Tensorway. Currently we're developing DL solutions for a variety of industries, as well as using state of the art models from a lot of DL areas (Computer Vision, NLP, etc)..

- Did you have any prior experience that helped you succeed in this competition?

Yes, this is my 6th gold medal. All of them are in Computer Vision, and 3 of them - in Medical Imaging.

- What made you decide to enter this competition?

I love Computer Vision, but even more - I love the application of Computer Vision in Medical Imaging.

- How much time did you spend on the competition?

About 100 hours in total.

- If part of a team, how did you decide to team up?

I competed solo.

- If you competed as part of a team, who did what?

I competed solo.

A3. Summary

4-6 sentences summarizing the most important aspects of your model and analysis, such as:

- The training method(s) you used (Convolutional Neural Network, XGBoost)

I used a combination of 2D U-Net models with post-processing to account for the 3D nature of the data.

To be more precise, several convolution and transformer based backbones were used to support such architecture: EffNet_m, DPN68, MaxVit_base.

3D images were split into crops of 512x512 with overlaps of 256x256. Crops were made along 3 axis: xy, xz and yz.

The models were trained with a combination of losses: Dice, Focal and CE with more weight to the boundaries of masks.

- The most important features

Pseudo labeling of additional data and composite loss with more weight to the boundaries of masks.

- The tool(s) you used

Code was written in Python with usage of the following modules: PyTorch, segmentation_models_pytorch.

- How long it takes to train your model

Roughly 1 week on RTX A6000 Ada.

A4. Features Selection / Engineering

- Did you use external data? (if permitted)

Yes, I uses external data available at https://human-organ-atlas.esrf.eu/. In particular, I used additional kidney and spleen data.

A5. Training Method(s)

- What training methods did you use?

Semantic segmentation training with hard augmentations, composite loss, big batches + EMA for training stability.

- Did you ensemble the models?

Yes, in my experience, it's beneficial to ensemble semantic segmentation models that produce logits of the same scale. Additionally, it's benifical to ensemble models on the logits-level as opposed to probabilities-level.

Both intuitions were true in this competition.

- If you did ensemble, how did you weight the different models?

Equal weights.

A6. Interesting findings

- What was the most important trick you used?

More data and composite loss. If I had to pick one, it would definitely be more data. In particular, training on soft-labels on new data (never the hard labels).

- What do you think set you apart from others in the competition?

Not overfitting to LB, and making the solution stable across different images and confidence thresholds.

Additionally, I adjusted to the private resolution by doing the 3D resize with the trilinear interpolation.

- Did you find any interesting relationships in the data that don't fit in the sections above?

Sure. Knowledge from kidneys transfer quite nicely to spleens, but not lungs, and especially brain or heart.

A7. Simple Features and Methods

Many customers are happy to trade off model performance for simplicity. With this in mind:

- Is there a subset of features that would get 90-95% of your final performance? Which features? *

Use slices only along the xy axis, take 2D U-Net with transformer backbone, composite loss, and crops from different organs not to overfit to a particular 3D image.

- What model that was most important?

Model based on transformer backbone (MaxVit base)

A8. Model Execution Time

Many customers care about how long the winning models take to train and generate predictions:

- How long does it take to train your model?

About 1 week on a single RTX A6000 Ada.

- How long does it take to generate predictions using your model?

For the ensemble - about 8 hours for both public + private datasets. For a single model - about 1/3 of that.

A9. References

Citations to references, websites, blog posts, and external sources of information where appropriate.

You can read more details about my solution on kaggle forum:

1) https://www.kaggle.com/competitions/blood-vessel-segmentation/discussion/475288