Recursive learning reinforced by redefining the train and validation volumes of an Encoder-Decoder segmentation model

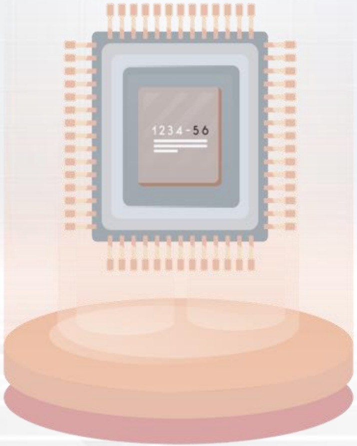
Antonio Vispi
Avionic Systems Engineer at Leonardo Helicopters
From Turin Polytechnic

antoniovispi1@gmail.com

INTRODUCTION

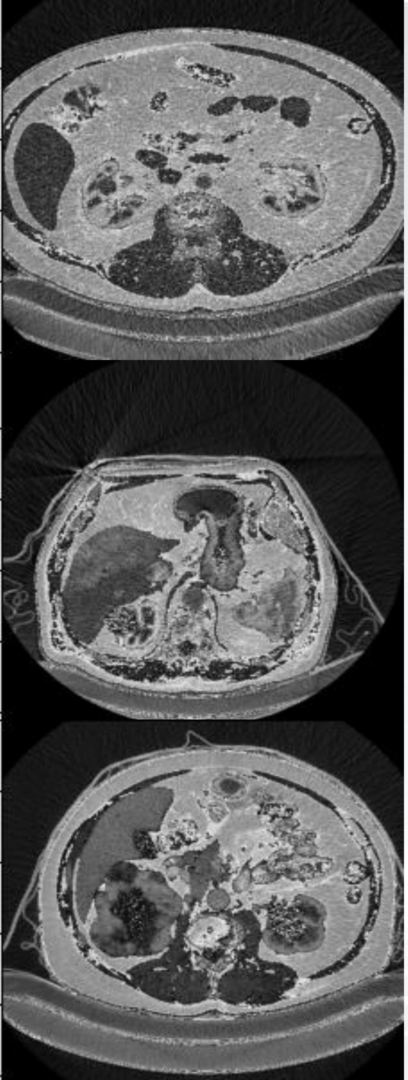
- This work addresses a critical need in the diagnosis of kidney cancer, one of the most prevalent and deadly cancers today. By accurately identifying lesions through segmentation, it paves the way for more targeted and successful treatments, ultimately improving patient outcomes.

- The final goal is to revolutionize the CT image segmentation process, specifically focusing on the precise delineation of the kidney and its potential pathological masses.



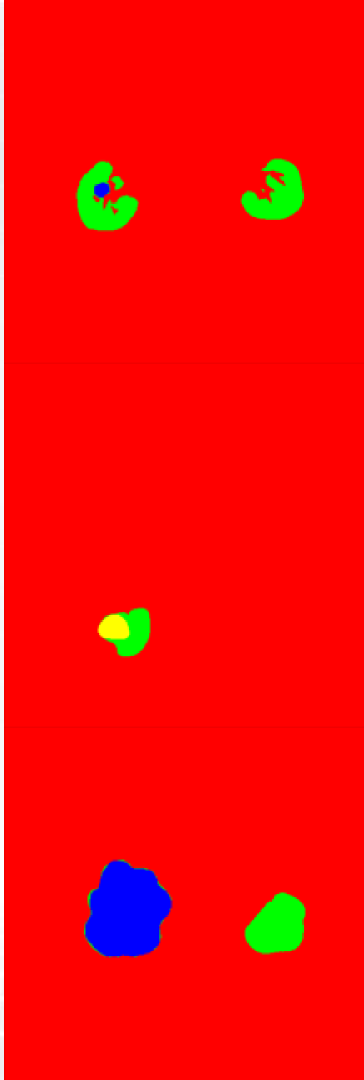
- This approach leverages Deep Learning, employing an Encoder-Decoder architecture with an EfficientNet-B5 encoder and a Unet decoder, thoughtfully customized for optimal performance.

- To reach good accuracy, it was used a cascading training method, at the beginning of each round a complete redefinition of training and validation data takes place. This strategy allowed the model to handle extensive data, enhancing its generalization capabilities.

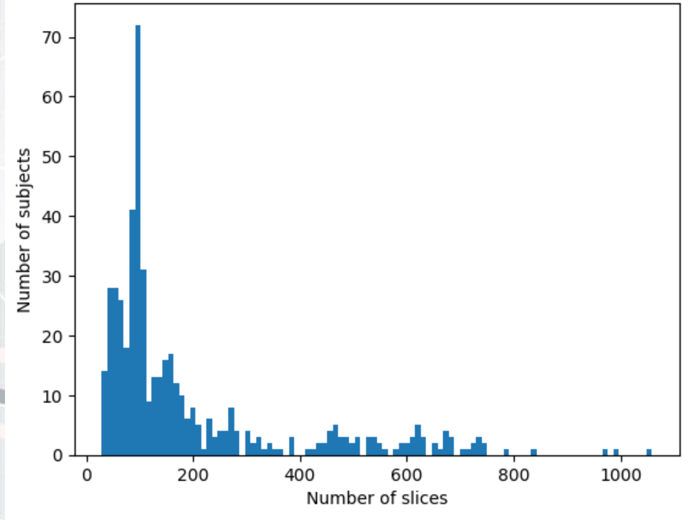


LEGEND

Background	
Kidney	
Tumor	
Cyst	

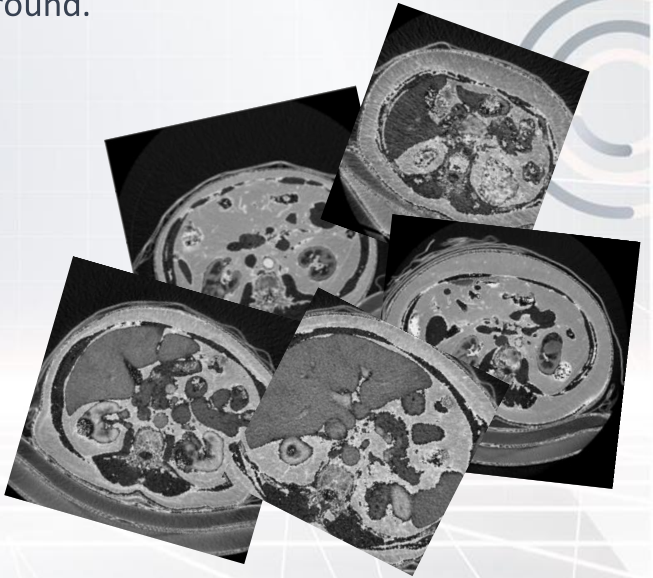


DATA PREPARATION



- The number of slices for each subject showed high variability.

- Slices were randomly selected from each subject in proportion to their individual slice counts, ensuring representation from all subjects and introducing diversity into the training and validation volumes, redefined each round.



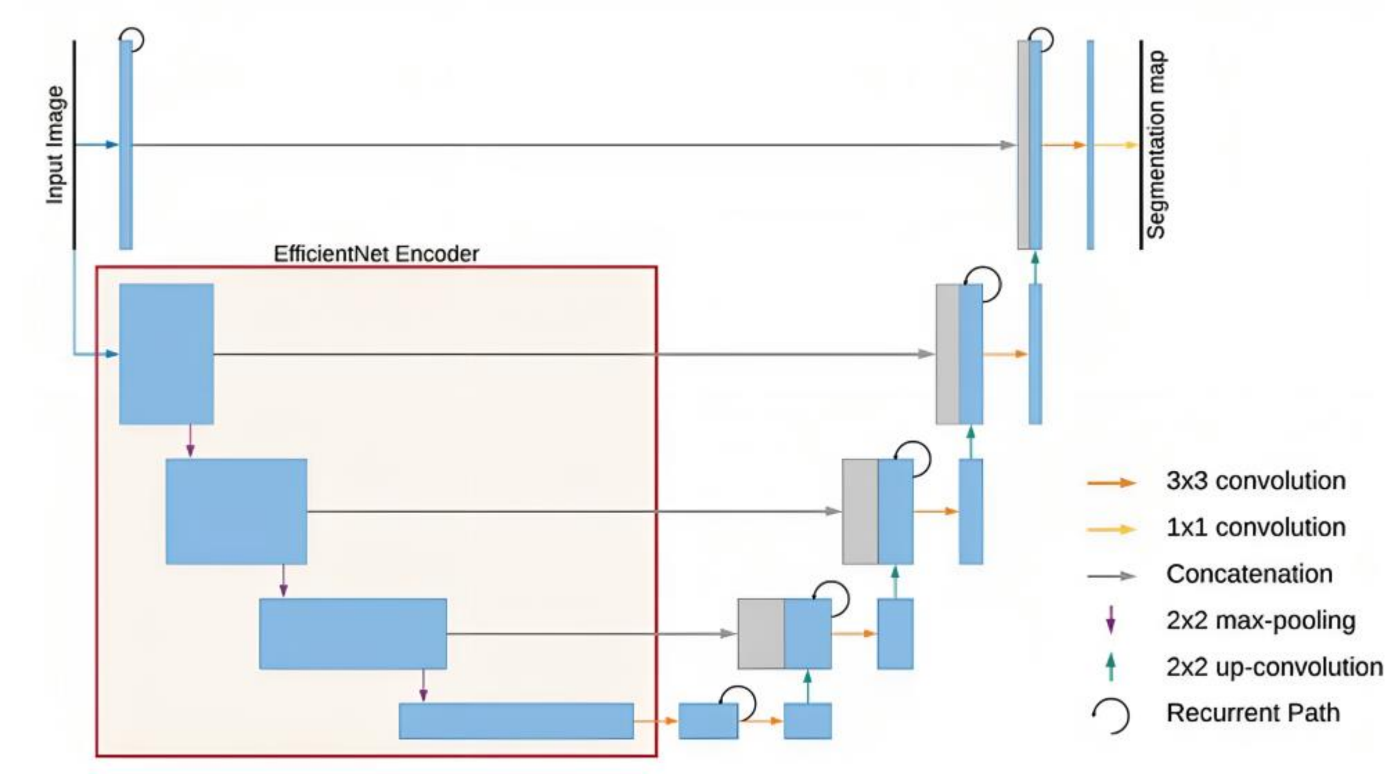
To combat overfitting, data augmentation techniques, such as rotations, zoom, dimension changes, and flips, were applied.

No normalizations or subject selections were performed during either the training or testing phases to preserve the model's generalization capacity.

MODEL ARCHITECTURE

- EfficientNet-B5 as the **encoder**: selected for its efficiency
- Unet as the **decoder**, with an increased number of filters

This architecture balances segmentation quality with computational efficiency, resulting in interesting performance.

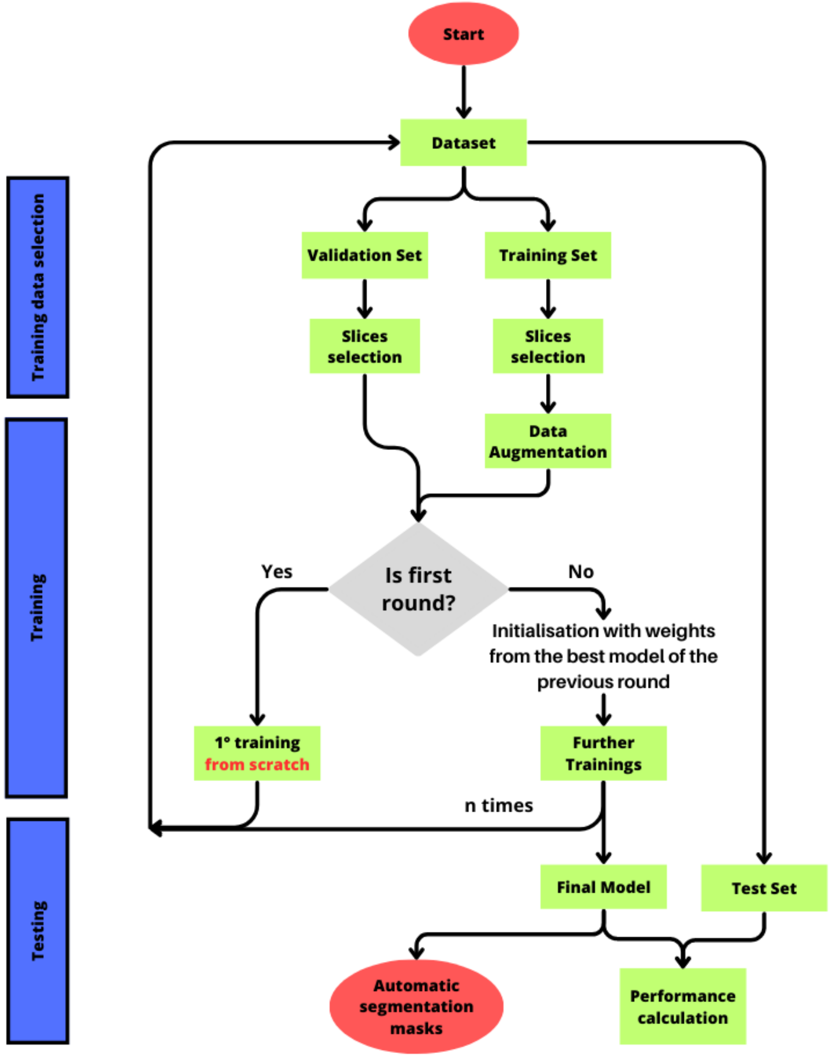


TRAINING ARCHITECTURE

Refined data selection mechanism

Ever-improving training

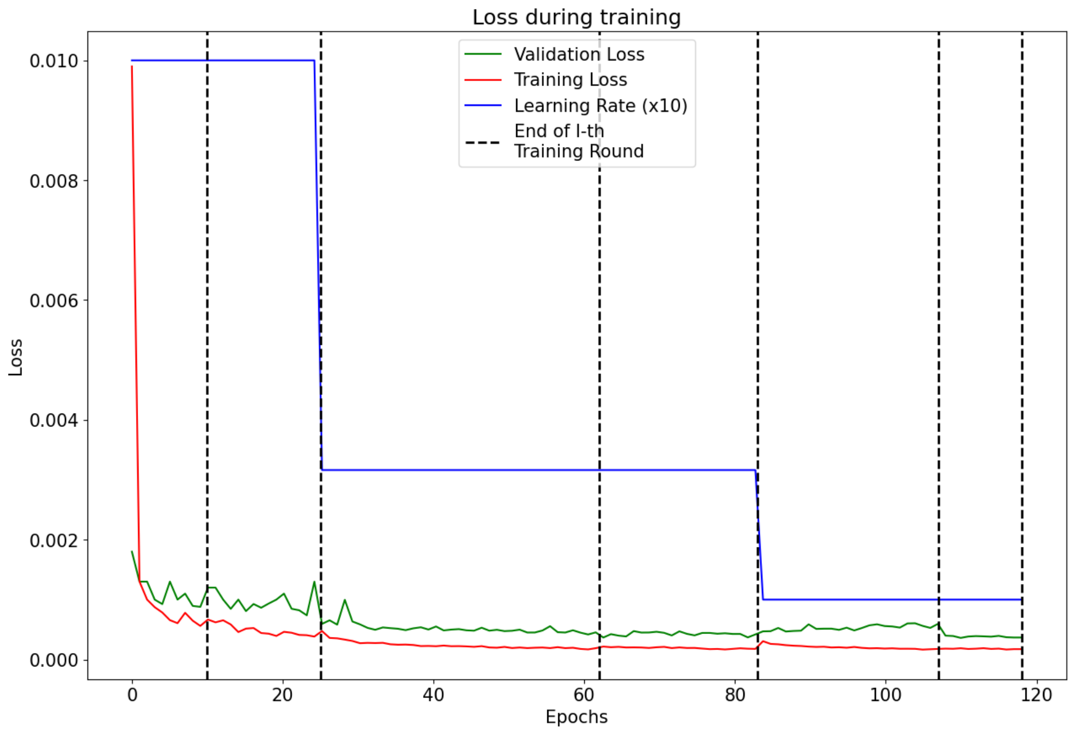
Testing



LOSSES

- The training process consisted of six successive rounds.
- The initialization of the next round consists of the best configuration of the entire previous round, forcing the model to improve during the various rounds.
- Each training round lasts approximately 10 hours.

- The inference time of the trained model lasts around 50 milliseconds per image with V100 GPU.



- The NADAM optimizer was utilized with default hyperparameters.
- The learning rate was manually adjusted promoting steady improvement.
- The best model from the final round, was used for the paper's results.

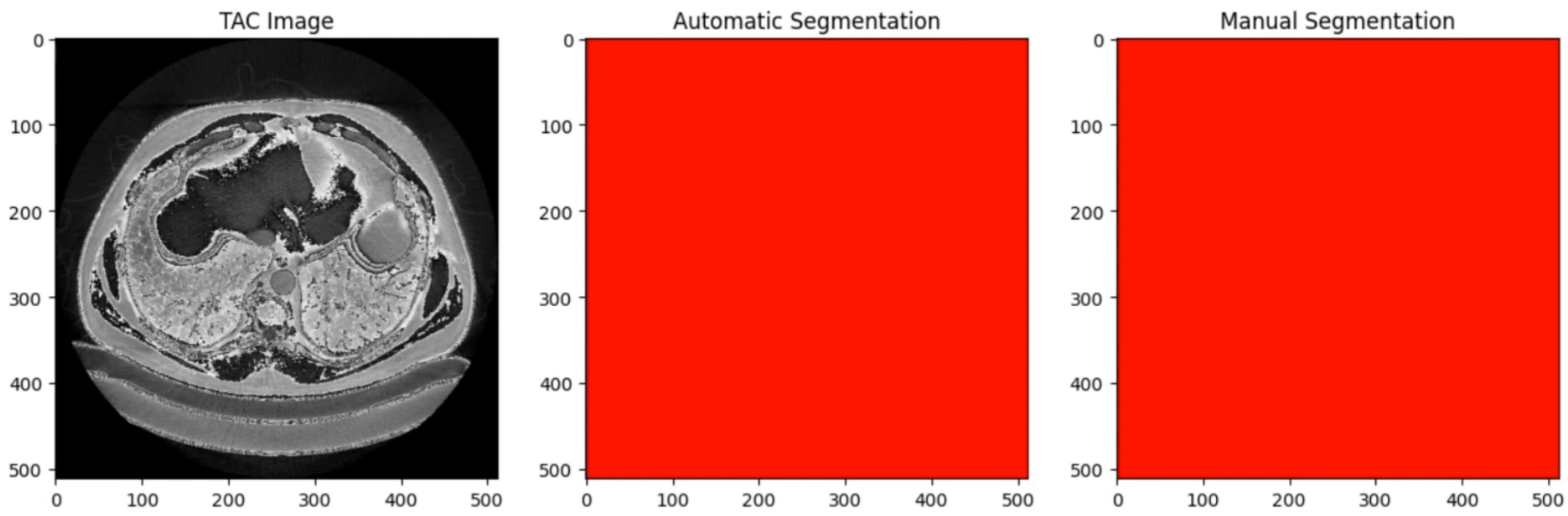
FINAL RESULTS

- Dice Similarity Coefficient (DSC): ranging from 0 (no overlap) to 1 (perfect overlap).
- Relative Volume Difference (RVD): 0 represents an optimal match.

Training Set	**Mean DSC (%) $\pm$ STD**	**Mean RVD**
Kidney + Tumor + Cyst	97.80 $\pm$ 0.8	0.0040 $\pm$ 0.020
Tumor + Cyst	80.98 $\pm$ 19.2	-0.0383 $\pm$ 0.622
Tumor	80.74 $\pm$ 20.4	-0.1503 $\pm$ 0.257
Test Set	**Mean DSC (%) $\pm$ STD**	**Mean RVD**
Kidney + Tumor + Cyst	97.71 $\pm$ 1.1	-0.0112 $\pm$ 0.017
Tumor + Cyst	81.39 $\pm$ 17.4	-0.0215 $\pm$ 0.659
Tumor	73.81 $\pm$ 24.6	-0.1065 $\pm$ 0.491

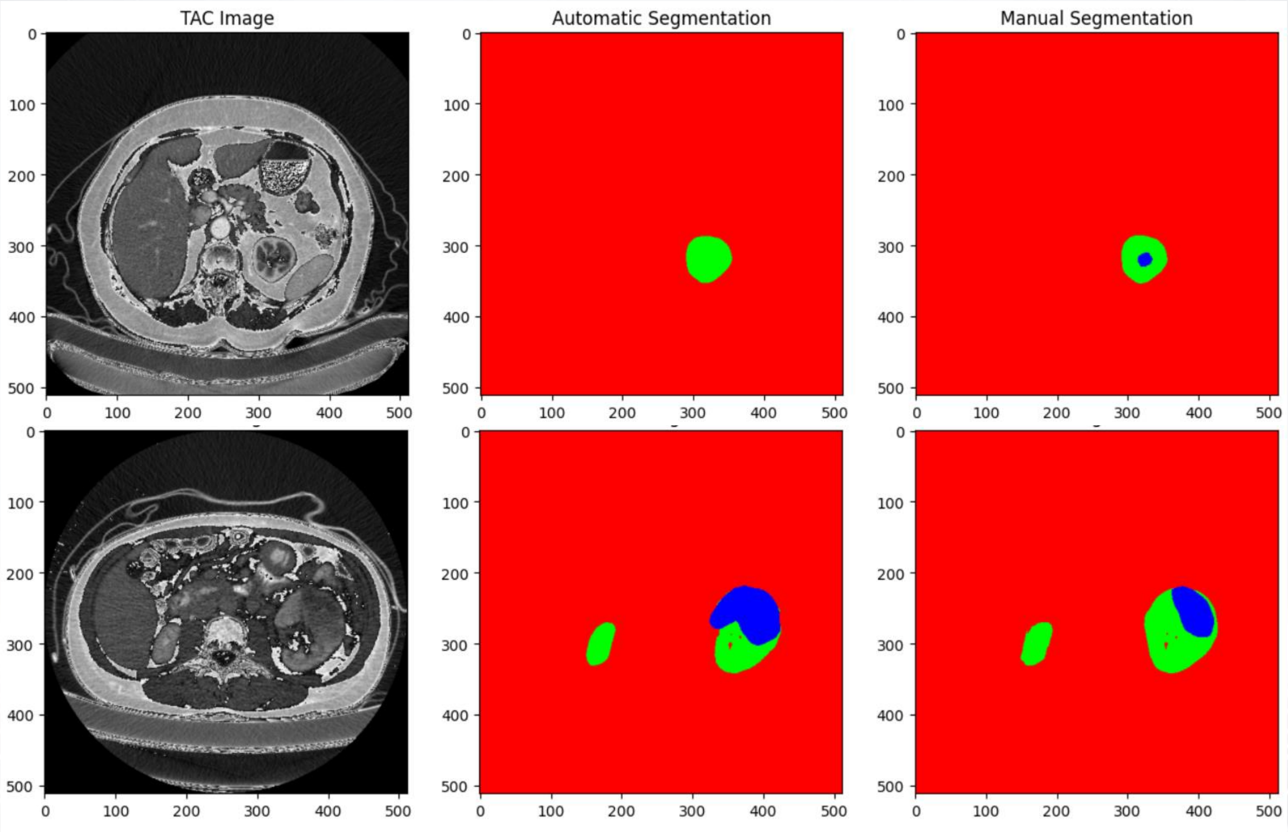
Place	**Team**	**Average Rank**	**Dice**	**Surface Dice**	**Tumor Dice**	**Kidney + Masses Dice**	**Masses Dice**	**Kidney + Masses SD**	**Masses SD**	**Tumor SD**
20	Antonio Vispi	20.0	0.719	0.572	0.570	0.941	0.645	0.869	0.453	0.393

Model generalization ability is not excellent, likely due to overfitting, sub-optimal hyperparameters, or architectural limitations.


TAC Image
Automatic Segmentation
Manual Segmentation
0
100
200
300
400
500

DRAWBACKS

Model tends to both over-segment and sub-segment, or sometimes even to omit the lesion.



CONCLUSIONS

- One key limitation is that the model makes inference on one image at a time without considering correlations with adjacent slices, potentially hindering segmentation accuracy.
- Another challenge is the nature of the processed CT images, which are noisy and make automatic segmentation difficult. Maybe using MR images could yield better results.
- The study suggests that deep learning-based automatic segmentation of the kidney and its pathological masses holds promise for improving medical diagnosis quality and efficiency.
- Through careful exploration of training configurations across multiple rounds, satisfactory results were achieved.

THANKS!

https://github.com/AntonioVispi

https://www.linkedin.com/in/antonio-vispi-ba215b243/