Kaggle competition

HuBMAP + HPA - Hacking the Human Body

Segment multi-organ functional tissue units

고민수, 박수영, 서동관, 윤태우

Summary

Problem

- Image’s size
- apply to model

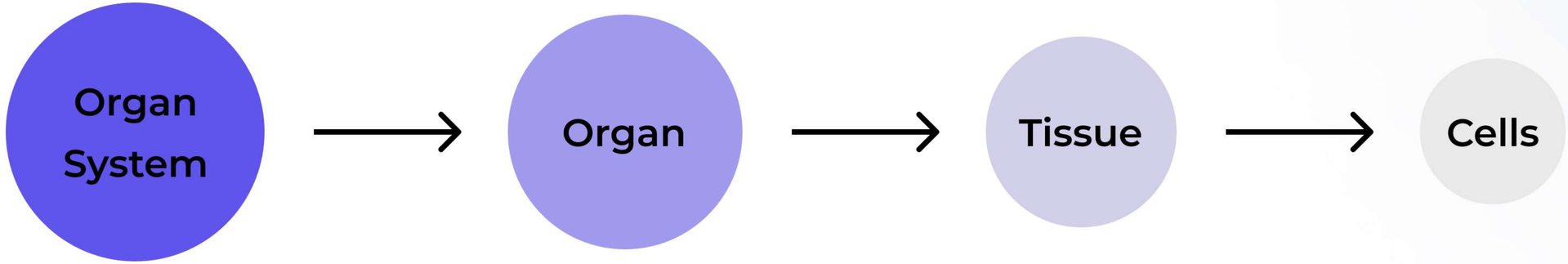
Works

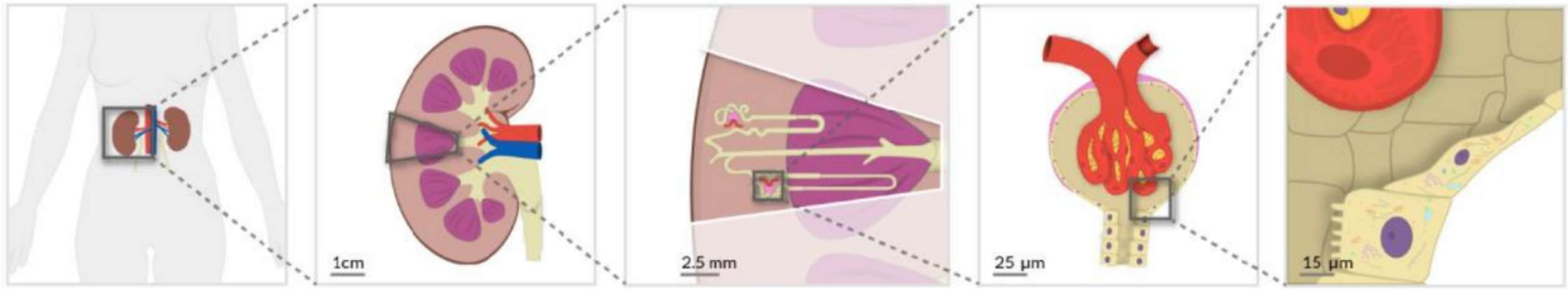
- crop
- model search

Problem

Discription

FTU : Functional Tissue units





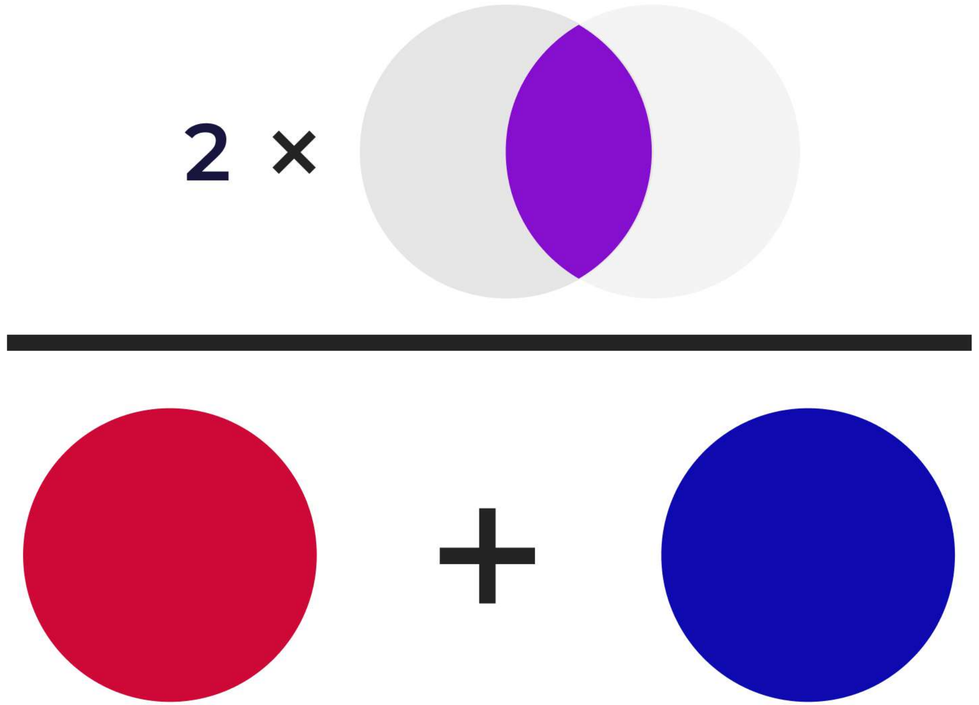
Body	Organ	Functional Tissue Unit	FTU Sub-structure(s)	Cellular
• Body	• Renal capsule	• Nephron	• Bowman's capsule	• Parietal epithelial cell
• Kidney (Left, Right)	• Renal pyramid	• Renal corpuscle	• Glomerulus	• Capillary endothelial cell
• Aorta	• Renal cortex	• Proximal convoluted tubule	• Efferent arteriole	• Mesangial cell
• Renal artery	• Renal medulla	• Loop of Henle	• Afferent arteriole	• Podocyte
• Renal vein	• Renal calyx	• Distal convoluted tubule		
• Ureter	• Renal pelvis	• Connecting tubule		

Construction and Usage of a Human Body Common Coordinate Framework Comprising Clinical, Semantic, and Spatial Ontologies(HuBMAP)

Evaluation & Judges

Model evaluation : Dice coefficient

The area of Overlap divided by the total number of pixels in both the images



Evaluation & Judges

Judges

Methodology (80 points)

- Are the statistical and modeling methods appropriate for the task?
- Are metrics provided that help interpret the results achieved by the segmentation methods?
- Is the presented characterization of FTUs useful for understanding individual differences?
- Is it possible to predict FTU area size distribution, given age and sex info across all organs?
- Did the team validate their methods and algorithm implementations and provide information on algorithm performance and limitations?
- Did the team document their method and code appropriately?
- Did the team develop a creative or novel method to segment FTUs?
- Did the team provide insights that would be useful for generating reference FTUs for inclusion into a Human Reference Atlas?

Evaluation & Judges

Judges

Diversity and Presentation (30 points)

- Does the team embrace diversity and equity, welcoming team members of different ages, genders, ethnicities, and with multiple backgrounds and perspectives?
- Did the authors effectively communicate the details of their method for segmenting FTUs, and the quality and limitations of their results?
- Are the important results easily understood by the average person?

EDA

Data

ㄴ Train / Test .csv

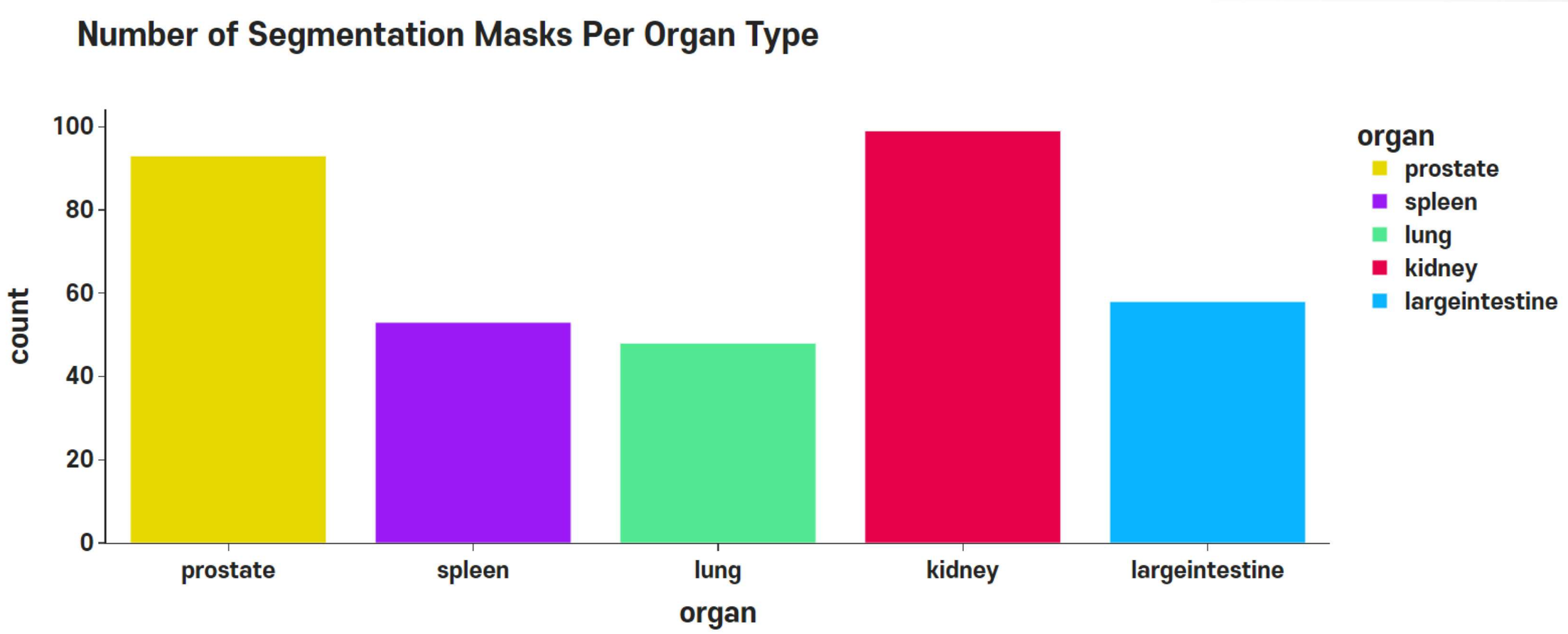
- Total data : 351
- id : image_id
- organ : [kidney, large intestine, lung, spleen, prostate
- img_height, img_widht : [2300 ~ 3000], Height and width are same
- data_sourcs : [HPA or HuBMAP], All data in dataframe are from HPA
- pixel_size : [0.4(μm)], The data in HuBMAP is different.
- tissue_thickness : [4(μm)], The data in HuBMAP is different.
- rle : Target (training set only)
- age : [20~84], Patient's age (training set only)
- sex : [Male, Female], Patient's sex (training set only)

ㄴSubmission .csv

- Total data : 1
- id : image_id
- rle : Target

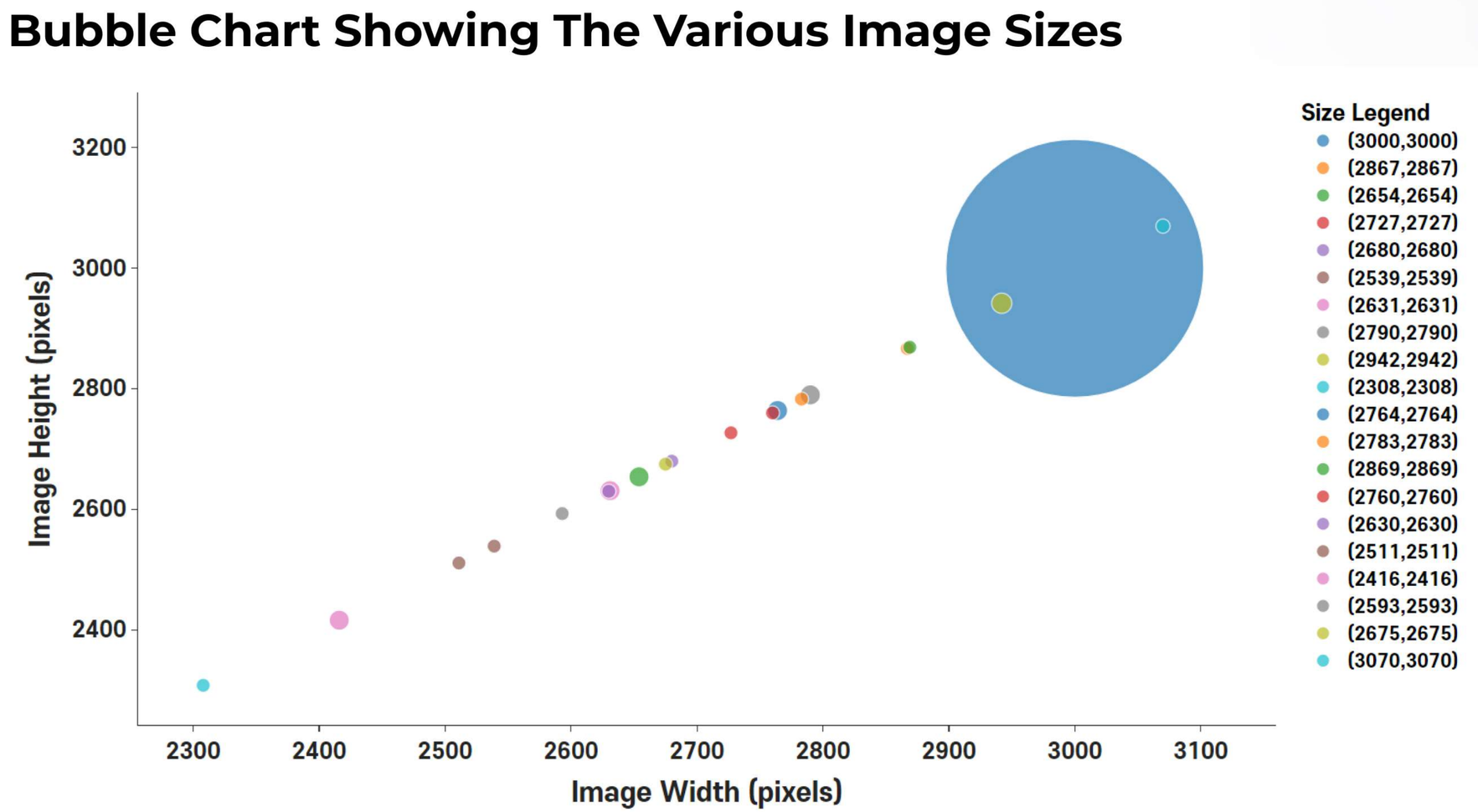
EDA

Frequency of data by organ



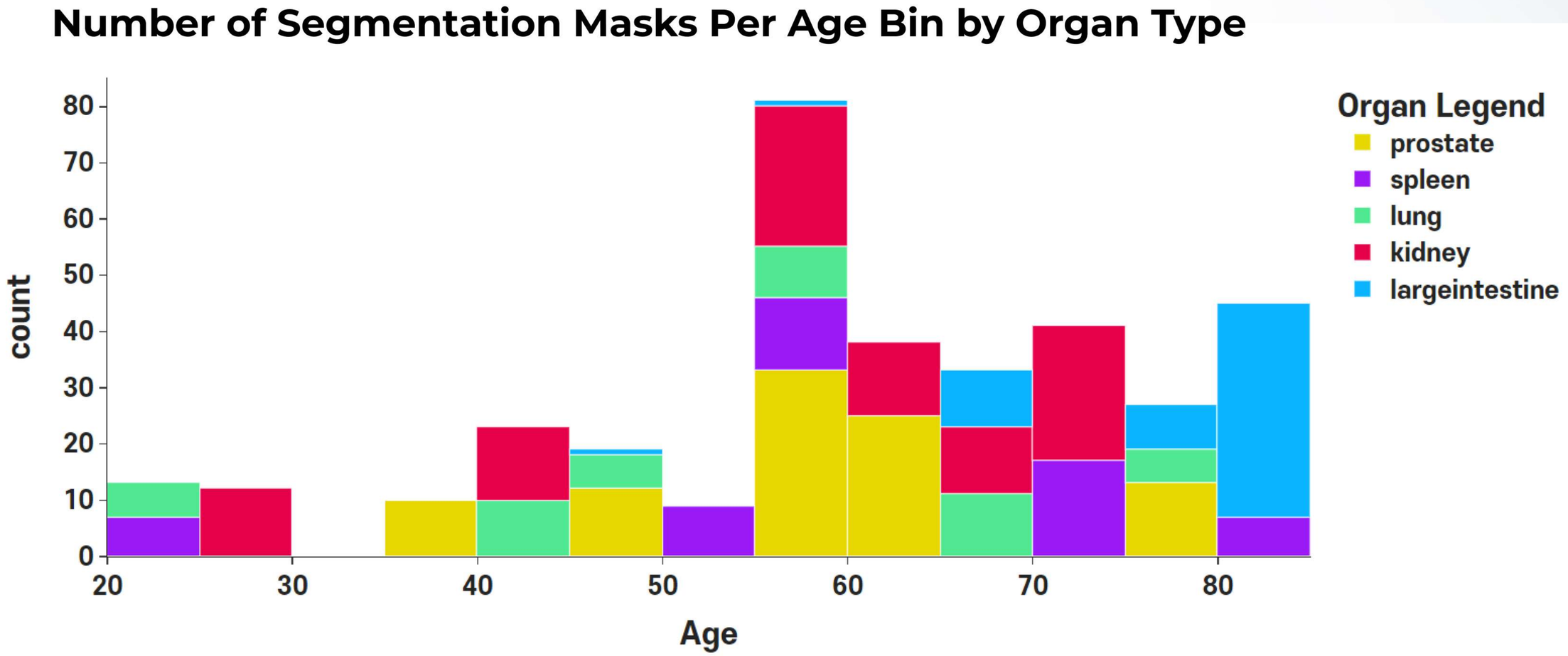
EDA

Image size distribution



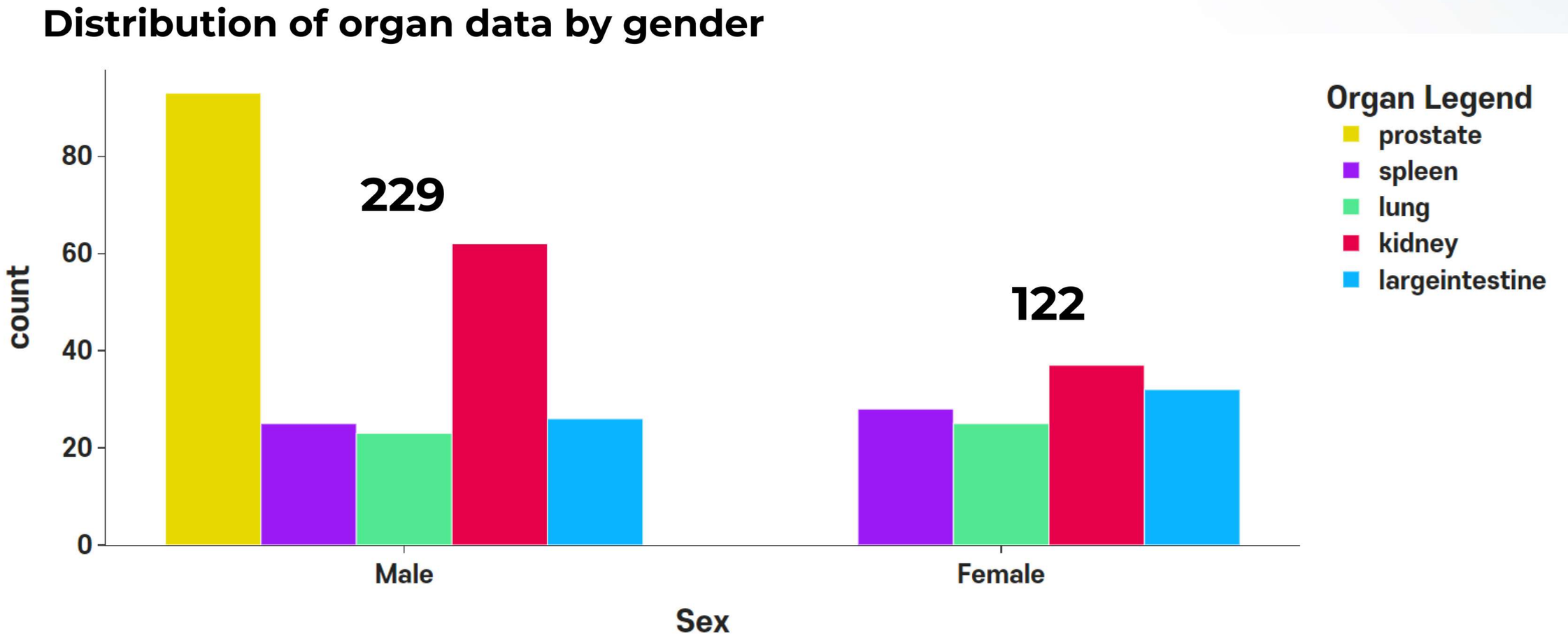
EDA

Age distribution



EDA

Sex distribution

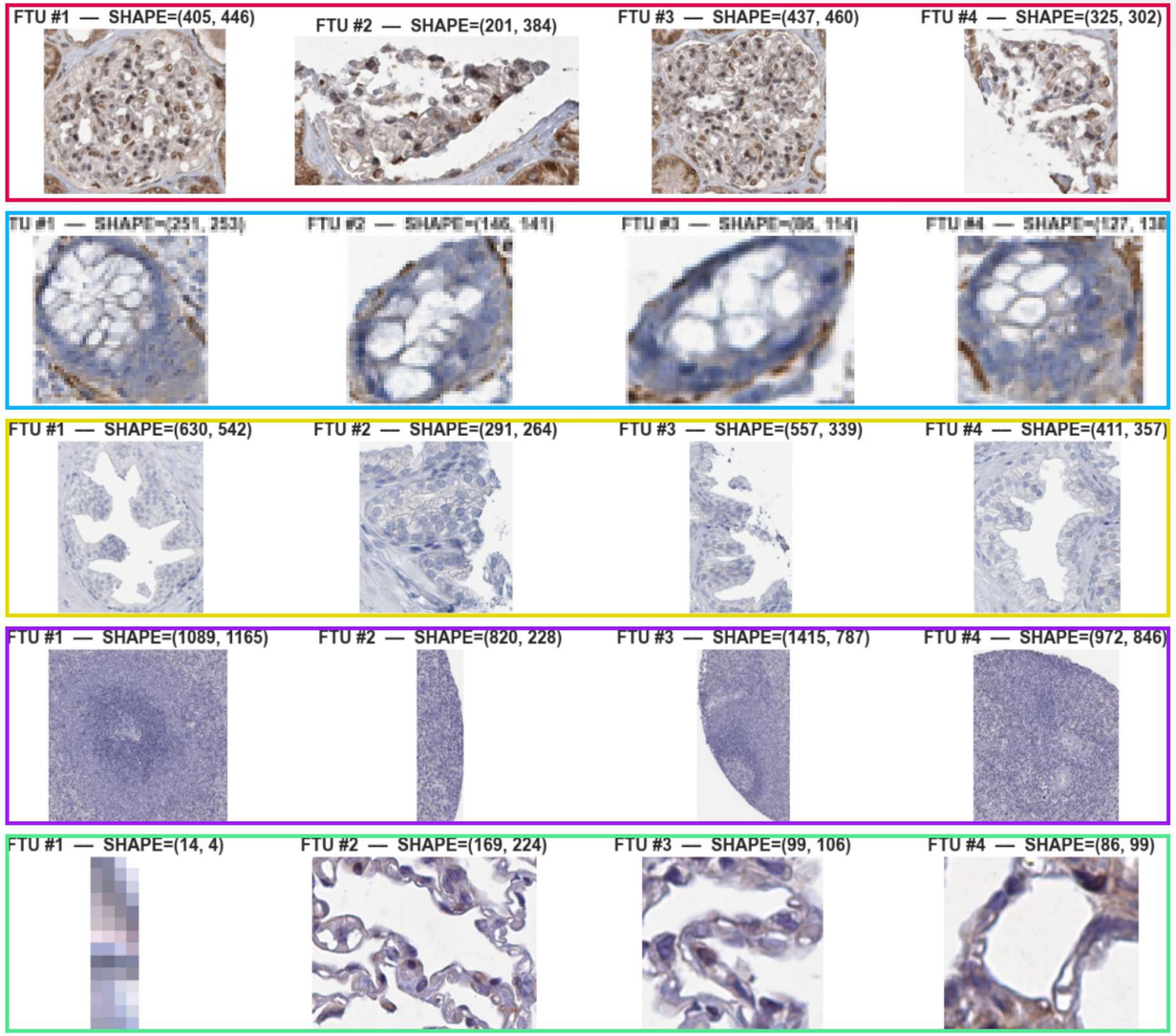


EDA

FTU bounding box Crop by organ

- **Crop FTU images with json**
  - Kidney FTU image counts : 337
  - largeintestine FTU image counts : 3117
  - lung FTU image counts : 191
  - prostate FTU image counts : 1097
  - spleen FTU image counts : 167

- **Remove very small images**
  - largeintestine FTU image counts : 14
  - lung FTU image counts : 3



EDA

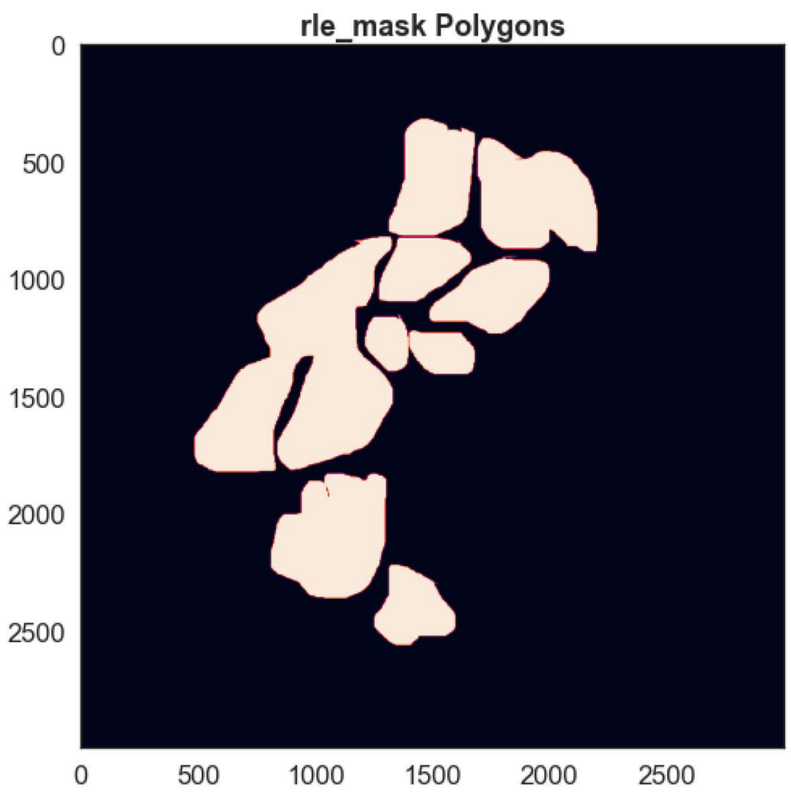
Average FTUs bounding box size by organ

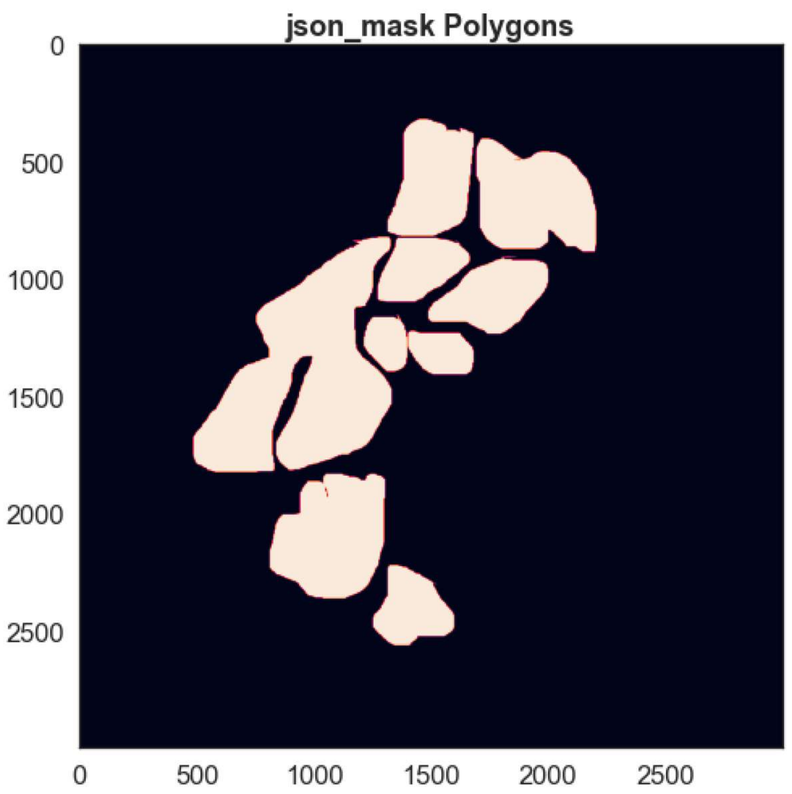
Organ	Avg_area	Avg_area_µm	Avg_height	Avg_width	Avg_shape_µm	Count
Kidney	106690.4	17070.4	456.5	303.5	(182.6, 121.4)	337
Largeintestine	50862.1	8137.9	357.0	143.5	(142.8, 57.4)	3103
Lung	119583.7	19133.3	299.5	140.5	(119.8, 56.2)	188
Prostate	193169.8	30907.1	421.5	433.5	(168.6, 173.4)	1097
Spleen	438419.0	70147.0	675.0	400.0	(270.0, 160.0)	167

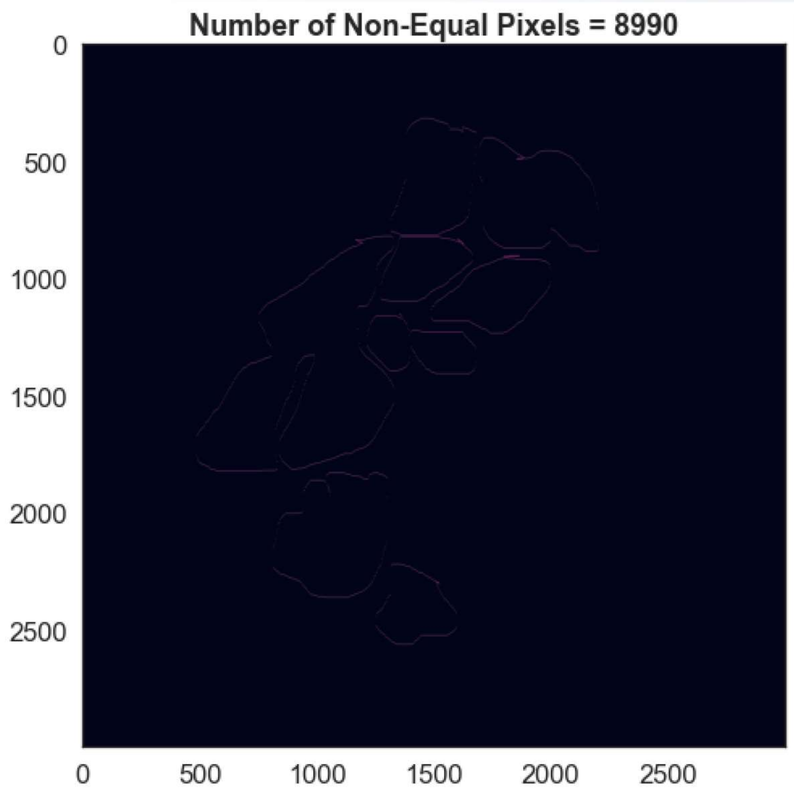
- **Is the presented characterization of FTUs useful for understanding individual differences?**
- **Is it possible to predict FTU area size distribution, given age and sex info across all organs?**

EDA

Difference between rle_maks and json_mask







Modeling

Baseline

Backbone : **Unet + EfficientnetB7**

Loss : **Dice_loss + Binary_crossentropy**

Image size : **3000**

Resize : **160**

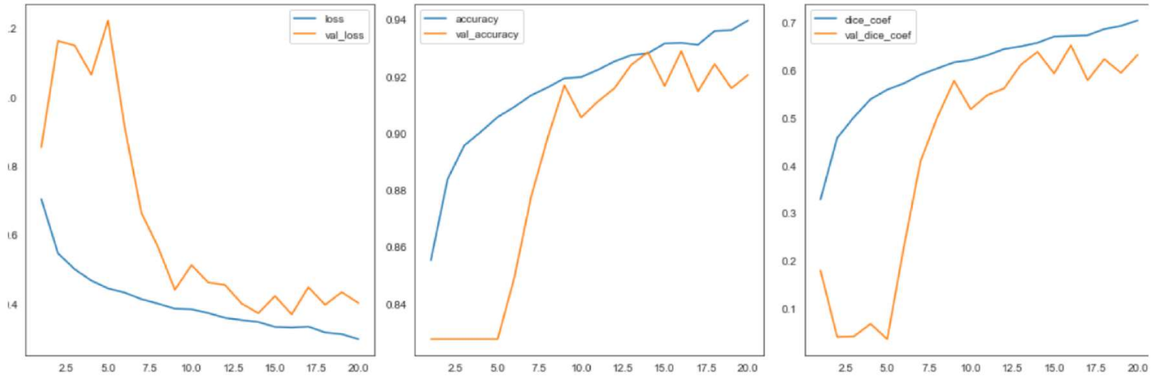
Number of Images :

- **prostate - 79 images**
- **spleen - 45 images**
- **lung - 40 images**
- **kidney - 84 images**
- **largeintestine - 49 images**

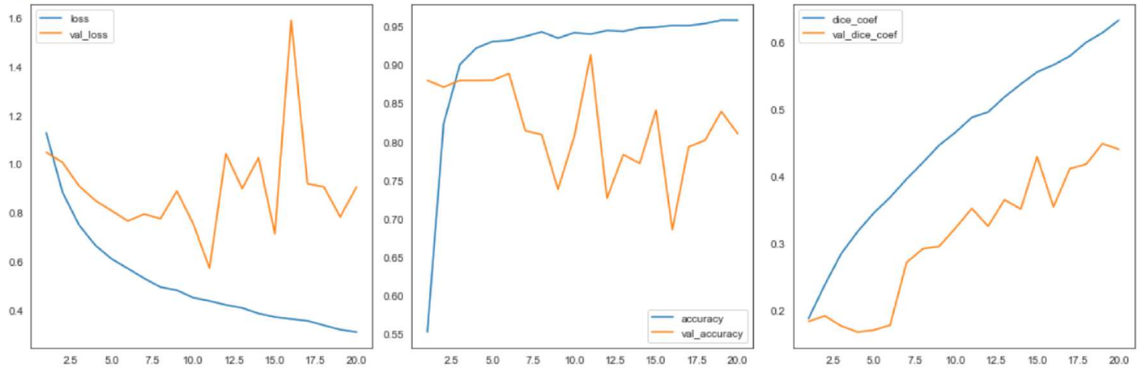
Augmentation : **Horizontal, Vertical Filp on all image**

Baseline result

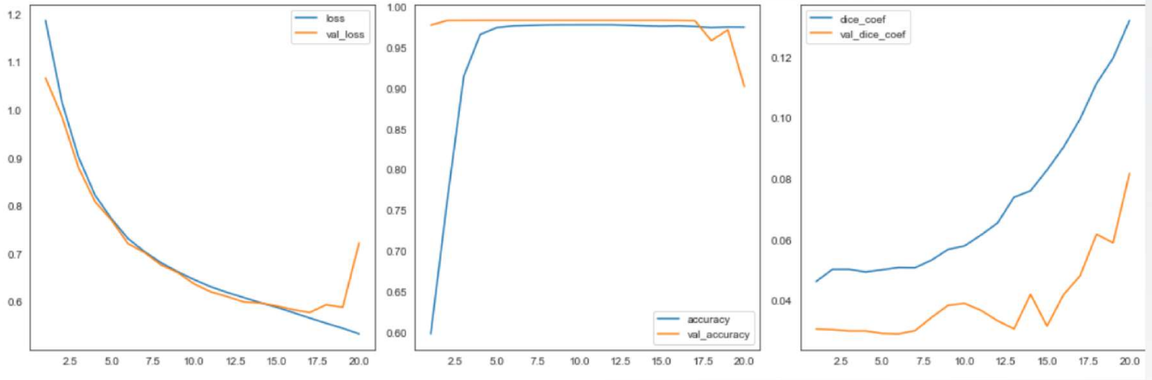
prostate



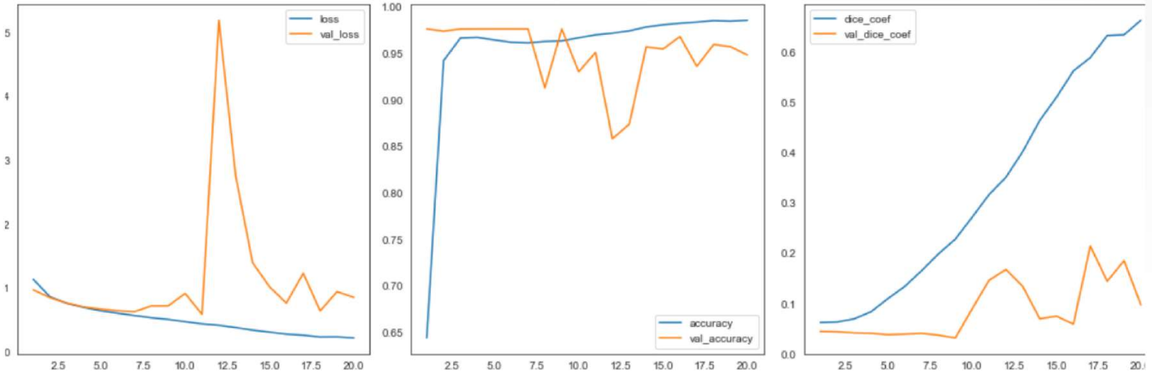
spleen



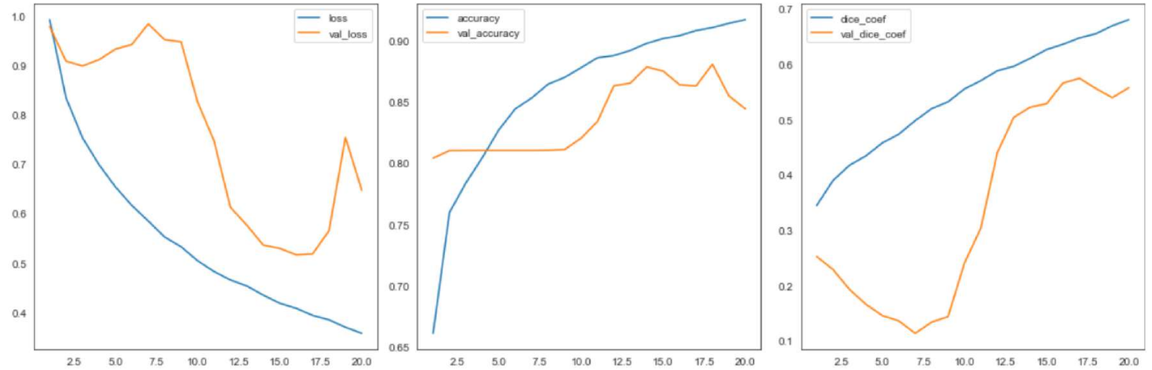
lung



kidney



largeintestine



Improvement strategy

Backbone : DeepLabV3+

Loss : Dice_loss + Weighted_BCE_loss (Each Organ)

Image size : 500 (3000 / 6)

Resize : 480

Number of Images :

- prostate - 79 * 34 images
- spleen - 45 * 34 images
- lung - 40 * 34 images
- kidney - 84 * 34 images
- largeintestine - 49 * 34 images

Augmentation : Horizontal, Vertical Filp on all image, color(stained tissue)

Improvement Strategy

Improvement strategy 1

Backbone : DeepLabV3+ -> Unet + EfficientnetB0

Loss : Dice_loss + Weighted_BCE_loss (Each Organ)

Image size : 500 (3000 / 6)

Resize : 480 -> 224

Number of Images :

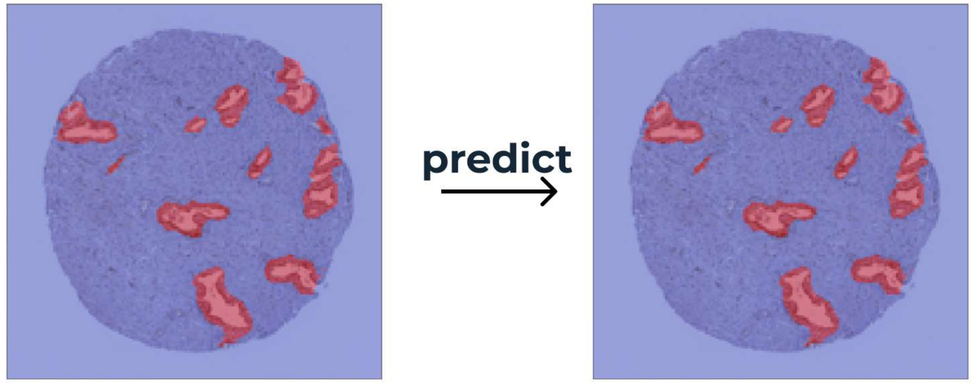
- prostate - 79 * 34 images
- spleen - 45 * 34 images
- lung - 40 * 34 images
- kidney - 84 * 34 images
- largeintestine - 49 * 34 images

Augmentation : Horizontal, Vertical Filp on all image, color(stained tissue)

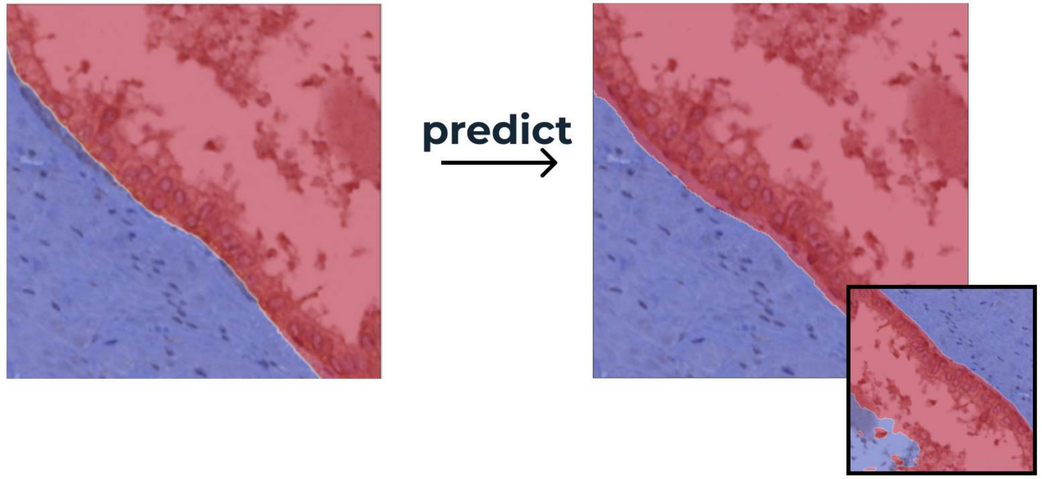
epochs : 60

Cropped image result

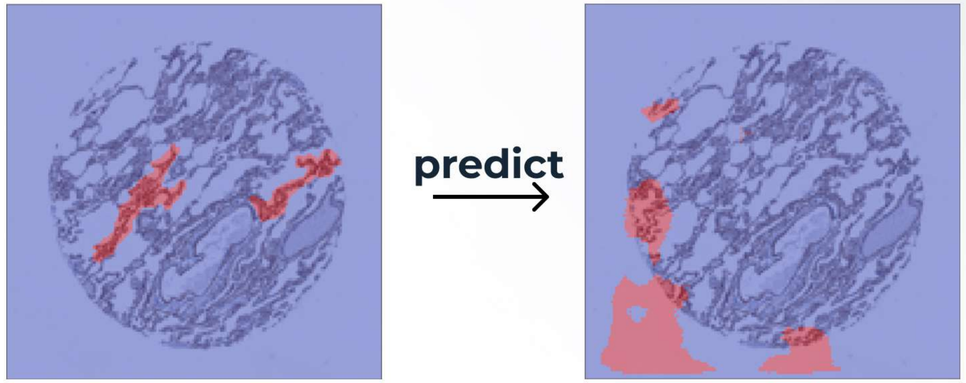
prostate_original



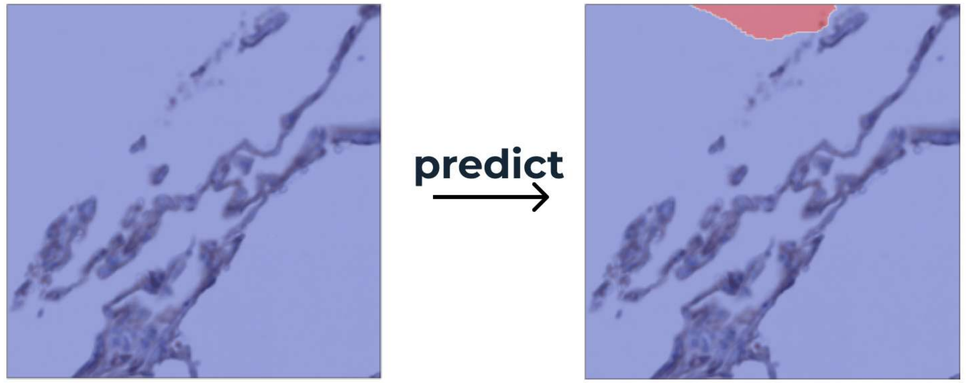
prostate_cropped



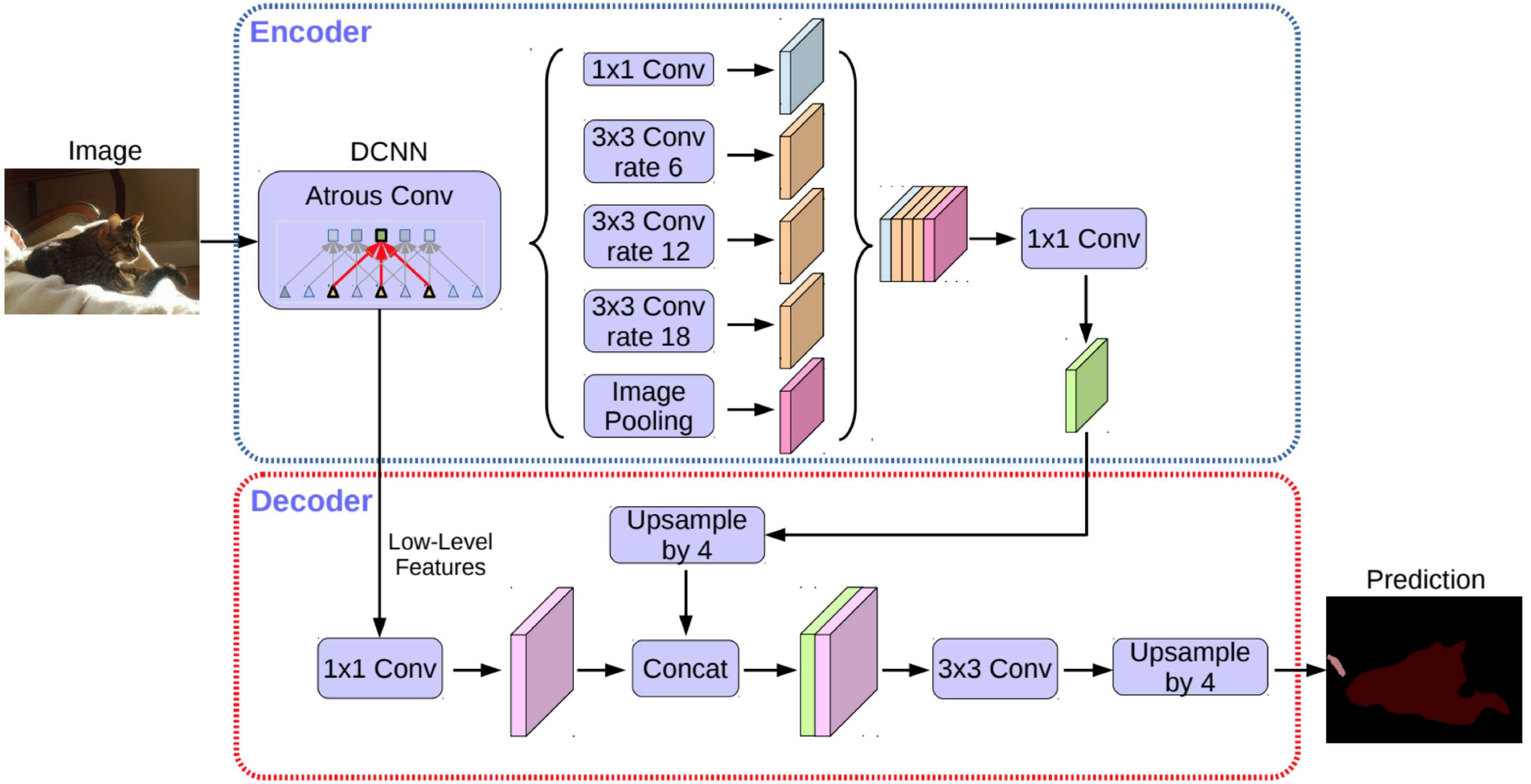
lung_original



lung_cropped



DeepLabV3


Encoder
Image
DCNN
Atrous Conv
1x1 Conv
3x3 Conv rate 6
3x3 Conv rate 12
3x3 Conv rate 18
Image Pooling
1x1 Conv
Decoder
Low-Level Features
Upsample by 4
1x1 Conv
Concat
3x3 Conv
Upsample by 4
Prediction