

HuBMAP + HPA - Hacking the Human Body Competition Progress Report 4

Presentation of a Report

By

Group C



Xi'an Jiaotong-Liverpool University

西交利物浦大學

Overview

- **Data Augmentation Strategy:**

1. Data augmentation strategies for some Top Ranking teams

- **Inference Trick:**

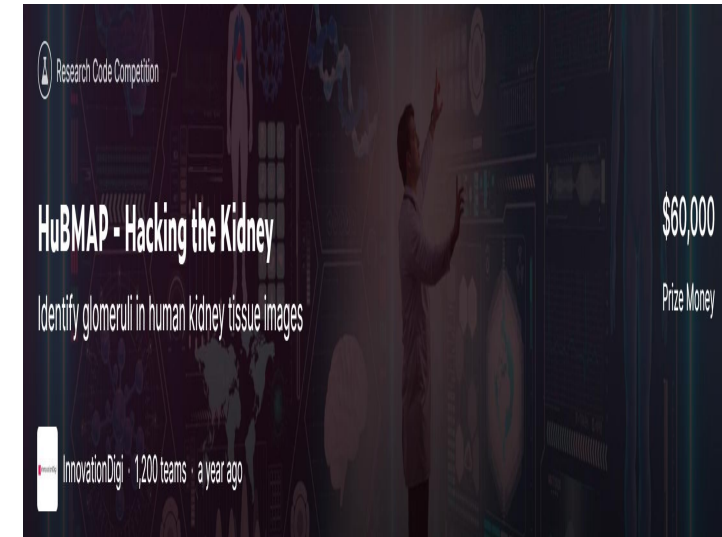
1. *Expansion tiles but only during inference*

- **Pre-training model:**

1. Pre-train the model on a dataset of similar competition (HuBMAP - Hacking the Kidney)

- **Future Plan:**

1. Continue to improve the UNext50 Model



Part 1: Data Augmentation Strategy

■ The 3rd Place's Strategy:

```
def get_aug(p=1.0):  
    return Compose([  
        HorizontalFlip(),  
        VerticalFlip(),  
        RandomRotate90(),  
        ShiftScaleRotate(shift_limit=0.0625, scale_limit=0.2, rotate_limit=15, p=0.9,  
                           border_mode=cv2.BORDER_REFLECT),  
        OneOf([  
            ElasticTransform(p=.3),  
            GaussianBlur(p=.3),  
            GaussNoise(p=.3),  
            OpticalDistortion(p=0.3),  
            GridDistortion(p=.1),  
            IAAPiecewiseAffine(p=0.3),  
        ], p=0.3),  
        OneOf([  
            HueSaturationValue(15,25,0),  
            CLAHE(clip_limit=2),  
            RandomBrightnessContrast(brightness_limit=0.3, contrast_limit=0.3),  
        ], p=0.3),  
    ], p=p)
```

[\[Inference\] - FastAI Baseline](#)
(version 14/16)

Succeeded 0.59

[\[Inference\] - FastAI Baseline](#)
(version 1/16)

Succeeded 0.56

11 days ago by [Juntuo Wang](#)

Notebook [\[Inference\] - FastAI Baseline](#) | Version 1

Part 1: Data Augmentation Strategy

■ The 17th Place's Strategy:

```
Horizontal/Vertical Flip (p=0.5)
RandomRotate90 (p=0.5)
Rotate (-40/+40, reflect101, p=0.5)
ShiftScaleRotate (scale: -0.2/0.1, p=0.5)
OneOf (p=0.5)
  RandomBrightnessContrast (B: 0.5, C: 0.1)
  HueSaturationValue (H: +-20, S: +-100, V: +-80)
One of (p=0.5)
  Cutout (holes: 100, size: 1/64, white RGB)
  GaussianNoise
One of (p=0.5)
  ElasticTransform
  GridDistortion (steps: 5, limit: 0.3)
  OpticalDistortion (distort: 0.5, shift: 0.0)
```

[Inference] - FastAI Baseline (version 15/16)	Succeeded	0.57
[Inference] - FastAI Baseline (version 1/16) 11 days ago by Juntuo Wang	Succeeded	0.56
Notebook [Inference] - FastAI Baseline Version 1		

Part 1: Data Augmentation Strategy

■ The Final Strategy:

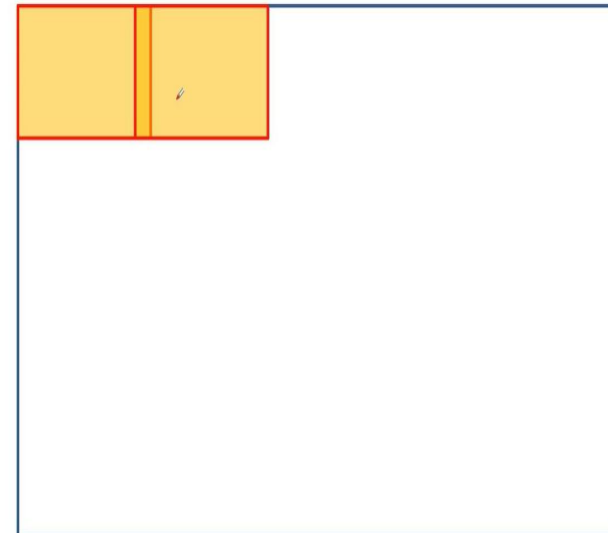
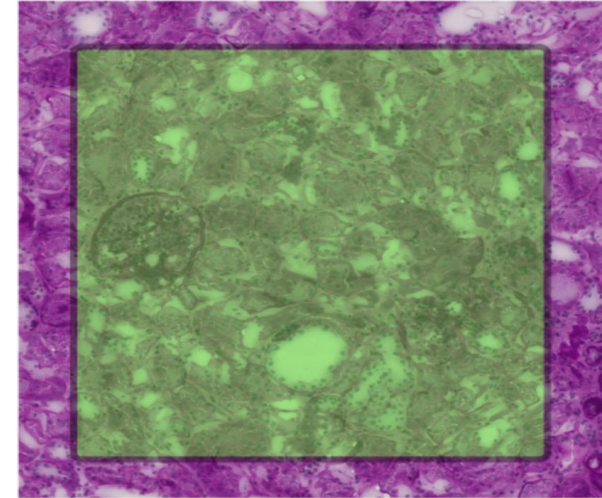
```
def get_aug(p=1.0):
    return Compose([
        HorizontalFlip(),
        VerticalFlip(),
        RandomRotate90(),
        Transpose(),
        ShiftScaleRotate(shift_limit=0.0625, scale_limit=0.2, rotate_limit=15, p=0.9,
                          border_mode=cv2.BORDER_REFLECT),
        OneOf([
            OpticalDistortion(p=0.3),
            GridDistortion(p=.1),
            GaussianBlur(p=.3),
            IAAPiecewiseAffine(p=0.3),
        ], p=0.3),
        OneOf([
            HueSaturationValue(10,15,10),
            CLAHE(clip_limit=2),
            RandomBrightnessContrast(),
            IAASharp(),
        ], p=0.3),
    ], p=p)
```

[Inference] - FastAI Baseline (version 4/16)	Succeeded	0.61
[Inference] - FastAI Baseline (version 1/16)	Succeeded	0.56
11 days ago by Juntuo Wang		
Notebook [Inference] - FastAI Baseline Version 1		

Part 2: Inference Trick:

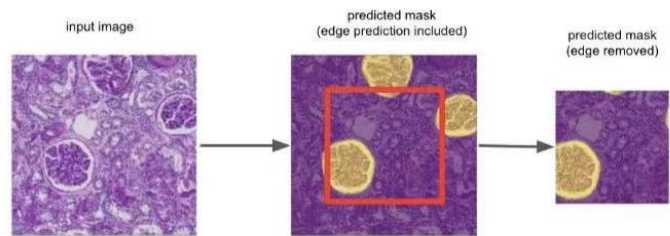
- **The idea from 3rd Place's solution:**
 - The basic idea is to eliminate edge effects. To achieve this, for each tile we run the model on, we expand the tile by a certain number of pixels in all 4 directions, but only use the center region's predictions which do not have edges.

Entire tile is run through the neural network, but only green region is used as inference output



Part 2: Inference Trick:

- Comparison of the Results:



The edge effect is eliminated by using the results in the middle of the predicted results

[\[Inference\] - FastAI Baseline](#)
(version 12/16)

2 days ago by [Juntuo Wang](#)

Notebook [Inference] - FastAI Baseline | Version 12

[\[Inference\] - FastAI Baseline](#)
(version 4/16)

6 days ago by [Juntuo Wang](#)

Succeeded

0.63



Succeeded

0.61



Part 3: Pre-Training Model :



- **Preprocessing the Dataset (Convert it to 256×256)**
- **Pre-train the Model**
- **Loading the Pre-trained Model and Continue to train the model on our own competition's dataset**

Part 3: Pre-Training Model :

Research Code Competition

HuBMAP - Hacking the Kidney

Identify glomeruli in human kidney tissue images

\$60,000

Prize Money

InnovationDigi · 1,200 teams · a year ago

■ Comparison of the Results:

[Inference] - FastAI Baseline (version 13/16) 2 days ago by Juntuo Wang Notebook [Inference] - FastAI Baseline Version 13		Succeeded	0.66	<input type="checkbox"/>	
[Inference] - FastAI Baseline (version 12/16) 2 days ago by Juntuo Wang Notebook [Inference] - FastAI Baseline Version 12		Succeeded	0.63	<input type="checkbox"/>	
17	Juntuo Wang 		0.66	13	17m



Your Best Entry!
Your most recent submission scored 0.66, which is an improvement of your previous score of 0.63. Great job!

Tweet this

Part 4: Further Improvement

- Continue to try to improve the UNext50 Model
- Try to Use Some New Model Architecture