

Simple can achieve 0.8

I hate mmsegmentation

남혜린, 주세환

INDEX

Introduction

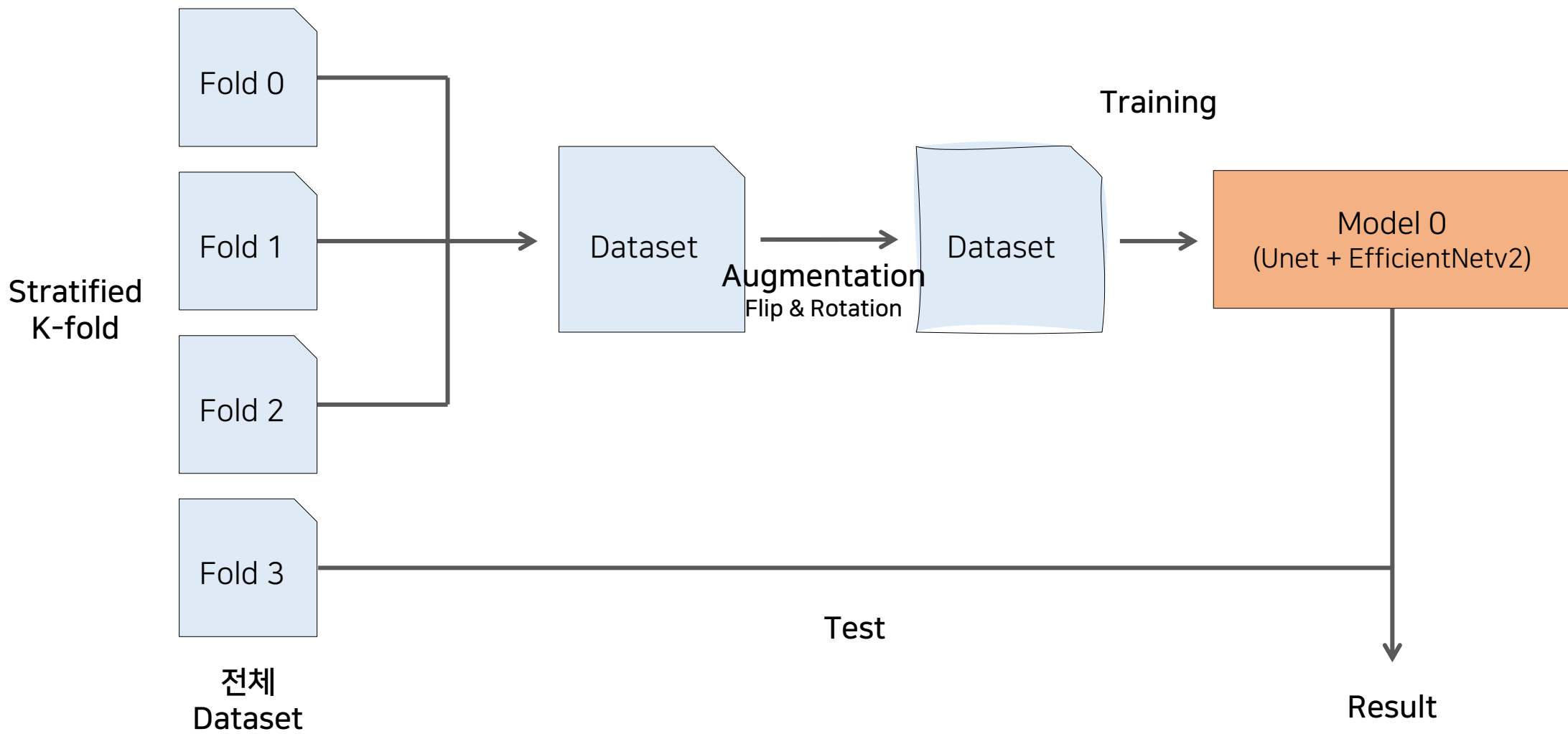
Methods

Results

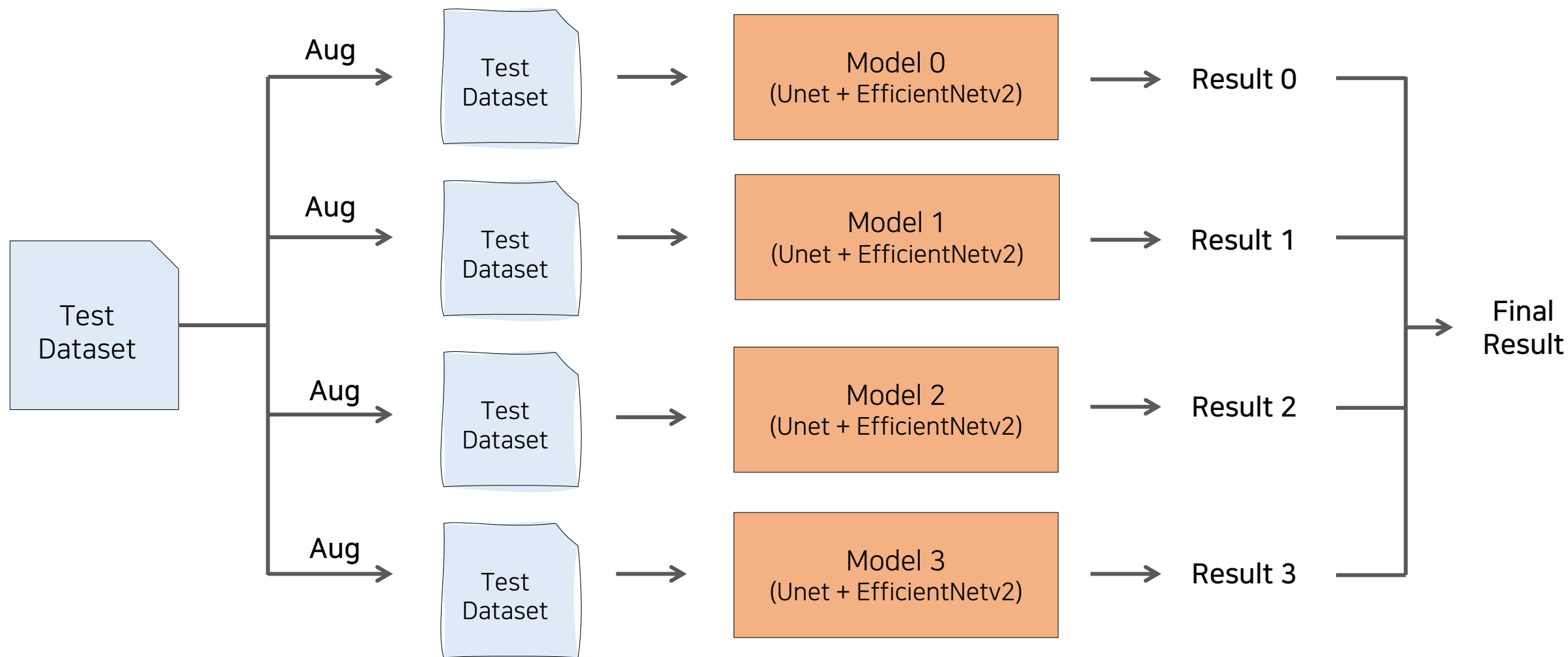
ETC

Introduction

Pipeline – Fold Training



Pipeline – Ensemble



Methods

1) 전처리 - DICOM to PNG

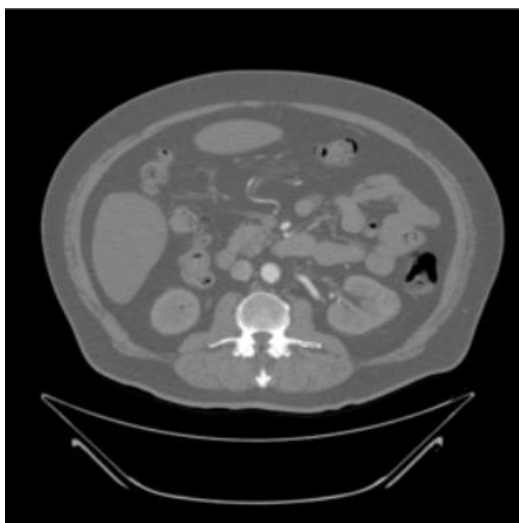
- DICOM file의 HU array를 불러와 -1024 ~ 3096 의 범위를 image로 저장하기 위해서 0 ~ 255 단위로 변환하여 PNG 로 저장

1) 전처리 - Stratified K fold

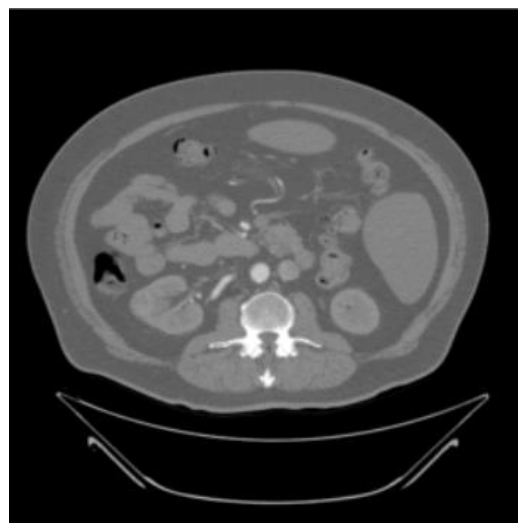
- Case 별로 비슷한 image가 들어있는 것을 보고 비슷한 image가 한 fold내에서 학습이 되지 않도록 case 단위로 4 fold로 split (k fold)
- 이 때 각 case 별로 존재하는 tumor와 kidney area를 균등하게 나눔 (Stratified)

1) 전처리 - Augmentation

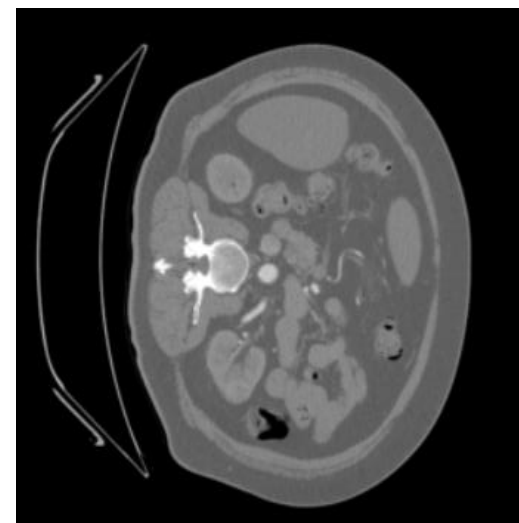
- Horizontal Flip과 RandomRotation을 동시에 사용할 때 성능이 가장 좋았음



Origin Data



Horizontal Flip



Rotation

2) 모델

- Encoder로는 EfficientNet v2를 Decoder로는 Unet을 사용함

Table 4. EfficientNetV2-S architecture – MBConv and Fused-MBConv blocks are described in Figure 2.

| Stage | Operator | Stride | #Channels | #Layers |
|-------|------------------------|--------|-----------|---------|
| 0 | Conv3x3 | 2 | 24 | 1 |
| 1 | Fused-MBConv1, k3x3 | 1 | 24 | 2 |
| 2 | Fused-MBConv4, k3x3 | 2 | 48 | 4 |
| 3 | Fused-MBConv4, k3x3 | 2 | 64 | 4 |
| 4 | MBConv4, k3x3, SE0.25 | 2 | 128 | 6 |
| 5 | MBConv6, k3x3, SE0.25 | 1 | 160 | 9 |
| 6 | MBConv6, k3x3, SE0.25 | 2 | 256 | 15 |
| 7 | Conv1x1 & Pooling & FC | - | 1280 | 1 |

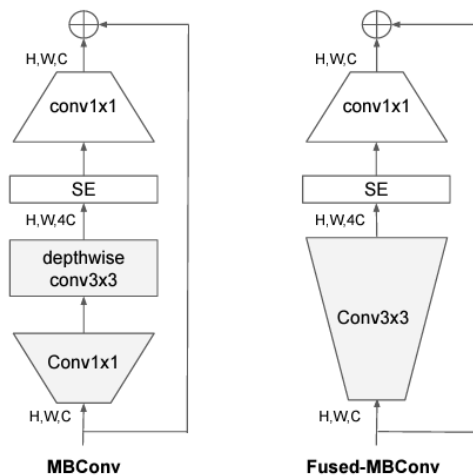
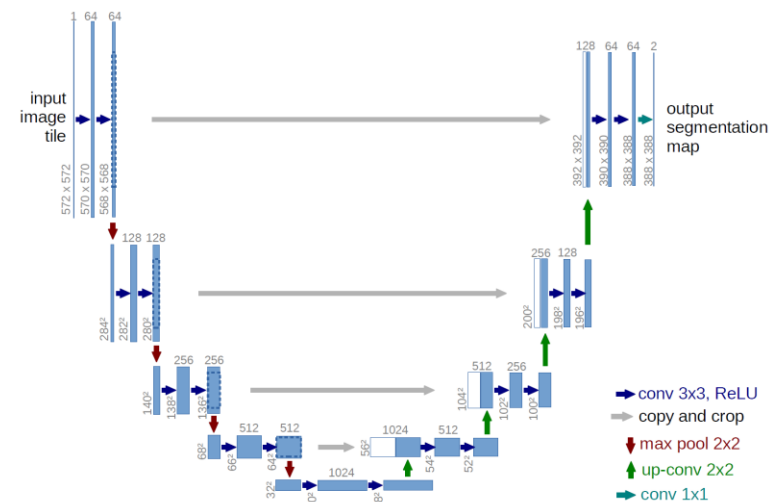


Figure 2. Structure of MBConv and Fused-MBConv.



<Encoder – EfficientNet v2>

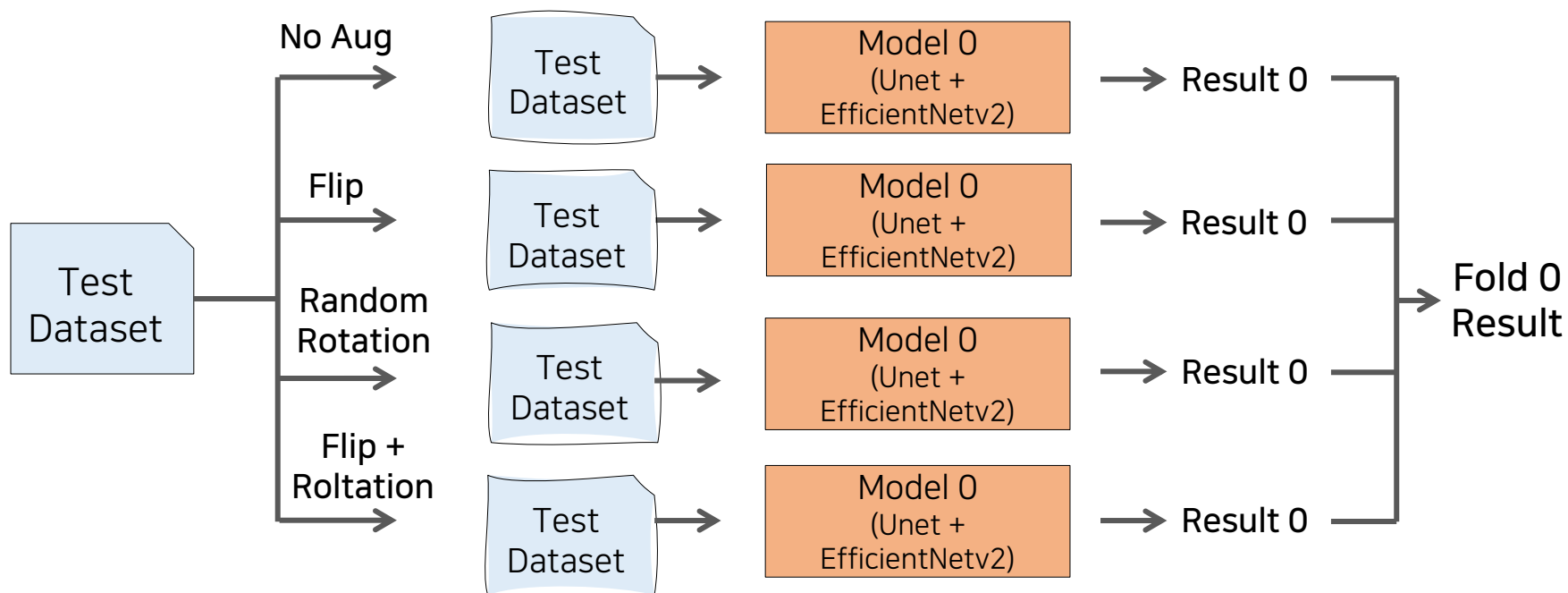
<Decoder – Unet>

2) 모델 - 최적화

- Loss function
 - $0.75 * \text{Focal} + 0.25 * \text{Dice}$
- Optimizer
 - Adam
- Lr scheduler
 - Step [30epoch, 40epoch]
- Hyperparameter
 - 50 epoch
 - lr – 0.001

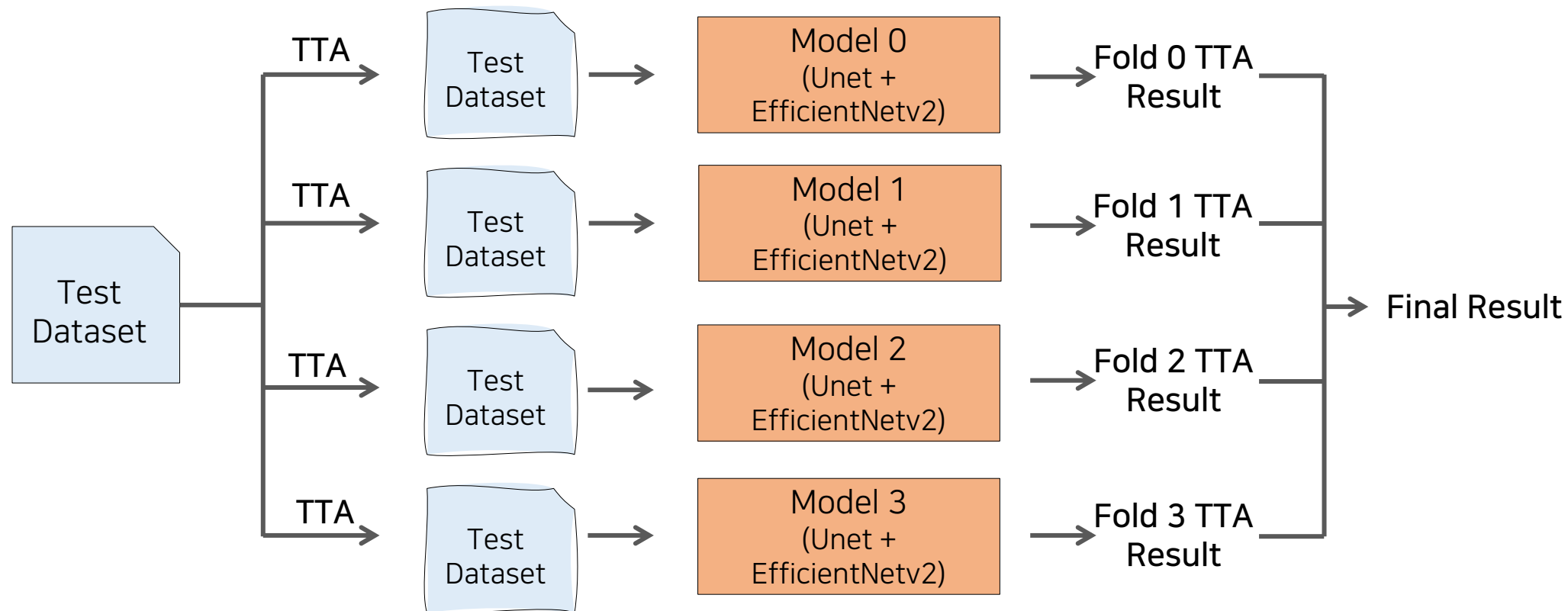
3) 후처리 - TTA

- TTA
 - 그림과 같이 TTA4를 적용
 - No Aug, Flip, Random Rotation, Flip + Random Rotation 을 적용한 image를 동일한 Model에 적용
 - 4개의 output을 Ensemble



3) 후처리 - Fold Ensemble

- TTA ensemble을 각 Model 0 ~ 3 까지 적용
- 각 Model 별 TTA 결과를 Fold Ensemble 하여 최종 output 추출



4) CV Metric

- 리더 보드 산출 방식이 case 단위로 DICE를 계산하여 동일하게 case 단위로 DICE 를 계산
- 그 결과 리더보드와 CV score가 유사하여 실험을 신뢰하고 할 수 있었음.

Results

Model search

| Backbone | 평균 dice | Kidney dice | Tumor dice |
|------------------------|--------------|--------------|--------------|
| Resnet34 | 0.477 | 0.888 | 0.065 |
| Efficientnet B0 | 0.538 | 0.854 | 0.222 |
| Efficientnet V2 | 0.726 | 0.941 | 0.512 |

| Decoder | 평균 dice | Kidney dice | Tumor dice |
|-------------|--------------|--------------|--------------|
| PAN | 0.626 | 0.915 | 0.337 |
| PSP | 0.642 | 0.926 | 0.358 |
| Unet | 0.726 | 0.941 | 0.512 |

Loss search

| Loss | 평균 dice | Kidney dice | Tumor dice |
|------------------------|---------|-------------|------------|
| BCE | 0.669 | 0.929 | 0.409 |
| 0.75 BCE + 0.25 DICE | 0.723 | 0.934 | 0.512 |
| Focal | 0.658 | 0.919 | 0.396 |
| Lovasz | 0.548 | 0.927 | 0.169 |
| Focal 0.75 + 0.25 DICE | 0.726 | 0.930 | 0.521 |

Augmentation Test

| Augmentation | 평균 dice | Kidney dice | Tumor dice |
|---------------------|---------|-------------|------------|
| Base | 0.713 | 0.929 | 0.496 |
| HU Normalization | 0.706 | 0.936 | 0.475 |
| Flip & Rotate | 0.801 | 0.943 | 0.658 |

Final Result

| | Score |
|----------------------|---------|
| Fold 0 | 0.77562 |
| Fold 0 + TTA 4 | 0.78747 |
| Fold Ensemble + TTA4 | 0.80098 |

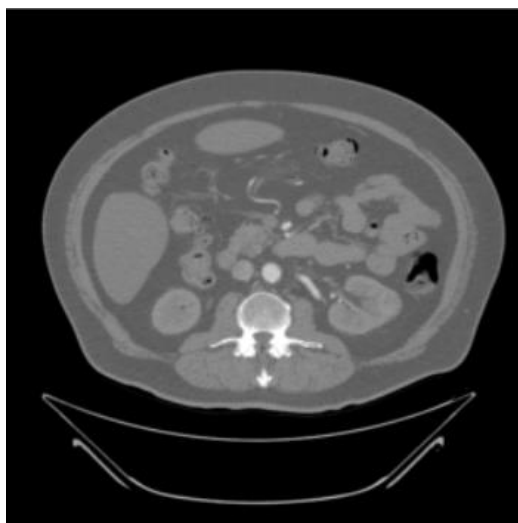
ETC

HU Preprocessing

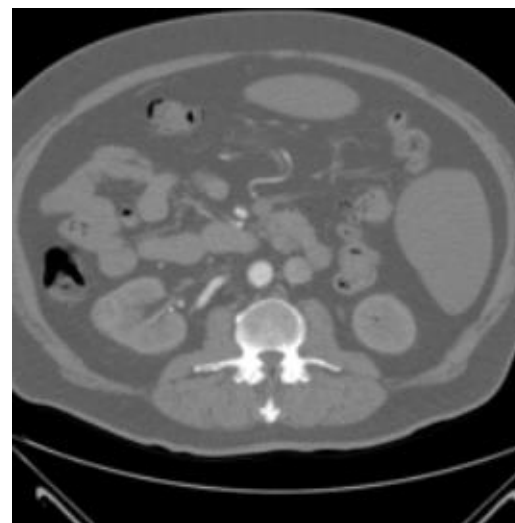
- Kidney와 tumor가 존재하는 값은 0~100 사이인 사실에 근거하여 HU 값을 -350 ~ 350으로 조정 (주변 background 고려)
- -350이하는 전부 -350으로 통일, 350 이상은 전부 350으로 통일
- 실험 결과 HU prerprocessing을 안했을 때와 비슷하거나 오히려 성능이 하락

CROP

- Kidney와 tumor의 주변부만 background 포함하여 crop하여 Model을 학습
- Model의 속도는 빨라졌으나 오히려 성능 하락



Origin Data



384 x 384
Center Crop

Swin Transformer

- CNN이 아닌 self-attention 을 이용한 Swin Transformer를 backbone으로 사용하여 실험
- 처음에는 score가 CNN 계열의 backbone 보다 높았으나, 다양한 기법을 적용하였을 때 CNN 계열의 backbone이 성능향상 폭이 컸음

Swin Transformer Decoder

| Decoder | 평균 dice | Kidney dice | Tumor dice | Public LB |
|----------------|---------|-------------|------------|-----------|
| DeeplabV3 | 0.85385 | 0.943 | 0.7647 | 0.6857 |
| DeeplabV3 plus | 0.87035 | 0.945 | 0.7957 | 0.71878 |
| Upernet | 0.9013 | 0.951 | 0.8516 | 0.74322 |

- Swin transformer는 image 단위로 dice 계산, CNN 계열의 backbone case 단위로 dice 계산

Loss finder

| Loss | 평균 dice | Kidney dice | Tumor dice |
|-------------------------------|---------------|---------------|---------------|
| BCE 0.9 + DICE 0.1 | 0.8249 | 0.8778 | 0.7621 |
| BCE 0.8 + DICE 0.1 | 0.8232 | 0.8889 | 0.7576 |
| BCE 0.7 + DICE 0.1 | 0.8116 | 0.8708 | 0.7524 |
| Lovasz | 0.8313 | 0.8901 | 0.7725 |
| Focal 0.75 + DICE 0.25 | 0.8476 | 0.9102 | 0.7850 |

- Swin transformer는 image 단위로 dice 계산, CNN 계열의 backbone case 단위로 dice 계산

Fold Training

- SwinTransformer-base Upernet

| | 평균 dice | Kidney dice | Tumor dice |
|--------|---------|-------------|------------|
| Fold 0 | 0.8249 | 0.9591 | 0.8341 |
| Fold 1 | 0.8232 | 0.8889 | 0.7576 |
| Fold 2 | 0.8116 | 0.8708 | 0.7524 |
| Fold 3 | 0.8313 | 0.8901 | 0.7725 |

Final Fold Ensemble Score: 0.76075

- Swin transformer는 image 단위로 dice 계산, CNN 계열의 backbone case 단위로 dice 계산

Thank You