

# DLCV Hw1 Report

R10942198 林仲偉

# 1-1

1. Draw the network architecture of method A or B

The network architecture of method B [Fig.1] consists of pre-trained Resnet50 and 1 layer MLP with input channel size = 1000.

2. Report accuracy of your model (A and B) on validation set.

Model A: 0.6376

Model B: **0.8928**

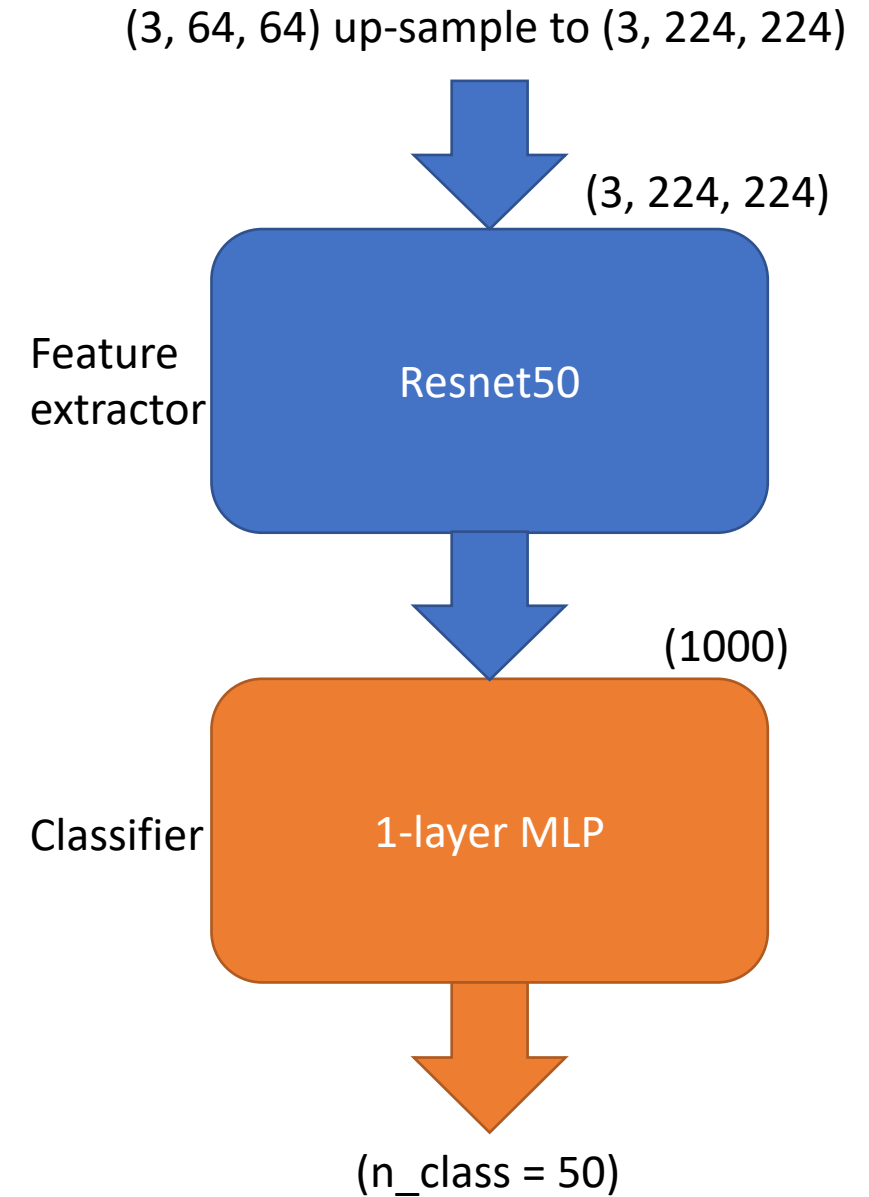


Fig.1. Network architecture of method B

# 1-1

## 3. Report your implementation details of model A.

- Optimizer: Adam
- Loss: CE loss
- Cross validation: 5-fold cross validation ensembling
- Learning rate: 5e-5
- Learning rate scheduled: StepLR
- Epoch: 50

## 4. Report your alternative model or method in B, and describe its difference from model A.

- Optimizer: Adam
- Loss: CE loss
- Cross validation: 5-fold cross validation ensembling
- Learning rate: 5e-5
- Learning rate scheduled: StepLR
- Epoch: 50
- Difference: model architecture
  - Use pre-train Resnet50 instead of regular CNN: adding residual during forwarding.

# 1-1

5. Visualize the learned visual representations of model A on the validation set by implementing PCA (Principal Component Analysis) on the output of the second last layer. Briefly explain your result of the PCA visualization.

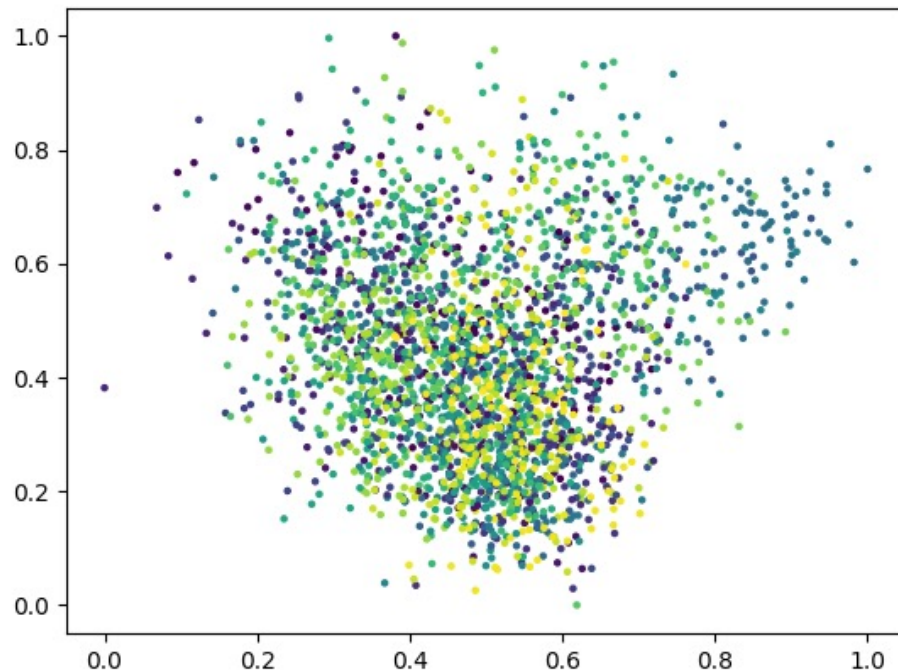


Fig.2. PCA(n\_component=2)

Explanation:

The clustering result is not very good since PCA is a linear dimensionality reducing algorithm. Moreover, the performance of model A also influence the result of PCA.

# 1-1

6. Visualize the learned visual representation of model A, again on the output of the second last layer, but using t-SNE (t-distributed Stochastic Neighbor Embedding) instead. Depict your visualization from three different epochs including the first one and the last one. Briefly explain the above results.

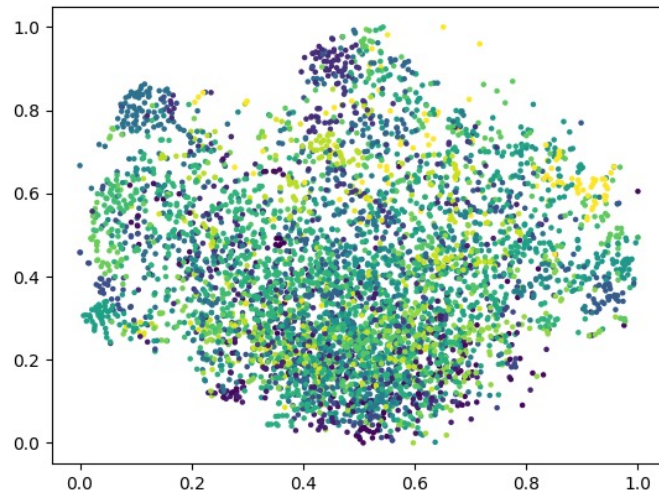


Fig.3(a). First-epoch t-SNE

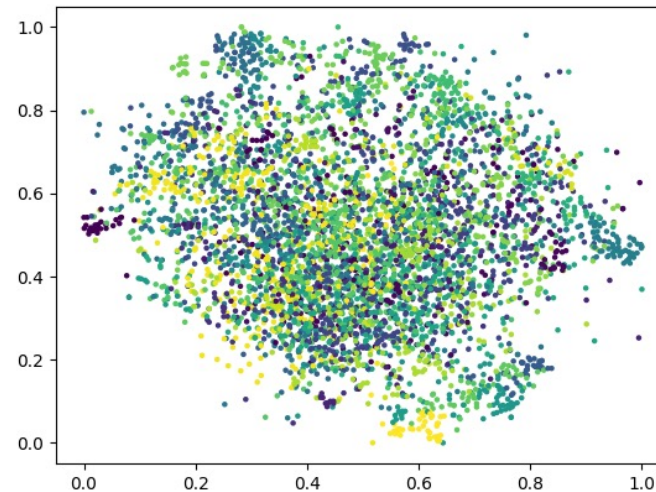


Fig.3(b). 5-epoch t-SNE

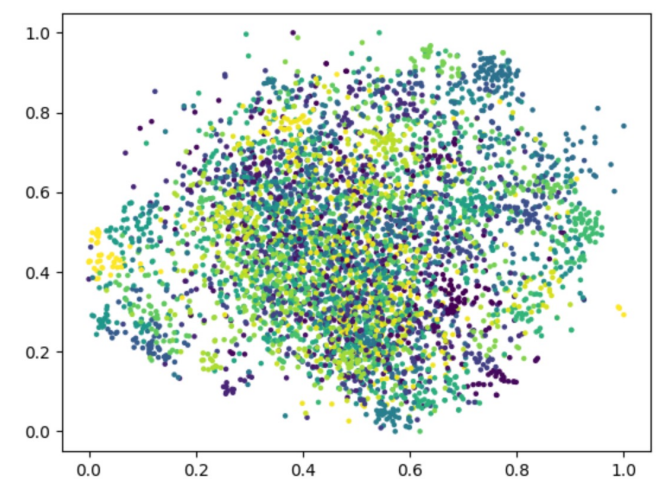


Fig.3(c). Last-epoch t-SNE

Explanation:

The clustering result is slightly better as the number of epoch increases.

However, since model A doesn't have good performance (in comparison to model B), the clustering result is not significantly greater than PCA.

# 1-2

1. Draw the network architecture of your VGG16-FCN32s model (model A).

Ref: <http://deanhhan.com/2018/07/26/vgg16/>

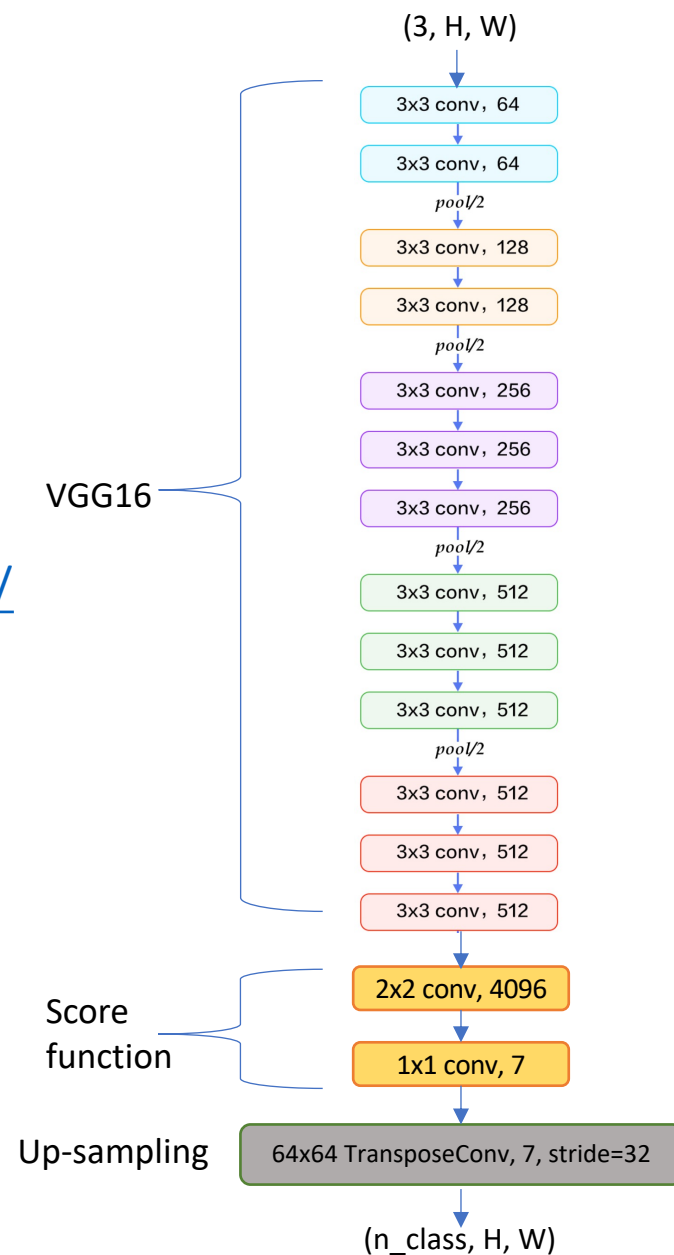
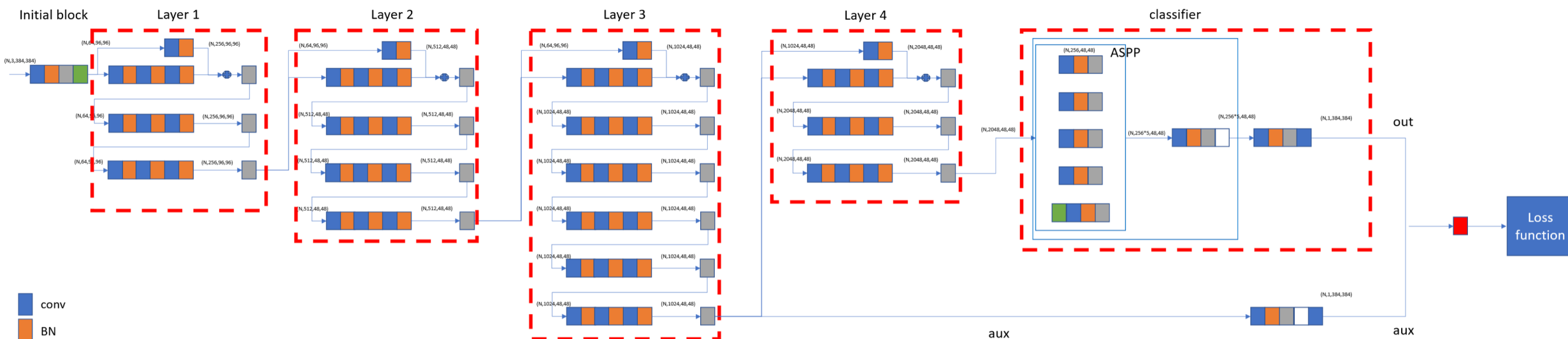


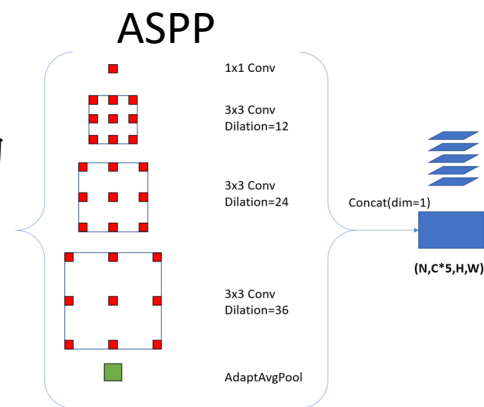
Fig.4(a). Network architecture of model A

# 1-2

2. Draw the network architecture of the improved model (model B: Deeplabv3\_resnet50) and explain it differs from your VGG16-FCN32s model.



Deeplabv3\_resnet50 多了一個 Atrous Spatial Pyramid Pooling (ASPP) module, 由一系列dilated convolution組成, 使其擁有更好的 up sampling 結果。



# 1-2

3. Report mIoUs of two models on the validation set.

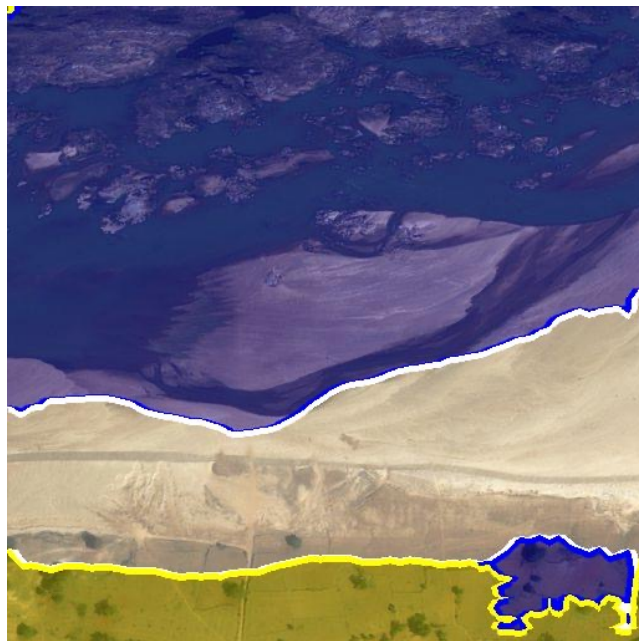
- Model A (VGG16+FCN32s): mIoU = 0.411476
- Model B (DeepLabV3\_Resnet50): mIoU= **0.747531**



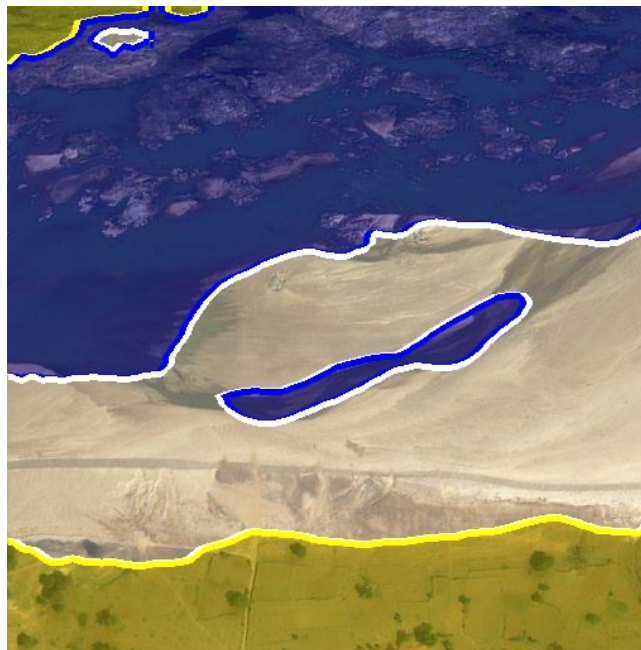
# 1-2

4. Show the predicted segmentation mask of “validation/0013\_sat.jpg”, “validation/0062\_sat.jpg”, “validation/0104\_sat.jpg” during the early, middle, and the final stage during the training process of the improved model.

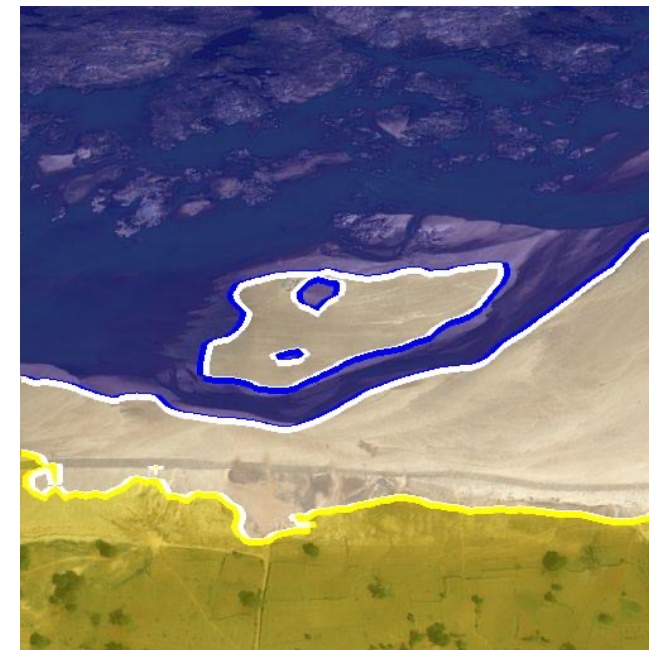
- validation/0013\_sat.jpg:



Early (epoch=0)



Middle (epoch=0)



final

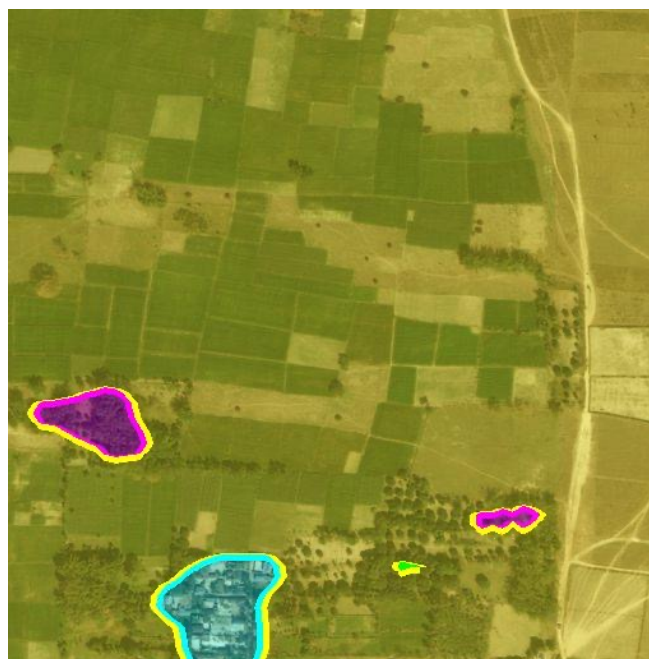
# 1-2

4. Show the predicted segmentation mask of “validation/0013\_sat.jpg”, “validation/0062\_sat.jpg”, “validation/0104\_sat.jpg” during the early, middle, and the final stage during the training process of the improved model.

- validation/0062\_sat.jpg:



Early (epoch=0)



Middle (epoch=5)

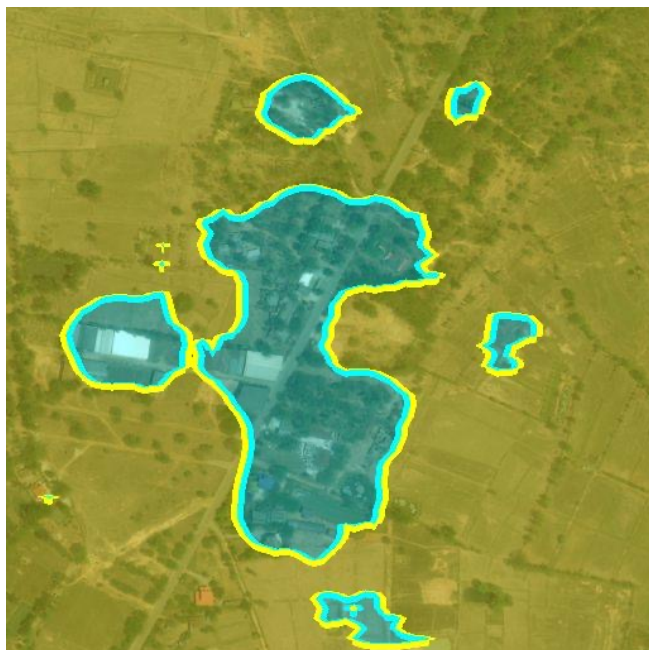


final

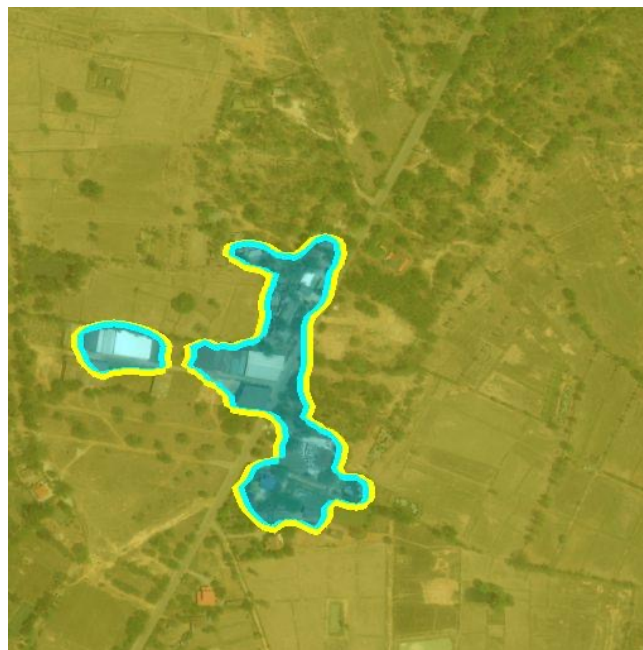
# 1-2

4. Show the predicted segmentation mask of “validation/0013\_sat.jpg”, “validation/0062\_sat.jpg”, “validation/0104\_sat.jpg” during the early, middle, and the final stage during the training process of the improved model.

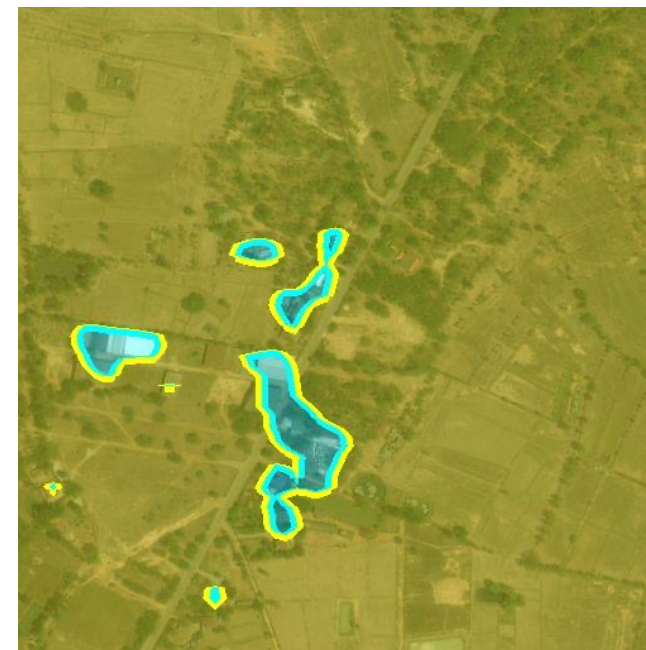
- validation/0104\_sat.jpg:



Early (epoch=0)



Middle (epoch=5)



final