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

```
(DLCV-hw1) zhongwei@bl530:/mnt/sda/Ben1211/DLCV Labs/hw1-benlin1211$ bash ./ev
al_1-1.sh "./result_1-1/val_A.csv" "./hw1_data/hw1_data/p1_data/val_gt.csv"
Acc: 0.6376
```

Model B: **0.8928**

```
(DLCV-hw1) zhongwei@bl530:/mnt/sda/Ben1211/DLCV Labs/hw1-benlin1211$ bash ./ev
al_1-1.sh "./result_1-1/val_B.csv" "./hw1_data/hw1_data/p1_data/val_gt.csv"
Acc: 0.8928
```

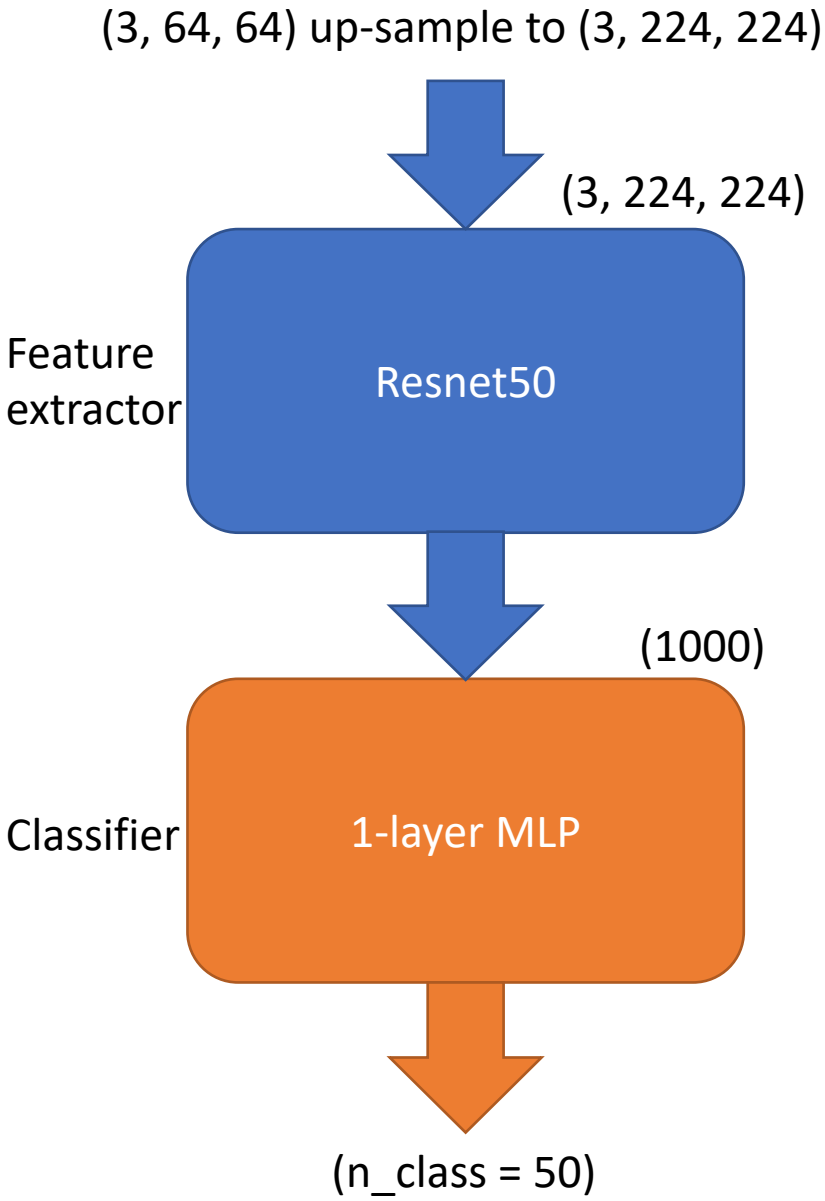


Fig.1. Network architecture of method B

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.

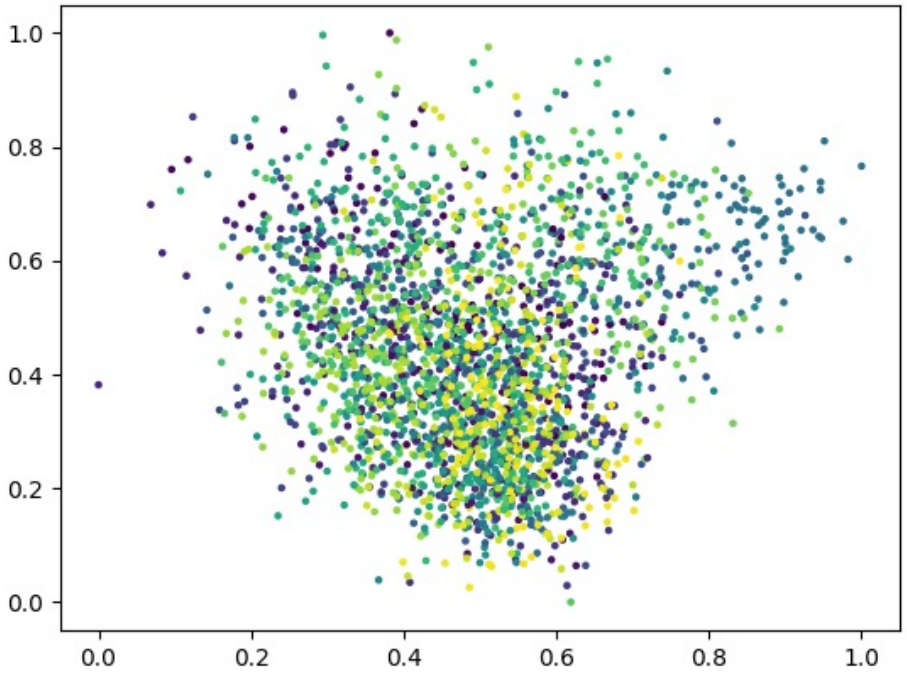


Fig.2. PCA(n_component=2)

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.

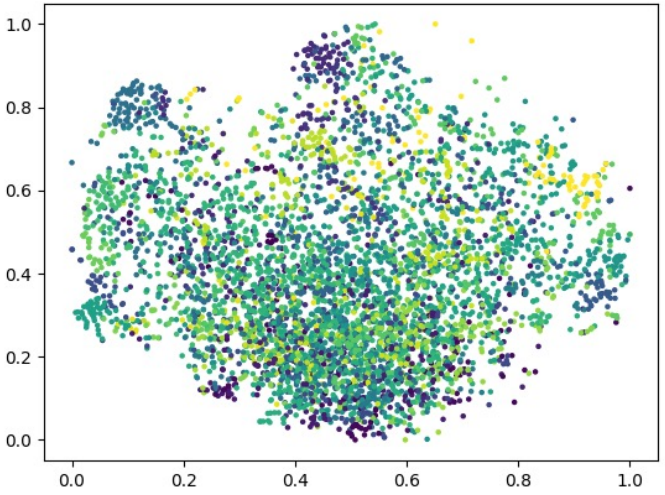


Fig.3(a). First-epoch t-SNE

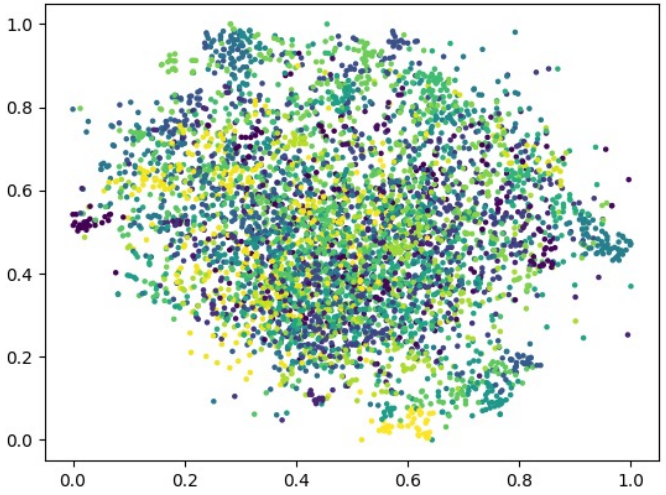


Fig.3(b). 5-epoch t-SNE

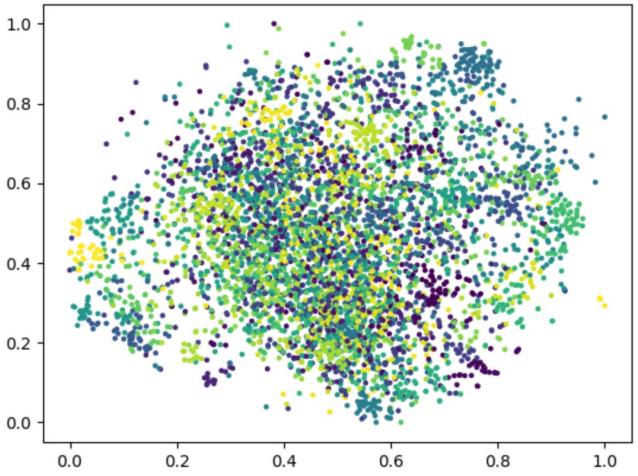


Fig.3(c). Last-epoch t-SNE

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://deanhan.com/2018/07/26/vgg16/

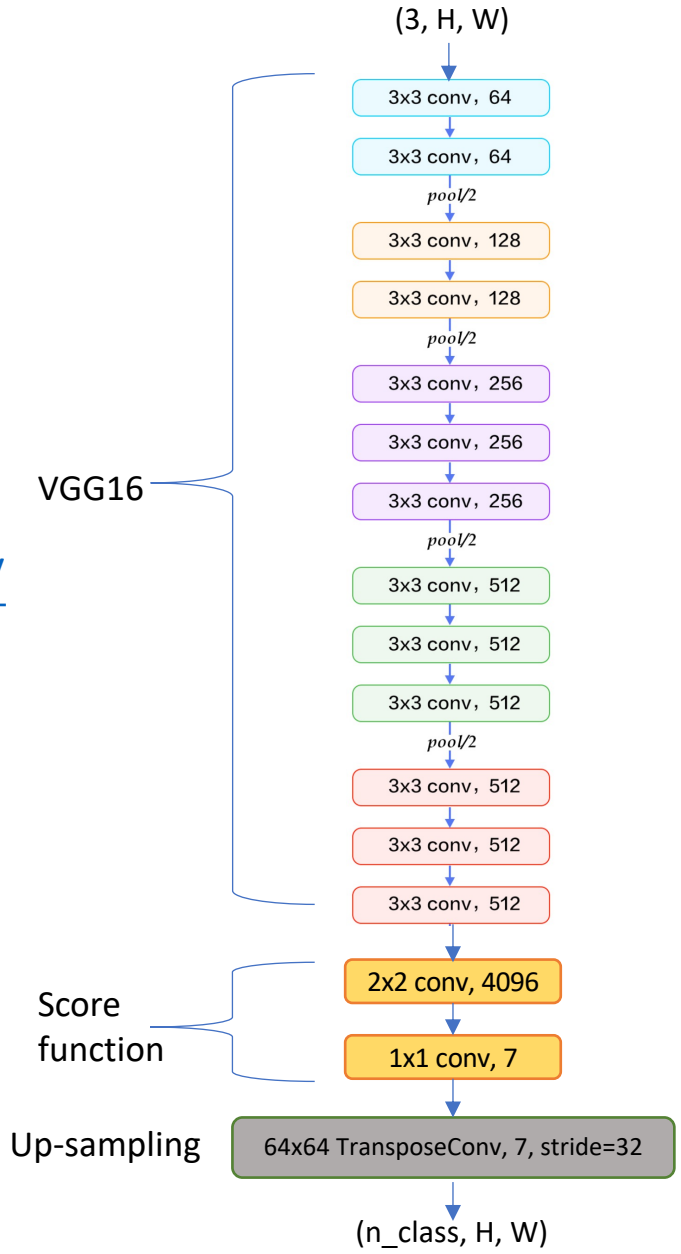
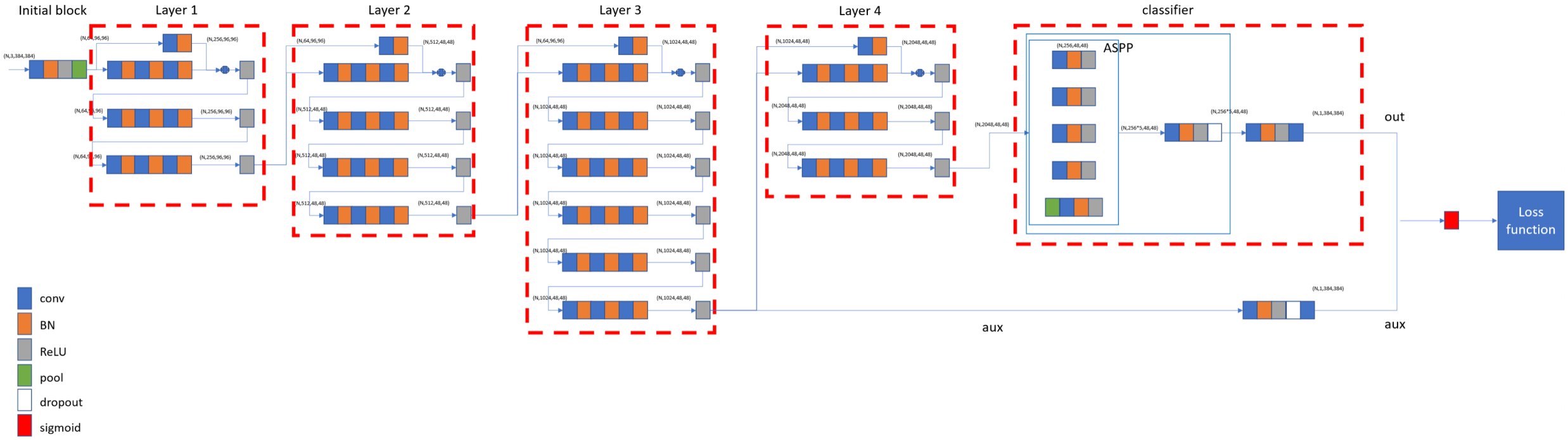


Fig.4(a). Network architecture of model A

1-2

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



Ref: https://blog.csdn.net/weixin_44816589/article/details/115266935

1-2

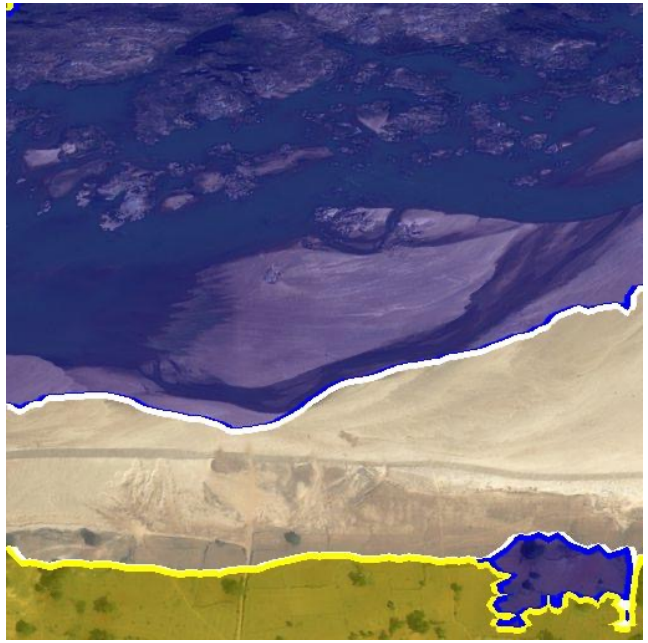
3. Report mIoUs of two models on the validation set.

- Model A (VGG16+FCN32s): mIoU = 0.411476
- Model B (DeepLabV3_Resnet50): mIoU= 0.725386

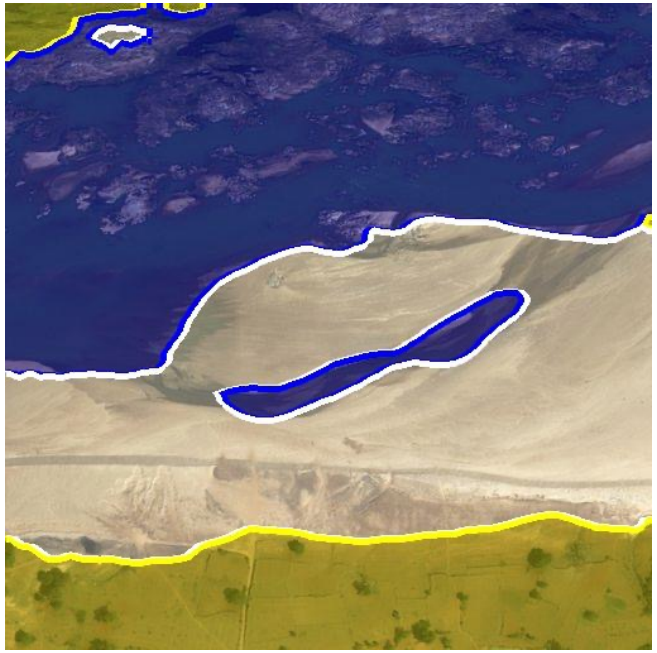
1-2

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)



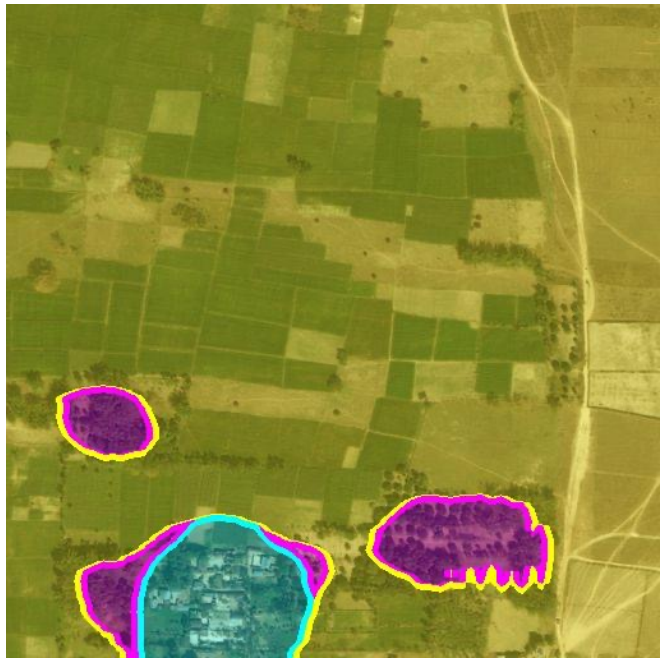
Middle (epoch=0)

final

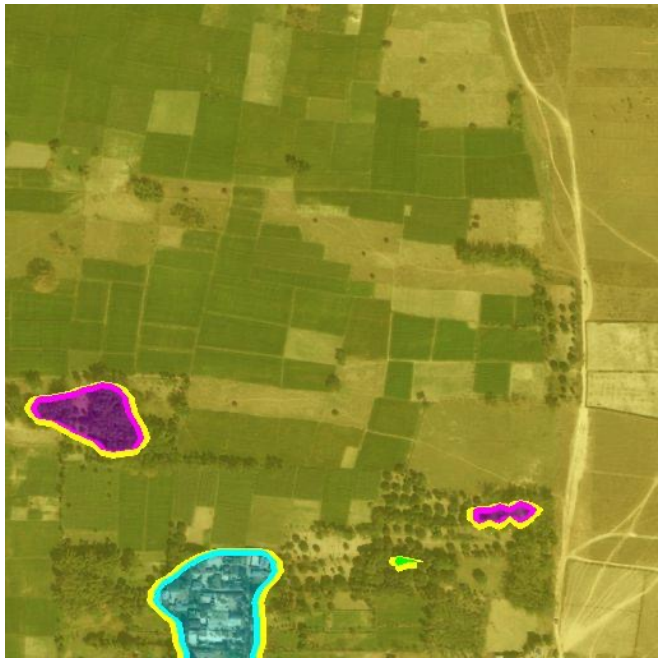
1-2

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)



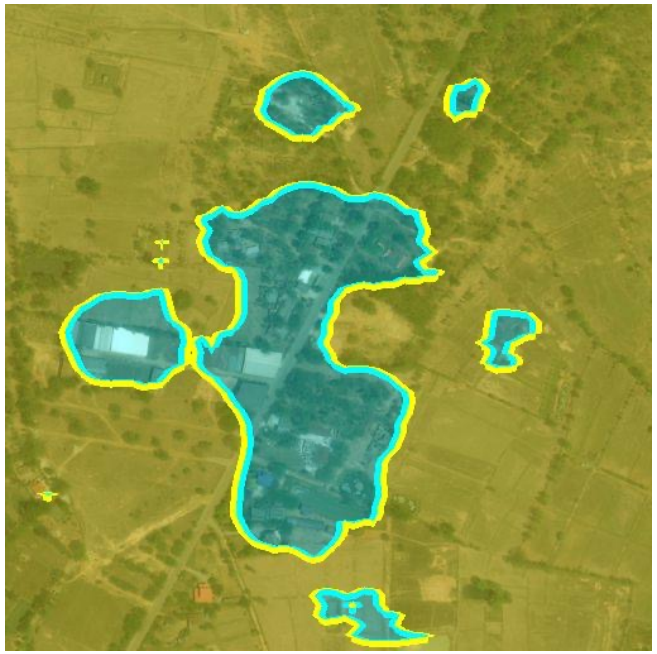
Middle (epoch=5)

final

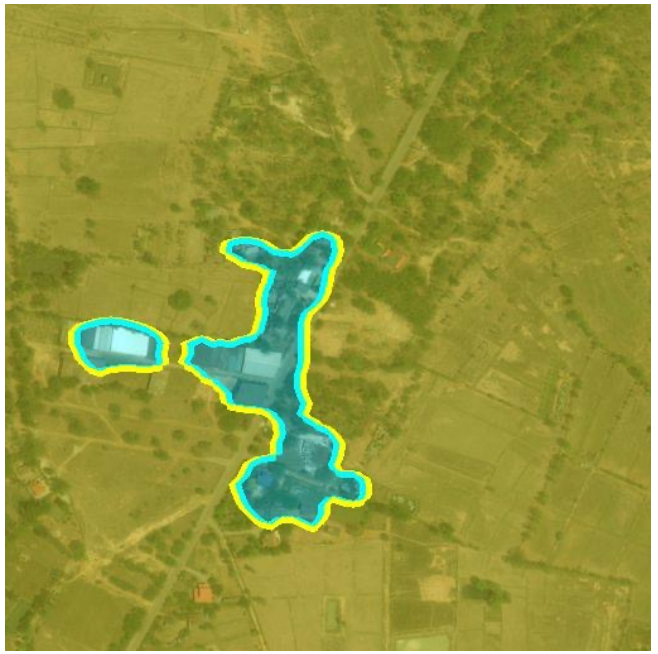
1-2

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.

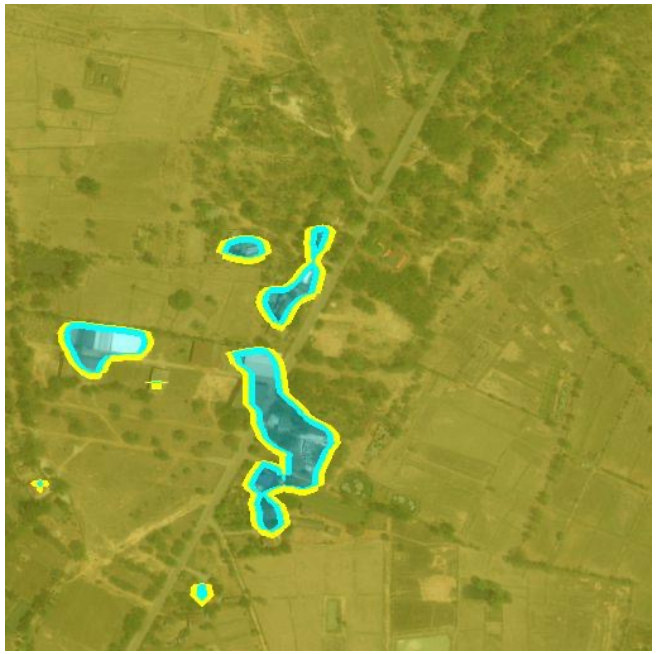
- validation/0104_sat.jpg:



Early (epoch=0)



Middle (epoch=5)



final