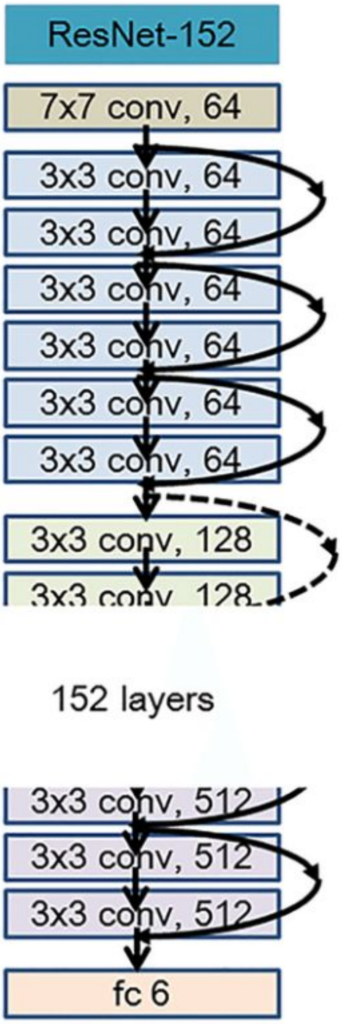
DLCV HW1 Report

Problem 1

1. Model of method B
   Resnet-152

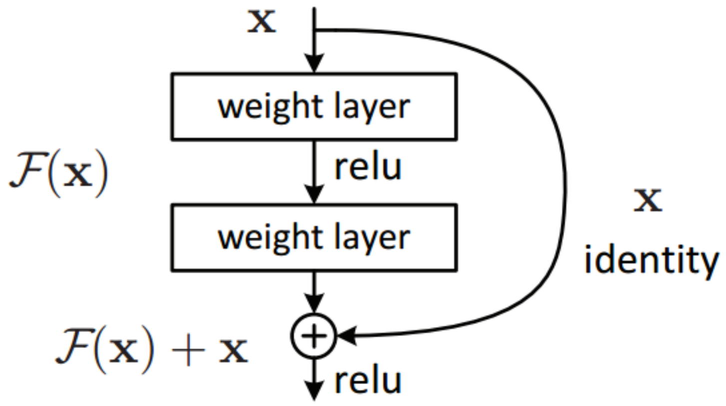


2. Accuracy on validation set

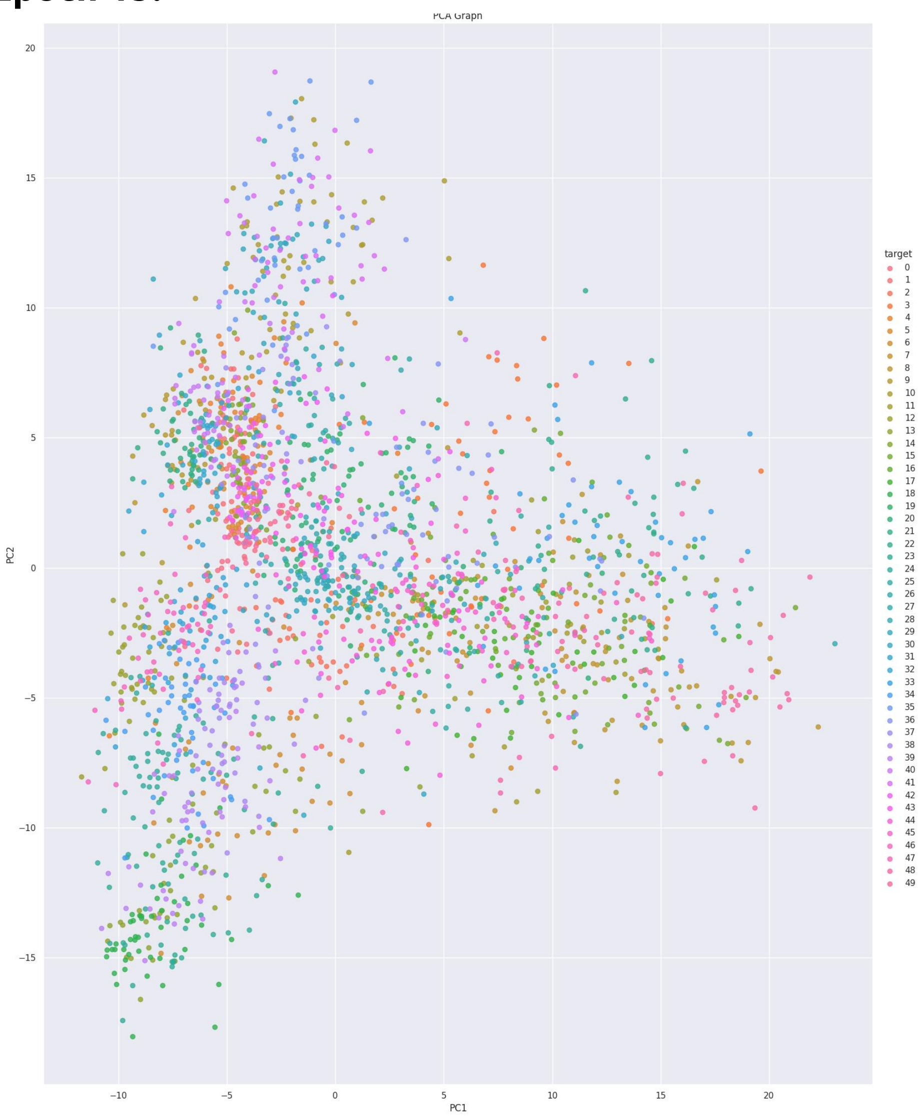
Model A	Model B
56.44%	90.72%

3. Implementation details of model A
   optimizer: **Adam** with lr = 2e-3
   lr_decay: Exponential decay
   loss function: **CrossEntropyLoss()**
   cross validation: **1-fold**
   batch normalization: Add **BatchNorm2d** in CNN block

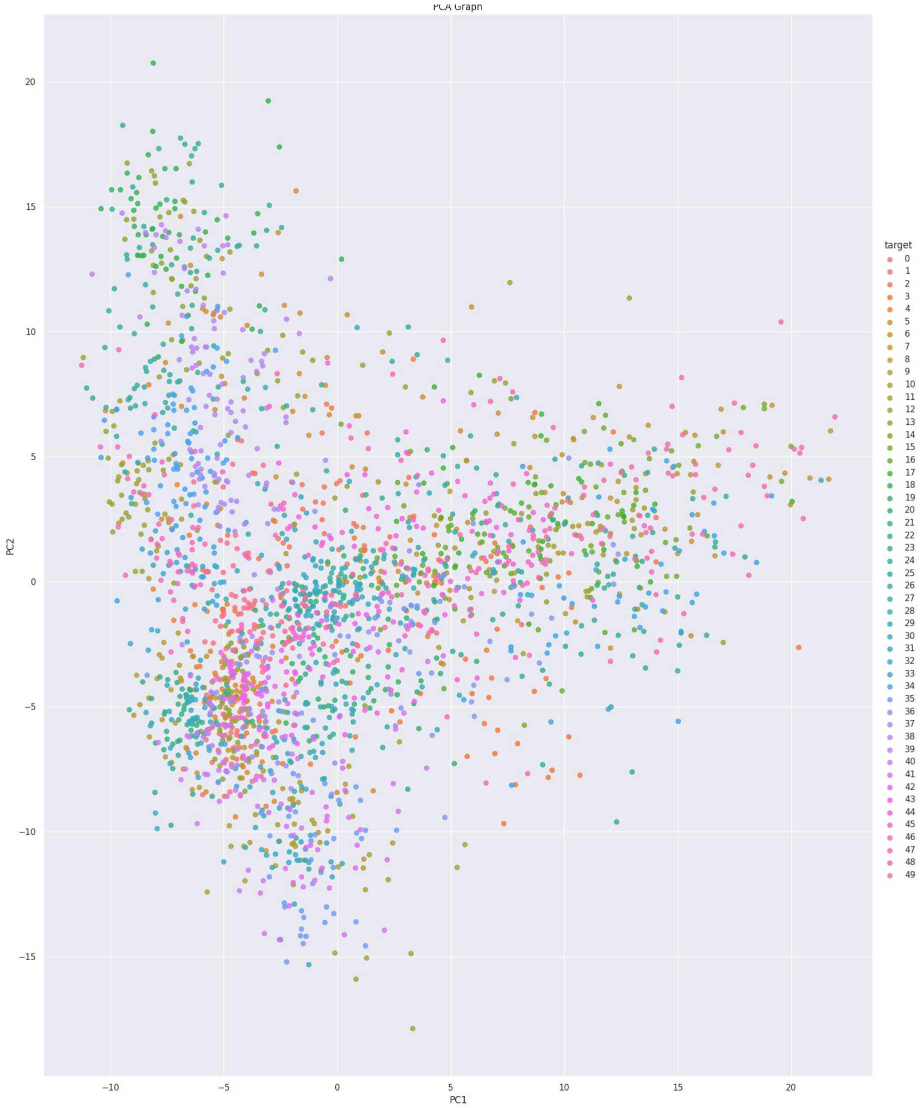
4. Method in B: Resnet152
   Resnet 引進了**Residual Block**(x跳過layer直接輸入到下一層的activation function)的機制使得原本在深層CNN容易發生的gradient vanishment/gradient explosion的問題得到緩解



5. **PCA result**
   Epoch 15:



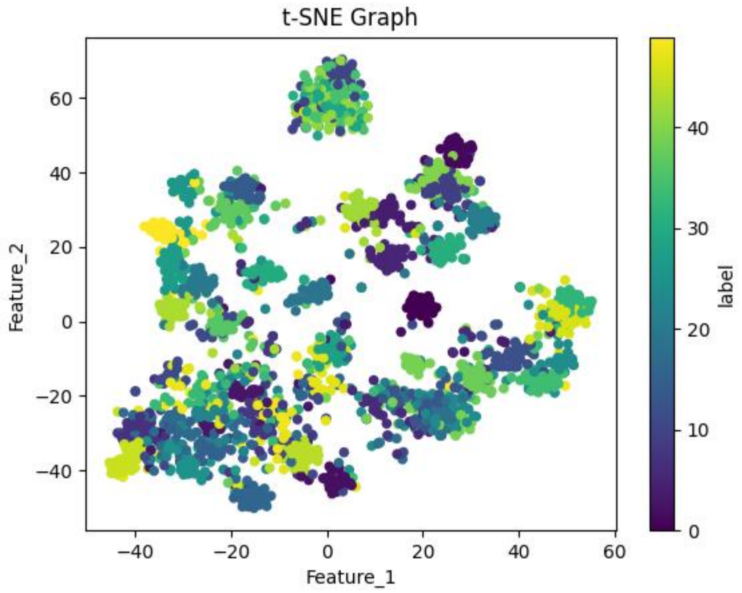
Epoch 50(last epoch):



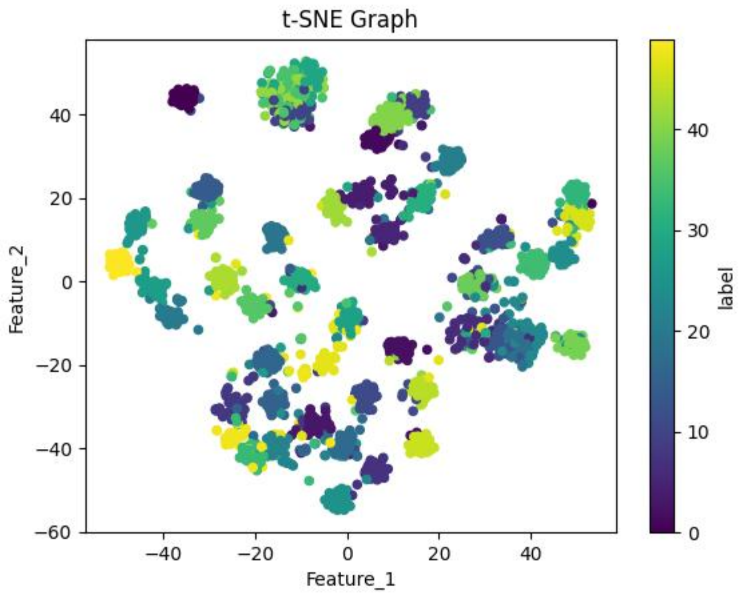
PCA主要是要做到降維的目的同時希望留下最重要的特徵，可以 看到在2-dim的圖中，模型在Epoch 15時還不能很好的分類不同class的feature但到epoch 50時已經可以把同class的feature拉近並把不同class的feature推遠

6. **t-SNE result**

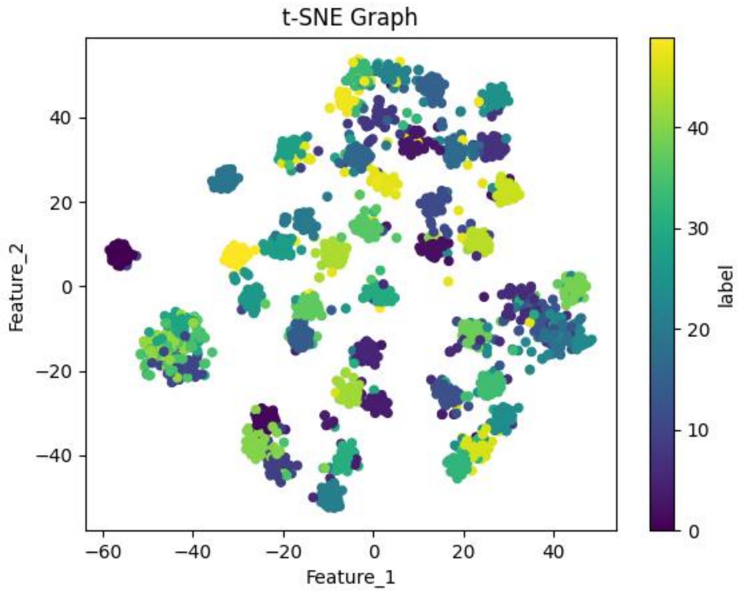
Epoch 1:



Epoch 15:



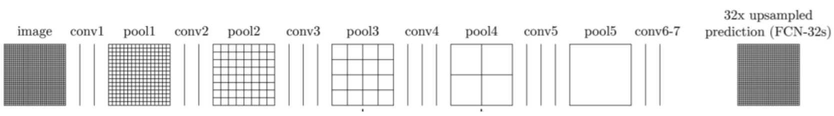
Epoch 50(last epoch):



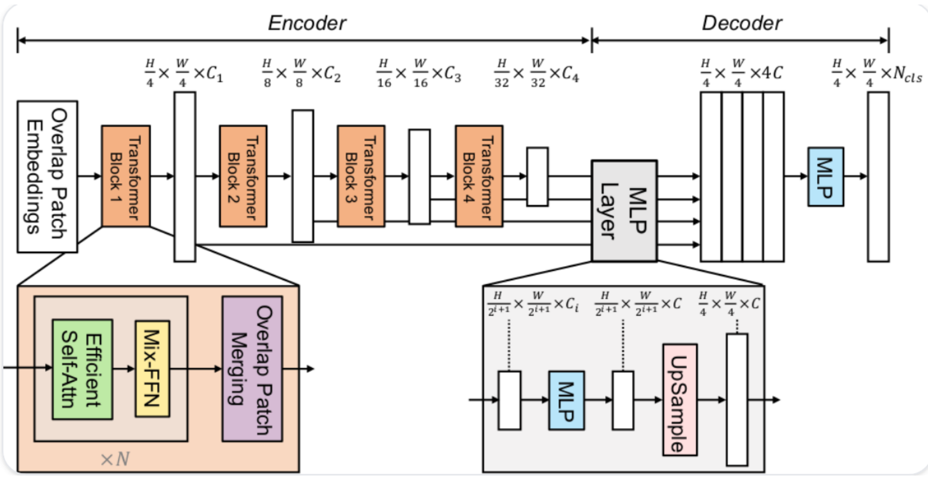
t-SNE是一種非線性的降維法，常用於模型的視覺化觀察，
可以看到Epoch數越大相同class的點分布越集中不同class的點分布則會散開

Problem 2

1. **VGG16-FCN32**



2. **Segformer**



Differs from VGG16-FCN32:
Segformer在encoder的部分是使用transformer encoder(self-attention)並且decoder方面使用輕量的MLP layer與VGG使用CNN當backbone十分不同

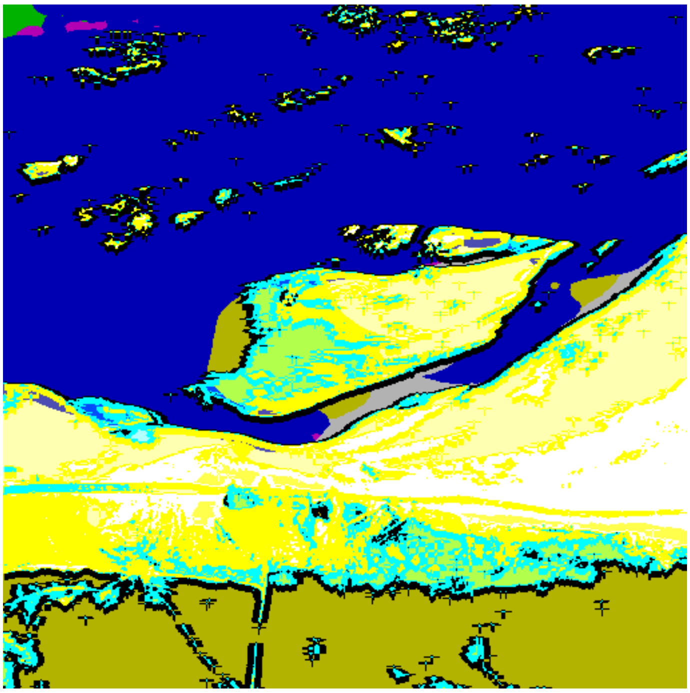
3. mIoU of two models on the validation set

Model A	Model B
70.20%	75.68%

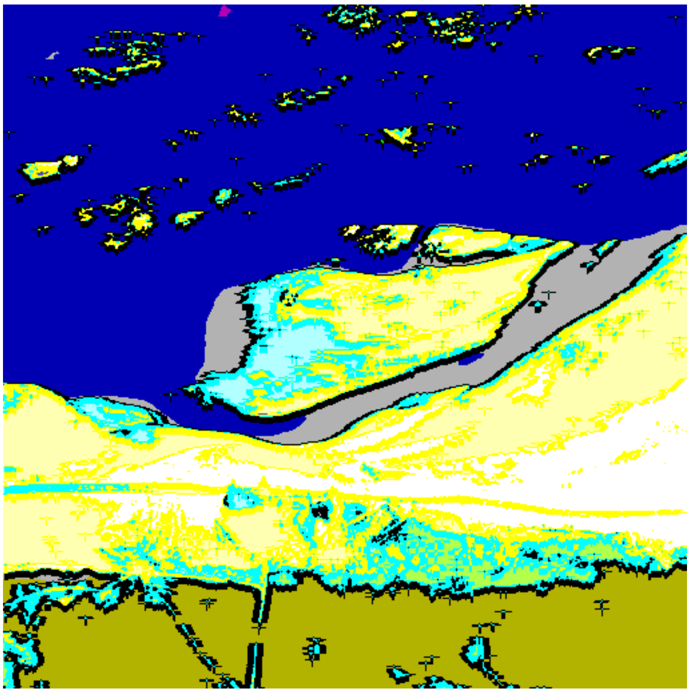
4. segmentation masks

0013_sat.jpg

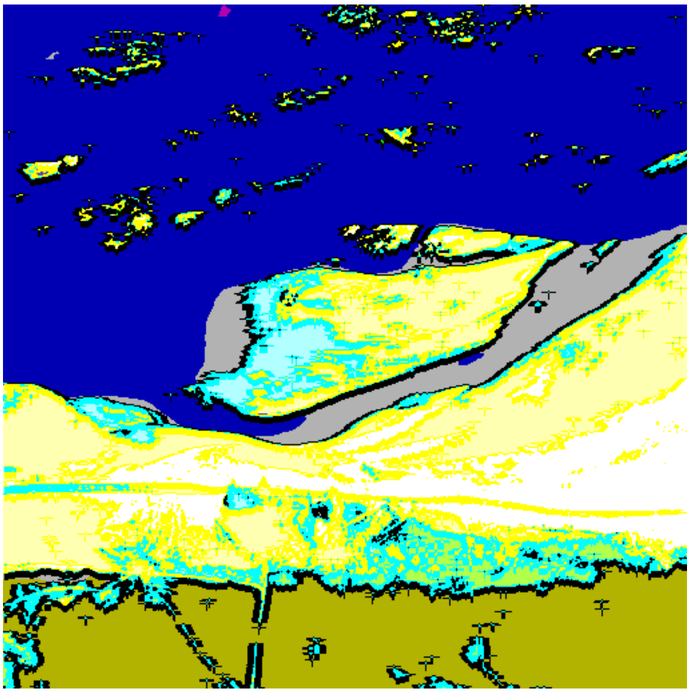
Epoch 1:



Epoch 10:

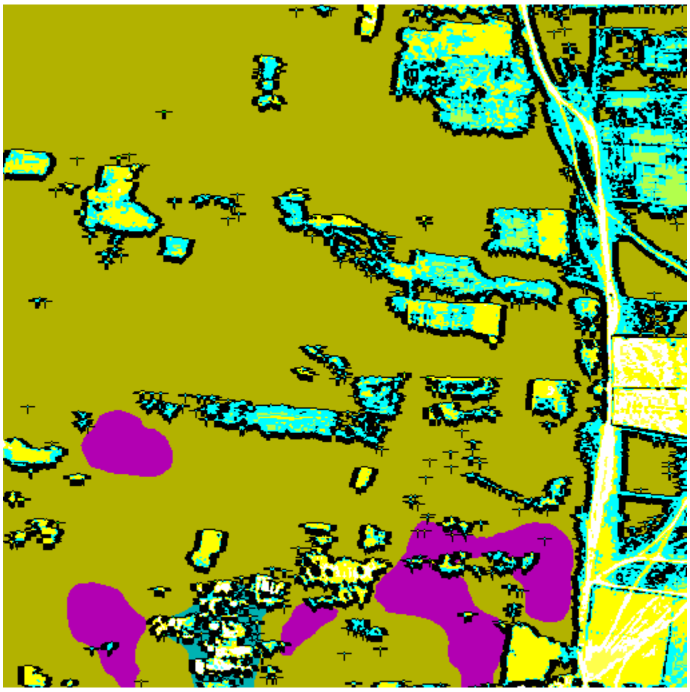


Epoch 20:

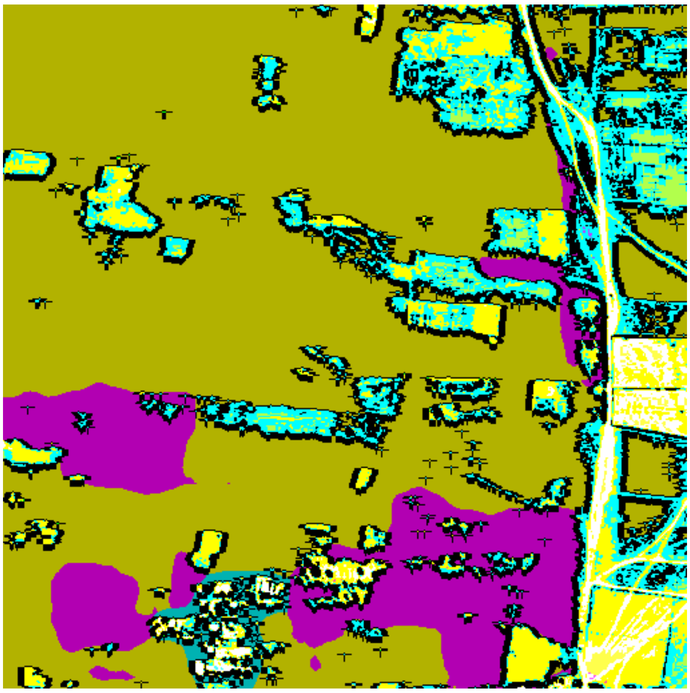


0062_sat.jpg

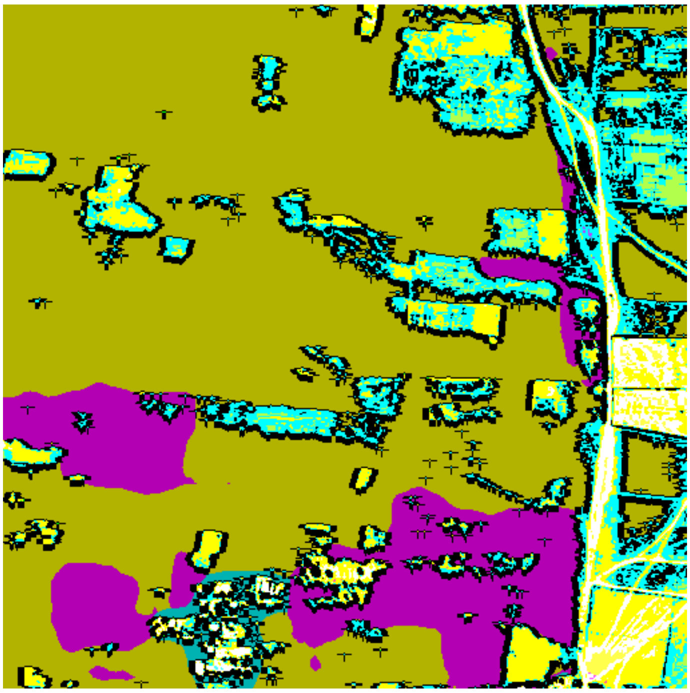
Epoch 1:



Epoch 10:

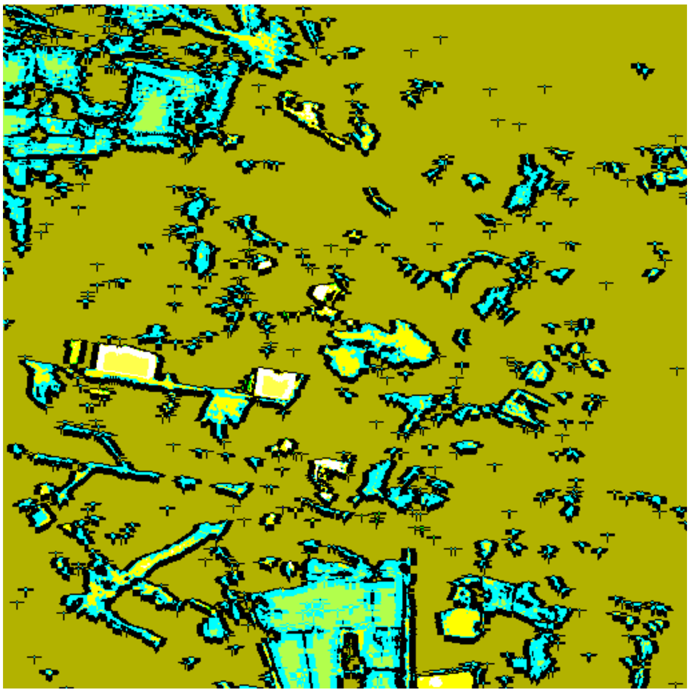


Epoch 20:

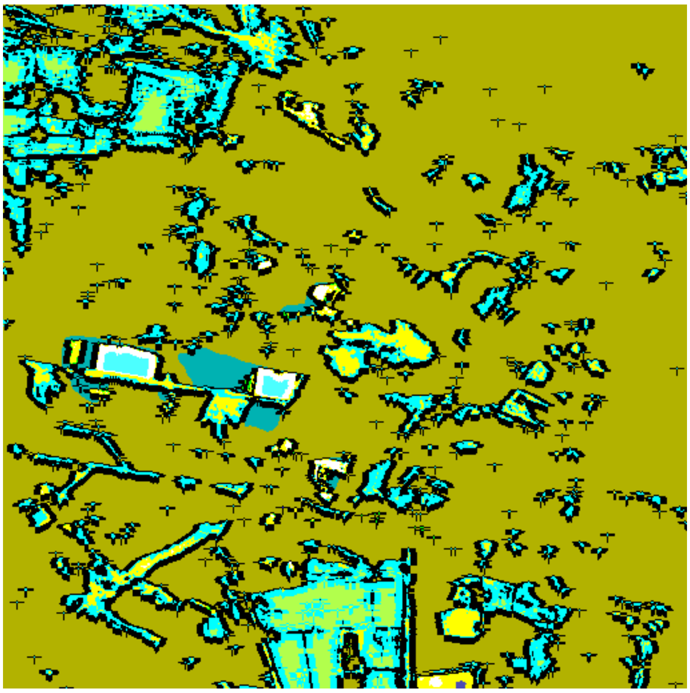


0104_sat.jpg

Epoch 1:



Epoch 10:



Epoch 20: