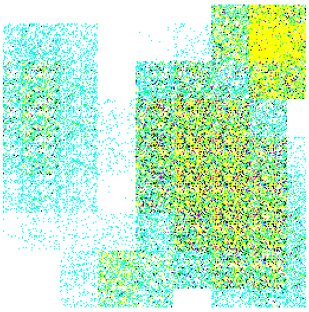
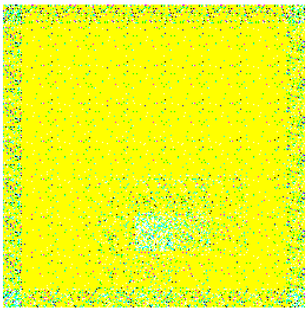
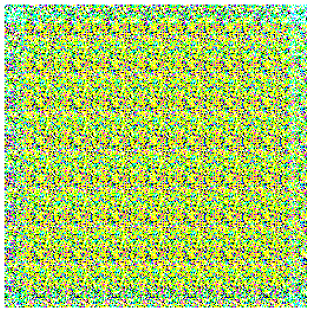
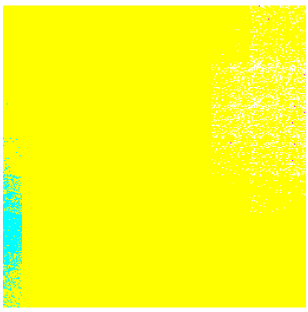
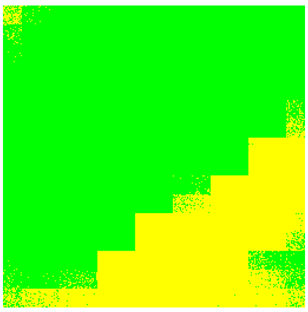
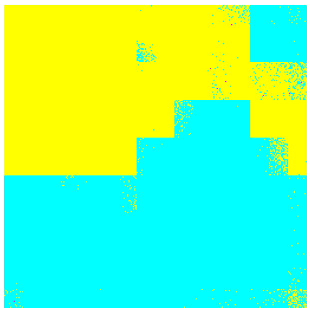
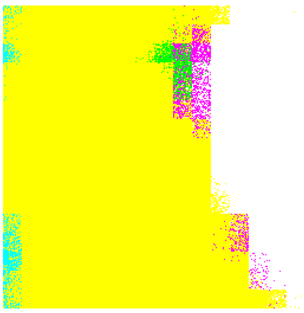
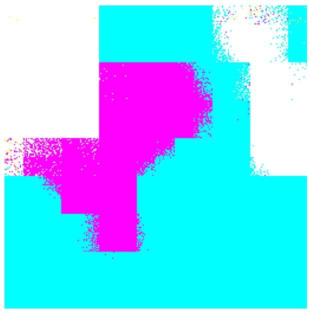
DLCV HW3 - REPORT

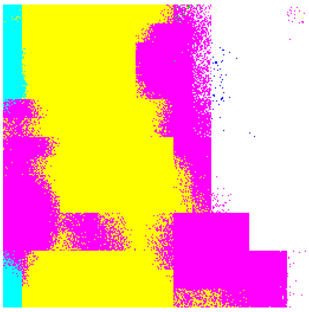
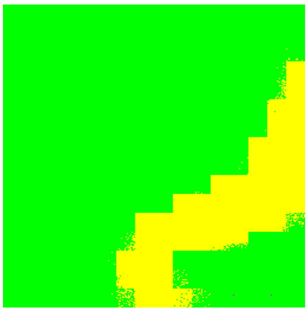
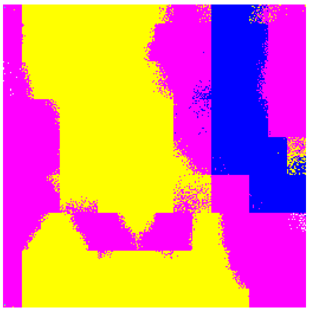
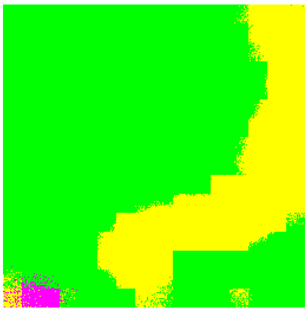
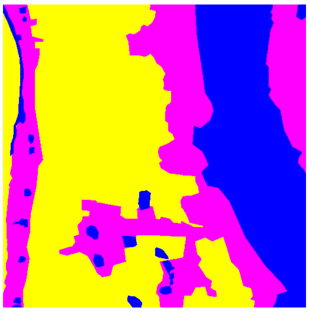
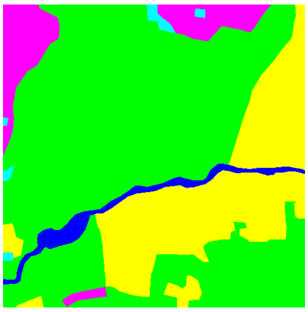
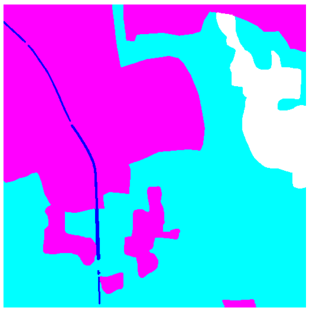
Name: 張鈞閔 Dep.:電機博二 Student ID:D05921027

1. (5%) Print the network architecture of your VGG16-FCN32s model.

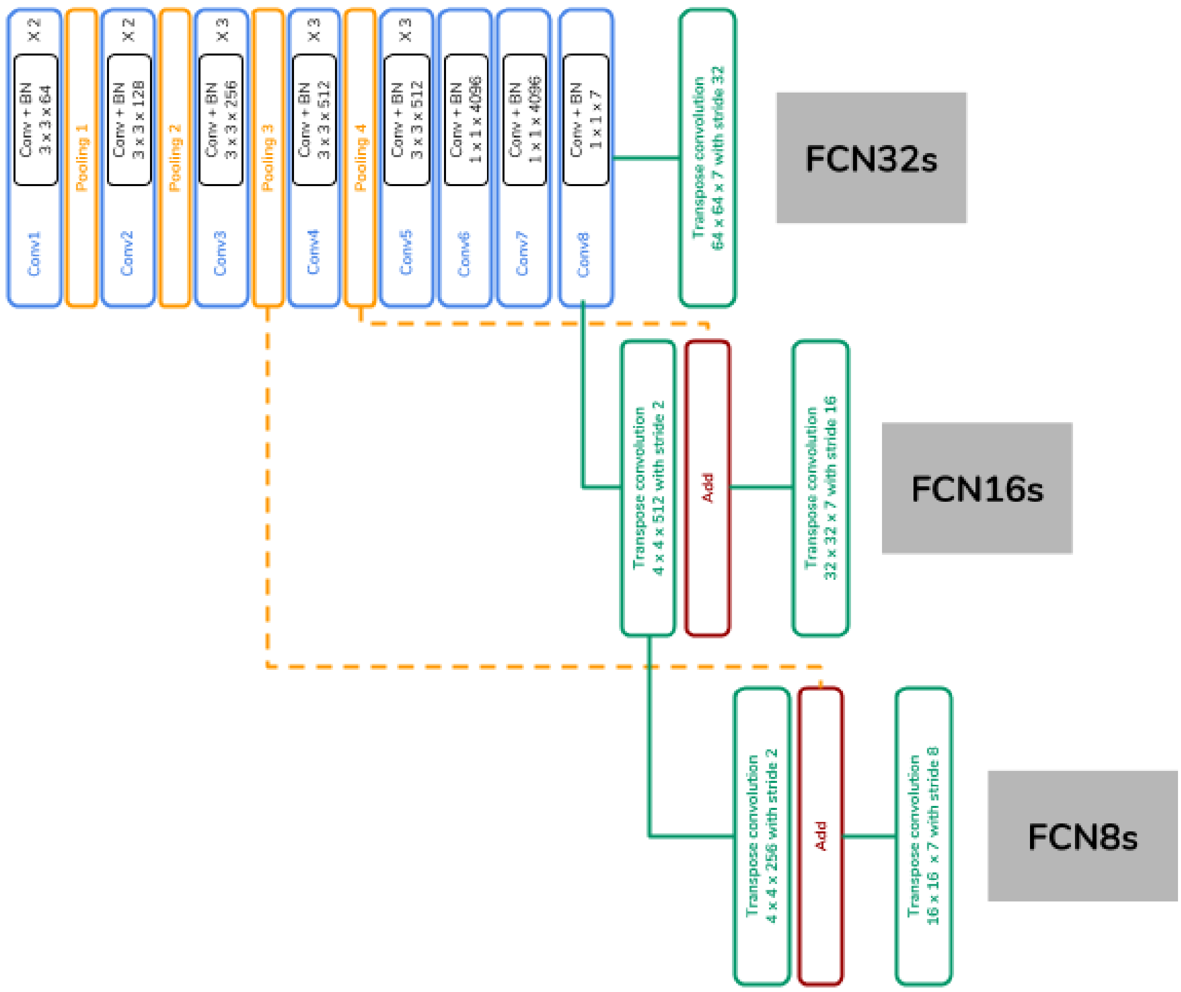
Use ReLU as the activation function in hidden layers; Softmax is applied in the output layer.

2. (10%) Show the predicted segmentation mask of validation/0008_sat.jpg, validation/0097_sat.jpg, validation/0107_sat.jpg during the early, middle, and the final stage during the training stage. (For example, results of 1st, 10th, 20th epoch)

Epoch	0008_sat	0097_sat	0107_sat
1			
20			
40			

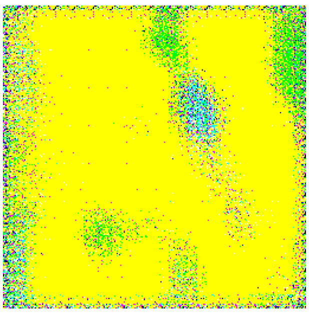
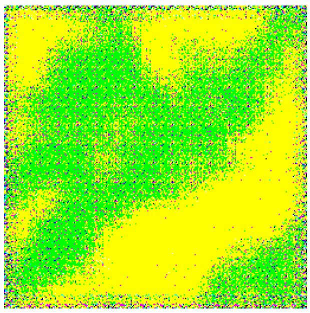
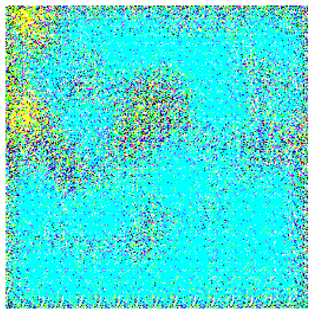
60			
80			
Final			
Ground truth			

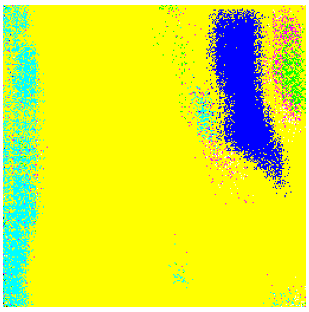
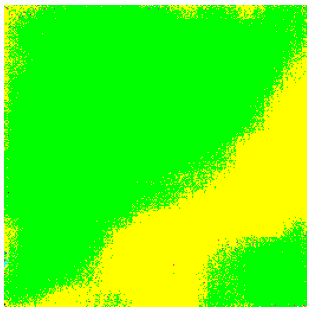
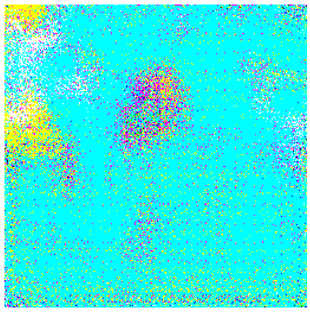
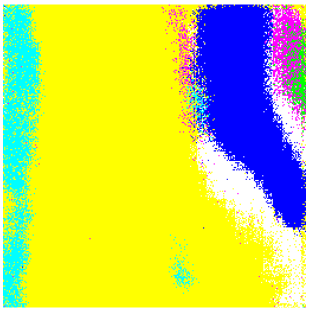
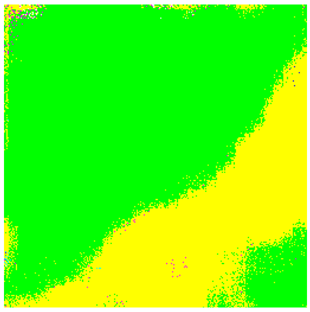
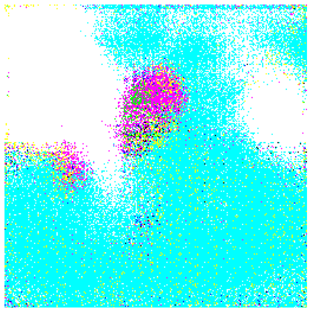
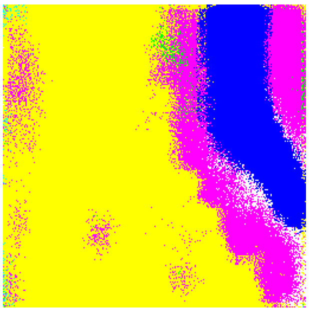
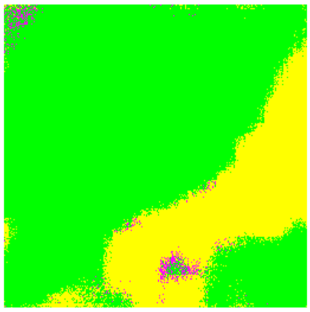
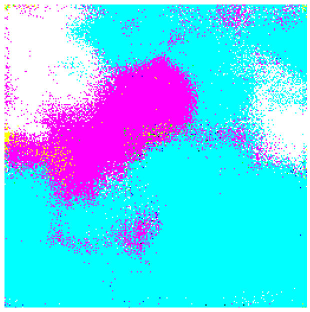
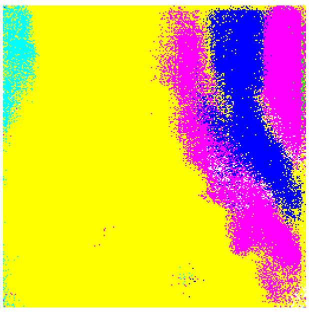
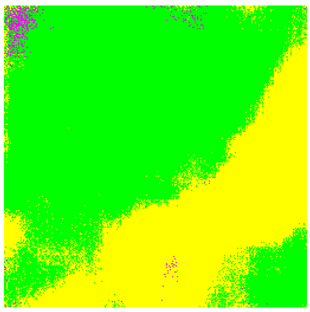
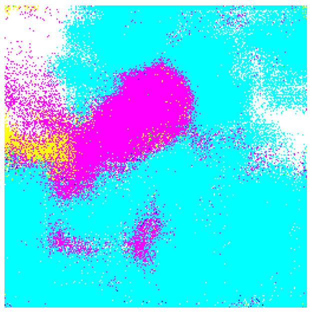
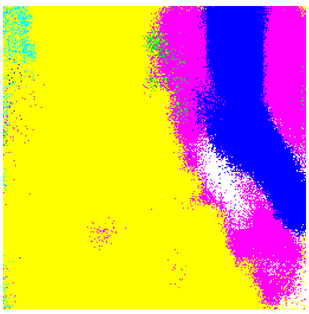
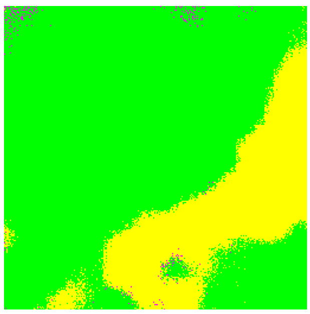
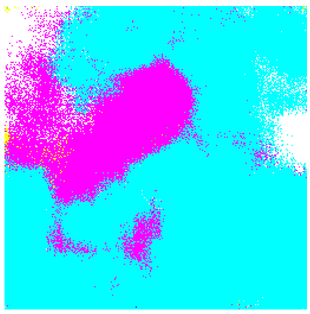
3. (15%) Implement an improved model which performs better than your baseline model. Print the network architecture of this model.

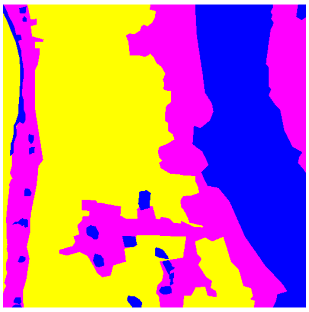
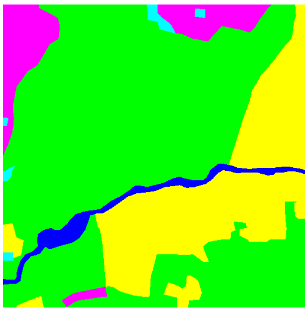
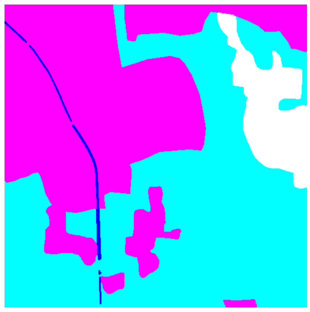


Use ReLU activation functions in every hidden layers; Softmax is applid in output layers. “Add” layer implies element-wise addition.

4. (10%) Show the predicted segmentation mask of validation/0008_sat.jpg, validation/0097_sat.jpg, validation/0107_sat.jpg during the early, middle, and the final stage during the training process of this improved model.

Epoch	0008_sat	0097_sat	0107_sat
1			

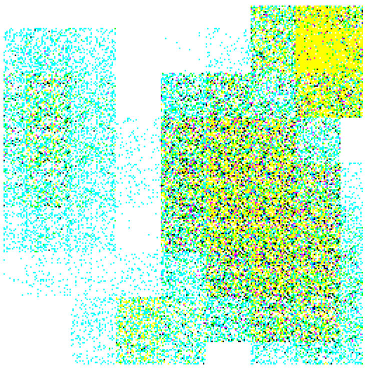
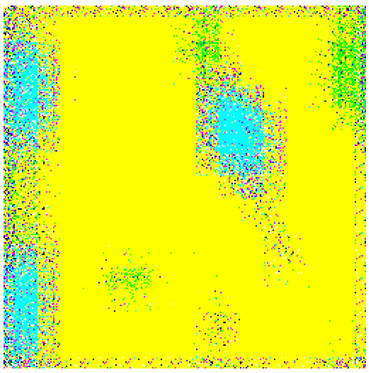
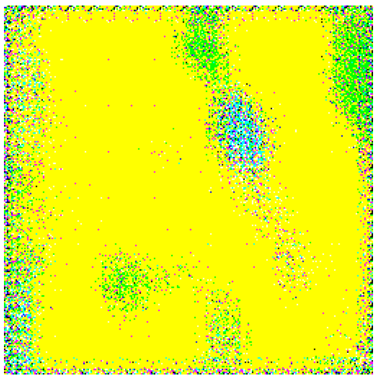
10			
20			
30			
40			
Final			

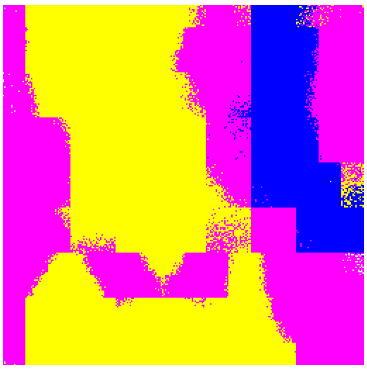
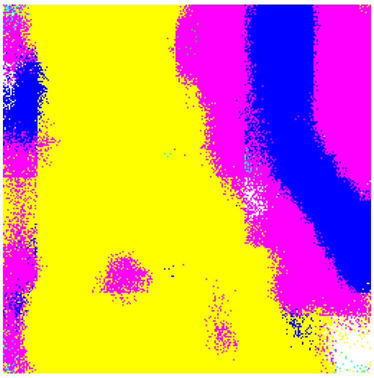
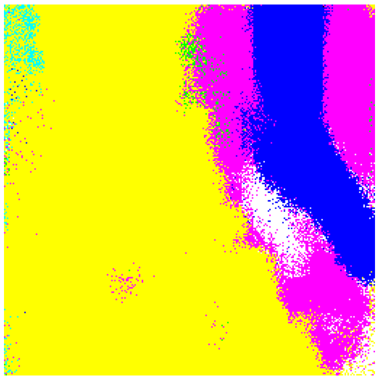
Ground truth			

5. (15%) Report mIoU score of both models on the validation set. Discuss the reason why the improved model performs better than the baseline one. You may conduct some experiments and show some evidences to support your discussion.

	FCN32s	FCN16s	FCN8s
mIoU	0.656	0.668	0.673

By comparing the predicted segmentation mask of validation/0008_sat.jpg by FCN8s, FCN16s, and FCN32s, we found that
(1) FCN32s has the worst granularity, due to its largest upsampling rate (transpose_conv2d with stride 32), and the resulting mask cannot smoothly present the twists and turns of the mask. For example, there are several conspicuous right angles at the boundary of the blue and purple region on the right-hand side.
(2) Both FCN16s and FCN8s slightly improved the granularity issue since they incorporate relatively higher-resolution feature maps that preserve more detailed context. Besides, their upsampling rates are also less than FCN32s and have smoother segmentation results. As a result, it is no surprise that the improved model, FCN8s, outperforms the baseline model.

FCN32s	FCN16s	FCN8s
 after 1 epoch	 after 1 epoch	 after 1 epoch

 finalized	 finalized	 finalized

6. (5%) [bonus] Caculate the result of d/dw G(w):

objective function:

$$G(\boldsymbol{w}) = -\sum_n \left[t^{(n)} \log \mathrm{x}(\boldsymbol{z}^{(n)};\boldsymbol{w}) + (1-t^n)\log\left(1-\mathrm{x}(\boldsymbol{z}^{(n)};\boldsymbol{w})\right)\right] \geq 0$$

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} G(\boldsymbol{w})$$ choose the weights that minimise the network's surprise about the training data

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{w}}G(\boldsymbol{w}) = \sum_n \frac{\mathrm{d}G(\boldsymbol{w})}{\mathrm{d}x^{(n)}}\frac{\mathrm{d}x^{(n)}}{\mathrm{d}\boldsymbol{w}} = -\sum_n (t^{(n)} - x^{(n)})\boldsymbol{z}^{(n)}$$ = prediction error x feature

$$\boldsymbol{w} \leftarrow \boldsymbol{w} - \eta\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{w}}G(\boldsymbol{w})$$ iteratively step down the objective (gradient points up hill) 39

$$x(z,w) = \frac{1}{1+e^{-zw}}$$

DATE

for a given n,

NO.

$$G(w) = t^{(n)} \cdot \log x(z^{(n)};w) + (1-t^{(n)}) \log (1-x(z^{(n)};w))$$

omit (n),

$$G(w) = t \cdot \log (x(z;w)) + (1-t) \log (1-x(z;w))$$

$$\frac{\partial G}{\partial w} = t \frac{1}{x(z;w)} \frac{\partial x}{\partial w} + (1-t) \cdot \frac{-1}{1-x(z;w)} \cdot \frac{\partial x}{\partial w}$$

$$\frac{\partial x}{\partial w} = \frac{ze^{-zw}}{(1+e^{-zw})^2}$$

$$= t\cancel{(1+e^{-zw})} \frac{ze^{-zw}}{(1+e^{-zw})^{\cancel{2}}} - (1-t) \cdot \frac{\cancel{1+e^{-zw}}}{\cancel{e^{-zw}}} \cdot \frac{z\cancel{e^{-zw}}}{(1+e^{-zw})^{\cancel{2}}}$$

$$= \frac{zte^{-zw} + tz - z}{1+e^{-zw}}$$

$$= \left(\frac{te^{-zw} + t - 1}{1+e^{-zw}}\right) z$$

$$= \left[\frac{t(e^{-zw}+1) - 1}{1+e^{-zw}}\right] z$$

$$= (t-x)z$$

for every n:

$$\underline{G(w) = \sum_n (t^{(n)} - x^{(n)}) z^{(n)}} \quad \#$$