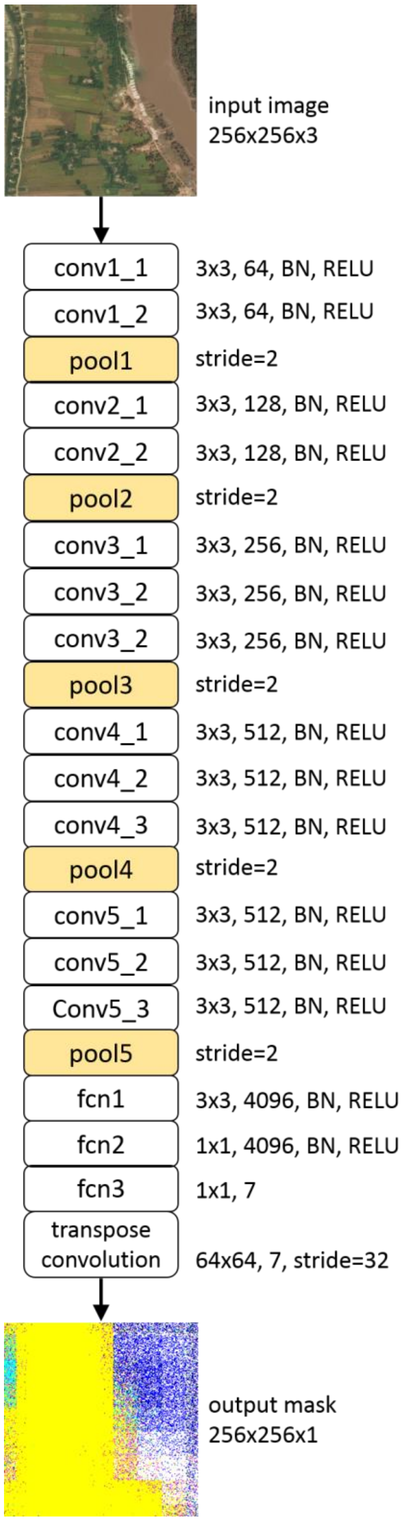
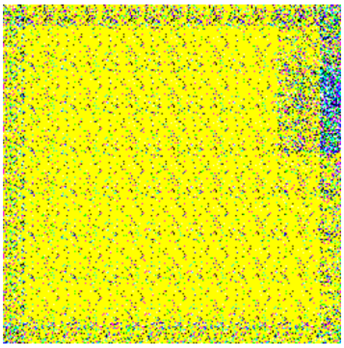
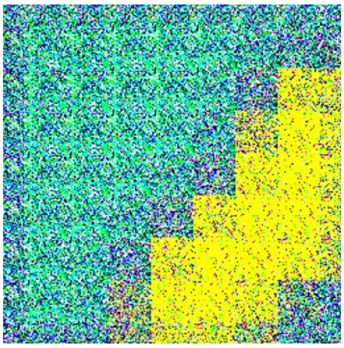
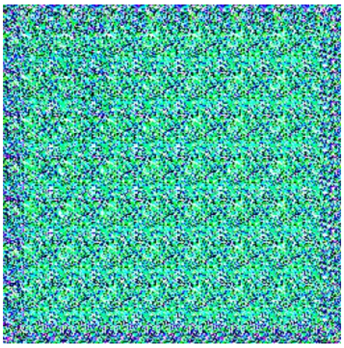
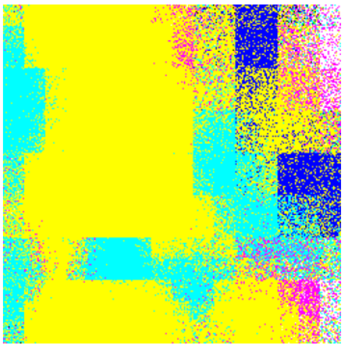
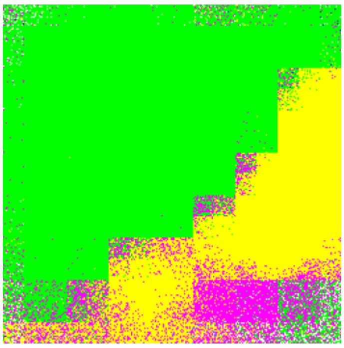
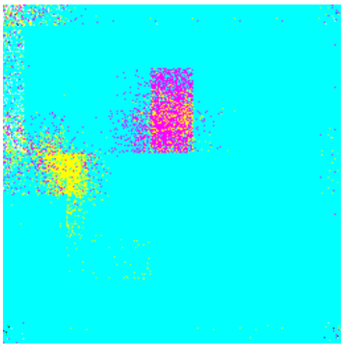
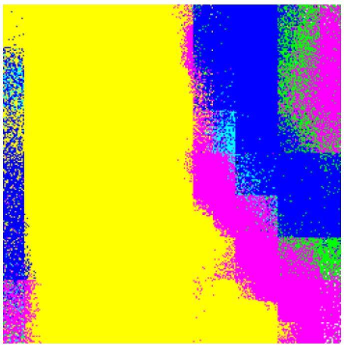
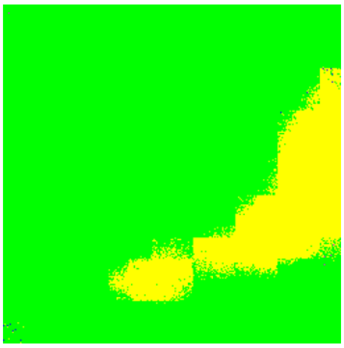
DLCV Sprint 2018 HW3

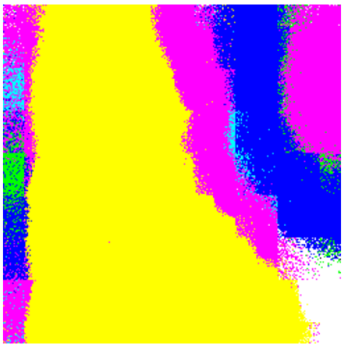
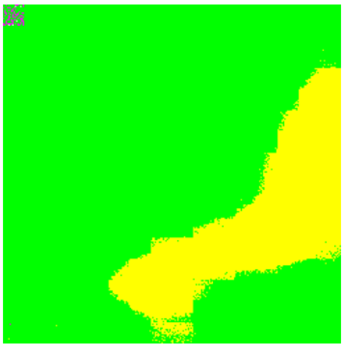
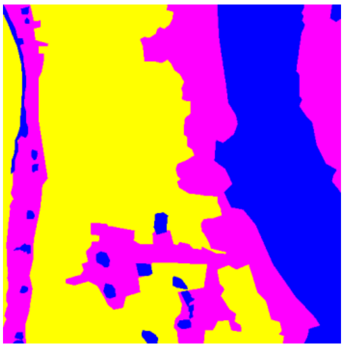
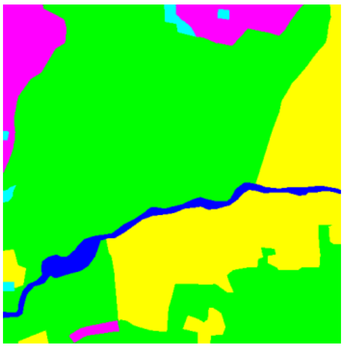
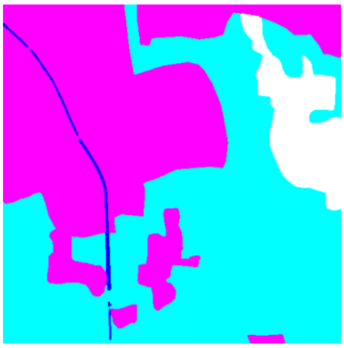
d05921018 林家慶

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

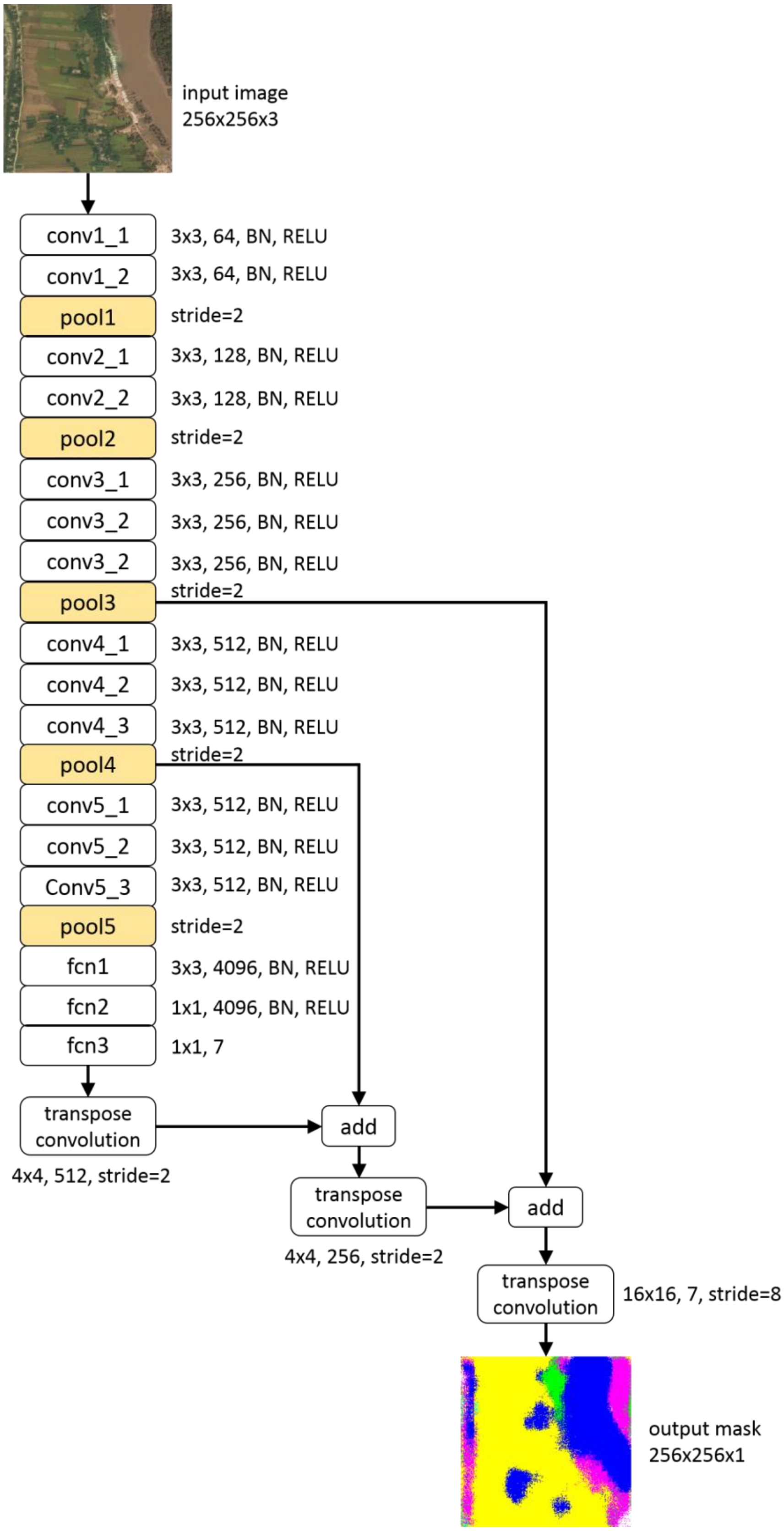


2. (10%) Show the predicted segmentation mask of "validation/0008_sat.jpg", "validation/0097_sat.jpg", and "validation/0107_sat.jpg" during the early, middle, and the final stage during the training stage.

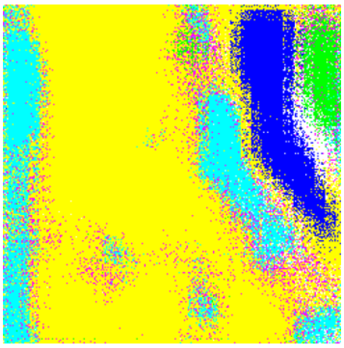
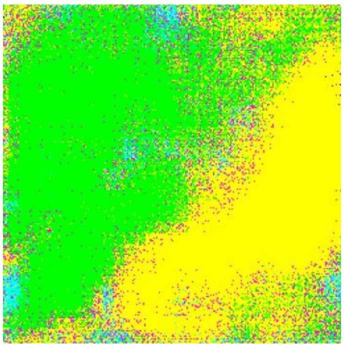
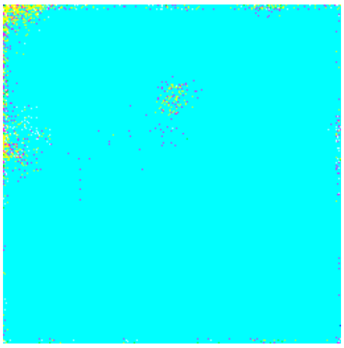
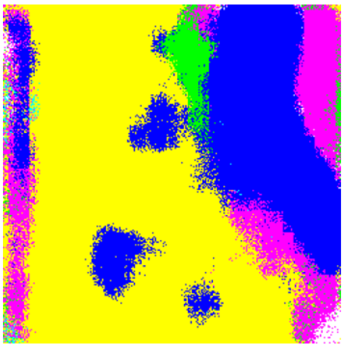
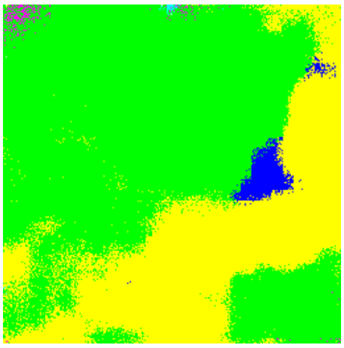
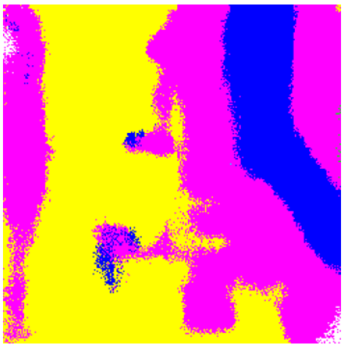
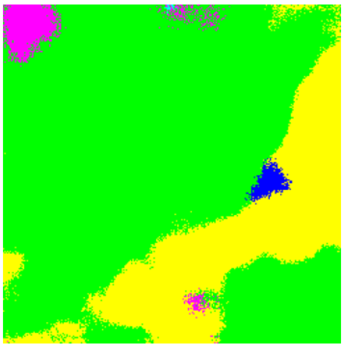
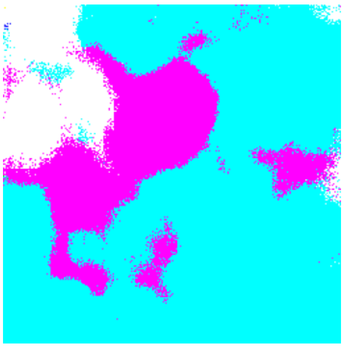
	0008	0097	0107
Satellite			
Epoch 1			
Epoch 10			
Epoch 20			

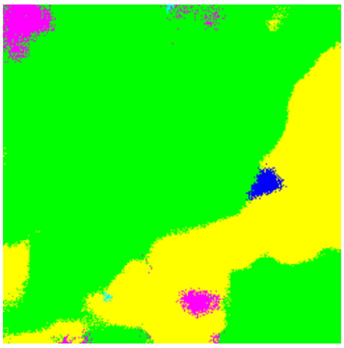
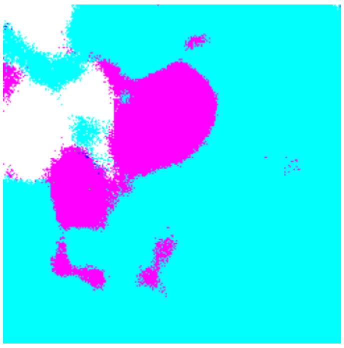
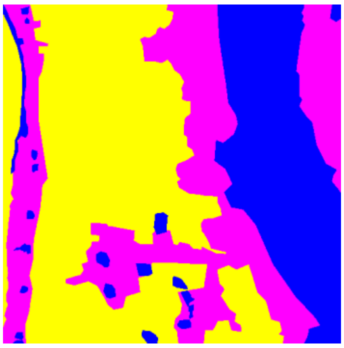
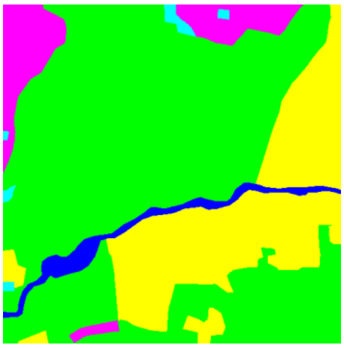
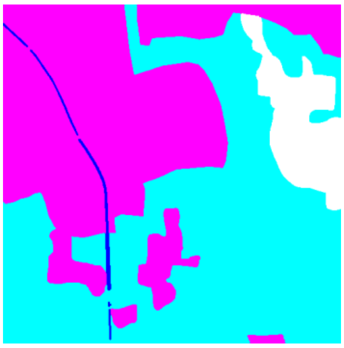
Epoch 30 (Final)			
Ground Truth			

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



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.

	0008	0097	0107
Satellite			
Epoch 1			
Epoch 10			
Epoch 20			

Epoch 30 (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 reasoning.

Performance summary on the validation set:

	VGG16_FCN32s	**VGG16_FCN8s**
class #0	0.69703	0.73903
class #1	0.87118	0.88106
class #2	0.30488	0.37889
class #3	0.80328	0.80991
class #4	0.70292	0.73472
class #5	0.68593	0.68962
mean_iou	**0.677539**	**0.705539**

As can be seen from the predicted masks during training processes, VGG16_FCN32s produces more coarse masks due to its large upsampling rate (transpose convolution with stride 32). As for VGG16_FCN8s, it produce masks with finer granuralities due to its smaller upsampling rate at the final stage (transpose convolution with stride 8). Also, since VGG16_FCN8s also take intermediate layer outputs (pool4 and pool3) into account, which can be seen as some kind of context module as used in dilated convolutional segmentation, hence it can produce masks with higher qualities.

(bonus)

$$x(\underline{z}^{(n)};\underline{w}) = \frac{1}{1+e^{-\underline{w}^T\underline{z}^{(n)}}} \Rightarrow 1 - x(\underline{z}^{(n)};\underline{w}) = \frac{e^{-\underline{w}^T\underline{z}^{(n)}}}{1+e^{-\underline{w}^T\underline{z}^{(n)}}}$$

Also, $\dfrac{d\,x(\underline{z}^{(n)};\underline{w})}{d\underline{w}} = \dfrac{e^{-\underline{w}^T\underline{z}^{(n)}}\;\underline{z}^{(n)}}{[1+e^{-\underline{w}^T\underline{z}^{(n)}}]^2}$

For each n, let $G^{(n)}(\underline{w}) = t^{(n)} \ln x(\underline{z}^{(n)};\underline{w}) + (1-t^{(n)}) \ln(1 - x(\underline{z}^{(n)};\underline{w}))$

$$\Rightarrow \frac{d}{d\underline{w}} G^{(n)}(\underline{w}) = \frac{d\,G^{(n)}(\underline{w})}{d\,x(\underline{z}^{(n)};\underline{w})} \cdot \frac{d\,x(\underline{z}^{(n)};\underline{w})}{d\underline{w}}$$

$$= \left[\frac{t^{(n)}}{x(\underline{z}^{(n)};\underline{w})} + \frac{t^{(n)}-1}{1-x(\underline{z}^{(n)};\underline{w})}\right] \cdot \frac{d\,x(\underline{z}^{(n)};\underline{w})}{d\underline{w}}$$

$$= \left[t^{(n)}\left(1+e^{-\underline{w}^T\underline{z}^{(n)}}\right) + \frac{(t^{(n)}-1)(1+e^{-\underline{w}^T\underline{z}^{(n)}})}{e^{-\underline{w}^T\underline{z}^{(n)}}}\right] \cdot \frac{e^{-\underline{w}^T\underline{z}^{(n)}}\;\underline{z}^{(n)}}{\left[1+e^{-\underline{w}^T\underline{z}^{(n)}}\right]^2}$$

$$= \left[t^{(n)}\left(\frac{e^{-\underline{w}^T\underline{z}^{(n)}}}{1+e^{-\underline{w}^T\underline{z}^{(n)}}}\right) + (t^{(n)}-1)\left(\frac{1}{1+e^{-\underline{w}^T\underline{z}^{(n)}}}\right)\right] \underline{z}^{(n)}$$

$$= \left[t^{(n)}\left(1 - x(\underline{z}^{(n)};\underline{w})\right) + \left(t^{(n)}-1\right) x(\underline{z}^{(n)};\underline{w})\right] \underline{z}^{(n)}$$

$$= \left[t^{(n)} - x(\underline{z}^{(n)};\underline{w})\right] \underline{z}^{(n)}$$

$$\Rightarrow \frac{d}{d\underline{w}} G(\underline{w}) = -\sum_n \left(\frac{d}{d\underline{w}} G^{(n)}(\underline{w})\right) = -\sum_n \left(t^{(n)} - x(\underline{z}^{(n)};\underline{w})\right) \underline{z}^{(n)}$$

#