Image Segmentation via UNet

Presentation Summary

- Background
- Approach
  - Set-up
  - Model
  - Loss Function
- Results

NYU TANDON SCHOOL OF ENGINEERING

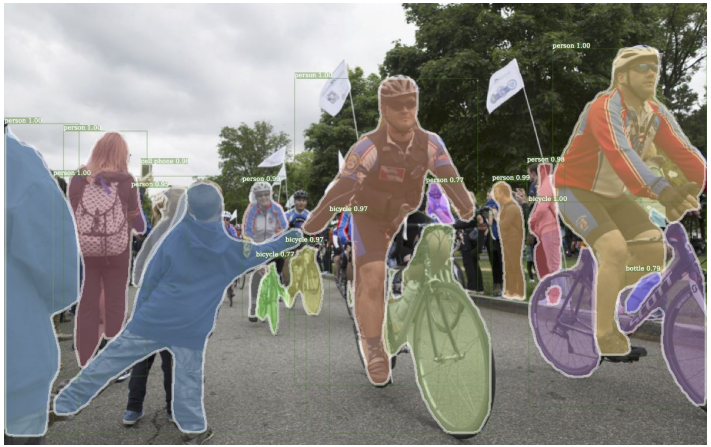
Background

Applications:

- Surveillance Systems
- Guidance Camera for Autonomous Driving
- Medical Imaging
- People Counting & Flow control

[3]

[2]



[1]

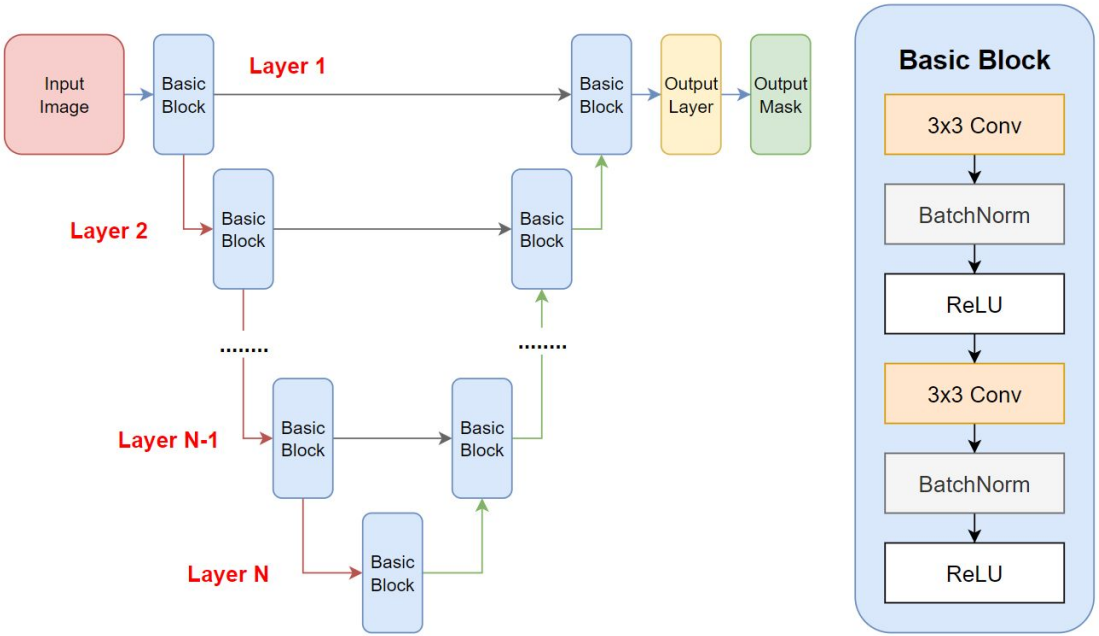
NYU TANDON SCHOOL OF ENGINEERING

Approach: Set-up

- Hardware: NVIDIA GeForce RTX 2060, Intel(R) Core(™) i7-9700 CPU @ 3.00GHz
- Platform: Windows 10 conda environment
- Framework: python=3.8 cuda=11.3 PyTorch=1.11.0 opencv=4.0.1

  For detail framework please visit out github ./conda_environment.yml
- Dataset: PennFudanPed [5]
- Functionalities: Detect Human
- Limitations: Accuracy drops when real-world data implemented.

Approach: UNet Model



NYU TANDON SCHOOL OF ENGINEERING

Fig 1. Model Overview

Approach: Simply stacking more layers?

“Overly deep” Network might have higher training and validation errors

Too many Layers will cause Model Overfitting

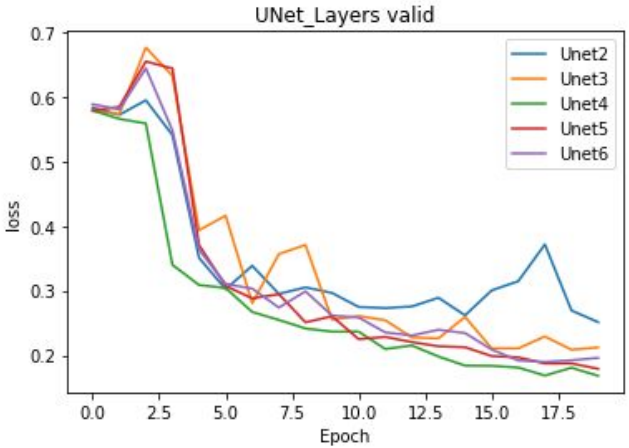


Fig 2. UNet_Layers valid

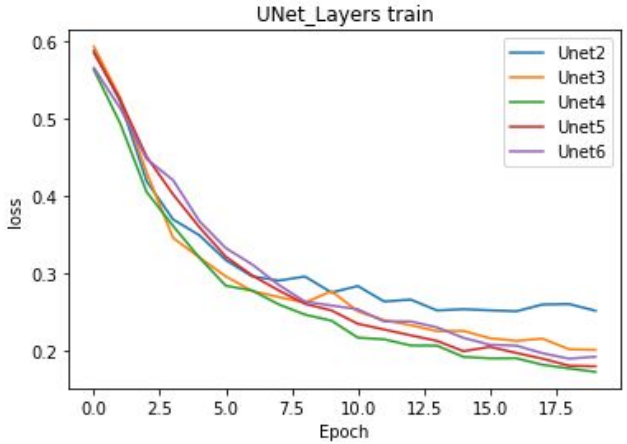


Fig 3. UNet_Layers train

As the result, we pick UNet 4-Layer to train our dataset.

NYU TANDON SCHOOL OF ENGINEERING

Approach: Loss Function

- $$DICE = \frac{2 \times |A \cap B|}{|A| + |B|} = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i + \sum_i^N g_i}$$

- ***DICE Loss = 1 - DICE***
- ***Example:***

$$DICE = \frac{2\sum_{i=0}^{25*80} 1}{\sum_{i=0}^{75*80} 1 + \sum_{i=0}^{50*80} 1} \approx 0.400$$

DICE Loss = 1 - DICE = 1 - 0.4 = 0.6

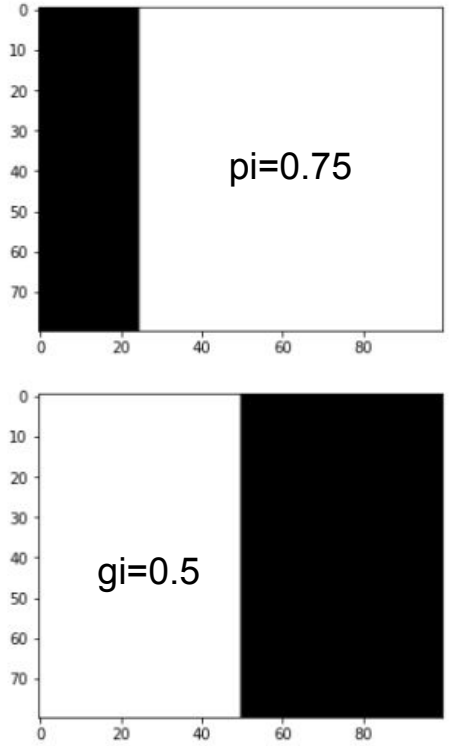


Fig 4. DICE visualized

Result: Loss

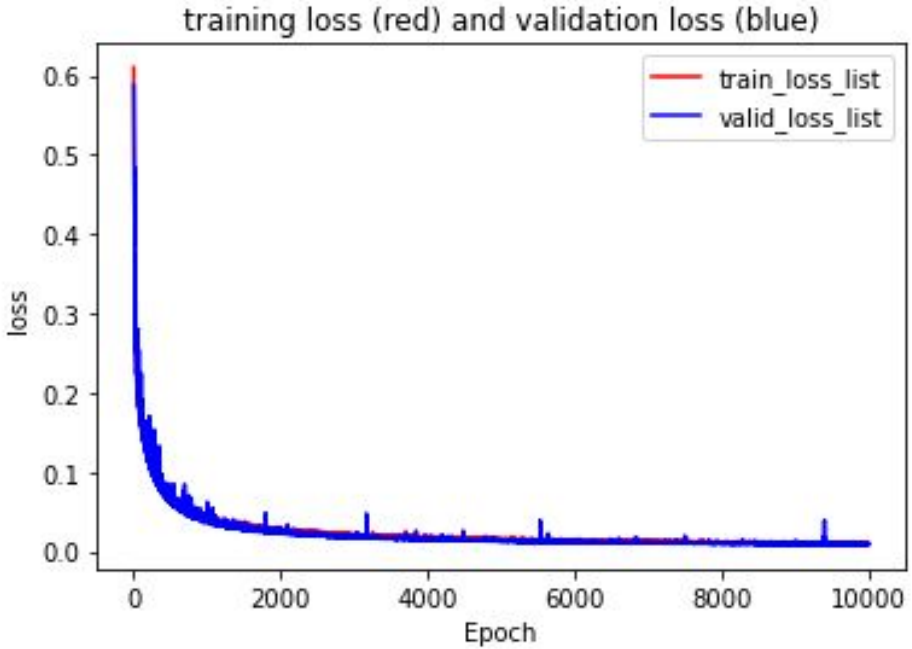


Fig 5. DICE visualized

NYU TANDON SCHOOL OF ENGINEERING

Result

Fig 6. Visualized Results

Reference

Project GitHub repo with documented code

https://github.com/popkarthb/UNet-Image-Segmentation-Deep-Learning

[1] Gruosso, Monica, et al. “Human Segmentation in Surveillance Video with Deep Learning.” *SpringerLink*, Springer US, 6 Sept. 2020, https://link.springer.com/article/10.1007/s11042-020-09425-0.

[2]He, Kaiming. “CVPR'17 Tutorial on Deep Learning for Objects and Scenes.” *CVPR'17 Tutorial*, http://deeplearning.csail.mit.edu/slide_cvpr2018/laurens_cvpr18tutorial.pdf.

[3] “Detectron.” *Meta AI*, https://ai.facebook.com/tools/detectron/.

[4] Ronneberger, Olaf, et al. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” *ArXiv.org*, 18 May 2015, https://arxiv.org/abs/1505.04597.

[5]Wang, Liming, et al. “Penn-Fudan Database for Pedestrian Detection and Segmentation.” *Pedestrian Detection Database*, University of Pennsylvania, Fudan University, Jan. 2007, https://www.cis.upenn.edu/~jshi/ped_html/.

NYU | TANDON SCHOOL OF ENGINEERING