CS 6476 Project 6

Vidit Pokharna
vidit@gatech.edu
vpokharna3
903772087

Parts 4 & 5: mIoU of different models (9 points)

Add each of the following (keeping the changes as you move to the next row):

	Training mIoU	Validation mIoU
Simple Segmentation Net (no pretrained weights)	0.429	0.318
+ ImageNet-Pretrained backbone	0.462	0.425
+ Data augmentation	0.451	0.430
ImageNet-Pretrained PSPNet w/ Data Aug. without PPM	0.588	0.601
+ PSPNet with PPM	0.592	0.608
+ PSPNet with auxiliary loss	0.613	0.624

Parts 4 & 5 (5 points)

1. ImageNet-Pretrained backbone: Using a pretrained backbone provides the model with a good initial feature representation. This helps the model to generalize better from the start, which is likely why both the training and validation mIoU scores increase when moving from a simple segmentation net without pretrained weights to one with an ImageNet-pretrained backbone.

2. Data augmentation: Data augmentation techniques create variations of the training data, which helps in preventing overfitting and enhances the model's ability to generalize to new, unseen data. This can explain the slight improvement in the validation mIoU, even though the training mIoU sees a marginal drop possibly due to the increased difficulty for the model to fit to the more complex, augmented data.

3. PSPNet: The PSPNet incorporates a pyramid pooling module that aggregates context at different scales. This enables the model to capture details as well as the big picture, which is important for understanding the scene in segmentation tasks. The significant jump in both training and validation mIoU after incorporating PSPNet indicates its effectiveness in capturing multi-scale contextual information.

4. PSPNet with PPM: The addition of the PPM to PSPNet further enhances the multi-scale context aggregation by pooling features at different scales and then upscaling them to the original image size to achieve precise localization. The improvement in mIoU scores reflects the model's increased capability to understand and segment images at varying resolutions.

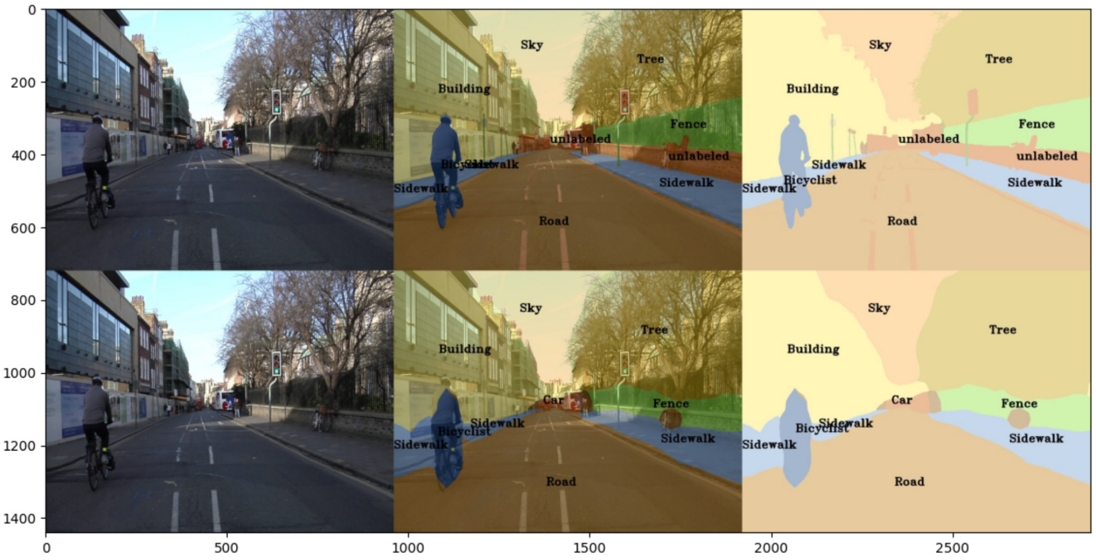
5. PSPNet with auxiliary loss: Auxiliary loss functions are often added to help with the training of deep networks, acting as a form of regularization and helping the gradient to propagate back through the network. This can improve the learning of the intermediate layers, leading to a better overall model and thus the slight increase in mIoU scores. It suggests that the model's ability to learn and generalize has been enhanced by the additional guidance provided during training.

Parts 4 & 5: Per class IoUs (1 points)

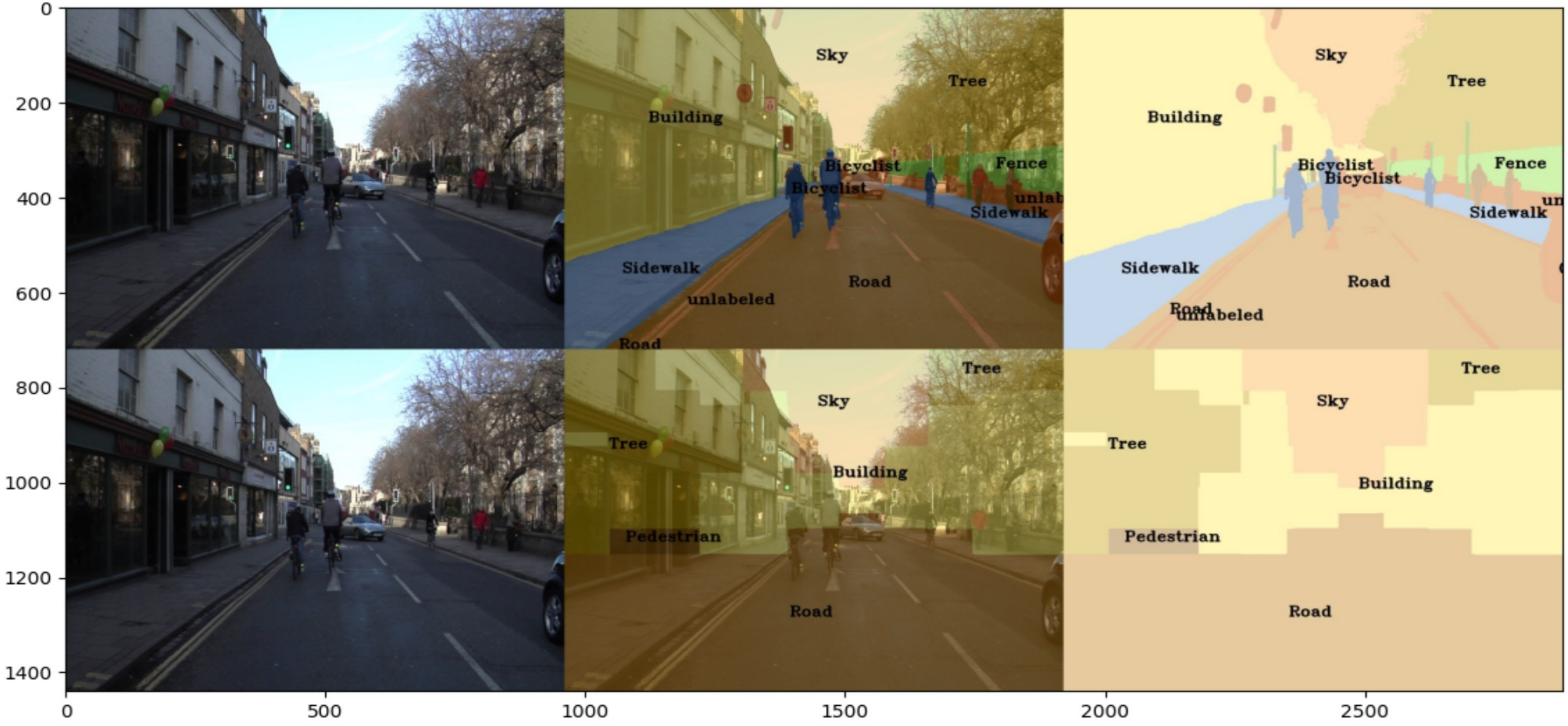
Class Index	Class name	Simple Segmentation Net Class IoU	PSPNet Class IoU
0	Building	0.4290	0.8869
1	Tree	0.2443	0.8514
2	Sky	0.6978	0.9189
3	Car	0.2948	0.8418
4	SignSymbol	0.0000	0.0000
5	Road	0.7026	0.9341
6	Pedestrian	0.0000	0.3774
7	Fence	0.0000	0.4252
8	Column_Pole	0.0000	0.0065
9	Sidewalk	0.0533	0.7760
10	Bicyclist	0.0000	0.6362

Parts 4 & 5: Most difficult classes (3 points)

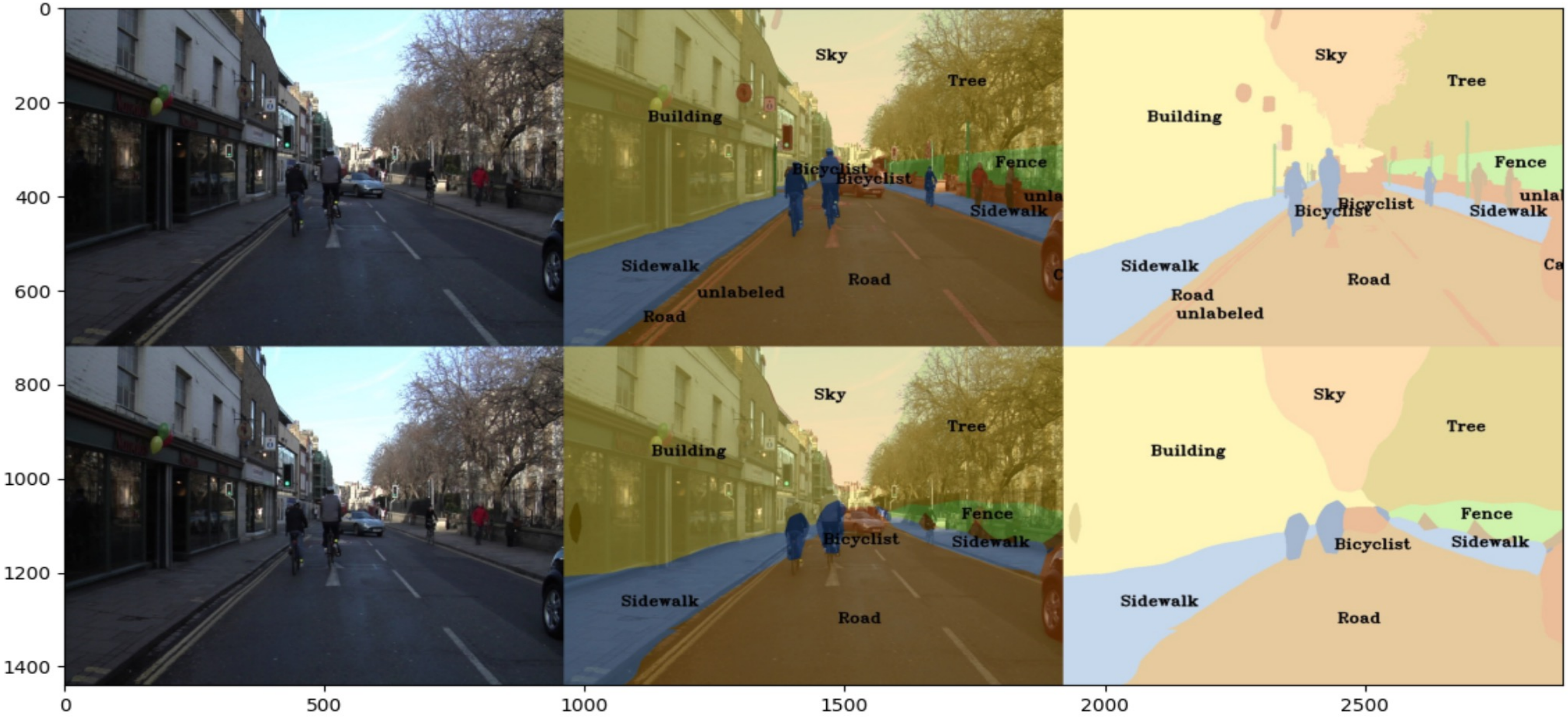
Sign symbols and **column poles** typically present a challenge for semantic segmentation models due to their small size and thin structure, which can be easily missed or inaccurately segmented, particularly in lower-resolution images or if the model lacks fine-grained detail in its feature maps. The variability in appearance of sign symbols, with their diverse shapes, colors, and designs, and column poles, with variations in size, shape, and color, complicates the task of learning a consistent representation, especially when these objects are partially occluded by urban clutter like pedestrians and vehicles. Additionally, their tendency to blend into the background due to similar colors and textures makes it hard to accurately segment their edges and shapes. This issue is further exacerbated by occlusions and varying lighting conditions, such as sunlight glare on reflective surfaces, which obscure critical features necessary for precise segmentation.



Part 4: Simple segmentation net qualitative results (1 point)



Part 5: PSPNet qualitative results (1 point)



Extra Credit: Transfer Learning (1 point)

Report your model's IoU for the Kitti Dataset.

Extra Credit: Transfer Learning (1 point)

Compare the training loss generated when training on Kitti dataset and Camvid dataset. Which decreases at a faster rate? If Camvid or Kitti training loss decreases at a faster rate than the other, why do you think this happened? Or, if the loss decreases at a similar rate, why do you think that is so?

Extra Credit : WandB (3 points)

Add the link to your WandB project and make sure the project is public