

1 Noise Experiments

Noise Introduction Evaluation on Cityscapes dataset: For testing the robustness of each algorithm, we add Gaussian Noise with different standard deviations, to the images in the Cityscapes dataset to observe the effect of noise on all three algorithms. For Momentum Adapt experiments, HRNetV2-W18 is used for both networks in AuxAdapt and Momentum Adapt. Unsuprisingly, As shown by Figure 1, the performance of all three models decreases as we increase the noise. The other trend is that by increasing noise levels, AuxAdapt starts to have lower and lower mIoU until it performance is lower than the network without adaptation. However, Momentum Adapt increases mIoU compare to other two models. For Temporal Consistency, while both method start at a similar position to no-adaptation model, their distances grow increasingly from the base model with more noise. Overall, according to the figures, Momentum Adapt is more capable of handling noise compare to AuxAdapt, especially for higher levels of noise.

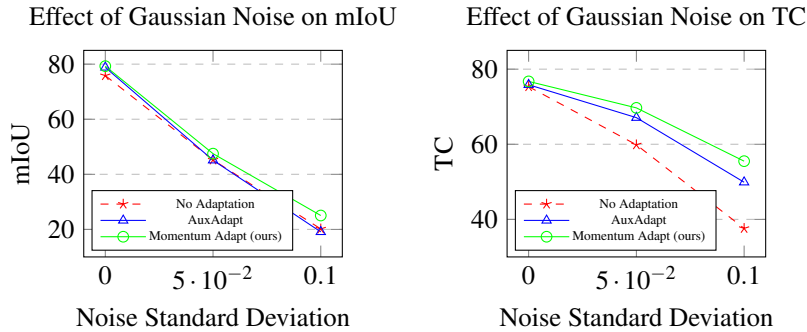


Figure 1: The effect of Noise on TC and mIoU for No adaptation, AuxAdapt, and Momentum Adapt.

Similar to Figure 1, in Figure 2, similar trends can be seen. For this set of experiments, HRNetV2-w48 and HRNetV2-w18 are employed as the MainNet and the AuxNet, respectively. For Aux-Momentum, the differences of performance compared to AuxAdapt is even more significant than Momentum Adapt.

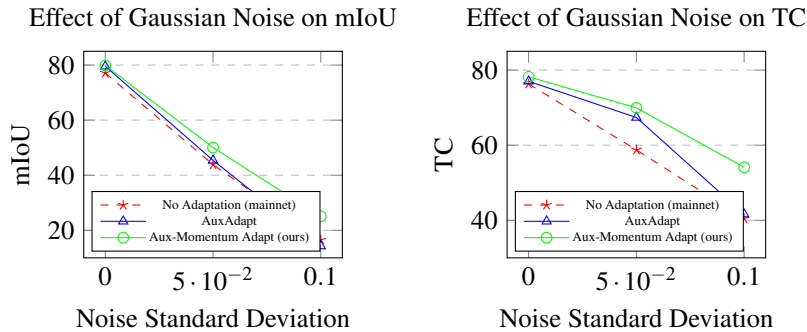


Figure 2: The effect of Noise on TC and mIoU for No adaptation, AuxAdapt, and Momentum Adapt.

2 Qualitative Comparison

Figure 3 shows a sample comparison between AuxAdapt and Momentum Adapt for Cityscapes dataset.

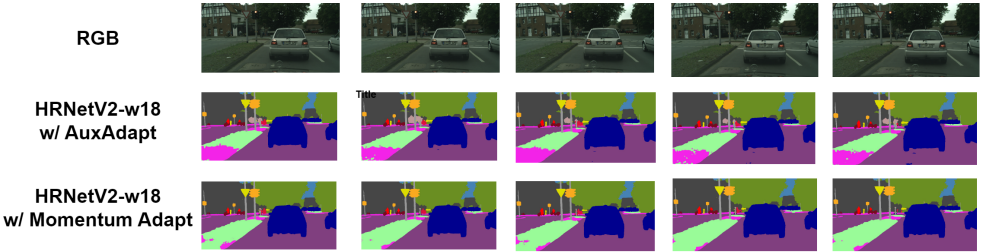


Figure 3: Top row: RGB frames from Cityscapes dataset. Middle row: output of AuxAdapt when both networks are HRNetV2-w18. Bottom row: Output of Momentum Adapt with the same network, showing more continuity in the grass area.

Figure 4 shows the comparison between AuxAdapt and Aux-Momentum Adapt for SceneNet RGBD dataset, where AuxNet is HRNetV2-w18s and the MainNet is HRNetV2-W18. It shows the collapse of AuxAdapt after few iteration.

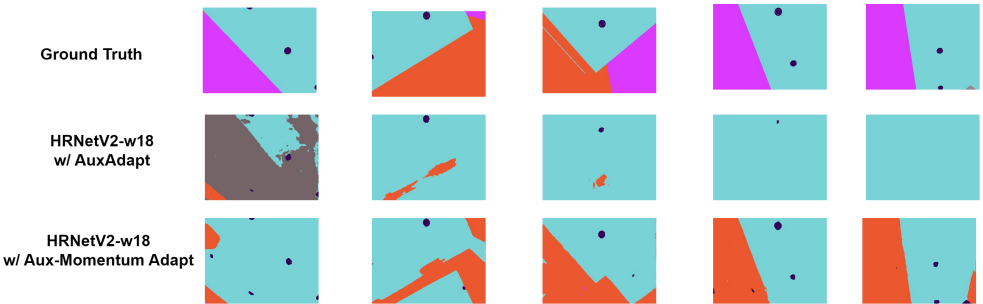


Figure 4: Top row: ground truth labels for frames in SceneNet RGBD dataset. Middle row: output of AuxAdapt when AuxNet is HRNetV2-w18s and MainNet is HRNetV2-w18. The prediction collapses to a constant output after a few iteration. Bottom row: Output of Aux-Momentum Adapt with the same networks as AuxAdapt. The purple class is glass which is miscalssified as the wall class.