Spatial Feedback Learning to Improve Semantic Segmentation in Hot Weather(Supplementary)

Shyam Nandan Rai[1]
shyam.nandan@research.iiit.ac.in
Vineeth N Balasubramanian[2]
vineethnb@iith.ac.in
Anbumani Subramanian[3]
anbumani.subramanian@intel.com
C. V. Jawahar[1]
jawahar@iiit.ac.in

[1] IIIT Hyderabad
[2] IIT Hyderabad
[3] Intel Corporation

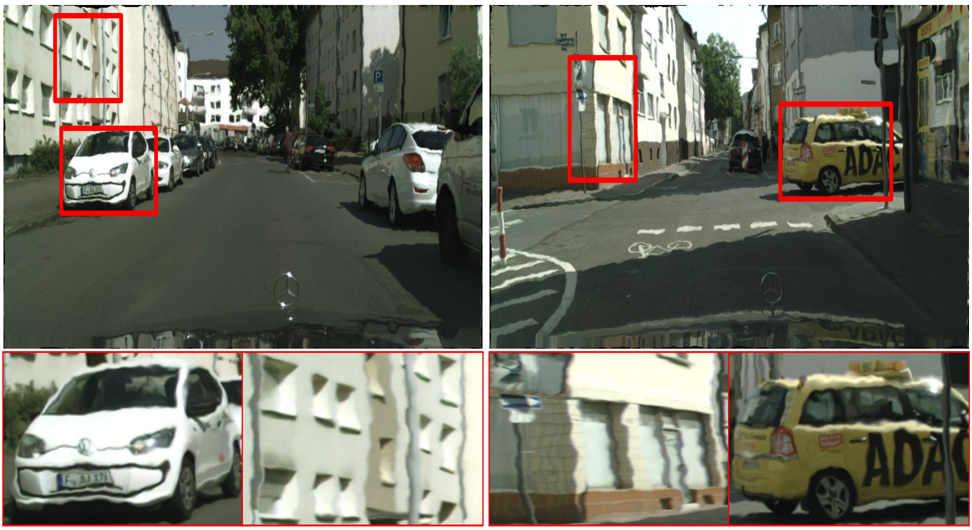

Figure 1: Examples of simulated Cityscapes turbulent images. Image patches in red boxes show the geometrical distortions caused by atmospheric turbulence.

1 Simulation Parameters For Turbulent Images

We create our turbulent dataset consisting of Cityscapes dataset by using Schwartzman *et al*. [?] method. This approach by passes the expensive process of 3D simulation of atmospheric turbulence and performed the same by using a series of 2D image transformation, that was computationally low. The method uses a virtual camera with focal distance of 300 mm and lens diameter as 5.35 cm with pixel size of $4e-3$ mm. The virtual system is mounted at a height of 4m. The value of C_n^2 is $9e-14m^{-2/3}$, which is a parameter to measure atmospheric turbulence. Figure 1, shows the simulated atmospheric turbulent images.

Feedback Input	$I_R^{i_t-1}$	$I_{SP}^{i_t-1} - I_{SP}^{i_t-2}$	$C(I_R^{i_t-1}, I_{SP}^{i_t-1} - I_{SP}^{i_t-2})$	$M(I_R^{i_t-1}, I_{SP}^{i_t-1} - I_{SP}^{i_t-2})$	$M(I_R^{i_t-1}, abs(I_{SP}^{i_t-1} - I_{SP}^{i_t-2}))$
mIoU	56.57	56.91	57.16	57.30	**57.45**

Table 1: **Feedback Input Analysis:** mIoU over various feedback inputs. We find that the multiplication of previously restored image with the absolute difference of output segmentation probabilities performs better than concatenation. M and C denote the multiplication and concatenation of inputs.

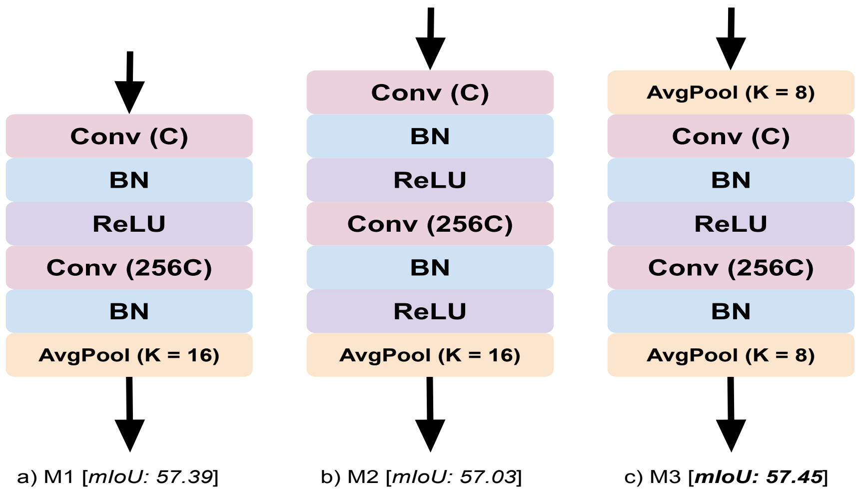

Figure 2: Shows various feedback architecture and its mIoU used in feedback framework. Conv. and C represent a convolutional layer with C output channels. AvgPool and K is an average pooling layer with kernel size K. BN stands for Batch normalization.

2 Additional Analysis on Feedback Framework

2.1 Feedback Input

We show how the appropriate semantic input to the feedback network improve its performance. Table 1 shows that the performance of feedback network in restoring turbulent images. We find that the multiplication of previously restored image $I_R^{i_t-1}$ with the absolute difference of output semantic segmentation probability prediction $abs(I_{SP}^{i_t-1} - I_{SP}^{i_t-2})$ gives better mIoU over concatenation. Also, we can observe individually $I_R^{i_t-1}$ or $I_{SP}^{i_t-1} - I_{SP}^{i_t-2}$ do not give adequate input to the feedback network.

2.2 Feedback Network Architecture

We propose 3 feedback architectures: $M1$, $M2$, and $M3$, which takes the feedback input and adds its output in the restoration network. Among all the models, $M3$ performed the best because it gradually averages the feature by adding average pooling layers at the head and tail of the feedback network instead of adding only an average pooling at the tail of the network showed in $M1$ and $M2$. We also observe that adding a ReLU layer before average pooling in $M2$ reduces the performance of the network as it increases sparsity in the output due to which sufficient feedback information is not passed.

References

[1] Armin Schwartzman, Marina Alterman, Rotem Zamir, and Yoav Y Schechner. Turbulence-induced 2d correlated image distortion. In *ICCP*. IEEE, 2017.