# $C^2$AIR: Consolidated Compact Aerial Image Haze Removal
Supplementary Material

Ashutosh Kulkarni, Shruti S. Phutke, Santosh Kumar Vipparthi, and Subrahmanyam Murala

## Overview

The supplementary material contains:

- Loss Functions
- Network Architectures Included in Ablation Studies
- More Qualitative Results

## 1 Loss Functions

We train the proposed network using $L_1$ loss ($\mathbb{L}_1$), and perceptual loss ($\mathbb{L}_p$). The detailed equations of these loss functions are:

### 1.1 $L_1$ Loss

$\mathbb{L}_1$ loss aims at reduction of the per pixel difference between the output image ($I_O$) and ground truth ($I_{GT}$), which can be formulated as:

$$\mathbb{L}_1 = \|I_O - I_{GT}\|_1 \tag{1}$$

### 1.2 Perceptual Loss

The perceptual loss ($\mathbb{L}_P$) aims at improving the perceptual quality of the output by calculating the differences between output and ground truth at various feature levels. The pretrained layers of VGG-16 [1] model are used as feature space for the loss calculation. The perceptual loss can be represented mathematically as:

$$\mathbb{L}_P = \left\|\Phi_{i\in(3,8,15)}\left(I_O\right) - \Phi_{i\in(3,8,15)}\left(I_{GT}\right)\right\|_1 \tag{2}$$

where, $\Phi_i(\bullet)$ is $i^{th}$ layer of VGG-16 model.

## 2 Network Architectures Included in Ablation Studies

Here, we provide the architectural diagrams of each network setting. Figure S 1 represents the network configurations with additive, concatenation-based and without query modulation. Figure S 2 shows the network diagrams of various fusion settings. of the main manuscript. Figure S 3 displays the various types of Feed-Forward blocks.

## 3 More Qualitative Results

We provide more qualitative results on RICE dataset for aerial haze removal in Figure S 4. We compare the qualitative results with state-of-the-art methods USID [2], MSBDN [3], TSDNet [4], RefineDNet [5], SPA-GAN [6], UFormer [7] and AIDNet [8]. Further, in Figure S 5, we provide the qualitative results in comparison with existing state-of-the-art methods TACL [9] and CLUIE [10] for underwater image enhancement on UIEB dataset. As seen from the results, the proposed method is able to maintain more spatial content and color balance.

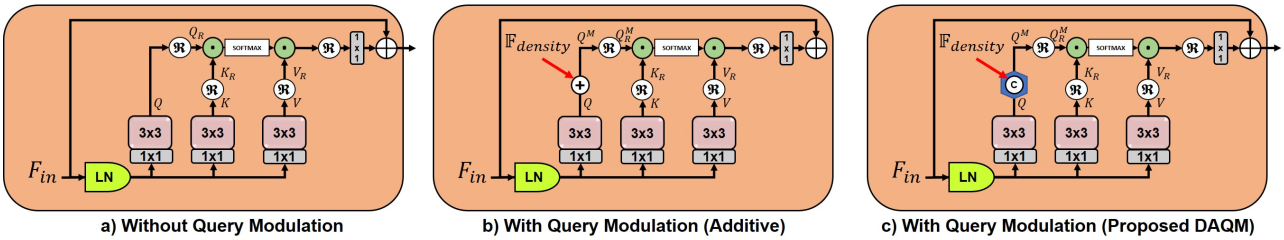


Figure S 1: Network architectures of different types of query modulation mechanisms considered in the ablation study.

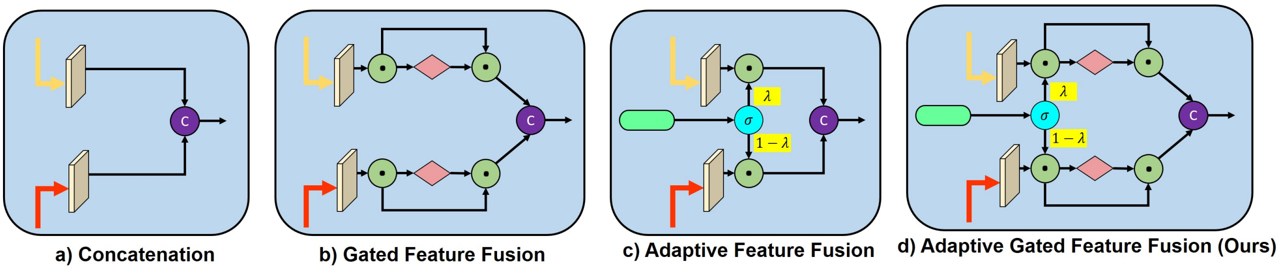


Figure S 2: Network architectures of different types of fusion mechanisms considered in the ablation study.

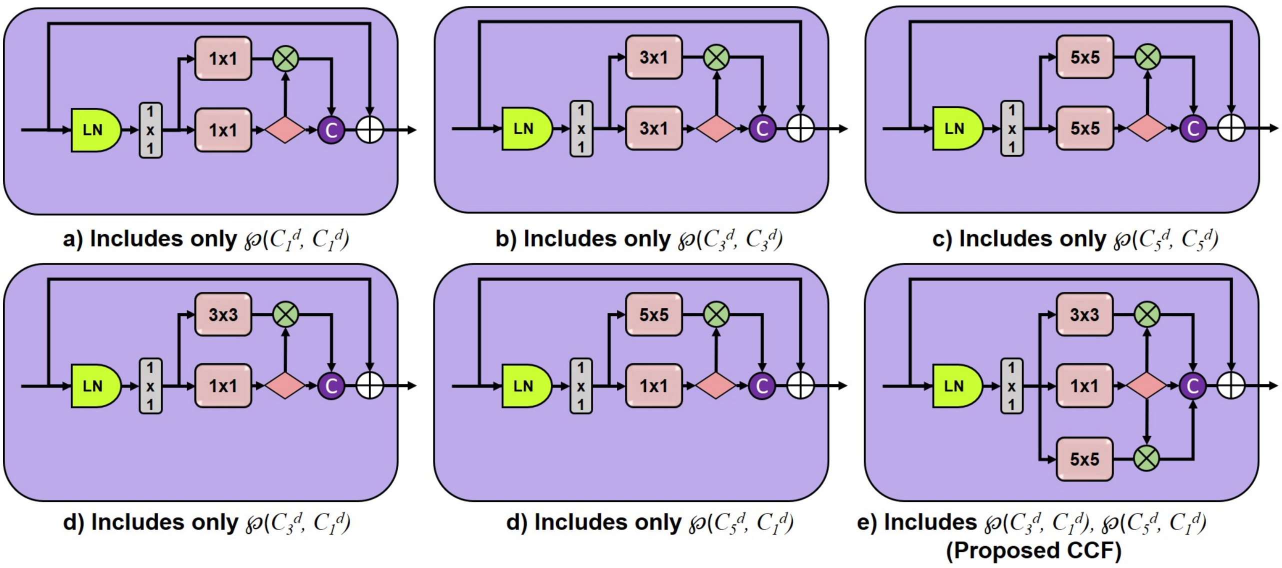


Figure S 3: Network architectures of different types of Feed-Forward blocks considered in the ablation study.


Figure S 4: Qualitative results comparison with existing state-of-the-art methods on RICE dataset for aerial image dehazing.

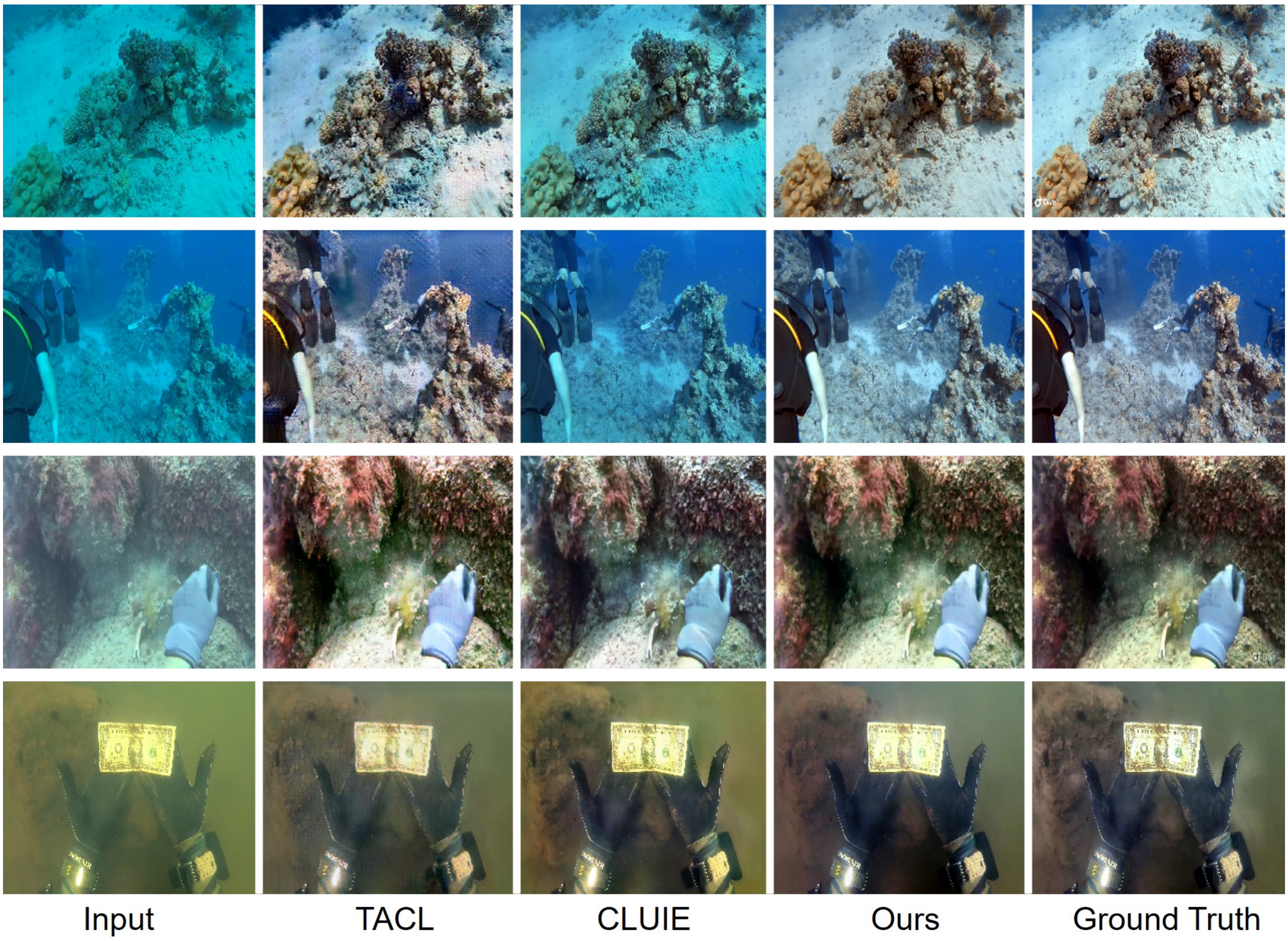


Figure S 5: Qualitative results comparison with existing state-of-the-art methods on the UIEB dataset for underwater image enhancement.

# References


[1] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[2] J. Li, Y. Li, L. Zhuo, L. Kuang, and T. Yu, "Usid-net: Unsupervised single image dehazing network via disentangled representations," *IEEE Transactions on Multimedia*, pp. 1–1, 2022.

[3] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2157–2167.

[4] R. W. Liu, Y. Guo, Y. Lu, K. T. Chui, and B. B. Gupta, "Deep network-enabled haze visibility enhancement for visual iot-driven intelligent transportation systems," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2022.

[5] S. Zhao, L. Zhang, Y. Shen, and Y. Zhou, "Refinednet: A weakly supervised refinement framework for single image dehazing," *IEEE Transactions on Image Processing*, vol. 30, pp. 3391–3404, 2021.

[6] H. Pan, "Cloud removal for remote sensing imagery via spatial attention generative adversarial network," *arXiv preprint arXiv:2009.13015*, 2020.

[7] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general u-shaped transformer for image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 17 683–17 693.

[8] A. Kulkarni and S. Murala, "Aerial image dehazing with attentive deformable transformers," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 6305–6314.

[9] R. Liu, Z. Jiang, S. Yang, and X. Fan, "Twin adversarial contrastive learning for underwater image enhancement and beyond," *IEEE Transactions on Image Processing*, vol. 31, pp. 4922–4936, 2022.

[10] K. Li, L. Wu, Q. Qi, W. Liu, X. Gao, L. Zhou, and D. Song, "Beyond single reference for training: underwater image enhancement via comparative learning," *IEEE Transactions on Circuits and Systems for Video Technology*, 2022.