GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing

*Xiaohong Liu, *Yongrui Ma, Zhihao Shi, Jun Chen
McMaster University

ICCV 2019 Seoul, Korea

McMaster University

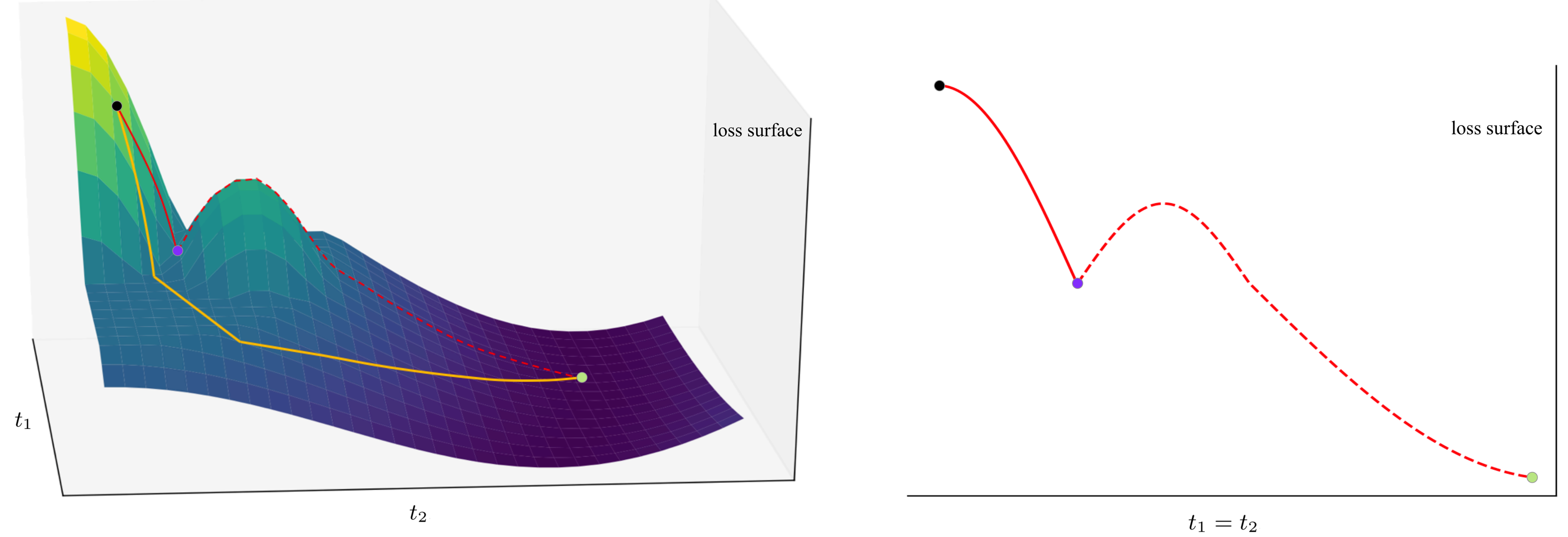
Abstract

We propose an end-to-end trainable Convolutional Neural Network (CNN), named GridDehazeNet, for single image dehazing. The GridDehazeNet consists of three modules: pre-processing, backbone, and post-processing. The trainable pre-processing module can generate learned inputs with better diversity and more pertinent features as compared to those derived inputs produced by hand-selected pre-processing methods. The backbone module implements a novel attention-based multi-scale estimation on a grid network, which can effectively alleviate the bottleneck issue often encountered in the conventional multi-scale approach. The post-processing module helps to reduce the artifacts in the final output. Experimental results indicate that the GridDehazeNet outperforms the state-of-the-arts on both synthetic and real-world images. The proposed hazing method does not rely on the atmosphere scattering model, and we provide an explanation as to why it is not necessarily beneficial to take advantage of the dimension reduction offered by the atmosphere scattering model for image dehazing, even if only the dehazing results on synthetic images are concerned.

Proposed Method

The proposed GridDehazeNet is an end-to-end trainable network with three important features:

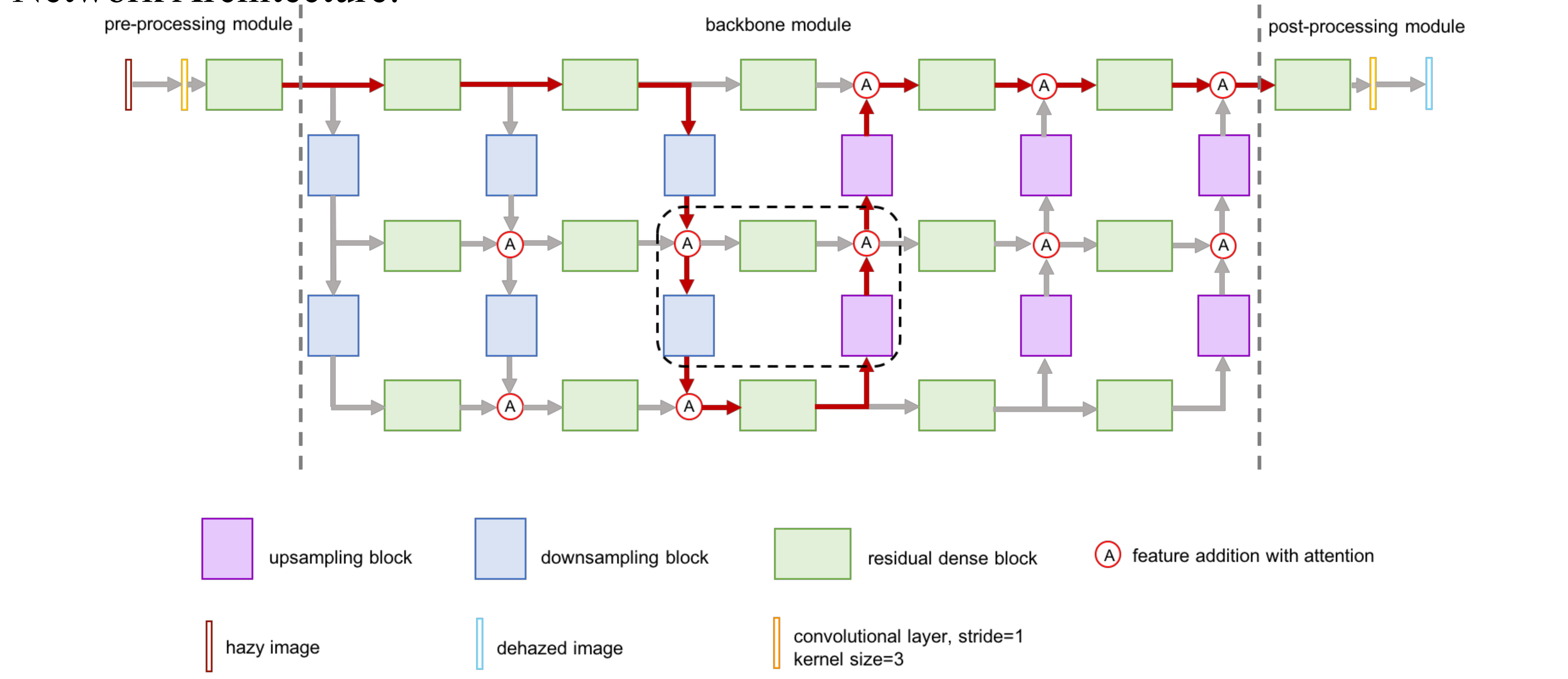
- No reliance on atmosphere scattering model: avoid creating an undesirable loss surface that can be potentially caused by the ASM.



(a) Loss surface (b) Constrained Loss surface

- Trainable pre-processing module: highlight different aspects of a hazy image and make the relevant feature information more evidently exposed.
- Attention-based multi-scale estimation: build dense connections across different scales based on channel-wise attention for more flexible information exchange and aggregation.

Network Architecture:



Experimental Results

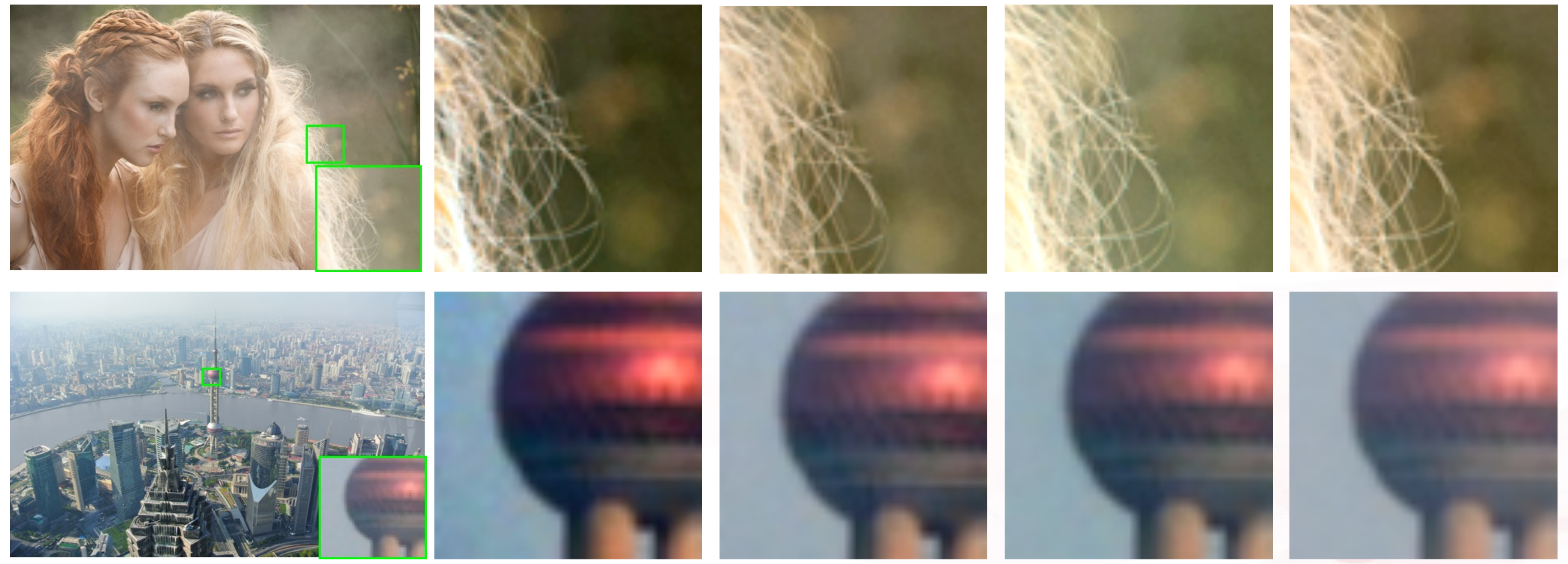
- Quantitative comparisons on SOTS and Sun RGB-D for different methods.

Method	Indoor		Outdoor		Sun RGB-D	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DCP	16.61	0.8546	19.14	0.8605	15.18	0.8191
DehazeNet	19.82	0.8209	24.75	0.9269	23.05	0.8870
MSCNN	19.84	0.8327	22.06	0.9078	23.85	0.9095
AOD-Net	20.51	0.8162	24.14	0.9198	22.51	0.8918
GFN	24.91	0.9186	28.29	0.9621	25.35	0.9250
Ours	**32.16**	**0.9836**	**30.86**	**0.9819**	**28.67**	**0.9599**

- Qualitative comparisons on SOTS indoor (1st row) and outdoor (2nd row) datasets.

Hazy input DCP DehazeNet GFN Ours Ground truth

- Qualitative comparisons on real captured images. Obviously, DCP suffers from severe color distortion (ball of the 2nd row), DehazeNet removes haze effect incompletely, and GFN has limited ability to deal with dense haze and causes color distortion in some cases (girl's hair of the 1st row). Compared with these methods, the proposed GridDehazeNet is more effective in haze removal.



Hazy input DCP DehazeNet GFN Ours

- Quantitative comparisons on SOTS for different estimation strategies. *Indirect* estimates the transmission map and global atmosphere intensity to produce dehazed image. On the contrary, *Direct* generates dehazed one directly from the hazy input.

Estimation	Indoor		Outdoor		SUN RGB-D	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Indirect	30.33	0.9160	30.12	0.9729	27.82	0.9477
Direct	**32.16**	**0.9836**	**30.86**	**0.9819**	**28.67**	**0.9599**

- Quantitative comparisons on SOTS for different types of inputs including the *Original* (original hazy input), *Derived* (human selected pre-processing features such as white balance, contrast enhancement, gamma correction, original and gray-scale hazy image), *Learned* (our learned features). The learned features offer significant diversity gain and have clear advantages over the derived features.

Input	Indoor		Outdoor	
	PSNR	SSIM	PSNR	SSIM
Original	31.48	0.9820	30.33	0.9808
Derived	30.21	0.9799	30.32	0.9778
Learned	**32.16**	**0.9836**	**30.86**	**0.9819**