Bridging the gap between low level vision and high level tasks

任文琦

中国科学院信息工程研究所

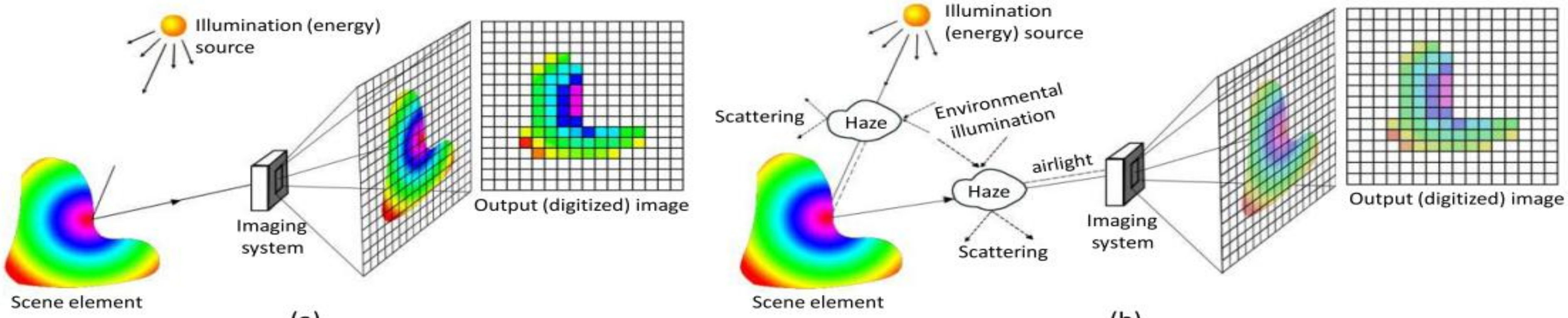
VALSE 2019-09-18

Outline

- Gated fusion network for single image dehazing, CVPR'18
- Benchmarks: RESIDE (dehazing), MPID (deraining)
  - Evaluate current low-level vision algorithms in terms of high-level tasks
  - (Dehazing/Deraining) + Object detection, TIP'19, CVPR'19
- Semi-supervised dehazing/deraining, TIP'19, CVIU'19

Introduction

- Hazy images
  - Low visibility: distance between an object and the observer increases
  - Faint colors: atmosphere color replaces the color of the object



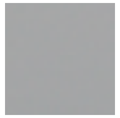
(a) (b)

[1] A fast single image haze removal algorithm using color attenuation prior (Zhu et al. TIP 2015)

Introduction

- Hazy imaging model

t(x): Transmission
d(x): Scene depth
β: medium extinction coefficient



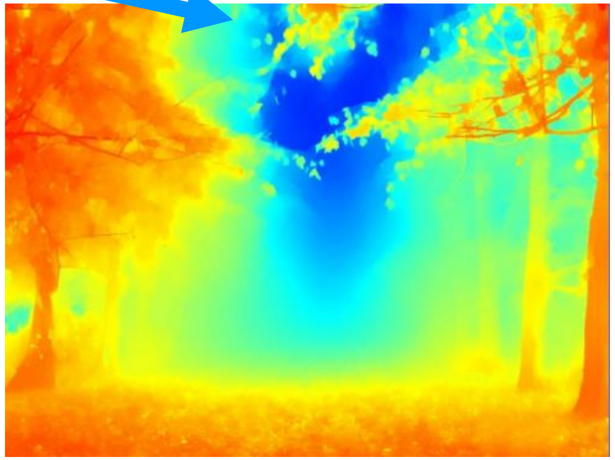
Atmospheric light

$$\boldsymbol{I}(x) = \boldsymbol{J}(x)t(x) + \boldsymbol{A}(1 - t(x))$$

$$t(x) = e^{-\beta d(x)}$$

Hazy image

Scene

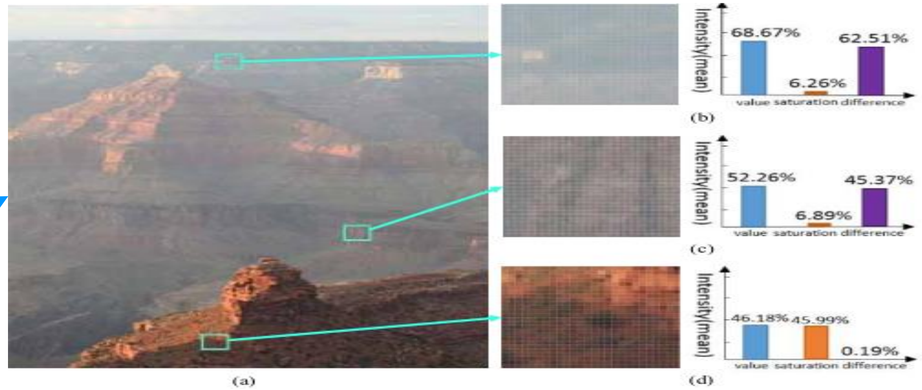


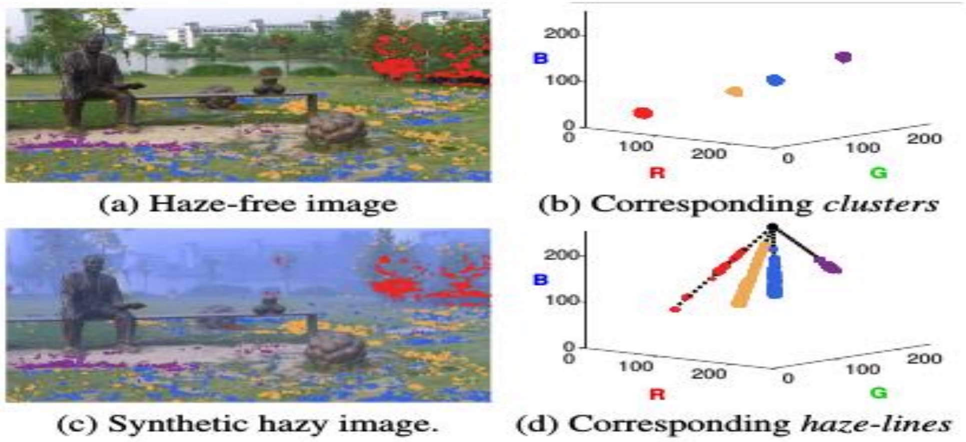
Transmission

Koschmieder, H.: Theorie der horizontalen sichtweite. Beitrage zur Physik der freien Atmosphare (1924)

Related work

- Maximize local contrast, CVPR’08
- Dark channel prior, CVPR’09
- Maximize local saturation, CVPR’14
- Color Attenuation Prior, TIP’15
- Non-local Prior, CVPR’16





Related work

- Multi-scale CNN, ECCV’16
- DehazeNet, TIP’16
- AOD-Net, ICCV’17
- Fusion Network, CVPR’18
- Densely Connected Network, CVPR’18
- CGAN, CVPR’18
- Proximal Dehaze-Net, ECCV’18
- ……

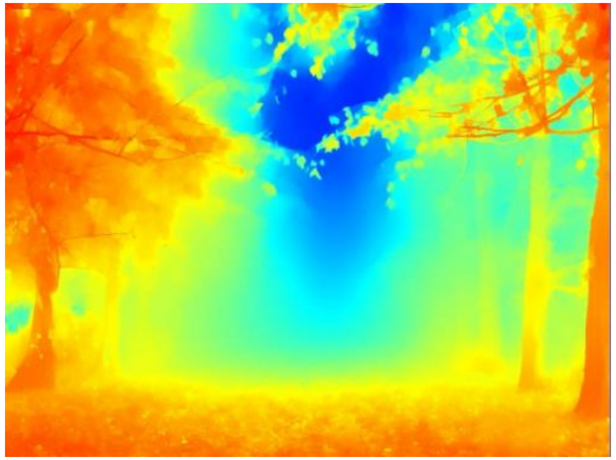
Gated Fusion Network for Single Image Dehazing

W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, M.-H. Yang

CVPR 2018

Motivation

$\boldsymbol{I}(x)$



$t(x)$

$\boldsymbol{J}(x)$

$$\boldsymbol{I}(x) = \boldsymbol{J}(x)t(x) + \boldsymbol{A}(1 - t(x))$$

Motivation

- End-to-end dehazing network

Input

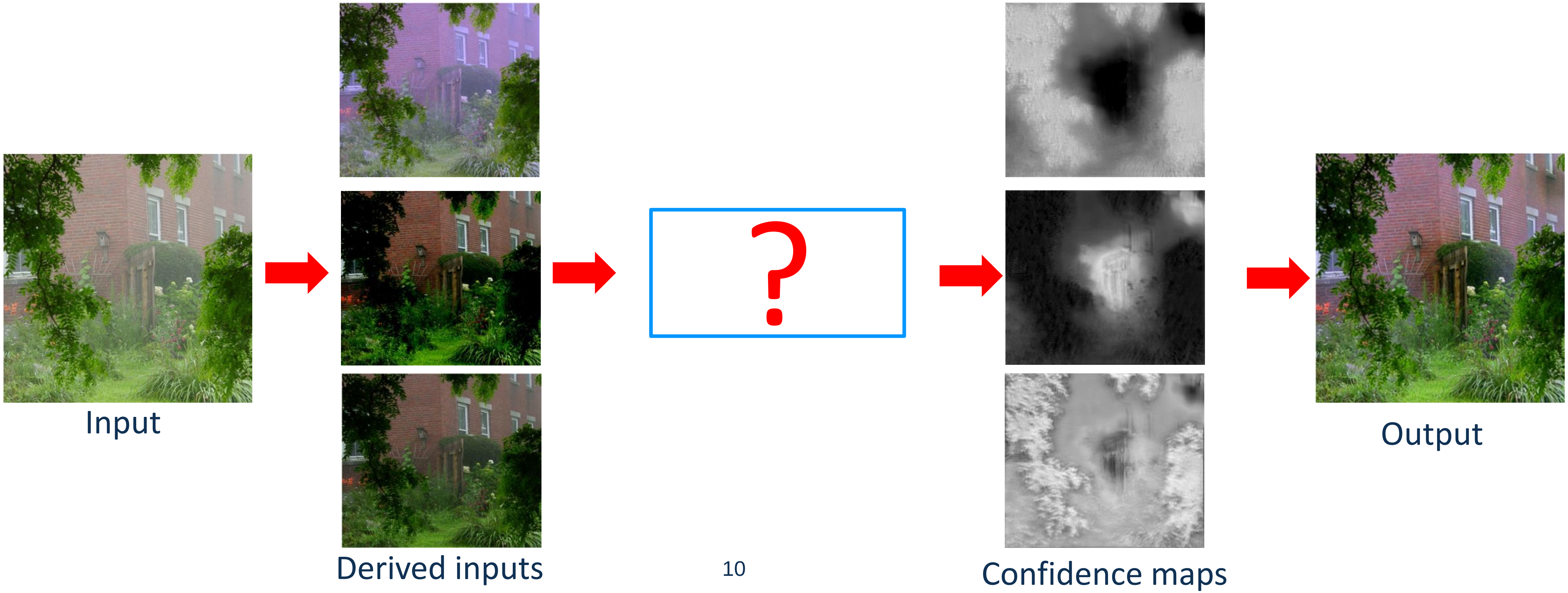
Network

Output

Motivation

Two major factors in hazy images:

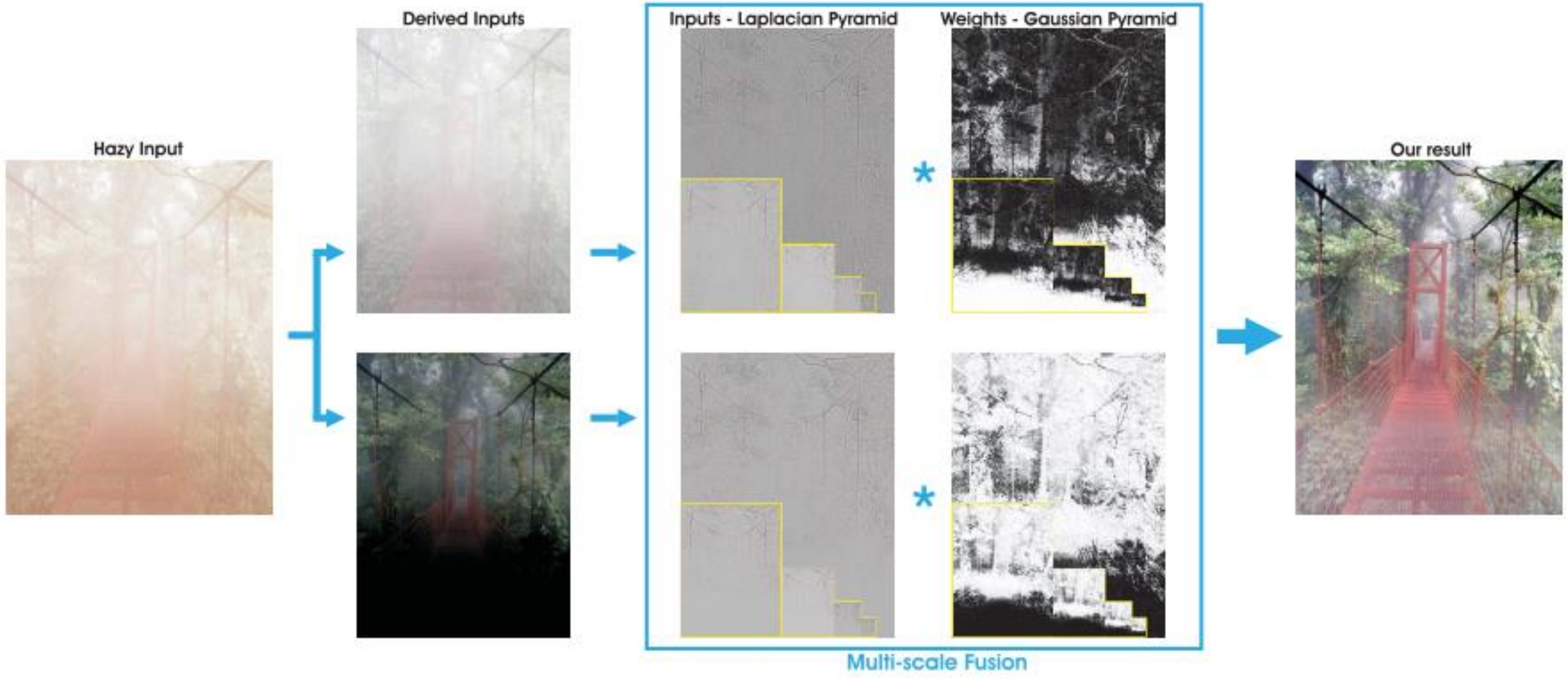
- Color cast introduced by the atmospheric light (White Balance)
- Lack of visibility due to attenuation (Gamma Correct, Contrast Enhance)



Motivation

Two major factors in hazy images:

- Color cast introduced by the atmospheric light (White Balance)
- Lack of visibility due to attenuation (Contrast Enhance)

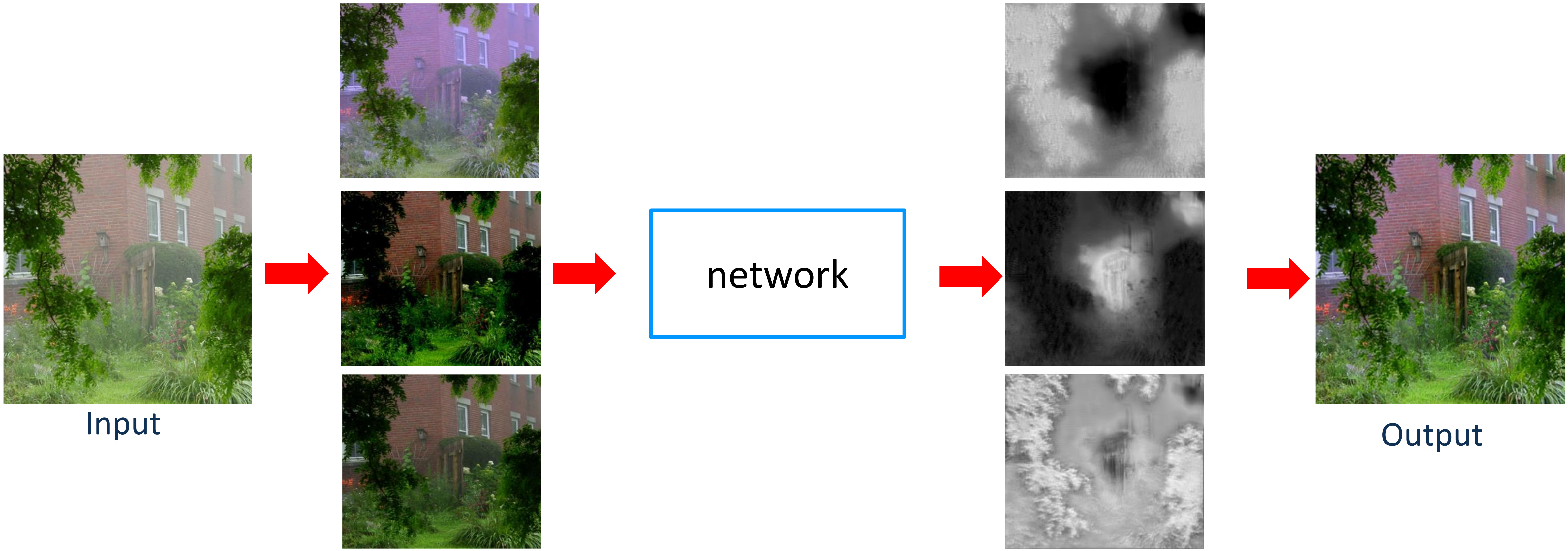


Codruta Orniana Ancuti and Cosmin Ancuti, Single Image Dehazing by Multi-Scale Fusion, TIP 2013

Motivation

Two major factors in hazy images:

- Color cast introduced by the atmospheric light (White Balance)
- Lack of visibility due to attenuation (Gamma Correct, Contrast Enhance)

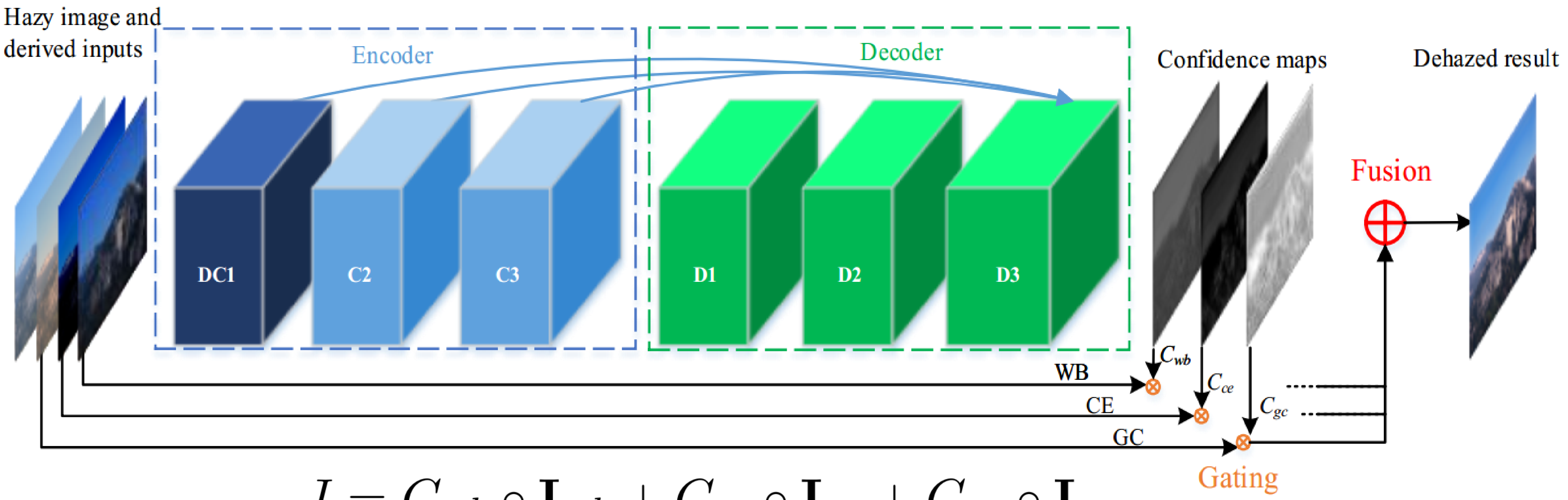


Derived inputs

- **White Balanced:** aims to eliminate chromatic casts caused by the atmospheric color
- **Contrast enhance:** extract visible information (denser haze regions)
- **Gamma correct:** extract visible information (light haze regions)

Input | White Balanced | Contrast Enhance | Gamma Correct

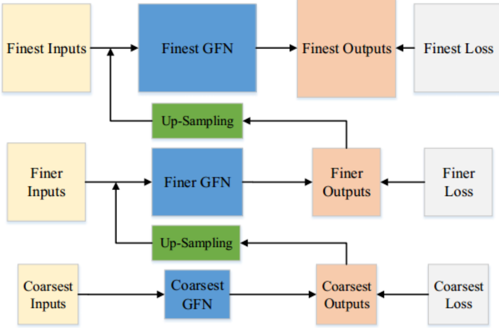
Network

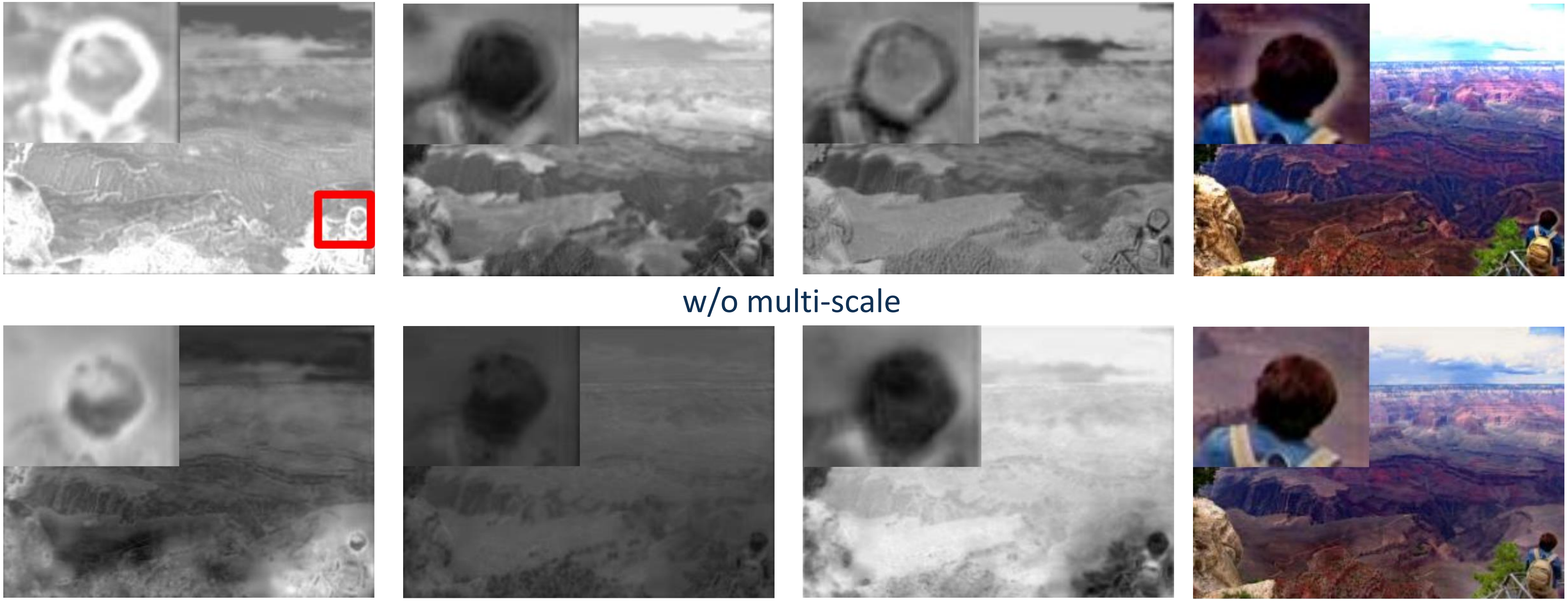


$$J = C_{wb} \circ \mathbf{I}_{wb} + C_{ce} \circ \mathbf{I}_{ce} + C_{gc} \circ \mathbf{I}_{gc}$$

- Use dilated convolution to enlarge receptive fields in the encoder
- Skip shortcuts are connected from the encoder to decoder
- Three derived inputs are weighted by the three confidence maps learned by our network
- Use adversarial loss and multi-scale to further improve results

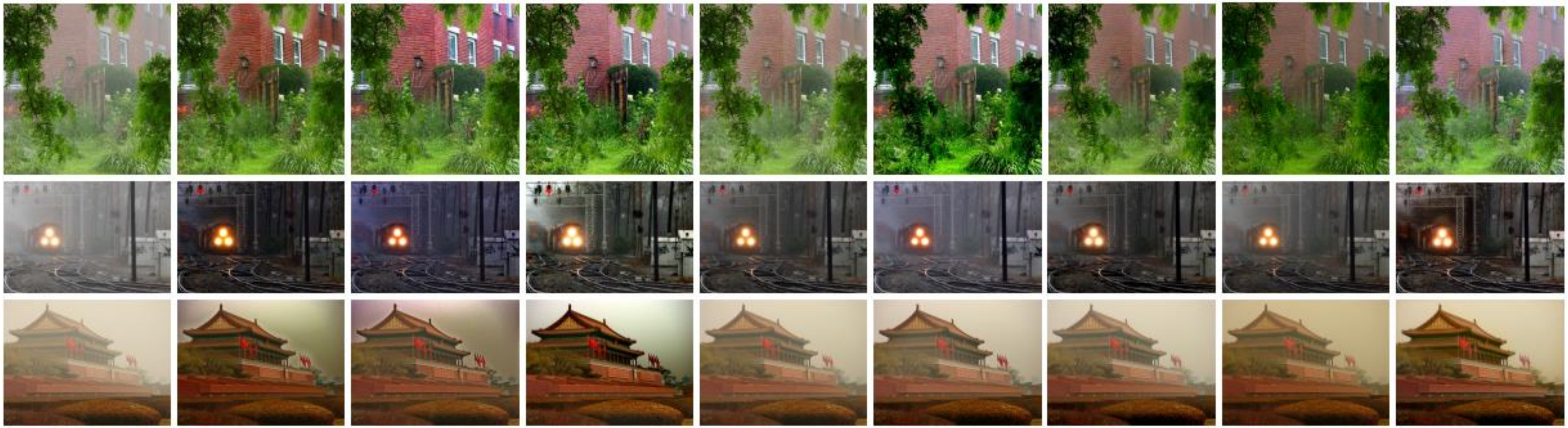
Multi-Scale Refinement





Results

SOTS Set	DCP	CAP	NLD	MSCNN	DehazeNet	AOD-Net	Ours
PSNR	16.62	19.05	17.29	17.57	21.24	19.06	**22.30**
SSIM	0.82	0.84	0.75	0.81	0.85	0.85	**0.88**



(a) Hazy inputs (b) DCP [13] (c) BCCR [23] (d) NLD [2] (e) CAP [44] (f) MSCNN [32] (g) DehazeNet [3] (h) AOD-Net [19] (i) GFN

Figure 6. Qualitative comparison of different methods on real-world images. Please zoom-in to see the differences.

Results: Derived inputs

- More inputs (e.g., other parameters) may be better for final dehazing
  - Original input **(O)**
  - White Balanced **(WB)**
  - Contrast Enhance **(CE)**
  - Gamma Correct **(GC)**

	O	**O+CE+GC**	**O+WB+CE**	**O+WB+GC**	**O+WB+GC+CE**
PSNR	19.16	18.99	19.32	21.02	22.41
SSIM	0.76	0.80	0.79	0.81	0.81

Gated Fusion Network for Single Image Dehazing

- Demonstrate the effectiveness of a gated fusion network for single image dehazing by leveraging the derived inputs.

- Learn the confidence maps to combine three derived input images into a single one by keeping only the most significant features of them.

- Train the proposed model with a multi-scale approach to eliminate the halo artifacts that hurt image dehazing.

Code available at: https://github.com/rwenqi/GFN-dehazing

Comprehensive Benchmark Analysis

REalistic Single-Image DEhazing (**RESIDE**) TIP’19

Multi-Purpose Image Deraining (**MPID**) CVPR’19

492 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 28, NO. 1, JANUARY 2019

Benchmarking Single-Image Dehazing and Beyond

Boyi Li, Wenqi Ren, *Member, IEEE*, Dengpan Fu, Dacheng Tao, *Fellow, IEEE*,
Dan Feng, *Associate Member*, Wenjun Zeng, *Fellow, IEEE*, and Zhangyang Wang, *Member, IEEE*

Single Image Deraining: A Comprehensive Benchmark Analysis

Siyuan Li[1]*, Iago Breno Araujo[2]*, Wenqi Ren[3]†, Zhangyang Wang[4]†, Eric K. Tokuda[2],
Roberto Hirata Junior[2], Roberto Cesar-Junior[2], Jiawan Zhang[1], Xiaojie Guo[1], Xiaochun Cao[3]
[1]*Tianjin University* [2]*University of Sao Paulo* [3]*SKLOIS, IIE, CAS* [4]*Texas A&M University*
https://github.com/lsy17096535/Single-Image-Deraining

Evaluation criteria in existing algorithms

- Synthetic images: PSNR/SSIM
  - Small scale images
  - insufficient for human perception quality and machine vision effectiveness
- Real images: visual comparison
  - Show about ten real images
  - No-reference metrics

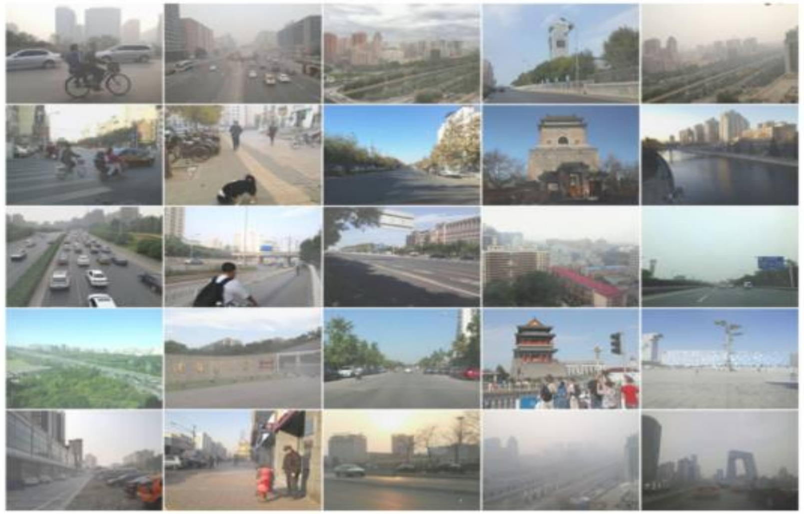
Examples in RESIDE

Three different sets of evaluation criteria:

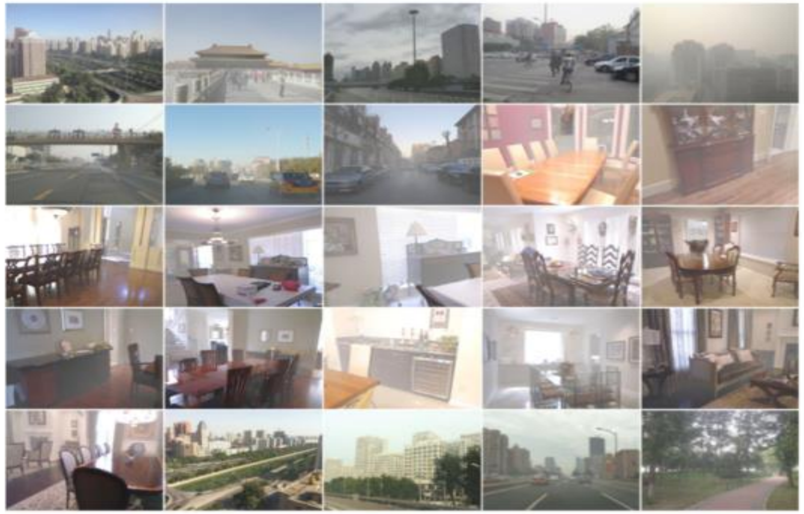
- objective (PNSR, SSIM + no-reference metrics),
- subjective (human rating),
- task-driven (whether or how well dehazed results benefits machine vision, e.g., object detection)

RESIDE(Standard)	
Subset	*Number of Images*
Indoor Training Set (ITS)	13,990
Synthetic Objective Testing Set (SOTS)	500
Hybrid Subjective Testing Set (HSTS)	20
RESIDE-β	
Subset	*Number of Images*
Outdoor Training Set (OTS)	313,950
Real-world Task-driven Testing Set (RTTS)	4,322

(a) ITS



(b) OTS



(c) SOTS

(d) RTTS

(e) HSTS. Top row: 10 synthetic hazy images; Bottom row: 10 realistic hazy images.

Examples in MPID: Multi-Purpose Image Deraining

Rain streak | Raindrops | Rain and mist

$$\mathbf{R}_s = \mathbf{B} + \mathbf{S}. \quad (1)$$

(a) Synthetic rainy images

$$\mathbf{R}_m = \mathbf{B} \odot t + A\,(1 - t) + \mathbf{S}, \quad (3)$$

Rain streak | Raindrops | Rain and mist

(b) Real-world rainy images

Examples in MPID: Multi-Purpose Image Deraining

(c) Rain in driving (RID)

(d) Rain in surveillance (RIS)

Table 2. Object Statistics in RID and RIS sets.

	Categories	*Car*	*Person*	*Bus*	*Bicycle*	*Motorcycle*
2495	**RID Set**	7332	1135	613	268	968
	Categories	*Car*	*Person*	*Bus*	*Truck*	*Motorcycle*
2048	**RIS Set**	11415	2687	488	673	275

RESIDE Result Analysis: Objective/Visual Quality

Table 4. Average subjective scores, as well as full- and no-reference evaluations results, of dehazing results on HSTS.

	DCP [12]	FVR [41]	BCCR [24]	GRM [6]	CAP [46]	NLD [3]	DehazeNet [5]	MSCNN [32]	AOD-Net [16]
	Synthetic images								
Clearness	1.26	0.18	0.62	0.75	0.50	1	0.29	1.22	0.86
Authenticity	0.78	0.14	0.50	0.95	0.86	1	1.94	0.54	1.41
PSNR	17.27	15.68	16.61	20.48	22.88	18.92	26.94	20.53	23.41
SSIM	0.7210	0.7157	0.6947	0.7631	0.8223	0.7411	0.8758	0.7893	0.8616
SSEQ	86.15	85.68	85.60	78.43	85.32	86.28	86.01	85.56	86.75
BLIINDS-II	90.70	87.65	91.05	82.30	85.75	85.30	87.15	88.70	87.50
	Real-world images								
Clearness	0.39	0.46	0.45	0.75	1	0.54	1.16	1.29	1.05
Authenticity	0.17	0.20	0.18	0.62	1	0.15	1.03	1.27	1.07
SSEQ	68.65	67.75	66.63	70.19	67.67	67.96	68.34	68.44	70.05
BLIINDS-II	69.35	72.10	68.55	79.60	63.55	70.80	60.35	62.65	74.75

- PSNR and SSIM appear to be less reliable metrics for dehazing perceptual quality, and are especially poor to reflect "clearness"
- There is certain inconsistency (domain gap) between synthetic and real-world data
- CNN-based dehazing show promising real-world performance (even training data has domain gap)
- MSCNN and AOD-Net achieve good trade-off on clearness v.s. authenticity for real-world dehazing
- Standard no-reference metrics are only roughly aligned with human subjective perception in dehazing

24

Benchmark Result Analysis: “Detection as a Metric”

- We propose a task-driven metric that captures more high-level semantics, and the object detection performance on the dehazed/derained images as a brand-new evaluation criterion for dehazing/deraining realistic images.

RESIDE Result Analysis: “Detection as a Metric”

TABLE VIII

ALL DETECTION RESULTS ON RTTS(IN %), PLEASE NOTE THAT, THE MODEL USED IN FRCNN IS TRAINED ON VOC2007_TRAINVAL DATASET, WHILE THE MODELS USED IN YOLO-V2 AND SSDS ARE TRAINED ON VOC2007_TRAINVAL + VOC2012_TRAINVAL.

		Haze	DCP [9]	FVR [10]	BCCR [11]	GRM [12]	CAP [13]	NLD [14]	DehazeNet [15]	MSCNN [16]	AOD [17]
mAP	FRCNN [52]	37.58	40.58	35.01	41.56	28.90	39.63	40.03	40.54	41.34	37.47
	YOLO-V2 [53]	40.37	39.81	38.06	40.65	29.41	39.80	39.93	40.10	40.76	40.53
	SSD-300 [54]	50.26	49.40	47.04	51.57	35.59	50.31	49.84	50.14	51.82	49.77
	SSD-512 [54]	55.55	55.71	52.29	57.17	39.18	55.70	54.99	55.40	56.88	55.29
Person	FRCNN [52]	60.84	61.54	57.72	64.51	50.22	61.29	60.53	61.40	61.43	61.22
	YOLO-V2 [53]	61.24	61.14	60.00	61.16	50.13	61.24	60.49	61.16	61.30	61.20
	SSD-300 [54]	68.60	68.18	66.36	69.12	53.91	68.78	66.96	68.18	69.20	68.28
	SSD-512 [54]	72.58	72.72	69.45	73.34	56.74	72.50	71.20	72.34	73.13	72.62
Bicycle	FRCNN [52]	40.72	40.77	38.76	44.57	30.71	40.48	40.21	40.68	41.69	40.33
	YOLO-V2 [53]	44.63	43.39	40.08	43.66	28.81	42.65	43.56	42.34	43.53	44.55
	SSD-300 [54]	54.92	51.36	49.35	53.33	34.48	53.38	53.42	53.08	55.73	54.18
	SSD-512 [54]	58.45	56.70	54.57	58.57	36.70	57.49	56.38	57.50	58.76	57.91
Car	FRCNN [52]	35.18	42.15	34.74	42.69	26.30	41.52	42.30	41.74	42.61	35.13
	YOLO-V2 [53]	39.39	38.93	37.22	39.88	29.91	39.03	38.96	39.35	40.00	39.49
	SSD-300 [54]	54.14	54.98	50.81	56.32	40.21	55.08	54.98	55.27	56.32	54.62
	SSD-512 [54]	63.05	64.95	61.54	65.80	47.79	64.15	65.04	64.21	65.22	64.05
Bus	FRCNN [52]	20.90	24.18	19.06	24.66	14.81	24.74	23.74	25.20	25.25	20.56
	YOLO-V2 [53]	20.57	19.34	19.42	20.01	12.86	18.90	18.22	19.07	19.63	19.09
	SSD-300 [54]	30.13	30.87	30.98	33.70	19.72	30.90	30.43	30.86	32.26	29.42
	SSD-512 [54]	34.60	36.51	33.47	37.69	22.81	35.47	34.31	35.18	37.42	34.13
Motorbike	FRCNN [52]	30.24	34.25	24.78	34.34	22.44	30.10	33.36	33.70	35.72	30.09
	YOLO-V2 [53]	37.84	36.23	33.59	38.54	25.33	37.10	38.40	38.59	39.33	38.31
	SSD-300 [54]	43.48	41.61	37.72	45.38	29.63	43.41	43.40	43.30	45.60	42.35
	SSD-512 [54]	49.08	47.69	42.40	50.46	31.85	48.89	48.04	47.79	49.87	47.76

TABLE IX

AVERAGE NO-REFERENCE METRICS OF DEHAZED RESULTS ON RTTS.

	DCP [9]	FVR [10]	BCCR [11]	GRM [12]	CAP [13]	NLD [14]	DehazeNet [15]	MSCNN [16]	AOD-Net [17]
SSEQ	62.87	63.59	63.31	58.64	60.66	59.37	60.01	62.31	65.35
BLIINDS-II	68.34	67.68	74.07	54.54	65.15	68.32	52.54	56.59	71.05

MPID Result Analysis: Objective/Visual Quality

Full- and no-reference evaluations on **synthetic** rainy images

	Degraded	GMM [27]	JORDER [32]	DDN [6]	CGAN [33]	DID-MDN [8]	DeRaindrop [7]
	rain streak						
PSNR	25.95	26.88	26.26	**29.39**	21.86	26.80	/
SSIM	0.7565	0.7674	**0.8089**	0.7854	0.6277	0.8028	/
SSEQ	70.24	67.46	73.70	**75.95**	70.02	60.05	/
NIQE	5.4529	4.4248	4.2337	**3.9834**	4.6189	4.8122	/
BLINDS-II	78.89	75.95	84.21	**91.71**	79.29	67.90	/
	raindrops						
PSNR	25.40	24.85	27.52	25.23	21.35	24.76	**31.57**
SSIM	0.8403	0.7808	0.8239	0.8366	0.7306	0.7930	**0.9023**
SSEQ	78.48	64.73	**84.32**	77.62	63.15	58.42	72.42
NIQE	3.8126	5.1098	4.3278	4.1462	**3.3551**	4.1192	5.0047
BLINDS-II	92.50	75.95	88.05	91.95	73.85	64.70	**96.45**
	rain and mist						
PSNR	26.84	29.37	30.37	**32.98**	22.44	28.77	/
SSIM	0.8520	0.8960	0.9262	**0.9350**	0.7636	0.8430	/
SSEQ	**72.37**	65.39	70.55	69.80	68.71	65.33	/
NIQE	3.4548	3.2117	2.8595	2.9970	**2.8336**	3.0871	/
BLINDS-II	82.95	74.90	83.75	**85.75**	80.20	76.35	/

No-reference evaluations on **real** rainy images

	Degraded	GMM [27]	JORDER [32]	DDN [6]	CGAN [33]	DID-MDN [8]	DeRaindrop [7]
	rain streak						
SSEQ	65.77	61.63	**64.00**	63.51	59.32	55.11	/
NIQE	3.5236	**3.2117**	3.5371	3.5811	3.5374	5.1255	/
BLINDS-II	78.04	75.54	82.62	**85.81**	78.42	66.65	/
	raindrops						
SSEQ	78.23	64.77	69.26	67.62	62.18	60.65	**79.83**
NIQE	3.8229	4.3801	3.6579	3.8290	4.4692	4.5631	**3.5953**
BLINDS-II	84.46	71.21	80.04	77.75	66.29	66.63	**87.13**
	rain and mist						
SSEQ	73.86	59.51	65.18	64.56	**70.04**	63.85	/
NIQE	3.2602	4.4808	3.3238	3.7261	**2.9532**	3.2260	/
BLINDS-II	84.00	62.70	78.62	81.67	**84.91**	76.08	/

- There is certain inconsistency (domain gap) between synthetic and real-world data

MPID Result Analysis: “Detection as a Metric”

Detection results (mAP) on the **RID** and **RIS** sets.

		Rainy	JORDER [32]	DDN [6]	CGAN [33]	DID-MDN [8]	DeRaindrop [7]
RID	FRCNN [44]	16.52	16.97	18.36	**23.42**	16.11	15.58
	YOLO-V3 [45]	**27.84**	26.72	26.20	23.75	24.62	24.96
	SSD-512 [46]	**17.71**	17.06	16.93	16.71	16.70	16.69
	RetinaNet [47]	**23.92**	21.71	21.60	19.28	20.08	19.73
RIS	FRCNN [44]	**22.68**	21.41	20.76	18.02	18.93	19.97
	YOLO-V3 [45]	**23.27**	20.45	21.80	18.71	21.50	20.43
	SSD-512 [46]	8.19	7.94	**8.29**	7.10	8.21	8.13
	RetinaNet [47]	**12.81**	10.71	10.39	9.36	10.33	10.85

A New Benchmark for Single Image Dehazing

Dataset, code, results are available at:

RESIDE: https://sites.google.com/view/reside-dehaze-datasets

MPID: https://github.com/lsy17096535/Single-Image-Deraining

Semi-Supervised Image Dehazing

Lerenhan Li, Yunlong Dong, Wenqi Ren, Jinshan Pan, Changxin Gao, Nong Sang, Ming-Hsuan Yang

TIP 2019, accept

Proposed semi-supervised dehazing network

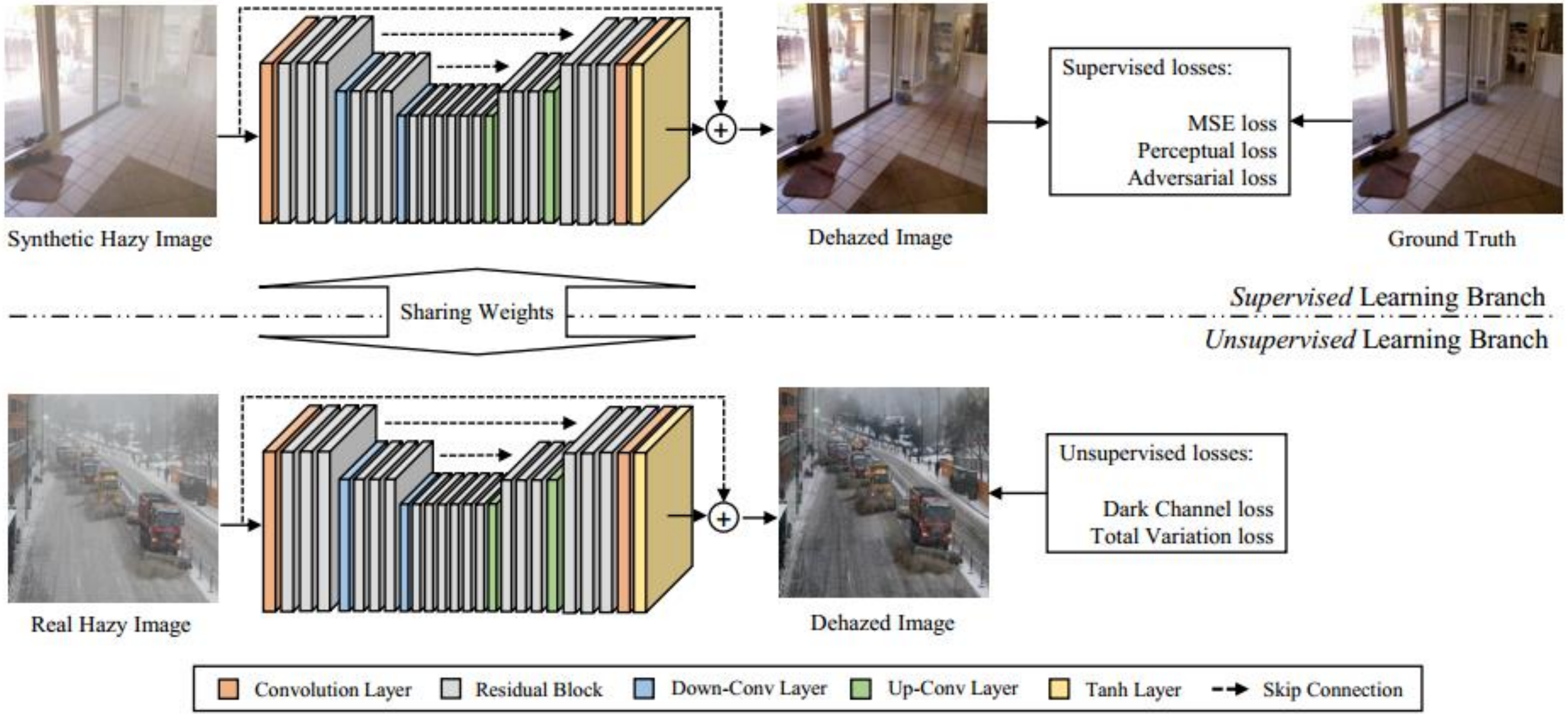


Fig. 2. Proposed semi-supervised learning framework for single image dehazing. The proposed method consists of two branches sharing the same weights. The supervised branch is trained using labeled synthetic data and loss functions based on mean squared, perceptual, and adversarial errors. The unsupervised branch is trained using unlabeled real data and loss functions based on dark channel loss and total variation.

Training details

- **Supervised loss on synthetic images:**
  - Euclidean loss of images and features between dehazed results and ground truths

$$L_c = \frac{1}{N_l}\sum_{i=1}^{N_l}\left\|\mathbf{J}_i - \hat{\mathbf{J}}_i\right\|_2,$$

$$L_p = \frac{1}{N_l}\sum_{i=1}^{N_l}\left\|\mathbf{F}_{J_i} - \mathbf{F}_{\hat{J}_i}\right\|_2,$$

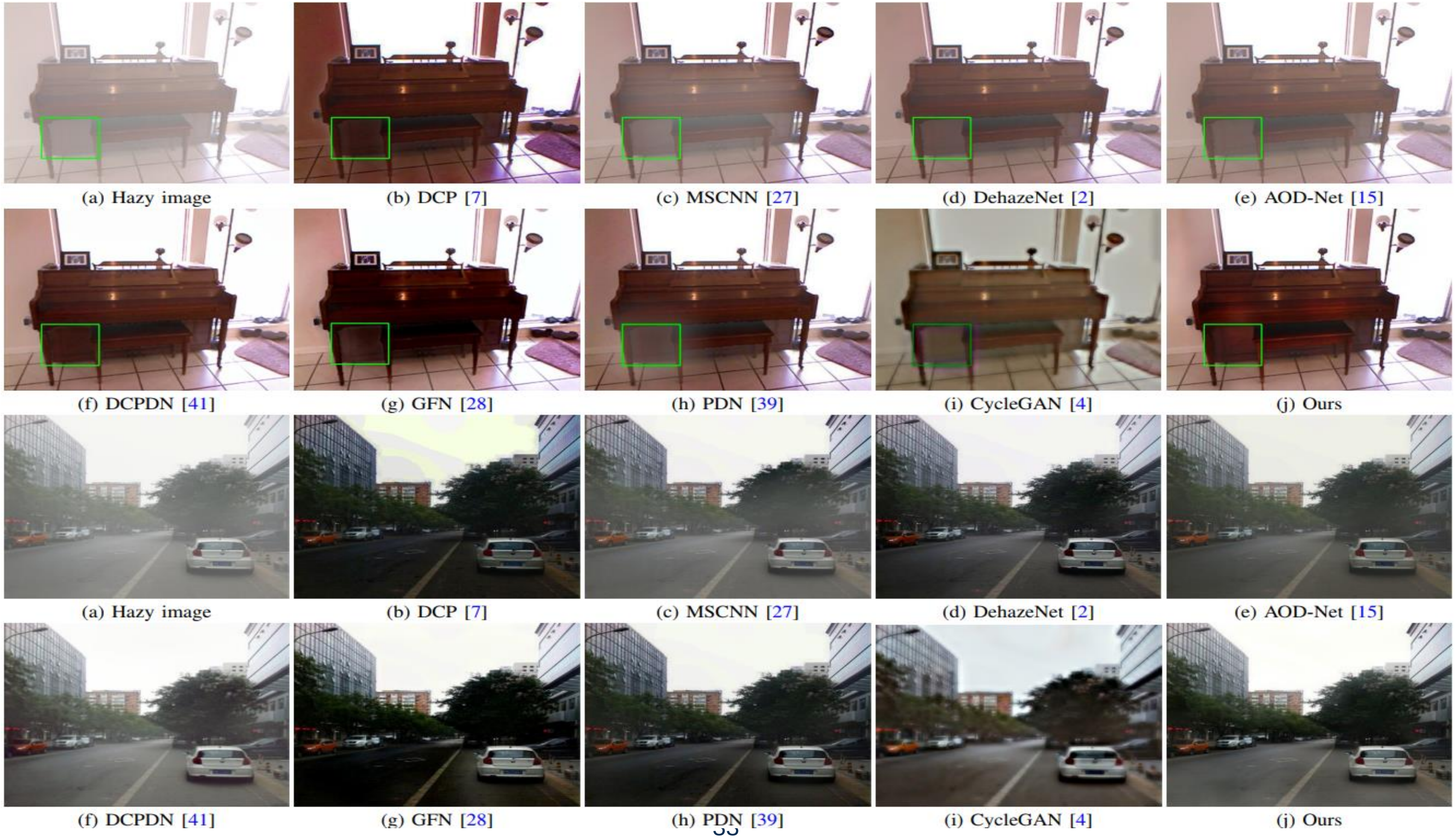
- **Unsupervised loss on real images**:
  - Total variation loss

$$L_t = \frac{1}{N_u}\sum_{i=1}^{N_u}(\|\nabla_h \mathbf{J}_i\|_1 + \|\nabla_v \mathbf{J}_i\|_1),$$

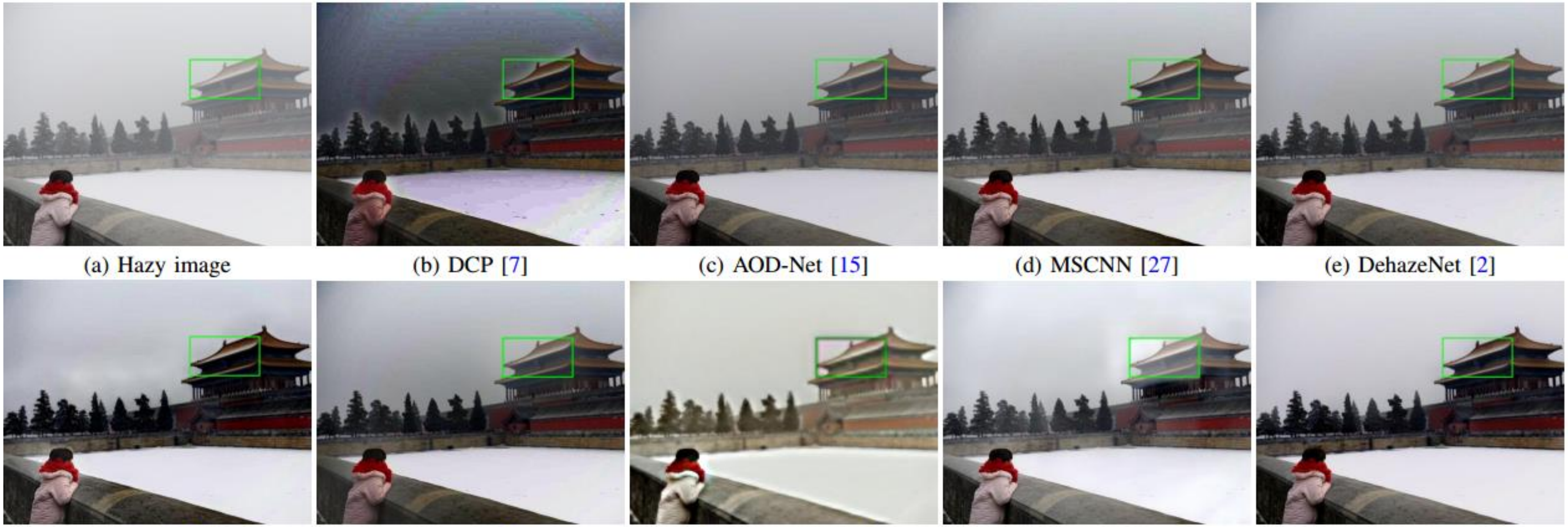
- Dark channel loss

$$L_d = \frac{1}{N_u}\sum_{i=1}^{N_u}\|\mathbf{D}_{J_i}\|_1, \qquad D(I) = \min_{y\in N(x)}\left[\min_{c\in\{r,g,b\}} I^c(y)\right],$$

Results: Synthetic images



Results: Real-world images


(a) Hazy image (b) DCP [7] (c) AOD-Net [15] (d) MSCNN [27] (e) DehazeNet [2]
(f) PDN [39] (g) GFN [28] (h) CycleGAN [4] (i) *baseline* (j) Ours

Results: Real-world images

TABLE V

OBJECT DETECTION RESULTS ON THE RTTS [16] DATASET. WE APPLY FASTER R-CNN TO DETECT OBJECTS OF INTERESTS ON DEHAZED IMAGES. FASTER R-CNN IS TRAINED ON THE VOC2007 [5] DATASET. THE DETECTION TASK FAVORS THE PROPOSED METHOD MOST AMONG THE OTHER ALGORITHMS.

	mAP (%)
Hazy	37.58
DCP [7]	39.63
MSCNN [27]	41.34
DehazeNet [2]	40.54
AOD-Net [15]	37.47
GFN [28]	58.11
DCPDN [41]	61.28
CycleGAN [4]	42.53
Ours	**62.61**

TABLE VIII

QUANTITATIVE EVALUATIONS WITH DIFFERENT AMOUNT OF LABELED DATA. THE PERFORMANCE ON THE SYNTHETIC DATASET ARE ROBUST WHILE THE PERFORMANCE ON THE REAL DATASET ARE SENSITIVE TO THE UNLABELED DATA.

Amount of the unlabeled data	SOTS (PSNR/SSIM)	RTTS (mAP, %)
0	23.65/0.86	53.48
500	24.37/0.88	58.97
1000	24.41/0.89	60.79
2000	24.44/0.89	62.61

Object detection results on the RTTS dataset

A New Benchmark for Single Image Dehazing

Dataset, code, results are available at:

https://sites.google.com/view/lerenhanli/homepage/semi_su_dehazing

Fast Single Image Rain Removal via a Deep Decomposition-Composition Network

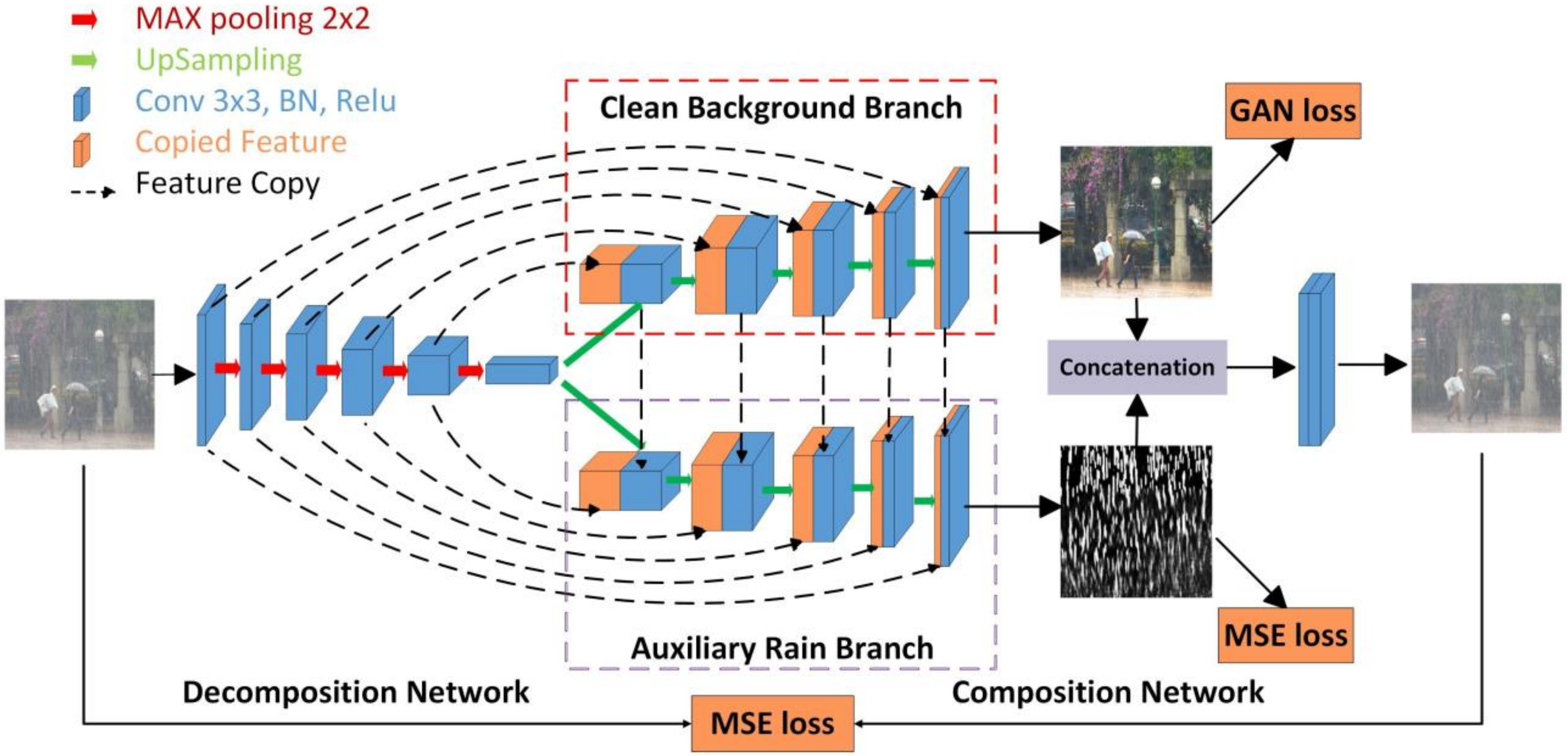
Siyuan Li, Wenqi Ren, Jiawan Zhang, Jinke Yu and Xiaojie Guo

CVIU 2019

Decomposition-Composition Network

Decomposition Net: O = B + R

Composition Net: B + R = O’ ≈ O



Training details of the decomposition net

- **Pre-train on synthetic images:** 10400 triplets [rainy image, clean background, rain layer]
  - paired image-to-image mapping: Euclidean loss of background and rain layer

$$\mathcal{L}_{\mathrm{B}} = \frac{1}{N}\sum_{i=1}^{N}\|f_b(\mathbf{O}^i) - \mathbf{B}^i\|_F^2, \qquad \mathcal{L}_{\mathrm{R}} = \frac{1}{N}\sum_{i=1}^{N}\|f_r(\mathbf{O}^i) - \mathbf{R}^i\|_F^2,$$

- **Fine-tune on real images**: 240 real-world samples
  - GAN adversarial loss

$$\mathcal{L}_{\mathrm{ADV}} = \mathop{\mathbb{E}}_{\mathbf{I}\sim p(I)}\left[\log D(\mathbf{I})\right] + \mathop{\mathbb{E}}_{\mathbf{O}\sim p(O)}\left[\log\left(1 - D(G(\mathbf{O}))\right)\right],$$

Layer	Kernel dimension	Stride	Output size
Input	-	-	224×224
Conv1	$64 \times 4 \times 4$	2	112×112
Conv2	$128 \times 4 \times 4$	2	56×56
Conv3	$256 \times 4 \times 4$	2	28×28
Conv4	$512 \times 4 \times 4$	1	27×27
Conv5	$1 \times 4 \times 4$	1	26×26
Sigmoid	-	-	-

Training details of the composition net

$$\mathbf{O} = \mathbf{B} + \mathbf{R}.$$

$$\mathbf{O} = \left(1 - \sum_{i=0}^{n} \alpha_i\right)\mathbf{B} + \alpha_0 \mathbf{A} + \sum_{i=1}^{n} \alpha_i \mathbf{R}^i, \; s.t. \; \alpha_i \geq 0, \; \sum_{i=0}^{n} \alpha_i \leq 1,$$

- Quadratic training cost function:

$$\mathcal{L}_{\mathbf{O}} = \frac{1}{N} \sum_{i=1}^{N} \| f(\mathbf{O}) - \mathbf{O} \|_F^2.$$

Results: synthetic images

(a) GMM (b) JORDER (c) DDN (d) ID-CGAN (e) Ours w/o CN (f) Our DDC (g) GrounTruth

Results: Real-world images

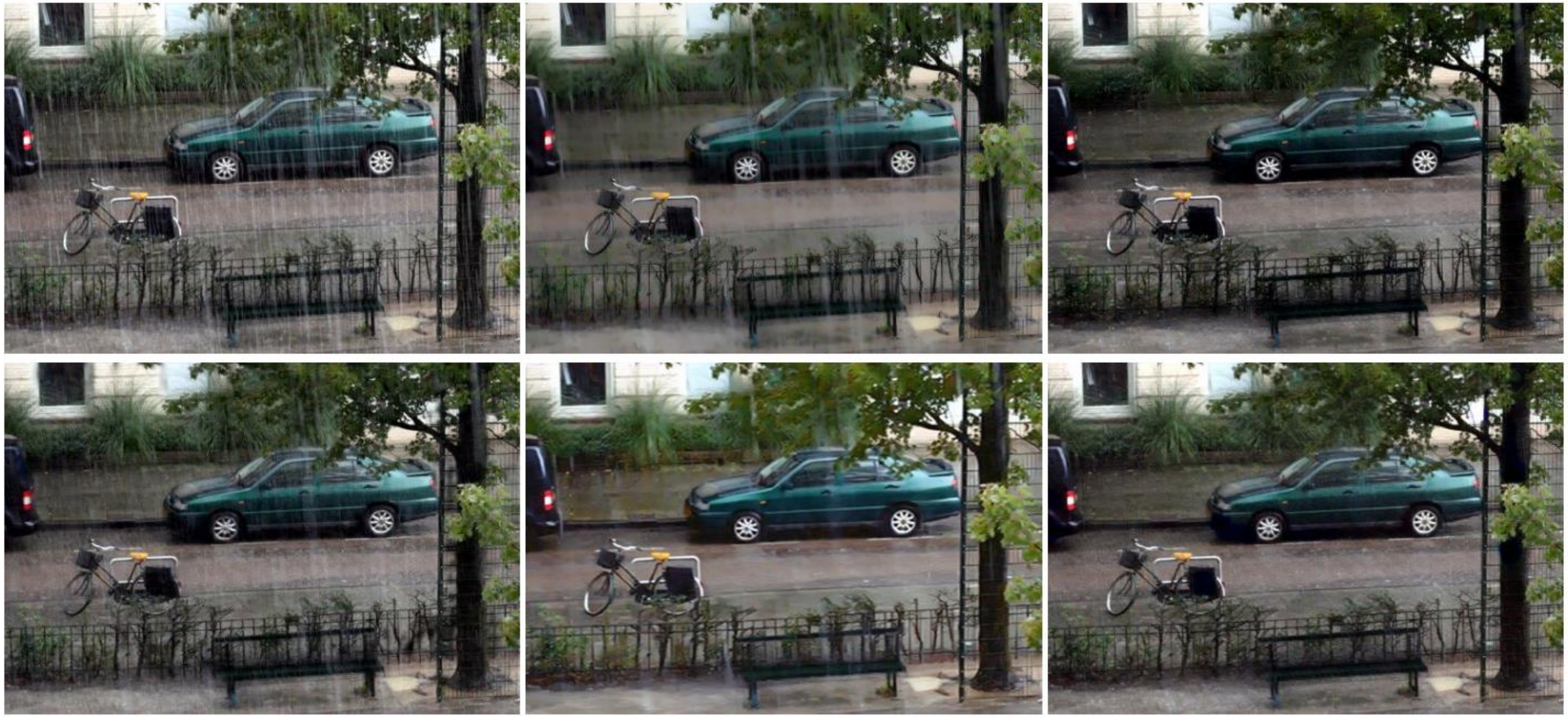


Figure 4: Visual comparison. The upper row contains the input, results by GMM and DDN, while the lower row contains the results by JORDER, ID-CGAN and our DDC-Net, respectively. Please zoom-in the results to see more details and differences.

Results: Real-world images

Figure 5: Visual comparison. The upper row contains the input, results by GMM and DDN, while the lower row contains the results by JORDER, ID-CGAN and our DDC-Net, respectively. Please zoom-in the results to see more details and differences.

Results: Real-world images

Figure 7: Visual comparison. The upper row contains the input, results by GMM and DDN, while the lower row contains the results by JORDER, ID-CGAN and our DDC-Net, respectively. Please zoom-in the results to see more details and differences.

Many unsolved, efforts ongoing…

How to get more and better training data?

I. Improving hazy image synthesis (including fog, smoke, haze…)
- ❑ Indoor depth is accurate, but content has mismatch
- ❑ Outdoor depth estimation is insufficiently accurate for synthesizing haze
- ❑ … and even the atmospheric model itself is *only an approximation*
- ❑ **Ongoing efforts:** developing photo-realistic rendering approaches of generating better hazy images from clean ones, e.g., GAN-based style transfer

II. Go beyond {clean, corrupted} pairs
- ❑ An unsupervised domain adaption or semi-supervised training perspective: we have included 4,322 unannotated realistic hazy images in RESIDE.
- ❑ Signal-level unsupervised prior (loss function): TV norm, no-reference IQA…

More tailored and credible evaluation metrics?

I. More reliable no-reference image quality assessment metrics in dehazing

II. More “task-specific” image quality assessment metrics?

45