NTIRE 2018 Challenge on Image Dehazing: Methods and Results

Cosmin Ancuti Codruta O. Ancuti Radu Timofte Luc Van Gool Lei Zhang Ming-Hsuan Yang Vishal M. Patel He Zhang Vishwanath A. Sindagi Ruhao Zhao Xiaoping Ma Yong Qin Limin Jia Klaus Friedel Sehwan Ki Hyeonjun Sim Jae-Seok Choi Soo Ye Kim Soomin Seo Saehun Kim Munchurl Kim Ranjan Mondal Sanchayan Santra Bhabatosh Chanda Jinlin Liu Kangfu Mei Juncheng Li Luyao Faming Fang Aiwen Jiang Xiaochao Qu Ting Liu Pengfei Wang Biao Sun Jiangfan Deng Yuhang Zhao Ming Hong Jingying Huang Yizhi Chen Erin Chen Xiaoli Yu Tingting Wu Anil Genc Deniz Engin Hazim Kemal Ekenel Wenzhe Liu Tong Tong Gen Li Qinquan Gao Zhan Li Daofa Tang Yuling Chen Ziying Huo Aitor Alvarez-Gila Adrian Galdran Alessandro Bria Javier Vazquez-Corral Marcelo Bertalmo H. Seckin Demir Omer Faruk Adil Huynh Xuan Phung Xin Jin Jiale Chen Chaowei Shan Zhibo Chen

Abstract

This paper reviews the first challenge on image dehazing (restoration of rich details in hazy image) with focus on proposed solutions and results. The challenge had 2 tracks. Track 1 employed the indoor images (using I-HAZE dataset), while Track 2 outdoor images (using O-HAZE dataset). The hazy images have been captured in presence of real haze, generated by professional haze machines. I-HAZE dataset contains 35 scenes that correspond to ***indoor*** *domestic environments, with objects with different colors and specularities. O-HAZE contains 45 different* ***outdoor*** *scenes depicting the same visual content recorded in haze-free and hazy conditions, under the same illumination parameters. The dehazing process was learnable through provided pairs of haze-free and hazy train images. Each track had* ~ 120 *registered participants and 21 teams competed in the final testing phase. They gauge the state-of-the-art in image dehazing.*

1. Introduction

Haze is a common atmospheric phenomena produced by small floating particles that reduce the visibility of distant objects due to light scattering and attenuation. This results in a loss of local contrast for distant objects, in the addition of noise to the image, and in a selective attenuation of the light spectrum. Image dehazing is a challenging ill-posed problem that has drawn a significant attention in the last few years.

In the last decade a significant amount of literature focused on single image dehazing research. The performance of the top methods continuously improved [29, 11, 31, 14, 23, 1, 20, 30, 2, 24, 7] showing that the field reaches maturity. Despite this growing interest, the field lacks standardized benchmarks to allow for evaluating objectively and quantitatively the performance of the existing dehazing techniques.

Basically, major issue preventing further developments is related to the impossibility to reliably asses the dehazing performance of a given algorithm, due to the absence of reference haze-free images (ground-truth). A key problem in collecting pairs of hazy and haze-free ground-truth images lies in the need to capture both images with identical scene illumination.

In general the existing dehazing quality metrics are restricted to non-reference image quality metrics (NRIQA) [21]. For instance, the Fog Aware Density Evaluator (FADE)[8] estimates the visibility of a hazy/foggy scene from a single image without corresponding ground-truth. Unfortunately, due to the absence of the reference (haze-free) images in real-life scenarios, none of these approaches has been generally accepted by the dehazing community.

C. Ancuti (cosmin.ancuti@upt.ro, University Politehnica Timisoara), C. O. Ancuti, R. Timofte, L. Van Gool, L. Zhang and M.-H. Yang are the NTIRE 2018 organizers, while the other authors participated in the challenge.
Appendix A contains the authors' teams and affiliations.
NTIRE webpage: http://www.vision.ee.ethz.ch/ntire18/

Recent works synthesize hazy images, using the optical model and known depth to synthesize the haze effect. For instance, FRIDA [32] dataset designed for Advanced Driver Assistance Systems (ADAS) is a synthetic image database with 66 computer graphics generated roads scenes. D-HAZY [3] is a dataset of 1400+ images of real complex scenes has been derived from the *Middleburry* and the *NYU-Depth V2* datasets. The depth map associated to each high quality image has been used to yield synthesized hazy images based on the simplified optical model. Khoury *et al*. [16] introduces the CHIC (Color Hazy Image for Comparison) database, providing hazy and haze-free images in real scenes captured under controlled illumination. The dataset however only considers two indoor scenes, thereby failing to cover a large variation of textures and scene depth.

The NTIRE 2018 challenge is a step forward in benchmarking single image dehazing. It uses two novel dataset (I-HAZE and O-HAZE) that can be employed as a more representative benchmark to assess dehazing algorithms in outdoor scenes, based on ground truth images. The I-HAZE consists from 35 hazy images (with haze generated in a controlled indoor environment) and their corresponding ground truth (haze-free) images of the same scene. The O-HAZE dataset includes 45 hazy images and corresponding ground truth (haze-free) images. The haze has been produced using a professional haze/fog generator that imitates the real conditions of haze scenes.

2. NTIRE 2018 Challenge

The objectives of the NTIRE 2018 challenge on single image dehazing are: (i) to gauge and push the state-of-the-art in image dehazing; (ii) to compare different solutions; and (iii) to promote novel I-HAZE and O-HAZE datasets with real haze and ground truth haze-free images.

2.1. I-HAZE dataset

I-HAZE [4] dataset contains 35 scenes that correspond to **indoor** domestic environments, with objects with different colors and specularities. Besides the domestic objects, all the scenes contains a color checker chart (Macbeth color checker). We used a classical Macbeth color checker with the size 11 by 8.25 inches with 24 squares of painted samples (4×6 grid).

After carefully setting each scene, we first recorded the ground truth (haze-free image) and then immediately started introducing haze in the scene. We used two professional fog/haze machines (LSM1500 PRO 1500 W) to generate a dense vapor. These fog generators use cast or platen type aluminum heat exchangers, which causes evaporation of the water-based fog liquid. The generated particles (since are water droplets) have approximately the same diameter size of 1 - 10 microns as the atmospheric haze. Before shooting the hazy scene, we used a fan to obtain in a relatively short period of time a homogeneous haze distribution in the entire room (room kept isolated as much as possible by closing all the doors and windows). The entire process to generate haze took approximately 1 minute. Waiting approximately another 5-10 minutes, we obtained a homogeneous distribution of the haze. The distances between the camera and the target objects ranged form 3 to 10 meters. The recordings were performed during the daytime in relatively short intervals (20-30 minutes per scene recording) with natural lightning and when the light remained relatively constant (either smooth cloudy days or when the sun beams did not hit directly the room windows).

To capture haze-free and hazy images, we used a setup that includes a tripod and a Sony A5000 camera that was remotely controlled (Sony RM-VPR1). We acquired JPG and ARW (RAW) 5456×3632 images, with 24 bit depth. The cameras were set on manual mode and we kept the camera still (on a tripod) over the entire shooting session of the scene. The camera was calibrated in haze-free scene, and then we kept the same parameters for the hazy scene. For each scene, the camera settings was calibrated by manually adjusting the aperture (F-stop), shutter-speed (exposure-time), ISO speed and the white-balance. Setting the three parameters aperture-exposure-ISO was realized using both the built-in light-meter of the camera and an external exponometer Sekonic. For the white-balance we used the gray-card, targeting a middle gray (18% gray). The calibration process was straight-forward, since it just required to set the white-balance in manual mode and to place the gray-card in front of the subject. In practice, we placed the gray-card in the center of the scene, two meters away from the camera.

2.2. O-HAZE Datasets

The O-HAZE [5] database has been derived from 45 various **outdoor** scenes in presence or absence of haze. Our dataset allows to investigate the contribution of the haze over the scene visibility by analyzing the scene objects radiance starting from the camera proximity to a maximum distance of 30m.

The recording process was similar with the one used to record the I-HAZE dataset. However, the entire recording phase for the O-HAZE was much more difficult. Since we aimed for outdoor conditions similar to the ones encountered in hazy days, the recording period spread over more than 8 weeks during the autumn season. We recorded the scenes during cloudy days, in the morning or in the sunset, and only when the wind speed was below 3 km/h (to limit

http://vision.middlebury.edu/stereo/data/scenes2014/
http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html
http://www.vision.ee.ethz.ch/ntire18/i-haze/
http://www.vision.ee.ethz.ch/ntire18/o-haze/

fast spreading of the haze in the scene). The absence of wind was the condition hardest to meet, and explains why the recording of 45 scenes took more than 8 weeks.

Each scene acquisition started with a manual adjustment of the camera settings. The same parameters were adopted to capture the haze-free and hazy scene. Those parameters include the shutter-speed (exposure-time), the aperture (F-stop), the ISO and white-balance. The similarity between hazy and haze-free acquisition settings is confirmed by the fact that the closer regions (that in general are less distorted by haze) have similar appearance (in terms of color and visibility) in the pair of hazy and haze-free images associated to a given scene.

After carefully checking all the conditions above mentioned and after the setting procedure we placed in each scene a color checker (Macbeth color checker) to allow for post-processing of the recorded images. We used a classical Macbeth color checker with the size 11 by 8.25 inches with 24 squares of painted samples (4×6 grid).

We chose special (haze) liquid with higher density in order to simulate the effect occurring with water haze over larger distances than the investigated 20-30 meters. The generation of haze took approximately 3-5 minutes. After haze generation, we used a fan to spread the haze as uniformly as possible in the scene.

2.3. Tracks

For each track of NTIRE 2018 challenge we created a Codalab competition. To access the data and submit their HR image results to the CodaLab evaluation server each participant had to register.

Track 1: Indoor facilitates the objective evaluation of single image dehazing techniques for the 35 indoor scenes of the I-HAZE dataset.

Track 2: Outdoor facilitates the objective evaluation of single image dehazing techniques for the 45 indoor scenes of the O-HAZE dataset.

Challenge phases *(1) Development (training) phase:* the participants got train data (hazy and haze-free images) (25 sets for the indoor track and 35 for the outdoor track); *(2) Validation phase:* the participants received 5 additional sets of images for both tracks and had the opportunity to test their solutions on the hazy validation images and to receive immediate feedback by uploading their results to the server. A validation leaderboard is available; *(3) Final evaluation (test) phase:* the participants got the hazy test images (5 sets for each track) and had to submit both their dehazed image and a description of their methods before the challenge deadline. One week later the final results were made available to the participants.

Evaluation protocol The Peak Signal-to-Noise Ratio (PSNR) measured in deciBels (dB) and the Structural Similarity index (SSIM) [34] computed between an image result and the ground truth are the quantitative measures. The higher the score is the better the restoration fidelity to the ground truth image.

3. Challenge Results

From more than 115 registered participants on average per each track, 19 teams entered in the final phase and submitted results, codes/executables, and factsheets. Table 1 reports the final scoring results of the challenge and Table 2 shows the runtimes and the major details for each entry. Section 4 describes briefly the methods for each team while in the Appendix A are the team members and affiliations.

Team	user	Track 1: Indoor		Track 2: Outdoor	
		PSNR	**SSIM**	**PSNR**	**SSIM**
Scarlet Knights[36]	**rutgerssprinter**	24.973_1	0.881	24.029_3	0.775
BJTU	**Team T-brain**	22.866_3	0.857	24.598_1	0.777
FKS	**fks**	22.909_2	0.864	23.877_4	0.775
KAIST-VICLAB[27]	**hjsim**	20.354_7	0.829	24.232_2	0.687
Ranjanisi[22]	**ranjanisi**	20.911_5	0.751	23.180_6	0.705
KAIST-VICLAB[17]	**Team KAIST-VICLAB**	22.421_4	0.852	22.705_9	0.707
Ranjanisi[22]	**ecsuiplab**	19.860_8	0.747	22.997_8	0.701
Dq-hisfriends	**liujl09**	17.622_{12}	0.817	23.207_5	0.770
CLFStudio	**Team CLFStudio**	20.549_6	0.803	20.230_{13}	0.722
Mt.Phoenix	**MTLab**			23.124_7	0.755
Xdhm	**hjy_2**	18.542_9	0.808		
Mt.Phoenix	**qxc**	16.653_{13}	0.787	22.080_{10}	0.731
SiMiT-Lab[10]	**Team DeAn**	18.305_{10}	0.809	19.628_{14}	0.674
Xdhm	**Tia**	18.254_{11}	0.800		
IVLab	**Team IVLab**			21.750_{11}	0.717
CLEAR	**lizhangray**			20.291_{12}	0.683
Dehazing-by-retinex[12]	**shuffle**	16.544_{14}	0.787	17.547_{17}	0.652
ASELSAN	**DemirAdil**	15.921_{15}	0.738	18.123_{16}	0.675
AHappyFaceI	**Team HCILab**	15.147_{17}	0.751	18.494_{15}	0.669
Xdhm	**xdhm**	15.292_{16}	0.746		
Xdhm	**hjy**	14.787_{18}	0.748		
IMCL	**IMCL**			16.527_{18}	0.616
hazy images	*baseline*	12.876	0.704	15.784	0.634

Table 1. NTIRE 2018 Challenge dehazing results and final rankings on indoor (I-HAZE) and outdoor (O-HAZE) test data.

Architectures and main ideas In general the proposed methods (excepting ASELSAN, CLEAR and Dehazing-by-retinex), use the end-to-end deep learning and employ the GPU(s) for both training and testing. Scarlet Knights Team built a Dense-Residual encoder-decoder structure with multi-level pyramid pooling module for estimating the dehazed image. The network is optimized using the traditional L2 loss and a newly introduced perceptual loss. KAIST-VICLAB designs an architecture based on pix2pixGAN model [15] with the U-Net as generator and patchGAN [15] as discriminator. Similarly FKS employs an architecture based on U-net [25]. The Dehazing-by-retinex uses an unsupervised technique, Retinex, on the inverted intensities of a hazy image. However, the CPU-based techniques are the slowest ranking below top ten in terms of PSNR and SSIM performances.

Restoration fidelity In PSNR terms the Scarlet Knights, BJTU and FKS are the winner teams of NTIRE 2018 dehazing challenge. Scarlet Knights achieves more than 24dB in both tracks. In SSIM terms Scarlet Knights obtained the best score for the Track 1 (0.881). On Track 2, BJTU got

https://competitions.codalab.org

	Runtime [s]				
Team	**Track1: Indoor**	**Track2: Outdoor**	**Platform**	**CPU/GPU (at runtime)**	**Architecture**
Scarlet Knights	10.0	10.0	pytorch	Titan X	Multi-scale Dense-residual Net
BJTU	6.4	9.9	Python 2.7, Pytorch	Titan X	DenseNet
FKS	11.5	10.3	Keras with tensorflow	Nvidia GTX1080Ti	
KAIST-VICLAB	3.9	4.8	Matlab, Python 3.6, Pytorch	Titan X	CBEGAN
Ranjanisi	5.7	4.8	Keras with tensorflow backend	Nvidia GeForce GTX 745	Fully Convolutional Net (FCN)
KAIST-VICLAB	9.6	5.2	Python 2.7, Tensorflow	CUDA v8.0, CuDNN 6	pix2pixGAN
CLFStudio	10.0	10.0	pytorch	Titan X?	
Dq-hisfriends	6.0	8.0	Tensorflow	Nvidia GTX1080Ti	
Mt.Phoenix	19.7	17.4	pytorch	Nvidia GTX1080Ti	
Xdhm	1.7		pytorch		pix2pix and IRCNN
SiMiT-Lab	8.0	8.0	Tensorflow	Titan X	Cycle-GAN
IVLab			Tensorflow	Titan X	
CLEAR	207.4	382.0	Matlab	(CPU)	
Dehazing-by-retinex	37.0	37.0	C++	(CPU)	
ASELSAN	30.0	30.0	Matlab	(CPU)	
AHappyFaceI	21.8	18.7	Tensorflow; caffe2 with python	Titan X	
IMCL		4.5	Tensorflow	Nvidia GTX1080Ti	DDC-GAN

Table 2. Reported runtimes per image on I-HAZE and O-HAZE test data and details from the factsheets

the best score of 0.777, but the top 3 SSIM scores are within 0.002.

Runtime / efficiency KAIST-VICLAB solution is the most efficient, it gives the best trade-off between runtime and quality of the results. It runs in 3.85s (Track 1) and 4.8s (Track 2) on Intel Xeon @3.0GHz (2 processors), 96GB RAM, Nvidia Titan X Pascal while being only 0.3dB below the best reported result on Track 2.

Train data I-HAZE [4] with 35 indoor set of images and O-HAZE [5] with 45 set of outdoor images were used by all the competitors that in general found the amount of data sufficient for training their model, especially after data augmentation [33] (by operations such as flipping, rotation, scaling).

Conclusions By analyzing the challenge methods and their results we can draw several conclusions. (i) The proposed solutions have a degree of novelty and go beyond the published state-of-the-art methods. (ii) In general the best solutions performed the best for both tracks and for both measures (PSNR and SSIM). (iii) There were no significant differences between the indoor and outdoor results. (iv) The evaluation based on SSIM is questionable since there is only a small variation of the SSIM results (especially for the top methods on Track 2).

4. Challenge Methods and Teams

4.1. Scarlet Knights Team

Scarlet Knights Team proposes multi-scale single image dehazing using perceptual pyramid deep network [36, 35], that aims to directly learn the mapping between input hazy image and its corresponding clean target image. The method combines dense block with residual block to maximize the information flow along features from different levels. The proposed network consists of a novel Dense-Residual encoder-decoder structure with multi-level pyramid pooling module for estimating the dehazed image. This network is optimized using the traditional L2 loss and a newly introduced perceptual loss. During inference, a multi-scale testing protocol is leveraged to boost the final results.

For indoor sample, the method leverages a two-scale strategy for testing. Basically, there are created two sets of overlapping-patches from the test image. For the first set, the size of the overlapping patch is chosen such that width is larger than the height (2048×1024). For the second set, the size of the overlapping patch is chosen such that width is equal to the height (2048×2048). Patches in both sets are forwarded through the network to obtain the dehazed results (patches). The output patches in each set are then merged appropriately to form the entire output image - one each for a set. The final output is computed as the mean of the two output images from the two sets.

For outdoor sample, the method follows a slightly different strategy that involves two scales. Two sets of overlapping-patches are created from the test image. For the first set, the size of the overlapping patch is chosen such that width is greater than the height (3072×1536). The second set consists of a single image that is obtained by downsampling the input test image to a resolution of 1024×1024. The patches in the first set are forwarded through the network and the resulting patches are merged into a single image and resized to the original resolution. The image in the second set is forwarded through the network and the results is upsampled to the original resolution. In addition, the results from version 2 to original resolution are upsampled. The final output is computed as the mean of the two output images from the two sets.

4.2. BJTU Team

BJTU Team introduces a novel dehazing technique (see Fig. 2) named DND (Dense Net for Dehaze). To get more

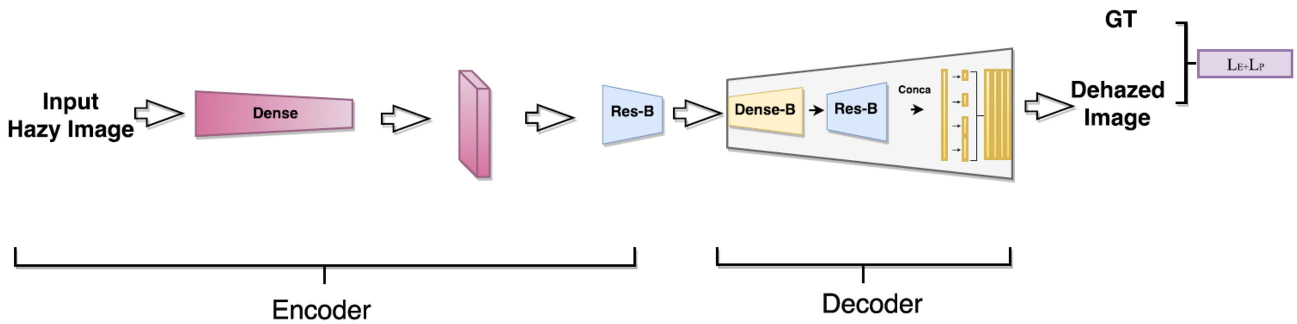

Figure 1. Scarlet Knights Team: An overview of the proposed dehazing method. It consists of: three-block dense block as encoder, dense-residual block as decoder and multi-scale pyramid module at the end of the decoder. It is optimized using Euclidean loss and perceptual loss, where perceptual loss is evaluated on ReLU 1-1 and ReLU 2-1 of pretrained vgg-16.

details from the train data, the method crops the images into multi-sizes. In the proposed method, the dense network is used to solve the gradient vanishing problem. Additionally the approach adds some transition blocks to down/upsample to learn more features. Through the dense blocks and transition blocks, the dehazed images are addressed by two loss functions to achieve better results. The learning rate is set as 0.0005 and the size of images as 736 for the consideration of GPU memory. The images are tested in size of 1024. And then, the dehazed results are resized into real sizes. Normal dense net can catch the local features, but it is hard to get the global features. The novelty of the proposed method is that through downsampling and upsampling, the approach can blur the local features and easily catch features of the whole images.

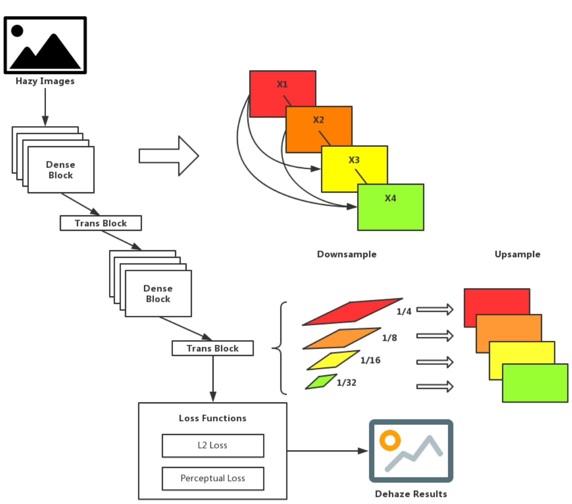

Figure 2. BJTU Team: overview of the proposed method.

4.3. FKS Team

FKS Team proposes a solution that uses a three staged approach. The first stage tries to reconstruct depth information for the input image. This first model is applied to the whole image. The global view of the image helps in depth estimation. This depth information is then added to the input image as additional feature map. The second stage works on smaller crops (512×512) of the stage one output. A second model is trained on the dehazing of such smaller tiles. The final part of the solution breaks down the full resolution image into smaller overlapping tiles. Every tile gets dehazed by the second model. To get better results test data augmentation [33] is used. Every tile is flipped and rotated by 90 degree to get multiple predictions for every tile. The tiles are then transformed back to the correct orientation and are composed to the final image. A weighted border is used on every tile to mitigate border artifacts.

Two deep convolutional networks build the foundation of the solution (see Fig. 3). The first one operates on the whole (scaled down) image, the second one on small image patches (crops) of high resolution. Both CNNs use similar architecture based on U-Net [25]. The depth prediction network is trained on a input size of 384×384. Several conv-blocks reduce the resolution down to 6×6. The deconvolution blocks of the network bring the resolution back to 384×384.

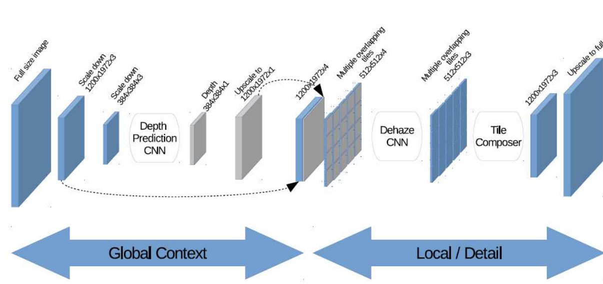

Figure 3. FKS Team: An overview of the method.

4.4. KAIST-VICLAB Team

The KAIST-VICLAB Team proposes an extremely large receptive field size GAN using a simple down-scaling and up-scaling. In general, in the case of dehazing or fog re-

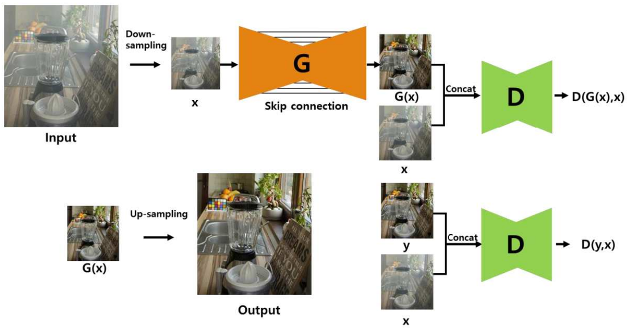

Figure 4. KAIST-VICLAB Team: Diagram of Model A. Model B can be represented by replacing D with patchGAN.

moval, the entire image should be input to a deep-network at once in order to estimate the atmosphere light of the whole image. However, since the sizes of training images and testing images are very large, general size of original patch captures only small part (flat wall or floor) of a picture, which makes network learning ineffective. Therefore, the team enlarges the effective receptive field size of the deep-network using down-scaling of the input images to feed to the networks and the up-scaling of the output dehazed images. The team proposes two GAN models[27, 17] for single image dehazing. The first model A is a conditional boundary equilibrium GAN (CBEGAN). Generator of the CBEGAN is U-net [15], and Discriminator of CBEGAN is a conditional auto-encoder network (input: generated dehazing image and input hazy image) which is modified from the discriminator of BEGAN [6]. Instead of the last transposed convolution, the upsampling block is used (bilinear upsampling layer convolution) for reducing the checkerboard convolution. The second model Bs architecture is based-on pix2pixGAN model [15]. Hence, generator is U-net like above CBEGAN, but discriminator is a patchGAN [15], which has a score map at the end, not a single value. While original images are downsampled to provide wide spatial information to the Generator, patchGAN architecture concentrates on the narrow part of the output. It also helps to make discriminator a lot shallower than the generator in order to balance convergence of generator and discriminator.

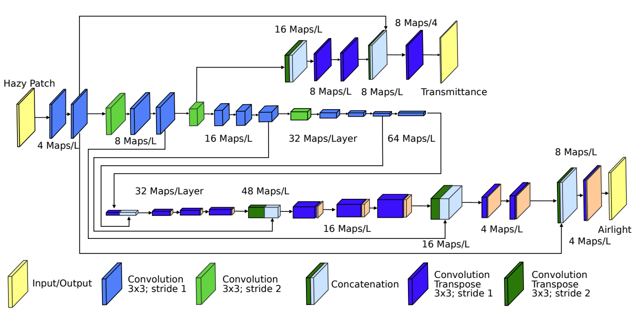

Figure 5. Ranjanisi Team: solution diagram.

4.5. Ranjanisi Team

The Ranjanisi Team first modifies the haze model of [18] and then they introduced a Fully Convolutional Neural Network (FCN) [19] for image dehazing (see Fig. 5). The method [22] is based on a network that jointly estimates scene transmittance and airlight. The network is trained using a custom designed loss, called bi-directional consistency loss, that tries to minimize the error to reconstruct the hazy image from clear image and the clear image from hazy image. Apart from that, to minimize the dependence of the network on the scale of the training data, the training and inference are performed in multiple levels.

4.6. Dq-hisfriends Team

The Dq-hisfriends Team trains a simplified U-Net [25], with only 7 layers. They use two loss functions, L1 loss function and the proposed difference gradient loss. Training using only L1 loss function shows strong artifacts. The gradient difference loss helps eliminate the artifacts and improves the final results significantly. The team randomly sample 512×512 patches from the training dataset and resize or rotate them to augment our dataset. The training of the network employs these patches. The output of the proposed method looks natural with no artifacts. The proposed method is significantly better than classic dehazing method such as dark channel and non-local dehazing method.

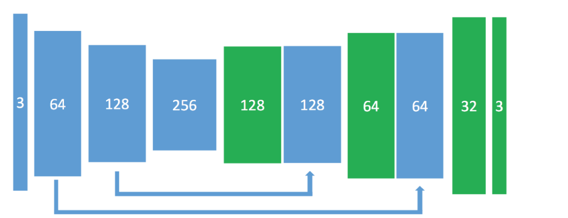

Figure 6. Dq-hisfriends Team: architecture with 7 used convolutional layers. The digits denotes numbers of channels.

4.7. CLFStudio Team

The CLFStudio team proposed a novel dehazing technique based on Progressive Feature Fusion Network (PFFNet). As shown on Figure 7, the PFFNet could be divided into four parts: (i) Feature Enhance Blocks (FEBs) handles the haze image that is given to the network; (ii) ResBlocks extracts features from the feature maps; (iii) Haze Restore Blocks(HRBs) restores the haze-free image from the feature maps; (iv) in order to keep detail and to accelerate learning speed, Progressive Feature Fusion (PFF) is used in the proposed architecture.

4.8. Mt.Phoenix Team

Mt.Phoenix Team proposed a Convolutional Neural Network (CNN) with a structure similar to U-Net [25]. For

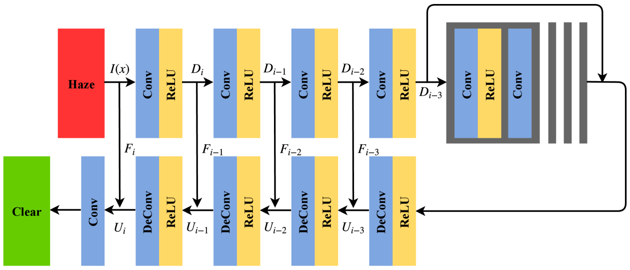

Figure 7. CLFStudio Team: overview of the proposed method.

the encoder part, several layers of convolution, batch normalization, PReLU are used to extract relevant features. The decoder part also utilizes several layers of convolution, batch normalization, PReLU layers to remove haze and reconstruct a clean image. The skip connections between encoder and decoder help feature and gradient propagation and ease the training difficulty. A variant of GAN [13] called Conditional Generative Adversarial Networks (cGANs) is used to train the CNN. The generator and discriminator of cGANs take as input not only a random vector or a generated fake image but also a condition image. Several other losses are also used. The L2 loss between pixel values, perceptual loss calculated using VGG features and the total variation loss calculated using the generated image are combined with the adversarial loss. The method is used for both indoor and outdoor tracks.

4.9. Xdhm Team

The Xdhm Team uses pix2pix [15] and IRCNN [37] to solve the dehazing problem. The dataset employed to train pix2pix model is generated using the team's hazing-pic-builder. The size of images used to train pix2pix is 128×128, but to generate haze-free images the team uses bigger sizes. These output images are resized to the original size, and in the final step the proposed method uses IRCNN to denoise the result.

4.10. SiMiT-Lab Team

SiMiT-Lab team introduces an end-to-end network, *Cycle-Dehaze*, which is an enhanced version of CycleGAN [38]. In order to increase visual quality metrics, Cycle-Dehaze combines cycle-consistency loss of CycleGAN [38] with the perceptual-consistency loss inspired by EnhanceNet [26]. Perceptual-consistency loss is calculated in a similar way with cycle-consistency loss between the original and cyclic image. However it compares the images with the combination of high and low-level features extracted from $2nd$ and $5th$ pooling layers of *VGG16* [28] architecture, instead of $L1 - norm$. Figure 8 illustrates an overview of Cycle-Dehaze. In order to feed the input images into our network, they are resized to 256×256 pixel resolution via bicubic downscaling. After obtaining low-resolution dehazed images, Laplacian pyramid is utilized to provide better high-resolution images.

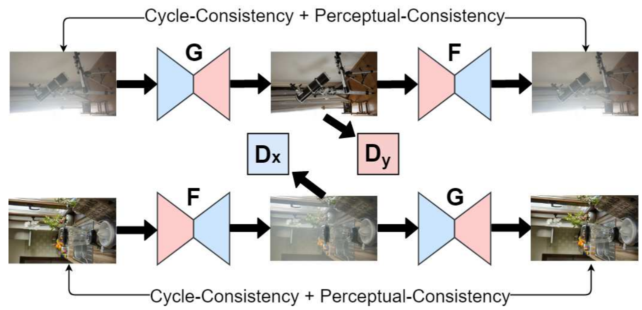

Figure 8. SiMiT-Lab Team: Cycle-Dehaze Network where G & F refers to the generators, and D_x & D_y to the discriminators.

4.11. IVLab Team

IVLab team proposes an efficient multi-scale densely connected dehazing network. Firstly, the proposed network is designed based on a novel re-formulated retinex to dehazing formula, which helps to optimize the pipeline from hazy images to clear images directly. Secondly, the proposed network is utilized to capture different haze droplet concentrations and size distribution, simultaneously. Thirdly, the intensity-based adaptive high contrast histogram equalization technique is used to pre-processing the hazy images, which significantly improves the de-haze performance. Finally, comparing with the most existing methods, the proposed network outperforms state-of-the-art with far fewer parameters.

4.12. CLEAR Team

CLEAR Team introduces an improved image dehazing method based on dark channel prior [14] with color fidelity (see Fig. 9). First, the initial transmission map estimated from the dark channel prior is refined using Prewitt operator and mean filters. Second, a piecewise constrained function is proposed to keep color fidelity. Experimental results on real world images show that the proposed approach is effective in haze removal and preventing the color distortion from excessive dehazing.

4.13. Dehazing-by-retinex Team

The Dehazing-by-retinex team's approach builds on their recent work [12], which shows, both theoretically and experimentally, that applying Retinex on the inverted intensities of a hazy image, and then inverting the results, can lead to the effective removal of haze. While [12] shows that the described relation of duality between Retinex and dehazing algorithms holds for a variety of Retinex methods, in this challenge the team applies the Light Random Spray Retinex (LRSR) implementation for both indoor and outdoor tracks,

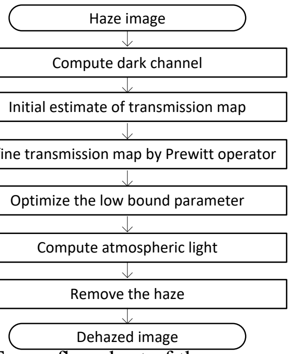

Figure 9. CLEAR Team: flowchart of the proposed method.

and adds a median filtering (with a disk element of size 3) as a postprocessing step.

4.14. ASELSAN Team

The proposed method improves on dark channel prior [14] using a more robust estimation of the haze color. The initial preprocessing step estimates the haze color by using intensity and saturation priors. Then, the effect of the haze color is reduced by using these estimations. This step provides not only a better color estimation but also a better estimation of the transmission map. The results show that the proposed method gives superior results than the original dark channel prior dehazing method.

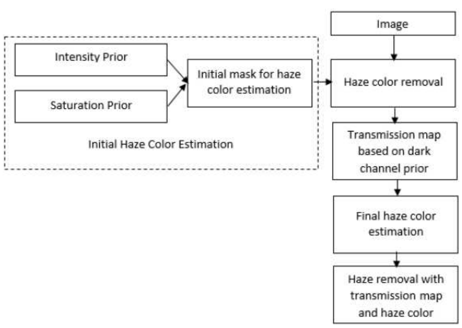

Figure 10. ASELSAN Team: overview of the proposed method.

4.15. AHappyFaceI Team

The team introduces a method (see Fig. 11) based on the CNN architecture of Dong *et al*. [9]. It uses two groups of convolutional layers. The input is added to the middle layer and the output layer. The input image is divided into smaller patches with a size of 64×64 for training and 512×512 for testing.

4.16. IMCL Team

IMCL Team's method decomposes the hazy image into two components, namely non-haze background image and haze image. An adversarial training mechanism is employed to synthesize non-haze background image and haze

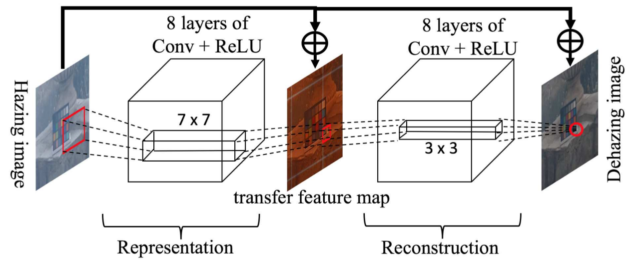

Figure 11. AHappyFaceI Team: overview of the proposed method.

image in a Dual-Cross manner, which makes the two adversarial branches interact with each other, archiving a win-win result ultimately. The proposed Decomposed Dual-Cross Generative Adversarial Network (DDC-GAN) (see Fig. 12) shows significant performance improvement. The proposed method does not just simply apply CNNs as a ready-made tool to solve such a complicated computer vision problem. On the contrary, as stated above, the method redefine such an issue of haze removal from a completely different perspective, to treat image dehazing as a signal decomposition problem. Then, based on the above motivation, the method is based on a novel dual-cross network structure: twisting two generators (one for background generation, another one for haze generation) together, which forms a closed loop structure. Therefore, the two generators can provide a dual-cross informative signals to each other while preserving independent domain features during forward propagation. During backward propagation, by optimizing two generators' targets together, the Generator-B and Generator-H mutually benefit each other through the symmetric feedback signals, which leads to realistic background and haze results simultaneously and achieves a win-win result ultimately.

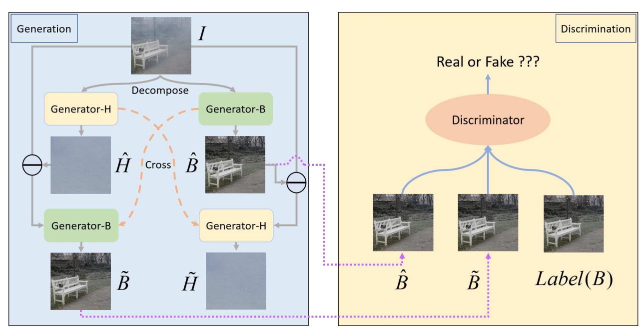

Figure 12. IMCL Team: overview of the proposed method.

Acknowledgments

We thank the NTIRE 2018 sponsors: Alibaba Group, NVIDIA Corp., SenseTime Group Ltd., Huawei Technologies Co. Ltd., Code Ocean, Google Inc., Disney Research, Amazon.com, Inc., and ETH Zurich.

A. Teams and affiliations

NTIRE2018 team

Title: NTIRE 2018 Challenge on Image Dehazing
Members: *Cosmin Ancuti* [1]*(cosmin.ancuti@upt.ro)*, Codruta O. Ancuti [1], Radu Timofte [2,3], Luc Van Gool [2,4], Lei Zhang [5], Ming-Hsuan Yang [6]
Affiliations:
[1] University Politehnica Timisoara, Romania
[2] Computer Vision Lab, ETH Zurich, Switzerland
[3] Merantix GmbH, Germany
[4] ESAT, KU Leuven, Belgium
[5] Polytechnic University of Hong Kong, China
[6] University of California at Merced, US

AHappyFaceI team

Title: Enhancing Image De-hazing Using Double Residual Convolutional Networks
Members: *Huynh Xuan Phung (phunghx@gmail.com)*
Affiliation:
HoChiMinh City University of Technology and Education, Vietnam

ASELSAN team

Title: Robust Haze Color Estimation for Dark Channel-based Dehazing
Members: *H. Seckin Demir (hsdemir@aselsan.com.tr)* and Omer Faruk Adil
Affiliation:
ASELSAN Inc., Ankara, Turkey

BJTU team

Title: Dense Net for Dehaze by BJTU
Members: *Ruhao Zhao (420325536@qq.com)*, Xiaoping Ma, Yong Qin and Limin Jia
Affiliation:
The State Key Lab of Rail Traffic Control and Safety in Beijing Jiaotong University, China

CLEAR team

Title: A Color Fidelity DCP-based Dehaze Tool
Members: *Zhan Li*[1,2] *(lizhangray@qq.com)*, Daofa Tang[3,2], Yuling Chen[1,2], Ziying Huo[1,2]
Affiliation:
[1] Department of Computer Science, Jinan University, Guangzhou, China
[2] Sino-France Joint Laboratory for Astrometry, Dynamics and Space Science, Jinan University, Guangzhou, China
[3] China United Network Telecommunications Corporation Limited Guangdong Branch, Guangzhou, China

CLFStudio team

Title: Progressive Feature Fusion Network for Image Dehazing
Members: *Kangfu Mei*[1] *(kfm@clfstudio.com)*, Juncheng Li[2], Luyao[1], Faming Fang[2], and Aiwen Jiang[1]
Affiliation:
[1] Jiangxi Normal University, China
[2] East China Normal University

Dehazing by retinex team

Title: Retinex on Inverted Intensities for Dehazing
Members: *Aitor Alvarez-Gila* [1,2] *(aitor.alvarez@tecnalia.com)*, Adrian Galdran [3], Alessandro Bria [4], Javier Vazquez-Corral [5], Marcelo Bertalmo [5]
Affiliation:
[1]Tecnalia, Spain
[2]Computer Vision Center/Universitat Autonoma de Barcelona, Spain
[3]INESC-TEC Porto, Portugal
[4]University of Cassino, Italy
[5]Universitat Pompeu Fabra, Spain

Dq-hisfriends team

Title: Training a convolutional neural network using difference gradient loss
Members: *Jinlin Liu (jl-liu13@mails.tsinghua.edu.cn)*
Affiliation:
Tsinghua University, Beijing, China

FKS team

Title: Global Depth Local Dehaze
Members: *Klaus Friedel (ntire@friedelprivat.de)*
Affiliation:
Bronnackerstr. 42, 73527 Schwabisch Gmund, Germany

IMCL team

Title: A Decomposed Dual-Cross Generative Adversarial Network for Image Dehazing
Members: *Xin Jin (jinxustc@mail.ustc.edu.cn)*, Jiale Chen, Chaowei Shan , Zhibo Chen
Affiliation:
USTC (University of Science and Technology of China)

IVLab team

Title: Multi-scale densely connected network for single image haze removal
Members: *Wenzhe Liu (wzliu@imperial-vision.com)*, Tong Tong, Gen Li, Qinquan Gao
Affiliation:
Imperial Vision Technology

KAIST-VICLAB team

Title: Extremely Large Receptive Field Size GAN for Single Image Dehazing
Members: Hyeonjun Sim, *Sehwan Ki (shki@kaist.ac.kr)*, Jae-Seok Choi, Sooye Kim, Soomin Seo, Saehun Kim, Munchurl Kim
Affiliation:
Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea

Mt.Phoenix team

Title: DehazeNetGan
Members: *Xiaochao Qu (qxc@meitu.com)*, Ting Liu, Pengfei Wang, Biao Sun, Jiangfan Deng, Yuhang Zhao
Affiliation:
Meitu, Inc.

Ranjanisi team

Title: Image Dehazing by Joint Estimation of Transmittance and Airlight using Bi-Directional Consistency Loss Minimized FCN
Members: *Ranjan Mondal (ranjan.rev@gmail.com)*, Sanchayan Santra, Bhabatosh Chanda
Affiliation:
Indian Statistical Institute, Kolkata, India

Scarlet Knights team

Title: Multi-scale single image dehazing using perceptual pyramid deep network
Members: *Vishal M. Patel (vishal.m.patel@rutgers.edu)*, He Zhang, Vishwanath A. Sindagi
Affiliation:
Department of Electrical and Computer Engineering Rutgers University, Piscataway, NJ 08854, US

SiMiT-Lab team

Title: Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing
Members: *Anil Genc (genca16@itu.edu.tr)*, Deniz Engin, Hazim Kemal Ekenel
Affiliation:
SiMiT-Lab, Istanbul Technical University, Turkey

Xdhm team

Title: Joing transmission and atmosphere estimate for single image dehaze
Members: *Ming Hong (1184307675@qq.com)*, Jingying Huang, Yizhi Chen, Erin Chen, Xiaoli Yu, Tingting Wu
Affiliation:
Xiamen University, China

References

[1] C. Ancuti and C. Ancuti. Single image dehazing by multi-scale fusion. *IEEE Transactions on Image Processing*, 22(8):3271–3282, 2013. 1

[2] C. Ancuti, C. O. Ancuti, A. Bovik, and C. De Vleeschouwer. Night time dehazing by fusion. *IEEE ICIP*, 2016. 1

[3] C. Ancuti, C. O. Ancuti, and C. De Vleeschouwer. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. *IEEE ICIP*, 2016. 2

[4] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer. I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images. In *arXiv:1804.05091*, 2018. 2, 4

[5] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 2, 4

[6] D. Berthelot, T. Schumm, and L. Mets. Began: Boundary equilibrium generative adversarial networks. In *arXiv preprint 1703*, 2017. 6

[7] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 2016. 1

[8] L. K. Choi, J. You, and A. C. Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *In IEEE Trans. on Image Processing*, 2015. 1

[9] C. Dong, C. C. Loy, K. He, and X. Tang. *Learning a Deep Convolutional Network for Image Super-Resolution*, pages 184–199. Springer International Publishing, Cham, 2014. 8

[10] D. Engin, A. Genc, and H. Ekenel. Cycle-dehaze: Enhanced cyclegan for single image dehazing. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 3

[11] R. Fattal. Single image dehazing. *SIGGRAPH*, 2008. 1

[12] A. Galdran, A. Alvarez-Gila, A. Bria, J. Vazquez-Corral, and M. Bertalmo. On the duality between retinex and image dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 3, 7

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Gen-

erative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014. 7

[14] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *In IEEE CVPR*, 2009. 1, 7, 8

[15] P. Isola. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 3, 6, 7

[16] J. E. Khoury(B), J.-B. Thomas, and A. Mansouri. A color image database for haze model and dehazing methods evaluation. *ICISP*, 2016. 2

[17] S. Ki, H. Sim, S. Y. Kim, J.-S. Choi, S. Kim, and M. Kim. Fully end-to-end learning based conditional boundary equilibrium gan with receptive field sizes enlarged for single ultra-high resolution image dehazing. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 3, 6

[18] H. Koschmieder. Theorie der horizontalen sichtweite. In *Beitrage zur Physik der freien Atmosphare*, 1924. 6

[19] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 6

[20] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. *In IEEE Int. Conf. on Computer Vision*, 2013. 1

[21] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *In IEEE Trans. on Image Processing*, 2012. 1

[22] R. Mondal, S. Santra, and B. Chanda. Image dehazing by joint estimation of transmittance and airlight using bi-directional consistency loss minimized FCN. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 3, 6

[23] K. Nishino, L. Kratz, and S. Lombardi. Bayesian defogging. *International Journal on Computer Vision*, 2012. 1

[24] W. Ren, S. Liu, H. Zhang, X. C. J. Pan, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. *Proc. European Conf. Computer Vision*, 2016. 1

[25] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. 3, 5, 6

[26] M. S. Sajjadi, B. Schölkopf, and M. Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. In *IEEE International Conference on Computer Vision (ICCV)*, 2017. 7

[27] H. Sim, S. Ki, S. Y. Kim, J.-S. Choi, S. Kim, and M. Kim. High-resolution image dehazing with respect to training losses and receptive field sizes. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 3, 6

[28] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*, 2015. 7

[29] R. T. Tan. Visibility in bad weather from a single image. *In IEEE Conference on Computer Vision and Pattern Recognition*, 2008. 1

[30] K. Tang, J. Yang, and J. Wang. Investigating haze-relevant features in a learning framework for image dehazing. *In IEEE Conference on Computer Vision and Pattern Recognition*, 2014. 1

[31] J.-P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. *In IEEE ICCV*, 2009. 1

[32] J.-P. Tarel, N. Hautire, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intelligent Transportation Systems Magazine*, 2012. 2

[33] R. Timofte, R. Rothe, and L. Van Gool. Seven ways to improve example-based single image super resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 4, 5

[34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004. 3

[35] H. Zhang and V. M. Patel. Densely connected pyramid dehazing network. In *CVPR*, 2018. 4

[36] H. Zhang, V. Sindagi, and V. Patel. Multi-scale single image dehazing using perceptual pyramid deep network. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. 3, 4

[37] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3929–3938, 2017. 7

[38] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *IEEE International Conference on Computer Vision (ICCV)*, 2017. 7