



INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv



Single Image Dehazing Using Deep Learning

Cahyo Adhi Hartanto^{a,*}, Laksmi Rahadiani^a

^aFaculty of Computer Science, Universitas Indonesia, Depok, Indonesia
Corresponding author: *cahyo.adhi91@ui.ac.id

Abstract— Many real-world situations such as bad weather may result in hazy environments. Images captured in these hazy conditions will have low image quality due to microparticles in the air. The microparticles light to scatter and absorb, resulting in hazy images with various effects. In recent years, image dehazing has been researched in depth to handle images captured in these conditions. Various methods were developed, from traditional methods to deep learning methods. Traditional methods focus more on the use of statistical prior. These statistical prior have weaknesses in certain conditions. This paper proposes a novel architecture based on PDR-Net by using a pyramid dilated convolution and pre-processing modules, processing modules, post-processing modules, and attention applications. The proposed network is trained to minimize L1 loss and perceptual loss with the O-Haze dataset. To evaluate our architecture's result, we used structural similarity index measure (SSIM), *peak signal-to-noise ratio* (PSNR), and color difference as an objective assessment and psychovisual experiment as a subjective assessment. Our architecture obtained better results than the previous method using the O-Haze dataset with an SSIM of 0.798, a PSNR of 25.39, but not better on the color difference. The SSIM and PSNR results were strengthened by using subjective assessments and 65 respondents, most of whom chose the results of the restoration of the image produced by our architecture.

Keywords— Single image dehazing; deep learning; image restoration; image quality assessment.

*Manuscript received 15 Nov. 2020; revised 15 Dec. 2020; accepted 2 Feb. 2021. Date of publication 31 Mar. 2021.
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.*



I. INTRODUCTION

The atmospheric light around the scene object dramatically influences the image quality captured by the camera. One of the factors that affect light is foggy weather, which results in decreased image quality [1]. This natural phenomenon occurs because the components or microparticles scatter in the air or the earth's atmosphere, consisting of dust, smoke, etc. These particles cause low contrast and color distortion due to scattered light [2]. Furthermore, hazy conditions also affect computer vision areas, such as automatic navigation systems for automatic vehicle navigation systems, causing the system cannot correctly pick up conditions around the vehicle, such as lane detection, other vehicle positions, and pedestrian detection [3]. The above problems cause image dehazing to become a hot topic of discussion in recent years.

Image dehazing is one of the most challenging sub-fields in computer vision. This is because hazy conditions can be unique in various regions. Hazy image conditions are closely related to the attenuation process that includes both absorption and scattering. The attenuation that occurs can be described using the transmission. Thus, the hazy-free image can be obtained by estimating the transmission. Various methods

have developed from traditional methods based on the atmosphere scattering model equation to deep learning methods. Several traditional methods, such as DCP developed by He, Sun, and Tang [4], then improvised DCP by Meng *et al* [5] and other traditional methods, will be described in Section II.

In general, the steps of the traditional method for image dehazing include: (1) predicting transmission map (t); (2) estimating global atmospheric light; (3) predicting hazy-free images based on the parameters in steps 1 and 2 [6]. The traditional method has a weakness because the prediction results are obtained based on statistical processes, highly dependent on light conditions and haze concentration. The statistical process produces low hazy-free images when the scene objects have a large airtight region.

Deep learning is overgrowing nowadays. Various deep learning architectures have been developed for single image dehazing, including DehazeNet [7], MSCNN [8], FFA-Net [9], GridDehaze-Net [10] and PDR-Net [11]. The advantage of deep learning is that it can directly predict the transmission map from a hazy image. In this paper, we developed an architecture that takes the concept from the PDR-Net [11]. These developments include pyramid dilated convolution,

dilated convolution, and application of attention. We implement dilated convolution to pyramid dilated modules to simplify the architecture, support the exponential expansion of the receptive field without losing resolution, and improve performance. Dilated convolution requires the same computations even though it has a larger receptive field [12]. Attention has been used in various deep learning architectures and has successfully resulted in better performance. We implemented attention mechanisms to retain useful information and give more weight to important information. The attention mechanism was adopted from the FFA-Net [9]. The architecture we have developed also adopts the ResNet skip connection architecture [13].

II. MATERIAL AND METHOD

A. Related Works

A hazy image can be described using an atmospheric scattering model that can be formally written as follows.

$$I(x) = J(x)t(x) + A(x)(1 - t(x)) \quad (1)$$

x indicates the pixel position, A indicates the global atmospheric light, $I(x)$ represents the hazy image, $J(x)$ represents the hazy-free image and t represents the hazy-image transmission map. t can be described mathematically with the assumption that the media passed is homogeneous. t can be described in exponential form as follows.

$$t(x) = e^{-\beta d(x)} \quad (2)$$

d is the object's distance to be taken, and β is the sum of the absorption and scattering coefficients.

Traditional methods focus more on the use of contrast, saturation, and dark channels [14]. For example, the method introduced by [4], namely DCP (dark channel prior), is an image restoration using channel values that have value close to zero as a recovery reference. DCP has weaknesses in several conditions, such as image conditions with airtight scene objects. DCP was further developed by [5] by adding L1-norm regularization. The addition has proven to improve the haze-free image quality because it reduces artifacts. In 2014, Fattal [15] developed a method using color lines combined with a Markov Random Field Model to remove noise and artifacts. The method of combining several different input sizes was introduced [16]. They used three inputs, 20 x

20, 80 x 80 from the original size of 800 x 800, and the Laplacian result. Multi-scale fusion is a reliable solution for various conditions, either day or night.

With technological developments, deep learning can overcome traditional methods and computer vision problems. In 2016, Cai *et al* [7] introduced Dehaze-Net that consist of four main parts: feature extraction, multi-scale mapping, local extremum, and non-linear regression. Besides, they also introduced BRELU, to solve ReLU problems that are not suitable for regression problems. In the same year, Ren *et al* [17] developed a multi-scale CNN. There are two parts of the method being developed, namely coarse-scale network and fine-scale network. A coarse-scale network will predict the transmission map and the results will be modified by a fine-scale network which will combine the input with coarse-scale network results. In 2019, Liu *et al* [10] introduced GridDehazeNet with three main parts: pre-processing module, backbone, and post-processing modules. Besides, GridDehazeNet also involved attention-based multi-scale, which aims to capture information from multiple scales. In 2019, Qin [9] introduced FFA-Net that used attention-based features to retain information from shallow layers to deep layers and learn different-level features. The attention-based feature consists of two parts, namely channel attention, and pixel attention mechanism. In 2020, Li *et al* [11] developed the PDR-Net, which has two main parts: haze removal subnetwork and refinement subnetwork. Haze removal subnetwork aims to remove hazy first, and a refinement subnetwork enhances the results. PDR-Net architecture is an architecture that is the basis for architectural development in this research.

B. Proposed Method

The architecture consists of two subnetworks, namely haze removal subnetwork and refinement subnetwork. Each convolutional layer parameter is denoted as "kernel size x output feature maps x dilation rate". Our architecture uses zero paddings to reduce boundary artifacts. Following Gridach and Voiculescu [18], we use summation operation because it shows better performance than concatenation operation and works better in capturing information. The haze removal network and refinement network consist of several connections because the two subnetworks implement pyramid dilated convolution [18]. Fig. 1 and Fig. 2 illustrate the proposed architecture in detail.

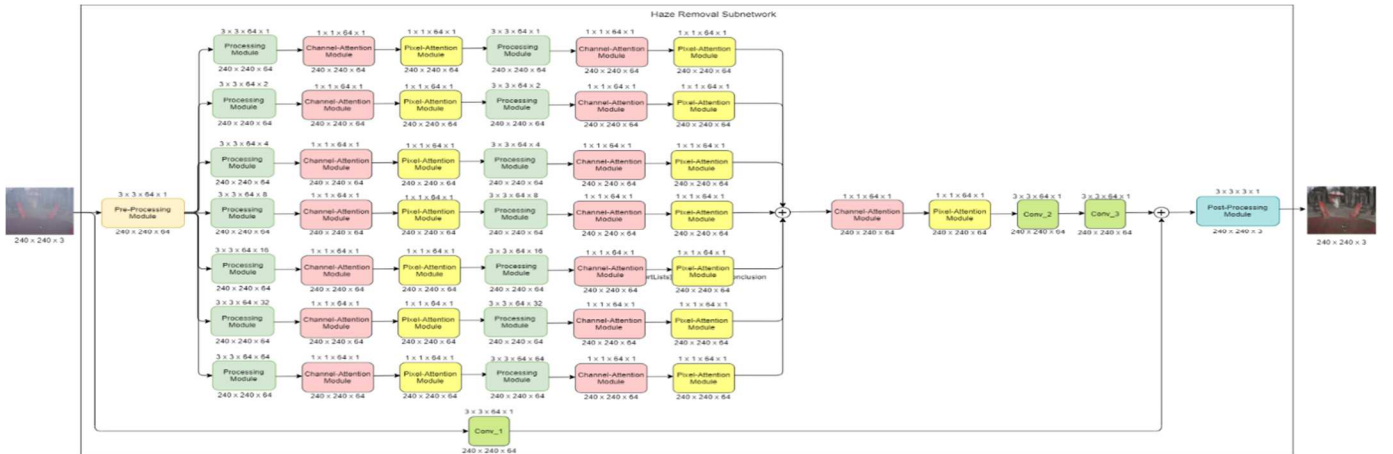


Fig. 1 An overview of the haze removal subnetwork

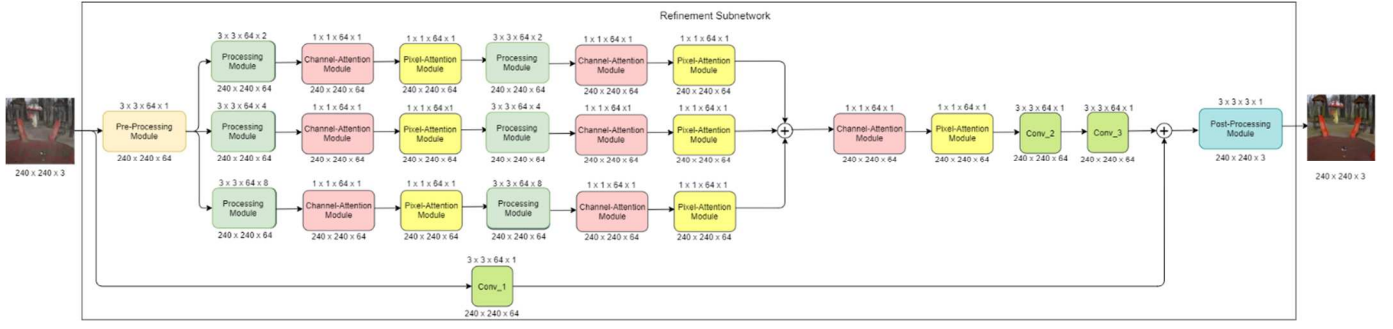


Fig. 2 An overview of the refinement subnetwork

1) *Dilated Convolution*: Dilated convolution was first developed for wavelet decomposition [19]. The idea of dilated convolution is to insert zero values between the pixels of convolutional filters. The important parameter in the dilated convolution is the dilation rate that indicates kernel gaps. If the dilation rate is one, it means the layer is still in standard convolution. In general, if the dilation rate is n , the pixel value is skipped in $n-1$ pixel. The advantage of dilated convolution is enlarging the receptive field without requiring additional parameters and computationally efficient [18]. Dilated convolution has been applied in the field of computer vision, including image segmentation [20], object detection [21], and simple deep learning architecture [22].

2) *Pyramid Dilated Convolution*: The receptive field plays an essential role because it shows the amount of information used. Deep learning architectures often use a pooling layer or stride convolution to expand the receptive field. The use of these two layers often causes failure because spatial information is often lost. This problem can be

overcome by gradually increasing the dilation rate in the dilated convolution [18].

3) *Haze Removal and Refinement Processing Module*: The module consists of 11 convolutional layers followed by the ReLU activation function except "Conv_11". In this module, we apply the skip connection, which is implemented in ResNet [13]. Skip connection works to maintain information and makes forward and backward passes more accessible. Skip connection allows thin haze region information and low-frequency information to be passed [9]. "Conv_9 + ReLU" aims to refine several convolution layers features before the features are proceeded by the attention layer. "Conv_11" does not use an activation function because it uses for the latent reconstruction, which can be obtained from the sum of the main connection features and the skip connection. The refinement processing module's skip connection is not as complicated as the haze removal processing module and does not have an attention module. Fig. 3 illustrates the haze removal-processing and refinement processing module.

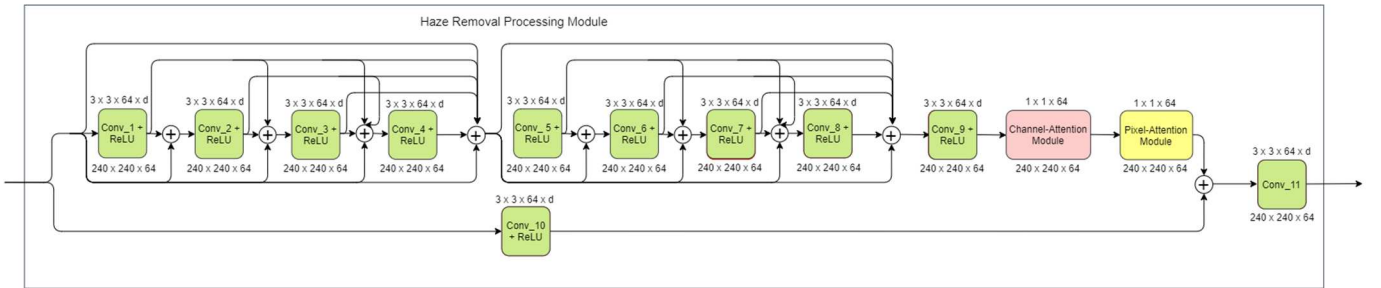


Fig. 3 An overview of the haze removal processing module

4) Pre-processing and Post-processing Module:

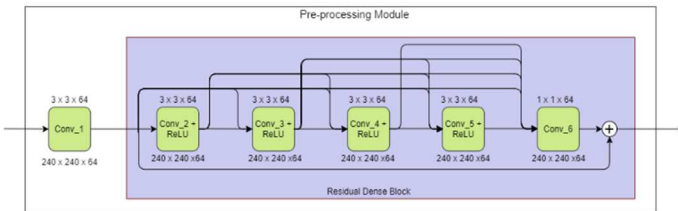


Fig. 4 An overview of the pre-processing module

The idea of the pre-processing module was adopted from the GridDehaze-Net architecture [10]. The module consists of a convolutional layer without an activation function followed by a residual dense block (RDB). The pre-processing module in the haze removal subnetwork and refinement subnetwork

produces 64 feature maps and aims to make pre-processing more efficient and relevant. The post-processing module is preceded by a residual dense block (RDB), followed by a convolutional layer without an activation function. Fig. 4 illustrates the pre-processing module.

5) Channel Attention Module:

The modules are adopted from FFA-Net [9]. The received information is channel-wise global spatial, which is processed first using global average pooling. The channel-attention (CA) module assigns a different weight to each channel, thus providing additional information. Global average pooling results are processed using two convolution layers, followed by ReLU in the first and sigmoid layers in the second layer to obtain different weights. Fig. 5 illustrates the channel-

attention module. Global average pooling is represented as follows.

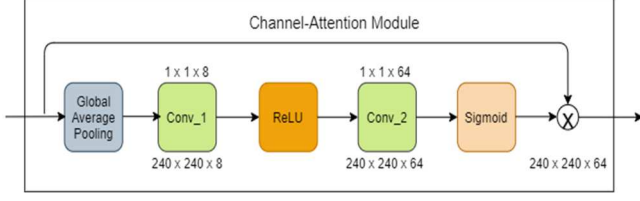


Fig. 5 An overview of the Channel-Attention Module

$$g_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_c(i, j) \quad (3)$$

The result of the operation is $C \times 1 \times 1$, where C is the number of channels. $X_c(i, j)$ denotes X on the c -th channel and is located in i, j . H_p is the global average pooling. The next operation of the CA module can be formulated as follows.

$$CA_c = \sigma(\text{Conv}(\delta(\text{Conv}(g_c)))) \quad (4)$$

$$\text{output} = CA_c \otimes F_c \quad (5)$$

Where δ is ReLU and σ is sigmoid. The final step is an element-wise multiply operation between the input F_c and the weight of CA_c , which produces 64 feature maps.

1) *Pixel-Attention Module*: Similar to the channel-attention module, this module is also adopted from the FFA-Net. The module aims to create an architecture that can better capture information on thick hazed pixels and high-frequency images. The module structure is almost the same as CA module, but does not have global average pooling, and the two convolutional layers directly receive information input. The output PA operation before the element-wise operation is $1 \times H \times W$ and can be formulated as follows (with an example input I).

$$PA = \sigma(\text{Conv}(\delta(\text{Conv}(I)))) \quad (6)$$

where δ is ReLU, and σ is sigmoid. The final step is the same as for the channel-attention module, namely the element-wise multiply operation between PA's input F and weight. The output results is 64 feature maps can be formulated as follows.

$$\text{output} = F \otimes PA \quad (7)$$

2) *Loss Function*: Deep learning architecture can get optimum results if the chosen loss function is right. L1 loss is preferred over L2 because L1 has better performance than L2, and some traditional methods that use L2 produce blurry results [23]. L1 loss represented as follows.

$$L_1 = \frac{1}{N} \sum_{x=1}^N \sum_{i=1}^3 F_s(I_i(x) - G_i(x)) \quad (8)$$

with

$$F_s(e) = \begin{cases} 0.5e^2, & \text{if } |e| < 1, \\ |e| - 0.5, & \text{otherwise.} \end{cases} \quad (9)$$

where N denotes the total pixels, $I_i(x)$ indicates the color intensity of pixel x in the dehazed image (hazy-free image), and $G_i(x)$ is the ground truth. Network optimization is also done by applying perceptual loss. Perceptual loss aims to minimize perceptual differences between dehazed images and ground-truth images (high-level differences), strengthen fine features, and retain color information [24]. We extracted the

features of the three activation layer of VGG16. The perceptual loss represented as follows.

$$L_p = \sum_{j=1}^3 \frac{1}{C_j H_j W_j} \|\phi_j(I) - \phi_j(G)\| \quad (10)$$

where $C_j H_j W_j$ denotes feature maps of the dehazed image and ground truth images, while ϕ_j denotes the perceptual feature of the VGG16. We combined these two loss functions and represented as follows.

$$L_{total} = L_1 + \lambda L_p \quad (11)$$

λ is a parameter that functions to adjust the weight between L1 loss and perceptual loss. We set λ to 0.04.

3) *Quality Measure*: The resulting image restoration results are measured using three metrics: Structural Similarity (SSIM) [25], Peak Signal to Noise Ratio (PSNR) [26], and CIE Color Difference Metric [27]. The higher SSIM value shows that the resulting image is structurally close to the ground-truth. SSIM value ranges between $[-1, 1]$. The PSNR is a quality measurement of the ratio of signal to noise between two images. The range of the PSNR is between $[0, \infty]$. A high PSNR value indicates better image quality. The color difference is a metric to assess the color difference between two images. Color difference considers chroma and hue for blue color performance and scaling factor for gray color performance. The range of color differences between 0 and 100.

Additionally, we also used a psychovisual experiment by distributing questionnaires randomly. The questionnaire is divided into two parts. The first part consists of 11 slides from the test set A with details of each slide consisting of 3 images. All images are placed on a 50% gray background (#7F7F7F). In the first part, respondents are asked to choose one of the more visually pleasing images between the right and left images. The right image is the result of our architecture, while the left image is the result of MSCNN. The second part consists of 4 slides from test set B. Each slide consists of 2 images. The left image is a hazy-image, and the right image is the result of our architecture. Respondents are asked to give ratings from 1 to 5, where 1 is the worst quality, and 5 is the best quality. Fig. 6 and Fig. 7 gives an overview of our psychovisual experiment.



Fig. 6 An overview of the first part of psychovisual experiment

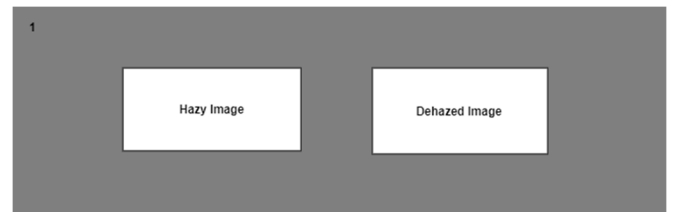


Fig. 7 An overview of the second part of psychovisual experiment

III. RESULTS AND DISCUSSION

A. O-Haze Dataset

Ancuti [28] in 2018 revealed that O-Haze consists of 45 pairs of original outdoor images. The image is taken in cloudy conditions, either in the morning or at sunset, and only with wind speeds below 3 km/h. The training and testing data distribution is the same as the original paper, namely 34 for training and 11 for test set A. The data was augmented randomly rotated by 90, 180, and 270, and horizontal flip to get more general results. Besides, we also took four random images without ground-truth from Google Image to test our architecture ability to perform single image dehazing in real-world conditions. The real-condition image is used to construct test set B.

B. Training Details

The architecture was trained using ADAM with epsilon 1e-8 and exponential decay (β_1, β_2) 0.9 and 0.999 with a batch size of 1. We adopt an annealing strategy for determining the learning rate every step. We initialize the initial learning rate

with a value of 1×10^{-4} . The annealing strategy will bring the learning rate closer to 0 as the step increase. The implementation uses the cosine function and represented as follows.













































$$\eta_t = \frac{1}{2} \left(1 + \cos \left(\frac{t\pi}{T} \right) \right) \eta \quad (12)$$

where η is initial learning rate, T is batch size and t is steps.

C. Objective and Subjective Results

Table I is the result of our comparison between our architecture results with MSCNN [17]. In the test set A1, our dehazed image has a dark color in the hazy region so that the road conditions that should be gray turn black. While the test set A 2-5, the dehazing image still has a thin hazy in some areas, even though the object is very clear. In test set A 6, the haze is not visible, but the color still does not resemble the ground truth. Overall, our architecture produces colors and structures that almost resemble ground-truth, even though some images still have noise.

TABLE I
THE COMPARISON BETWEEN OUR ARCHITECTURE AND MSCNN IN TEST SET A

	Set 1	Set 6	Set 10	Set 19	Set 20	Set 21	Set 27	Set 30	Set 33	Set 41	Set 42
Hazy											
Ren et al											
Ours											
GT											

Meanwhile, the MSCNN results in the A1 test set produced a more natural image in terms of color even though a thin haze was still visible. In test set A 2-11, MSCNN produced an image with visible objects, but still covered by thin hazy and artifacts so that the resulting colors do not resemble ground-truth. Table II and Table III provides information on PSNR, SSIM, and color difference.

TABLE II
THE QUANTITATIVE RESULTS OF A METHOD COMPARISON BETWEEN IN THE TEST SET A

Set	Ren et al			Ours		
	SSIM	PSNR	ΔE_{00}^*	SSIM	PSNR	ΔE_{00}^*
1	0.81	14.93	18.17	0.74	16.91	19.60
6	0.72	16.10	13.20	0.71	18.45	17.45
10	0.80	15.22	13.70	0.81	22.68	17.53
19	0.83	15.82	12.94	0.76	21.00	16.94
20	0.63	13.17	20.98	0.68	18.12	21.87
21	0.73	14.49	20.26	0.79	23.39	26.27
27	0.71	17.00	14.16	0.79	22.52	16.50
30	0.82	18.62	12.66	0.81	23.03	14.5
33	0.88	19.52	10.87	0.84	26.84	16.30
41	0.88	19.71	12.34	0.86	24.26	11.06
42	0.72	15.57	12.87	0.73	21.43	18.51

TABLE III
THE RESULTS OF A METHOD COMPARISON BETWEEN MSCNN ON THE ENTIRE O-HAZE

Metrics	Ren et al	Our
SSIM	0.765	0.798
PSNR	19.068	25.39
ΔE_{00}^*	14.670	18.63

To verify that our architectural results were better than MSCNN, we conduct a subjective assessment of the 65 respondents shown in Fig. 8. Based on Fig. 8, respondents prefer the results of our architectural restoration in the test set A 2-11, while in the test set A1, respondents prefer the results and dark color that produced by our results. The number of assessments in the A 4 test set is slightly different because the haze conditions in the restoration results from both methods still have many hazy regions of the restoration from MSCNN due to the loss of some details. However, our models show less hazy region than MSCNN results. In the test set B, the respondents give various ratings from 1 to 5. The average value of each image's assessment is shown in Fig. 9 and shows that the real-condition of restoration images have medium to good quality.

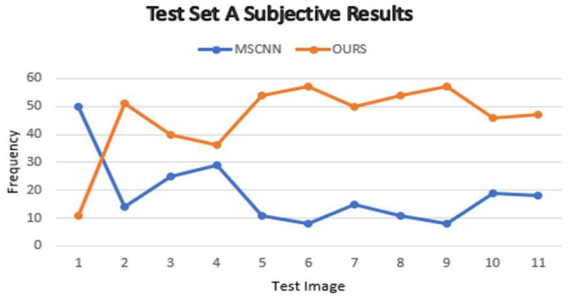


Fig. 8 The subjective results of Test Set A

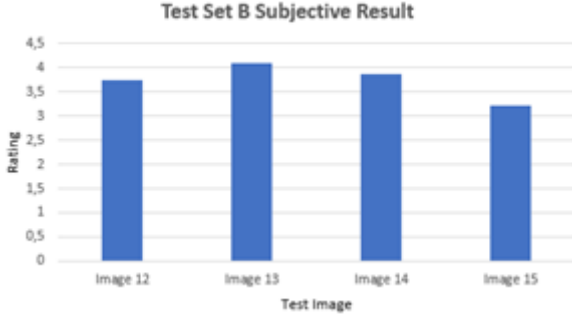


Fig. 9 The subjective results of Test Set B

Respondents are also asked to provide comments regarding the resulting image results. Some respondents considered the resulting image to be inconsistent and less sharp. They also commented that some images are less natural, less clear colors, and do not do well in image restoration in thick hazy areas. Besides, they also argue that there are missing objects such as the top of the building in the image test set B 4.

D. Evaluation on Real Image

The sky region is the most challenging thing in single image dehazing because the sky and haze region have the same color. In Table IV, our architecture can perform restoration on non-airlight parts well. It can be seen that our architecture can find all objects in all real-conditions test images which covered by haze. However, the result of image restoration in test set B 2-3, shows that the sky area slightly shifts towards yellow/ red and distorted. This is because the O-Haze dataset has less data, so that the architecture overfit the color scheme, structure, and appearance of O-Haze.

TABLE IV
THE RESULTS OF OUR ARCHITECTURE RESULTS IN REAL-CONDITION IMAGE

Set	Hazy Image	Ours	Set	Hazy Image	Ours
1			3		
2			4		

IV. CONCLUSION

In this paper, we develop a PDR-Net based architecture. Our network is trained end-to-end and does not rely on transmission maps and atmospheric light. We implemented a pyramid-dilated convolution in the architecture to maintain spatial information over a wide range of receptive fields. The

architecture consists of a pre-processing module, processing module, channel-attention module, pixel-attention module, and post-processing module. The network that we have developed is trained to minimize the L1 loss and perceptual loss functions. The experimental results show the best performance for the O-Haze test data. The quantitative results are supported by our psychovisual experiment, where on average, the respondents prefer the restoration results with our architecture over other methods. However, our architecture still has a weakness and needs improvement to be applied in real-world image restoration.

REFERENCES

- [1] H. Fu, B. Wu, Y. Shao, and H. Zhang, "Perception Oriented Haze Image Definition Restoration by Basing on Physical Optics Perception Oriented Haze Image Definition Restoration by Basing," *IEEE Photonics J.*, vol. 10, no. 3, pp. 1–16, 2018.
- [2] Y. Liang, K. Zhao, W. Zhang, and Y. Li, "Research on single image haze removal algorithm based on parameter optimization search of linear model," *2018 Int. Conf. Adv. Mechatron. Syst.*, pp. 326–331, 2018.
- [3] X. Min, G. Zhai, K. Gu, X. Yang, and X. Guan, "Objective Quality Evaluation of Dehazed Images," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–14, 2018.
- [4] K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 2351–2353, 2011.
- [5] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient Image Dehazing with Boundary Constraint and Contextual Regularization," *IEEE Int. Conf. Comput. Vis.*, 2013.
- [6] B. Li *et al.*, "Benchmarking Single Image Dehazing and Beyond," *IEEE Trans. Image Process.*, vol. PP, no. 8, p. 1, 2018.
- [7] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An End-to-End System for Single Image," *Springer Int. Publ.*, pp. 1–13, 2016.
- [8] H. Zhang, V. Sindagi, and V. M. Patel, "Multi-scale Single Image Dehazing using Perceptual Pyramid Deep Network."
- [9] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature Fusion Attention Network for Single Image Dehazing," *AAAI Conf. Artif. Intell.*, 2020.
- [10] X. Liu, Y. M. Ma, Z. Shi, and C. Jun, "GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing," *CVPR 2019*, pp. 7314–7323, 2019.
- [11] C. Li, C. Guo, J. Guo, P. Han, H. Fu, and S. Member, "PDR-Net: Perception-Inspired Single Image Dehazing Network With Refinement," vol. 22, no. 3, pp. 704–716, 2020.
- [12] F. Yu and V. Koltun, "Multi-scale Context Aggregation by Dilated Convolutions," *ICLR*, 2016.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016.
- [14] H. Park, D. Park, D. K. Han, and H. Ko, "Single Image Haze Removal Using Novel Estimation of Atmospheric Light and Transmission," *Int. Conf. Image Process.*, no. 2, pp. 4502–4506, 2014.
- [15] R. Fattal, "Dehazing Using Color-Lines," *ACM Trans. Graph.*, vol. 34, no. 1, pp. 1–14, 2014.
- [16] C. Ancuti, C. O. Ancuti, and A. C. Bovik, "Night time dehazing by fusion," *Int. Conf. Image Process.*, 2016.
- [17] W. Ren, S. Liu, H. Zhang, J. Pan, and X. C. B., "Single Image Dehazing via Multi-scale Convolutional Neural Networks," *Springer Int. Publ.*, vol. 1, pp. 154–169, 2016.
- [18] M. Gridach and I. Voiculescu, "Oxendonet: A Dilated Convolutional Neural Networks for Endoscopic Artefact Segmentation," *EndoCV2020*, pp. 2–5, 2020.
- [19] M. Holschneider and R. Kronland-martinet, "A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform," *Springer*, 1990, no. August 2014, 1989.
- [20] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *CVPR 2015*, pp. 3431–3440, 2015.
- [21] L. Yan, B. Zhong, and W. Song, "Region-based Fully Convolutional Networks for Vertical Corner Line Detection," *2017 Int. Symp. Intell. Signal Process. Commun. Syst.*, pp. 159–163, 2017.
- [22] L. Chen and A. L. Yuille, "Attention to Scale: Scale-aware Semantic Image Segmentation," 2016.

- [23] H. Zhang, V. Sindagi, and V. M. Patel, "Multi-scale Single Image Dehazing using Perceptual Pyramid Deep Network," *2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 1015–101509, 2018.
- [24] Y. Qu, Y. Chen, J. Huang, and Y. Xie, "Enhanced Pix2pix Dehazing Network," *CVPR 2017*, pp. 8160–8168, 2017.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment : From Error Visibility to Structural Similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
- [26] Z. Kotevski and P. Mitrevski, "Experimental Comparison of PSNR and SSIM Metrics," *ICT Innov. 2009*, 2009.
- [27] M. R. Luo, G. Cui, and B. Rigg, "The Development of the CIE 2000 Colour-Difference Formula ;," *CIE*, pp. 340–350, 2001.
- [28] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE : a dehazing benchmark with real hazy and haze-free outdoor images," 2018.