

Image Dehazing Algorithm Based on Conditional Generation against Network

Yuan Yao, Xiangyang Li, Jinhui Meng

College of Physical Science and Technology, Central China Normal University, Wuhan Hubei
Email: kent_xyli@163.com

Received: Nov. 9th, 2019; accepted: Nov. 25th, 2019; published: Dec. 2nd, 2019

Abstract

This paper improves an image dehazing algorithm based on conditional generation against networks. Using 32 Layer Tiramisu instead of U-Net in the generator network model can reduce training parameters and improve parameter utilization. The final layer of the discriminator network uses the Sigmoid function to complete the normalization. The network model is trained using the weighted sum of loss function, using a score to evaluate the ability of defogging, saving the model and parameters with higher scores during the training process, and finally using the network model with the highest score and parameters to perform the dehazing test on the outdoor real image. The training set uses real fog-free images and synthetic foggy images both indoors and outdoors. The test results show that the proposed model has improved subjective visual effects, image detail and color compared with the traditional dehazing method.

Keywords

Dehazing, Deep Learning, Conditional Generation against Networks, Convolutional Neural Network

基于条件生成对抗网络的图像去雾算法

姚 远, 李向阳, 孟金慧

华中师范大学物理科学与技术学院, 湖北 武汉
Email: kent_xyli@163.com

收稿日期: 2019年11月9日; 录用日期: 2019年11月25日; 发布日期: 2019年12月2日

摘 要

本文改进了一种基于条件生成对抗网络的图像去雾算法。在生成器网络模型中使用32层Tiramisu代替U-Net可以减少训练参数, 提高参数利用率。判别器网络的最后一层使用Sigmoid函数完成归一化。使用损失函数加权的方式训练网络模型, 使用一种评价去雾能力的分数, 保存训练过程中分数较高的模型

及参数,最后使用分数最高的网络模型及参数对室外真实图像进行去雾测试。训练集使用了室内和室外的真实无雾图像以及合成有雾图像,测试结果显示本文提出的模型和传统去雾方法相比,主观视觉效果、图像细节、色彩方面都有提高改善。

关键词

图像去雾,深度学习,条件生成式对抗网络,卷积神经网络

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

雾是由于大气光被吸收和散射造成的自然现象。在雾的影响下,直射光线被散射以后造成图像对比度、可见度以及色彩保真度的下降。相机拍摄会由于大气散射光与物体反射光造成图像表面亮度和色彩失真。此外,相机从场景点接收到的信号随着距离增大而衰减。雾带来的图像对比度和可见度下降问题对于某些行业影响很大,比如自动驾驶、交通监控、卫星遥感等领域。

经典的图像去雾算法主要分为基于图像增强去雾算法和基于大气散射模型去雾算法两大类。基于图像增强去雾算法主要有直方图均衡化,小波变换,Retinex去雾算法等。Tan等观察到通常无雾图像对比度比含雾图像高,提出增大图像局部区域对比度去雾算法[1]。基于大气散射模型的去雾算法根据图像景深信息计算出每个像素上的透射系数,由每个像素上的透射系数计算出透射图,利用透射图复原出无雾图像。Kopf等通过计算多张图像或物理三维模型的景深信息复原图像[2]。He等通过观察大量的图像以及实验研究,提出了基于暗通道先验的单张图像去雾算法[3]。He方法的缺点是在高亮度区域(如天空)容易出现色彩失真。在He去雾算法的基础上,Meng等提出限制透射率函数边界来更精确的估计出透射率[4]。Berman等提出了非局部区域去雾方法,他们观察到无雾图像中的颜色可以被RGB空间域100种不同颜色簇近似表示,而有雾图像的颜色像素会在RGB空间域形成雾线,利用这种雾线估计出每个像素的透射率[5]。以上方法都基于一个或多个雾相关特征的先验信息,使用深度学习算法计算雾相关特征可以提高去雾算法的鲁棒性。Cai提出的DehazeNet是一种能提取出图像相关特征的卷积神经网络(Convolutional Neural Network, CNN),这种模型可以计算透射率[6]。在DehazeNet上更进一步,Ren等使用多尺度深度卷积神经网络来计算透射率[7]。Goodfellow等人提出的生成对抗网络(Generative Adversarial Networks, GANs)框架可以将随机输入噪声输出逼真图像[8]。条件生成对抗网络(Conditional Generation Against Networks, cGAN)通过优化目标函数来学习从输入图像和随机噪声中生成清晰图像[9]。cGAN在超分辨率、图像修复和样式转移等图像处理领域得到了广泛应用。Zhang等联合使用三种U-Net神经网络模型引导图像去雾[10],第一个模型融合使用对抗损失与欧几里德损失的cGAN来估计透射率图;第二个模型先估计出雾相关特征,然后将雾相关特征与透射率联合在一起;第三个模型使用欧几里德损失函数和感知损失函数联合在一起输出无雾图像。

本文使用Tiramisu代替U-Net作为生成器网络模型,减少训练参数,提高参数利用率。判别器网络的最后一层借助Sigmoid函数完成归一化。使用损失函数加权的方式训练网络模型,提出一种评价去雾能力的分数,保存训练过程中分数较高的模型及参数,进而使用分数最高的网络模型及参数对室外真实图像进行去雾训练测试。

2. 相关理论与改进算法

2.1. 相关理论

大气散射模型通常被用来描述在大气光和雾霾共同作用下的成像原理[11] [12]:

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

其中, 等式右边的第一部分 $J(x)t(x)$ 为直接衰减项, 第二部分 $A(x)(1-t(x))$ 为大气光散射模型。 x 代表图像中的像素坐标, $I(x)$ 表示相机拍摄或人眼直接观察到的有雾图像, $J(x)$ 表示无雾图像, $A(x)$ 表示大气光值, $t(x)$ 表示介质透射率图。当大气光在图像中是均匀数值时, 透射率图可以表示为:

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

式中 β 代表衰减系数, $d(x)$ 代表场景深度。通过估计透射率图, 无雾图像可以使用下面等式表示:

$$J(x) = \frac{I(x) - A(x)(1-t(x))}{t(x)} \quad (3)$$

2.2. 条件生成对抗网络

cGAN 的损失函数由生成器和判别器的极小极大值对抗表示, 用下式可以表示对抗过程[13]:

$$\min_G \max_D (D, G) = E_{(x,y)} [\log(D(x,y))] + E_{(x,y)} [\log(1 - D(x, G(x,z)))] \quad (4)$$

其中 D 和 G 分别表示判别器和生成器。

输入有雾图像表示为 x , x 对应的无雾目标图像表示为 y , 给定一个随机矢量噪声 z , 生成器 G 生成与 x 相似的图像。本文使用一种评价去雾能力的分数, 保存训练过程中分数较高的模型及参数。输入单张真实含雾图像通过条件生成对抗网络模型后可输出去雾图像。

下面详细介绍该条件生成对抗网络的各个模块。本文使用 32 层 Tiramisu 取代 U-Net。这个模型的编码器和解码器端各有 5 个密度块, 连接层有 1 个密度块。在编码器端, 每一个密度块后紧跟一个下采样层。下采样层包含 (BatchNorm-Relu-Convolution) 操作。与编码器端相似, 在解码器端, 每一个密度块后紧跟一个上采样层。上采样层仅有一个逆卷积操作。编码器和解码器端的每个密度块各有 4 层。连接端的密度块有 15 层, 上采样层增长率设置为 12。图像的空间维度在经过一个下采样层以后空间维度减半, 经过一个上采样层以后空间维度翻倍。Simon 等提出的 100 层 Tiramisu 语义分割模型具有参数效率高的优点[14]。生成器网络结构如图 1。本文使用与 DCGANs 相同的判别器网络 (Patch GAN) [15]。此模型仅对比目标图像与生成图像区域上的不同, 而不是逐个像素对比。由于 CNN 在输出的感知域大于一个像素 (对应于原图像中的像素块), 因此使用填充控制感知域。本文使用 Pix2Pix GAN 中的 70×70 判别器, 在最后的特征图集上应用逐个像素对比方式。在特征图的有效感知域上覆盖一块图像区域用来消除图像中的大量伪影。在判别器网络的最后一层利用 Sigmoid 函数来完成特征映射, 以便使判别结果可归一化为 [0,1] [16], 判别器网络结构如图 2。

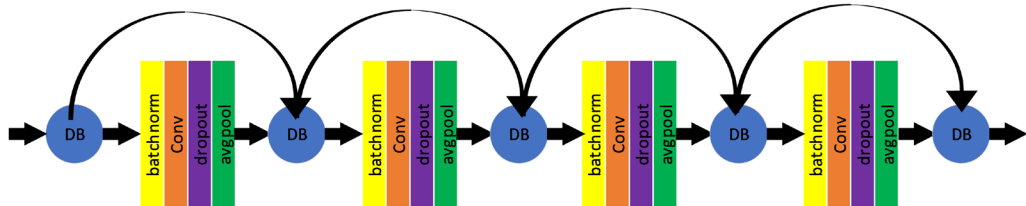


Figure 1. Generator network structure
图 1. 生成器网络结构

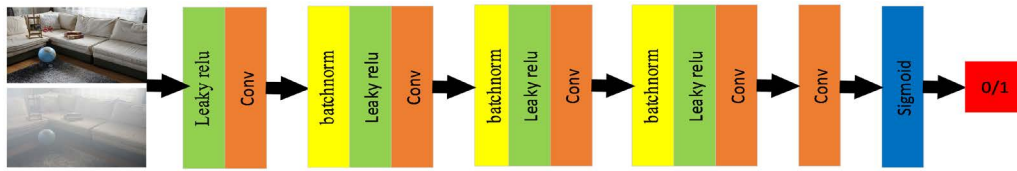


Figure 2. Discriminator network structure
图 2. 判别器网络结构

用来优化网络模型的损失函数为三个权重的加权组合，总损失的公式为：

$$Loss_{total} = W_{gan} \times L_{Adv} + W_{L1} \times L_{L1} + W_{vgg} \times L_{vgg} \quad (5)$$

其中 L_{Adv} 损失定义为：

$$L_{Adv} = E_{(x,y)} [\log(D(x,y))] + E_{(x,z)} [\log(1 - D(x,G(x,z)))] \quad (6)$$

在模型每次迭代中，判别器更新一次，生成器更新一次。最后将结果与权 W_{gan} 相乘加到总损失函数。如文献[13]所述：利用 L1 损失和对抗损失加权来减少输出图像中的伪影。目标图像 y 和生成图像 $G(x,z)$ 之间的 L1 损失计算如下：

$$L_{L1} = E_{x,y,z} [\|y - G(x,z)\|_1] \quad (7)$$

将结果与权 W_{L1} 相乘加到总损失函数。生成图像与目标图像通过一个预训练的 VGG-19 网络[17]，两张图像在通过池化 4 层后得到其 L2 损失，特征感知损失加入总损失函数，感知损失定义如下：

$$L_{vgg} = \frac{1}{CHW} \sum_{c=1}^C \sum_{w=1}^W \sum_{h=1}^H \|V(G(x,z)^{c,h,w}) - V(y^{c,h,w})\|_2^2 \quad (8)$$

其中 C, H, W 分别代表输出的通道，宽度和高度。感知损失值是一个常数，对于池化 4 层的输出为 $1e-5$ 。V 表示由 VGG-19 网络模型执行的非线性 CNN 变换。此结果与权 W_{vgg} 相乘加到总损失函数。

3. 实验方法

3.1. 数据集

为提高算法的鲁棒性，本文使用两个不同的数据集合并训练。首先使用 NYU-Depth V2 数据集合成含雾图像得到一个含有大量真实含雾/真实不含雾图像对的数据集，NYU-Depth V2 数据集由来自各种室内场景的视频序列组成，使用 Microsoft Kinect 的 RGB 深度摄像机记录[18]。该数据集包含有 1449 张室内图像和其对应的图像深度信息对(RGB-D)。为模拟浓雾图像，本文给传输系数一个服从平均分布[0.2,0.4]的随机值，利用 RGB-D 图像生成透射图(t)来获得图像中每一像素的透射系数，大气光值(A)设置为 1 然后利用等式(1)生成含雾图像。本文注意到 NYU 深度数据集只有室内场景，一些室外场景的特定特征(例如天空)在 NYU 数据集中是没有的。为了解决室外场景问题，本文使用 Codruta 等创建的室外场景数据库(名为 O-HAZE)，O-HAZE 数据集包含 45 对室外含雾/不含雾的图像[19]。将这些图像和之前生成的图像放到一起建立数据集并把数据集中所有图像统一调整大小为(256,256,3)。此数据集共有 1494 张大小为(256,256,3)的图像，图 3 为真实清晰图像和生成含雾图像对比。

3.2. 实验过程

式 5 中使用的权值分别为 $W_{vgg} = 9, W_{gan} = 3, W_{L1} = 80$ 。使用 Adam 优化器[20]，学习率的值设为 0.001。算法实现为使用 Google 公司的 Colab 平台：15 GB 内存，Nvidia Tesla K80 GPU。测试推理使用 Nvidia

公司的 Jetson TX2 平台。数据集被分为训练集，验证集，测试集。从数据集中随机选择 1200 张数据集的图像水平翻转后加到训练集。此时训练集共 2400 张图像，每次训练时图像顺序被随机打乱。数据集中剩余的 294 张图像再随机分为 94 张图像的验证集和 200 张图像的测试集。在每次迭代中，生成器更新一次，判别器更新一次。因此，每次迭代的每一次更新包含两个步骤。



Figure 3. Synthetic fog contrasts with real clear images
图 3. 合成有雾与真实清晰图像对比

4. 实验结果与综合分析

最大峰值信噪比(Peak Signal to Noise Ratio, PSNR)测量了算法将噪声(雾)从噪声图像中去除的能力，两张相同的图像会有一个无限大的 PSNR 值，更高的 PSNR 值意味着去雾能力更好。结构相似性(Structural Similarity Index Measure, SSIM)测量了两张图像的相似程度。完全相同的两张图像 SSIM 值为 1。本文定义了图像 PSNR 和 SSIM 的加权去雾分数作为一个拥有两种属性的指标。

PSNR 的权重值(W_{PSNR})设置为 0.05，SSIM 的权重值(W_{SSIM})设置为 1。一旦去雾分数在验证集中得到一个较高值时就保存模型，每次保存最大的 10 个检查点。本文还测试了表现较好的几个模型在真实含雾图像上的去雾能力，定量对比了这些模型在测试集上得到的 PSNR 和 SSIM 值，本文选择了文献[3] [5] [6] 提出的模型对比分析本文提出模型的去雾能力，表 1 列出了 PSNR，SSIM 对比值。

Table 1. Quantitative comparison of different dehazing methods
表 1. 不同去雾方法的定量对比

Model	He <i>et al.</i>	Berman <i>et al.</i>	Cai <i>et al.</i>	Ours method
PSNR	13.83	11.48	10.55	14.52
SSIM	0.71	0.50	0.69	0.89

从表 1 可以看出，在没有大片天空区域的情况下，He 算法去雾图像在结构相似度和最大峰值信噪比两个值较高。Cai 算法虽然比 Berman 算法的结构相似度值较高，但是从图 4 可以看出 Cai 算法的去雾图像看起来去雾效果并不好，而 Berman 算法的结构相似度值虽然较低，但是图 4 中去雾图像却比 Cai 算法的图像清晰。很明显，本文提出的算法模型与真实图像在结构相似度和最大峰值信噪比值都有很大提升，去雾图像看起来细节丰富，层次鲜明。为了证明本文模型的鲁棒性，本文使用了不包含在数据集中的真实有雾图像上进行了评估，图 4 列出了原始含雾图像以及对应的去雾图像，为本文算法和前文提到去雾算法去雾图像的主观对比。图 4 的 d 行图像中看出本文的算法在重雾条件下可以还原出较多图像细节，可以清晰的看到后面楼房的窗户。从 a 行与 c 行图像对比可以看出本文的算法可以比其他方法提取出相同或更多的细节，输出色彩很真实。从第 b 行和 d 行可以看出，图像的分辨率比其他图像有很大提高。

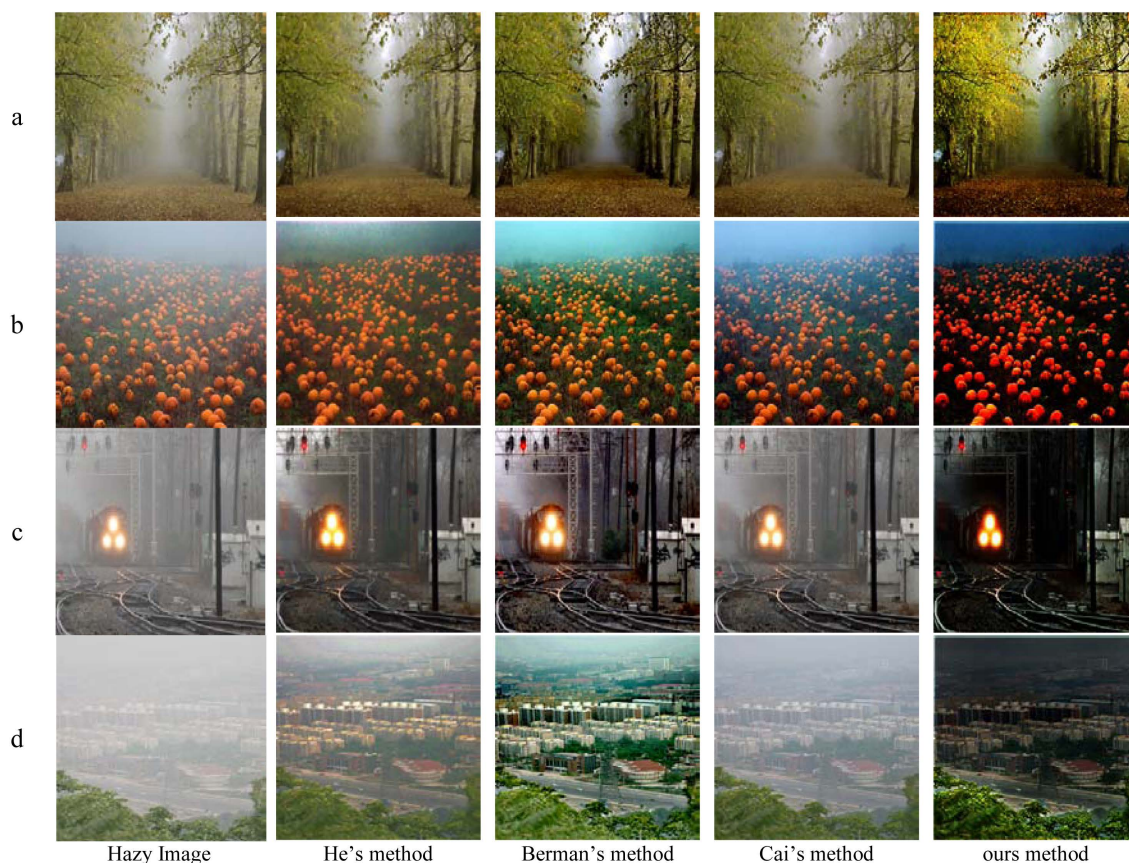


Figure 4. Defogging image comparison
图 4. 去雾图像对比

5. 结论

本文改进一个基于条件生成对抗网络的图像去雾算法模型，无需估计图像大气光和透射图以及额外图像深度信息，将输入的含雾图像直接输出去雾图像。通过将 U-Net 模型替换为 Tiramisu 模型提高参数利用率提取出图像更多信息。训练过程中使用的图像数据集分为室内和室外图像，提高去雾模型的鲁棒性，可以适应常见的图像。最后提供了本文模型与其他去雾算法的客观对比与主观对比，结果证明本文算法可以较好的复原出浓雾图像的框架和细节，定量分析显示本模型生成图像与原始图像结构相似度达到 0.89。综上所述，本文提出去雾模型可以很好的实现去雾效果，还原出图像的更多细节，对于雾天物体识别和雾天条件下监控有一些帮助。

参考文献

- [1] Tan, R.T. (2008) Visibility in Bad Weather from a Single Image. *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, 23-28 June 2008, 1-8.
- [2] Kopf, J., Neubert, B., Chen, B., Cohen, M., Cohen-Or, D., Deussen, O., Uyttendaele, M. and Lischinski, D. (2008) Deep Photo: Model-Based Photograph Enhancement and Viewing. *ACM Transactions on Graphics*, **27**, Article No. 116. <https://doi.org/10.1145/1457515.1409069>
- [3] He, K.M., Sun, J. and Tang, X.O. (2010) Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 2341-2353. <https://doi.org/10.1109/TPAMI.2010.168>
- [4] Meng, G.F., Wang, Y., Duan, J.Y., Xiang, S.M. and Pan, C.H. (2013) Efficient Image Dehazing with Boundary Constraint and Contextual Regularization. *Proceedings of the IEEE International Conference on Computer Vision*, Sydney,

- 1-8 December 2013, 617-624. <https://doi.org/10.1109/ICCV.2013.82>
- [5] Berman, D., Avidan, S., *et al.* (2016) Non-Local Image Dehazing. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 26 June-1 July 2016, 1674-1682. <https://doi.org/10.1109/CVPR.2016.185>
- [6] Cai, B., Xu, X.M., Jia, K., Qing, C.M. and Tao, D.C. (2016) Dehazenet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing*, **25**, 5187-5198. <https://doi.org/10.1109/TIP.2016.2598681>
- [7] Ren, W.Q., Liu, S., Zhang, H., Pan, J.S., Cao, X.C. and Yang, M.-H. (2016) Single Image Dehazing via Multi-Scale Convolutional Neural Networks. In: *European Conference on Computer Vision*, Springer, Cham, 154-169. https://doi.org/10.1007/978-3-319-46475-6_10
- [8] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014) Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, Montreal, 8-13 December 2014, 2672-2680.
- [9] Isola, P., Zhu, J.-Y., Zhou, T.H. and Efros, A.A. (2017) Image-to-Image Translation with Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 1125-1134. <https://doi.org/10.1109/CVPR.2017.632>
- [10] Zhang, H., Sindagi, V. and Patel, V.M. (2017) Joint Transmission Map Estimation and Dehazing Using Deep Networks.
- [11] McCartney, E.J. (1976) *Optics of the Atmosphere: Scattering by Molecules and Particles*. John Wiley and Sons, Inc., New York, 421.
- [12] Narasimhan, S.G. and Nayar, S.K. (2003) Contrast Restoration of Weather Degraded Images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **25**, 713-724. <https://doi.org/10.1109/TPAMI.2003.1201821>
- [13] Mirza, M. and Osindero, S. (2014) Conditional Generative Adversarial Nets.
- [14] Jégou, S., Drozdal, M., Vazquez, D., Romero, A. and Bengio, Y. (2017) The One Hundred Layers Tiramisu: Fully Convolutional Densenets for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, 21-26 July 2017, 11-19. <https://doi.org/10.1109/CVPRW.2017.156>
- [15] Radford, A., Metz, L. and Chintala, S. (2015) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.
- [16] 梁毓明, 张路遥, 卢明建, 杨国亮. 基于条件生成对抗网络的图像去雾算法[J]. 光子学报, 2019, 48(5): 0510002.
- [17] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [18] Silberman, P.K.N., Hoiem, D. and Fergus, R. (2012) Indoor Segmentation and Support Inference from Rgb-d Images. In: *European Conference on Computer Vision*, Springer, Berlin, Heidelberg, 746-760. https://doi.org/10.1007/978-3-642-33715-4_54
- [19] Ancuti, C.O., Ancuti, C., Timofte, R. and De Vleeschouwer, C. (2018) O-Haze: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, 18-22 June 2018, 754-762. <https://doi.org/10.1109/CVPRW.2018.00119>
- [20] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization.