

姓名： 史君宝

学号： 2152118

一、引言

在严重的雾霾条件下，大气中的悬浮颗粒如黄昏和烟雾会极大地吸收和散射光线，从而导致图像质量退化。这些退化反过来可能会影响许多计算机视觉系统的性能，如我们常见的分类和检测任务。

为了克服雾霾造成的退化，之前发表的很多文献中都已经提出了基于图像和视频的去雾算法。今天我们关注的是论文 Densely Connected Pyramid Dehazing Network。我们将利用作者提供的代码复现密集连接的金字塔去雾网络，进一步了解图像增强领域的一个小应用。

二、理论知识

2.1 去雾基本理论

在图像增强领域，对于有雾图像还原到无雾图像，有一个基本的理论模型。由于雾的存在导致的图像退化(大气散射模型)的数学形式为：

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

在公式中 I 为观测到的有雾图像， J 为真实场景的图像， A 为全局大气光，表示环境光的强度， t 为透射度， z 为像素位置。当大气光 A 均匀时，透射度 t 可以表示为 $t(z) = e^{-\beta d(z)}$ ，其中 β 表示大气的衰减系数， d 为场景深度。在我们的去雾领域中，问题就是对于给定的单幅有雾图像，即给定 I ，目标是通过计算估计 J ，还原出真实场景的图像。

改变上面的公式我们可以知道，一旦估计出透射图和大气光，就可以恢复去雾图像，求解过程如下：

$$\hat{J}(z) = \frac{I(z) - \hat{A}(z)(1 - \hat{t}(z))}{\hat{t}(z)}. \quad (2)$$

2.2 密集连接的金字塔去雾网络

传统的去雾方法已经有基于先验的方法和基于学习的方法。基于先验的方法通常利用不同的先验来表征，如暗原色先验，对比颜色线和雾线先验；而基于学习的方法中有基于卷积神经网络(cnn)的方法，试图直接从训练数据中学习。但是这些方法都有一些因素阻碍了性能。

论文中提出了一种新的图像去雾架构来解决端到端学习图像去雾问题，被称为密集连接金字塔去雾网络(DCPDN)。它通过深度学习框架提供的数学运算模块，将基本理论计算公式直接嵌入到网络中，实现端到端的学习。

本文的创新：

(1) 提出了一种新的端到端的联合优化去雾网络。将基本公式通过数学运算模块直接嵌入优化框架来实现的。

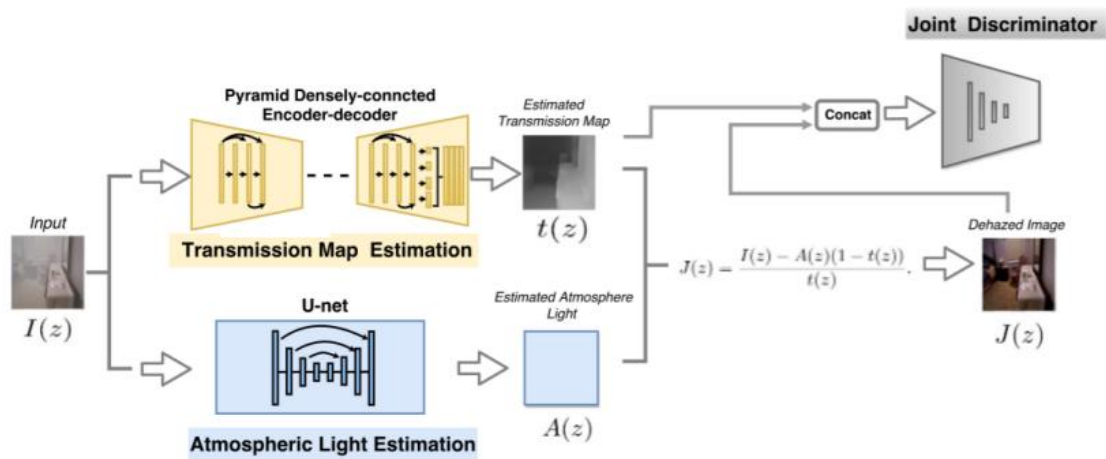
(2) 为了准确估计传输映射，提出一种边缘保持的金字塔密集连接编码器-解码器网络。进一步，通过新提出的边缘保持损失函数进行优化。

(3) 由于估计的传输映射和去雾图像的结构高度相关，利用 GAN 框架内的联合判别器来确定成对样本(即传输映射和去雾图像)是否来自数据分布。

2.3 网络架构

论文中所提到的 DCPDN 网络架构由四个模块组成：

- (1) 金字塔密集连接的透射度估计网络。
- (2) 大气光估计网络。
- (3) 利用公式的去雾模块。
- (4) 联合判别器。



本文使用了一种密集连接的编码器-解码器结构，它利用 CNN 多层的特征，采用密集块作为基本结构。使用密集块可以最大化沿着这些特征的信息流，并通过连接所有层来保证更好的收敛。此外，采用多级金字塔池化模块，通过考虑全局结构信息来优化。

为了利用 dense-net 的预定义权重，作者采用了预训练 dense-net121 的第一层 Conv 和前三层密集块及其相应的下采样操作过渡块作为我们的编码器结构。编码部分末尾的特征大小是输入大小的 1/32。为了将传输映射重建为原始分辨率，我们将五个密集块与精细上采样过渡块堆叠作为解码模块。此外，还将相同维度对应的特征进行了拼接。在密集连接编码器-解码器结构中，还采用了多级金字塔池化块，以确保不同尺度的特征嵌入到最终结果中，架构中采用四层池化操作，池化大小分别为 1/32、1/16、1/8 和 1/4。下面就是所提到的金字塔密集连接传输图估计网络的架构。

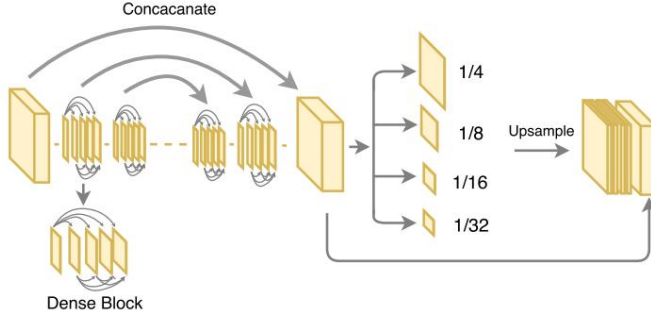


Figure 3: 金字塔密集连接传输图估计网络综述。

2.4 联合鉴别器 GAN

为了优化输出，并确保估计的传输映射 $G_t(I)$ 和去雾图像 $G_d(I)$ 分别与它们对应的真实值 t 和 J 不可区分，论文中还使用了具有新颖联合鉴别器的 GAN。由之前的公式我们可以知道，透射度和去雾图像之间的结构信息是高度相关的，使用联合判别器可以学习联合分布，用来决定传输映射和去雾图像对是真是假。

联合鉴别器的优化函数是：

$$\begin{aligned} \min_{G_t, G_d} \max_{D_{joint}} & \mathbb{E}_{I \sim p_{data(I)}} [\log(1 - D_{joint}(G_t(I)))] + \\ & \mathbb{E}_{I \sim p_{data(I)}} [\log(1 - D_{joint}(G_d(I)))] + \\ & \mathbb{E}_{t, J \sim p_{data(t, J)}} [\log D_{joint}(t, J)]. \end{aligned} \quad (3)$$

2.5 总损失函数

所提出的 DCPDN 架构使用以下四个损失函数进行训练：

$$L = L^t + L^a + L^d + \lambda_j L^j, \quad (7)$$

其中 L^t 由保边损失 L^E 组成， L^a 由预测大气光的传统 L2 损失组成， L^d 代表去雾损失，也只由 L2 损失组成。 L^j 记为联合鉴别器损失。定义如下：

$$L^j = -\log(D_{joint}(G_t(I))) - \log(D_{joint}(G_d(I))). \quad (8)$$

这里我们着重提及一下保边损失：

由于使用欧氏损失 (L2 loss) 往往会模糊最终结果，导致细节的丢失。为了有效地解决这个问题，文中提出了一种新的保边损失，其动机是以下两点观察。

- (1) 边缘对应于图像强度中的不连续性，因此可以用图像的梯度来表征。
- (2) 边缘和轮廓等低层特征可以在 CNN 结构的浅层 (前几层) 中捕获。可以实现边缘信息在最终的特征映射中得到保留。

因此，作者提出了一个新的边缘保持损失函数，由三个不同的部分组成：L2 损失、双向梯度损失和特征边缘损失。

$$L^E = \lambda_{E, l_2} L_{E, l_2} + \lambda_{E, g} L_{E, g} + \lambda_{E, f} L_{E, f}, \quad (4)$$

$$L_{E,g} = \sum_{w,h} \|(H_x(G_t(I)))_{w,h} - (H_x(t))_{w,h}\|_2 + \|(H_y(G_t(I)))_{w,h} - (H_y(t))_{w,h}\|_2, \quad (5)$$

$$L_{E,f} = \sum_{c_1,w_1,h_1} \|(V_1(G_t(I)))_{c_1,w_1,h_1} - (V_1(t))_{c_1,w_1,h_1}\|_2 + \sum_{c_2,w_2,h_2} \|(V_2(G_t(I)))_{c_2,w_2,h_2} - (V_2(t))_{c_2,w_2,h_2}\|_2, \quad (6)$$

三、复现过程

3.1 代码获取和训练集准备

在作者的 github 主页 <https://github.com/hezhangsprinter/DCPDN> 中我们可以轻松地获得源代码，并且作者已经准备好了充足的相应的训练集合训练数据供以我们使用，我们去对应的网盘中下载就可以了。

3.2 训练过程

我们按照作者的指示，准备好相应的系统环境，其中系统环境如下所示：

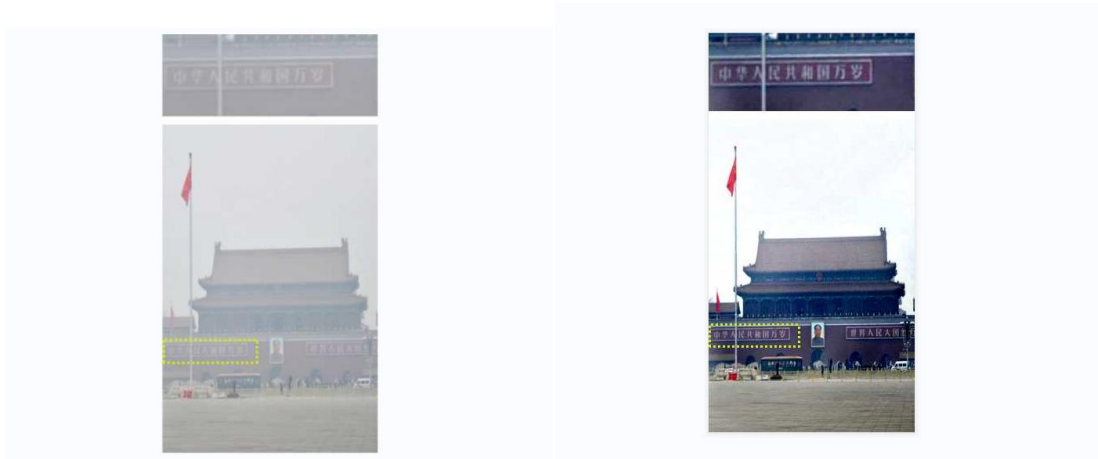
Prerequisites:

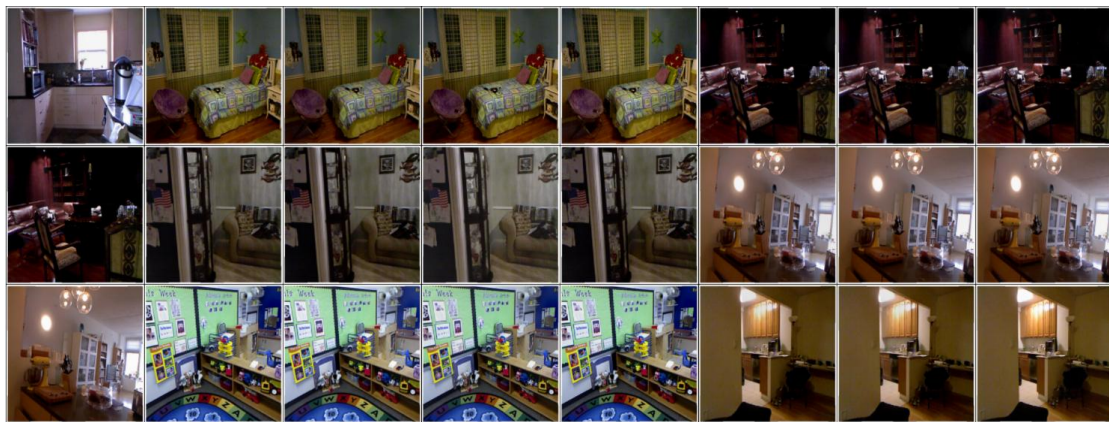
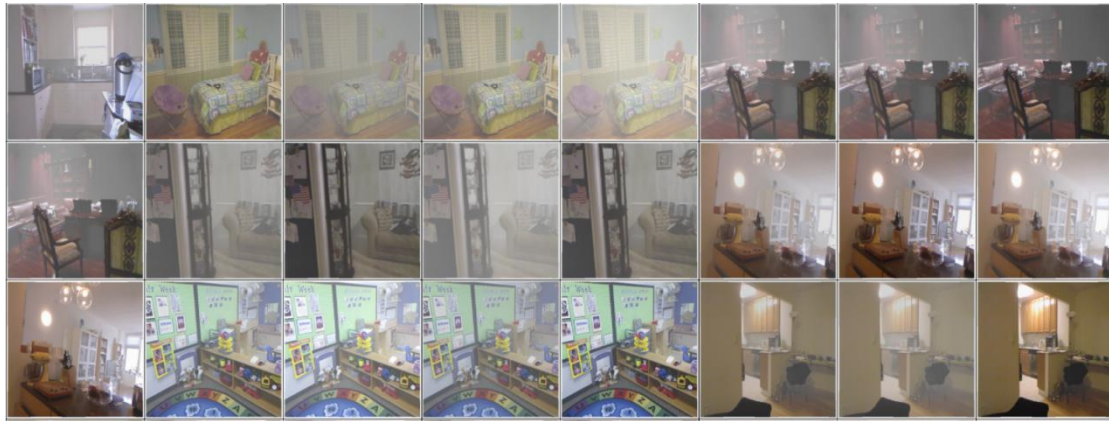
1. Linux
2. Python 2 or 3
3. CPU or NVIDIA GPU + CUDA CuDNN (CUDA 8.0)

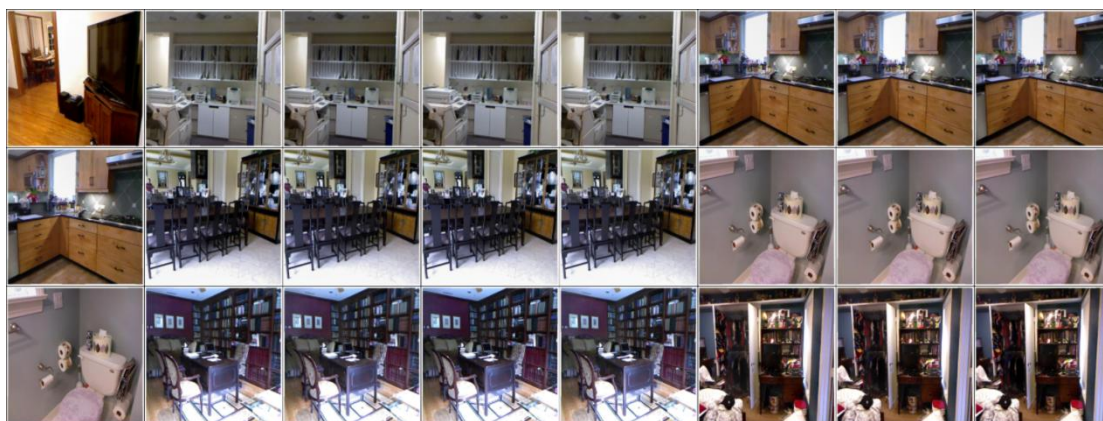
我们按照作者的指示首先运行 `demo.py` 进行预训练，然后运行 `train.py` 进行训练模型。同时我们自己生成了相应的有雾图片用来测试。

3.3 训练结果

下面是去雾图前后对比效果：







四、总结

图像增强也属于人工智能的一个重要领域。这个项目也是去雾问题中一个比较经典的论文了，对于我们的应用也比较有现实意义，因此我选择了复现这个模型。在复现 DPCDN 去雾模型的过程中，我经历了一系列的挑战，深入了解了 DPCDN 模型的工作原理，虽然主要的复现过程中主要是读论文、查资料和配环境。但是按照作者的指示下一步的实现也收获了很多。

通过这个复现项目，我不仅加深了对 DPCDN 去雾模型的理解，还提升了自己的实践能力和解决问题的能力。这个经历对我的学术和职业发展都有积极的影响，我期待将来能够在深度学习和计算机视觉领域做出更多的贡献。

五、参考资料

- [1] Zhang H , Patel V M .Densely Connected Pyramid Dehazing Network[J]. IEEE, 2018. DOI:10.1109/CVPR.2018.00337.
- [2] D. Berman, S. Avidan, et al. Non-local image dehazing. In CVPR, pages 1674 - 1682, 2016.
- [3] X. Di and V. M. Patel. Face Synthesis from Visual Attributes via Sketch using Conditional VAEs and GANs. ArXiv eprints, Dec. 2018.
- [4] R. Fattal. Dehazing using color-lines. volume 34, New York, NY, USA, 2014. ACM.
- [5] Y. Zhu and S. Newsam. Densenet for dense flow. In ICIP, 2017.
- [6] https://blog.csdn.net/qq_32734095/article/details/89059949