

深度网络图像去雾对比研究

李明宇

(大连工程学院)

DOI: 10.65285/ZRKXLT.V1I1.001

摘要: 目的: 比较不同编码—解码结构在浓雾图像恢复中的表现, 分析多尺度特征融合与自注意力机制对去雾质量的影响。方法: 基于 Dense-Haze 成对图像开展探索性实验, 构建 ResNet-50 编码器—解码器、U-Net 及嵌入多头自注意力的 U-Net 模型; 统一将图像缩放至 256×256 像素, 批量大小为 4, 学习率为 1×10^{-4} , 训练 30 轮, 以均方误差为损失函数, 并从收敛趋势、残余雾、结构细节、色彩恢复和伪影等方面进行比较。结果: 3 种模型的训练损失均呈下降趋势。ResNet 模型虽收敛较快, 但输出存在模糊、亮度不足和棋盘状伪影; U-Net 借助跳跃连接改善了结构与色彩恢复; 加入自注意力后, 所示样例中的大范围雾分布处理更均衡, 整体自然度相对较好。结论: 多尺度特征传递是浓雾图像恢复的重要基础, 全局依赖建模具有进一步改善视觉一致性的潜力; 由于原实验未保留独立测试集及峰值信噪比、结构相似性等完整记录, 相关结论仍需通过统一划分、重复训练和外部数据验证。

关键词: 图像去雾; U-Net; 残差网络; 自注意力; Dense-Haze

Comparative Study of Deep Networks for Image Dehazing

Mingyu Li

Dalian Institute of Technology

Abstract: Objective: To compare the performance of different encoder-decoder structures in dense-haze image restoration, and analyze the influences of multi-scale feature fusion and self-attention mechanism on dehazing quality. Methods: Exploratory experiments are carried out based on paired Dense-Haze images. Three models are constructed, including ResNet-50 encoder-decoder, U-Net, and U-Net embedded with multi-head self-attention. All images are resized to 256×256 pixels, with a batch size of 4, a learning rate of 1×10^{-4} and 30 training epochs. Mean squared error is adopted as the loss function. The models are compared in terms of convergence trend, residual haze, structural details, color restoration and artifacts. Results: The training loss of the three models shows a downward trend. Although the ResNet-50 model converges fast, its outputs suffer from blurriness, insufficient brightness and checkerboard artifacts. U-Net achieves better structural and color restoration by virtue of skip connections. The U-Net with self-attention realizes more balanced processing for large-range haze distribution in samples and obtains relatively good overall naturalness. Conclusion: Multi-scale feature transfer serves as an important foundation for dense-haze image restoration, and global dependency modeling has the potential to further improve visual consistency. Since independent test set and complete metrics such as PSNR and SSIM are not preserved in original experiments, the conclusions need to be verified by unified dataset splitting, repeated training and external data.

Keywords: image dehazing; U-Net; residual network; self-attention; Dense-Haze

前言:

雾霾条件下, 悬浮颗粒对光线产生吸收与散射, 使获取图像出现对比度下降、颜色偏移和边缘细节衰减。图像去雾旨在由退化观测恢复场景辐射信息, 是典型的不适定逆问题。传统方法通常依据大气散射模型估计透射率与大气光, 其中暗通道先验通过无雾图像的局部统计特征约束恢复过程^[1]。此类方法具有一定可解释性, 但在天空区域、非均匀浓雾或不满足先验的场景中容易出现颜色失真和光晕。

深度学习方法能够从成对数据中直接学习雾图到清晰图的映射。DehazeNet 将物理模型中的透射率估计转化为可

学习网络^[2], AOD-Net 则通过重参数化实现端到端恢复^[3]。

近年来, 多尺度网络与注意力机制进一步用于缓解空间信息损失和长距离依赖建模不足的问题^[8]。然而, 在小规模真实浓雾数据上, 训练损失较低并不必然对应更好的视觉恢复, 网络的下采样方式、特征传递路径和输出重建策略均会显著影响伪影与细节。

残差网络 (residual network, ResNet) 通过恒等映射改善深层网络优化^[4]; U-Net 依靠对称编码器—解码器及跳跃连接传递多尺度细节^[5]; 自注意力机制则可在全局范围建立特征位置之间的关联^[6]。三者代表了特征提取、局部细节传递和全局上下文建模的不同侧重。基于此, 本文在统一

的基础训练条件下比较 ResNet 编码器—解码器、U-Net 和 U-Net+ 自注意力 3 种方案, 分析损失变化与视觉结果不一致的原因, 并提出面向规范验证的改进路径。本文不将现有结构组合表述为新算法, 研究重点在于可复核的结构对照与局限性分析。

一、材料与方法

(一) 数据集与任务定义

实验采用 Dense-Haze 数据集。该数据集面向真实浓雾场景, 提供浓雾图像及同场景无雾参考图像, 可用于监督式单幅图像去雾研究^[7]。官方挑战训练数据包含 55 对真实拍摄图像^[9]。本文沿用原实验留存的成对输入与监督训练流程, 将雾图作为网络输入, 将对应无雾图作为回归目标。

(二) 实验环境与统一设置

实验在 Google Colaboratory 的 NVIDIA T4 GPU 环境下完成, 软件环境为 Python 与 PyTorch。为保证实验结果的可比性, 3 种模型采用统一的实验参数: 以 Dense-Haze 成对图像为数据集, 以浓雾图像为输入、对应无雾图像为监督目标, 图像统一调整为 256×256 像素的 RGB 三通道, 批量大小设为 4, 训练 30 轮, 初始学习率为 1×10^{-4} , 并采用均方误差 (MSE) 作为损失函数。

(三) 网络结构

1. ResNet 编码器—解码器

第一种模型采用 ImageNet 预训练 ResNet-50 作为编码器, 去除全局平均池化层和全连接层, 提取约 8×8 、2048 通道的深层特征。解码器由转置卷积、批归一化和 ReLU 激活组成, 逐步恢复至 256×256 像素, 并通过 3 通道卷积和 Sigmoid 输出去雾图像。该结构具有较强的特征提取能力, 但连续下采样会造成空间细节损失, 且缺少浅层特征的直接传递。

2. U-Net

第二种模型采用 U 形编码器—解码器。编码端通道数依次为 64、128、256 和 512, 利用最大池化逐级降低空间分辨率; 瓶颈层通道数为 1024; 解码端通过转置卷积上采样, 并与编码端同尺度特征拼接。跳跃连接将边缘、纹理等浅层信息直接送入解码阶段, 可减少仅依赖压缩特征重建导致的细节丢失^[5]。输出层使用 1×1 卷积映射至 3 通道并经 Sigmoid 函数归一化。

3. U-Net+ 自注意力

第三种模型以 U-Net 为骨干, 在瓶颈层加入多头自注意力模块。瓶颈特征首先展平为空间序列, 经层归一化后输入 8 头注意力层, 注意力输出与原序列残差相加, 再恢复为二维特征图。多头机制可从不同子空间学习位置间关系^[6], 用于补充卷积在长距离依赖建模方面的不足。注意力层的 dropout 设为 0.1, 其余编码、解码及训练参数与 U-Net 保持一致。

(四) 观察指标与分析方法

本文从训练过程和图像表现两个层面比较。训练过程主要观察 MSE 曲线的总体下降趋势、波动幅度与异常尖峰;

图像表现采用同一场景并列观察, 重点考察残余雾、亮度与对比度、色彩偏移、目标轮廓、纹理保留以及棋盘状伪影。

二、结果

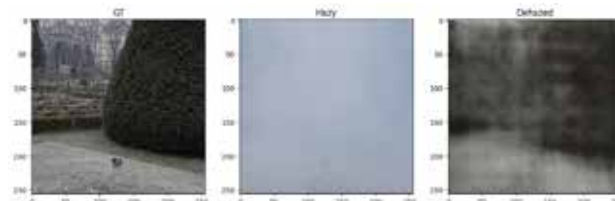
(一) 训练损失变化

3 种模型的 MSE 曲线均随迭代总体下降, 表明网络能够从成对样本中学习基本的去雾映射。ResNet 编码器—解码器曲线下降相对平稳, 后期损失集中于较低区间; U-Net 与 U-Net+ 自注意力曲线波动更明显, 并出现少量尖峰, 但总体趋势仍持续下降。原实验日志分别保留了 ResNet 在第 29 轮第 400 次迭代时 0.012604、U-Net 在相同标记时 0.060630, 以及注意力模型在第 8 轮第 100 次迭代时 0.040030 的记录。

(二) 去雾结果的定性比较

在同一浓雾场景中, ResNet 模型虽能提高部分区域的可见性, 但整体偏暗, 轮廓和纹理较为模糊, 并存在块状纹理。U-Net 借助跳跃连接改善了座椅、道路边界和植被等结构的恢复, 但仍存在残余雾和纹理噪声。U-Net+ 自注意力对大范围雾层处理更均衡, 场景层次和整体色彩较为自然, 但植被区域仍存在过度增强和纹理不稳定现象。

(a) ResNet 编码器—解码器



(b) U-Net



(c) U-Net+ 自注意力



图 4 同一场景的去雾结果对比 (Figure 4. Comparison of dehazing results on the same scene)

(三) 综合比较

综合训练曲线与留存图像可见, 较低的像素损失并不意味着更优的视觉恢复。ResNet 模型虽易于优化, 但低分辨率瓶颈限制了细节重建, 并存在边缘模糊、整体偏暗和棋盘伪影; U-Net 通过多尺度特征融合改善了轮廓、纹理及色彩表现, 但仍存在残余雾和局部噪声; 引入自注意力后, U-Net 的场景层次和色彩表现更为均衡, 但局部纹理仍存在

不稳定现象。

三、讨论

(一) 多尺度特征传递对细节恢复的作用

ResNet-50 原本面向图像分类,其深层特征强调语义不变性而非像素级精确定位。本文模型将输入压缩至约 8×8 后再依赖转置卷积恢复图像,浅层边缘与纹理无法直接到达解码端,因而即使训练 MSE 较低,也可能通过输出平滑的平均化结果降低像素误差。多级转置卷积还可能在核大小、步幅与重叠模式不匹配时形成周期性响应,表现为棋盘状伪影。

U-Net 的同尺度跳跃连接缩短了浅层信息到输出端的路径,使解码器能够同时利用局部纹理与高层语义。所示结果中,座椅轮廓、道路边界和植被结构的改善与这一机制一致。由此可见,对于去雾这类图像到图像回归任务,强编码器并不能替代多尺度空间信息传递;网络设计应在感受野、语义抽象与细节保真之间取得平衡。

(二) 自注意力的潜在贡献与代价

卷积网络的感受野虽可随深度增加而扩大,但对远距离空间关系的建模能力仍有限。瓶颈层引入自注意力机制可建立不同空间位置间的全局关联,有助于改善非均匀及大范围浓雾的特征建模。实验结果表明,注意力模型对整体雾层和场景层次的处理较为均衡,但现有样例尚不足以证明其具有稳定增益。此外,自注意力计算开销较大,将其置于低分辨率瓶颈层可降低计算成本,但可能限制细粒度特征表达。

(三) 损失函数与感知质量的不一致

MSE 对较大的像素偏差惩罚较强,易促使网络输出条件均值,从而产生过度平滑。浓雾区域本身缺少可辨细节,模型仅依赖 MSE 时更可能优先降低整体颜色误差,而忽略边缘锐度、局部对比度和感知自然度。这解释了 ResNet 模型训练损失较低但视觉结果不理想的现象。

(四) 研究局限与验证方案

本研究的主要局限包括:第一,原实验未保留实际使用样本数量、训练集—验证集—测试集划分及随机种子,无法排除训练样例重用造成的乐观判断;第二,3 种模型的初始化条件并不完全对称,ResNet 使用 ImageNet 预训练权重,而 U-Net 及其注意力变体从头训练;第三,仅保留部分训练日志和定性图像,缺少逐图 PSNR、SSIM、参数量、推理时间及显存消耗;第四,训练仅运行 30 轮,且未采用早停、学习率调度与多次重复实验,曲线波动可能来自小批量训练或数据顺序;第五,图像统一拉伸至正方形可能引入几何畸变。

结语:

本文在 Dense-Haze 成对图像上对 ResNet 编码器—解码器、U-Net 和 U-Net+ 自注意力 3 种结构进行了探索性比较。结果表明,3 种模型均能学习基本去雾映射,但网络结构对输出质量的影响不能由训练 MSE 单独解释。ResNet 模型虽

收敛较快,却因空间压缩和缺少浅层特征传递而出现模糊、亮度不足及棋盘状伪影;U-Net 通过跳跃连接改善了轮廓、纹理和色彩;在所示样例中,瓶颈层自注意力进一步增强了大范围雾分布处理的整体一致性。

由此认为,多尺度特征融合是浓雾图像恢复的基础,全局上下文建模具有补充价值。受独立测试记录与量化指标缺失限制,当前结论应视为方法筛选依据,而非性能排名。后续需通过规范数据划分、多次重复、客观指标、消融实验和外部验证,确定注意力模块的真实增益及其计算代价。

参考文献:

- [1]He, K., Sun, J. and Tang, X. (2011) Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 2341–2353.
- [2]Cai, B., Xu, X., Jia, K., Qing, C. and Tao, D. (2016) DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing*, 25, 5187–5198.
- [3]Li, B., Peng, X., Wang, Z., Xu, J. and Feng, D. (2017) AOD-Net: All-In-One Dehazing Network. *Proceedings of the IEEE International Conference on Computer Vision*, 4770–4778.
- [4]He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 770–778.
- [5]Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*, Munich, 234–241.
- [6]Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- [7]Ancuti, C.O., Ancuti, C., Sbert, M. and Timofte, R. (2019) Dense Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images. *Proceedings of the IEEE International Conference on Image Processing*, 1014–1018.
- [8]Liu, X., Ma, Y., Shi, Z. and Chen, J. (2019) GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 7314–7323.
- [9]Ancuti, C.O., Ancuti, C., Timofte, R., et al. (2019) NTIRE 2019 Image Dehazing Challenge Report. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, 2241–2253.

作者简介: 李明宇 (2003.04–)、男、汉族、河北省衡水市、学士学位、研究方向: 计算机视觉与图像处理。