

基于边缘信息的轻量化图像去雾卷积神经网络

常可铮¹ 李耀成¹ 杨得武¹

CHANG Kezheng LI Yaocheng YANG Dewu

摘要

由于在采用 CCD/CMOS 传感器进行图像采集时,可能会存在严重的信息退化,单图像去雾处理是一个具有挑战性的不适定问题。然而,现有的基于深度学习的单图像去雾方法只采用清晰和有雾的图像对来指导去雾网络的训练,忽视了图像内存在的非均匀的细节信息,从而导致去雾网络的处理过程是全局的,导致某些原有信息丢失,同时对于非均匀雾气也不能很好地去除。还有部分算法采用先验信息对图像去雾进行引导,如 DCP(暗通道先验)、颜色先验、物理模型先验等。这些方法可以为图像去雾提供一定的先验信息,从而增强图像去雾效果。此外,还有部分方法针对去雾任务来设计网络模型和函数,这些算法通常在去雾指标峰值信噪比、结构相似度上能够取得较好的成绩,然而恢复无雾图像的主观评价往往不尽如人意,真实的雾气图像去雾中并不能去除非均匀的雾气,同时导致图像存在一定的颜色失真。对此,提出一种基于边缘对应峰值信噪比的先验方法,通过检测图像的边缘信息,评估相应的峰值信噪比,从而估测图像中雾气存在的主要区域,以边缘信息作为先验来引导图像去雾网络。此外,还设计了一种基于边缘先验引导模块的可变形卷积网络,利用图像的雾气特征进行引导,从而增强去雾网络恢复图像能力。所提出的算法在合成和真实数据集的广泛实验上均展现了较好的结果。

关键词

图像去雾;边缘先验;可变形卷积神经网络;真实雾气

doi: 10.3969/j.issn.1672-9528.2024.09.015

0 引言

在现代的社会和工作中,尤其是需要在复杂环境中工作的时候,视野往往会被雾气遮挡。这一问题会导致边缘端智能设备应用时,如带有目标检测功能的行车记录仪,或者智能监控摄像头就有可能受雾气影响,从而导致许多在户外环境中部署的基于视觉的算法都具有较差的效果,如图像任务中的目标检测任务,进而使得智能装备工作时存在安全隐患。因此,提高模糊图像的可见性,对提高各种重要的计算机视觉任务的性能具有重要意义。而现阶段的去雾算法中有些网络的去雾效果并不能很好满足应用需求,同时由于现阶段有许多去雾网络为了达到更好的性能而引入复杂的结构,这种措施对提高模型性能可能有一定的帮助,但是会导致模型推理时计算量太大,整个网络对算力的要求更难以满足,从而无法应用到边缘设备上。本文提出了一种以物理模型引导图像去雾网络的 EGC 卷积,不仅利用雾气的物理特征来提高去雾能力,还使得整体网络可以设计得更加轻量化,进一步降低算力需求,因此本文的网络设计有着实际应用的价值和意义。

雾霾的存在在很大程度上降低了室外场景的能见度和对比度,这会导致下游视觉任务的表现变得并不理想,如自动驾驶任务和监控中的目标检测和图像分割任务。通过大气散射模型,可以将雾霾形成用公式进行描述:

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

早期的去雾方法多是基于传统算法来实现,例如基于直方图的去雾方法、暗通道先验(DCP)、Retinex 图像增强算法等,但是效果并不理想,因为仅仅从雾气图像中很难提取出有用的特征信息。有人提出了同一场景提取多种天气图像的多图像去雾方法。此外,还有提取图像的深度信息,基于给定的深度信息进行去雾的算法。近年来,已经有很多基于卷积神经网络的端到端去雾方法被提出,通过去雾网络模型直接学习模糊到清晰的图像转换来简化去雾问题,利用数据引导的方式可以大大提高图像去雾的效果。大多数现有的方法通常采用清晰图像与有雾图像成对作为训练对象,通过基于 L_1 或 L_2 损失函数来指导更新去雾网络,同时不需要进行任何正则化处理。但是只有这两种损失函数仍无法有效地处理图像的细节特征,往往会导致恢复图像的颜色失真和部分细节丢失。之前的论文里有的方法是专注于通过显著增加去雾模型的深度或宽度来提高去雾性能,这种方法不考虑内存

1. 西北机电工程研究所 陕西咸阳 712000

占用和计算开销,导致它们不能在资源有约束的场景下部署,例如嵌入式设备或移动端,如三叉戟网络(TDN)^[1]。

为了实现网络去雾性能和资源占用量之间的最佳权衡,本文提出了一种轻量化去雾网络。考虑到去雾任务需要在移动端部署的需求,设计了一种基于深度可分离卷积构成的边缘先验引导卷积模块(EGC),利用少量的计算先得出图像中边缘特征对应的峰值信噪比值,从而利用这一参数引导可变形卷积神经网络,并利用残差结构保留原有图像信息。这一去雾模型可以有效从非均匀雾气提取边缘特征图,并且可变形卷积能够对图像边缘细节进行更精细的复原处理,同时针对不同雾气浓度区域的计算参数量也不同,可以使得计算参数和模型去雾性能达到平衡。

1 相关工作

单幅图像去雾任务中有两种常用的方法:一是通过先验信息引导的方法,二是利用数据驱动网络学习的方法。基于先验的方法通常依赖于物理大气散射模型手工设计先验模型去雾的方法,如对比度最大化方法、暗通道先验(DCP)^[2]、颜色衰减先验和非局部先验算法^[3]等。虽然这些基于先验的方法在人工合成数据集上取得了较为良好的结果,但是这些算法依赖于人为假设和相对特定的目标场景而设计,导致了设计的网络具有较低的鲁棒性,并且在真实雾气图像的去雾处理上往往不能达到预期的效果^[4]。

利用数据驱动网络学习的方法通常使用深度神经网络来估计大气散射模型中的参数(投射系数和大气光照强度),以实现单张图像去雾。例如AODnet^[5],这种网络估计了一种与透射系数和大气光强度相关的新的物理模型参数。然而,这种方法可能会因为误差累积而导致伪影,从而导致出现不准确的估计偏差,使得恢复图像和清晰图像之间存在较大的重建误差。此外,能够真实反映现实环境中的大气光照强度和投射系数是非常复杂和困难的。因此,这些年来,有许多论文抛弃了大气散射模型,提出了端到端的图像去雾算法。例如,Qin等人提出的一种特征融合注意机制网络^[6],通过处理不同类型的信息来增强特征提取网络的灵活性,该网络只使用考虑有雾气图像与无雾图像之间的 L_1 损失误差。此外,还有部分算法引入了encoder和decoder框架^[7],或者提出正负样本进行对比训练的方法^[8],这些方法通常需要较大的训练样本和较多的计算参数量,从而导致网络模型不能实现轻量化。

2 本文提出的方法

2.1 边缘先验

提取边缘特征信息这一方法通常会在图像超分领域使用,通过构建一个边缘查找表,将图像块的边缘信息映射到每个子网络的PSNR性能及其计算消耗。在推理阶段,每个

图像块将被送入不同子网络,以寻求更佳的计算复杂度和性能的均衡,根据不同图像块的性能会随超分模型的大小而变化。受到一个大型网络对每组数据进行PSNR推断这一启发,本文构建了边缘信息到PSNR值的对应转换关系,通过计算得到边缘信息所对应的PSNR值,进而引导网络模型。具体来说,则是先针对每一个图像块,通过索贝尔边缘检测得到边缘信息值的均值和最大最小值,其中 A 表示图像的灰度图,而灰度图与对应的 X 方向和 Y 方向的梯度矩阵进行卷积,就可得出对应方向上的边缘检测图像灰度值:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A \quad (2)$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (3)$$

通过逐点提取图像中每一个像素的环绕相邻像素的特征数据,根据两个矢量方向分别进行卷积计算,再通过均方差来计算该点确切的边缘检测得分,如下所示:

$$G = \sqrt{G_x^2 + G_y^2} \quad (4)$$

$[\min(E), \max(E)]$ 整个区间会根据得分水平来划分,通过离散化处理将整个区间平均分成 K 个子区间,区间范围如下:

$$s_k = [\min(E) + \frac{\max(E) - \min(E)}{K} \cdot (k-1), \min(E) + \frac{\max(E) + \min(E)}{K} \cdot k] \quad (5)$$

再通过一个PSNR计算子网,分别计算第 k 个子区间内的平均PSNR值,并分别将这 k 个区间的PSNR值和边缘检测得分进行对应,构建出边缘检测值和峰值信噪比的对照关系。

2.2 边缘先验引导卷积(EGC)

由于目标的尺度、姿态、视点和结构变形引起的种种复杂因素的存在,在机器视觉图像任务中通常采用可变形卷积来解决这类难题。本文针对雾气的非均匀性同样采用这一结构设计了去雾网络。可变形卷积^[9]不同于传统卷积的地方在于其通过引导子网络模块可以改变网络输入的采样方式,通过学习不同模块内的特征信息可以得到不同的卷积输入数据,这一方法在还原物体边缘信息上具有先天优势,因此本文基于边缘引导模块构建了一个基于PSNR值的可变形卷积模块。卷积输入先给到可变形offset卷积中,通过可变形卷积获得偏移引导量,再根据该偏移引导来改变输入特征的横坐标和纵坐标,从而得到一个非矩形的不规则采样框,之后将采样得到的参数进行卷积,得到输出。图1为基于可变形卷积的边缘先验引导卷积模块(EGC)。

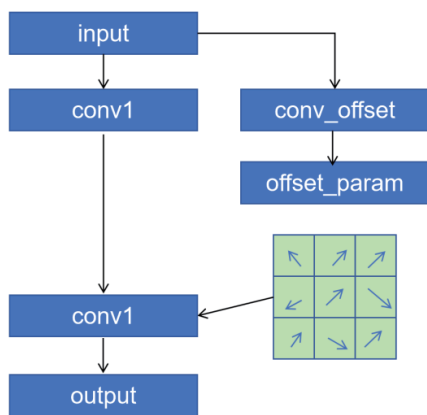


图1 边缘先验引导卷积模块

$$y(P_0) = \sum W(P_n) \cdot X(P_0 + P_n + \Delta P_n) \quad (6)$$

可变形卷积训练的时候,用于生成输出特征的卷积核和用于生成偏移量的卷积核是成对进行更新的,因此每一次反向传播都是对相同组的卷积参数进行反向传播学习。同时,同一层卷积采用的可变形偏移量是通用的,这也能够在一定程度上保留通道之间的相关性特征。

本文设计的去雾网络结构图如图2所示。

Conv2d-1	[-1, 6, 416, 416]	24
Conv2d-2	[-1, 6, 416, 416]	60
BatchNorm2d-3	[-1, 6, 416, 416]	12
ReLU6-4	[-1, 6, 416, 416]	0
ConvBNReLU-5	[-1, 6, 416, 416]	0
Conv2d-6	[-1, 12, 416, 416]	84
BatchNorm2d-7	[-1, 12, 416, 416]	24
ConvBNReLU-8	[-1, 12, 416, 416]	0
Residual_only-9	[-1, 12, 416, 416]	0
Conv2d-10	[-1, 12, 416, 416]	120
BatchNorm2d-11	[-1, 12, 416, 416]	24
ReLU6-12	[-1, 12, 416, 416]	0
ConvBNReLU-13	[-1, 12, 416, 416]	0
Conv2d-14	[-1, 12, 416, 416]	156
BatchNorm2d-15	[-1, 12, 416, 416]	24
ConvBNReLU-16	[-1, 12, 416, 416]	0
Residual_only-17	[-1, 12, 416, 416]	0
Conv2d-18	[-1, 12, 416, 416]	120
BatchNorm2d-19	[-1, 12, 416, 416]	24
ReLU6-20	[-1, 12, 416, 416]	0
ConvBNReLU-21	[-1, 12, 416, 416]	0
Conv2d-22	[-1, 12, 416, 416]	156
BatchNorm2d-23	[-1, 12, 416, 416]	24
ConvBNReLU-24	[-1, 12, 416, 416]	0
Residual_only-25	[-1, 12, 416, 416]	0
Conv2d-26	[-1, 12, 416, 416]	120
BatchNorm2d-27	[-1, 12, 416, 416]	24
ReLU6-28	[-1, 12, 416, 416]	0
ConvBNReLU-29	[-1, 12, 416, 416]	0
Conv2d-30	[-1, 12, 416, 416]	156
BatchNorm2d-31	[-1, 12, 416, 416]	24
ConvBNReLU-32	[-1, 12, 416, 416]	0
Residual_only-33	[-1, 12, 416, 416]	0
Conv2d-34	[-1, 12, 416, 416]	120
BatchNorm2d-35	[-1, 12, 416, 416]	24
ReLU6-36	[-1, 12, 416, 416]	0
ConvBNReLU-37	[-1, 12, 416, 416]	0
Conv2d-38	[-1, 12, 416, 416]	156
BatchNorm2d-39	[-1, 12, 416, 416]	24
ConvBNReLU-40	[-1, 12, 416, 416]	0
Residual_only-41	[-1, 12, 416, 416]	0
Conv2d-42	[-1, 24, 416, 416]	240
BatchNorm2d-43	[-1, 24, 416, 416]	48
ReLU6-44	[-1, 24, 416, 416]	0
ConvBNReLU-45	[-1, 24, 416, 416]	0
Conv2d-46	[-1, 6, 416, 416]	150
BatchNorm2d-47	[-1, 6, 416, 416]	12
ConvBNReLU-48	[-1, 6, 416, 416]	0
Residual_only-49	[-1, 6, 416, 416]	0
Conv2d-50	[-1, 3, 416, 416]	21

Total params: 1,971
Trainable params: 1,971
Non-trainable params: 0

图2 去雾网络结构

图3为PyTorch自带的网络模型生成函数,由图3可知本文设计的去雾网络采用了6层隐含层残差结构,其中隐含层中包括边缘先验引导卷积计算,用来引导去雾处理,并且在每层网络之间引入了残差结构,网络的输入层为逐点卷积,卷积核大小为1,从而扩展了通道数。输出层同样使用卷积核大小为1的逐点卷积,用作图像复原功能,并与原图叠加获得最终去雾结果。

Input size (MB): 1.98
Forward/backward pass size (MB): 772.38
Params size (MB): 0.01
Estimated Total Size (MB): 774.37
Total GFLOPS: 728.912M
Total params: 1.971K

图3 模型参数量

模型参数量仅为1971,并且计算量仅为0.7 GFLOPS。该模型大小远远小于轻量化网络mobilenetV2的2.496 GFLOPS,可以部署到算力匮乏的边缘端设备上。

2.3 损失函数

损失函数是机器学习模型中用于量化大部分神经网络模型预测值与实际观测值之间差异程度的函数,也是用于参数更新的训练期望指标之一。选择合适的损失函数对于确保网络模型训练后能够产生精确且符合预期的结果至关重要。不同的任务和问题具有不同的数据分布和特性,因此需要针对性地选择合适的损失函数。在针对去雾图像处理任务设计的损失函数中, L_2 损失函数(也称为均方误差损失函数)是一个常见的组成部分,它有助于模型更好地捕捉并减少预测图像与实际清晰图像之间的像素级差异,从而提升去雾效果。

$$L_2(\hat{y}, y) = \sum_{i=0}^m (y^i - \hat{y}^i)^2 \quad (7)$$

针对图像去雾算法的特点^[10],参考大多数的图像去雾论文^[11-13]中遇到的颜色特征失真问题,损失函数对于不同的图像处理任务都有不同的针对特性和鲁棒性,本设计引入了余弦损失函数、 L_2 损失函数和物理模型损失函数。

L_{phy} 损失函数是基于物理模型对 t 和 A 进行更新,该损失函数由一组有雾图像和对应的透射系数 t 和一组对应的大气光照强度组成,根据训练的每一轮结果计算均方误差。

$$L_{phy}(i) = \frac{1}{N} \sum_{j=1}^N (|T(x_j) - t_j|^2 + |A(x_j) - t_j|^2) \quad (8)$$

根据不同损失函数对模型更新参数的差异和多次训练结果进行特征加权的调整,得出一个加权损失函数:

$$Loss(i) = 0.55L_2(i) + 0.25L_{phy}(i) + 0.2L_{cos}(i) \quad (9)$$

3 实验

数据集:采用RESIDE的ITS和OTS作为训练集,这是一个新的大规模数据集,用于训练和测试单图像去雾算法。RESIDE在数据源(合成与真实世界)、内容(室内与室外场景)和评估选项方面都是全面和多样化的。ITS包含11万张合成模糊图像,使用现有室内深度数据集的图像,OTS包含了313 950张由收集到的真实世界户外场景合成的图像,然而数据并没有深度信息。合成数据集是通过大气散射模型来估计每幅图像的深度图,用它来合成户外模糊图像。

在训练中,为了验证边缘检测先验去雾架构的性能,利用PSNR和SSIM来定量评价去雾性能。对于训练数据,从RESIDE的图像中随机抽取10 000个无雾模块。对于每个模块,均匀采样10个随机传输 $t \in (0,1)$,生成10个雾气斑块。因此,总共生成了10万个合成雾气块用于训练。每一层卷积

权值通过从高斯分布（均值 $\mu=0$ ，标准差 $\sigma=0.001$ ）中随机抽取来初始化，并将偏差设置为 0。学习率每 100 000 次迭代从 $5e-3$ 下降到 $1.125e-4$ ，并在 Nvidia RTX2060S 上训练 20 轮，图 4 和图 5 为室内和室外的去雾图像。

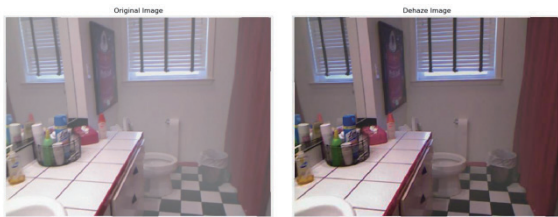


图 4 室内去雾效果



图 5 室外去雾效果

4 结论

本文提出了一种新的单幅图像去模糊的深度学习方法，采用的是轻量化可变形卷积网络模型。受传统的雾气相关特征和去雾方法的启发以及现有轻量化卷积网络算法的启发，基于投射系数估计设计了一个具有特殊去雾功能的可训练的端到端系统，其中特征提取层和物理模型参数估计层与传统的卷积神经网络有区别。本文在现有的去雾卷积网络的研究基础上，通过分析雾气图像的形成原因，以及图像在雾气中的成像原理，提出了基于物理边缘信息引导的轻量化去雾可变形卷积神经网络，同时针对去雾颜色失真的问题提出了一种新的去雾损失函数，在合成和真实数据集上训练网络，在室内和室外图像上均展现了较好的结果。同时，较少的模型参数量使得该轻量化去雾卷积模型能够在拥有不多计算资源的边缘设备上机进行部署。

参考文献：

- [1] LIU J, WU H, XIE Y, et al. Trident Dehazing Network[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 14-19.
- [2] HE K, SUN J, TANG X. Single image haze removal using dark channel prior[J]. IEEE transactions on pattern analysis & machine intelligence, 2010, 33: 2341-2353.
- [3] ZHU Q, MAI J, SHAO L. A fast single image haze removal algorithm using color attenuation prior[J]. IEEE transactions on image processing, 2015, 24: 3522-3533.
- [4] JOHNSON D J, RAHMAN Z, WOODSELL G A. A multiscale retinex for bridging the gap between color images and the

human observation of scenes[J]. IEEE transactions on image processing, 1997, 6: 965-976.

- [5] LI B, PENG X, WANG Z, et al. AOD-Net: all-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 2380-7504.
- [6] WANG W, YUAN X, WU X, et al. An efficient method for image dehazing[C]//2016 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2016: 2381-8549.
- [7] PARTHASARATHY S, SANKARAN P. An automated multi scale retinex with color restoration for image enhancement[C]//2012 National Conference on Communications (NCC). Piscataway: IEEE, 2012: 16157444.
- [8] WU H, QU Y, LIN S, et al. Contrastive learning for compact single image dehazing[EB/OL]. (2021-04-19)[2024-04-13]. <https://arxiv.org/abs/2104.09367>.
- [9] QIN L, XIE Y, LIU X, et al. An end-to-end 12-lead electrocardiogram diagnosis system based on deformable convolutional neural network with good antinoise ability[J]. IEEE transactions on instrumentation and measurement, 2021, 70: 2508313.
- [10] HORÉ A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition. Piscataway: IEEE, 2010: 9506273.
- [11] KIM G, HA S, KWON J. Adaptive patch based convolutional neural network for robust dehazing[C]//2018 25th IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2018: 2845-2849.
- [12] CAI B, XU X, JIA K, et al. DehazeNet: an end-to-end system for single image haze removal[J]. IEEE transactions on image processing, 2016, 25: 5187-5198.
- [13] WAN L, EIGEN D, FERGUS R. End-to-end integration of a convolutional network, deformable parts model and non-maximum suppression[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2018: 5857689.

【作者简介】

常可铮（1997—），男，河南唐河人，硕士研究生，助理研究员，研究方向：深度学习图像处理、模型量化及 FPGA 硬件加速。

李耀成（1997—），男，陕西韩城人，硕士研究生，助理研究员，研究方向：FPGA 总线接口设计、FPGA 图像采集和处理。

杨得武（1997—），男，甘肃白银人，硕士研究生，助理研究员，研究方向：硬件电路设计。

（收稿日期：2024-06-19）