

基于深度学习的单幅图像去雾研究进展

贾童瑶¹, 卓 力^{1,2}, 李嘉锋^{1,2}, 张 菁^{1,2}

(1. 北京工业大学信息学部, 北京 100124; 2. 北京工业大学计算智能与智能系统北京市重点实验室, 北京 100124)

摘要: 户外视觉系统极易受到雾霾等恶劣天气影响, 采集到的图像/视频质量严重下降, 这不仅影响人眼的主观感受, 也给后续的智能分析带来严峻挑战. 近年来, 学者们将深度学习应用于图像去雾领域, 取得了诸多的研究成果. 但是雾霾图像场景复杂多变、降质因素众多, 这对去雾算法的泛化能力提出了很高的要求. 本文主要总结了近年来基于深度学习的单幅图像去雾技术研究进展. 从先验知识和物理模型、映射关系建模、数据样本、知识迁移学习等角度出发, 介绍了现有算法的研究思路、具体特点、优势与不足. 尤其侧重于近两年来新出现的训练策略和网络结构, 如元学习、小样本学习、域自适应、Transformer等. 另外, 本文在公共数据集上对比了各种代表性去雾算法的主客观性能、模型复杂度等, 尤其是分析了去雾后的图像对于后续目标检测任务的影响, 更全面地评价了现有算法性能的优劣, 并探讨了未来可能的研究方向.

关键词: 单幅图像去雾; 深度学习; 无监督学习; 域泛化

基金项目: 国家自然科学基金(No.61871006, No.61971016); 北京市教育委员会科学研究计划项目资助(No.KZ202210005007, No.KZ201910005007); 北京市自然科学基金-丰台轨道交通前沿研究联合基金(No.L211017); 北京市教育委员会科技计划一般项目(No.KM202110005027)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2023)01-0231-15

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220838

Research Advances on Deep Learning Based Single Image Dehazing

JIA Tong-yao¹, ZHUO Li^{1,2}, LI Jia-feng^{1,2}, ZHANG Jing^{1,2}

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China)

Abstract: Vision-based outdoor systems are highly susceptible to severe weather such as haze. The quality of the collected images in the hazy environments is seriously degraded, which affects subjective perception and brings challenges to the subsequent intelligent processing tasks. In recent years, deep learning has been applied to single image dehazing and achieved promising results. However, the hazy scenes are complex and unpredictable, which puts a high demand on the generalization ability of the dehazing methods. In this paper, we summarize the recent deep-learning-based single-image dehazing methods. The advantages and disadvantages of these methods are analyzed in terms of network mapping relationships, learning methods, training datasets, and knowledge transfer. In particular, we focus on new training strategies and network structures that have emerged in the last few years, such as meta-learning, few-shot learning, domain adaption, and Transformer. In addition, the subjective and objective performances of various representative dehazing methods are compared on several public datasets. Further, the impact of the dehazed images on the performance of subsequent object detection tasks is analyzed and evaluated comprehensively. We also provide the computational complexity and running time of these methods. Finally, the conclusions and future tendency of single-image dehazing are drawn.

Key words: single-image dehazing; deep learning; unsupervised learning; domain generalization

Foundation Item(s): National Natural Science Foundation of China (No.61871006, No.61971016); R&D Program of Beijing Municipal Education Commission (No.KZ202210005007, No.KZ201910005007); Beijing Natural Science Foundation (No.L211017); General Program of Beijing Municipal Education Commission (No.KM202110005027)

1 引言

自动驾驶、视频监控、军事侦察等户外视觉系统采

集到的图像/视频极易受到恶劣天气的影响, 不仅主观感受差, 而且会对后续的目标检测、跟踪、分类与识别

等智能化分析处理任务造成严重影响. 雾霾是一种常见的图像降质因素, 去雾技术通过对雾霾进行去除, 可以有效提升图像的主观感受. 从客观角度来看, 增强后的图像有助于提升后续智能化分析处理任务的性能. 因此, 图像去雾成为近年来工业界和学术界的研究热点^[1].

2012年以来, 以卷积神经网络(Convolutional Neural Network, CNN)为代表的深度学习在机器视觉、自然语言处理和语音处理等领域取得了突破性的进展, 并逐步被应用在图像去雾领域. 大量的研究表明, 与传统的图像去雾算法相比, 基于深度学习的图像去雾算法以大数据为依托, 充分利用深度学习强大的特征学习和上下文信息提取能力, 从中自动学习蕴含的丰富知识, 可以从大数据中自动学习到雾霾图像-清晰图像之间的复杂映射关系, 获得了性能上的大幅提升. 基于深度学习的图像去雾也因此成为图像去雾领域的主流研究方向, 取得了重要的进展.

近年来, 出现了不少图像去雾研究综述^[2-4]. 与现有的综述相比, 本文更侧重于近两年来出现的新思路、训练策略和网络结构, 如元学习、小样本学习、域自适应、Transformer等. 另外, 本文在公共数据集上对比了各种代表性去雾算法的主客观性能、模型复杂度等, 尤其是分析了去雾后的图像对于后续目标检测任务的影响, 更全面地评价了现有算法的优劣.

2 基于深度学习的单幅图像去雾

早期的基于深度学习的图像去雾算法借鉴了传统算法^[5-7]中基于物理模型和先验知识的思路进行图像去雾. 随着深度学习的迅猛发展和图像去雾数据集的不断推出, 基于深度学习的图像去雾算法取得了众多研究成果. 有学者利用CNN进行模型参数预测以实现图像去雾, 但是这种算法容易叠加参数的预测误差. 因此学者们尝试直接利用CNN学习端到端的雾霾图像到清晰图像的映射关系模型. 但是, 由于在训练过程中往往需要大量低质量-高质量图像样本对, 而在实际应用中成对样本往往难以获得, 低质量图像往往通过物理模型对高质量图像降质得到, 这种人为仿真得到的图像无法很好地描述实际的图像降质过程, 因此训练出来的模型泛化能力差, 用于处理真实图像时, 往往会失效.

针对上述问题, 研究者们从两个角度开展了研究工作: 一是使用非成对样本开展无监督、自监督学习的图像去雾研究, 降低对成对样本的依赖; 二是将知识蒸馏、元学习、域自适应等机器学习领域的最新研究成果应用于图像去雾中, 提升网络的泛化能力, 提高实际图像的去雾效果.

本文按照基于深度学习的图像去雾方法的发展进程, 对现有的方法进行了分类, 如图1所示. 总的来说,

可以分成基于物理模型和先验知识的方法、基于像素域端到端映射的方法、基于非成对低质量-高质量样本对的无监督和自监督学习方法以及基于域知识迁移学习的方法等.

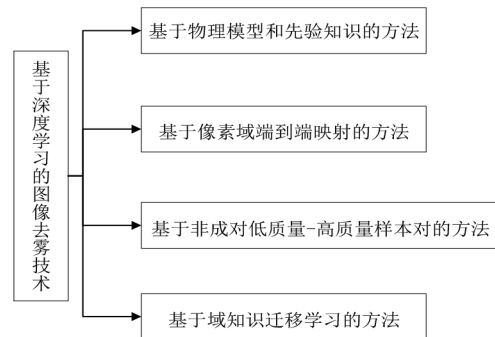


图1 基于深度学习的单幅图像去雾算法分类

2.1 基于物理模型和先验知识的图像去雾

这类方法根据雾霾图像成像模型或者降质退化模型, 利用先验知识或通过学习的方法估计模型中的参数, 然后根据退化模型计算得到去雾图像.

经典雾霾成像模型是大气光单散射模型(Atmospheric Scattering Model, ASM). 大气光单散射现象如图2所示, 光线经过一个散射媒介之后, 其原方向的光线强度会受到衰减, 并且其能量会发散到其他方向^[8-10]. 因此, 在一个有雾的环境中, 相机或者人眼接收到的某个物体(或场景)的光来源于两个部分: (1)来自于该物体(或场景)本身, 这个部分的光强受到散射媒介的影响会有衰弱; (2)来自大气中其他光源散射至相机或人眼的光强.

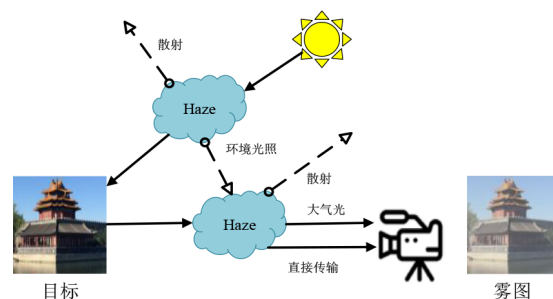


图2 大气光单散射现象示意图

基于大气光单散射现象的雾霾图像退化模型如式(1)所示:

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

其中, $I(x)$ 是有雾图像, $J(x)$ 是物体(或场景)的原始辐射, A 是全局大气光值, $t(x)$ 被称作介质透射率且 $t(x) = e^{-\beta d(x)}$, β 为全散射系数, d 为场景深度. 由式(1)可知, 同时求解 $J(x)$, $t(x)$ 和 A 是一个欠适定的问题, 往往需要利用各种先验知识来先估计透射图 $t(x)$, 并以此

求出其他未知量。

目前,图像去雾领域最具代表性的先验知识为暗通道先验(Dark Channel Prior, DCP). He 等人^[5]统计了大量的无雾图像,发现在图像的大部分区域内,存在一些像素点在至少一个颜色通道中具有非常低的值.暗通道先验可以用数学公式表达为

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} J^c(y) \right) \quad (2)$$

据此预测图像透射图,利用大气光散射模型得到最终的去雾结果。

但是,暗通道先验对大片天空等区域不够鲁棒,当雾霾图像存在大片天空区域时,处理效果并不理想. Ju 等人^[11]通过在 ASM 中引入光吸收系数设计了增强型大气光散射模型(Enhanced Atmospheric Scattering Model, EASM),结合灰色世界假设实现了图像去雾任务. Ju 等人^[12]提出区域线先验(Region Line Prior, RLP),结合 ASM 实现了图像去雾任务. 这些方法均在一定程度上提升了去雾效果。

随着深度学习的发展,学者们尝试利用 CNN 优化基于先验的图像去雾算法. Chen 等人^[13]针对基于固定图像块的暗通道先验图像去雾存在过饱和和颜色失真的缺点,提出了一种自适应图像块尺寸选择网络(Patch Map Selection Network, PMS-Net),以此提升模型的去雾效果. 虽然去雾效果有一定提升,但是算法性能仍受限于先验知识的可靠性. 为了避免先验知识的不可靠问题,学者们尝试利用深度学习强大的非线性建模能力进行模型参数的预测,再利用大气光散射模型实现图像去雾. 此类方法结合了模型和深度学习的优势,受到广大学者的关注。

2016年, Cai 等人^[14]提出了 DehazeNet 去雾网络,这是早期具有代表性的利用深度学习进行去雾的算法,目前经常作为基准对比方法. 该网络包含结合传统手工特征的特征提取层、多尺度映射层、局部极值层以及非线性回归层,通过学习雾霾退化模型中的介质透射率 $t(x)$ 进行去雾. 计算时,假设大气光值 A 为固定经验值与实际大气光值之间会有差异,通过退化模型求得得到的去雾图像也相应地会产生偏差。

为了解决这一问题, Li 等人^[15]在式(1)的基础上,将介质透射率 $t(x)$ 和大气光值 A 统一到一个变量 $K(x)$ 中,只需要求解一个 $K(x)$ 就可以实现图像增强. 具体计算方式如下:

$$J(x) = K(x)I(x) - K(x) + b \quad (3)$$

其中,

$$K(x) = \frac{\frac{1}{t(x)}(I(x) - A) + (A - b)}{I(x) - 1} \quad (4)$$

其中, b 为常量, $K(x)$ 由一体化去雾网络(All-in-One Dehazing Network, AOD-Net)来求解,网络仅包含 5 个卷积层,计算复杂度低,去雾效果有了进一步的提升。

与这一思想类似, Zhang 等人^[16]提出了一种快速的多尺度端到端去雾网络(Fast Accurate Multi-scale End-to-end Dehazing Network, FAMED-Net),该网络仅使用 3 个尺度的编码器和一个融合模块预测模型参数,再利用大气光散射模型实现图像去雾,但是该网络性能仍然受限于模型参数预测精度。

Zhang 等人^[17]提出了一种密集连接的金字塔去雾网络(Densely Connected Pyramid Dehazing Network, DCPDN). 该网络包括生成和判别两个子网络,其中生成网络可以端到端地联合学习介质透射率 $t(x)$ 和大气光值 A , 并利用式(1)得到去雾后的图像 $J(x)$, 判别网络用于判断 $J(x)$ 是否符合真实高质量图像样本的概率分布. 该网络通过两个子网络的对抗学习,达到纳什均衡。

上述这些算法在设计网络结构时未能有效地针对雾霾的降质过程设计相应的策略,导致一些特殊场景中的去雾图像存在伪影和雾霾残差。

为此, Li 等人^[18]提出了一种用于单图像去雾的水平感知渐进网络(Level-Aware Progressive Network, LAP-Net). 利用深度卷积网络渐进的估计大气光 A 、透射图 $t(x)$ 和清晰图像 $J(x)$, 且在每个阶段都用不同级别的雾霾图像进行监督. 通过这种设计,每个阶段都可以关注一个具有特定雾霾水平的区域,并恢复清晰的细节. 最后,设计了一种自适应集成策略对各个阶段去雾结果进行加权融合得到最终的去雾结果,以突出每个阶段的最佳恢复区域. 与之类似, Liu 等人^[19]利用 CNN 通过迭代来估计全局大气光 A 、透射图 $t(x)$ 和清晰图像 $J(x)$, 以使去雾网络能很好地学习到去雾过程,获得更好的去雾效果。

根据雾霾图像形成过程, Li 等人^[20]提出了一种以任务为导向的图像去雾网络,整体网络包括一个编解码器和一个由加雾过程衍生的空间变异递归神经网络. 此外,他们开发了一种多级去雾算法,通过逐步过滤雾霾残差来进一步提高去雾性能。

为了从输入的雾霾图像本身中探索有用信息指导网络的去雾过程, Bai 等人^[21]尝试开发了一个深度预去雾模块,通过参数预测,利用大气光散射模型从雾霾图像本身生成中间去雾结果. 为了更好地利用生成的参考图像,他们进一步开发了渐进式特征融合模块,对特征进行渐进式聚合,使生成的参考图像中有用的特征可以用于指导后续的图像恢复模块以获得更好的去雾效果。

上述这些基于成像物理模型和先验知识的去雾算法仅仅是利用深度学习或先验知识估计出模型参数,本质上讲属于传统方法的扩展. 受限于成像物理模型的准确性和先验知识的表达能力以及估计模型参数时

引入的误差,这种方法的去雾效果虽比传统方法有了一定程度的提升,但是未能取得大的突破.

2.2 基于像素域映射的图像去雾

随着深度学习的不断发展和应用,越来越多的学者尝试直接在像素域建立雾霾图像与清晰图像之间的端到端映射关系模型.这种方法往往以大量成对的低质量-高质量图像样本作为训练数据,利用CNN强大的非线性建模能力建立二者之间的映射模型,实现方式简单,而且可以取得很好的性能.其核心在于如何设计CNN网络架构以及有效的网络训练策略.

基于像素域映射的图像去雾算法借鉴图像分割领域常用的编码-解码结构^[22]对图像中的信息进行挖掘,结合注意力机制、特征融合等策略,提升特征的表达能

力,从而提升去雾性能.近年来,随着Transformer不断在图像分类、图像分割等领域取得进展,学者们将其应用到图像去雾中,取得了不错的性能.下面分别进行介绍.

2.2.1 全卷积编码-解码器结构

图3所示的是一个典型的全卷积编码-解码器结构图,主要包括编码器和解码器两部分.编码器用于对输入图像进行特征提取,解码器利用编码器得到的特征重构目标图像.为了充分利用图像各个层级的特征,常常在编码器和解码器间加入跳跃连接.与普通CNN网络相比,编码-解码结构能够更好地进行特征提取和表达,有效提升网络利用率^[22],在图像去雾领域得到了广泛应用.

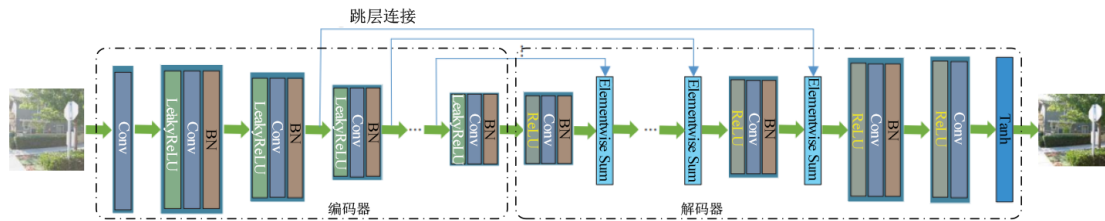


图3 编码-解码器结构示意图

生成对抗网络(Generative Adversarial Network, GAN)是2014年Goodfellow等人^[23]提出的一种网络结构,如图4所示.它包含生成器(generator)和判别器(discriminator)两个子网络,均可以利用CNN加以实现.生成器用于获得真实数据样本的特征分布,并据此

生成新的数据样本.判别器是一个二分类器,用于判别输入的是真实数据还是生成的样本.GAN建立了生成器和判别器的非合作博弈关系,其优化过程是一个极小极大博弈问题,通过迭代交替更新达到纳什均衡,从而训练出最优的网络模型.

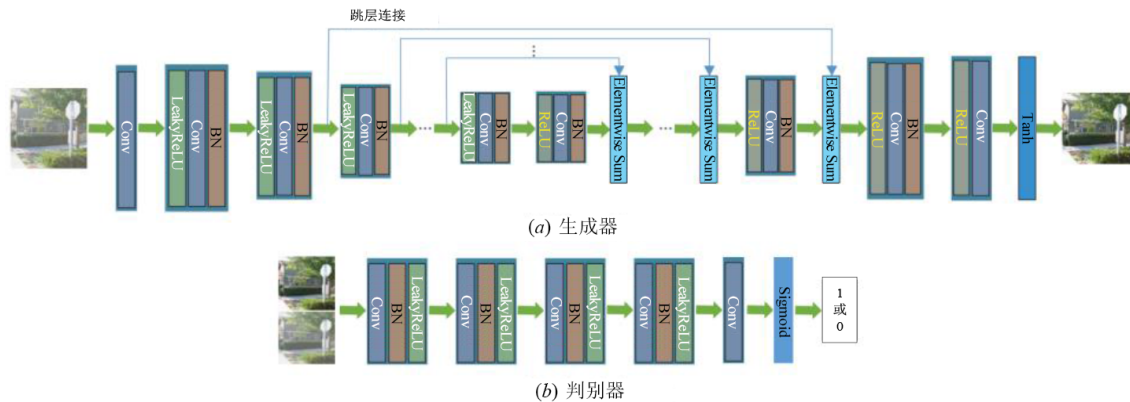


图4 GAN基本网络结构图

近年来GAN不断演进,出现了多种网络架构,如深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Network, DCGAN)^[24]、条件生成对抗网络(Conditional Generative Adversarial Network, CGAN)^[25]、Wasserstein距离生成对抗网络(Wasserstein Generative Adversarial Network, WGAN)^[26]等,并被应用到各个领域,用于解决图像分类/识别任务中的小样本数据集扩

充、图像风格迁移以及图像增强等问题.

Li等人^[27]基于CGAN架构,提出了一种像素域的端到端图像去雾网络.生成网络采用编码-解码结构,用于学习有雾图像到清晰图像的映射,使用对称层的跳跃连接(skip connection)来突破解码过程中的信息瓶颈,并使用求和方法(summation method)捕获更多有用信息;判别网络则使用了较为简单的结构,输出1或0

表示判断图像的真或假.该方法结合了感知损失和 L_1 正则化损失,提升了去雾效果.但是网络结构并未针对图像去雾进行设计,去雾能力还有待进一步提升.

Ren 等人^[28]基于编码-解码器架构,提出了一种基于门限融合网络(Gated Fusion Network, GFN)的图像去雾方法.该方法针对输入的雾霾图像,分别提取图像白平衡、对比度增强和伽马矫正特征,将雾霾图像和提取的特征图像作为门限融合 CNN 的输入,利用 CNN 对雾霾图像的每个特征图像估计一个权重矩阵,再利用权重矩阵对所有的特征图像进行融合,获得最终的去雾图像.虽然他们针对雾霾图像特性做了预处理,但是算法性能仍受限于预处理的有效性 & 合成的训练数据规模.

Qu 等人^[29]将图像去雾问题简化为图像到图像的转换问题,并提出了增强型去雾网络(Enhanced Pix2pix Dehazing Network, EPDN),在不依赖物理散射模型的情况下生成无雾图像. EPDN 由多分辨率生成器模块、增强器模块和多尺度判别器模块组成.多分辨率生成器对雾霾图像在两个尺度上进行特征提取;增强模块用于恢复去雾图像的颜色和细节信息;多尺度判别器用于对生成的去雾结果进行鉴别.虽然算法在主客观结果上都有了一定提升,但是对真实雾霾图像进行处理时,会存在过增强现象.

为了更有效地挖掘及利用各个阶段得到的特征, Liu 等人^[30]设计了一种新型双残差连接网络(Dual Residual Networks, DuRN),允许块中的第一个操作与后续块中的第二个操作交互,并利用成对操作的潜力进行图像去雾任务. Liu 等人^[31]进一步提出了 GridDehazeNet 网络结构,通过独特的网格格式结构,并利用网络注意力机制进行多尺度特征融合,充分融合底层和高层特征,网络取得了较好的映射能力.

Dong 等人^[32]提出了一种基于 U-Net 架构的具有密集特征融合的多尺度增强去雾网络(Multi-Scale Boosted Dehazing Network, MSBDN),通过一个增强解码器来逐步恢复无雾霾图像.为了解决在 U-Net 架构中保留空间信息的问题,他们设计了一个使用反投影反馈方案的密集特征融合模块.结果表明,密集特征融合模块可以同时弥补高分辨率特征中缺失的空间信息,并利用非相邻特征.但是算法的模型复杂、参数量大,而且在下采样过程中容易丢失细节信息.

Qin 等人^[33]去除了上下采样操作,提出了一种端到端特征融合注意力网络(Feature Fusion Attention Network, FFA-Net)来直接恢复无雾霾图像.该方法的主要思想是自适应地学习特征权重,给重要特征赋予更多的权重.在每一个残差块后加入特征注意力,并且对各个组的特征进行加权自适应选择,提升网络的映射能力.该

模型在合成数据集上取得了很好的客观指标,常被作为基准对比方法.在此基础上, Wu 等人^[34]首次将有监督对比损失(Contrastive Learning, CL)引入图像去雾领域,通过拉大去雾图像与雾霾图像距离、缩小去雾图像与清晰图像距离,进一步提升图像去雾性能.

Zheng 等人^[35]为了使去雾模型能够处理高分辨率图像,尤其是超高清或 4K 分辨率的图像,提出了一种超高分辨率图像去雾模型.首先在低分辨率图像上使用特征提取块重建双边系数.再利用回归的仿射双边网格,在全分辨率特征的指导下生成高质量的特征图.此外,为了提供更丰富的颜色和边缘信息,他们考虑了所有的 RGB 通道,从而使去雾网络可以更好地恢复细节.

2.2.2 Transformer 结构

Transformer 最初是针对自然语言处理任务^[36]提出的,通过多头自注意力机制和前馈 MLP(Multi-Layer Perceptron)层堆叠,捕获单词之间的非局部交互. Dosovitski 等人^[37]创新性地提出了用于图像识别的 Transformer 模型(Vision Transformer, ViT).最近的研究工作探索了 Transformer 应用于低层视觉领域的可行性,例如,图像处理 Transformer(Image Processing Transformer, IPT)^[38].与 ViT 类似, IPT 直接将通用的 Transformer 应用于图像块中.

近期,一些研究者认为当使用 ViT 中基于图像块的架构时,复杂自注意力机制的必要性值得商榷. MLP-Mixer^[39]采用了一个简单标记混合的 MLP 来取代 ViT 中的自注意力机制,构建了一个全 MLP 的体系结构.文献[40]提出了 gMLP,通过在视觉符号上应用了一个空间门控单元达到了与 Transformer 相当的性能. ResMLP^[41]采用仿射变换作为加速层归一化的替代方法,构建了一个残差架构,在图像分类领域取得了很好的效果.

受以上工作的启发, Tu 等人^[42]设计了第一个基于 MLP 的通用底层视觉类 U-Net 骨干网络(Multi-AXIs MLP, MAXIM)用于图像去雾.该网络平衡了局部和全局算子的使用,在任意尺寸图片上都具有全局感受野并且只需要线性复杂度.但是该算法依赖于强大的算力,模型参数量巨大,而且仅在合成数据域上拥有强大的映射能力,对于真实雾霾图像的处理效果仍有待于进一步的提升.

总的来看,基于像素域端到端映射的图像去雾方法实现简单,通过针对性地设计网络结构以及有效的特征加权融合策略,在合成数据集上取得了很不错的客观指标.但是这类方法存在两个重要缺陷.

(1) 网络训练需要以大量的低质量-高质量图像样本对作为支撑,实际应用中低质量-高质量图像数据样本对获取困难.目前普遍的做法是以仿真的方式由已

知的高质量图像,经过退化模型得到低质量图像,组成训练样本对。很显然,图像质量退化模型往往无法很好地模拟真实的降质过程,采用仿真数据训练得到的去雾模型在处理真实图像时常会失效,使得研究成果难以真正得到应用。

(2)网络模型的泛化能力差。在一个数据集上训练的模型应用于另一个数据集上,由于样本分布不一致,去雾后的图像效果难以尽如人意。

为此,研究者们尝试从两个角度出发,来解决上述问题:一是利用实际采集的非成对图像样本对网络进行训练,提升对真实图像的处理性能;二是利用域知识迁移学习的思路,提升网络模型的泛化能力,从而提高对实际图像的去雾效果。

2.3 基于非成对样本的图像去雾

根据是否利用高质量图像样本训练,基于非成对样本的图像去雾方法又可以分为基于无监督学习和基于自监督学习的方法。

2.3.1 基于无监督学习的方法

这类方法普遍基于生成对抗网络(Generative Adversarial Network, GAN)架构,采用数据驱动的方式,从大量无标注、非成对的数据中分别学习低质图像和高质图像的特征,用于图像去雾。

Zhu 等人^[43]提出的循环一致对抗网络(Cycle-consistent Adversarial Networks, CycleGAN)是一种比较具有代表性的基于非成对样本的网络结构,该网络是面向图像风格迁移任务设计的。整体架构如图5所示,包含了两个生成器(G, F)和两个判别器(D_x, D_y)。生成器 G 负责将 X 域图像映射到 Y 域,生成器 F 负责将 Y 域图像映射到 X 域;判别器 D_x 用于判断输入到判别器的图像是否属于 X 域。假设存在一对非成对样本 $\{x_i, y_i\}, x_i \in X, y_i \in Y$,以正向训练为例, x_i 用于训练 D_x ,标签为真, $G(x_i)$ 用于训练 D_y ,标签为假,此时判别器 D_y 可以监督生成器 G 的训练;通过优化输入 x_i 与 $F(G(x_i))$ 之间进行 L_1 范数损失,可以同时监督生成器 G 和 F 的训练。这个损失称为循环一致性损失;反向训练时同理。通过正反向交替训练可以达到训练生成器 G 和 F 的目的。

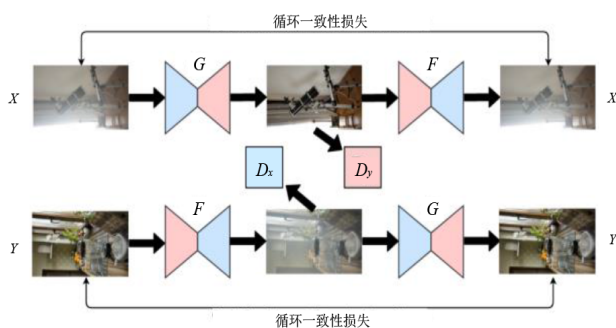


图5 CycleGAN的整体架构

采用类似思想的方法还有对偶学习生成对抗网络^[44](Dual learning Generative Adversarial Network, DualGAN)和学习跨域关系的生成对抗网络^[45](Discover cross-domain relations Generative Adversarial Network, DiscoGAN)等。

Engin 等人^[46]基于CycleGAN非成对样本无监督学习的思想,提出了CycleDehaze网络,用于图像去雾。该网络在循环一致性损失中增加了感知损失,并且通过基于拉普拉斯金字塔的上采样方法获得高分辨率图像,利用大量的无标记图像样本进行无监督训练,使得网络更适用于真实雾霾图像。但是由于缺乏有力的约束,去雾结果常常不尽如人意。

Li 等人^[47]尝试使用一对神经网络之间的对抗博弈,集成图像去雾以及雾霾密度估计两个任务。生成器同时学习恢复无雾霾图像和雾霾密度。他们还将两个任务以串联和并联的方式组合,以支持在不同级别上共享信息。生成器的架构隐式地形成了一个允许特征选择的去雾模型集合。多尺度鉴别器通过学习检测去雾伪影和去雾图像与雾霾空间变化之间差异来与生成器进行竞争。另外,他们设计了任务驱动的训练方式使模型得到的去雾结果可以更好的适应后续目标检测任务。

本文提出了一种新的无监督图像去雾算法USID-Net^[48],基于解纠缠表示区分雾霾图像的内容和雾霾信息。为了获得更真实的去雾霾图像,利用八度卷积设计了编码器来有效地捕捉雾霾特征信息。此外,考虑到性能和内存存储之间的权衡,设计了一个紧凑的多尺度特征注意(Multi-scale Feature Attention, MFA)机制,集成了多尺度特征表示和注意机制,提升了去雾性能。

Zhao 等人^[49]将基于先验和基于学习的方法结合起来,提出了一个两阶段弱监督去雾框架RefinedNet,以发挥二者的优势。在第一阶段,采用暗通道先验恢复可见性;在第二阶段,细化第一阶段的初步去雾结果,通过非成对的雾霾和清晰图像的对抗学习来提高真实性。为了获得更优越的结果,还提出了一种有效的感知融合策略来混合不同的去雾输出,可以有效提升去雾效果。

2.3.2 基于自监督学习的方法

自监督学习利用图像的先验知识或者统计分布规律训练网络模型,无需标注,网络可以学习到图像本身包含的结构信息。

零样本学习(Zero Shot Learning, ZSL)是一种有效的自监督学习方式,直接从输入图像本身学习对任务有用的特征。这种学习方式能够有效缓解合成数据集与真实场景的域迁移问题,尤其适合于真实标签难以

获得的真实场景图像去雾任务. 但是由于缺乏任何参考, 零样本学习非常具有挑战性. 为此, 许多学者开展了广泛的研究工作, 将零样本学习应用于图像去雾中, 并取得了重要的研究进展.

Li 等人^[50]提出了一种基于零样本学习的图像去雾方法 (Zero-shot Image Dehazing, ZID). 利用解耦思想根据先验模型将雾霾图像视为清晰图像、透射图和大气光的融合. 利用单个雾霾图像进行学习和推理, 不遵循在大规模数据集上训练深度模型的传统范式. 这能够避免数据收集和使用合成雾霾图像来解决现实世界图像的域转移问题. 但是, 对一张图像需要 500 轮次迭代, 而且去雾结果中常常存在大量伪像.

Kar 等人^[51]利用 Koschmieder 模型, 提出了一种零样本图像恢复模型. 通过维持图像降质前后的 Koschmieder 模型参数关系以及使用一些无参考损失来实现零样本网络优化. 所提出的零样本框架在图像去雾、水下图像恢复和低照度图像复原任务上均取得了良好效果. 但是, 由于需要对输入图像进行 1 000 轮迭代, 训练效率不高, 且处理效果仍有待提升.

虽然有学者尝试仅用单个降质图像来训练图像去雾模型, 但总的来看, 由于缺少有力的监督信息, 现有基于自监督学习的去雾方法普遍存在网络训练效率低下、有用信息难以获取或过度依赖于降质建模等问题.

对比来看, 无监督学习虽然减轻了收集成对样本数据的压力, 但是由于缺少必要的监督信息, 会导致网络模型在训练过程中不易收敛. 而自监督学习方法旨在利用图像自身特性, 虽然在一定程度上降低了对训练数据的依赖, 但是需要加入强有力的先验知识约束, 去雾性能仍有待于进一步提升.

2.4 基于域知识迁移学习的图像去雾

近年来, 机器学习领域的各种新思路、新方法不断涌现, 如知识蒸馏、集成学习、元学习等, 学者们将这些训练策略应用于图像去雾中, 有效提升了去雾性能. 这类方法的基本思路是利用从其他域/任务中学到的知识, 来指导、辅助模型的训练和学习, 从而提升模型的泛化能力, 对于真实图像能取得较好的去雾效果.

“知识蒸馏”的概念是 Hinton 等人^[52]于 2015 年提出来的, 其基本思想是将复杂的教师模型学习到的知识迁移到轻型的学生模型上, 指导学生模型的训练和学习. 研究表明, 与教师模型相比, 学生模型能以较低的模型复杂度, 获得略逊于甚至于相当的性能.

受此启发, Hong 等人^[53]提出了一种基于知识蒸馏的去雾网络 (Knowledge-Distill Dehazing Network, KDDN), 该网络通过异构任务学习图像去雾知识. 教师网络与学生网络使用相同的编码-解码网络, 教师网络学习清晰图像在网络各层的特征表达, 用于指导学

生 (去雾) 网络学习雾霾-清晰图像映射关系模型. 此外, 作者还设计了一个用于学生图像去雾网络的空间加权信道注意残差块, 以自适应地学习内容感知的信道级权重, 并更加关注密集雾霾区域重构的特征.

元学习是“迁移学习”中的一种, 可以模拟人类的学习过程. 模型通过不断适应每个具体任务学习到“元”知识, 使其具备了一种抽象的学习能力, 可以快速适应新的目标任务. 元学习在小样本学习中得到了广泛应用, 并取得了不错的性能. 据此, 本研究团队^[54]提出了一种元注意去雾网络 (Meta-Attention Dehazing Network, MADN), 可以从雾霾图像中直接恢复清晰图像. 结合并行操作和增强模块. 元注意模块根据当前输入的雾霾图像自动选择最合适的去雾网络结构. 研究表明, 该方法可以获得更优的主客观去雾效果.

Li 等人^[55]提出了一种有监督和无监督相结合的半监督图像去雾网络 (Semi-Supervised image dehazing network, SSL), 该网络包含有监督学习分支和无监督学习分支两个部分, 两部分共享权重. 在有监督分支中使用合成数据集进行网络训练, 神经网络受到监督损失函数的约束, 从有标注的样本中学习知识. 而无监督分支则使用无标签的真实数据, 并通过稀疏的暗通道和先验梯度来充分利用清晰图像的属性约束网络. 二个分支学到的知识相结合, 充分利用学到的知识, 提升去雾效果. 研究表明, 与无监督学习相比, 这种学习方式可以充分发挥合成数据和真实数据的作用, 从而获得更优的去雾性能.

但是由于收集到的真实场景下的雾霾图像数量仍十分有限, 模型对真实图像的处理能力仍然有待于进一步提升. 因此很多学者借鉴域泛化、域自适应等新的研究思路, 提升模型的泛化能力.

域泛化 (Domain Generalization, DG) 的目标是学习一个模型, 该模型在一个未知的新域中也可以取得很好的性能. 域自适应 (Domain Adaptation, DA) 旨在减少不同域之间的差异, 利用现有的源域数据对模型进行训练, 使其在给定目标域上的性能最大化. 域自适应和域泛化的区别在于, 域自适应在模型训练过程中可以访问目标域数据, 而域泛化则不可访问这些数据. 这使得域泛化比域自适应更具挑战性, 但在实际应用中更有意义.

为了缓解合成数据和真实图像域之间的差异, Shao 等人^[56]提出了一种端到端的域自适应网络框架 (Domain Adaptation Dehazing Network, DADN). 该网络由一个图像转换模块和两个图像去雾模块组成. 具体而言, 首先采用双向变换网络将图像从一个域变换到另一个域, 以弥补合成域和真实域之间的差距. 然后, 使用变换前后的图像来训练所提出的两个去雾网络.

在此阶段,通过利用清晰图像的特性(如暗通道先验和图像梯度平滑)将真实图像纳入到去雾训练中,以进一步提升真实图像的去雾效果.目前,上述两种方法常被作为基准对比方法.

Zhang 等人^[57]将去雾任务转化为一个半监督的域迁移问题.通过两个辅助域迁移任务来捕捉真实图像的雾霾特性,并将合成的雾霾图像与真实世界的雾霾图像对齐,以减少域差距.此外,作者使用差分优化策略来搜索最优的去雾网络架构,进一步提升了网络的泛化能力,对真实图像能取得更优的去雾效果.

Chen 等人^[58]指出上述算法未能合理利用先验知识,而且模型对于真实雾霾图像的处理性能仍然十分有限.基于此,作者基于各种先验知识分别设计多个先验损失函数,用于网络的优化训练.具体来说,对于在合成数据集上预训练的模型,利用先验损失以无监督的方式使用真实雾霾图像对网络进行微调,以提升模型对于真实图像的去雾效果.

上述的研究工作将元学习、零样本学习、域自适应等机器学习领域最新的研究成果应用于图像去雾领域,极大地提升了真实雾霾图像的去雾效果,在解决模型的泛化能力方面取得了重要的研究进展.这为开展图像去雾的研究工作提供了新的思路,已经成为目前

新的研究趋势.然而,利用上述新思想解决图像去雾问题仍存在一些挑战亟待解决,如元学习训练稳定性差、需要大量的训练样本;零样本学习在测试阶段需要多次迭代,对于实时任务有很大限制;域自适应算法对于域间差距较大的样本泛化能力仍然十分有限,等等.

3 图像去雾常用数据集及评价指标

3.1 图像去雾常用数据集

随着基于深度学习的图像去雾研究的不断深入,图像去雾数据集不断被推出,极大地促进了该领域的研究进展.表1所示的是目前最常用的8个去雾数据集,表中给出了发布时间、数据规模和图像样本数据生成方式.

可以看出,除了RESIDE(REALISTIC SINGLE IMAGE DE-hazing)和MDID(Multi-Degraded Image Dataset)数据集,其他数据集的规模普遍较小.HazeRD数据集^[59]仅包含15张清晰图像和75张相应具有不同程度的雾霾的合成图像.I-HAZE^[60]包含35对雾霾和相应的无雾室内图像,其中雾霾图像是使用专业雾霾机器产生的真实雾霾条件下拍摄的.O-HAZE^[61]包含45对户外场景的雾霾和清晰图像.Dense-HAZE^[62]和NH-HAZE^[63]数据集分别包含55对户外场景的雾霾和清晰图像.

表1 图像去雾常用数据集

数据集	时间	规模	生成方式	数据集链接
HazeRD ^[59]	2017	75	模型合成	https://labsites.rochester.edu/gsharma/research/computer-vision/hazerd/
I-HAZE ^[60]	2018	35	真实	https://data.vision.ee.ethz.ch/cvl/ntire18/i-haze/
O-HAZE ^[61]	2018	45	真实	https://data.vision.ee.ethz.ch/cvl/ntire18/o-haze/
Dense-HAZE ^[62]	2019	55	真实	https://data.vision.ee.ethz.ch/cvl/ntire19/dense-haze/
NH-HAZE ^[63]	2020	55	真实	https://data.vision.ee.ethz.ch/cvl/ntire20/nh-haze/
RESIDE ^[64]	2019	87 125	模型合成/真实	https://sites.google.com/view/reside-dehaze-datasets/reside-standard?authuser=3D0
MDID ^[65]	2020	30 346	模型合成	—
BeDDE ^[66]	2020	208	真实	https://github.com/xiaofeng94/BeDDE-for-defogging

RESIDE数据集是2019年Li等人^[64]提出的,训练集包括室内(Indoor Training Set, ITS)和室外(Outdoor Training Set, OTS)合成数据集,分别包含13 990对和14 427对图像.此外测试数据集包括合成目标测试集(Synthetic Objective Testing Set, SOTS),未注释的真实世界朦胧图像(Unannotated Real-world Hazy Images, URHI)数据集,真实世界任务驱动测试集(Real-world Task-driven Testing Set, RTTS)以及混合主观测试数据集(Hybrid Subjective Testing Set, HSTS).这些数据集广泛用于各种算法的定量和定性比较实验.SOTS分别由500对室内和室外雾霾图像组成.URHI拥有超过4 000张真实世界的雾霾图像.RTTS有4 322张带有标签的雾霾图像,用于后续的目标检测任务.HSTS数据集包含10张合成图和10张真实图用于主观测试.

上述数据集只考虑了理想状态下单一的雾霾降质因素,未考虑实际采集的雾霾图像也会存在模糊、噪声等其他降质因素.本研究团队^[65]提出的MDID数据集包含30 346个图像对,是根据NYU Depth V2 dataset^[67]构建的复合降质数据集,包括雾霾、模糊、噪声等三种降质类型.

由于实际应用场景中成对的雾霾-清晰图像很难获得,因此上述数据集样本对中的雾霾图像多为通过模型仿真生成.通过对同一场景进行长时间拍摄及后续的对齐处理,Zhao等人^[66]在2020年提出了BeDDE(Benchmark Dataset for Dehazing Evaluation)数据集.共包含208对自然图像,每对图像由自然雾霾图像和对齐良好的清晰参考组成,原始图像来自23个城市.这一数据集的发布为开展真实雾霾图像的去雾研究提供了

重要的数据基础,但是不足之处是数据集规模仍比较有限.

3.2 图像去雾评价指标

图像去雾性能的评价包括主观和客观两种. 主观评价是指通过观察者的主观视觉感受来判断图像质量,可以直接反映人的主观感受,是最为准确、可靠的质量评价方法. 但是,主观评价方法存在操作难度大、成本高等局限性^[68]. 因此,在验证算法性能优劣时,往往采用客观评价指标.

客观评价根据人眼视觉系统感受建立数学模型. 相较于主观评价,客观评价具有可重复性、成本低等优点. 根据对参考图像的需求,客观评价方法可分为全参考、半参考和无参考等三种^[68]. 图像去雾领域常用的客观评价指标如表 2 所示,其中峰值信噪比(Peak Signal to Noise Ratio, PSNR)、结构相似度(Structural SIMilarity, SSIM)是两种最具代表性的全参考评价指标. PSNR 计算方式如下:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_J^2}{\text{MSE}} \right) \quad (5)$$

其中, MAX_J^2 为图像可取到的最大像素值, MSE 为去雾图像 \hat{J} 和对应真实图像 J 的均方误差. SSIM 从亮度、对比度和结构三个方面来衡量图像的相似度. SSIM 的值范围为 $[0, 1]$, 该值越大图像失真率就越小. 其计算方式如下:

$$\text{SSIM}(J, \hat{J}) = \left[l(J, \hat{J})^\alpha \right] \left[c(J, \hat{J})^\beta \right] \left[s(J, \hat{J})^\gamma \right] \quad (6)$$

其中, l, c, s 分别表示图像 J 和图像 \hat{J} 的亮度、对比度和饱和度的比较. 在实际的工程计算中,超参数 α, β 和 γ 一般设为 1.

表 2 图像去雾常用评价指标

指标	全参考	无参考	应用领域
PSNR	√		通用
SSIM	√		通用
BRISQUE ^[69]		√	通用
NIQE ^[70]		√	通用
MetaQA ^[71]		√	通用
VI ^[66]	√		去雾
RI ^[66]	√		去雾
DHQI ^[72]		√	去雾

此外,无参考评价指标如 BRISQUE^[69], NIQE^[70], MetaQA^[71] 等也用于图像去雾质量评价. BRISQUE^[69] 的基本思路是从图像中提取 MSCN (Mean Subtracted Contrast Normalized) 系数,然后将 MSCN 系数拟合成非对称广义高斯分布,利用失真对分布造成的参数影响来提取高斯分布特征,最后使用 SVR 实现特征到分数的回归计算,得到图像质量的评估结果. NIQE^[70] 基于空间域自然场景统计(NSS)模型,将失真图像的质量表

示为模型统计量与失真图像统计量之间的简单距离度量. MetaQA^[71] 借助元学习的思想来学习各种失真所共享的先验知识,提升质量评价模型的泛化能力.

Zhao 等人^[66] 借助收集的 BeDDE 数据集,提出了两个全参考图像去雾评价指标,能见度指数(Visibility Index, VI)和真实度指数(Realness Index, RI),专门用于去雾图像的质量评价. 其中,VI 利用图像与其清晰参考之间的相似性来评估图像质量,RI 利用去雾后的图像与清晰参考图在特征空间的相似性来评估去雾图像的真实性. 也有学者针对不同类型的图像去雾,设计了一些专用的质量评价指标. 如无参考去雾质量指数(DeHazing Quality Index, DHQI)^[72], 通过提取和融合去除雾霾特征、结构保护特征、过度增强等三组去雾关键特征对去雾结果进行评价.

雾霾会严重影响户外视觉系统性能,因此去雾后目标检测与跟踪等视觉任务的精度也越来越多地被用于去雾方法的性能评价中. 如图 6 所示,雾霾会导致目标的错检、漏检等问题. 因此去雾算法的性能评价不仅包括上述的主客观评价指标,去雾后目标检测与跟踪的精度也越来越多地被用作性能评价准则.



(a) 雾霾图像目标检测结果 (b) 去雾图像目标检测结果

图 6 真实世界图像的目标检测结果

3.3 代表性算法的性能对比

为了更清晰地展示各类方法的性能,本节将分别给出各个阶段比较具有代表性的去雾算法的客观指标和主观视觉对比结果. 表 3 分别给出了 10 种典型的深度学习去雾算法在合成户外数据集 SOTS 和 HazeRD 上的 PSNR 和 SSIM 指标对比结果以及在 HSTS 数据集上的 DHQI 指标对比结果,标红数据表示最优结果. 图 7 和图 8 分别展示了各个算法在合成雾霾图像和真实雾霾图像的处理结果. 可以看出,早期基于物理模型的 DehazeNet^[14] 和 AOD-Net^[15] 等算法受限于成像模型的表达能力,去雾效果较差;而基于像素域映射的算法 EPDN^[29], FFA-Net^[33], MADN^[54] 则明显取得了更好的客观指标,这不仅得益于 CNN 网络强大的映射能力,同时也受助于大规模合成数据集的支撑. 但是上述

表3 图像去雾代表性算法的客观指标对比

	算法	DehazeNet ^[14]	AOD-Net ^[15]	EPDN ^[29]	FFA-Net ^[33]	CycleDehaze ^[46]	RefineDNet ^[49]	ZID ^[50]	MADN ^[54]	SSL ^[55]	DADN ^[56]
数据集	源码链接	https://github.com/caibolun/DehazeNet	https://github.com/weber0522bb/AODnet-by-pytorch	https://github.com/ErinChen1/EPDN	https://github.com/zhilin007/FFA-Net	https://github.com/niranjangavade98/CycleDehaze-Pytorch	https://github.com/xiaofeng94/RefineDNet-for-dehazing	https://github.com/liboyun/ZID	https://github.com/TongyJia/MADN	https://github.com/Prevalenter/semi-dehazing	https://github.com/HUSTSYJ/DA_dahazing
SOTS-I	PSNR	21.14	19.06	21.55	36.39	17.81	24.23	19.83	26.69	21.87	25.82
	SSIM	0.847 2	0.850 4	0.907 1	0.988 6	0.809 5	0.943 1	0.835 3	0.932 3	0.874 3	0.928 8
SOTS-O	PSNR	24.75	22.71	22.32	33.57	19.95	20.61	19.83	28.13	25.36	26.72
	SSIM	0.927 0	0.911 2	0.868 3	0.984 0	0.885 8	0.879 8	0.835 3	0.957 5	0.921 0	0.922 0
HazeRD	PSNR	15.87	16.85	15.60	16.74	15.65	17.39	12.84	17.50	17.34	16.54
	SSIM	0.787 6	0.792 5	0.749 2	0.793 5	0.756 9	0.855 3	0.505 0	0.750 6	0.802 9	0.763 6
HSTS	DHQI	40.327 0	59.702 0	70.316 0	68.872 9	—	43.694 0	50.562 8	63.500 1	40.327 0	63.803 4

注:加粗数据表示最优结果。

有监督图像去雾算法对真实雾霾图像的主观处理效果并不理想,存在去雾不彻底或过度增强问题。目前基于物理模型的无监督图像去雾算法 CycleDehaze^[46], RefineDNet^[49]和自监督图像去雾算法 ZID^[50]仍处于探索阶段。由于使用非成对的数据或未使用数据对网络进行训练,网络的鲁棒性十分有限,对合成和真实数据集的处理效果并不尽如人意。而基于域知识迁移学习的半监督去雾算法 SSL^[55]和 DADN^[56],可以学习来自其他域的知识,虽然在合成数据上的指标不如有监督算法,但是其对真实雾霾图像的泛化能力非常突出。

表4所示的是分别采用10种去雾算法对图像处理后的目标检测精度对比结果。可以看出,部分去雾算法如 EPDN, ZID 等处理后的图像给后续的目标检测任务带来了负面影响,这主要是因为这些去雾图像存在信息丢失等问题。而面向真实雾霾图像处理的算法如 SSL, DADN 等则可以有效提升后续视觉任务的性能。

表5给出了各个算法的模型参数量及运行时间对比。所有算法都是在—台带有 NVIDIA GeForce RTX 2080Ti GPU 的台式计算机上实现,运行时间为在 500 张 512×512 像素图像的平均测试时间。可以看出,随着去雾处理性能的不不断提升,去雾算法的模型参数量、运行时间都有所增加。也就是说,去雾性能的提升是以计算复杂度为代价的,很显然,这不利于算法的实际推广应用。

4 思考与展望

近年来,基于深度学习的图像去雾研究取得了重

要进展。基于物理模型和先验知识的图像去雾方法受限于成像/退化模型的表达能力,往往会忽视掉一些重要的特征信息,导致图像去雾效果难以尽如人意。基于像素域端到端映射的方法可以通过网络自动学习、提取到与图像去雾任务相关的特征,通过多层的非线性变换建立低质量图像与高质量图像之间的映射关系模型。由于利用了大数据进行训练,因此该方法特征表达能力远超过传统方法,大大提升了合成雾霾图像的去雾效果,但是对真实图像的去雾效果难以令人满意,严重影响了算法的实际应用。近期的研究工作则尝试通过各种知识迁移学习策略或在训练中引入真实雾霾图像,以提升算法对真实雾霾图像的处理性能。

尽管深度学习在图像去雾领域取得了一定的进展,但是还存在—些问题。在未来的研究工作中,可以从以下几个方面开展。

(1)“域鸿沟”问题

如前所述,现有图像去雾的研究中,普遍通过仿真方式得到训练样本,与真实图像之间存在严重的“域鸿沟”,导致训练出来的模型应用到真实图像时性能下降严重,甚至完全失效。基于非成对样本、以半监督学习方式训练图像去雾模型,存在训练不稳定、内容不一致等问题。如何设计有效的机制和策略,真正跨越这—“域鸿沟”,是未来图像去雾值得深入研究的重要方向。

(2)知识迁移

如何设计高效的知识迁移策略,提高真实雾霾图像的去雾性能仍是未来—段时间的研究热点。近期,

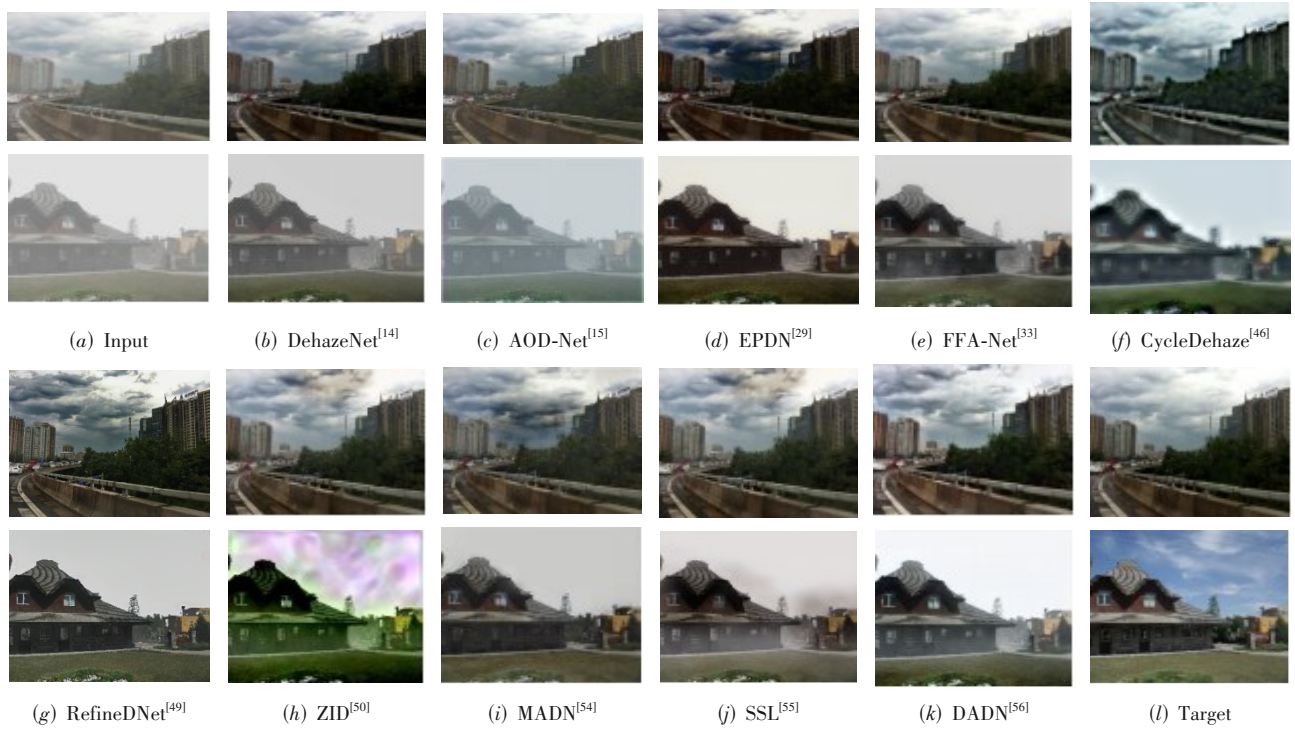


图7 合成数据集主观结果对比

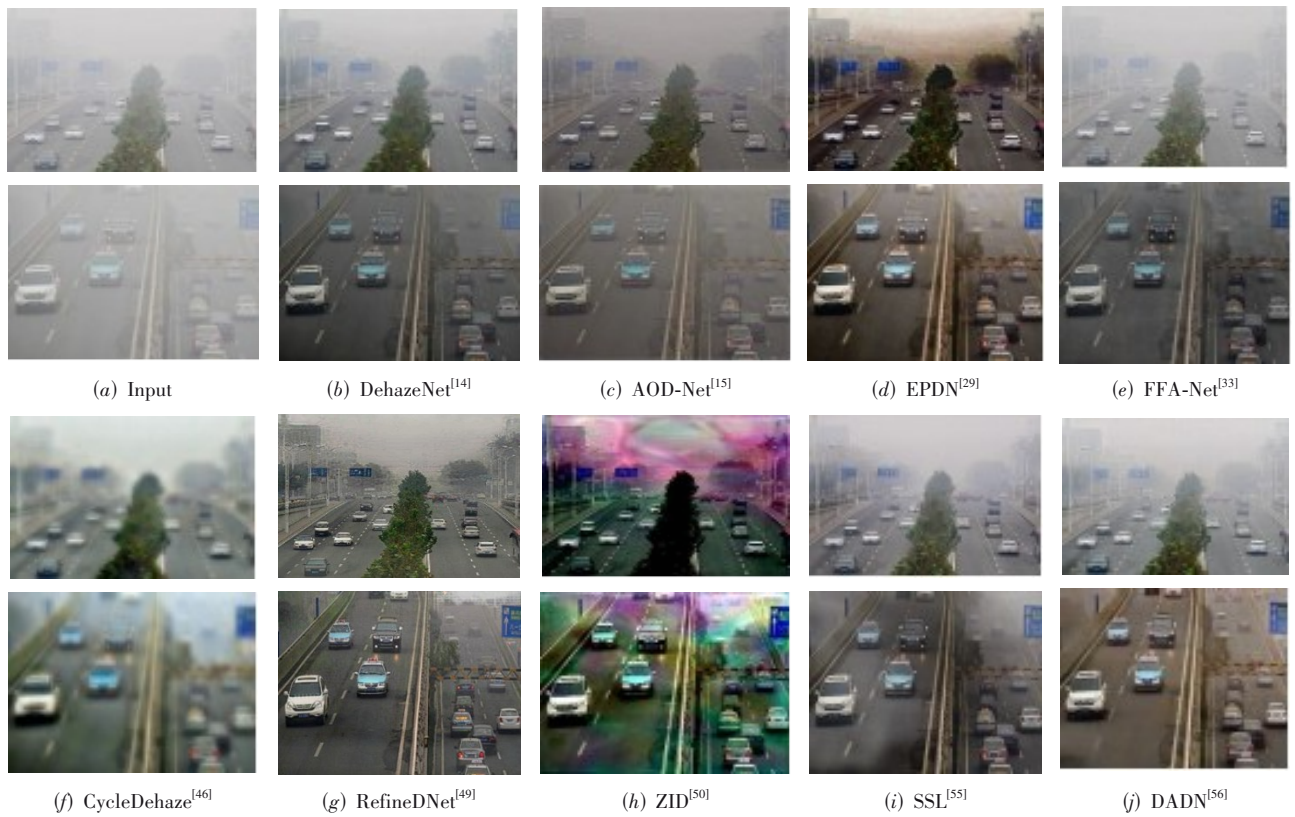


图8 真实雾霾图像去雾结果对比

元学习、域泛化、域自适应、零样本学习等机器学习领域的新成果不断出现,将这些新思路应用于图像去雾

领域,值得开展深入的研究,有望取得创新性的研究成果.

表4 RTTS数据集上的目标检测结果

模型	Person	Bicycle	Car	Motorbike	Bus	mAP
Hazy	0.721	0.441	0.617	0.473	0.364	0.523
DehazeNet ^[14]	0.722	0.424	0.622	0.485	0.375	0.526
AOD-Net ^[15]	0.720	0.448	0.629	0.498	0.376	0.534
EPDN ^[29]	0.706	0.419	0.598	0.458	0.347	0.506
FFA-Net ^[33]	0.727	0.452	0.626	0.500	0.374	0.536
CycleDehaze ^[46]	0.684	0.421	0.581	0.424	0.325	0.487
RefineDNet ^[49]	0.726	0.459	0.625	0.499	0.231	0.536
ZID ^[50]	0.584	0.336	0.468	0.383	0.320	0.418
MADN ^[54]	0.726	0.454	0.628	0.499	0.374	0.536
SSL ^[55]	0.722	0.440	0.625	0.500	0.379	0.533
DADN ^[56]	0.716	0.431	0.653	0.495	0.397	0.538

注:加粗数据表示最优结果。

表5 算法模型参数量及运行速度对比

模型	运行平台	模型参数量/M	运行时间/s
DehazeNet ^[14]	Matlab(C)	0.008	1.339 9
AOD-Net ^[15]	PyTorch(G)	0.002	0.002 2
EPDN ^[29]	PyTorch(G)	17.379	0.025 4
FFA-Net ^[33]	PyTorch(G)	4.456	0.127 9
CycleDehaze ^[46]	Tensorflow(G)	11.383	99.000 0
RefineDNet ^[49]	PyTorch(G)	65.795	0.454 1
ZID ^[50]	PyTorch(G)	40.406	59.990 0
MADN ^[54]	PyTorch(G)	0.605	0.018 2
SSL ^[55]	PyTorch(G)	9.231	0.077 7
DADN ^[56]	PyTorch(G)	54.591	0.035 0

(3)模型复杂度

虽然一些研究工作尝试设计轻型网络结构以满足户外视觉系统的需求,但模型的性能也受到了限制.如何在算法处理性能和速度之间获得良好的折中,可以开展深入的研究.

(4)Transformer结构

Transformer是目前计算机视觉领域的新兴研究热点,各种Transformer网络结构不断出现,并被应用于语义分割、目标检测等各种视觉任务中,获得了不错的性能.未来可以考虑将Transformer应用于图像去雾中,利用其强大的图像内部结构表达能力,提升去雾性能.

(5)去雾后图像质量的评价

如何对去雾后的图像质量进行评价,从而有效指导去雾算法的设计一直是图像去雾领域的难题.有学者尝试开展了专门的去雾图像质量评价方法研究,但是目前还缺乏权威、被普遍认可的评价准则.未来,结合具体的后续智能化分析处理任务设计图像去雾算法性能评价准则,也是一个值得探索的研究方向.

5 结论

单图像去雾是计算机视觉和图像处理领域的重要研究方向.本文综述了基于深度学习的图像去雾最新研究进展,并在公共数据集上对部分算法的性能进行对比分析,最后对基于深度学习的图像去雾技术研究中的困难和挑战进行了分析,对未来的发展趋势进行了思考与展望.

参考文献

- [1] LI J F, ZHUO L, ZHANG H, et al. Effective data-driven technology for efficient vision-based outdoor industrial systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4344-4354.
- [2] 郭璠,蔡自兴,谢斌,等.图像去雾技术研究综述与展望[J].计算机应用,2010,30(9):2417-2421.
GUO F, CAI Z X, XIE B, et al. Review and prospect of image dehazing techniques[J]. Journal of Computer Applications, 2010, 30(9): 2417-2421. (in Chinese)
- [3] 沈琛,曹风云,杨雪洁.夜间图像去雾研究综述与展望[J].电子测量与仪器学报,2020,34(11):101-114.
SHEN C, CAO F Y, YANG X J. Review and prospect of nighttime haze removal[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(11): 101-114. (in Chinese)
- [4] 郑凤仙,王夏黎,何丹丹,等.单幅图像去雾算法研究综述[J].计算机工程与应用,2022,58(3):1-14.
ZHENG F X, WANG X L, HE D D, et al. Survey of single image defogging algorithm[J]. Computer Engineering and Applications, 2022, 58(3): 1-14. (in Chinese)
- [5] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [6] ZHU Q S, MAI J M, SHAO L. A fast single image haze removal algorithm using color attenuation prior[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3522-3533.
- [7] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1674-1682.
- [8] MCCARTNEY E J, HALL F F. Optics of the atmosphere: Scattering by molecules and particles[J]. Physics Today, 1977, 30(5): 76-77.
- [9] NAYAR S K, NARASIMHAN S G. Vision in bad weather [C]//Proceedings of the Seventh IEEE International Conference on Computer Vision. Kerkyra: IEEE, 1999: 820-827.

- [10] NARASIMHAN S G, NAYAR S K. Removing weather effects from monochrome images[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai: IEEE, 2001: II.
- [11] JU M Y, DING C, REN W Q, et al. IDE: Image dehazing and exposure using an enhanced atmospheric scattering model[J]. IEEE Transactions on Image Processing, 2021, 30: 2180-2192.
- [12] JU M Y, DING C, GUO C A, et al. IDRLP: Image dehazing using region line prior[J]. IEEE Transactions on Image Processing, 2021, 30: 9043-9057.
- [13] CHEN W T, DING J J, KUO S Y. PMS-Net: Robust haze removal based on patch map for single images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 11673-11681.
- [14] CAI B L, XU X M, JIA K, et al. DehazeNet: An end-to-end system for single image haze removal[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [15] LI B Y, PENG X L, WANG Z Y, et al. AOD-Net: All-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 4780-4788.
- [16] ZHANG J, TAO D C. FAMED-net: A fast and accurate multi-scale end-to-end dehazing network[J]. IEEE Transactions on Image Processing, 2020, 29: 72-84.
- [17] ZHANG H, PATEL V M. Densely connected pyramid dehazing network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3194-3203.
- [18] LI Y N, MIAO Q G, OUYANG W L, et al. LAP-Net: Level-aware progressive network for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 3275-3284.
- [19] LIU Y, PAN J S, REN J, et al. Learning deep priors for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 2492-2500.
- [20] LI R D, PAN J S, HE M, et al. Task-oriented network for image dehazing[J]. IEEE Transactions on Image Processing, 2020, 29: 6523-6534.
- [21] BAI H R, PAN J S, XIANG X G, et al. Self-guided image dehazing using progressive feature fusion[J]. IEEE Transactions on Image Processing, 2022, 31: 1217-1229.
- [22] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [23] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [24] RADFORD A, METZ L, CHINTALA S, et al. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//The 4th International Conference on Learning Representations. San Juan: International Conference on Learning Representations, 2016: 149803.
- [25] MIRZA M, OSINDERO S. Conditional generative adversarial nets. (2014-11-06)[2022-07]. <https://arxiv.org/abs/1411.1784>.
- [26] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein GAN. (2017-01-26)[2022-07]. <https://arxiv.org/abs/1701.07875>.
- [27] LI R D, PAN J S, LI Z C, et al. Single image dehazing via conditional generative adversarial network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8202-8211.
- [28] REN W Q, MA L, ZHANG J W, et al. Gated fusion network for single image dehazing[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3253-3261.
- [29] QU Y Y, CHEN Y Z, HUANG J Y, et al. Enhanced Pix2pix dehazing network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 8152-8160.
- [30] LIU X, SUGANUMA M, SUN Z, et al. Dual residual networks leveraging the potential of paired operations for image restoration[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 7000-7009.
- [31] LIU X H, MA Y R, SHI Z H, et al. GridDehazeNet: attention-based multi-scale network for image dehazing[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 7313-7322.
- [32] DONG H, PAN J S, XIANG L, et al. Multi-scale boosted dehazing network with dense feature fusion[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 2154-2164.
- [33] QIN X, WANG Z L, BAI Y C, et al. FFA-net: Feature fusion attention network for single image dehazing[C]//The Thirty-Fourth AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2020, 34(7): 11908-11915.
- [34] WU H Y, QU Y Y, LIN S H, et al. Contrastive learning for compact single image dehazing[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition.

- Nashville: IEEE, 2021: 10546-10555.
- [35] ZHENG Z R, REN W Q, CAO X C, et al. Ultra-high-definition image dehazing via multi-guided bilateral learning [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 16180-16189.
- [36] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//31st Annual Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 6000-6010.
- [37] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[EB/OL]. (2020-10-22) [2022-07]. <https://arxiv.org/abs/2010.11929>.
- [38] CHEN H T, WANG Y H, GUO T Y, et al. Pre-trained image processing transformer[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 12294-12305.
- [39] TOLSTIKHIN I, HOULSBY N, KOLESNIKOV A, et al. MLP-mixer: An all-MLP architecture for vision[EB/OL]. (2021-05-04)[2022-07]. <https://arxiv.org/abs/2105.01601>.
- [40] LIU H, DAI Z, SO D R, et al. Pay attention to MLPs[C]//Advances in Neural Information Processing Systems. Virtual Conference: NeurIPS, 2021, 34: 9204-9215.
- [41] TOUVRON H, BOJANOWSKI P, CARON M, et al. ResMLP: Feedforward networks for image classification with data-efficient training[EB/OL]. (2021-05-07) [2022-07]. <https://arxiv.org/abs/2105.03404>.
- [42] TU Z Z, TALEBI H, ZHANG H, et al. MAXIM: multi-axis MLP for image processing[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 5759-5770.
- [43] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2242-2251.
- [44] YI Z L, ZHANG H, TAN P, et al. DualGAN: Unsupervised dual learning for image-to-image translation[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2868-2876.
- [45] KIM T, CHA M, KIM H, et al. Learning to discover cross-domain relations with generative adversarial networks[C]//Proceedings of the 34th International Conference on Machine Learning. Sydney: JMLR. org, 2017: 1857-1865.
- [46] ENGIN D, GENC A, EKENEL H K. Cycle-dehaze: Enhanced CycleGAN for single image dehazing[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 938-9388.
- [47] LI Y N, LIU Y H, YAN Q X, et al. Deep dehazing network with latent ensembling architecture and adversarial learning[J]. IEEE Transactions on Image Processing, 2021, 30: 1354-1368.
- [48] LI J F, LI Y P, ZHUO L, et al. USID-Net: Unsupervised single image dehazing network via disentangled representations[J/OL]. IEEE Transactions on Multimedia. DOI: 10.1109/TMM.2022.3163554.
- [49] ZHAO S Y, ZHANG L, SHEN Y, et al. RefineDNet: A weakly supervised refinement framework for single image dehazing[J]. IEEE Transactions on Image Processing, 2021, 30: 3391-3404.
- [50] LI B Y, GOU Y B, LIU J Z, et al. Zero-shot image dehazing[J]. IEEE Transactions on Image Processing, 2020, 29: 8457-8466.
- [51] KAR A, DHARA S K, SEN D, et al. Zero-shot single image restoration through controlled perturbation of koschmieder's model[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 16200-16210.
- [52] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. (2015-03-09) [2022-07]. <https://arxiv.org/abs/1503.02531>.
- [53] HONG M, XIE Y, LI C H, et al. Distilling image dehazing with heterogeneous task imitation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 3459-3468.
- [54] JIA T Y, LI J F, ZHUO L, et al. Effective meta-attention dehazing networks for vision-based outdoor industrial systems[J]. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1511-1520.
- [55] LI L, DONG Y L, REN W Q, et al. Semi-supervised image dehazing[J]. IEEE Transactions on Image Processing, 2020, 29: 2766-2779.
- [56] SHAO Y J, LI L, REN W Q, et al. Domain adaptation for image dehazing[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 2805-2814.
- [57] ZHANG K S, LI Y N. Single image dehazing via semi-supervised domain translation and architecture search[J]. IEEE Signal Processing Letters, 2021, 28: 2127-2131.
- [58] CHEN Z Y, WANG Y C, YANG Y, et al. PSD: principled synthetic-to-real dehazing guided by physical priors [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 7176-7185.
- [59] ZHANG Y F, DING L, SHARMA G. HazeRD: An out-

door scene dataset and benchmark for single image dehazing[C]//2017 IEEE International Conference on Image Processing. Beijing: IEEE, 2017: 3205-3209.

- [60] ANCUTI C, ANCUTI C O, TIMOFTE R, et al. I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images[C]//International Conference on Advanced Concepts for Intelligent Vision Systems. Poitiers: Springer, 2018: 620-631.
- [61] ANCUTI C O, ANCUTI C, TIMOFTE R, et al. O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 867-8678.
- [62] ANCUTI C O, ANCUTI C, SBERT M, et al. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images[C]//2019 IEEE International Conference on Image Processing. Taipei: IEEE, 2019: 1014-1018.
- [63] ANCUTI C O, ANCUTI C, TIMOFTE R. NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020: 1798-1805.
- [64] LI B Y, REN W Q, FU D P, et al. Benchmarking single-image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.
- [65] WANG K, ZHUO L, LI J F, et al. Learning an enhancement convolutional neural network for multi-degraded images[J]. Sensing and Imaging, 2020, 21(1): 1-15.
- [66] ZHAO S Y, ZHANG L, HUANG S Y, et al. Dehazing evaluation: Real-world benchmark datasets, criteria, and baselines[J]. IEEE Transactions on Image Processing, 2020, 29: 6947-6962.
- [67] SILBERMAN N, HOIEM D, KOHLI P, et al. Indoor segmentation and support inference from RGBD images[C]// European Conference on Computer Vision. Florence: Springer, 2012: 746-760.
- [68] 韩翰, 卓力, 张菁, 等. 基于深度学习的无参考图像质量评价综述[J]. 测控技术, 2022, 41(4): 1-10.
HAN H, ZHUO L, ZHANG J, et al. Summary of no-reference image quality assessment based on deep learning[J]. Measurement & Control Technology, 2022, 41(4): 1-10. (in Chinese)
- [69] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain[J]. IEEE Transactions on Image Processing, 2012, 21(12): 4695-4708.
- [70] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Mak-

ing a “completely blind” image quality analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3): 209-212.

- [71] ZHU H C, LI L D, WU J J, et al. MetaIQA: Deep meta-learning for no-reference image quality assessment[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 14131-14140.
- [72] MIN X K, ZHAI G T, GU K, et al. Objective quality evaluation of dehazed images[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(8): 2879-2892.

作者简介



贾童瑶 女. 北京工业大学博士生. 研究方向为图像复原/增强.
E-mail: jty@emails.bjut.edu.cn



卓力(通讯作者) 女. 博士, 北京工业大学教授、博士生导师. 主要研究方向为图像/视频的编码与传输、多媒体大数据处理等.
E-mail: zhuoli@bjut.edu.cn



李嘉锋 男. 博士, 北京工业大学副教授. 主要研究方向为图像恢复/增强、目标检测和目标跟踪等.
E-mail: lijiafeng@bjut.edu.cn



张菁 女. 博士, 北京工业大学教授、博士生导师. 主要研究方向为图像处理、图像识别、图像检索等.
E-mail: zhj@bjut.edu.cn