

◇ “复杂视频监控场景图像分析与数据挖掘”专题

【编者按】随着当前安全形势日益严峻，利用先进的图像分析、数据挖掘等技术实现复杂视频监控场景的智能感知、智能决策等智能化分析，对提高高效、准确、实时的监控服务质量具有重要意义，亟需利用相关理论方法开展技术研究。本专题聚焦复杂视频监控场景图像分析与数据挖掘领域的前沿热点问题，围绕恶劣环境降质图像增强、多摄像机网络监控视频处理、遥感建筑物智能三维测图、视频数据深度挖掘分析、典型区域景象匹配适配、人员情绪智能分析处理等领域问题，诚邀相关基础理论研究、实际工程应用和前沿科学问题探究等方面的国内专家、学者和科研人员积极投稿。

基于在线知识蒸馏的物理感知图像去雾算法

兰云伟¹, 崔智高^{1*}, 李晓阳², 苏延召¹, 王念¹, 李爱华¹

(1. 火箭军工程大学, 陕西 西安 710025;

2. 北部战区海军参谋部 信息保障室, 山东 青岛 266077)

摘要: 为解决传统图像模型在去雾时会产生伪影、光晕等问题, 结合深度学习, 提出了一种基于在线知识蒸馏的物理感知图像去雾算法。算法首先通过物理感知单元构建的多尺度网络获取丰富的共享特征, 然后采用基于模型的方式和端到端的方式分别生成去雾图像, 最后采用特征聚合块来融合两种方式生成的去雾图像。此外, 算法采用在线知识蒸馏的学习策略来联合优化网络。实验结果表明: 与其他去雾算法相比, 所提算法在合成图像和真实场景图像上均取得了优异性能, 在合成有雾室外数据集 SOTS 上的峰值信噪比和结构相似度分别为 23.59 dB 和 0.937。

关键词: 图像去雾; 图像增强; 物理模型; 深度学习; 在线知识蒸馏

中图分类号: TJ410

文献标志码: A

文章编号: 2097-2741(2024)03-038-007

Physically-aware Image Dehazing Algorithm Based on Online Knowledge Distillation

LAN Yunwei¹, CUI Zhigao^{1*}, LI Xiaoyang², SU Yanzhao¹,
WANG Nian¹, LI Aihua¹

(1. Rocket Force University of Engineering, Xi'an 710025, Shaanxi;

2. Information Support Office, Staff Department of Navy, Northern Theater, Qingdao 266077, Shandong)

ABSTRACT: To solve problems of artifacts and halos generated by traditional image models in dehazing, a physically aware image dehazing algorithm was proposed based on online knowledge distillation and deep learning. First, shared features were acquired by a multi-scale network con-

收稿日期: 2024-04-01

基金项目: 陕西省自然科学基金面上项目 (2023-JC-YB-501)

第一作者: 兰云伟 (1995—), 男, 博士研究生, 主要从事图像处理研究。E-mail: 1223827018@qq.com

* **通信作者:** 崔智高 (1984—), 男, 副教授, 主要从事智能视频监控研究。E-mail: cuizg@126.com

引用格式: 兰云伟, 崔智高, 李晓阳, 等. 基于在线知识蒸馏的物理感知图像去雾算法[J]. 火箭军工程大学学报, 2024, 38 (3): 38-44.

structed by physically-aware units. Then, dehazed images were generated by model-based and end-to-end approaches. Finally, feature aggregation blocks were employed to fuse the dehazed images generated by the two approaches. Additionally, the learning strategy of online knowledge distillation was introduced to jointly optimize the network. Results showed that the proposed algorithm achieves excellent performance on both synthetic and real scene images than other dehazing algorithms. On the SOTS outdoor dataset, it achieves a PSNR of 23.59 dB and SSIM of 0.937.

KEYWORDS: image dehazing; image enhancement; physical model; deep learning; online knowledge distillation

雾霾会降低图像的可视性和对比度并改变其颜色, 从而导致目标检测、自动驾驶和图像理解等高级计算机视觉任务的性能下降。因此, 图像去雾技术已成为计算机视觉领域的一个重要课题。雾霾图像的效果受无雾图像、像素位置、大气光照和透射图的影响, 但由于大气光照和透射图均不确定, 所以无法直接从输入有雾图像中恢复无雾图像。为解决该问题, 通常基于先验的方法通过对清晰图像的统计规则来估计透射图和全局大气光照, 进而实现图像去雾, 如暗通道先验 (Dark Channel Prior, DCP) 图像去雾算法^[1]。

随着深度学习在计算机视觉领域的广泛应用, 一些方法通过深度学习的方式, 可以更准确地估计大气光和透射图, 进而实现图像去雾, 此类方法又称基于模型优化的方法。例如, 密集连接金字塔去雾网络 (Densely Connected Pyramid Dehazing Network, DCPDN)^[2]将大气散射模型嵌入网络, 通过联合估计大气光和透射图直接还原去雾图像。文献[3]提出一种两阶段的弱监督去雾算法 RefineDNet, 将图像去雾分解为可见性恢复和真实性改善 2 个子任务, 该算法分别使用嵌入的先验知识和深度学习来处理 2 个子任务, 从而对生成的去雾图像进行感知融合并提高算法对真实有雾图像的去雾能力。文献[4]利用生成对抗网络和大气散射模型, 提出了一种非配对的图像去雾算法。然而, 由于大气散射模型是一个理想模型, 无法完全模拟雾霾的形成过程, 所以上述算法仍然会受到图像颜色和光照变化的影响。

为解决上述问题, 研究者提出了基于端到端学习的图像去雾方法, 不需要估计上述中间参数, 而是直接建立清晰图像与雾霾输入图像之间的映射关系来实现图像去雾。文献[5]构建了一个多尺度网络 MSBDN (Multi-Scale Boosted Dehazing Network), 并利用特征融合机制将不同尺度的特征进行结合以提取全局和局部特征。文献[6]提出

一种基于生成器和多尺度判别器之间的对抗训练算法 EPDN (Enhanced Pix2pix Dehazing Network) 进行去雾。然而, 由于没有任何额外知识辅助训练, 上述基于端到端学习的图像去雾算法在真实场景中的模型泛化能力较差。因此, Shao 等^[7]首先建立双向翻译网络并提出了一种领域自适应去雾算法, 在应用于真实世界的雾霾图像时能有效去雾。Chen 等^[8]利用物理先验来指导网络训练, 提出了一种物理先验引导的图像去雾算法, 在多个真实场景中获得了高对比度的去雾结果。

与现有的去雾算法不同, 本文提出了一种基于在线知识蒸馏的物理感知图像去雾算法 (Physically-aware Image Dehazing Algorithm Based on Online Knowledge Distillation, PAIDA)。算法首先通过 2 次卷积对雾霾图像进行预处理, 并通过特征共享网络获取丰富的特征。然后, 将这些共享特征送入在线蒸馏网络, 通过监督训练实现知识蒸馏。此外, 考虑到 2 个去雾分支在图像对比度和色彩保真度方面具有互补优势, 算法提出采用特征聚合块 (Feature Aggregation Block, FAB) 来融合 2 个去雾分支的优势。在在线蒸馏网络中, FAB 充当教师网络, 2 个去雾分支充当学生网络, 实现了教师网络和学生网络的联合训练和优化。本文将这种知识蒸馏方式称为在线知识蒸馏。

1 本文方法

由于缺乏对雾霾的有效感知, 现有基于端到端学习的去雾算法无法在真实场景中有效去雾。考虑到大气散射模型可模拟雾霾的形成过程, 对网络训练具有重要的指导作用, 本文提出了一种基于在线知识蒸馏的物理感知图像去雾算法 PAIDA, 原理如图 1 所示。图中: t 为大气光; A 为生成的大气光; J 为端到端分支生成的去雾图像; J_{phy} 为基于模型分支生成的去雾图像; Cat 表示对

图像进行通道维度的拼接操作。

PAIDA 算法的整体架构可分为 3 个部分：预处理网络、特征共享网络 and 在线蒸馏网络。预处理网络由 2 个 3×3 的卷积组成，用于提取输入有雾图像的特征。在 2 次卷积操作后分别进行批量归一化和 ReLU 激活函数处理以获取预处理后的特征 F_{pre} ， F_{pre} 的大小为 $32 \times 256 \times 256$ 。

1.1 特征共享网络

为有效提取雾霾相关特征并获取初步去雾图像，算法将预处理后的特征 F_{pre} 送入特征共享网络。如图 1 所示，特征共享网络首先通过 2 个下采样块将 F_{pre} 的特征大小从 $32 \times 256 \times 256$ 下采样为 $64 \times 128 \times 128$ ，再进一步下采样为 $128 \times 64 \times 64$ ，形成不同感受野的特征，然后，通过多尺度网络提取局部特征（如结构和边缘）和全局特征（如颜色和纹理）。为了提高每个尺度的特征表示能力，算法提出物理感知单元并将其作为特征共享网络的基础结构。相邻尺度的图像特征经过网络提取后，通过拼接和 2 个上采样块恢复到前一尺度，从而得到最终的共享特征。

物理感知单元的结构如图 2 所示。首先，对送

入的图像特征 M 分别使用卷积和全局平均池化进行特征提取；然后，通过卷积、ReLU 激活函数和 Sigmoid 激活函数进行特征提取，得到估计的透射图 $\tilde{t}$ 。经过全局平均池化提取的特征同样经过卷积、ReLU 激活函数和 Sigmoid 激活函数进行特征提取，得到估计的大气光照 $\tilde{A}$ 。最后将输入特征 M 、大气光照 $\tilde{A}$ 和透射图 $\tilde{t}$ 代入大气散射模型，可得到估计的清晰图像 $\tilde{J}$ 。

1.2 在线蒸馏网络

为有效结合基于模型方式和端到端学习方式的互补优势，算法基于这 2 种方式可得到 2 幅初步去雾图像，并构建特征聚合块用于在线知识蒸馏。如图 1 所示，在线知识蒸馏网络可分为 2 部分：学生网络和教师网络。学生网络是一个多分支架构，包括 2 个去雾分支：端到端的分支和基于模型的分支。在端到端的分支中，由于特征共享网络已经有效地提取了雾霾相关特征，因此，算法使用 2 次卷积生成去雾图像 J ，其中，第 1 次卷积之后是批量归一化和 ReLU 激活函数，第 2 次卷积之后是批量归一化和 Tanh 激活函数。对于基于模型的分支，算法使用 2 次卷积生成透射图 t ，并使用

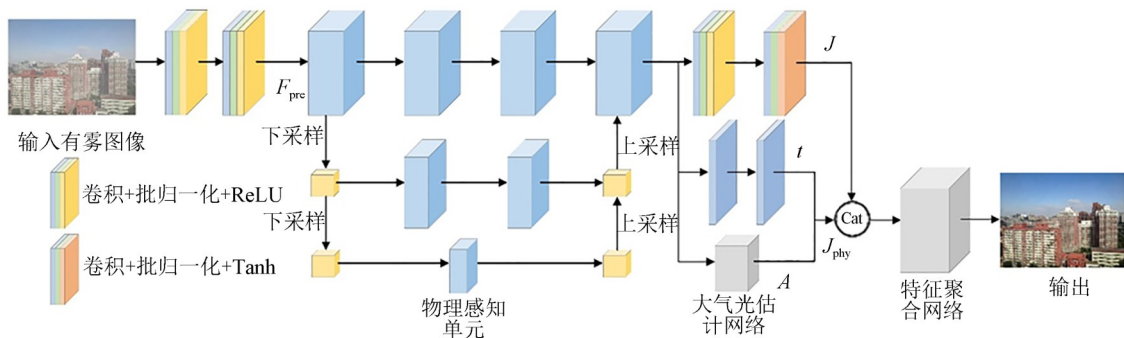


图 1 基于在线知识蒸馏的物理感知图像去雾算法原理

Fig. 1 Principles of physically-aware image dehazing algorithm based on online knowledge distillation

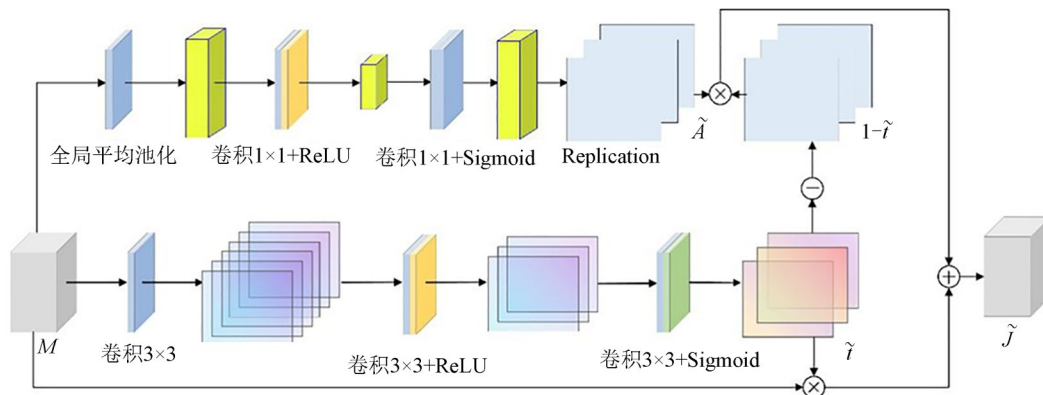


图 2 物理感知单元结构

Fig. 2 Schematic of physically-aware units

DCPDN中的大气光估计网络ANet生成大气光 A 。通过将生成的 A 和 t 代入大气散射模型, 可得到另一去雾图像 J_{phy} 。 J 和 J_{phy} 的生成方式不同, 各有优势: J_{phy} 具有良好的去雾效果, 但容易造成色彩或光照失真, 从而降低去雾图像的质量; 相比之下, J 具有更好的色彩保真度, 但局部区域存在残留雾。因此, 算法提出特征聚合块, 以融合基于模型的方式和端到端学习的互补优点。

此外, 传统的知识蒸馏方法使用预训练网络作为教师网络, 本文则将学生分支的去雾图像进行结合并联合训练, 通过在线知识蒸馏构建一个教师网络。

如图3所示, FAB由4个并行的 1×1 卷积(点卷积)和1个门控网络组成。首先, J 和 J_{phy} 通过拼接, 作为FAB的输入; 然后, 采用多个并行的点卷积和不同核大小的池化层, 同时提取局部和全局特征。点卷积加上池化核大小为 7×7 (5×5 , 3×3)的池化层, 等同于直接使用卷积核大小为 7×7 (5×5 , 3×3)的传统并行卷积。与传统并行卷积相比, 点卷积后加池化层的结构可以在不使用大卷积核的情况下减少模型参数。

此外, 算法将多尺度特征进行拼接, 并将其输入一个门控网络 (3×3 卷积), 以生成2个注意力图 α_J 和 α_{phy} 。然后, 2个注意力图通过线性组合对输入特征 (J 和 J_{phy}) 进行加权, 从而得到优化的去雾图像。该过程可表示为

$$J_{FAB} = J \otimes \alpha_J \oplus J_{phy} \otimes \alpha_{phy} \quad (1)$$

式中: J_{FAB} 表示FAB生成的去雾图像; α_J 和 α_{phy} 表示门控网络生成的注意力图; J 和 J_{phy} 表示学生分支生成的2幅去雾图像。

如上所述, FAB的输出结合了端到端学习的方式和基于模型方式的互补优势, 可以作为教师对2个去雾分支(作为学生)进行反向优化。

1.3 损失函数

将像素损失和特征损失相结合能够有效改善网络训练效果。因此, 所提算法的损失函数包含 L_1 损失、负结构相似度 (Structural Similarity Index Measure, SSIM) 损失和蒸馏损失。

像素到像素的损失可以快速匹配有雾图像和清晰图像之间的特征分布, 因此, 算法采用 L_1 损失 (标准差误差) 对网络进行训练。与均方误差损失不同, L_1 损失能够更稳定地加速网络训练, 可表示为

$$L_1 = \|G_T - J_{FAB}\|_1 \quad (2)$$

式中: G_T 表示清晰图像。

负SSIM损失可以有效地匹配亮度、对比度和结构, 从而提高清晰图像和去雾图像之间的结构相似性。因此, 算法同时采用负SSIM损失来优化网络训练, 可表示为

$$L_{SSIM} = -SSIM(J_{FAB}, G_T) \quad (3)$$

式中: L_{SSIM} 为负SSIM损失。

在PAIDA算法中, 由FAB得到的去雾图像作为教师, 对学生(2个去雾分支)进行反向优化。为此, 算法最小化教师和学生去雾图像之间距离, 并将其定义为蒸馏损失。该损失可表示为

$$L_{dist} = \|J_{FAB} - J\|_1 + \|J_{FAB} - J_{phy}\|_1 \quad (4)$$

2 实验结果

为评估所提算法在合成和真实数据集上的性能, 本文将PAIDA算法与几种先进的图像去雾算法进行定量和定性比较, 包括DCP、DCPDN、EPDN、PSD^[8] (Principled Synthetic-to-red Dehazing) 和MSBDN。除DCP是传统的基于先验信息的图像去雾算法外, 其他算法均为基于深度学

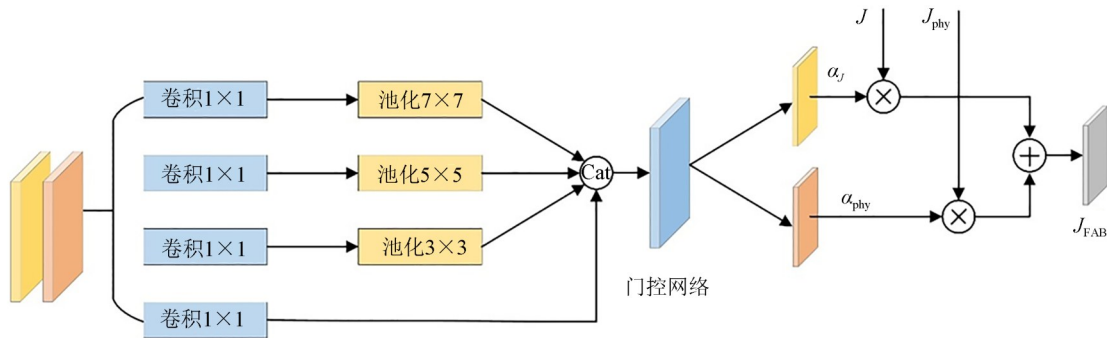


图3 特征聚合块

Fig. 3 Feature aggregation blocks

习的图像去雾算法。

2.1 实验设置

算法采用一台 PC 主机进行模型训练和测试, 其 CPU 为 Intel Core i9-7900X, 主频为 3.30 GHz, 内存为 64 GB, GPU 为一张 NVIDIA RTX 3 090 显卡。所有实验均在 Pytorch 深度学习框架下进行设计和实现。

在训练过程中, 随机裁剪训练图像为 256×256 , 批处理大小设置为 4, 训练 30 个回合; 采用衰减系数为默认值的 Adam 优化器来加速训练进程; 初始学习率设置为 0.001, 每 5 个回合降低为原来的一半。所提算法在真实单幅图像去雾数据集 (Realistic Single Image Dehazing, RESIDE)^[9] 的室内训练集 (Indoor Training Set, ITS) 上进行训练, 并在合成有雾数据集 (Synthetic Objective Testing Set, SOTS) 和真实有雾数据集 (Unannotated Real Hazy Images, URHI) 上进行测试, 以充分验证所提算法在真实有雾场景中的去雾效果和有效性。

2.2 合成图像实验结果

图 4 为本文方法与其他算法在 SOTS 室外合成数据集上的实验结果对比。

从图 4 可以看出: 基于先验信息的 DCP 算法存在色彩过饱和和亮度异常的问题, 这是因为用于估计透射图和大气光的单边假设不适用于各种场景; 基于模型优化的算法 DCPDN 可获得高对比度的结果, 但会造成光照失真; 相比之下, 基于端到端学习的算法 EPDN 能有效对图像进行去

雾处理, 但会降低去雾图像的亮度, 降低图像的视觉效果; 另一种算法 PSD 可以改善图像对比度, 但会导致严重的亮度过饱和; 只有 MSBDN 和本文所提算法能获得高质量的去雾结果, 其纹理更加清晰, 色彩保真度较好。

为进一步验证 PAIDA 在合成数据集 SOTS 上的性能, 本文采用峰值信噪比 (Peak Signal-to-noise Ratio, PSNR) 和结构相似度 (SSIM) 对其进行定量比较。实验结果如表 1 所示。

表 1 SOTS 数据集定量实验结果

Tab. 1 Quantitative experimental results on the SOTS dataset

方法	PSNR/dB		SSIM	
	室内	室外	室内	室外
DCP	19.95	20.44	0.872	0.898
DCPDN	20.85	20.08	0.875	0.896
EPDN	25.09	20.31	0.932	0.902
PSD	16.32	15.15	0.729	0.771
MSBDN	32.67	23.16	0.983	0.936
本文方法	30.68	23.59	0.988	0.937

对于室内数据集, 从表 1 可以看出: DCP 的 PSNR 和 SSIM 分别为 19.95 dB 和 0.872, 表现较差, 意味着该方法产生的伪影和色彩失真严重影响了去雾图像的质量; 大多数基于深度学习的方法均能获得较高的 PSNR 和 SSIM, 这是因为它们

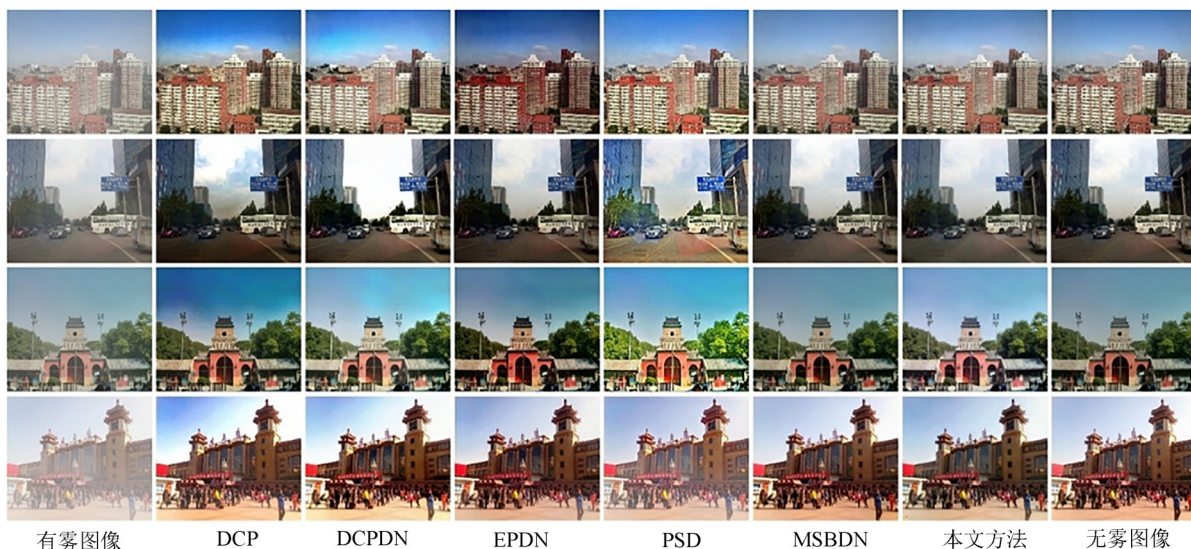


图 4 SOTS 数据集实验结果

Fig. 4 Experimental results on the SOTS dataset

均采用室内合成数据集作为训练集; 只有 PSD 在室内合成场景中表现不佳。

对于室外数据集, 从表 1 可以看出: DCP 取得了与室内场景相似的结果; 相比之下, 由于室内和室外场景的深度不同, 基于深度学习的方法在应用于室外图像时, PSNR 和 SSIM 都急剧下降; 然而, 在嵌入大气散射模型的帮助下, 所提出的算法与排名第 2 的 MSBDN 相比, PSNR 从 23.16 dB 提高到 23.59 dB, SSIM 提高了 0.001。

需要指出: 之所以 MSBDN 在室内和室外图像上均取得了较好的结果, 是因为它在 RESIDE 的室内和室外图像数据集上进行训练; 相比之下, 所提算法仅在室内图像数据集上进行训练, 但在室内和室外图像上均能取得较好的结果, 表明其泛化能力较强。

2.3 真实图像实验结果

为进一步展示所提算法的泛化能力, 本文在 URHI 数据集上的真实有雾图像上对这些算法进行结果比较。如图 5 所示。

从图 5 可以看出: 基于先验信息的算法 DCP 仍能有效地对图像进行去雾处理, 但会导致明显的色彩失真; 基于深度学习的去雾算法去雾不彻底, 如 DCPDN 会导致光照失真并在局部区域留下残留雾, MSBDN 在这些场景中的去雾性能较差, 这也表明这 2 种算法在真实场景中的泛化能力

较差; 另一种算法 PSD 也无法有效去雾, 且容易引起光照和颜色的变化; 只有 EPDN 和本文提出的 PAIDA 能有效去雾, 并能恢复大部分纹理, 然而 EPDN 存在较严重的色彩过饱和现象; 相比之下, 所提出的 PAIDA 结合了基于模型的方式和端到端学习方式的互补优势, 通过在线知识蒸馏, 获得了特征更鲜明、色彩更自然的去雾图像。

2.4 消融实验

为验证所提 PAIDA 算法的有效性, 本文进行了消融研究, 以评估以下关键模块的性能: 物理感知单元、特征聚合块、多尺度特征共享网络、基于模型的学生分支的知识蒸馏以及端到端的学生分支的知识蒸馏。

本文在实验中构建了以下变体: 变体 A, 未采用物理感知单元 (本文将其替换为残差密集块); 变体 B, 未采用特征聚合块 (本文将其替换为拼接和卷积操作); 变体 C, 未采用多尺度特征共享网络 (本文将其替换为由物理感知单元组成的单尺度网络); 变体 D, 未采用基于模型的学生分支知识蒸馏; 变体 E, 未采用端到端的学生分支的知识蒸馏; 变体 F, 本文方法。本文在 ITS 数据集上对这些变体进行了 30 回合的训练, 并在 SOTS 户外数据集上对训练后的变体进行了测试, 通过量化比较来评估每个变体的性能, 结果如表 2 所示。

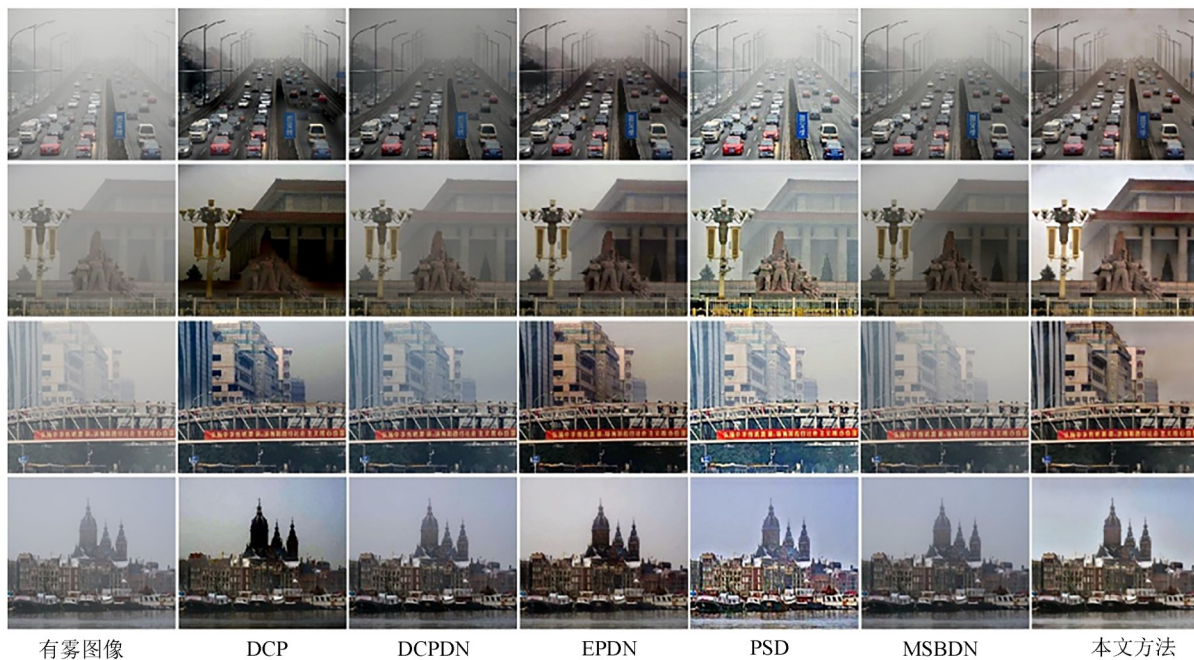


图 5 URHI 数据集实验结果

Fig. 5 Experimental results on the URHI dataset

表 2 SOTS 数据集消融实验结果

Tab. 2 Ablation experimental results on the SOTS dataset

编号	PSNR/dB	SSIM
变体 A	22.15	0.912
变体 B	23.04	0.926
变体 C	20.46	0.754
变体 D	22.68	0.913
变体 E	23.24	0.923
变体 F	23.59	0.937

从表 2 可以看出, 所提算法的 PSNR 和 SSIM 均表现优异。此外, 通过比较变体 F 和相应的变体, 可以发现上述每个模块均对提出的 PAIDA 的去雾性能做出贡献。

3 结论

本文提出了一种基于在线知识蒸馏的物理感知图像去雾算法。该算法构建了一个多分支网络, 其中: 由物理感知单元构建的多尺度网络作为共享网络, 以提取图像特征; 而基于大气散射模型的图像去雾分支和端到端的图像去雾分支分别作为学生网络, 用于生成初步的去雾图像。在此基础上, 采用有效的特征聚合块聚合 2 个初步的去雾图像, 并对学生分支网络进行反向优化, 从而得到更好的去雾结果, 并进一步提升特征聚合块的输出。所提算法通过在线知识蒸馏方式, 利用最终生成的去雾图像作为知识反向传递给每个去雾分支, 使学生分支网络能够进一步学习, 从而实现整个教师网络和学生网络的联合优化。与传统知识蒸馏方式相比, 所提算法不依赖预训练的教师模型, 可有效改善图像去雾算法的泛化能力。

参考文献:

[1] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine In-

telligence, 2011, 33(12): 2341-2353.

- [2] ZHANG H, PATEL V M. Densely connected pyramid dehazing network[C]//2018 IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2018: 3194-3203.
- [3] 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.
- [4] YANG Y, WANG C Y, LIU R S, et al. Self-augmented unpaired image dehazing via density and depth decomposition[C]//2022 IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 2027-2036.
- [5] 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 (CVPR). Piscataway, NJ, USA: IEEE, 2020: 2154-2164.
- [6] 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 (CVPR). Piscataway, NJ, USA: IEEE, 2019: 8152-8160.
- [7] 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 (CVPR). Piscataway, NJ, USA: IEEE, 2020: 2805-2814.
- [8] 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 (CVPR). Piscataway, HJ, USA: IEEE, 2021: 7176-7185.
- [9] 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.