

基于改进 U-Net 的多尺度层级融合去雾网络

季云云¹, 熊亮²

(1. 马鞍山学院 大数据与人工智能学院, 安徽 马鞍山 243100;

2. 国轩高科动力能源有限公司, 安徽 合肥 230041)

摘要: 计算机视觉任务中, 雾霾的存在会导致图像出现质量下降问题, 现有的去雾方法因全局与局部特征协同建模不足及多尺度雾浓度自适应处理欠缺, 常出现去雾残留和色彩失真。基于改进 U-Net 的多尺度层级融合去雾网络以多尺度输入捕获层级特征, 结合分层次特征融合优化信息传递。相邻层嵌入特征增强块自适应聚焦关键区域, 跨层交叉融合实现多尺度互补, 促使网络内部信息充分流动。实验表明, 该方法在合成与真实数据集上均优于对比方法, 尤其在细节恢复、色彩保真和真实场景泛化性方面表现突出。

关键词: 图像去雾; U-Net; 多尺度; 特征融合

中图分类号: TP391.4

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2025.11.006

引用格式: 季云云, 熊亮. 基于改进 U-Net 的多尺度层级融合去雾网络 [J]. 网络安全与数据治理, 2025, 44(11): 30-37.

Multi-scale hierarchical fusion dehazing network based on enhanced U-Net

Ji Yunyun¹, Xiong Liang²

(1. School of Big Data and Artificial Intelligence, Ma'an Shan University, Ma'an Shan 243100, China;

2. Gotion High-Tech Power Energy Co., Ltd., Hefei 230041, China)

Abstract: In computer vision tasks, the presence of haze leads to a degradation in image quality. Existing dehazing methods frequently suffer from residual haze and color distortion due to insufficient joint modeling of global and local features and the lack of adaptive processing for multi-scale haze concentrations. The proposed multi-scale hierarchical fusion dehazing network based on enhanced U-Net captures hierarchical features through multi-scale inputs and refines information transmission via hierarchical feature fusion. Adjacent layers embed feature enhancement blocks to adaptively focus on key regions, while cross-layer cross-fusion achieves multi-scale complementarity, facilitating full internal information flow. Experiments demonstrate that the method outperforms comparative approaches on both synthetic and real-world datasets, particularly in detail recovery, color fidelity, and generalization to real scenes.

Key words: image dehazing; U-Net; multi-scale; feature fusion

0 引言

光线在传播过程中会因大气中悬浮的微小颗粒和尘埃而发生多次散射, 导致捕获的图像质量退化, 出现模糊、对比度下降、色彩失真等问题^[1]。在计算机视觉领域中, 模型训练离不开清晰图像的支持, 尤其是涉及语义分割、目标检测等相关任务, 因此清晰图像的获取尤为重要。图像去雾旨在通过一系列处理, 将模糊的雾霾图像还原成清晰的无雾图像, 从而为相关任务提供技术支持。

图像去雾技术根据核心原理可分为三类: 基于图像

增强、基于物理模型及基于深度学习的方法^[2]。早期图像去雾技术主要依赖于图像增强, 通过调整图像的对比度或频域从而在视觉上改善清晰度, 但未能从根本上消除雾霾。例如, Wu 等人提出利用小波变换方法将图像分解为高频与低频分量^[3], 其中高频分量包含了边缘纹理等信息, 通过处理高频系数增强图像细节, 提升图像清晰度; Oakley 等人提出直方图均衡化^[4], 通过调整图像的灰度直方图分布, 增强图像对比度; Zhang 等人基于 Retinex 理论^[5], 将去雾看作是解耦反射分量和光照分量的过程, 通过分离光照干扰以恢复场景反射特性, 从而

改善图像质量。然而这类方法在处理非均匀雾霾或复杂场景时易出现色彩失真或过增强等问题。相比之下,基于物理模型的去雾方法,则是从大气光的成像原理出发,用大气散射模型描述雾霾图像的成因,如式(1)所示:

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

其中 I 表示雾霾图像, J 表示无雾图像, t 表示传输率, A 表示大气光。由于 t 和 A 是未知变量,求解无雾图像 J 则成为一个病态问题,因此这类方法往往需要引入先验知识。经典的有 He 等人^[6]提出的暗通道先验,指出无雾的清晰图像在 RGB 颜色空间上至少有一个通道上的像素值接近于 0; Berman 等人^[7]提出“雾线”先验,指出 RGB 颜色空间中颜色相似的像素点会在有雾的情况由原先的簇状聚集转变成直线; Zhu 等人^[8]提出颜色衰减先验,指出像素的亮度和饱和度的差值和雾霾浓度呈正相关。引入这些先验知识,再结合大气散射模型,可求解出无雾图像 J 。但这些先验知识往往都是依赖于特定条件或场景,通过研究大量数据总结得出,一旦场景发生改变,先验可能会失效,进而影响最终的去雾效果。如 He 等人提出的暗通道先验,当图像中存在大面积的天空区域时,去雾后会产生严重的失真和 Halo 效应^[9],如图 1 所示。随着技术的发展,研究者为了避免这一局限性,开始将深度学习应用于去雾任务,前期主要采用两段式,如 Cai 等人^[10]提出的 DehazeNet 网络,先利用卷积计算透射率 t ,再结合物理模型求解出无雾图像。尽管级联估计能够逐步细化透射率与大气光参数,但该过程易发生参数误差叠加^[11]。于是,研究者们进一步提出端到端去雾方法,如 Li 等人^[12]提出的 AOD-Uet,直接将雾图作为输入,无雾图像作为输出,利用神经网络计算二者之间的映射关系,摆脱了物理模型的束缚。



(a) 雾图 (b) 去雾结果

图 1 暗通道去雾结果

目前,基于深度学习的端到端去雾方法已经成为主流,通过数据驱动学习雾霾图像到清晰图像间的映射关系,其中基于编码器-解码器的网络结构,由于其可扩展性以及低复杂性,在去雾任务中受到广泛青睐。然而自然场景下的雾霾分布呈现显著的空间非均匀性,传统的编码器-解码器结构处理这种复杂场景时存在明显

局限性:一是编码器在下采样过程中,虽然能够提取高层语义特征,但往往会丢失大量高频细节信息,尤其是在雾霾浓度较高的远景区域,图像的纹理和边缘细节容易被过度平滑;二是解码器在上采样阶段,难以精确重建与深度相关的雾霾分布特征,尤其是在景深变化剧烈的区域,容易出现不自然的过渡效果;三是网络信息传递单向,缺乏不同层级特征的动态交互,导致全局与局部特征协同不足。这些问题最终导致重建图像出现伪影、色彩失真以及对对比度不足,从而削弱了去雾效果^[13]。

为解决上述问题,本文提出了一种基于改进 U-Net 的多尺度层级融合去雾网络 (MHF-EUN),主要贡献如下:

(1) 针对编码器下采样过程中高频细节丢失问题,基于 U-Net 改进结构,引入多尺度输入机制,通过不同尺度图像捕获层级特征,补充高频细节信息。

(2) 针对解码器上采样过程中特征重建问题,提出了特征增强块 (Feature Enhancement Blocks, FEB),通过融合高效通道注意力 (Efficient Channel Attention, ECA) 机制和像素注意力 (Pixel Attention, PA) 机制,精准聚焦景深剧烈变化区域的关键特征,避免空间信息丢失。

(3) 针对全局与局部特征协同欠缺问题,提出了多尺度特征融合模块 (Multi-scale Feature Fusion Module, MSFM),以自适应特征交互机制替代传统线性连接,促进不同层级特征动态交互,实现全局与局部特征的充分协同。

1 多尺度层级融合去雾网络

1.1 整体结构

考虑到在单一编码器-解码器结构中多次下采样会导致信息丢失,基于 U-Net 网络结构提出了一种多尺度层级融合结构,整体结构如图 2 所示。在编码阶段,对于多个不同尺度的输入图像,通过预处理模块 (Pre-processing Module, PCM) 进行浅层特征提取,随后经编码器模块 (Encoder Module, EM) 进一步提取深层特征。考虑到原始 U-Net 仅在通道维度通过跳跃连接拼接同尺度特征,缺乏跨层级的深度交互机制,难以充分挖掘多尺度上下文信息,从而削弱了对全局结构与局部细节间关联的建模能力,因此在网络结构中采用了分层次特征融合,对相邻层级间的特征,使用层级特征融合块 (Hierarchical Feature Fusion Blocks, HFFB) 进行融合处理,再经过 MSFM 对不同层级输出特征进行融合处理,实现特征在不同尺度上下文之间的信息传递,最后经解码器模块 (Decoder Module, DM) 处理后输出得到对应的清晰无雾图像。

1.2 预处理模块和层级特征融合块

为了让不同尺度的输入图像参与到网络信息传递,首

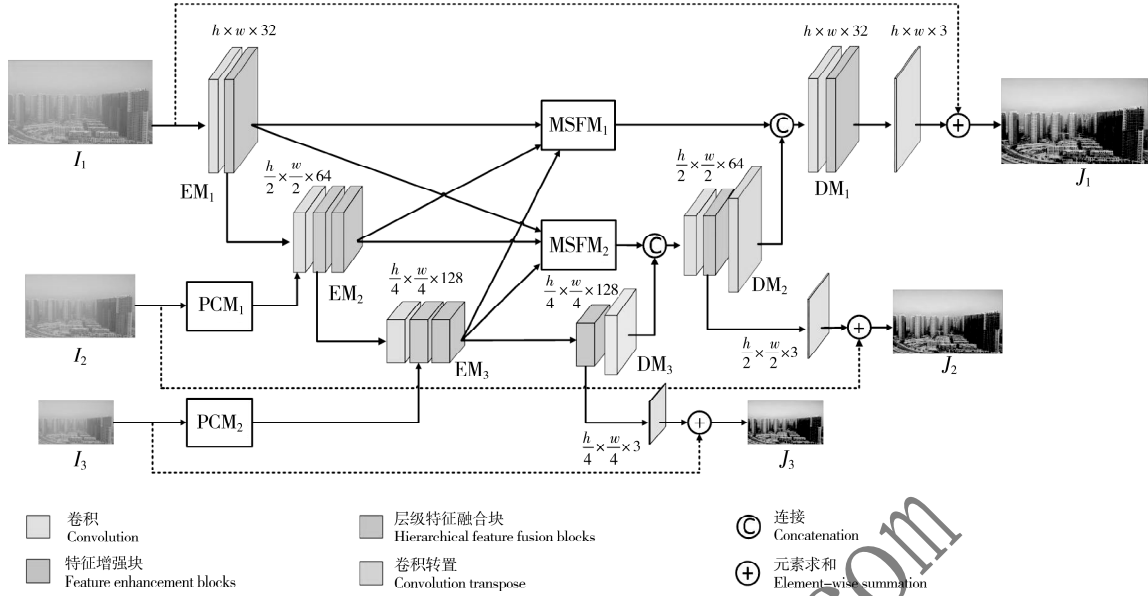


图2 网络结构图

先使用 PCM 分别对中、低分辨率图像进行浅层特征提取,其具体结构如图3所示。每个 PCM 包含 2 个 3×3 卷积层和 1×1 卷积层,进行浅层特征提取后与原始输入 I_k 拼接,最终由一个 1×1 卷积层加工处理得到输出特征图 O_k^{PCM} 。

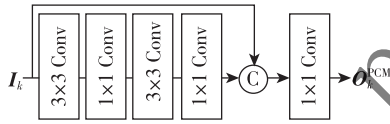


图3 PCM 结构示意图

在解码器块中,通过 HFFB 将 PCM 输出特征图 O_k^{PCM} 和相邻层级的编码器块输出结果 O_{k-1}^{EM} 进行融合处理,具体结构如图4所示。 O_{k-1}^{EM} 首先经步长为 2 的 3×3 卷积层下采样处理,使其与 O_k^{PCM} 的尺度匹配。随后将下采样后的特征与 O_k^{PCM} 逐元素相乘,结果由 3×3 卷积层进一步加工后与 O_k^{PCM} 逐元素相加,完成特征的最终融合。

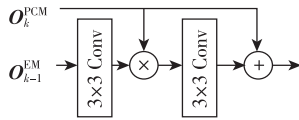


图4 HFFB 结构示意图

1.3 整体结构特征增强块

为了增强特征的表达能力,在网络的第二、三层编码块中引入了 FEB。该模块由 8 个残差注意力块级联构成,每个残差注意力块整合了 ECA 和 PA,其详细结构如

图5所示。相比于传统的通道注意力机制,ECA 使用自适应 1×1 卷积核代替全连接层,避免了传统通道注意力机制使用全连接导致的参数过大问题。此外 ECA 能够根据输入特征的通道数动态调整卷积核大小 k ,从而在不同复杂程度的通道上自适应地调整感受野范围。

首先对输入特征图 $X \in \mathbf{R}^{H \times W \times C}$ 进行全局平均池化操作,将空间维度压缩为一个标量,得到全局特征向量 $G \in \mathbf{R}^C$ 。

$$G = \text{GlobalAvgPool}(X) \quad (2)$$

然后,使用大小为 k 的一维卷积操作对全局特征向量 G 进行通道间交互建模,得到特征向量 $Z \in \mathbf{R}^C$ 。

$$k = \left\lfloor \frac{\log_2(C+b)}{\gamma} \right\rfloor_{\text{odd}} \quad (3)$$

$$Z = \text{Conv}_{1 \times 1}(G, k) \quad (4)$$

其中, $\text{Conv}_{1 \times 1}$ 表示一维卷积操作, C 是通道数, b 和 γ 为超参数, $\lfloor \cdot \rfloor_{\text{odd}}$ 表示向下取最接近的奇数。接着通过 Sigmoid 激活函数处理特征向量 Z ,生成通道注意力权重,将通道注意力权重与原始输入特征图 X 逐元素相乘,得到增强后的特征图 $X' \in \mathbf{R}^{H \times W \times C}$:

$$X' = \sigma(Z) \odot X \quad (5)$$

其中, $\odot$ 表示逐元素相乘操作, σ 表示 Sigmoid 激活函数。

针对自然雾图像中雾霾浓度随景深非均匀分布的特性,利用 PA 实现景深自适应加权,强化近景细节、恢复远景结构。通过一个卷积操作 Conv 和 Sigmoid 激活函数生成归一化的像素权重图 X^* ,其中权重图的高响应区域自动对应图像中雾浓度较高的远景区域。随后通过逐元

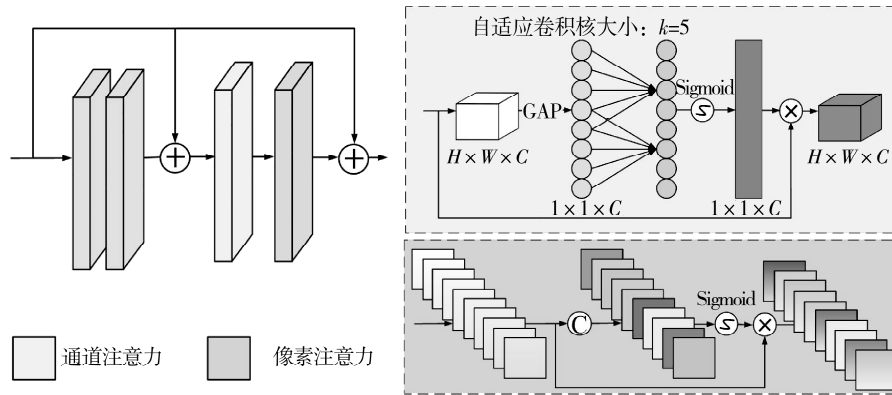


图5 残差注意力块结构示意图

素相乘, 实现特征图的景深自适应增强, 在突出高雾浓度区域特征的同时, 通过残差连接保留原始特征图的细节信息。

$$\mathbf{X}^* = \sigma(\text{Conv}(\mathbf{X}')) \quad (6)$$

$$\mathbf{X}^{\text{PA}} = \mathbf{X}' \otimes \mathbf{X}^* \quad (7)$$

1.4 多尺度特征融合模块

在传统的多尺度去雾网络中, 为了实现信息在网络中的全流动, 通常采用堆叠子网络的方式来处理不同尺度的输入图像^[14-15]。然而, 尽管这种架构能够有效提升模型性能, 但由于多个子网络的存在, 容易导致网络整体参数量过大。为了在单编码器-解码器结构中实现多尺度特征在不同层级之间的有效信息流动, 引入了MSFM, 其结构如图6所示。该模块通过融合各层级的特征图, 使得解码器能够在不同层级动态地获取和利用多尺度特征信息。在特征融合过程中, 若采用直接拼接的方式, 不仅会导致特征维度显著增加, 还难以有效利用不同特征之间的互补性。因此, 在MSFM中引入了交叉融合机制^[16]来实现不同层级特征的深度整合。

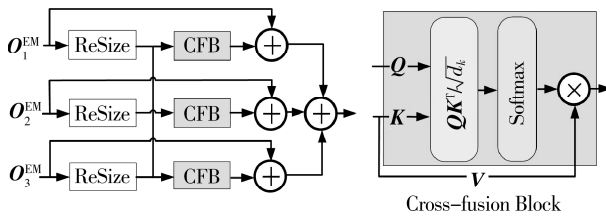


图6 MSFM 结构示意图

具体而言, 首先通过上采样或下采样的方式, 将不同层级编码器块的输出 $\mathbf{O}_k^{\text{EM}}$ 进行特征对齐, 输入到交叉融合块 (Cross-fusion Blocks, CFB) 中, 通过 1×1 卷积层分别生成查询每个 $\mathbf{O}_i^{\text{EM}}$ 对应的查询向量 $\mathbf{Q}_i$ 、键向量 $\mathbf{K}_i$ 和值向量 $\mathbf{V}_i$, 如式 (8)~式 (10) 所示:

$$\mathbf{Q}_i = \text{Conv}_{1 \times 1}^{\mathbf{Q}_i}(\mathbf{O}_i^{\text{EM}}) \quad (8)$$

$$\mathbf{K}_i = \text{Conv}_{1 \times 1}^{\mathbf{K}_i}(\mathbf{O}_i^{\text{EM}}) \quad (9)$$

$$\mathbf{V}_i = \text{Conv}_{1 \times 1}^{\mathbf{V}_i}(\mathbf{O}_i^{\text{EM}}) \quad (10)$$

随后, 计算每个 $\mathbf{O}_i^{\text{EM}}$ 与其他编码器块输出之间的注意力权重。对于 $\mathbf{O}_1^{\text{EM}}$, 使用式 (11) 和式 (12) 分别计算其与 $\mathbf{O}_2^{\text{EM}}$ 和 $\mathbf{O}_3^{\text{EM}}$ 的注意力权重 $\mathbf{A}_{12}$ 和 $\mathbf{A}_{13}$ 。

$$\mathbf{A}_{12} = \text{Softmax}\left(\frac{\mathbf{Q}_1 \cdot \mathbf{K}_2^T}{\sqrt{d_k}}\right) \quad (11)$$

$$\mathbf{A}_{13} = \text{Softmax}\left(\frac{\mathbf{Q}_1 \cdot \mathbf{K}_3^T}{\sqrt{d_k}}\right) \quad (12)$$

其中, d_k 是键向量的维度。根据式 (13), 计算 $\mathbf{O}_1^{\text{EM}}$ 对应的上下文向量 Context_1 , 并将其与 $\mathbf{O}_1^{\text{EM}}$ 相加, 得到更新后的特征表示 $\mathbf{O}_1^{\text{EM}'}$ 。

$$\text{Context}_1 = \mathbf{A}_{12} \cdot \mathbf{V}_2 + \mathbf{A}_{13} \cdot \mathbf{V}_3 \quad (13)$$

$$\mathbf{O}_1^{\text{EM}'} = \mathbf{O}_1^{\text{EM}} + \text{Context}_1 \quad (14)$$

重复上述步骤得到 $\mathbf{O}_2^{\text{EM}'}$ 和 $\mathbf{O}_3^{\text{EM}'}$ 。将更新后的 $\mathbf{O}_1^{\text{EM}'}$ 、 $\mathbf{O}_2^{\text{EM}'}$ 、 $\mathbf{O}_3^{\text{EM}'}$ 融合, 得到多尺度特征融合后的结果 $\mathbf{O}_k^{\text{MSFM}}$:

$$\mathbf{O}_k^{\text{MSFM}} = \mathbf{O}_1^{\text{EM}'} + \mathbf{O}_2^{\text{EM}'} + \mathbf{O}_3^{\text{EM}'} \quad (15)$$

将融合后的特征 $\mathbf{O}_k^{\text{MSFM}}$ 和解码器块 DM_{k+1} 上采样后的结果拼接后, 输入传递到解码器块 DM_k 中, 经解码器处理后得到输出 $\mathbf{O}_k^{\text{DM}}$, 随后将输入特征 $\mathbf{I}_k$ 与 $\mathbf{O}_k^{\text{DM}}$ 进行残差连接, 最终输出无雾图像 J_k 。

1.5 损失函数

为了在去雾任务中获得更加精细的结果, 损失函数包含了像素损失、感知损失和结构损失三种损失, 其中像素重建损失 (Pixel Reconstruction Loss, $\mathcal{L}_{\text{rec}}$) 通过 L_1 范数确保去雾图像与真实无雾图像之间的像素级一致性。这一损失函数的引入, 有助于提高模型在像素层面的准确性。同时, 为了使去雾结果更符合人类的视觉感知, 引入了感知损失 (Perceived Loss, $\mathcal{L}_{\text{per}}$), 通过评估去雾输

出与无雾图像之间的特征级差异, 确保增强的视觉真实感和自然度。此外, 为了进一步优化去雾图像的结构信息, 还引入了结构损失 (Structural Loss, $\mathcal{L}_{str}$), 通过评估去雾图像与无雾图像之间的结构相似性, 确保生成图像在结构层面上的完整性。最终损失函数如式 (19) 所示。

$$\mathcal{L}_{rec} = \lambda_{L1} \|J - J_{gt}\| \quad (16)$$

$$\mathcal{L}_{per} = \lambda_{per} \|\varphi(J) - \varphi(J_{gt})\| \quad (17)$$

$$\mathcal{L}_{str} = \lambda_{str} (1 - SSIM(J, J_{gt})) \quad (18)$$

$$\mathcal{L}_{total} = \lambda_{L1} \mathcal{L}_{rec} + \lambda_{per} \mathcal{L}_{per} + \lambda_{str} \mathcal{L}_{str} \quad (19)$$

其中, $\varphi(\cdot)$ 代表从预训练的 VGG19 网络中提取的特征, $SSIM(J, J_{gt})$ 表示去雾图像 J 和真实无雾图像 J_{gt} 之间的结构相似性指数。 λ_{L1} 、 λ_{per} 和 λ_{str} 分别是用来平衡各种损失的超参数, 实验中分别设置为 1。

2 实验结果及分析

2.1 数据集

自然雾图难以获取成对的无雾图像, 且现有的真实雾霾数据集规模有限。例如, O-HAZE 数据集^[17] 包含 45 对室外有雾图像及其对应的清晰图像, NH-HAZE^[18] 则包含了 55 对, 而 DENSE-HAZE^[19] 提供了 33 对由雾机生成的真实有雾图像及其对应的清晰图像。相比之下, RESIDE^[20] 则涵盖了 110 500 张不同雾浓度、不同场景的合成有雾图像, 能为模型训练提供丰富的样本支撑, 因此使用 RESIDE 作为训练集。测试集则选用 SOTS-outdoor 与 RTTS 数据集, 分别用于评估模型在合成和真实雾图上的性能。其中, SOTS-outdoor 数据集是来自 RESIDE 数据集的测试子集, 包含 500 对合成的室外雾图与对应的清晰

图像; RTTS (Real-World Task-Driven Testing Set) 是由 RESIDE 数据团队发布的一个包含了 100 张不同自然场景的真实雾霾图像数据集。

2.2 实施细节

实验中使用的操作系统为 Ubuntu 18.04.6 LTS, 采用 Adam 优化器以及余弦退火策略^[21] 动态调整学习率 (初始学习率为 2×10^{-4}), batch size 设置为 16, 图像裁剪为 256×256 图像块, 并以 50% 的概率随机进行水平翻转以增强泛化能力。整个流程在 PyTorch 中实现, 并在两块 NVIDIA 3090 GPU 上执行。

2.3 实验结果

为了全面评估提出的 MHF-EUN 模型性能以及鲁棒性, 分别在合成雾霾图像和真实雾霾图像上, 将其与当前几种最先进的算法进行比较, 包括 USID-Net^[22]、D4^[23]、MSBDN^[24]、Dehamer^[25] 和 C2PNet^[26]。

2.3.1 基于合成雾图

首先基于合成雾图 SOTS-outdoor 数据集展开实验, 图 7 展示了 MHF-EUN 和其他去雾算法的部分对比结果, 其中第一列为合成雾图, 第二列为对应的清晰无雾图像。为了便于观察, 标记了一些对比明显的区域。可以直观地看出, 其中 C2PNet 和 MSBDN 去雾后的结果整体清晰度较低, 尤其在远景建筑边缘的细节上表现不足, 如第二组和第五组图所示。在色彩还原方面, D4 和 Dehamer 方法存在明显缺陷, 尤其是在处理大面积天空区域时, 容易出现色彩失真问题。相比之下, USID-Net 的去雾结果与 MHF-EUN 较为接近, 但在细节清晰度上仍略逊一

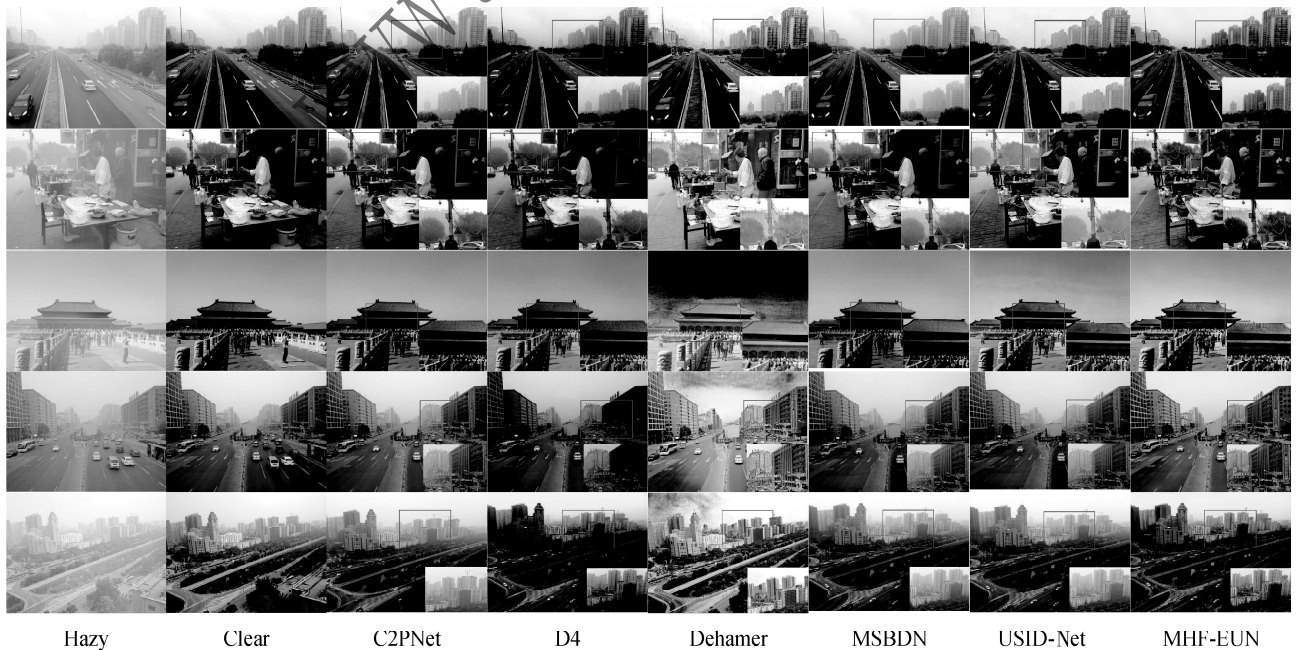


图 7 基于 SOTS-outdoor 数据集的去雾结果对比

筹，如在第三组和第五组图中，MHF-EUN 方法能够更完整地还原建筑细节，而 USID-Net 的去雾结果则稍显模糊。总体来看，无论是在近景区域还是远景区域，MHF-EUN 结果的清晰度和色彩还原方面均优于其他方法。

为了进一步验证 MHF-EUN 的去雾表现，使用一系列客观评价指标进行量化评估，包括峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）^[27]、结构相似性指数（Structural Similarity Index Measure, SSIM）^[28]和特征相似性指数（Feature Similarity Index, FSIM）^[29]，对 SOTS-outdoor 数据集上的去雾结果进行了系统性评估，具体结果如表 1 所示，其中最佳结果以粗体突出显示。

表 1 基于 SOTS-outdoor 数据集的客观质量评价结果

算法	SOTS-outdoor		
	PSNR/dB	SSIM	FSIM
C2PNet	21.75	0.826	0.976
D4	20.68	0.912	0.852
Dehamer	14.02	0.417	0.747
MSBDN	18.56	0.605	0.912
USID-Net	9.72	0.274	0.694
MHF-EUN	28.94	0.975	0.957

客观评价结果表明，MHF-EUN 在去雾性能上具有显著优势。具体而言，其 PSNR 达到 28.94 dB，远超次优方

法 C2PNet (21.75 dB) 7.19 dB，展现出更强的去雾能力；在 SSIM 指标上，MHF-EUN 以 0.975 的成绩领先 D4 (0.912) 0.063，验证了其在图像结构信息保留方面的优越性。尽管 C2PNet 在 FSIM 指标上略高，但综合考虑 PSNR 和 SSIM 结果，MHF-EUN 方法在去雾性能上仍展现出更优的综合表现。

2.3.2 基于自然雾图

在真实场景中雾霾分布往往更加复杂，局部浓雾、光照变化等因素对算法的鲁棒性提出了更高要求。为了验证 MHF-EUN 在自然场景中的泛化能力，进一步在 RTTS 真实雾图数据集上进行了实验，图 8 展示了 MHF-EUN 与其他算法的对比结果。相比之下，C2PNet 和 MSBDN 的去雾结果雾霾残留较多，D4 和 Dehamer 在处理光照度较低或天空面积较大的区域时，容易出现图像对比度、色彩失真等问题，而 USID-Net 虽然整体清晰度表现不错，但在远景部分的边缘纹理等细节保留上，相较于 MHF-EUN 仍有不足。从这些结果可以看出，MHF-EUN 方法在自然雾图上依旧表现出色。

为了进一步验证 MHF-EUN 在自然雾图上的有效性，继续采用客观质量评价指标对去雾效果进行定量评估。由于自然雾图缺乏成对的清晰无雾图像，采用 BRISQUE^[30]、NIQE^[31]和 FADE^[32]这些无参考质量评价指标，其中 FADE 专门用于量化图像中的雾霾密度。具体评估结果如表 2 所示，表现最佳的指标以粗体突出显示。

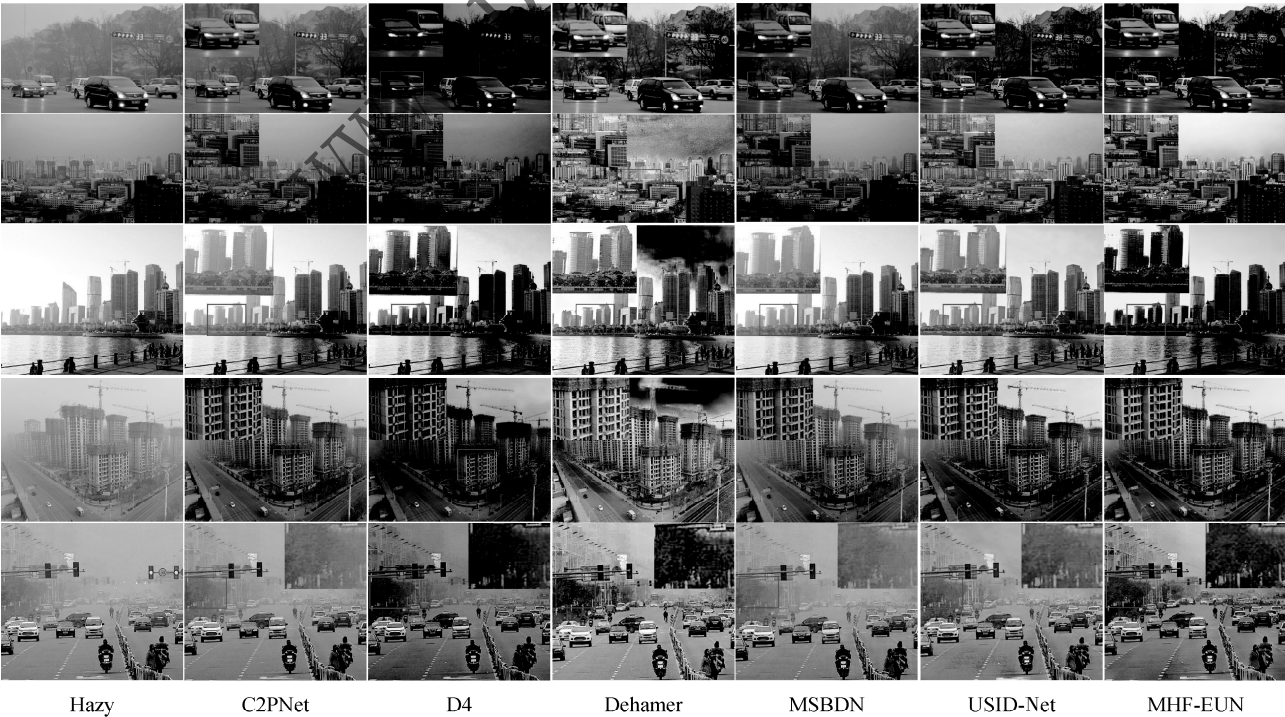


图 8 基于 RTTS 数据集的去雾结果对比

表2 基于 RTTS 数据集的客观质量评价结果

算法	RTTS		
	BRISQUE	NIQE	FADE
C2PNet	0.526	14.49	2.075
D4	0.551	14.09	0.952
Dehamer	0.510	15.28	0.998
MSBDN	0.532	14.96	1.726
USID-Net	0.497	15.14	1.468
MHF-EUN	0.481	13.27	0.865

从表2的结果可以看出, MHF-EUN 在所有指标上均取得了最佳。具体而言, 其 FADE (0.865) 值显著低于其他方法, 表明其在去除雾霾方面表现出色, 能够有效降低图像中的雾霾密度。且 BRISQUE (0.481) 和 NIQE (13.27) 值均为最低, 进一步验证了其在减少失真和提升视觉质量上的优势。相比之下, 其他方法在某些指标上表现欠佳。例如 Dehamer 的 NIQE 值 (15.28) 较高, 其去雾结果可能存在视觉伪影, 而 C2PNet 和 MSBDN 的 FADE 值分别为 2.075 和 1.726, 明显高于 MHF-EUN, 揭示了它们在去除雾霾方面的不足, 这一结果与主观视觉评估结果相符。

综上所述, MHF-EUN 在合成和自然雾图数据集上均优于对比方法。尽管自然场景中复杂的雾霾分布带来了更大挑战, 但所提出的方法在多个指标上仍保持卓越性能, 尤其在细节还原、色彩保真度和对比度恢复方面表现出色, 凸显了其强大的泛化能力和实用潜力。

3 结论

本文针对 U-Net 架构在细节保留和跨尺度融合方面的局限性, 提出了一种多尺度层级融合去雾网络。该网络首先通过 PCM 处理多分辨率输入图像, 提取不同尺度的浅层特征; 随后在 EM 中通过特征增强块, 引导模型自适应聚焦关键区域, 并采用交叉融合机制实现不同尺度特征的高效整合。此外, 为优化网络性能, 设计了融合像素重建、感知特征和结构相似性的复合损失函数。在 SOTS-outdoor 合成数据集和 RTTS 真实数据集上的实验表明, 该方法在主客观评价指标上均显著优于现有算法, 不仅能够有效恢复图像细节和色彩, 还展现出优异的泛化能力。未来研究将着重探索轻量化网络设计, 以进一步提升算法的实用性。

参考文献

[1] YUAN F, ZHOU Y, XIA X, et al. A confidence prior for image dehazing [J]. Pattern Recognition, 2021, 119 (2): 108076.
[2] 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.

[3] WU R, GONG L, LIU J, et al. A low distortion image defogging method based on histogram equalization in wavelet domain [J]. Journal of Image Processing Theory and Applications, 2020, 3 (1): 1–10.
[4] OAKLEY J P, BU H. Correction of simple contrast loss in color images [J]. IEEE Transactions on Image Processing, 2007, 16 (2): 511–522.
[5] ZHANG T, ZHU W, LI Y, et al. Improved image enhancement method based on retinex algorithm [C]//International Symposium on Artificial Intelligence and Robotics. Cham: Springer International Publishing, 2018: 291–299.
[6] HE K, SUN J, TANG X. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33 (12): 2341–2353.
[7] BERMAN D, AVIDAN S. Non-local image dehazing [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1674–1682.
[8] 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 (11): 3522–3533.
[9] LI Y, TAN R T, BROWN M S. Nighttime haze removal with glow and multiple light colors [C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 226–234.
[10] 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 (11): 5187–5198.
[11] LIU X, MA Y, SHI Z, et al. GridDehazeNet: attention-based multi-scale network for image dehazing [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 7314–7323.
[12] LI B, PENG X, WANG Z, et al. An all-in-one network for dehazing and beyond [J]. arXiv preprint arXiv: 1707.06543, 2017.
[13] KULKARNI A, MURALA S. Aerial image dehazing with attentive deformable transformers [C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023: 6305–6314.
[14] HUANG X, LI Y, POURSAEED O, et al. Stacked generative adversarial networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5077–5086.
[15] HOWARD A A, MURPHY S H, AHMED S E, et al. Stacked networks improve physics-informed training: applications to neural networks and deep operator networks [J]. arXiv preprint arXiv: 2311.06483, 2023.

- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems, 2017: 5998 – 6008.
- [17] ANCUTI C O, ANCUTI C, TIMOFTE R, et al. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018: 754 – 762.
- [18] ANCUTI C O, ANCUTI C, TIMOFTE R. NH-HAZE: an image dehazing benchmark with non-homogeneous hazy and haze-free images [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 444 – 445.
- [19] 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 (ICIP). IEEE, 2019: 1014 – 1018.
- [20] LI B, REN W, FU D, et al. Benchmarking single-image dehazing and beyond [J]. IEEE Transactions on Image Processing, 2018, 28 (1): 492 – 505.
- [21] LOSHCHILOV I, HUTTER F. SGDR: stochastic gradient descent with warm restarts [J]. arXiv preprint arXiv: 1608.03983, 2016.
- [22] LI J, LI Y, ZHUO L, et al. USID-Net: unsupervised single image dehazing network via disentangled representations [J]. IEEE Transactions on Multimedia, 2022, 25: 3587 – 3601.
- [23] YANG Y, WANG C, LIU R, et al. Self-augmented unpaired image dehazing via density and depth decomposition [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2037 – 2046.
- [24] DONG H, PAN J, XIANG L, et al. Multi-scale boosted dehazing network with dense feature fusion [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 2157 – 2167.
- [25] GUO C L, YAN Q, ANWAR S, et al. Image dehazing transformer with transmission-aware 3d position embedding [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5812 – 5820.
- [26] ZHENG Y, ZHAN J, HE S, et al. Curricular contrastive regularization for physics-aware single image dehazing [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 5785 – 5794.
- [27] SAMAJDAR T, QURAIISHI M I. Analysis and evaluation of image quality metrics [C]//Information Systems Design and Intelligent Applications: Proceedings of Second International Conference. Springer India, 2015, 2: 369 – 378.
- [28] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13 (4): 600 – 612.
- [29] ZHANG L, ZHANG L, MOU X, et al. FSIM: a feature similarity index for image quality assessment [J]. IEEE Transactions on Image Processing, 2011, 20 (8): 2378 – 2386.
- [30] 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.
- [31] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer [J]. IEEE Signal Processing Letters, 2012, 20 (3): 209 – 212.
- [32] CHOI L K, YOU J, BOVIK A C. Referenceless prediction of perceptual fog density and perceptual image defogging [J]. IEEE Transactions on Image Processing, 2015, 24 (11): 3888 – 3901.

(收稿日期: 2025 – 07 – 13)

作者简介:

季云云 (1997 –), 女, 硕士, 助教, 主要研究方向: 图像处理。

熊亮 (1997 –), 男, 硕士, 工程师, 主要研究方向: 图像处理。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com