

层级特征交互与增强感受野双分支遥感图像去雾网络

孙航^{1,3}, 方帅领¹, 但志平^{1,3}, 任东^{1,2}, 余梅^{1,2}, 孙水发^{1,3}

1. 三峡大学 计算机与信息学院, 宜昌 443002;

2. 三峡大学 湖北省农田环境监测工程技术研究中心, 宜昌 443002;

3. 三峡大学 湖北省水电工程智能视觉监测重点实验室, 宜昌 443002

摘要:近年来, 深度学习的去雾方法在图像去雾领域取得了显著的成绩。然而, 大多数基于U型网络的去雾方法将编码层特征直接传递到对应解码层, 缺乏浅层和深层特征之间的信息交互。此外, 基于非U型网络的去雾方法存在感受野受限问题, 无法有效的利用上下文信息。从而导致这些方法在场景尺度变化较大的遥感图像去雾中无法取得理想效果。为此, 本文提出了一种层级特征交互与增强感受野的双分支遥感图像去雾网络, 该方法包含层级特征交互子网和多尺度信息提取子网。其中, 层级特征交互子网利用层级特征交互融合模块, 逐层的在浅层特征中引入语义信息, 深层特征中引入空间细节信息, 从而增强编码层中不同层级特征之间的信息交互。多尺度信息提取子网利用多尺度残差空洞卷积模块, 融合不同感受野的特征, 从而获取对于遥感图像去雾至关重要的上下文信息。在两个公开数据集上的实验结果表明, 本文提出的去雾方法相比现有的9种优秀的去雾算法, 取得了最好的客观评价指标和视觉效果。

关键词:深度学习, 遥感图像去雾, 层级特征交互, 感受野, 双分支

中图分类号: TP751/P2

引用格式: 孙航, 方帅领, 但志平, 任东, 余梅, 孙水发. 2023. 层级特征交互与增强感受野双分支遥感图像去雾网络. 遥感学报, 27(12): 2831-2846

Sun H, Fang S L, Dan Z P, Ren D, Yu M and Sun S F. 2023. A two-branch remote sensing image dehazing network based on hierarchical feature interaction and enhanced receptive field. National Remote Sensing Bulletin, 27(12): 2831-2846[DOI:10.11834/jrs.20232333]

1 引言

随着遥感技术的快速发展, 遥感图像被广泛应用于气象、测绘、灾害监测与预警等领域。然而, 受雾霾等天气的影响, 悬浮在空气中的粒子, 如水蒸气、云和雾霾, 会导致遥感卫星和无人机在获取遥感图像的过程中出现图像模糊、色彩失真、对比度下降等问题, 严重影响其他高阶任务(Kim等, 2022; Li和chen, 2021; 魏秀参等, 2022; 杨胜天等, 2021; 赵伍迪等, 2021)的处理决策。因此, 对遥感图像有效去雾以及恢复图像细节信息具有重要意义。

在早期的图像去雾任务中, 大多数去雾方法基于先验知识估计大气散射模型(Cox, 1977)的中间参数, 利用大气散射模型将有雾图像恢复成清晰图像。大气散射模型可以表示为

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

式中, x 表示图像中的像素位置, $I(x)$ 表示有雾图像, $J(x)$ 表示清晰图像, A 表示全球大气光值, $t(x)$ 表示传输图。在式(1)中, 全球大气光值和传输图是未知的, 因此基于先验知识进行图像去雾的关键是准确的估计全球大气光值和传输图。

基于先验知识的去雾算法, 如He等(2011)提出暗通道先验图像去雾, 暗通道先验假设在室

收稿日期: 2022-07-06; 预印本: 2023-02-21

基金项目: 湖北省自然科学基金青年基金(编号: 2021CFB004); 国家重点研发计划(编号: 2016YFD0800902); 国家自然科学基金面上项目(编号: 61871258)

第一作者简介: 孙航, 研究方向为计算机视觉、图像去雾、图像增强。E-mail: sunhang0418@whu.edu.cn

通信作者简介: 但志平, 研究方向为计算机视觉、自然语言处理。E-mail: zp_dan@ctgu.edu.cn

外图像的非天空局部区域中,至少有一个通道包含低强度的像素值。虽然基于暗通道先验的方法取得了良好的去雾效果,但是该方法在天空区域的去雾效果并不理想。Zhu等(2014)提出颜色衰减先验,通过对有雾图像的场景深度进行建模恢复深度信息,然后利用雾信息先验估计大气光值,该方法尽可能的保留了还原后图像的颜色信息。Tan(2008)提出基于马尔可夫随机场最大化局部对比度的方法,该方法假设清晰图像的局部对比度会高于有雾图像的局部对比度,并且局部区域内的传输图接近恒定值。虽然该方法通过最大化对比度的方式能达到去雾的目的,但是建立在清晰图像的对比度高于有雾图像的对比度等假设之上,因此导致适用范围十分有限。虽然基于大气散射模型的方法取得了显著的进展,但是这些方法均依赖先验知识的准确性。由于图像场景的复杂多变,从而导致基于先验知识恢复出的清晰图像质量不够理想。

随着深度学习的发展,越来越多的学者通过设计端到端的去雾网络模型,直接将输入网络的有雾图像转换为清晰图像,而不需要估计大气散射模型的中间参数。Qu等(2019)提出增强的pix2pix去雾网络,该网络独立于大气散射模型,通过设计多分辨率生成器、多尺度判别器和一个增强网络直接将有雾图像恢复成清晰图像。Dong等(2020)提出一种完全端到端的具有融合判别器的生成对抗网络进行图像去雾,该方法将频率信息作为附加先验信息融合到判别器中,使生成器网络模型恢复的清晰图像更加真实自然。Wu等(2021)提出基于对比学习的图像去雾网络,该方法使用一种新的对比正则化损失,能够在表示空间中,将网络恢复的清晰图像与正样本GT拉近,与输入网络的有雾图像(负样本)推远,从而在特征空间进一步约束网络,提升网络的去雾性能。虽然上述基于深度学习的端到端去雾方法取得了比基于先验知识更好的去雾效果,然而,这些基于U型结构设计的网络模型直接将编码阶段的特征传递到对应的解码层,缺乏浅层特征和更深层特征之间的信息交互。此外,基于U型结构设计的网络模型,在下采样的过程中,可能会破坏对于恢复图像重要的细节信息,从而使恢复的清晰图像缺乏细节纹理以及结构等信息。Qin等(2020)提出基于特征融合注意力的图像去雾网络,该方

法基于非U型结构设计网络模型,并且利用残差学习的特性堆叠大量的通道注意力和像素注意力块,从而实现了较好的去雾效果。Li和Chen(2021)提出从粗到细的两阶段遥感图像去雾网络,该方法的核心结构由密集残差块组成,通过堆叠大量的密集残差块结构增加网络的深度。Yu等(2021)提出基于集成学习的双分支非均匀图像去雾网络,该网络下分支的核心结构由残差块组成。然而,不管是堆叠特征注意力块还是不同的残差块,基于非U型结构设计的网络模型中,在最基本的核心结构中仅能获得固定的感受野,没有充分利用对于场景尺度变化较大遥感图像至关重要的上下文信息,进而影响恢复图像的质量。

针对以上问题,本文提出了一种基于层级特征交互与增强感受野的双分支遥感图像去雾网络。该网络包含两个分支,上分支基于U型结构设计网络,由于在U型网络中,编码阶段的浅层特征分辨率更高,包含更多位置以及空间细节信息,而深层特征具有更多的场景、语义信息,但分辨率低,包含的细节信息相对较少。为了在浅层特征中引入场景和语义信息,在深层特征中引入空间细节信息。本文在上分支网络中提出层级特征交互融合模块(HFIM),该模块能够将深层特征中的语义信息嵌入到浅层特征中,将浅层特征中的空间细节信息嵌入到深层特征中,从而增强编码阶段不同层级特征之间的信息交互。然而,基于U型结构设计的网络在下采样的过程中会丢失特征信息,并且会破坏输入图像的结构信息。因此,本文设计了下分支网络对上分支网络丢失的信息进行补充,并且提供具有不同感受野的上下文信息。下分支网络基于非U型结构设计,在下分支网络中,本文提出了多尺度残差空洞卷积模块(MRDCB),该模块作为下分支网络的核心结构,能够在不改变特征图尺寸的同时增加网络的感受野,融合不同尺度的上下文信息。此外,由于下分支网络不会对特征图进行下采样,从而尽可能的保留了输入图像的结构信息。

2 相关工作

在介绍本文提出的基于层级特征交互与增强感受野的双分支遥感图像去雾网络之前,本节先回顾近几年图像去雾算法的发展过程。

2.1 基于先验的图像去雾方法

目前, 图像去雾方法大致可以分为两类: 基于先验的图像去雾方法和基于深度学习的图像去雾方法。基于先验知识的去雾方法依赖大气散射模型, 将先验知识或假设作为额外的约束以此得到大气散射模型的中间参数, 即传输图与全球大气光值, 然后利用大气散射模型将有雾图像恢复成无雾图像。He等(2011)提出暗通道先验的去雾方法, 该方法假设图像可能至少在一个颜色通道中具有极低的强度。Zhu等(2014)提出了线性颜色衰减先验去雾方法, 通过对有雾图像的場景深度进行建模恢复深度信息, 从而实现图像去雾。Tang等(2014)使用随机森林算法学习回归模型, 该方法通过提取各种与雾相关的多尺度特征并进行自动评估, 从而获取有用的特征信息进行中间参数的估计。袁非牛等(2021)提出一种两阶段特征的小样本训练算法, 通过提取对雾敏感的两阶段特征, 预测更精确的传输图, 从而实现图像去雾。基于物理模型的图像去雾方法虽然取得了良好的性能, 然而, 由于现实场景的复杂多变, 基于先验知识的去雾方法会出现伪影、颜色失真等图像退化现象。

2.2 基于深度学习的图像去雾方法

随着深度卷积神经网络的发展, 深度学习在解决图像去雾问题方面表现出显著的能力。一些基于深度学习的方法(Cai等, 2016; Ren等, 2016; Zhang和Patel, 2018; Ren等, 2018; Song等, 2018)依然依赖于大气散射模型, 这些方法使用卷积神经网络来估计大气散射模型的中间参数, 然后利用大气散射模型进行图像去雾。Cai等(2016)提出了可训练的图像去雾网络模型, 该模型的输入是有雾图像, 输出是预测的传输图, 然后将传输图应用到大气散射模型中得到清晰图像。通过学习图像与其对应的传输图之间的映射关系, Ren等(2016)提出了一种基于多尺度卷积神经网络的图像去雾方法, 该方法由预测整体透射图的粗尺度网络和局部精修的细尺度网络组成。Zhang和Pate(2018)提出密集连接金字塔去雾网络, 该方法直接将大气散射模型嵌入到网络中, 通过网络估计出传输图和全球大气光值, 进而恢复出清晰图像。虽然通过卷积神经网络估计大气散射模型中间参数的方法取得了较好的效果, 但是这些方法依赖大气散射模型中间参数的准确估计, 当

发生不准确的估计时, 去雾效果并不理想。

近年来, 研究人员提出了许多不需要依赖大气散射模型的端到端图像去雾方法, 通过设计不同的卷积神经网络, 直接将有雾图像恢复成无雾图像, 而不需要估计任何中间参数。这些方法大致可以分为U型网络(Qu等, 2019; Dong等, 2020; Wu等, 2021; Huang和Chen, 2021)和非U型网络(Qin等, 2020; Li和Chen, 2021)。在基于U型网络的去雾方法中, (Qu等, 2019)提出了增强pix2pix去雾网络, 该方法设计了多分辨率生成器和多尺度判别器, 并且在多分辨率生成器末端加入了增强模块, 增强了图像在纹理和颜色上的恢复效果。Huang和Chen(2021)提出基于条件生成对抗网络的SAR图像先验的遥感图像去雾网络, 该网络将传感器数据SAR(Synthetic Aperture Radar)图像加入到网络中, 为恢复清晰图像提供辅助信息。Wu等(2021)提出基于对比学习的图像去雾网络, 该网络不仅采用清晰图像作为正样本来指导网络训练, 而且加入了有雾图像作为负样本来进一步约束网络。虽然基于U型网络的去雾方法能够取得较好的效果, 但是这些方法直接将编码层特征传递到对应解码层, 缺乏不同层级特征之间的交互。此外, 由于基于U型网络的方法会进行多次下采样操作, 因此, 不可避免的破坏输入图像的结构信息。在非U型网络中, Qin等(2020)提出一种特征融合注意网络FFA-Net, 该方法利用残差学习的特性堆叠大量的通道注意力和像素注意力块, 取得了当时state-of-the-art去雾性能。Li和Chen(2021)设计了一个先粗后细的两阶段图像去雾网络FCTF-Net, 该方法能够更好的利用前一阶段信息, 从而恢复出更精细的清晰图像。由于非U型网络不会对网络特征图进行下采样操作, 从而避免了由于下采样带来的信息损失。然而, 大多数基于非U型网络的去雾方法中的感受野是固定的, 从而不利于提取对于遥感图像去雾至关重要的多尺度上下文信息。

3 本文方法

本节, 首先介绍本文设计的基于深度学习图像去雾方法的整体算法流程, 如图1所示。在训练阶段, 通过计算网络生成的清晰图像和GT图像的损失以及反向传播, 不断调整网络中的参数, 直到网络模型收敛。在推理阶段, 将训练阶段的模型参数加载到网络中, 然后将遥感有雾图像送入

网络中，得到清晰的遥感图像。

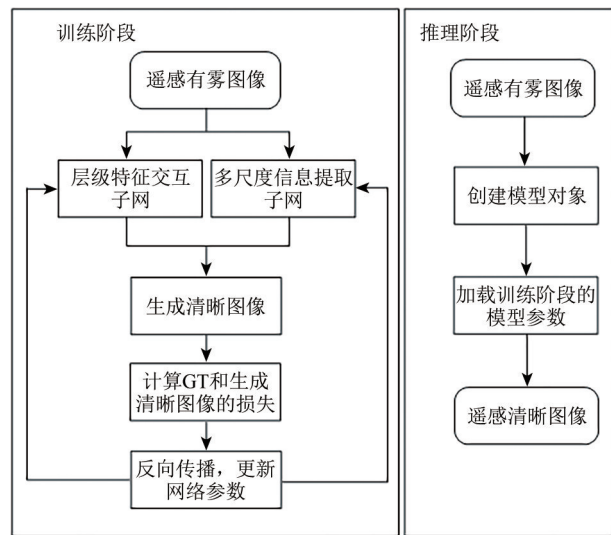


图1 算法流程图

Fig. 1 Algorithm flow chart

3.1 整体网络框架

在计算机视觉领域，有研究人员通过设计双分支网络来解决图像重压缩检测（Li 等，2017）

和图像文本匹配任务（Wang 等，2019）。通过双分支的架构，每个分支可以单独处理输入网络的图像，从而使网络能够从相同的输入中提取不同的信息。考虑到遥感图像中场景复杂多变，使用双分支网络能够提供比单分支网络更全面的信息，从而提高图像去雾的性能。因此，本文设计了一个双分支网络模型，其中包含层及特征交互子网和多尺度信息提取子网。

3.1.1 层级特征交互子网

层级特征交互子网是U型网络。在U型网络的编码层，特征图的尺寸会逐渐减小，并且提取的信息由浅层特征变成深层特征，浅层特征具有丰富的边缘纹理信息，但是缺乏更深层的语义信息，深层特征由于下采样导致含有较少的空间细节信息，但是具有丰富的场景和语义信息。因此，与传统U型网络不同，本文没有将编码层的特征图直接传递到解码层的对应层，而是将浅层空间细节信息和深层语义信息交互融合，然后再传递到对应解码层，如图2所示。

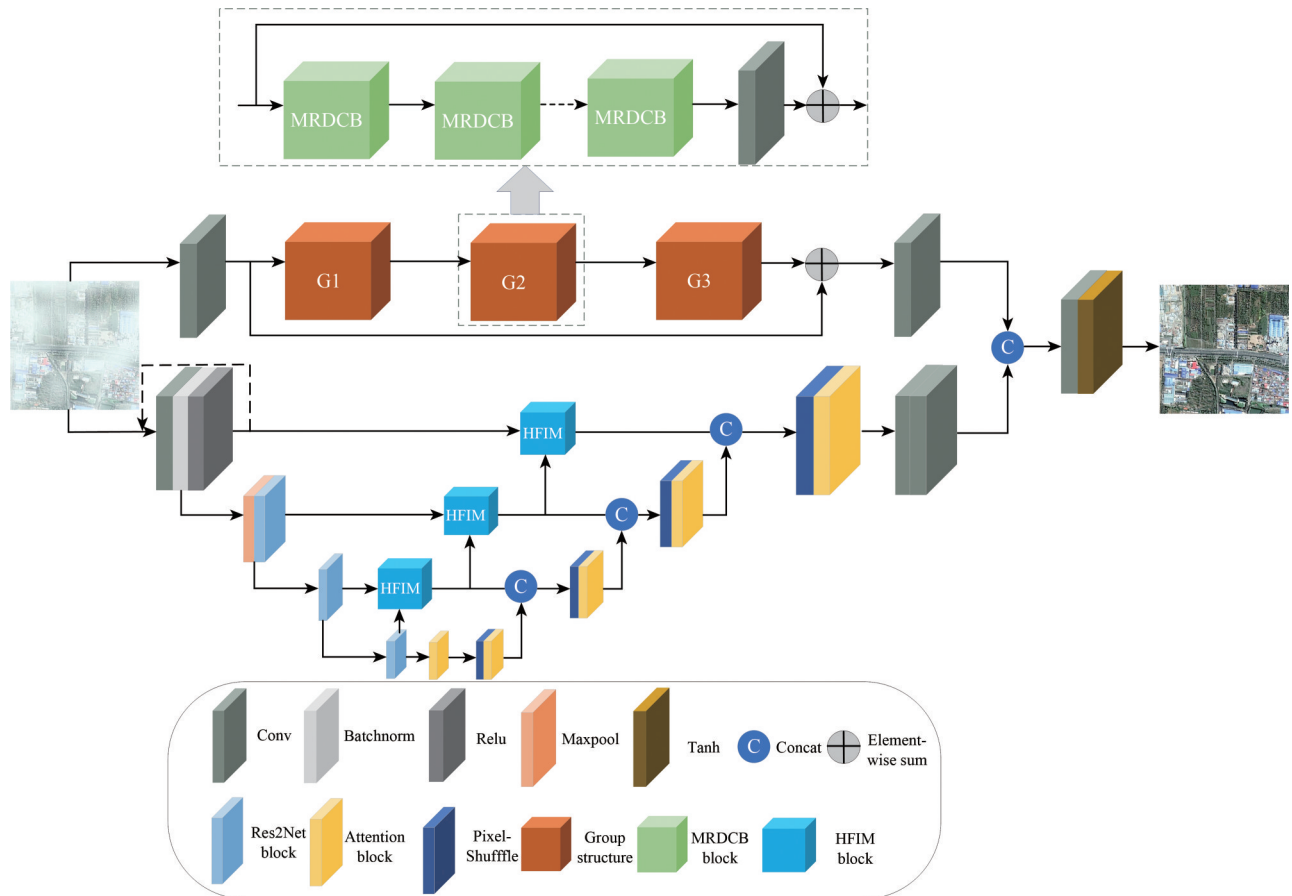


图2 整体网络框架

Fig. 2 Overall network framework

通过层级特征交互融合模块 (HFIM) 模块, 传递到解码层的特征图不仅具有当前层的空间细节信息, 而且融合了编码阶段更深层的语义信息。在融合的过程中, 本文提出了层级交互特征融合模块, 具体融合方式在 3.2 节详细介绍。此外, 受 Wu 等 (2020) 启发, 本文使用去除全连接层的 Res2Net (Gao 等, 2021) 作为编码层, 并加载其在 ImageNet (Deng 等, 2009) 数据集的预训练模型。由于 Res2Net 在 ImageNet 上的预训练模型具有鲁棒的特征表示, 因此在一定程度上能够缓解网络可能产生的过拟合问题。在 U 型网络的解码层, 本文使用 pixel-shuffle (Shi 等, 2016) 操作进行上采样, 并将特征图的大小逐渐恢复到原始分辨率。同时, 本文在 pixel-shuffle 操作之后使用注意力块 (Qin 等, 2020) 来对特征图的通道和空间信息进行校准, 从而使提出的网络在处理不同特征信息时更加灵活。

虽然层级特征交互子网融合了不同层级的特征, 提升网络的表达能力。但是 U 型网络在下采样的过程中可能会损失一些对于恢复图像很重要的细节信息。此外, 层级特征交互子网在编码阶段虽然能够获取多尺度上下文信息, 但是却是在损失一定空间信息的条件下获得的。针对以上考虑, 本文通过第二个分支多尺度信息提取子网来缓解这一问题。

3.1.2 多尺度信息提取子网

由于遥感图像中的场景尺度变化很大, 如同一张遥感图像中可能具有较大场景的建筑和较小场景的人。此外, 层级特征交互子网中多次的下采样操作, 可能导致遥感图像中较小的物体无法被重建。针对遥感图像特殊的特征结构, 本文通过多尺度信息提取子网来获取不同尺度的场景特征, 以实现多尺度上下文信息的提取, 同时在该子网中不使用下采样操作, 从而尽可能避免输入图像结构空间信息的损失。如图 1 所示, 具体的方式是在多尺度信息提取子网中引入多尺度残差空洞卷积块。多尺度残差空洞卷积块能够在不改变特征图尺寸的同时增加网络的感受野, 并且融合不同尺度的上下文信息。详细的说明在 3.3 节展示。

3.2 层级特征交互融合模块

在 U 型网络中, 大部分研究工作直接将编码层的特征传递到对应的解码层, 在浅层特征传递的过程中, 没有利用到深层特征。然而深层特征

相对浅层特征包含更多的全局信息和语义信息, 这些信息有利于在图像恢复过程中提供全局指导。因此, 本文将编码阶段的深层特征与浅层特征进行融合得到更丰富的内容特征, 然后再传递到对应的解码层, 从而使解码阶段获取更多的特征信息, 进而增强网络恢复清晰图像的能力。此外, 文献 (Zhang 和 Patel, 2018) 表明, 在浅层特征中引入更深层的语义信息, 或在深层特征中引入空间细节信息对特征融合更有效。因此, 本文提出了层级特征交互融合模块, 具体的做法是, 在 U 型网络的编码阶段, 将深层特征中的每个通道作为语义信息嵌入到浅层特征中作为指导。由于浅层特征中分辨率更高, 包含更多的空间位置信息, 因此将浅层特征中提取到的空间注意矩阵嵌入到深层特征中作为指导。通过使深层特征和浅层特征进行交互融合, 以此增强网络的表达能力, 从而恢复出高质量的清晰图像。具体做法如图 3 所示。

层级特征交互融合模块的输入可以表示为 $S = [s_1, s_2, \dots, s_n, \dots, s_N]$, $D = [d_1, d_2, \dots, d_m, \dots, d_M]$ 。式中, S 表示浅层特征, D 表示深层特征, N 与 M 表示输入特征的通道数。首先将浅层特征和深层特征经过卷积和反卷积操作, 调整其特征图的长宽为 H , W , 调整其通道数为 N 。如式 (2) 和式 (3) 所示:

$$S^* = \text{Conv}(S) \quad (2)$$

$$D^* = \text{Deconv}(D) \quad (3)$$

式中, S^* 与 D^* 均为长宽为 H , W , 通道数为 N 的特征图。为了提取深层特征中的语义信息, 本文将深层特征按照通道维度压缩, 其中每个通道代表一种语义信息。具体的做法是, 首先将深层特征经过全局平均池化函数, 如式 (4) 所示:

$$f_n = H_{\text{gap}}(d^*_n) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W d^*_n(i, j) \quad (4)$$

式中, $d^*_n(i, j)$ 表示第 n 个单通道特征图 d^*_n 在位置 (i, j) 处的值, H_{gap} 表示全局平均池化函数。特征图的形状从 $H \times W \times C$ 变成 $1 \times 1 \times C$ 。然后, 将通道向量 f 经过带有 Sigmoid 激活的简单门控机制 (Hu 等, 2020), 确保允许强调多个通道 (而不是强制执行单热激活)。整个过程可以表示为

$$\text{CRB} = \sigma \left(C_2 \left(\delta \left(C_1(f) \right) \right) \right) \quad (5)$$

式中, σ , δ 分别表示 Sigmoid 和 Relu 激活函数。 C_1 , C_2 表示两次卷积, 并且分别把特征图的通道数转换成 $N/8$ 和 N 。

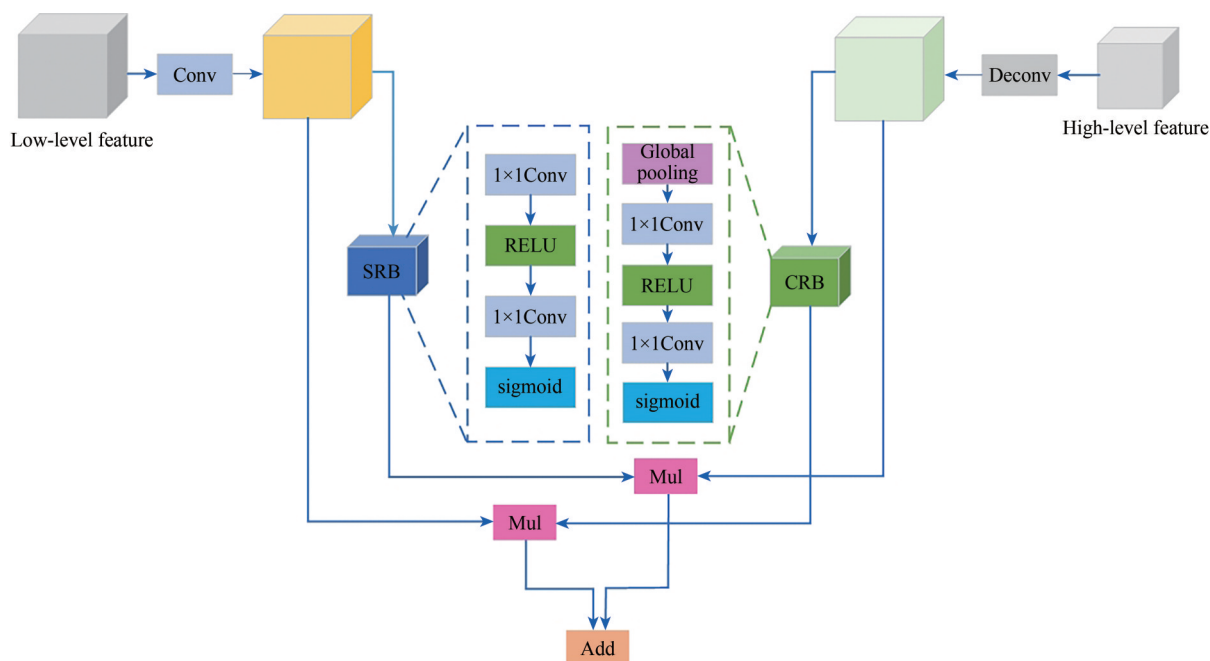


图3 层级特征交互融合模块

Fig. 3 Hierarchical feature interaction fusion module

为了提取浅层特征中的空间细节信息, 本文直接将浅层特征 S^* 输入到具有 Relu 和 Sigmoid 激活函数的两个卷积层中。如式 (6) 所示:

$$SRB = \sigma(C_4(\delta(C_3(S^*)))) \quad (6)$$

特征图的形状由 $H \times W \times C$ 变成 $1 \times H \times W$, 这是将从深层特征中提取到的语义信息嵌入到浅层特征中, 并且将浅层特征中提取到的空间细节信息嵌入到深层特征中。最后, 通过这两个加权特征进行融合。如式 (7) 所示:

$$R = SRB \otimes D^* + CRB \otimes S^* \quad (7)$$

经过上述推理和分析可知, 本文在浅层特征中引入了更深层的语义信息, 并且在深层特征中引入了空间细节信息, 然后再将浅层特征与深层特征进行融合。通过这种层级特征交互的方式, 有利于充分利用网络提取到的信息, 进而恢复出质量更好的清晰图像。

3.3 多尺度残差空洞卷积模块

在图像识别任务中, 通过堆叠残差块 (He 等, 2016), 能够构建更深层的网络。类似地, 在图像去雾任务中, FCTF (Li 和 chen, 2021) 将密集残差块作为基础模块, 通过堆叠密集残差块来构建深层次的两阶段去雾网络。FFA (Qin 等, 2020) 利用局部残差学习和特征注意力块来设计网络的基础模块, 并且取得了较好的去雾效果。然而,

这些基础模块中的感受野是固定的, 在场景变化复杂的遥感图像中, 不利于提取更多的上下文信息, 进而影响恢复图像的质量。因此, 本文设计了多尺度残差空洞卷积模块, 该模块作为多尺度信息提取子网的基础模块, 能够增强网络的感受野, 从而获取到更多对于遥感图像至关重要的上下文信息, 具体的方式如图4所示。

图4上半部分展示是 Resnet (He 等, 2016) 网络中的残差块结构, 残差块结构能够解决深层神经网络产生的退化问题, 该结构由两个 3×3 的卷积、一个激活层和一个跳跃连接组成。然而, 作为 Resnet 网络中最基本的模块, 每个残差块仅能获取当前特征图固定大小的感受野。由于遥感图像尺度场景变化大, 该残差块结构不利于遥感图像的特征提取。受 Yu 和 Koltun (2016) 启发, 空洞卷积能够在不改变特征图大小的同时, 获取特征图更大的感受野。因此, 本文在残差块结构中使用了不同空洞率的卷积, 如图4下半部分所示: $d(1,1)$ 表示在残差块中使用两个 3×3 的卷积, 每个卷积的空洞率都为1; $d(1,2)$ 表示在残差块中使用两个 3×3 的卷积, 第1个卷积的空洞率为1, 第2个卷积的空洞率为2; $d(1,2,3)$ 表示在残差块中使用3个的卷积, 第1个卷积的空洞率为1, 第2个卷积的空洞率为2, 第3个卷积的空洞率为3。不同残差块获取的感受野如图5所示。

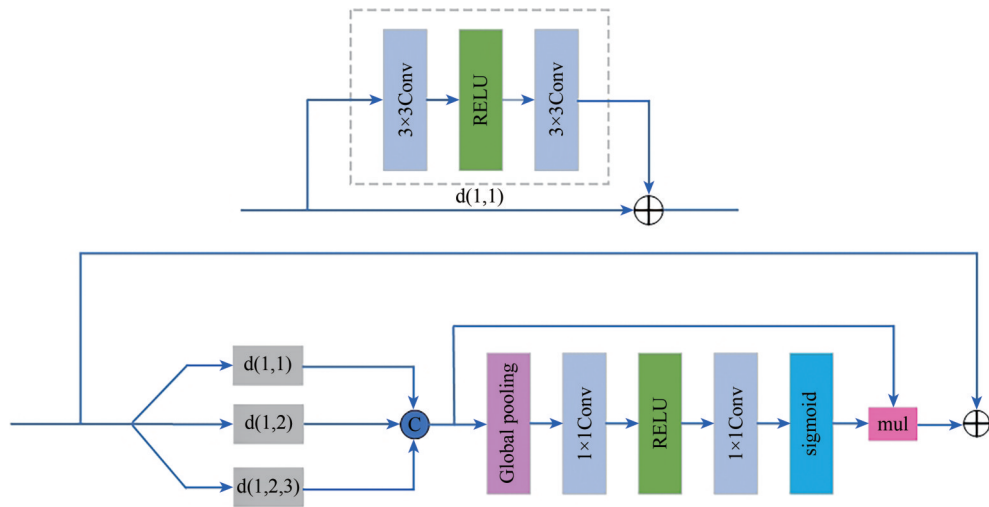


图4 多尺度残差空洞卷积模块

Fig. 4 Multiscale residual dilated convolution block

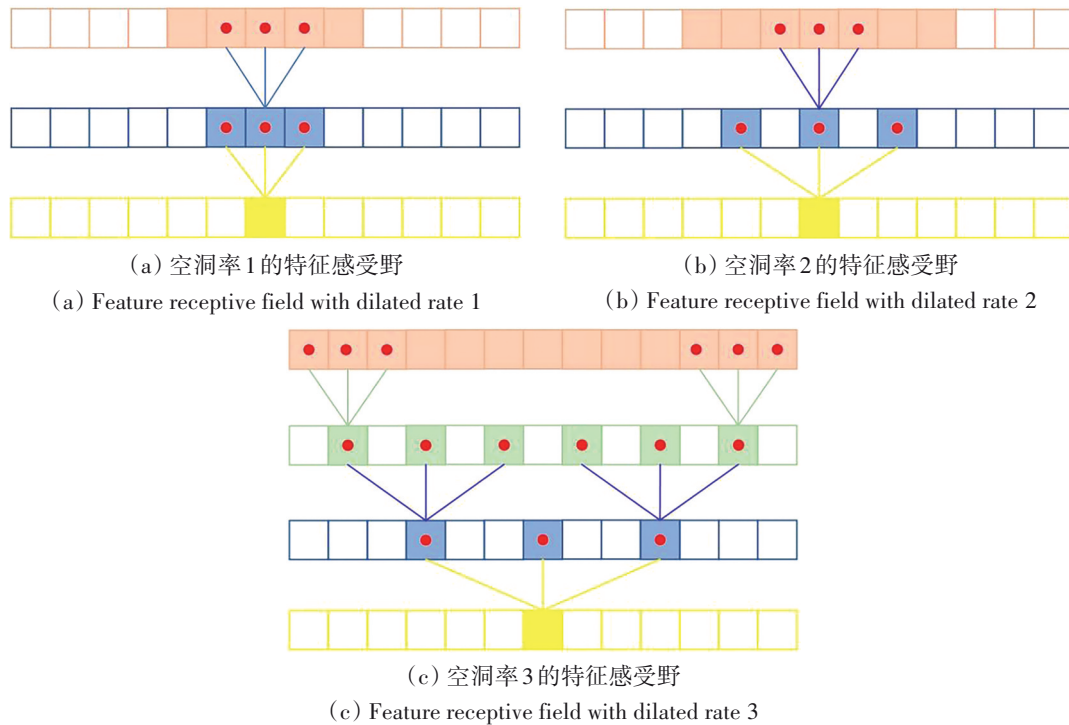


图5 不同残差块结构的感受野

Fig. 5 Sensory fields for different residual block structures

图5中黄色方块表示输出特征图中的某个点，红点表示卷积核。在5(a)中，黄色方块由上层特征图通过卷积核大小为3，空洞率为1的卷积得到，如图5(a)中的蓝色方块。蓝色方块同样由更上层特征图经过卷积核大小为3，空洞率为1的卷积得到，如图5(a)中的粉色方块。因此，在 $d(1,1)$ 残差块中，输出特征图的1个点对应输入特征图的5个点，即 $d(1,1)$ 残差块能够获取的感受野为5。在5(b)中，黄色方块由上层特征图通过卷

积核大小为3，空洞率为2的卷积得到，如图5(b)中的蓝色方块。蓝色方块同样由更上层特征图经过卷积核大小为3，空洞率为1的卷积得到，如图5(b)中的粉色方块。因此， $d(1,2)$ 残差块能够获取的感受野为7。为了进一步获取具有更大感受野的特征图，本文在 $d(1,2)$ 的基础上，又加入了一个卷积核大小为3，空洞率为3的卷积层，如图5(c)所示， $d(1,2,3)$ 残差块能够获取的感受野为13。

通过多尺度残差空洞卷积模块, 网络在基本的核心结构中能够获取具有不同感受野大小的特征图, 然后将不同感受野的特征图进行融合, 融合后的特征图具有更加丰富的多尺度上下文信息, 有利于在场景尺度变化较大的有雾遥感图像中恢复出清晰的无雾图像。此外, 本文在多尺度残差空洞卷积模块中加入了通道注意力模块 (Qin 等, 2020), 该注意力模块能够有助于减少信息冗余。

3.4 损失函数

为了训练所提出的网络, 本文使用平滑 L1 损失、感知损失、MS-SSIM 损失以及对抗损失作为网络的优化目标。其中, 将有雾图像表示为 I^{hazy} , 其对应的清晰图像表示为 I^{st} , 本文方法去雾后的图像表示为 I^* 。

由于平滑 L1 损失可以防止潜在的梯度爆炸, 且对异常值的敏感性低于 MSE 损失。因此, 本文使用平滑 L1 损失来约束网络, 如下所示:

$$L_{smooth} - L1 = \frac{1}{3N} \sum_{i=1}^N \sum_{c=1}^3 \alpha(I_c^*(i) - I_c^{st}(i)) \quad (8)$$

$$\alpha(e) = \begin{cases} 0.5e^2, & \text{if } |e| < 1 \\ |e| - 0.5, & \text{其他} \end{cases} \quad (9)$$

式中, $I_c^*(i)$ 与 $I_c^{st}(i)$ 分别表示去雾后图像和清晰图像第 c 个通道上像素点 i 的值。 N 表示每个通道中像素的个数。

Engin 等 (2018) 表明, 使用感知损失恢复的图像保留了更多的纹理信息。具体的做法是, 计算清晰图像和去雾后图像在 VGG16 (Simonyan 和 Zisserman, 2015) 预训练模型不同层的 L2 范数, 感知损失能够计算清晰图像和去雾后图像在不同特征空间的相似性。本文的感知一致损失定义如下:

$$L_p = \frac{1}{N} \sum_j \frac{1}{C_j H_j W_j} \|\varphi_j(I^*) - \varphi_j(I^{st})\| \quad (10)$$

式中, $\varphi_j(\cdot)$ 表示获取 VGG16 预训练模型中第 j 层的长度为 H_j , 宽度为 W_j , 通道数为 C_j 的特征图。表示在感知损失函数中使用 VGG16 预训练模型特征层的数量。在本文中, 使用 VGG16 预训练模型的第 3、5、15 层进行计算感知损失。

为了使去雾后的图像和清晰图像具有结构相似性, 本文采用了多尺度结构相似度损失 (MS-SSIM) 来约束网络。如式 (4) 所示:

$$L_{ms} - ssim = 1 - \prod_{m=1}^M \left(\frac{2\mu_p\mu_g + C_1}{\mu_p^2 + \mu_g^2 + C_1} \right)^{\beta_m} \left(\frac{2\sigma_{pg} + C_2}{\sigma_p^2 + \sigma_g^2 + C_2} \right)^{\gamma_m} \quad (11)$$

式中, μ_p , μ_g 分别表示去雾后图像和清晰图像的均值。 σ_p , σ_g 分别表示去雾后图像和清晰图像的标准差。 σ_{pg} 表示去雾后图像和清晰图像之间的协方差。 β_m , γ_m 表示两项之间的相对重要性, C_1 , C_2 为常数项, 防止除数为 0。

此外, 在恢复清晰图像过程中, 对抗损失被证明是有效的 (Ledig 等, 2017), 在本文使用的对抗损失如式 (5) 所示:

$$L_{adv} = \sum_{n=1}^N -\log D(I^*) \quad (12)$$

式中, $D(I^*)$ 表示将去雾后的图像 I^* 判断为清晰图像的概率。

整体的损失函数如式 (6) 所示:

$$L_{total} = L_{smooth} - L1 + \alpha_{ms-ssim} + \beta L_p + \gamma L_{adv} \quad (13)$$

式中, α , β , γ 为每个损失函数的超参数, 在本文中, 将 α , β , γ 分别设置为 $\alpha = 0.2$, $\beta = 0.001$, $\gamma = 0.005$ 。

4 实验

本节主要展示所提出的遥感图像去雾方法和目前 9 种优秀的去雾方法, 在公开遥感数据集上的定量以及定性的评估结果。对比的去雾方法包括: DCP (He 等, 2011), EPDN (Qu 等, 2019), (Huang 等, 2020), FFA (Qin 等, 2020), FCTF (Li 和 Chen, 等, 2021), DCRD (Huang 和 Chen, 2021), SkyGAN (Mehta 等, 2021), AECD (Wu 等, 2021), (Yu 等, 2021)。本节剩余部分将介绍实验设置、数据集、评估指标以及定量和定性的评价结果。最后, 通过消融实验验证各个模块的作用。

4.1 实验设置及数据集

实验设置: 本文代码是基于 Pytorch 框架实现, 在 Ubuntu 环境下使用 NVIDIA 2080Ti GPU 训练网络模型。采用 Adam 优化器进行优化网络, 其动量衰减指数 $\beta_1 = 0.9$, $\beta_2 = 0.999$ 。初始学习率设置为 0.001, 使用多步长衰减方式调整学习率, Gamma=0.5。在训练过程中, 为了对数据集进行增强, 本文将图像随机裁剪成 256×256 的尺寸, 然后进行 90° , 180° , 270° 的随机旋转、水平翻转、垂直翻转后再送入到网络中进行训练。

本文在公开遥感数据集 SateHaze1k (Huang 等, 2020) 和 RICE (Lin 等, 2019) 上评估各种去雾方法。在 SateHaze1k 数据集中, 包含 3 个子数据集, 每个子数据集中图像的雾浓度不同。Thin 子数据集中图像的雾浓度为薄雾, Moderate 子数据集中图像的雾浓度为中等雾, Thick 子数据集中图像的雾浓度为浓雾。在每个子数据集中, 训练集包含 320 张图像, 验证集包含 35 张图像, 测试集包含 45 张图像。牟范 (2021) 指出云和雾颗粒类型和粒子半径相同, 只是浓度范围有所不同。RICE 数据集是通过卫星拍摄的图像, 通过观察数据集中的有云图像可以发现云和雾极为相似。因此, 孙潇和徐金东 (2021) 将 RICE 数据集做为去雾对待。因此, 本文也将 RICE 数据集作为去雾数据集。RICE 数据集包含两个子数据集, RICE1 子数据集包含 500 对训练样本, 每对样本包含一张有云图像和一张无云图像。该子数据集由谷歌 Earth 收集, 通过设置云层是否显示, 来获得有云/无云图像。RICE2 子数据集是由 Landsat 8 OLI_TIRS 数据制作而成, 并且使用了具有地理参考的 LandsatLook 图像。LandsatLook 图像包括自然彩色图像和高质量图像。在 RICE2 子数据集中包含 736 组分辨率为 512×512 的图像, 每组包含一张有云图像、一张无云图像和一张云掩模图像。在 RICE 数据集中, 没有明确指出训练集和测试集, 为了验证各个方法的有效性, 本文将 RICE 数据集按照 9:1 的比例划分成训练集和测试集。此外, 由于两个数据集中的训练集均包含较少的图像, 因此在网络训练过程中, 本文使用数据增强的方式对数据集进行了扩充。

评价指标: 为了评估各种方法的性能, 本文采用峰值信噪比 (PSNR) 和结构相似性指数 (SSIM) 作为定量评估指标, 这两个指标通常用于图像去雾任务中评估算法恢复图像质量。此外, 本文还根据主观感受在 4.3 节定性分析展示了各种算法去雾后的可视化效果。

4.2 定量分析

本文提出的去雾方法和 9 种优秀的去雾方法定量评估如表 1 和表 2 所示, 如图表展示了所有评估去雾算法分别在 PSNR 和 SSIM 上的测试结果, 第 1 名为黑体, 第 2 名为斜体, 第 3 名为粗斜体。实验结果表明, 本文提出的方法在 Haze1k 数据集以及

RICE 数据集均取得了优异的性能, 并且在 PSNR 和 SSIM 方面取得了最好的效果。其中, 与基于 U 型结构设计网络模型的算法 EPDN、Huang、AECR 相比, 本文在 Thin fog 子数据集上的 PSNR 分别提高了 5.622 dB, 3.199 dB, 4.457 dB, SSIM 分别提高了 0.0464, 0.0224, 0.0507。与基于非 U 型结构设计网络模型的算法 FFA, FCTF 相比, 本文在 Moderate fog 子数据集上的 PSNR 分别提高了 1.668 dB, 3.977 dB, SSIM 分别提高了 0.0068, 0.0191。与近年来提出的遥感图像去雾算法 Huang, FCTF, DCRD, SkyGAN 相比, 本文在 Thick fog 子数据集上的 PSNR 分别提高了 0.064 dB, 5.588 dB, 6.104 dB, 1.707 dB, SSIM 分别提高了 0.0311, 0.0932, 0.095, 0.0026。与同样基于双分支设计网络模型的去雾方法 TBN 相比, 本文在 RICE2 子数据集上的 PSNR 提高了 0.824 dB, SSIM 提高了 0.0014。此外, 在对比算法中, (Huang 等, 2020) 提出在去雾网络中引入了雷达数据作为辅助信息来指导网络恢复出清晰图像, 并且取得了较好的性能提升。然而, 雷达数据的获取相对较难, 并且, 本文的方法在没有雷达数据的情况下也取得了比 (Huang 等, 2020) 更好的定量指标。

4.3 定性分析

本节在 Haze1k 和 RICE 的子数据集上对各种去雾方法进行定性分析。图 6 展示的是各去雾算法在 Thin fog 子测试集上去雾后的可视化视觉效果。如图所示, DCP 去雾方法恢复的清晰图像存在严重的曝光问题, 如图 6 (b) 中的第 1 张和第 2 张图片中的大部分区域。EPDN 去雾方法虽然没有出现图像曝光问题, 但是去雾后的图像仍有部分雾残留, 如图 6 (c) 第 3 张图像中的红色框区域。FFA 去雾方法的去雾能力较 EPDN 有所提升, 然而其细节恢复能力不如本文提出的方法, 如图 6 (d) 第 2 张图片中的红色框区域, FFA 去雾方法恢复的图像在该区域偏白, 而本文去雾方法恢复的图像在该区域更接近真实图像。FCTF 和 DCRD 是最近提出的针对遥感图像的去雾方法, 虽然其能够去除大部分的雾, 然而对于不同场景深度其去雾能力不同, 如对于图 6 第 3 张图片中的较大的草地区域取得了不错的去雾效果, 对于图 6 第 4 张图片红框区域中的较小的房屋, 其去雾能力相对较差。AECR 去雾方法取得了较好的去雾效果, 然而其恢复的图像

和真实有雾图像之间存在一定的色差，如图6（g）第2张和第3张图片中的绿色草地部分。TBN去雾方法在去雾后图像中，与清晰图像相比依然有少量雾残留，并且会产生色调偏离，如图6（h）第4张图片中的红色区域。本文去雾方法如图6（i）所

示，由于层级特征交互模块和多尺度残差空洞卷积模块的作用，相比其他7种去雾方法，本文去雾方法在达到去雾效果的同时，在主观视觉上如颜色、饱和度和清晰度方面最接近清晰图像。

表 1 各种算法在Haze1k数据集上的定量比较
Table 1 Quantitative comparison of various algorithms on the Haze1k dataset

方法	Thin fog		Moderate fog		Thick fog	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DCP(TPAMI, 2011)	13.450	0.7015	9.782	0.5916	10.896	0.5720
EPDN(CVPR, 2019)	21.740	0.8821	24.057	0.9103	19.731	0.7860
Huang.w/o SAR	21.474	0.8168	22.094	0.8274	22.120	0.7842
Huang.SAR(WACV, 2020)	24.163	0.9061	25.311	0.9264	25.073	0.8640
FFA(AAAI, 2020)	24.836	0.9200	26.503	0.9415	22.404	0.8487
FCTF(GSRL, 2021)	20.798	0.8802	24.194	0.9292	19.549	0.8019
DCRD(ICIP, 2021)	20.966	0.8769	24.714	0.9318	19.033	0.8001
SkyGAN(WACV, 2021)	25.381	0.9248	25.583	0.9035	23.430	0.8925
AECR(CVPR, 2021)	22.905	0.8778	24.035	0.8794	20.612	0.8122
TBN(CVPR, 2021)	25.001	0.9127	27.426	0.9447	22.230	0.8431
本文	27.362	0.9285	28.171	0.9483	25.137	0.8951

注：加黑、斜体、粗斜体分别代表对应列中的第1名,第2名和第3名。

表 2 各种算法在RICE数据集上的定量比较
Table 2 Quantitative comparison of various algorithms on the RICE dataset

方法	RICE1		RICE2		方法	RICE1		RICE2	
	PSNR	SSIM	PSNR	SSIM		PSNR	SSIM	PSNR	SSIM
DCP(TPAMI, 2011)	12.396	0.6286	12.476	0.5356	DCRD(ICIP, 2021)	32.310	0.9512	30.298	0.8890
EPDN(CVPR, 2019)	29.950	0.9549	27.479	0.8888	AECR(CVPR, 2021)	34.910	0.9507	33.759	0.8770
FFA(AAAI, 2020)	36.443	0.9633	34.820	0.8968	TBN(CVPR, 2021)	35.954	0.9679	34.543	0.9139
FCTF(GSRL, 2021)	32.063	0.9415	30.617	0.8595	本文	37.792	0.9684	35.367	0.9153

注：加黑、斜体、粗斜体分别代表对应列中的第1名,第2名和第3名。

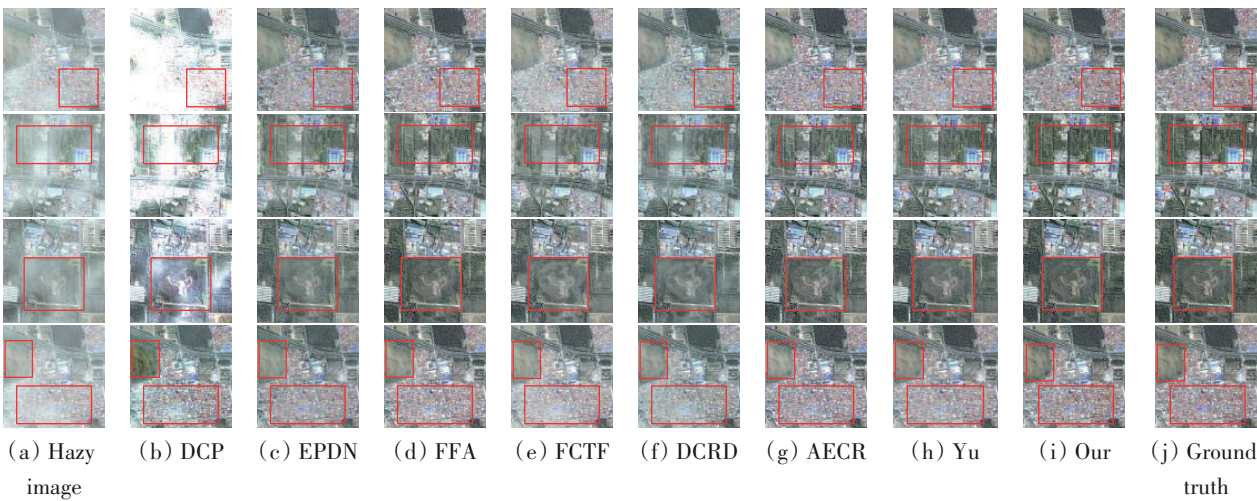


图 6 Thin fog测试集上可视化视觉结果
Fig. 6 Visualize visual results on the Thin fog subtest set

图7展示的是各去雾算法在 Moderation fog 子测试集上去雾后的可视化视觉效果。如图所示, 由于DCP去雾方法基于先验知识, 其固有的缺点是过于依赖先验知识, 而容易忽略图像本身的信息, 从而导致去雾效果大大降低, 如图7(b)中的图像出现了较为严重的曝光问题。EPDN去雾方法在中等雾测试集上, 取得了较好的去雾效果, 然而仍有少量雾残留, 如图7(c)第1张和第2张图像中的红色框区域。FFA去雾方法在雾分布均匀的原始有雾图像中, 其表现能力较雾分布不均匀的图像表现更好, 如图7前两张原始图像中雾分布相对均匀, FFA去雾方法仅有少量雾残留, 而对于后两张雾分布不均匀图像, 其去雾后的图像有大量雾

残留。FCTF和DCRD去雾方法的问题与FFA去雾方法相似, 但是, 其在雾分布不均匀图像的去雾能力相对FFA更好。AECR去雾方法在雾分布不均匀图像中相对FCTF和DCRD方法去雾能力进一步提升。TBN去雾方法在网络中同样使用了res2net在imagenet数据集上的预训练模型, 由于引入了额外的知识信息, 因此在雾分布不均匀的区域取得了不错的效果, 然而相对本文提出的方法, 在 Moderation fog 子测试集的表现不如本文方法。虽然本文方法恢复的清晰图像在视觉效果上与原图像更为接近, 同时不可否认的是, 对于雾浓度较大的区域, 本文方法恢复的图像质量仍有提升空间。

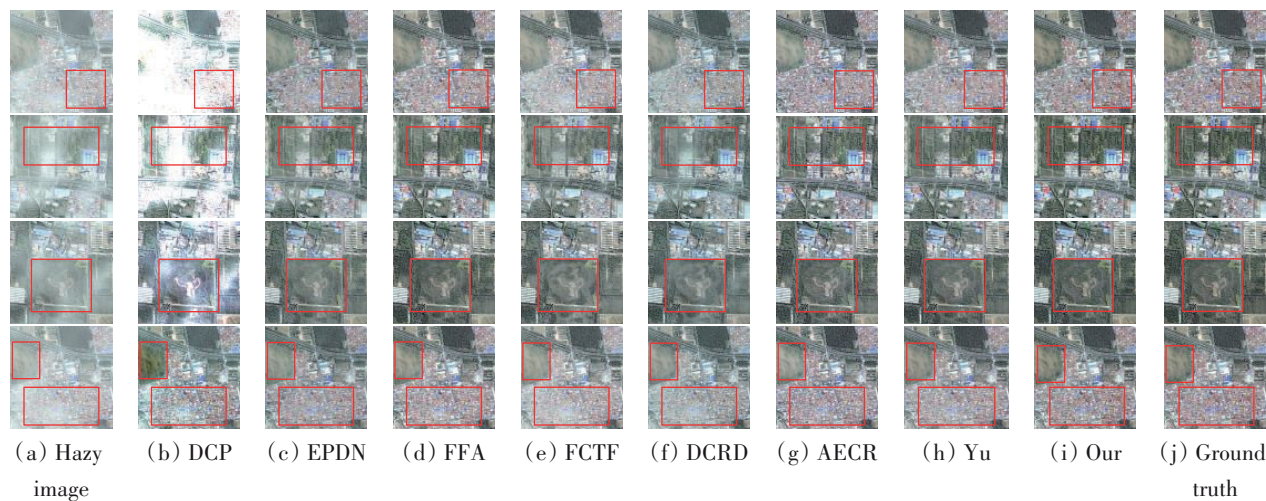


图7 Moderation fog测试集上可视化视觉结果

Fig. 7 Visualize visual results on the Moderation fog subtest set

图8展示的是各去雾算法在Thick fog子测试集上去雾后的可视化视觉效果。如图8所示, DCP去雾方法不可避免的在恢复后的图像上存在失真严重问题, 并且大部分区域出现了曝光现象。EPDN去雾方法能够去除大部分的雾, 但是部分区域去除的并不干净, 如图8(c)第3张图像中的红色框区域。FFA去雾方法在浓雾测试集上的表现要优于EPDN, 在图8相同位置的红色框区域, FFA去雾方法去除了更多的雾。FCTF和DCTD去雾方法在浓雾测试集去雾效果不佳, 如图8中第4张图像, 其恢复的图像缺少纹理细节信息, 并且恢复后的图像颜色整体偏白。AECR和TBN去雾方法的去雾能力相当, 但是TBN去雾方法在色彩还原与细节修复上要略优于AECR去雾方法, 如图8第1张和第3张对应位置的红色框区域。由图8(i)可以

看出, 本文提出的去雾方法可以有效去雾的同时, 保留了更多的细节特征, 特别是在色彩还原上相比其他去雾方法更为出色。

图9展示的是各去雾算法在RICE1子测试集上去雾后的可视化视觉效果。如图9所示, DCP去雾方法去雾后的图像丢失了大量内容信息。与Haze1k数据集类似, EPDN去雾方法在RICE1子测试集中恢复的图像依然有部分雾残留。如图9(c)第1张和第3张图像中的红色框区域。FFA在RICE1子测试集中的表现较好, 但是其在部分区域去雾后色调偏白, 如图9(d)第3张图像的右半部分区域。FCTF和DCRD去雾方法去雾后的图像在颜色和清晰图像具有一定偏差, 整体颜色更暗, 如图9对应的第2张图像和第3张图像中的红色框区域。AECR去雾方法具有良好的去雾效果, 然而其去雾

后的图像在部分区域容易引入杂色,如图9(g)第1张图像中的红色框区域。TBN去雾方法恢复的图像虽然在颜色上和清晰图像很相近,然而其去雾并不彻底,并且依然有少量雾残留,如图9(h)

第3张图像中红色框区域。与上述方法相比,本文提出的方法在色彩恢复上更好,并且有效去雾的同时,保留了更多的场景细节信息。

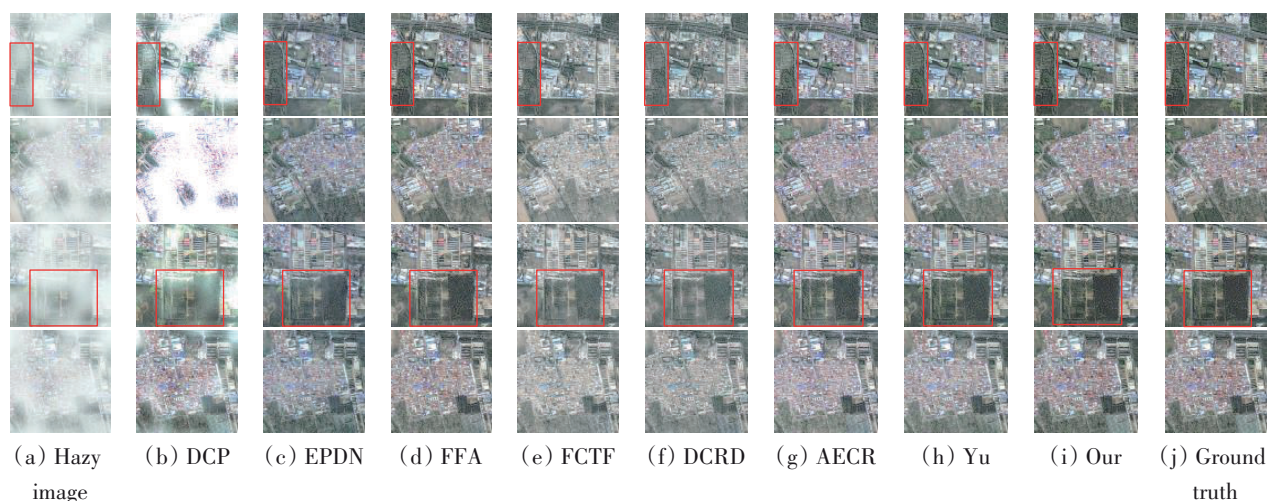


图8 Thick fog子测试集上可视化视觉结果

Fig. 8 Visualize visual results on the Thick fog subtest set

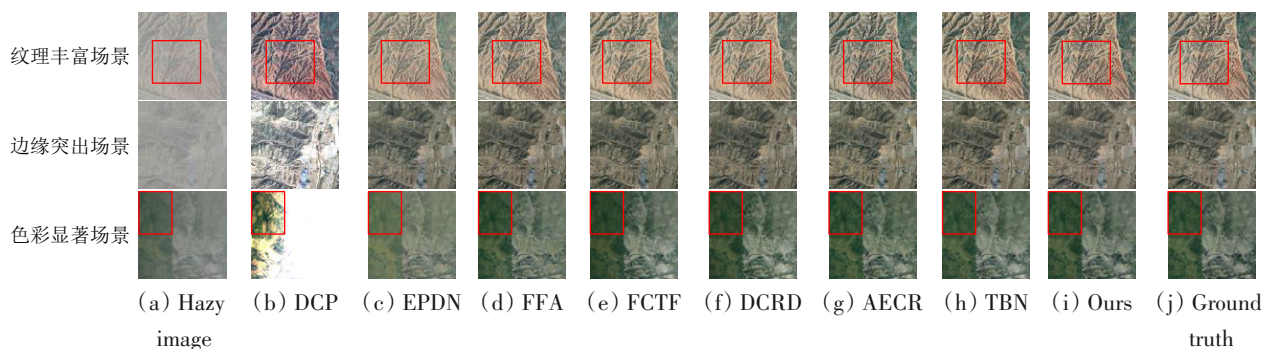


图9 RICE1子测试集上可视化视觉结果

Fig. 9 Visualize visual results on the RICE1 subtest set

图10展示的是各去雾算法在RICE2子测试集上去雾后的可视化视觉效果。如图10所示,DCP去雾方法恢复的图像存在严重的颜色失真问题。EPDN去雾方法去除了少量的雾,并且恢复的图像偏暗,如图10(c)中第2张图像。FFA取得了较好的去雾效果,但是恢复的图像缺乏纹理细节,如图10(d)中第2张图像缺少河流,而清晰图像中有明显的河流区域。FCTF和DCRD去雾方法在颜色鲜艳的图像中容易引入其他噪声,如图10中对应的第一行图像,并且恢复的图像在部分区域存在伪影,如图10中对应的第三行图像。AECR去雾方法对于图10第2张和第3张图像恢复的较好,然而对于第1张图像恢复的图像整体偏暗,并

且几乎没有纹理细节信息。TBN去雾方法在RICE2子测试集取得了较其他方法更为理想的效果,但是和本文提出的方法相比,本文提出的方法在图10第2行中的图像和清晰图像更接近,并且TBN去雾方法在对应的位置产生了伪影。

4.4 消融实验

为了验证提出的去雾方法中各个模块的有效性,本文按照网络结构和模块创新的思路设计了消融实验,共包含6个实验:(1) OUB (Ordinary Up Branch) 表示仅包含上分支的网络模型,其中上分支中仅采用普通的残差块结构。(2) ODB (Ordinary Down Branch) 表示仅包含下分支的网络模型,其中下分支为普通的U型网络,没有加入层

级特征融合模块。(3) OTB (Ordinary Two Branch) 表示双分支网络, 其中上分支网络采用普通的残差块结构, 下分支网络中采用普通的 U 型网络。(4) OTB_MRDCB (ordinary two branch, Multiscale residual dilated convolution block) 表示在 (3) 的基础上加入多尺度残差空洞卷积模块。(5) OTB_HFIM (Ordinary Two Branch, Hierarchical Feature

Interaction block) 表示在 (3) 的基础上加入层级特征交互模块。(6) OTB_MRDCB_HFIM (Ordinary Two Branch, Multiscale Residual Dilated Convolution Block, Hierarchical Feature Interaction Block) 表示在 (3) 的基础上既加入了多尺度残差空洞卷积模块, 又加入了层级特征交互模块, 该模型也是本文使用的最终模型。

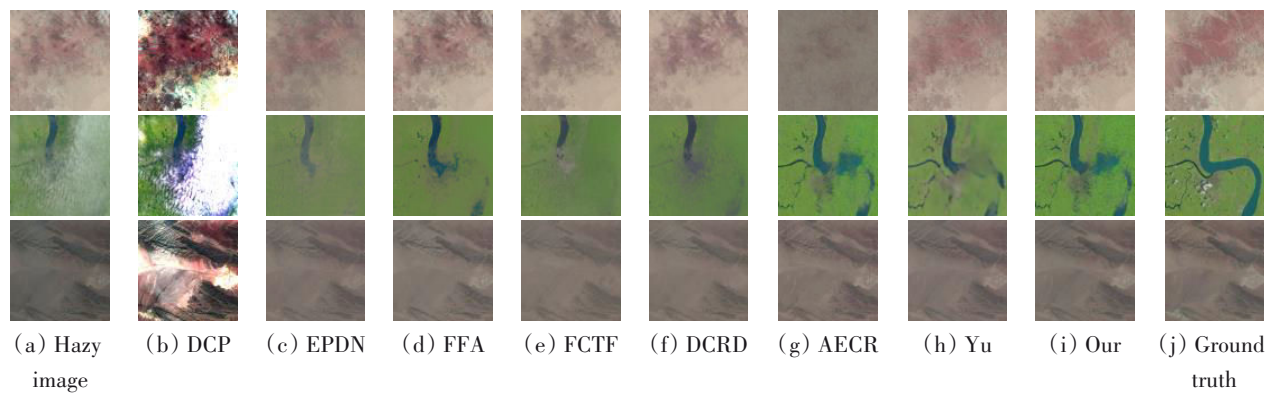


图 10 RICE2子测试集上可视化视觉结果

Fig. 10 Visualize visual results on the RICE2 subtest set

消融实验是在 Haze1k 的 Thick 子测试集上进行, 本文对比了上述 6 种实验, PSNR 和 SSIM 的结果如表 3 所示。通过比较表 3 前 3 个模型可以发现, 双分支网络可以在一定程度上能够提高模型在 PSNR 和 SSIM 方面的性能。此外, 通过观察第 3 个模型和第 4 个模型, 加入多尺度空洞卷积模块的模型在 PSNR 和 SSIM 上分别较未加入的模型提高了 1.501 dB 和 0.0064。这是由于多尺度空洞卷积模块在不改变输入图像结构信息的同时, 增强了网络的感受野, 并且融合了上下文信息。此外, 为了验证层级特征融合模块的有效性, 本文在普通的双分支网络中加入了层级特征融合模块, 通过对比表 3 的第 3 个模型和第 5 个模型, 可以发现, 层级特征融合模块在改善 PSNR 和 SSIM 方面起着重要作用, 增加的 PSNR 和 SSIM 也表明, 将编码阶段的深层语义信息逐步和浅层细节纹理信息融合对于提高去雾性能起着重要作用。

图 11 展示了在消融实验中各模型的可视化效果。如图 11 所示, 仅有上分支网络或下分支网络的模型虽然能够去除大部分的雾, 但仍有少量雾残留。而双分支网络较单分支网络能够进一步减少雾残留, 如图 11 对应的红色框区域。双分支网

络模型恢复的图像包含更少的雾, 并且细节恢复上更好, 如图 11 第二行中的红色框区域。此外, 加入多尺度空洞卷积模块的模型去雾效果进一步提升, 并且在颜色恢复上较双分支网络更好, 如图 11 第三行图像中的红色框区域。在双分支网络中加入层级特征融合模块的模型同样具有较好的细节恢复能力, 如图 11 (f) 第二行图像中的红色框区域, 然而, 该模型在局部区域的去雾能力略有不足, 如图 11 (f) 第一行图像中的红色框区域的右上方区域。因此, 本文在双分支网络模型的基础上加入了多尺度空洞卷积模块和层级特征融合模块, 并且也取得了最好的视觉效果, 本文方法去雾后的图像与清晰图像更接近。

表 3 在 Thick 子数据集上的 PSNR 和 SSIM 结果

Table 3 PSNR and SSIM results on the Thick subdataset

模型	PSNR	SSIM
OUB	22.243	0.8423
ODB	22.325	0.8465
OTB	22.489	0.8493
OTB_MRDCB	23.990	0.8557
OTB_HFIM	24.253	0.8574
OTB_MRDCB_HFIM	25.137	0.8951

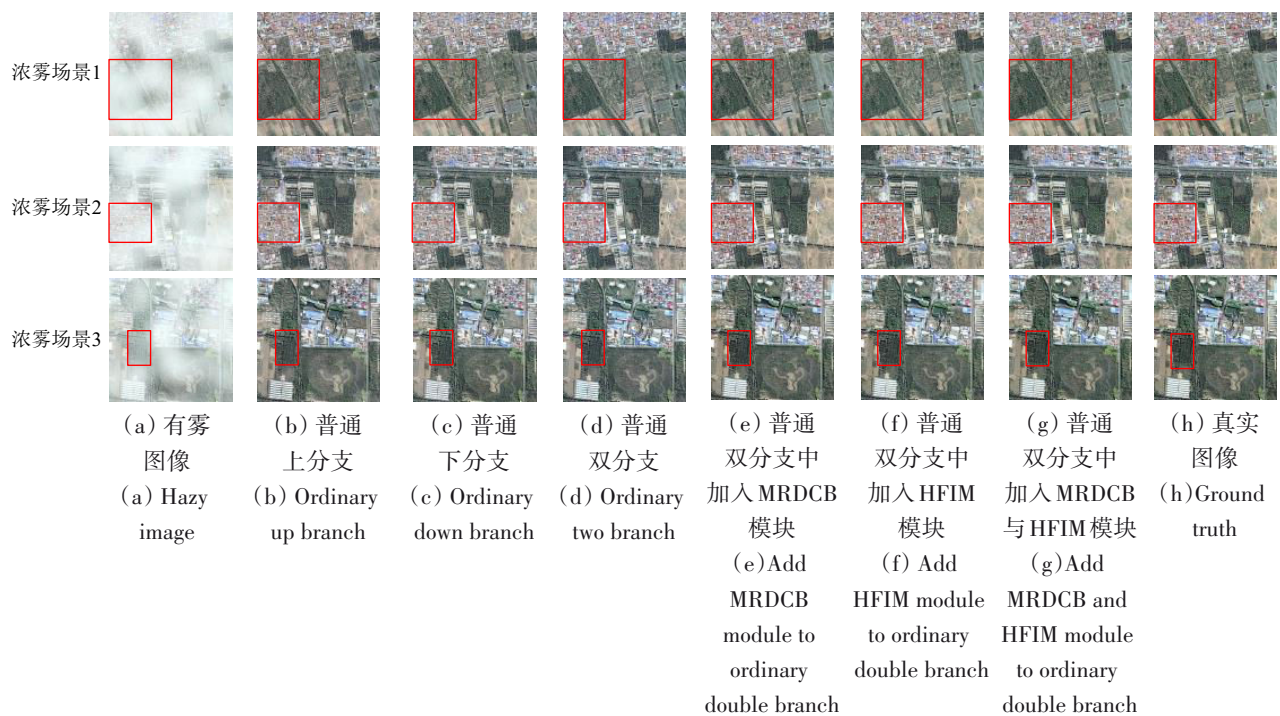


图 11 消融实验展示图

Fig. 11 Ablation experiment demonstration diagram

5 结 论

本文提出了一种基于层级特征交互与增强感受野的双分支遥感图像去雾网络，该网络主要包含层级特征交互子网与多尺度信息提取子网。其中，层级特征交互子网利用层级特征交互融合模块（HFIM），逐层将深层特征中的每个通道作为语义信息嵌入到浅层特征中作为指导，同时将浅层特征中提取到的空间注意矩阵嵌入到深层特征中作为指导。通过深层特征和浅层特征的交互融合，以此增强网络的表达能力，从而恢复出高质量的清晰图像。多尺度信息提取子网利用多尺度残差空洞卷积模块（MRDCB），获取具有不同感受野大小的特征图，然后将感受野不同的特征图进行融合，融合后的特征图具有更加丰富的多尺度上下文信息。与现有9种优秀的去雾方法相比，本文去雾方法在两种公开数据集中均取得了最佳性能。同时不可否认的是，本文去雾方法在浓雾上仍有一定的提升空间，未来，将重点研究浓雾图像去雾和恢复。

参考文献 (References)

Cai B L, Xu X M, Jia K, Qing C and Tao D C. 2016. DehazeNet: an end-to-end system for single image haze removal. IEEE Transac-

tions on Image Processing, 25(11): 5187-5198 [DOI: 10.1109/TIP.2016.2598681]

Cox L J. 1977. Optics of the atmosphere-scattering by molecules and particles. Optica Acta: International Journal of Optics, 24(7): 779-779 [DOI: 10.1080/713819629]

Deng J, Dong W, Socher R, Li L J, Li K and Fei-Fei L. 2009. ImageNet: a large-scale hierarchical image database//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]

Dong Y, Liu Y H, Zhang H, Chen S F and Qiao Y. 2020. FD-GAN: generative adversarial networks with fusion-discriminator for single image dehazing. Proceedings of the AAAI Conference on Artificial Intelligence, 34(7): 10729-10736 [DOI: 10.1609/aaai.v34i07.6701]

Engin D, Gen A and Ekenel H K. 2018. Cycle-Dehaze: enhanced CycleGAN for single image dehazing//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE: 825-833 [DOI: 10.1109/CVPRW.2018.00127]

Gao S H, Cheng M M, Zhao K, Zhang X Y, Yang M H and Torr P. 2021. Res2Net: a new multi-scale backbone architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(2): 652-662 [DOI: 10.1109/TPAMI.2019.2938758]

He K M, Sun J and Tang X O. 2011. Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(12): 2341-2353 [DOI: 10.1109/TPAMI.2010.168]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Hu J, Shen L, Albanie S, Sun G and Wu E H. 2020. Squeeze-and-

- excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(8): 2011-2023 [DOI: 10.1109/TPAMI.2019.2913372]
- Huang B H, Li Z, Yang C, Sun F C and Song Y X. 2020. Single satellite optical imagery dehazing using SAR image prior based on conditional generative adversarial networks//2020 IEEE Winter Conference on Applications of Computer Vision. Snowmass: IEEE: 1795-1802 [DOI: 10.1109/WACV45572.2020.9093471]
- Huang Y F and Chen X. 2021. Single remote sensing image dehazing using a dual-step cascaded residual dense network//2021 IEEE International Conference on Image Processing. Anchorage: IEEE: 3852-3856 [DOI: 10.1109/ICIP42928.2021.9506603]
- Kim U H, Kim S H and Kim J H. 2022. SimVODIS: simultaneous visual odometry, object detection, and instance segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1): 428-441 [DOI: 10.1109/TPAMI.2020.3007546]
- Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z H and Shi W Z. 2017. Photo-realistic single image super-resolution using a generative adversarial network//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 105-114 [DOI: 10.1109/CVPR.2017.19]
- Li B, Luo H, Zhang H X, Tan S Q and Ji Z Z. 2017. A multi-branch convolutional neural network for detecting double JPEG compression. *arXiv:1710.05477* [DOI: 10.48550/arXiv.1710.05477]
- Li Y F and Chen X. 2021. A coarse-to-fine two-stage attentive network for haze removal of remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 18(10): 1751-1755 [DOI: 10.1109/LGRS.2020.3006533]
- Li Z L, Wang L Y, Jiang S, Wu Y H and Zhang Q J. 2021. On orbit extraction method of ship target in SAR images based on ultra-lightweight network. *National Remote Sensing Bulletin*, 25(3): 765-775 (李宗凌, 汪路元, 蒋帅, 吴雨航, 张庆君. 2021. 超轻量网络的SAR图像舰船目标在轨提取. *遥感学报*, 25(3): 765-775) [DOI: 10.11834/jrs.20210160]
- Lin D Y, Xu G L, Wang X K, Wang Y, Sun X and Fu K. 2019. A remote sensing image dataset for cloud removal. *arXiv:1901.00600* [DOI: 10.48550/arXiv.1901.00600]
- Mehta A, Sinha H, Mandal M and Narang P. 2021. Domain-aware unsupervised hyperspectral reconstruction for aerial image dehazing//2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE: 413-422 [DOI: 10.1109/WACV48630.2021.00046]
- Mou F. 2021. Research on Thin Cloud Removal and Application of Remote Sensing Image Based on Deep Learning. Chengdu: University of Electronic Science and Technology of China (牟范. 2021. 基于深度学习的遥感影像去薄云及应用研究. 成都: 电子科技大学) [DOI: 10.27005/d.cnki.gdzku.2021.001113]
- Qin X, Wang Z L, Bai Y C, Xie X D and Jia H Z. 2020. FFA-Net: feature fusion attention network for single image dehazing. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(7): 11908-11915 [DOI: 10.1609/aaai.v34i07.6865]
- Qu Y Y, Chen Y Z, Huang J Y and Xie Y. 2019. Enhanced Pix2pix dehazing network//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE: 8152-8160 [DOI: 10.1109/CVPR.2019.00835]
- Ren W Q, Liu S, Zhang H, Pan J S, Cao X C and Yang M H. 2016. Single image dehazing via multi-scale convolutional neural networks//14th European Conference on Computer Vision. Amsterdam: Springer: 154-169 [DOI: 10.1007/978-3-319-46475-6_10]
- Ren W Q, Ma L, Zhang J W, Pan J S, Cao X C, Liu W and Yang M H. 2018. Gated fusion network for single image dehazing//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 3253-3261 [DOI: 10.1109/CVPR.2018.00343]
- Shi W Z, Caballero J, Huszar F, Totz J, Aitken A P, Bishop R, Rueckert D and Wang Z H. 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 1874-1883 [DOI: 10.1109/CVPR.2016.207]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556* [DOI: 10.48550/arXiv.1409.1556]
- Song Y F, Li J, Wang X G and Chen X W. 2018. Single image dehazing using ranking convolutional neural network. *IEEE Transactions on Multimedia*, 20(6): 1548-1560 [DOI: 10.1109/TMM.2017.2771472]
- Sun X and Xu J D. 2021. Remote sensing image dehazing method based on cascaded generative adversarial network. *Journal of Computer Applications*, 41(8): 2440-2444 (孙潇, 徐金东. 2021. 基于级联生成对抗网络的遥感图像去雾方法. *计算机应用*, 41(8): 2440-2444) [DOI: 10.11772/j.issn.1001-9081.2020101563]
- Tan R T. 2008. Visibility in bad weather from a single image//2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage: IEEE: 1-8 [DOI: 10.1109/CVPR.2008.4587643]
- Tang K T, Yang J C and Wang J. 2014. Investigating haze-relevant features in a learning framework for image dehazing//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE: 2995-3002 [DOI: 10.1109/CVPR.2014.383]
- Wang L W, Li Y, Huang J and Lazebnik S. 2019. Learning two-branch neural networks for image-text matching tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2): 394-407 [DOI: 10.1109/TPAMI.2018.2797921]
- Wei X S, Xu Y Y and Yang J. 2022. Review of webly-supervised fine-grained image recognition. *Journal of Image and Graphics*, 27(7): 2057-2077 (魏秀参, 许玉燕, 杨健. 2022. 网络监督数据下的细粒度图像识别综述. *中国图象图形学报*, 27(7): 2057-2077) [DOI: 10.11834/jig.210188]
- Wu H Y, Liu J, Xie Y, Qu Y Y and Ma L Z. 2020. Knowledge transfer dehazing network for NonHomogeneous dehazing//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE: 1975-1983 [DOI: 10.1109/CVPRW50498.2020.00247]
- Wu H Y, Qu Y Y, Lin S H, Zhou J, Qiao R Z, Zhang Z Z, Xie Y and Ma L Z. 2021. Contrastive learning for compact single image dehazing//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE: 10546-10555 [https://doi.org/10.1109/CVPR46437.2021.01041]
- Yang S T, Wang P F, Wang J, Lou H Z and Gong T L. 2021. River flow estimation method based on UAV aerial photogrammetry. *National Remote Sensing Bulletin*, 25(6): 1284-1293 (杨胜天, 王鹏飞, 王娟, 姜和震, 巩同梁. 2021. 结合无人机航空摄影测量的河道流量估算. *遥感学报*, 25(6): 1284-1293) [DOI: 10.11834/jrs.20209082]
- Yu F and Koltun V. 2016. Multi-scale context aggregation by dilated

- convolutions. arXiv:1511.07122 [DOI: 10.48550/arXiv.1511.07122]
- Yu Y K, Liu H, Fu M H, Chen J, Wang X Y and Wang K Y. 2021. A two-branch neural network for non-homogeneous dehazing via ensemble learning//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville: IEEE: 193-202 [DOI: 10.1109/CVPRW53098.2021.00028]
- Yuan F N, Li Z Q, Shi J T, Xia X and Li Y. 2021. Image defogging algorithm using a two-phase feature extraction strategy. *Journal of Image and Graphics*, 26(3): 568-580 (袁非牛, 李志强, 史劲亭, 夏雪, 李雅. 2021. 两阶段特征提取策略的图像去雾. *中国图象图形学报*, 26(3): 568-580) [DOI: 10.11834/jig.200057]
- Zhang H and Patel V M. 2018. Densely connected pyramid dehazing network//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 3194-3203 [DOI: 10.1109/CVPR.2018.00337]
- Zhang Z L, Zhang X Y, Peng C, Xue X Y and Sun J. 2018. ExFuse: enhancing feature fusion for semantic segmentation//15th European Conference on Computer Vision. Munich: Springer: 273-288 [DOI: 10.1007/978-3-030-01249-6_17]
- Zhao W D, Li S S, Li A, Zhang B and Chen J. 2021. Deep fusion of hyperspectral images and multi-source remote sensing data for classification with convolutional neural network. *National Remote Sensing Bulletin*, 25(7): 1489-1502 (赵伍迪, 李山山, 李安, 张兵, 陈俊. 2021. 结合深度学习的高光谱与多源遥感数据融合分类. *遥感学报*, 25(7): 1489-1502) [DOI: 10.11834/jrs.20219117]
- Zhu Q S, Mai J and Shao L. 2014. Single image dehazing using color attenuation prior//Proceedings of the British Machine Vision Conference 2014. Nottingham: BMVA [DOI: 10.5244/C.28.114]

A two-branch remote sensing image dehazing network based on hierarchical feature interaction and enhanced receptive field

SUN Hang^{1,3}, FANG Shuailing¹, DAN Zhiping^{1,3}, REN Dong^{1,2}, YU Mei^{1,2}, SUN Shuifa^{1,3}

1. College of Computer and Information Technology, China Three Gorges University, Yichang, 443002, China;

2. Hubei Engineering Technology Research Center for Farmland Environmental, China Three Gorges University, Yichang, 443002, China;

3. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang, 443002, China

Abstract: In recent years, deep learning-based dehazing methods have achieved remarkable results in the field of image dehazing. However, most dehazing methods based on U-shaped networks directly transfer the features of the encoding layer to the corresponding decoding layers, which lacks information interaction between the low- and high-level features. Meanwhile, the network model designed based on the U-shaped structure may destroy the detailed information important for the restored image in the process of downsampling. As a result, the restored clear image lacks detailed texture and structure information. In addition, the dehazing method based on non-U-shaped network has limited receptive field, which hinders its capability to effectively utilize contextual information. As a result, these methods cannot achieve ideal dehazing results in remote sensing images with large scene scale changes. Therefore, this study proposes a two-branch remote sensing image dehazing network based on hierarchical feature interaction and enhanced receptive field. This network includes hierarchical feature interaction sub-net and multi-scale information extraction sub-net. The hierarchical feature interaction sub-net uses the hierarchical feature interaction fusion module to introduce semantic information into low-level features and spatial details into high-level features layer by layer. This way enhances the information interaction between features at different levels in the encoding layer. The multi-scale information extraction sub-net uses the multi-scale residual dilated convolution module to fuse the features of different receptive fields for obtaining contextual information, which is crucial for remote sensing image dehazing. The experiment on two public datasets show that the dehazing method proposed in this study achieves the best evaluation compared with the existing nine excellent dehazing algorithms. Among them, in the three sub-test sets of the public remote sensing dataset Haze1k, the quantitative index PSNR values of this study reach 27.362, 28.171, and 25.137 dB. In the two sub-test sets of the public remote sensing dataset RICE, the quantitative index PSNR values of this study reach 37.79 and 35.367 dB. In addition, the method proposed in this study is the closest to ground truth in terms of subjective visual qualities such as color, saturation, and sharpness, while still achieving the dehazing effect. The following conclusions can be drawn: (1) through the proposed hierarchical feature interaction fusion module, the deep semantic information in the coding stage is gradually interactively fused with the shallow detailed texture information, which enhances the expressive ability of the network and restores clear images with higher quality. (2) Through the multi-scale residual dilated convolution module, the dehazing network proposed in this study can increase the receptive field of the network without changing the size of the feature map. The contextual information of different scales can also be fused. (3) In two public remote sensing image dehazing datasets, namely, Haze1k and RICE, the dehazing method proposed in this study outperforms nine recently proposed excellent dehazing algorithms in terms of objective evaluation indexes and subjective visual effects.

Key words: deep learning, remote sensing image dehazing, hierarchical feature interaction, receptive field, two-branch

Supported by Natural Science Foundation of Hubei Province of China (No. 2021CFB004); National Key Research and Development Program of China (No. 2016YFD0800902); Natural Science Foundation of China (No. 61871258)