



# 计算机科学

COMPUTER SCIENCE

## 基于动态卷积核的自适应图像去雾算法

刘哲, 梁宇栋, 李嘉莹

引用本文

刘哲, 梁宇栋, 李嘉莹. 基于动态卷积核的自适应图像去雾算法[J]. 计算机科学, 2023, 50(6): 200-208.

LIU Zhe, LIANG Yudong, LI Jiaying. [Adaptive Image Dehazing Algorithm Based on Dynamic Convolution Kernels](#) [J]. Computer Science, 2023, 50(6): 200-208.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

### [基于日志模板主题特征的日志异常检测](#)

LTTTFAD:Log Template Topic Feature-based Anomaly Detection

计算机科学, 2023, 50(6): 313-321. <https://doi.org/10.11896/jsjcx.220500020>

### [基于增强AST的图神经网络函数级代码漏洞检测方法](#)

Function Level Code Vulnerability Detection Method of Graph Neural Network Based on Extended AST

计算机科学, 2023, 50(6): 283-290. <https://doi.org/10.11896/jsjcx.220600131>

### [基于多特征嵌入的中文医学命名实体识别](#)

Chinese Medical Named Entity Recognition Based on Multi-feature Embedding

计算机科学, 2023, 50(6): 243-250. <https://doi.org/10.11896/jsjcx.220400115>

### [深度学习容器云平台下的GPU共享调度系统](#)

GPU Shared Scheduling System Under Deep Learning Container Cloud Platform

计算机科学, 2023, 50(6): 86-91. <https://doi.org/10.11896/jsjcx.220900110>

### [基于多事件语义增强的情感分析](#)

Sentiment Analysis Based on Multi-event Semantic Enhancement

计算机科学, 2023, 50(5): 238-247. <https://doi.org/10.11896/jsjcx.220400256>

# 基于动态卷积核的自适应图像去雾算法

刘 哲 梁宇栋 李嘉莹

山西大学计算机与信息技术学院 太原 030006

(202022407037@email.sxu.edu.cn)

**摘 要** 现有图像去雾方法普遍存在去雾不彻底、容易出现颜色失真等问题,基于传统深度学习模型的图像去雾方法多采用静态推理模式,在该模式下,模型对不同样本会采用同样的、固定的参数设置,从而抑制了模型的表达能力,影响图像的去雾效果。针对以上问题,文中提出了一种基于动态卷积核的自适应图像去雾算法,该算法包括编码网络、自适应特征增强网络和解码网络3个部分。文中采用动态卷积、密集残差、注意力机制设计了自适应特征增强网络,该网络主要包括动态残差组件和动态跨层特征融合组件。动态残差组件由动态密集残差模块、一个卷积层和双注意力模块构成,其中动态密集残差模块将动态卷积引入密集残差模块,同时设计了一个基于注意力的权重动态聚合子网络,动态地生成卷积核参数以达到样本自适应的目的,在减少信息丢失的同时增强了模型的表达能力;双注意力模块结合通道注意力和像素注意力,使模型更加关注图像通道之间的差异性以及雾霾分布不均匀的区域。动态跨层特征融合组件通过动态融合不同阶段的特征,来学习丰富的上下文信息,防止网络深层计算时遗忘网络的早期特征,同时极大地丰富了特征表示,有利于模型对无雾图像细节信息的恢复。在合成数据集和真实数据集上进行了大量实验,结果表明,所提方法不仅取得了较好的客观评价分数,而且重建了主观效果较好的去雾图像,超越了对比方法的性能。

**关键词:** 图像去雾;深度学习;动态神经网络;注意力机制;特征融合

**中图法分类号** TP391

## Adaptive Image Dehazing Algorithm Based on Dynamic Convolution Kernels

LIU Zhe, LIANG Yudong and LI Jiaying

School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China

**Abstract** Existing image dehazing methods generally have problems such as incomplete dehazing and color distortion. Image dehazing methods based on traditional deep learning models mostly use static inference during testing, which use the same and fixed parameters for different samples, thereby inhibiting the expressive ability of the model and decreasing the dehazing performance. Aiming at the above problems, this paper proposes an adaptive image dehazing algorithm based on dynamic convolution kernel. The proposed model includes three parts: encoding network, adaptive feature enhancement network and decoding network. This paper combines dynamic convolutions, dense residual connections, and attention mechanism to complete the adaptive feature enhancement network, which mainly includes dynamic residual components and dynamic skip-connected feature fusion components. The dynamic residual component is composed of a dynamic residual dense block, a convolutional layer and a dual attention module. The dynamic residual dense block introduces dynamic convolutions into the residual dense block, and an attention-based weight dynamic aggregator is designed at the same time, which dynamically generates adaptive convolution kernel parameters. The

到稿日期:2022-04-28 返修日期:2022-11-08

基金项目:国家自然科学基金(61802237,62272284,61906114);山西省研究生教育创新项目(2022Y127);山西省基础研究计划(自由探索类)项目(202203021221002,202203021211291);山西省自然科学基金(201901D211176,201901D211170,202103021223464);山西省高等学校科技创新项目(2019L0066);山西省科技重大专项计划(202101020101019);山西省重点研发计划(国际科技合作,201903D421041,能源与节能环保领域,202102070301019);山西省科技创新人才团队专项资助

This work was supported by the National Natural Science Foundation of China(61802237,62272284,61906114), Graduate Education Innovation Programs of Shanxi Province(2022Y127), Fundamental Research Program of Shanxi Province(202203021221002,202203021211291), Natural Science Foundation of Shanxi Province(201901D211176,201901D211170,202103021223464), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi(2019L0066), Science and Technology Major Project of Shanxi Province(202101020101019), Key R & D Program of Shanxi Province(201903D421041,202102070301019) and Special Fund for Science and Technology Innovation Teams of Shanxi Province.

通信作者:梁宇栋(liangyudong@sxu.edu.cn)

dynamic convolutions have reduced the loss of information and enhanced the expressive ability of the model. The dual attention module combines channel attention and pixel attention to make the model pay more attention to the differences between image channels and areas with uneven distribution of haze. The dynamic skip-connected feature fusion component learns rich contextual information by dynamically fusing the features of different stages via skip-connections, preventing the early features of the network from being forgotten when the information flows into deeper layers. Meanwhile, the feature representations are greatly enriched, which benefits the restorations of the details for fog-free images. Extensive experiments on synthetic datasets and real datasets show that our method not only achieves better objective evaluation scores, but also reconstructs dehazing images with better visual effects, surpassing the performance of compared methods.

**Keywords** Image dehazing, Deep learning, Dynamic neural network, Attention mechanism, Feature fusion

## 1 引言

随着工业的发展与碳排放量的加大,雾霾天气已经逐渐成为城市中一种常见的天气现象。雾霾成像条件会导致图像采集设备受到干扰,采集到的图像会产生严重的退化现象,出现低对比度、场景模糊、颜色失真、视觉效果差等情况,从而影响后续基于图像的智能分析等应用。因此,对图像去雾技术的研究有着广泛的实用价值与现实意义。

目前针对图像去雾问题已有大量解决方法,主要可以分为非基于学习的方法和基于学习的方法。其中,非基于学习的图像去雾算法包括基于传统图像处理操作的图像去雾算法和基于先验模型的图像去雾算法。基于传统图像处理操作的图像去雾算法,通过增强图像对比度来恢复无雾的清晰图像,相关图像处理操作如直方图均衡化<sup>[1]</sup>、Retinex算法<sup>[2]</sup>等。这类方法虽然能在一定程度上提升低质图像的视觉效果,但并没有考虑到有雾图像的退化机理,算法结果仍会出现颜色失真等情况。基于先验模型的图像去雾算法利用先验知识,同时结合物理模型来对雾霾图像进行恢复,如基于暗通道先验(DCP)<sup>[3]</sup>的方法等,当先验成立条件不满足时,这种方法的去雾效果往往不尽如人意。

基于学习尤其是基于深度学习的图像去雾算法<sup>[4-5]</sup>取得了当前最好的去雾性能,该类方法使用深度神经网络进行建模,同时在大规模数据集上进行训练,现已经成为低质图像恢复领域的主流方法。其中,残差网络,尤其是密集残差网络的使用,显著提升了网络的表达能力,并避免了网络过深而造成的信息丢失,极大地降低了训练的难度,提升了算法性能。残差网络<sup>[6]</sup>通过跳层连接在一定程度上缓解了梯度消失的问题,同时也加快了模型在训练过程中的收敛速度。密集残差模块(Residual Dense Block, RDB)<sup>[7]</sup>是在残差网络的基础上在不同网络层之间加入更加密集的残差连接,充分利用了多级特征之间的信息。然而,当前基于深度学习的图像去雾算法基本采用静态推理范式,即在推理时模型的参数固定。模型的静态推理并没有考虑不同输入雾图的特征,而是以一个固定的模型来处理不同的输入,这大大影响了模型的表达能力,同时也降低了模型的推理效率。

传统静态神经网络通过在大规模数据集上进行训练,最终得到了参数固定的、静态的网络模型,在测试时以训练好的固定的模型参数来应对不同输入样本,无法应对输入的多样

性。在图像去雾任务中,不同场景、不同时间下获得的有雾图像的雾特征分布是不同的,如果仍然采用传统静态的推理模式,即针对不同图像、不同雾分布采用相同的模型参数,模型将无法有效实现有雾图像到无雾图像的映射,去雾效果将会受到影响。因此,采用动态卷积网络可以避免传统静态推理模式带来的弊端,以一种输入样本自适应的方式动态调整网络参数,在推理阶段根据不同输入样本、不同的特征分布生成适合当前输入样本的映射关系,如调整卷积核参数,从而更好地调整模型参数以适应不同的有雾场景,极大地提升了去雾模型的表征能力。

真实有雾场景下容易出现雾霾分布不均匀、色偏严重等情况,而深度模型的静态推理模式无法针对不同的雾霾图像、同一雾霾图像的不同雾霾区域进行不同的去雾操作,极大降低了去雾算法的性能。针对该问题,本文提出了一种基于动态卷积核的自适应图像去雾算法,端到端地恢复无雾图像。为了增强去雾模型的表达能力,应对真实场景雾霾图多样、雾霾分布不均匀等情况,本文设计了一种动态密集残差模块(Dynamic Residual Dense Block, Dy-RDB),引入动态卷积核来构造动态卷积,可以在推理阶段根据不同的输入样本生成适合当前样本的卷积核参数,从而针对不同浓度分布的有雾图像调整模型参数,以增加较少模型参数量的代价实现了样本的自适应去雾。为了进一步融合不同阶段的特征,减小梯度爆炸和湮灭问题并增加模型容量,Dy-RDB使用密集残差设计,通过跨层连接来密集融合不同阶段的特征,加强了不同阶段特征间的紧密联系,提升了特征间的传递效率,避免了信息丢失;同时,兼顾了模型容量与网络性能,降低了网络的训练难度;提高了网络的推理效率,对于真实复杂雾霾场景具有不错的去雾效果。为了进一步提升去雾网络的性能并解决图像部分区域去雾不彻底的现象,本文设计了一种动态跨层特征融合组件(Dynamic Skip-Connected Feature Fusion Component, DSCFF),动态融合不同阶段的特征,能够自适应地利用多级特征互补地进行特征增强,从而帮助模型更好地去雾,使得去雾效果更加明显。有雾图像雾霾分布不均匀、雾霾图像不同的通道具有不同的特性,为了进一步提升模型对不同浓度的雾图像区域和不同的雾图像通道特征的自适应映射能力,本文还采用通道注意力和像素注意力相结合的双注意力机制,对不同浓度的有雾区域、不同的图像通道分配不同的权重,以赋予不同的关注程度,即对雾霾更浓的区域、更重要的

通道分配更大的权重。双注意力机制广义上属于动态模型的一种,能够与本文设计的基于动态卷积核的动态卷积操作协同增加模型对输入图像的自适应能力。密集连接通过跨层连接融合不同阶段的特征,缓解梯度爆炸和湮灭问题。动态卷积、密集连接和注意力机制共同增加了模型容量,有效提升了模型特征的代表能力。

本文的主要贡献如下:

(1)提出了一种基于动态卷积核的自适应图像去雾算法,设计了一个基于注意力的权重动态聚合子网络,能够根据不同雾霾浓度的样本动态地生成卷积核权重,自适应地重建去雾图像。

(2)设计了一种将深度模型不同阶段特征通过跨层连接动态融合的动态跨层特征融合组件,有效缓解了网络早期特征被遗忘的难题,实现了多阶段特征的有效融合和互补,增加了模型容量。

本文在采取较小计算量的情况下,与目前性能最好的图像去雾方法相比,取得了更好或相近的性能,并表现出了更好的跨数据集模型泛化性。

## 2 相关工作

### 2.1 图像去雾

基于传统图像处理操作的图像去雾算法,往往从提升图像对比度或减小色彩偏移的角度对有雾图像进行增强。为了更好地进行局部增强,Khan等<sup>[1]</sup>使用直方图分割算法将输入图像的直方图分为多个子直方图,并对子直方图对应的区域进行增强,在避免图像亮度饱和的同时得到了较为自然的效果。Li等<sup>[2]</sup>提出了一种基于GPU CUDA的多尺度Retinex理论的图像增强算法,提升了算法的执行效率,增强了去雾后图像的视觉效果。然而,传统的图像处理操作没有考虑雾霾图像退化的机理,只能在一定程度上改善图像的视觉效果,不能从根本上去除雾霾对图像的影响。

基于先验模型的图像去雾算法假设图像满足某种统计特性,从而进行图像去雾。该类方法中的部分方法基于大气散射模型,通过估计模型参数对有雾图像进行恢复,取得了较好的去雾效果。He等<sup>[3]</sup>提出了暗通道先验理论,通过简单估计大气光值和传输率来快速去雾。Arigela等<sup>[8]</sup>设计了非线性正弦函数来修正暗通道先验,初步获得了传输率图,同时利用变换函数为每个像素生成所需的非线性映射,同时保持局部对比度。Berman等<sup>[9]</sup>提出了一种基于非局部先验进行图像去雾的算法,通过在雾霾图像的RGB空间中生成雾线,来恢复出距离图和无雾图。这些方法虽然考虑到了雾图的退化机理,但在先验假设不成立的情况下,去雾的效果也会相应变差。

基于深度学习的图像去雾算法通过在大规模数据集上进行训练来恢复无雾图像。Cai等<sup>[4]</sup>通过一个端到端的卷积神经网络来估计传输率,再通过大气光散射模型恢复无雾图像。Ren等<sup>[5]</sup>提出了一个多尺度卷积神经网络,由粗到细地学习有雾图到传输率图之间的映射关系,再使用估计出的传输率图和大气光值,通过大气散射模型进行图像去雾。Ren等<sup>[10]</sup>

采用白平衡、对比度增强和伽马校正3种方法对雾图进行处理,并设计了一个编码-解码网络,用于处理预处理图像并提取特征,最后融合得到恢复的图像。Li等<sup>[11]</sup>没有直接估计传输率图和大气光值,而是将两者统一为参数 $K$ ,通过卷积神经网络估计参数 $K$ 再进行图像去雾。Zhang等<sup>[12]</sup>通过构建密集连接的金字塔网络和U-net网络来同时估计传输率和大气光值,再通过大气光散射模型进行图像恢复。Yi等<sup>[13]</sup>提出了一种多尺度拓扑网络,在融合不同尺度特征时自动进行特征选择以筛选重要信息,从而忽略干扰信息,实现渐进式图像去雾。Zhang等<sup>[14]</sup>通过两路并行且交互的网络同时进行有雾图像低频和高频信息的学习,通过融合两路的中间特征实现信息交互,从而使低频信息和高频信息的增强效果得到进一步提升。Wu等<sup>[15]</sup>提出了一种基于对比学习的去雾方法,将有雾图和无雾图分别作为负样本和正样本,拉近去雾结果与正样本之间的距离,增大与负样本之间的距离,进而提升去雾网络的性能。Yang等<sup>[16]</sup>设计了一种自增强图像去雾框架,将传输图分解为密度信息与深度信息,同时采用非成对图像用于雾霾的生成与去除,提升了模型的泛化性能。Liu等<sup>[17]</sup>利用一个自监督形式的辅助网络,在测试时帮助去雾网络调整参数,以提升模型对不同类雾图的去雾效果,同时以元学习的方式保证两个网络的目标一致。然而,这些基于深度学习的去雾方法往往并未重点关注有雾图像雾霾较重的局部区域,而是对图像不同区域、不同通道都采用相同的卷积处理。

### 2.2 动态神经网络

动态神经网络是能够根据输入自适应地调整映射关系的神经网络模型。注意力机制就是一种常见的动态神经网络技术。

基于样本自适应的动态神经网络可大致分为结构动态化网络和参数动态化网络两种类型。结构动态化网络往往可以动态调整模型的深度或宽度,通过这些动态结构可达到样本自适应的目的。Wu等<sup>[18]</sup>通过学习一个策略网络,以强化学习的方式来输出“保留”或者“丢弃”的决策,进而在推理时能够进行最佳的路径选择,在保持精度的同时降低模型的复杂度。Yang等<sup>[19]</sup>根据不同分辨率(即不同难易程度)的输入图像采用不同深度的网络结构,设计了一种分辨率自适应网络,为不同难度的样本采用不同深度的网络,在性能和计算效率之间达到了不错的平衡。与以上工作在网络结构上进行动态化的操作不同,参数动态化网络是在保持结构静态的同时只在卷积核等参数上进行动态化操作。Yang等<sup>[20]</sup>提出了一种条件卷积方法,通过将多个卷积核加权求和得到最终的卷积核,再对特征进行卷积操作,相比先卷积后加权的方式极大地提升了计算效率。Ma等<sup>[21]</sup>提出了权重网络,通过在注意力激活层后添加分组的全连接层来直接输出卷积核权重,在降低计算量的同时提升了模型性能。Chen等<sup>[22]</sup>提出了一种动态卷积操作,在没有增加网络深度和宽度的基础上,通过一种非线性的方式动态地聚合多个卷积核来得到最终的卷积核参数。近年来,随着注意力机制研究成果的大量出现<sup>[23-26]</sup>,基于注意力机制的图像去雾算法也越来越多。注意力机制本质上

是根据信息重要程度的不同来动态调整权重,以使模型更加关注重点信息,从而减少对不重要信息的关注。Liu 等<sup>[27]</sup>提出了一种基于注意力的多尺度去雾网络,通过融合不同尺度的特征实现高效去雾。Hong 等<sup>[28]</sup>采用知识蒸馏框架,提出了一种空间加权残差通道注意力模块,用于构建教师和学生网络,教师网络和学生网络分别完成异构任务,目的是将训练好的教师重建网络的中间层特征知识迁移给学生网络,以帮助学生网络更好地去雾。Qin 等<sup>[29]</sup>将通道注意力与像素注意力相结合,达到了较好的去雾效果。以上算法虽然采用注意力机制动态地对雾图重要的部分分配更大的权重,在一定程度上可以提升模型的去雾性能,但仍然无法充分发掘图像去雾网络的表达能力,不能灵活处理真实场景下雾浓度分布不均匀等情况,从而出现去雾不彻底、颜色失真和细节丢失等现象。本文引入动态卷积核,在推理阶段可以根据不同样本的不同特征自适应地、动态地调整卷积核参数,避免了上述方法带来的弊端。

### 3 本文方法

本文提出的基于动态卷积核的自适应图像去雾算法,

可以端到端地自适应学习图像去雾过程,其网络结构如图 1 所示。该网络主要包含 3 个部分:编码网络、自适应特征增强网络和解码网络。具体地,输入的有雾图像首先经过一个编码网络进行初步的特征提取与 4 倍下采样,编码网络由 4 个卷积层组成,其中后 3 个卷积层都采用 ReLU 作为激活函数。经过编码网络的特征图具有紧凑的特征表征,其维度为 128 维。自适应特征增强网络由 6 个动态残差组件(Dynamic Residual Components, DRC)和 2 个动态跨层特征融合组件(Dynamic Skip-Connected Feature Fusion Component, DSCFF)组成,可以对编码网络得到的特征图进行动态自适应增强。解码网络由 3 个卷积层构成,其中前两个卷积层进行反卷积操作,将自适应特征增强网络的输出特征图进行 4 倍上采样,最后一个卷积层恢复出最终的去雾图像。同时,在编码网络和解码网络间使用大残差跨层融合结构,即将相应特征进行级联 concat 操作。这样不仅可以在网络中保留丰富的上下文信息,还可以促进模型收敛,防止网络在深层出现梯度消失与退化。接下来将对各个模型组件和损失函数分别进行介绍。

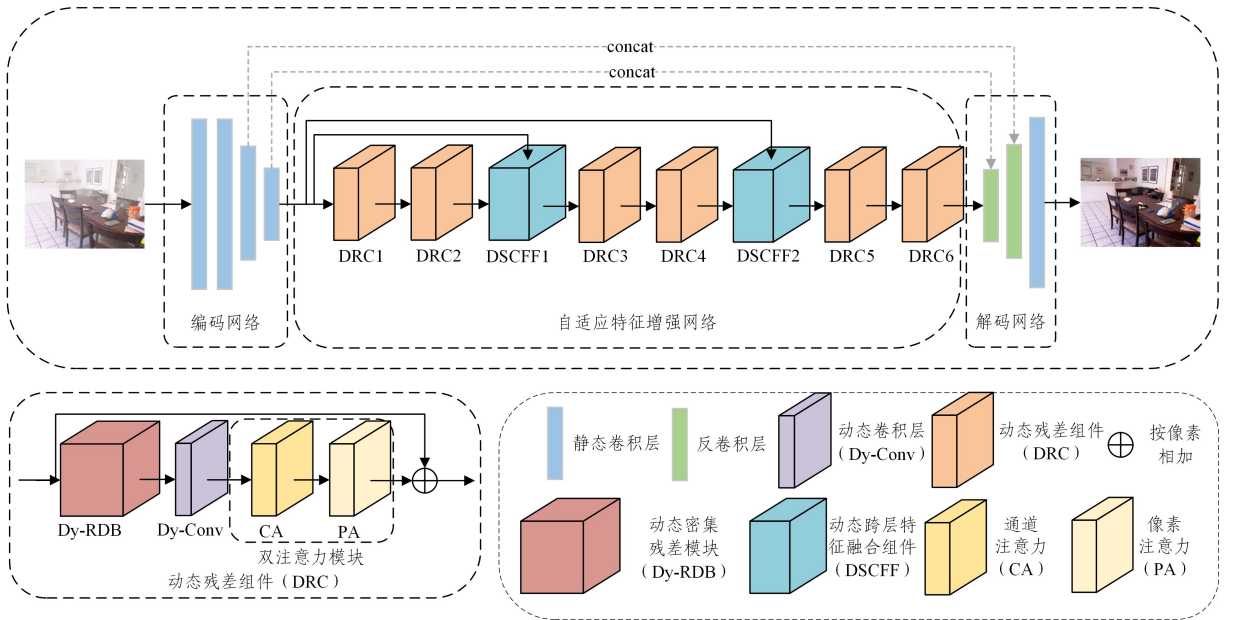


图 1 基于动态卷积核的自适应图像去雾模型

Fig. 1 Adaptive image dehazing model based on dynamic convolution kernel

#### 3.1 动态残差组件

为了有效提升去雾网络的性能,同时更好地应对真实场景雾霾分布不均匀的情况,本文设计了动态残差组件。如图 1 所示,动态残差组件主要由动态密集残差模块(Dynamic Residual Dense Block, Dy-RDB)、一个卷积层和双注意力模块构成。

##### 3.1.1 基于注意力的聚合动态卷积

为了增强传统卷积神经网络的表达能力,本文设计了一个基于注意力的权重动态聚合子网络来动态地生成卷积核权重。如图 2 所示,首先模型将特征图输入平均池化层中,再经过一个“全连接层-ReLU-全连接层”结构之后,将全连接层的输出输入 softmax 层中得到一组注意力权重  $\{A_1, A_2, \dots,$

$A_K\}$ ,最后将得到的这组注意力权重与  $K$  个卷积核权重进行动态聚合,从而得到最终的卷积核。在具体实现中, $K$  取值为 4,这  $K$  个卷积核采用标准正态分布进行初始化。这种动态卷积操作在不增加网络深度和宽度的基础上,仅增加较少的模型复杂度,达到了对给定的不同样本动态调整卷积核以适应样本多样性的目的。对于复杂多变的有雾场景,动态卷积可以对雾霾分布不均匀的有雾图像自适应地生成相应的卷积核,增加模型的灵活性。动态卷积的过程可表示为:

$$y = C(x, W(x)) = C(x, \sum_{k=1}^K A_k(x) W_k) \quad (1)$$

其中,  $C$  表示动态卷积操作,  $W(x)$  表示样本自适应生成的

卷积核参数,  $K$  表示生成的权重个数,  $A_k(x)$  表示第  $k$  个注意力权重,  $W_k$  表示第  $k$  个卷积核参数。动态卷积有效避免了传统卷积无法依据输入调整映射关系的弊端, 提升了算法性能。

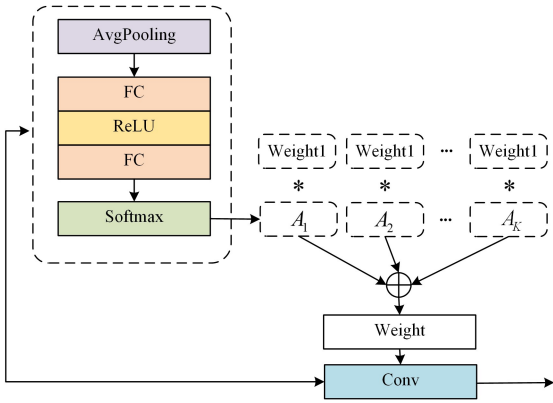


图2 动态卷积层结构

Fig. 2 Dynamic convolutional layer structures

### 3.1.2 动态密集残差模块

传统 RDB(Residual Dense Block)<sup>[7]</sup> 将残差结构、局部特征融合和密集跨层连接整合在一起, 更好地保证了信息的流动。其中密集跨层连接在一定程度上能够减少网络较浅层处随着网络层数增加而造成的信息损失, 能够保留多层次的信息, 增强模型的性能。但 RDB 固定的推理模式限制了其对复杂多样性样本的适应能力。针对上述弊端, 本文将动态卷积操作引入 RDB 中, 设计了一种基于动态卷积核的动态密集残差模块(Dy-RDB)。如图 3 所示, 该模块中局部的密集连接和特征融合不仅充分利用了多层次特征之间的信息, 还可以使得网络的训练更加稳定。

具体地, 输入的雾特征图首先经过 3 个密集连接的动态卷积层, 其中每个动态卷积层的特征都会与其余动态卷积层的特征进行密集连接; 然后融合后的特征会经过一个  $1 \times 1$  的动态卷积层; 最后将输出的特征图与最初的输入进行残差连接。经过动态密集残差模块输出的特征可以保留网络中每一层的信息, 同时对于不同的雾图可以动态自适应地生成参数, 极大地提升了去雾模型的代表能力。

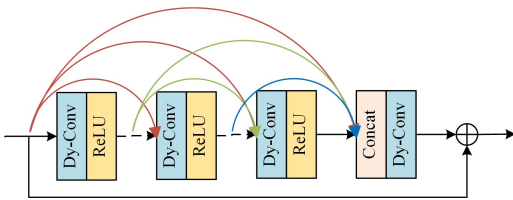


图3 动态密集残差模块

Fig. 3 Dynamic residual dense block

### 3.1.3 双注意力模块

有雾图像中常常雾霾分布不均匀, 现有模型对浓雾去雾效果往往不彻底。受文献[29]的启发, 本文采用通道注意力和像素注意力相结合的双注意力机制来实现更好的去雾效果。有雾图像不同通道、不同像素位置受到雾霾的影响不同,

对雾图不同通道、不同像素进行不同的操作能够更好地进行图像去雾。

如图 4 所示, 对于通道注意力, 特征图  $F$  首先经过平均池化来求各个通道的均值, 即将全局空间信息转化为通道描述信息, 得到  $1 \times 1 \times C$  维的特征。再将在空间维度压缩后的特征输入两个卷积层和 Sigmoid 激活函数中, 最后根据得到的一组通道注意力权重  $W_c$  在通道方向上对特征图进行分配。

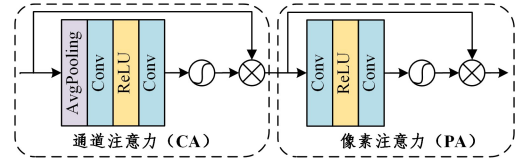


图4 双注意力模块

Fig. 4 Dual attention module

通道注意力权重的计算式如下:

$$W_c = \sigma(\text{Conv}(\delta(\text{Conv}(\text{AvgPool}(F)))))) \quad (2)$$

结合通道注意力机制的特征图的计算式如下:

$$F_c = W_c \otimes F \quad (3)$$

其中,  $\text{Conv}$  表示卷积层,  $\delta$  表示 ReLU 激活函数,  $\sigma$  表示 Sigmoid 激活函数,  $\otimes$  表示点乘操作,  $F_c$  表示经过通道注意力的输出特征。

类似于通道注意力, 像素注意力将通道注意力的输出  $F_c$  直接送入两个卷积层和 Sigmoid 激活函数中, 得到  $H \times W \times 1$  维的像素注意力权重  $W_p$ , 最后对像素分配注意力。

$$W_p = \sigma(\text{Conv}(\delta(\text{Conv}(F_c)))) \quad (4)$$

$$F_p = W_p \otimes F_c \quad (5)$$

其中,  $W_p$  表示像素注意力的权重,  $F_p$  表示结合像素注意力的输出特征。

通道与像素注意力会使网络更加关注浓雾区域和图像细节等较难重建的区域, 能够更好地建模有雾图像不均匀的雾霾分布。

### 3.2 动态跨层特征融合组件

动态跨层特征融合组件融合初级阶段的特征图并将其作为当前特征的补充信息, 在防止信息丢失的同时动态地提升了去雾模型的性能与效率。如图 5 所示, 动态跨层特征融合组件(DSCFF)分两支进行, 一个分支首先将当前特征图  $F$  与编码网络得到的初级特征图  $E(x)$  进行级联 concat 操作, 再经过两个动态卷积层使得到的特征充分融合, 增强了不同阶段的信息。一方面这两种不同阶段的信息可以互为补充, 另一方面也有效解决了初级阶段特征在网络深层出现信息遗忘的问题。动态跨层特征融合组件的另一个分支将当前特征图直接输入到一个动态卷积层中, 用于当前特征的信息提取与动态增强。最后将两个分支分别得到的特征图进行进一步的融合。该组件可以学习到丰富的上下文信息, 使得多阶段特征进行有效融合、互补, 防止早期特征被遗忘, 有效提升了模型的代表能力, 从而实现更加彻底高效的去雾效果。

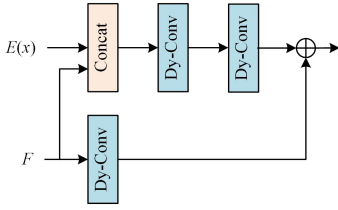


图5 动态跨层特征融合组件

Fig. 5 Dynamic skip-connected feature fusion component

### 3.3 损失函数

本文采用了重建损失  $L_c$  和感知损失  $L_p$  作为模型训练的损失函数。为了更好地恢复无雾清晰图像,在像素级别上缩小去雾图与 ground-truth 之间的差距,本文使用 L1 损失作为重建损失。重建损失  $L_c$  可表示为:

$$L_c = \|G(x) - J\|_1 \quad (6)$$

其中,  $G(x)$  表示去雾网络的输出结果图,  $J$  表示真实无雾图。

雾特征图中不仅具有丰富的高级语义信息,还包含着大量的域结构信息。为了使去雾结果与清晰图像 ground-truth 之间的相似性更高,本文采用感知损失来衡量两者特征之间的相似性,以缩小两者之间的域差异。本文使用预训练好的 VGG19 模型,在网络中分别提取去雾图和清晰图的相应特征图来计算感知损失。感知损失  $L_p$  可表示为:

$$L_p = \sum W_i * \|\Phi^i(G(x)) - \Phi^i(J)\|_1, i \in \{2, 7, 12, 21, 30\} \quad (7)$$

其中,  $\Phi^i(\cdot)$  表示 VGG19 网络中第  $i$  层输出的特征图,  $W_i$  表示第  $i$  层的权重。本文采用 VGG19 网络中的 2, 7, 12, 21 和 30 层输出的特征图,对应的权重分别为 1/32, 1/16, 1/8, 1/4 和 1。

故网络的总损失函数为:

$$L = L_c + \lambda L_p \quad (8)$$

其中,  $\lambda$  表示感知损失的权重系数,用于平衡两个损失的比例。

## 4 实验结果与分析

### 4.1 数据集和评价指标

本文采用了两个数据集对网络进行训练,包括 RESIDE 中的 ITS 室内合成雾图数据集(测试集为 SOTS-indoor)<sup>[30]</sup> 和 NH-HAZE 真实雾图数据集<sup>[31]</sup>。ITS 数据集集中有 1399 张清晰无雾图,每张无雾图对应 10 张根据大气光散射模型合成的不同浓度的雾图。SOTS-indoor 测试集中有 500 张室内雾图,本文采用的通用做法即使用 ITS 数据集进行训练,使用 SOTS 数据集进行测试。NH-HAZE 数据集包含 55 张无雾图及与其相对应的真实场景雾图,注意官方未公布 5 张验证集和 5 张测试集对应的清晰图像。为了公平比较,本文根据通用的做法,即采用官方划分的训练集的前 40 张图像进行训练,最后 5 张图像进行测试。

为了评估本文方法的去雾性能,本文在两个数据集上采用峰值信噪比(PSNR)和结构性相似性(SSIM)进行定量的客观评价。

### 4.2 实验设置

本实验基于 Pytorch1.6.0 框架,使用显卡 NVIDIA GeForce RTX 2080Ti 进行实验。训练时从每张图像上截取  $256 \times 256$  大小的图像块作为网络的输入,批处理大小为 6。在训练时网络采用 Adam 优化器,动量衰减指数  $\beta_1$  和  $\beta_2$  分别为 0.9 和 0.999,初始学习率为 0.0002。损失函数权重  $\lambda$  为 1。本文选择图像去雾领域多个经典或目前性能较优的方法进行定量、定性的比较,包括 DCP<sup>[3]</sup>, AOD-Net<sup>[11]</sup>, GridDehazeNet<sup>[27]</sup>, KDDN<sup>[28]</sup>, FFA<sup>[29]</sup>, MSTN<sup>[13]</sup>, DPRN<sup>[14]</sup> 和 AECR<sup>[15]</sup>。

### 4.3 合成雾图数据集实验

本实验在 SOTS 室内合成数据集上进行了去雾测试,并对去雾结果进行定量与定性分析。表 1 列出了本文方法与具有代表性或当下性能较好的方法的定量对比结果。从表中可以看出,本文方法在 PSNR 指标上达到了最优性能;相比基于深度学习的图像去雾方法,传统的 DCP 去雾方法的效果要差许多;相比性能较好的 FFA 和 MSTN 算法,本文方法在 PSNR 和 SSIM 上分别提升了 0.97 dB, 0.0013 和 1.99 dB, 0.0029;与 DPRN 相比,在 SSIM 指标相近的情况下,本文在 PSNR 指标上提高了 6.09 dB;除此以外,与近期性能突出的 AECR 算法相比,本文方法在 SSIM 指标上与其相近,但在 PSNR 指标上相比该算法提升了 0.19 dB。在算法复杂度方面,本文方法的参数量和近期性能较优的方法 FFA 和 AECR 的参数量相近;而在计算量 FLOPS 上,除了早期的 AOD-Net,本文方法远低于其他比较方法,显示出了本文方法的优越性。特别地,本文方法在计算量较少的情况下取得了更好的性能。

表 1 SOTS 室内合成数据集和 NH-HAZE 真实数据集去雾结果的定量比较

Table 1 Quantitative comparison of dehazing results on SOTS indoor synthetic dataset and NH-HAZE real dataset

Method	SOTS	NH-HAZE	Param Mb	FLOPS G
	PSNR/SSIM	PSNR/SSIM		
DCP	15.09/0.7649	10.57/0.5196	—	—
AOD-Net	19.82/0.8178	15.40/0.5693	0.002	0.12
GridDehazeNet	32.16/0.9836	13.80/0.5370	0.960	21.45
KDDN	34.72/0.9845	17.39/0.5897	5.990	30.33
FFA	36.39/0.9886	19.87/0.6915	4.680	288.34
MSTN	35.37/0.9870	18.42/0.6300	18.910	—
DPRN	31.27/0.9900	—	—	—
AECR	37.17/0.9901	19.88/0.7173	2.610	43.07
Our	<b>37.36/0.9899</b>	<b>20.14/0.7097</b>	4.820	6.16

本文从 SOTS 合成数据集选取了一些有雾图像进行去雾测试并进行了展示。本文方法与其他方法的去雾效果的对比如图 6 所示。DCP 方法恢复出的图像出现了明显的颜色失真现象; AOD-Net 容易出现去雾不彻底的现象; GridDehazeNet, KDDN, FFA 和 AECR 虽然比传统方法有更好的去雾效果,但当雾霾较重时,这些方法依然会出现雾残留的现象,如图像的远景处、墙壁、地板等区域。相比以上方法,本文方法具有最好的去雾视觉效果,并且在图像细节和色彩恢复上更加接近真实的无雾图像。



图 6 SOTS 合成数据集上不同方法的定性比较

Fig. 6 Qualitative comparison of different methods on SOTS synthetic dataset

#### 4.4 真实雾图数据集实验

真实有雾场景存在雾霾分布不均匀、去雾不彻底、色彩失真等问题,为了验证本文提出的模型所真实有雾场景下的去雾效果,本实验在 NH-HAZE 非均匀真实雾霾数据集上进行了测试。从表 1 中可以看出,相比其他所有方法,本文方法在 PSNR 上取得了最好的性能,相比 AECR,在 SSIM 相近的情况下本文方法的 PSNR 提升了 0.26 dB,体现了本文动态密集残差模块和动态跨层特征融合组件的有效性。AECR 模型采用的对比学习方法有望进一步提升本文方法的性能。

图 7 给出了各方法在 NH-HAZE 真实雾霾数据集上进行去雾的效果对比。可以看出,传统的 DCP 去雾方法会导致去雾后出现严重的色偏现象,整体呈现蓝色;AOD-Net 去雾的整体效果较差;GridDehazeNet 虽然可以去除近景处的部分雾,但对于远景处雾霾的去雾效果较差且不够彻底;KDDN 虽然比传统方法的去雾效果明显,但容易出现雾残留现象,如天空区域;FFA 在去雾后也存在色偏现象,如图像的地面区域;近期性能较优的 AECR 方法虽然缓解了一些色偏的问题,但相比真实无雾图在地面、树枝的色彩等方面仍然存在

颜色过于饱和等问题。相比以上算法,本文方法在恢复图像细节和避免去雾色偏的问题上都有更好的表现,对于近处和远处雾较浓的地面等区域有着较好的重建效果,可以恢复出更加清晰自然的无雾图像。这主要是因为本文方法针对真实有雾场景去雾不彻底的问题,通过基于动态卷积核的自适应动态网络和注意力机制相结合的方法,根据不同样本的雾浓度分布动态地生成相应的卷积核权重,同时采用注意力机制对雾图的不同雾浓度区域、不同通道分配不同的权重,即让网络更加关注较难去雾区域,从而对这些区域进行更加彻底的去雾。除此以外,图像去雾是个病态问题,不同无雾图像可以退化生成同一个有雾图像,针对浓雾图像区域在去雾后容易出现颜色失真、难以恢复出真实自然色彩的问题,本文通过堆叠动态密集残差模块和动态跨层特征融合组件,使用清晰无雾图像作为监督信息,动态地调整输入雾图与输出无雾图之间的映射关系参数,从而更好地融合深度网络不同阶段的特征,使得网络在训练时可以更加灵活、充分、准确地学习到无雾图到无雾图的映射关系,从而缓解处理图像去雾这一病态问题容易出现的色偏问题。

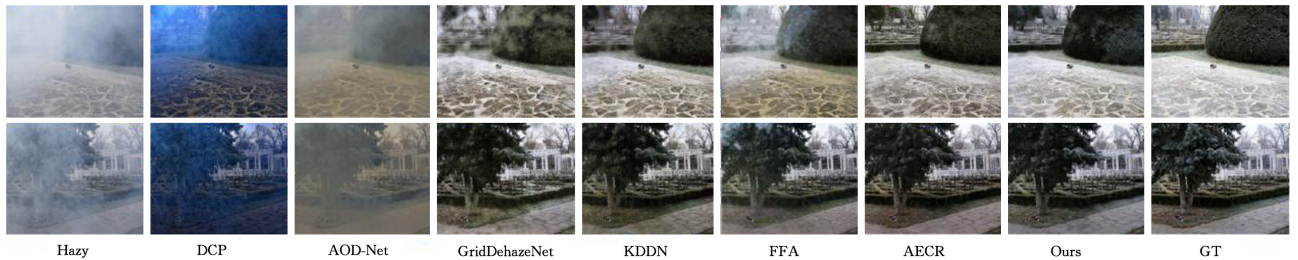


图 7 NH-HAZE 真实数据集上不同方法的定性比较

Fig. 7 Qualitative comparison of different methods on NH-HAZE real dataset

#### 4.5 跨数据集的模型泛化性实验

跨数据集模型泛化性能能够反应模型在遇到未知数据时的性能表现。本文探索了在合成雾图数据集 ITS 上进行训练并在真实雾图数据集 NH-HAZE 上进行测试时不同模型的

性能表现。如表 2 所列,在相同训练集和测试集的设置下,可以看出有方法在跨数据集的泛化性上表现普遍较差,现有基于全监督学习的去雾方法需要采集应用时的雾图数据进行模型训练才能取得较好的性能。即便如此,本文方法在

跨数据集实验上相比其他方法取得了最好的性能,从而证明了本文方法在跨数据集上有较好的泛化能力。

表2 跨数据集的模型泛化性实验结果

Table 2 Experimental results of model generalization across datasets

Method	PSNR	SSIM
AOD-Net	8.23	0.3759
GridDehazeNet	7.42	0.3568
KDDN	9.98	0.4016
FFA	10.02	0.3017
AEER	10.11	0.4252
Our	10.31	0.4327

#### 4.6 消融实验

为了证明本文所提出动态密集残差模块和动态跨层特征融合组件的有效性,本文在SOTS室内测试集上做了以下4个消融实验,对比了以下4种模型架构的性能:1)基准模型Baseline,使用传统的RDB替换动态密集残差模块,同时在网络中不使用动态跨层特征融合组件;2)Baseline+DSCFF,使用传统的RDB,在网络中使用动态跨层特征融合组件;3)Baseline+Dy-RDB,使用动态密集残差模块,不使用动态跨层特征融合组件;4)Ours:同时使用动态密集残差模块和动态跨层特征融合组件。

上述消融实验的结果如表3所列。从表中可以看出,仅使用动态跨层特征融合组件相比基准模型在PSNR和SSIM上分别能提升1.69dB和0.0045;仅使用动态密集残差模块相比基准模型在PSNR和SSIM上可以提升2.95dB和0.0059;同时使用动态跨层特征融合组件和动态密集残差模块实现了最好的性能,PSNR和SSIM分别可以提升3.63dB和0.0069。消融实验证明了本文提出的动态密集残差模块和动态跨层特征融合组件的有效性。

表3 不同模块对模型性能的影响

Table 3 Effects of different modules on model performance

Method	PSNR	SSIM
Baseline	33.73	0.9830
Baseline+DSCFF	35.42	0.9875
Baseline+Dy-RDB	36.68	0.9889
Ours	37.36	0.9899

**结束语** 本文提出了一种基于动态卷积核的自适应图像去雾算法,可以端到端地恢复出无雾清晰图像。在特征增强部分,提出了基于动态卷积设计的动态密集残差模块和动态跨层特征融合组件,能够根据输入的不同特征分布来自适应地调整卷积参数,增强了模型对输入的表达能力和适应性;对网络不同阶段的特征进行动态融合,防止网络早期特征被遗忘的问题,得到了融合不同阶段信息的有效特征表示,进一步增强了模型的表达能力。在合成数据集和真实数据集上的实验表明,本文方法在公开数据集上相比对比方法取得了更好的客观评价指标,获得了去雾更加彻底、失真现象更少的重建图像。本文从参数生成的角度实现了模型动态化,未来将会探索在图像去雾问题的研究中如何将参数动态化与网络结构动态化相结合,力争找到两者的最佳的平衡点以达到更好的去雾效果。

#### 参考文献

- [1] KHAN M F, KHAN E, ABBASI Z A. Segment dependent dynamic multi-histogram equalization for image contrast enhancement[J]. Digital Signal Processing, 2014, 25: 198-223.
- [2] LI H, XIE W, WANG X, et al. GPU Implementation of Multi-scale Retinex Image Enhancement Algorithm[C]// 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications(AICCSA). IEEE, 2016: 1-5.
- [3] 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.
- [4] 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.
- [5] REN W, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks[C]// European Conference on Computer Vision. Cham: Springer, 2016: 154-169.
- [6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [7] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 2472-2481.
- [8] ARIGELA S, ASARI V K. Enhancement of hazy color images using a self-tunable transformation function[C]// International Symposium on Visual Computing. Cham: Springer, 2014: 578-587.
- [9] BERMAN D, AVIDAN S. Non-local image dehazing[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1674-1682.
- [10] REN W, MA L, ZHANG J, et al. Gated fusion network for single image dehazing[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3253-3261.
- [11] LI B, PENG X, WANG Z, et al. Aod-net: All-in-one dehazing network[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017: 4770-4778.
- [12] ZHANG H, PATEL V M. Densely connected pyramid dehazing network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3194-3203.
- [13] YI Q, LI J, FANG F, et al. Efficient and accurate multi-scale topological network for single image dehazing[J]. IEEE Transactions on Multimedia, 2021, 24: 3114-3128.
- [14] ZHANG X, JIANG R, WANG T, et al. Single image dehazing via dual-path recurrent network[J]. IEEE Transactions on Image Processing, 2021, 30: 5211-5222.
- [15] WU H, QU Y, LIN S, et al. Contrastive learning for compact single image dehazing[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10551-10560.
- [16] 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.
- [17] LIU H, WU Z, LI L, et al. Towards Multi-Domain Single Image Dehazing via Test-Time Training [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022;5831-5840.
- [18] WU Z, NAGARAJAN T, KUMAR A, et al. Blockdrop: Dynamic inference paths in residual networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;8817-8826.
- [19] YANG L, HAN Y, CHEN X, et al. Resolution adaptive networks for efficient inference [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020;2369-2378.
- [20] YANG B, BENDER G, LE Q V, et al. Condconv: Conditionally parameterized convolutions for efficient inference[J]. Advances in Neural Information Processing Systems, 2019, 32; 1-11.
- [21] MA N, ZHANG X, HUANG J, et al. Weightnet: Revisiting the design space of weight networks[C]//European Conference on Computer Vision. Springer, Cham, 2020; 776-792.
- [22] CHEN Y, DAI X, LIU M, et al. Dynamic convolution: Attention over convolution kernels[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020; 11030-11039.
- [23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;7132-7141.
- [24] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision(ECCV). 2018;3-19.
- [25] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv:2010.11929, 2020.
- [26] GUO M H, XU T X, LIU J J, et al. Attention Mechanisms in Computer Vision: A Survey[J]. arXiv:2111.07624, 2021.
- [27] 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.
- [28] HONG M, XIE Y, LI C, et al. Distilling image dehazing with heterogeneous task imitation [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020;3462-3471.
- [29] QIN X, WANG Z, BAI Y, et al. FFA-Net: Feature fusion attention network for single image dehazing[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020;11908-11915.
- [30] 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.
- [31] 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.



**LIU Zhe**, born in 1998, master. His main research interests includes computer vision and image processing.



**LIANG Yudong**, born in 1988, associate professor. His main research interests includes computer vision, image processing, and deep learning based applications.

(责任编辑:喻黎)