

DOI:10.19651/j.cnki.emt.2519740

基于互补混合专家与一致性偏置路由的图像去雪算法^{*}

尤逸晖^{1,2} 张立新¹ 朱灵龙¹

(1. 无锡学院物联网工程学院 无锡 214105; 2. 南京信息工程大学计算机学院网络空间安全学院 南京 210044)

摘要: 单幅图像去雪是图像恢复领域的重要分支,其主要挑战在于雪粒遮挡和雪雾模糊会导致图像退化,从而影响下游视觉任务的效果。针对现有方法在特征建模单一与专家选择自适应性不足的问题,提出了一种基于互补混合专家与一致性偏置子网路由的单幅图像去雪模型 SynergyRestorer。首先,提出了互补混合专家解码器,通过专精与协作专家实现了多维度特征间的互补,增强模型的建模能力。同时,设计了一致性偏置子网路由,通过融合多源特征并引入一致性信号,动态权衡不同特征之间的协调与冲突,增强专家选择的判别能力与自适应能力。实验结果表明,所提方法在 CSD、Snow100K 和 SRRS 3 个主流基准数据集上的平均 PSNR 和 SSIM 分别达到 33.71 dB 与 0.950,验证了其在复杂雪景恢复任务中的有效性。

关键词: 深度学习;图像恢复;单幅图像去雪;混合专家模型

中图分类号: TP391.4;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Image desnowing algorithm based on a complementary mixture-of-experts and agreement-biased routing

You Yihui^{1,2} Zhang Lixin¹ Zhu Linglong¹

(1. School of Internet of Things Engineering, Wuxi University, Wuxi 214105, China;

2. School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Single-image desnowing is an important subtask in the field of image restoration. Its primary challenges lie in snow particle occlusion and snow-fog blur, which degrade image quality and affect the performance of downstream visual tasks. To address the limitations of existing methods in feature modeling and expert selection adaptability, a single-image desnowing model named SynergyRestorer was proposed. The model is based on a complementary mixture of experts and an agreement-biased sub-network routing scheme. A complementary mixture of experts decoder was designed to capture complementary information across multi-dimensional features by combining specialized and cooperative experts, thereby enhancing the model's representation capacity. An agreement-biased sub-network router was also introduced to fuse multi-source features and incorporate agreement signals. It dynamically balanced coordination and conflict among features, improving the discriminative and adaptive capacity of expert selection. Experimental results showed that the proposed method achieved an average PSNR of 33.71 dB and SSIM of 0.950 on three benchmark datasets: CSD, Snow100K and SRRS. The results validate its effectiveness in complex snowy scene restoration tasks.

Keywords: deep learning; image restoration; single-image desnowing; mixture of experts model

0 引言

我国北方,降雪是一种常见的气候现象,冬季频繁出现的降雪天气往往对后续视觉分析造成显著干扰。单幅图像去雪任务旨在从被雪粒和雪雾污染的图像中恢复出清晰、真实的背景图像。雪天图像往往同时受到多种类型的视觉

干扰,其前景和背景信息会因降雪而严重退化,图像细节与整体结构均易受到不同程度的破坏。这类复合性退化严重影响图像分类、目标检测和自动驾驶等下游任务在雪天条件下的准确性。

针对上述挑战,已有众多去雪方法被提出。早期去雪方法主要依赖于图像先验与滤波器技术,如 Zheng 等^[1]利

收稿日期:2025-08-30

^{*} 基金项目:国家自然科学基金青年科学基金(42305158)项目资助

用引导滤波器实现雪粒与背景的分隔和去除。随着深度学习的发展,数据驱动的去雪模型取得了显著进展。Liu 等^[2]提出的 DesnowNet 首次采用两阶段网络架构,分别估计雪粒掩码并恢复背景。针对雪景中的薄雾效应,Chen 等^[3]提出 JSTASR,将可微暗通道先验引入网络,有效提升了雪粒和雾状散射的去除能力。为进一步增强对复杂雪粒的多尺度和多频率特征建模,Chen 等^[4]后续提出 HDCWNet,在小波域实现精细的特征提取和雪迹分离,提升了复杂雪景下的恢复质量。Lin 等^[5]提出了轻量级的去雪网络 LMQFormer,利用拉普拉斯算子构建粗掩码先验,并通过掩码查询变换器实现对雪区域的高效建模。Shen 等^[6]提出 DTCWT 网络,通过小波分解与特征重构提升雪粒与背景的分隔效果,有效增强了图像细节的保留能力。

尽管近年来的图像去雪研究取得了显著进展,但正如前文所分析的,去雪任务具有高度复杂性。雪景图像中的干扰形式多样,既包括呈现高频特性的不透明雪粒,也包括造成低频模糊的半透明雪雾,以及覆盖范围较大的雪片遮挡。这种多维度的复杂退化特性决定了有效的去雪模型必须同时具备局部纹理感知能力、全局结构建模能力以及频域噪声分离能力。这一需求自然而然地引出了混合专家 (mixture of experts, MoE) 架构^[7]的应用,因为 MoE 能够为不同类型的特征分配最优的专家子网络。然而,现有 MoE 模型^[8]大多采用同质化专家结构,缺乏面向特征维度的建模专精性,且专家间独立工作,无法实现跨维度特征的协同建模。另一方面,传统 MoE^[8]的路由机制通常仅依赖单一特征源进行专家选择,缺乏对多源信息间一致性的判断,难以充分利用输入图像中的多源语义信息。

为解决上述问题,本文提出了一种异质化多专家协同机制和一致性偏置子网路由相结合的单幅图像去雪模型 SynergyRestorer。该模型在网络结构上主要有以下两点创新:

1) 针对传统 MoE 模型专家同质、缺乏分工与协作的问题,提出了互补混合专家解码器 (complementary mixture of experts decoder, CMED),实现了不同特征维度下的专家互补与信息融合,提升了模型对复杂雪景多维特征的适应能力。

2) 针对传统路由机制依赖单一特征、缺乏多源信息融合的问题,提出了一致性偏置子网路由 (agreement biased sub network router, ABSR),通过多源特征融合与一致性信号调控,增强了专家选择的准确性。

1 相关工作

1.1 去雪任务特性与挑战

与图像去雨、去雾等常见图像恢复任务相比,去雪任务面临更为复杂的挑战。如图 1(a)所示,雨痕呈线性分布且较为透明,图像背景仍可辨识。图 1(b)中的雾主要造成整体低对比度与模糊,但不遮挡物体结构。如图 1(c)所示,

雪粒通常不透明,且以不规则的形状分布于前景,容易遮挡物体边缘和结构信息。图 1(d)展示了雪雾天气,与雾相比,夹杂雪粒的雪雾使图像模糊与遮挡程度进一步加深。

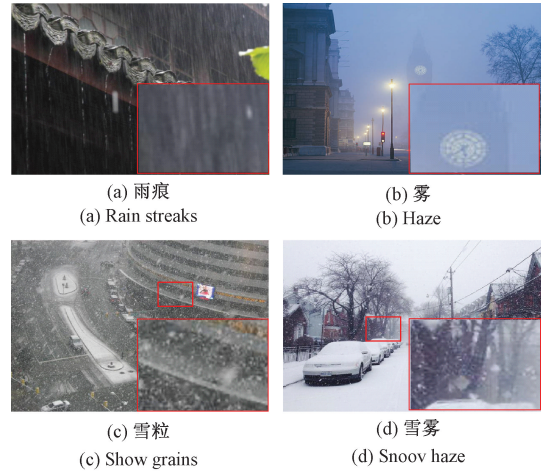


图 1 各种恶劣天气与雪天的对比

Fig. 1 Comparison of various severe weather with snow

1.2 ViT 在图像恢复中的应用

近年来, Vision Transformer (ViT)^[9]及其变体在图像去雨^[10]、去噪^[11]、去雾^[12]等底层视觉任务展现出强大潜力。Zamir 等^[13]提出的 Restormer 通过引入多分支深度卷积变换注意力模块与门控深度卷积前馈网络,显著提升了高分辨率图像恢复的效率与性能,在去雨、去模糊和去噪等多个任务上达到了新的 SOTA 表现。Chen 等^[14]提出了 HAT,通过融合通道、窗口和跨窗口注意力机制,显著提升了模型对输入像素的有效利用率和重建性能。Chen 等^[15]在系统性对比分析多种主流图像恢复网络结构的基础上,进一步提出了通用性与性能兼备的图像恢复骨干网络 X-Restormer,并在多项任务上验证了其优越的任务泛化能力。

1.3 混合专家模型

混合专家由 Jacobs 等^[7]提出,是提升深度网络容量与计算效率的经典策略。在计算机视觉领域,MoE 已广泛应用于高层任务,并逐步拓展至低层视觉领域。DRSformer^[16]引入混合专家特征补偿器,结合多专家并行和自适应门控机制,实现了多尺度特征提取优化,显著提升了复杂雨纹场景的处理能力。WM-MoE^[17]通过天气感知路由实现特征与专家的解耦分配,并融合多尺度专家以应对复杂天气。MoCE-IR^[18]则设计了嵌套式复杂度专家和复杂度感知路由策略,实现了高效的任務判别性学习与推理。

2 网络模型

2.1 网络总体结构

SynergyRestorer 结构如图 2 所示。给定一幅带有雪

痕的输入图像 $N \in \mathbf{R}^{H \times W \times 3}$, 首先通过一个卷积操作提取浅层特征 $F_{\text{shallow}} \in \mathbf{R}^{H \times W \times C}$, 其中 H 表示图像的高, W 表示图

像的宽, C 表示图像的通道数。随后, 浅层特征 F_{shallow} 被输入至一个四级的非对称编码器-解码器结构。

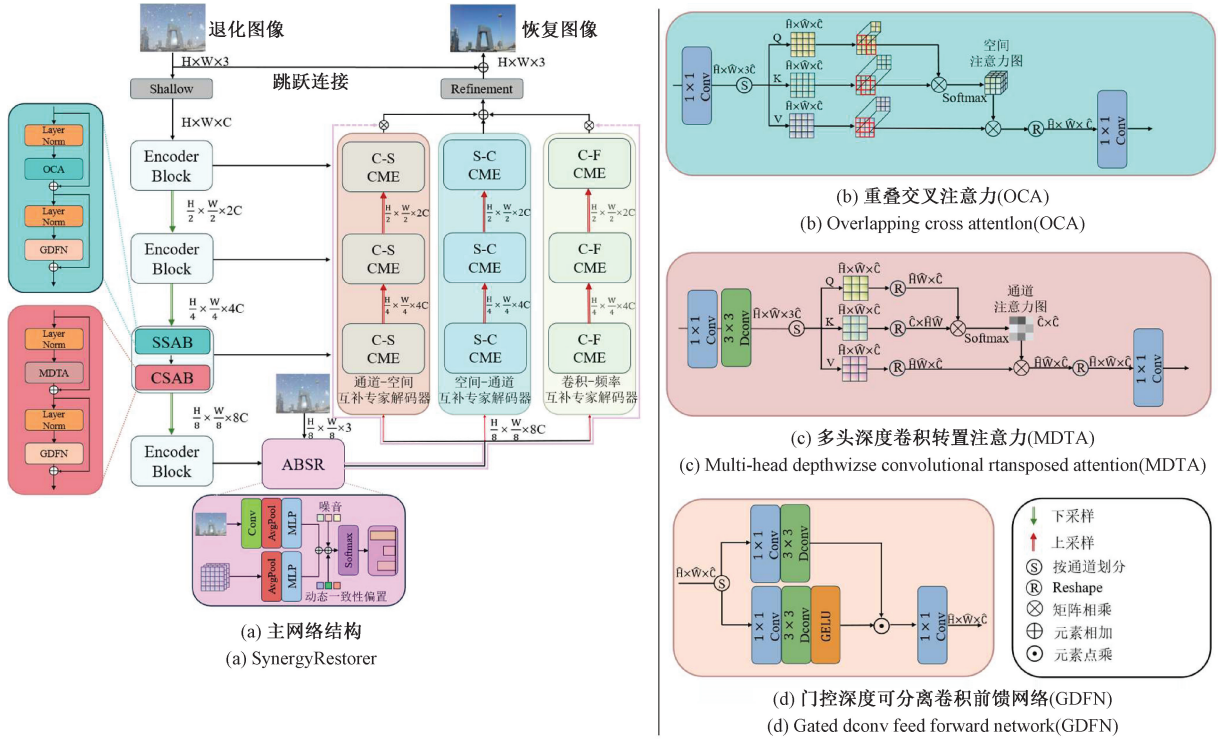


图2 网络结构图

Fig.2 Framework of overall network

编码器由级联的编码块及下采样模块构成, 其中编码块集成了通道自注意力模块(channel self-attention blocks, CSAB)、空间自注意力模块(spatial self-attention blocks, SSAB)及门控深度可分离卷积前馈网络(gated dconv feed forward network, GDFN), 同时保存各层级输出作为跳跃连接。下采样模块由 3×3 卷积和 pixel unshuffle 函数组成。特征在编码过程中逐层下采样、通道数翻倍。最终特征被送入由编码块构成的瓶颈层, 进行深层语义融合。

瓶颈层输出的特征与退化图像的浅层特征一同输入至一致性偏置子网路由, 动态选择两个不同的互补混合专家解码器并加权其输出。各 CMED 通过上采样与跳跃连接逐步恢复图像细节。上采样操作由一个 3×3 卷积与 pixel shuffle 函数实现。加权融合后的专家解码器输出经过由编码块构成的精炼模块优化。最终, 通过一个 3×3 卷积层对精炼后的特征进行处理, 生成残差图像 $R \in \mathbf{R}^{H \times W \times 3}$, 并将其与退化的输入图像 N 相加, 得到最终的恢复图像: $C = N + R$ 。

2.2 编码器模块

雪在图像中通常表现为局部异质分布的雪粒、雪条, 以及覆盖全局的雾状散射效应, 不仅遮蔽图像细节, 还显著降低对比度与清晰度。针对上述特性, 空间注意力机制

有助于强化模型对雪粒等局部扰动区域的感知能力, 而通道注意力机制则有助于建模受散射影响的全局语义特征, 从而实现更有效的雪痕建模与图像恢复。

本文受 X-Restormer 模型^[15] 启发, 该模型通过将 Restormer^[13] 中部分转置自注意力块替换为空间自注意力块, 显著提升了网络的空间信息交互能力。因此, 本文采用空间自注意力块模块(spatial self-attention block, SSAB)与通道自注意力模块串联的方式, 构成 SynergyRestorer 的编码器模块。

图 2(b) 展示了重叠交叉注意力(overlapping cross-attention, OCA)的具体结构。OCA 最早由 HAT^[14] 提出, 与 Swin Transformer^[19] 的传统窗口自注意力不同, 其通过构造查询与键值对之间的重叠窗口, 实现更广泛的局部特征交互。

具体而言, 对于输入特征图 $X \in \mathbf{R}^{\hat{H} \times \hat{W} \times \hat{C}}$, 首先通过 1×1 卷积分别生成 3 个投影特征图, 即查询特征图 X_Q 、键特征图 X_K 与值特征图 X_V 。随后, X_Q 被划分为大小为 $M \times M$ 的非重叠窗口, 而 X_K 与 X_V 则划分为大小为 $M_o \times M_o$ 的重叠窗口, 其中重叠窗口尺寸定义为:

$$M_o = (1 + \theta) \times M \quad (1)$$

其中, θ 为窗口重叠率, 本文设定为 0.5, 即重叠窗口尺寸为原窗口尺寸的 1.5 倍。

与传统基于移动窗口的自注意力 (shifted window multihead self-attention, SW-MSA) 相同, OCA 在每个窗口内计算空间自注意力, 注意力的计算公式为:

$$\text{Attention}_{\text{spatial}}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}} + B\right)\mathbf{V} \quad (2)$$

其中, $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 分别表示查询、键、值矩阵, d 表示注意力头的维度大小, B 表示相对位置编码。

图 2(c) 展示了多头深度卷积转置注意力 (multi-head depthwise convolutional transposed attention, MDTA) 的具体结构。对于输入特征图 $\mathbf{X} \in \mathbf{R}^{\hat{h} \times \hat{w} \times \hat{c}}$, 首先采用 1×1 卷积与 3×3 深度卷积分别提取逐像素的跨通道特征与空间局部上下文, 随后将特征图在通道上划分为特征矩阵查询 \mathbf{Q} 、键 \mathbf{K} 和值 \mathbf{V} 。接着, 将查询 \mathbf{Q} 和键 \mathbf{K} 特征矩阵 Reshape, 以通道维度计算转置注意力图 $\mathbf{A} \in \mathbf{R}^{\hat{c} \times \hat{c}}$, 并与值 \mathbf{V} 在通道纬度进行矩阵相乘。注意力的计算过程定义为:

$$\text{Attention}_{\text{channel}}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{K} \cdot \mathbf{Q}}{\alpha}\right)\mathbf{V} \quad (3)$$

其中, α 为可学习的尺度参数, 用于控制点积结果的数值大小, 以提升网络训练的稳定性。

图 2(d) 展示了 GDFN 模块的具体结构。对于输入特征图 $\mathbf{X} \in \mathbf{R}^{\hat{h} \times \hat{w} \times \hat{c}}$ 首先在通道维度划分为两路, 分别通过 1×1 卷积结合 3×3 深度卷积提取通道与空间上下文信息。其中一路经 GELU 激活增强非线性表达, 另一路保持线性。两路特征逐元素相乘, 实现动态门控以自适应强化重要信息。最后, 通过一个 1×1 卷积将特征通道数恢复至原始维度。

2.3 一致性偏置子网路由

针对传统专家路由仅依赖单一特征、难以兼顾多源特征的问题, 本文提出了 ABSR。如图 2(a) 所示, ABSR 同时接收深层特征嵌入与全局图像嵌入两路信息, 经卷积和池化提取空间及通道上下文特征, 随后通过多层感知器映射为专家评分 logits。ABSR 显式引入一致性信号作为偏置, 通过对两路评分结果逐元素相乘, 动态感知并处理两路评分之间的协调与冲突。当两路评分趋于一致时, 一致性信号被增强, 反之则被抑制, 从而提升路由机制对多源信息的自适应协调能力。一致性信号计算为:

$$\text{Agreement Signal} = \tanh\left(\frac{\text{logits}_{\text{deep}}}{2}\right) \times \tanh\left(\frac{\text{logits}_{\text{img}}}{2}\right) \quad (4)$$

一致性信号作为动态偏置项, 与随机噪声和专家评分相加, 最终经 Softmax 获得动态路由分布。

在训练过程中, 为同时兼顾专家资源分配的均衡性与整体计算效率, ABSR 引入了在 MoCE-IR^[18] 中提出的辅助损失函数。专家路由不仅关注每个专家在当前批次中被选中的频率, 还考虑其对应的计算复杂度。首先定义第 i 个专家在批次 \mathcal{B} 中的重要性为其被分配的总门控权重之

和, 记作:

$$\text{Imp}_i(x) = \left(\sum_{x \in \mathcal{B}} \text{Softmax}(Wx)_i\right) \cdot b_i \cdot \tau \quad (5)$$

其中, W 为路由器的权重, $\text{Softmax}(Wx)_i$ 表示第 i 个专家在输入 x 中的门控得分, 系数 τ 为缩放超参数, b_i 为该专家的复杂度偏置, 定义为:

$$b_i = \frac{p_i}{p_{\text{sum}}} \quad (6)$$

其中, p_i 表示第 i 个专家的可学习参数数量, p_{sum} 为所有专家的参数量总和。最终, 重要性损失与负载均衡损失^[18]共同构成了辅助损失函数:

$$L_{\text{aux}}(x) = \frac{1}{2}\text{CV}(\text{Imp}(x))^2 + \frac{1}{2}\text{CV}(\text{Load}(x))^2 \quad (7)$$

其中, $\text{CV}(\cdot)$ 表示变异系数 (coefficient of variation, CV), 衡量专家间分配的均衡程度。

2.4 编码器模块

为解决传统混合专家模型内部设计同质化、缺乏特征维度互补的问题, 本文提出了多解码器协同的去雪解码器框架, 通过动态路由机制以实现对输入特征的自适应处理。该框架集成了 3 个并行且各具侧重的 CMED, 分别建模通道注意力、空间注意力和多尺度卷积特征。每个互补混合专家 (complementary mixture of experts, CME) 内部通过专精专家与协作专家的协同分工实现特征维度互补。

如图 3 所示, CME 首先接收来自编码器的特征, 由相应的专精专家, 包括 CSAB、SSAB 或多尺度模块 (multi-scale block, MSB) 提取专精特征 x_s 。MSB 通过多分支卷积并行处理, 专注于捕捉多尺度空间结构。专精特征 x_s 一方面通过标准的专家路由器动态激活 Top-K 个多尺度互补专家 (multi-scale complementary experts, MSCE), 另一方面送入协作专家生成互补调制特征 x_c 。协作专家根据专精专家类型动态配置, 实现特征维度互补。当专精专家为 MSB 时, 协作专家采用频率卷积模块对专精特征 x_s 先执行二维快速傅里叶变换 (two-dimensional fast Fourier transform, FFT-2D), 进入频域后通过卷积对不同频率分量进行滤波操作, 随后经逆傅里叶变换 (inverse two-dimensional fast Fourier transform, InvFFT-2D) 转回空间域, 生成互补调制 x_c , 以弥补专精专家中卷积操作在全局频率信息感知上的不足。类似地, 若专精专家为 CSAB, 则协作专家相应配置为 SSAB (反之亦然), 从而实现通道与空间语义的有效互补。

每个 CME 包含 4 个结构一致但卷积核尺寸不同 ($1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7$) 的 MSCE。每个被激活的 MSCE 接收专精特征 x'_s 和协作专家生成的互补调制特征 x_c 作为输入。专精特征 x'_s 经过线性层和卷积提取深度特征后, 与互补调制特征 x_c 生成的调制权重逐元素相乘, 形成调制特征。该特征经线性层恢复通道维度后, 与原始残差相加, 得到该 MSCE 的输出。所有激活的专家输出根据门控权重加权融合, 并与调制信号 x_c 叠加后, 经线性层生成最终

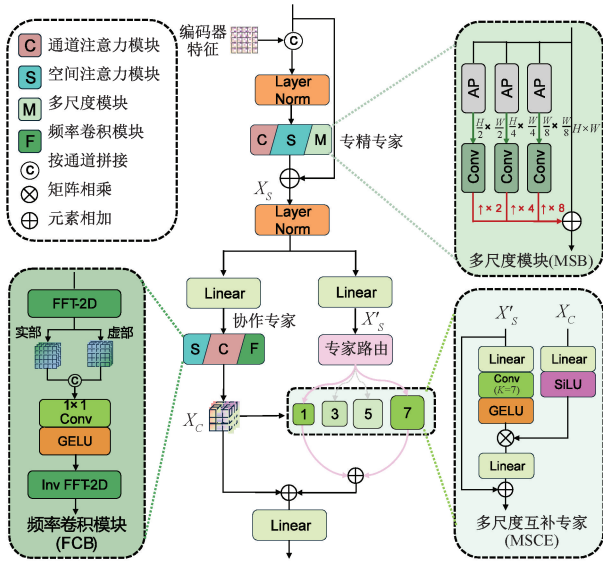


图3 互补混合专家

Fig.3 Complementary mixture of experts

特征 x_{out} 。

2.5 解码器网络

本文采用复合损失函数来优化网络,复合损失函数由 L_1 损失(平均绝对误差损失)和辅助损失(ABSR 辅助损失)两部分组成。 L_1 损失用于约束重建图像与无雪参考图像在像素空间的一致性,辅助损失则用于动态平衡专家路由分配。复合损失函数定义为:

$$L_{total} = L_1 + L_{aux} \quad (8)$$

L_1 损失定义为:

$$L_1 = \frac{1}{N} \sum_{i=1}^N |y_i - g_i| \quad (9)$$

其中, y_i 表示模型输出的第 i 个像素, g_i 表示无雪参考图像的第 i 个像素, N 为像素总数。

3 实验设置与结果分析

3.1 数据集与评价指标

本文在多个公开基准数据集上进行了去雪实验,包括 Snow100K^[2]、SRRS^[3] 和 CSD^[4]。Snow100K 数据集包含 10 万张合成雪景图像(基于大、中、小 3 种尺寸的雪粒掩膜生成,划分为 Snow100K-L, Snow100K-M, Snow100K-S 3 个子集)和 1 329 张真实雪景图像。SRRS 数据集包含 15 000 张合成雪景图像和 1 000 张真实世界雪景图像。CSD 数据集包含 10 000 张合成雪图像,考虑了雪粒的多样性并应用高斯模糊以增强真实感。依据先前研究的实验设置^[4],本文对这 3 个数据集分级随机划分出 8 000 对图像进行训练,2 000 对图像进行测试。

在性能评估方面,本文采用图像恢复领域广泛使用的峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似度指数(structural similarity index measure, SSIM)作为评

估指标。两项指标均在 YCbCr 色彩空间的 Y 通道上进行计算,以衡量模型在图像视觉质量与结构一致性方面的表现。

3.2 实验环境及设置

本文提出的 SynergyRestorer 网络采用非对称四级级编码器-解码器设计。编码器包含 4 个层级,各层级自顶向下的编码器块数量依次设定为 2、4、4、4。每个专家解码器包含 3 个层级,分别配置 2、4、4 个解码器模块。在 OCA 和 MDTA 模块中,通道与空间的注意力头数量均随网络深度逐级设置为 1、2、4、8。其中,OCA 模块的计算窗口尺寸固定为 8,窗口重叠率设定为 0.5。网络在不同层级的特征图通道数则相应设置为 32、64、128 及 256。在 CME 与 ASBR 中,Top-K 均设为 2。

在所有实验中,各数据集的训练轮数均为 200 轮。优化过程中使用 AdamW 优化器 ($\beta_1 = 0.9, \beta_2 = 0.999$, weight decay=0.0001),初始学习率为 2×10^{-4} ,并通过余弦退火策略逐步降低至 1×10^{-6} 。此外,Patch size 设置为 256×256 像素,Batch size 设置为 2。训练过程运行于 Ubuntu 20.04.6 LTS 操作系统,采用 Nvidia GeForce RTX 4070 Ti Super 进行模型训练。模型实现基于 PyTorch 框架。在数据预处理中,采用水平翻转和垂直翻转进行数据增强。

3.3 合成雪图对比

为验证本文方法的有效性,本文将所提出的 SynergyRestorer 与多种主流单幅图像去雪算法进行了定量与可视化对比,包括 MGF^[1]、DesnowNet^[2]、InvDN^[20]、JSTASR^[3]、DTCWT^[6]、DesnowGAN^[21]、DDMSNet^[22]、HDCWNet^[4]、SmartAssign^[23]、DAN-Net^[24]、InvDSNet^[25]、DARDNet^[26]、LMQFormer^[5] 和 MSP-former^[27]。定量结果如表 1 所示。

在 Snow100K、SRRS 和 CSD3 个主流数据集上,SynergyRestorer 均表现出优越的性能,平均 PSNR 达 33.71 dB,平均 SSIM 达 0.950,在所有对比方法中排名第一。与当前综合性能领先的 MSP-former 相比,SynergyRestorer 在平均 PSNR 上提升了 1.06 dB,这在图像恢复任务中属于显著增益。在更具挑战性的 Snow100K 数据集中,SynergyRestorer 同样以 34.66 dB 的 PSNR 取得最优结果。特别是在中雪(M)与大雪(L)这类严重遮挡场景下,其 PSNR 分别达到了 35.32 与 36.91 dB,明显优于其他对比模型。

为客观评估算法性能,图 4 展示了合成雪图去雪对比结果。针对车牌区域的对比分析表明,JSTASR 和 HDCWNet 均存在明显的雪痕残留或车牌字符模糊的问题,而 SynergyRestorer 能够清晰地还原字符细节,无雪痕残留,显著提升了图像的可读性。针对机车纹理区域的对比分析表明,LMQFormer 能较好地去除大部分雪痕,但在机车金属表面的光影还原方面仍显不足,尤其是高光区域

表 1 合成数据集上的评估指标
Table 1 Evaluation metrics on synthetic dataset

算法	Snow100K			SRRS		CSD		Average	
	PSNR	SSIM	S/M/L	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MGF	22.41	0.77	24.32/22.99/19.95	15.78	0.74	13.98	0.67	17.39	0.727
DesnowNet	30.11	0.93	32.33/30.86/27.16	20.38	0.84	20.13	0.81	23.54	0.860
InvDN	27.99	0.81	28.83/28.44/26.74	26.49	0.88	27.46	0.86	27.31	0.850
JSTASR	28.59	0.86	31.40/29.11/25.32	25.82	0.89	27.96	0.88	27.46	0.877
DTCWT	28.39	0.90	—	27.52	0.88	29.89	0.91	28.60	0.897
DesnowGAN	31.11	0.95	33.43/31.87/28.06	—	—	27.09	0.88	29.10	0.915
DDMSNet	32.03	0.92	34.34/32.89/28.85	27.03	0.90	28.79	0.91	29.28	0.910
HDCWNet	31.10	0.95	32.84/31.75/28.75	27.78	0.92	29.06	0.91	29.31	0.927
SmartAssign	29.45	0.92	—	30.53	0.93	32.50	0.95	30.83	0.933
DAN-Net	32.48	0.96	—	29.34	0.95	30.82	0.95	30.88	0.950
InvDSNet	32.41	0.93	34.39/33.17/29.69	29.25	0.95	31.85	0.96	31.17	0.947
DARDNet	32.03	0.95	—	30.08	0.95	32.16	0.96	31.42	0.950
LMQFormer	31.96	0.95	33.81/32.67/29.41	31.04	0.87	32.64	0.96	31.88	0.927
MSP-former	33.43	0.96	—	30.76	0.95	33.75	0.96	32.65	0.956
本文	34.66	0.96	31.77/35.32/36.91	31.22	0.94	35.26	0.95	33.71	0.950

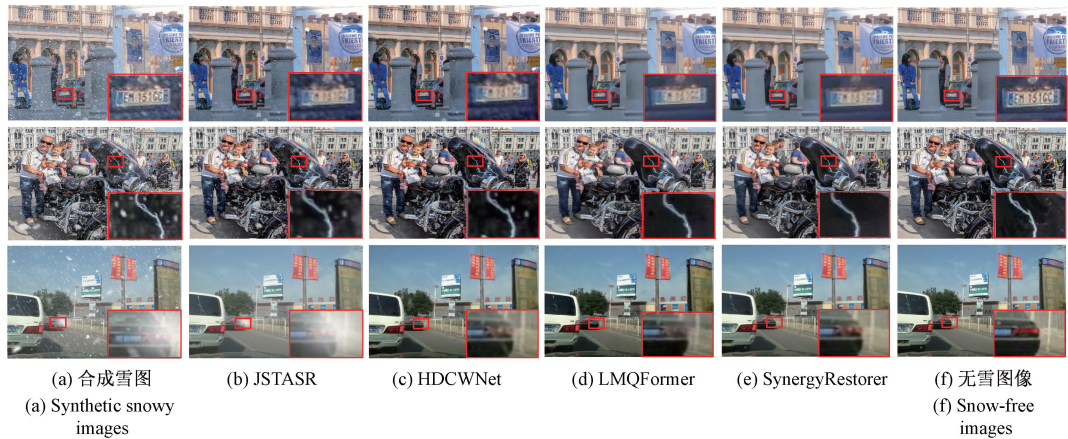


图 4 合成雪图去雪对比结果

Fig. 4 Comparison results of snow removal on synthetic snow images

边缘模糊。相较之下, SynergyRestorer 不仅去除了雪痕, 还准确还原了机车金属表面的光泽变化与纹理细节, 细节更清晰自然。

3.4 真实雪图对比

图 5 展示了不同模型在真实雪图上的去雪结果。在人物区域, HDCWNet 与 SmartAssign 虽然能去除部分雪粒, 但人脸与衣物纹理仍显模糊, 细节难以辨识; 相比之下, SynergyRestorer 能够有效去除雪痕, 同时清晰恢复面部轮廓与衣物纹理, 整体真实感更强。在招牌区域, HDCWNet 结果中仍残留大量雪痕, 文字边缘模糊; SmartAssign 和 LMQFormer 虽去除了大量雪痕, 但模糊问题依旧存在。而 SynergyRestorer 则在去除雪痕的同

时, 较清晰地恢复了被遮挡的文字与结构细节, 显著提升了图像的可读性。

为进一步验证在真实极端场景下的鲁棒性, 于 Snow100K 真实子集中选取两类极端样例: 大雪场景以及夜间弱光且密集飘雪。

在大雪场景下, HDCWNet 与 SmartAssign 仅能去除部分雪雾, 整体背景仍显模糊; LMQFormer 虽在一定程度上减弱了雾感, 但人物边缘与远景轮廓依然不够清晰。而 SynergyRestorer 不仅能同时消除薄雾与细小雪粒, 还能有效保留衣物褶皱等细节。

在夜间弱光样本中, 所有对比模型均能有效去除大部分明显雪粒, 然而在放大观察背景细节时可以发现, 其他



图5 真实雪图去雪对比结果

Fig. 5 Comparison results of snow removal on real snow images

模型对远处暗部区域的细微雪痕仍存在残留。相比之下, SynergyRestorer 能够彻底清除这些难以察觉的雪痕。

3.5 主观质量评测

为评估不同去雪方法的视觉感知效果,开展了主观质量评测(mean opinion score, MOS)实验。从 Snow100 K 真实子集中选取 30 幅真实雪景图像,涵盖多种雪量与不同环境。对比方法包括 HDCWNet、SmartAssign、LMQFormer 3 种主流算法。通过调查问卷的形式向 22 名受试者进行盲评,从细节恢复、伪影抑制和整体质量 3 个维度进行 5 分制评分。

主观评价结果如表 2 所示。SynergyRestorer 在 3 个维度均取得最高得分,整体质量 MOS 值为 3.85,细节恢复与伪影抑制分别达 3.89 和 3.80。统计检验证实本方法显著优于其他对比方法。在大雪和中雪场景中,本方法能有效保持边缘细节并抑制伪影,仅在高饱和区域存在轻微色度偏差。主客观结果一致表明,SynergyRestorer 在视觉感知质量方面具有明显优势。

表2 真实雪景主观质量评测实验结果

Table 2 MOS results on real snow scenes

算法	细节恢复	伪影抑制	整体质量
HDCWNet	3.72±0.58	3.70±0.60	3.71±0.54
SmartAssign	3.78±0.55	3.71±0.57	3.75±0.52
LMQFormer	3.86±0.51	3.74±0.53	3.80±0.49
本文	3.89±0.47	3.80±0.52	3.85±0.45

3.6 模型效率对比

为评估模型效率,本文在计算复杂度与推理速度方面

进行了对比分析。如表 3 所示,SynergyRestorer 的参数量(13.07 M)略高于对比模型。但在计算量方面,其 FLOPs 为 19.29 G,仅高于轻量型方法 DARDNet(12.24 G),而显著低于 InvDSNet 与 DaN-Net(降低约 69%),反映出整体计算负担处于合理范围。

表3 计算复杂度与推理速度对比结果

Table 3 Comparison results of computational complexity and runtimes

算法	Parameters/M	Runtimes/s	FLOPs/G
InvDSNet	6.94	0.031	62.29
DaN-Net	2.73	—	62.72
DARDNet	5.48	—	12.24
LMQFormer	2.18	0.042	22.47
本文	13.07	0.068	19.29

在推理速度方面,SynergyRestorer(0.068 s)略慢于 InvDSNet(0.031 s)和 LMQFormer(0.042 s),说明运行效率仍有提升空间。尽管如此,SynergyRestorer 在参数量适度增加的同时实现了精度显著提升,较好平衡了性能与复杂度。

3.7 消融实验与分析

为全面评估本文提出的 SynergyRestorer 的核心设计的有效性,在 SRRS 数据集上设计了 4 组消融实验。实验过程中,Patch Size 均设为 128×128。

表 4 展示了 ABSR 消融实验结果。结果表明,移除图像嵌入输入或一致性信号后,PSNR 和 SSIM 均出现不同

程度的下降。相比之下,完整的 ABSR 获得了最佳性能,表明多源特征融合和一致性信号约束对提升专家路由效果具有显著贡献。

表 4 一致性偏置子网路由消融实验结果

Table 4 Ablation results of the ABSR

方法配置	PSNR	SSIM
移除图像嵌入输入	30.97	0.942
移除一致性信号	31.01	0.942
ABSR(本文)	31.17	0.946

表 5 展示了 ABSR 中 Top-K 专家激活策略的有效性。结果表明,随着激活解码器专家数量由 1 增加到 2,PSNR 和 SSIM 显著提升,而进一步增加至 3 时,性能提升趋于饱和且计算量显著增加。故本文采用 $K=2$,兼顾了去雪效果与计算效率。

表 5 ABSR 专家选择策略 (Top-K) 消融实验

Table 5 Ablation results of the ABSR expert selection strategy (Top-K)

Top-K 配置	PSNR	SSIM	FLOPs/G
$K=1$	30.86	0.933	14.57
$K=2$ (本文)	31.17	0.946	19.29
$K=3$	31.18	0.946	23.92

表 6 展示了互补混合专家模块 (CME) 的消融实验结果。将卷积核尺寸统一为 (3,3,3,3) 时,性能略降,表明多尺度配置 (1,3,5,7) 有助于建模不同尺度雪痕。移除协作专家或采用同质结构 (即无专精-协作分工) 均显著降低性能,验证了专精-协作分工和维度互补策略对去雪效果的重要性。

表 7 对比了不同解码器配置下的性能,仅采用单一类

表 6 互补混合专家消融实验结果

Table 6 Ablation results of the complementary mixture of experts

方法配置	PSNR	SSIM
卷积核尺寸连续 (3,3,3,3)	31.12	0.940
卷积核尺寸递增 (1,3,5,7)	31.17	0.946
移除协作专家	30.88	0.934
同质化专家设计	30.94	0.940
完整 CME 设计(本文)	31.17	0.946

型解码器时,模型去雪性能均低于互补策略。采用解码器互补策略后,显著提升了 PSNR 和 SSIM,表明多解码器互补策略的有效性。

表 7 不同解码器配置消融实验结果

Table 7 Ablation results of different decoder configurations

方法配置	PSNR	SSIM
单一 CA 解码器	30.67	0.936
单一 SA 解码器	30.34	0.933
单一 MSF 解码器	30.46	0.931
多解码器互补策略(本文)	31.17	0.946

为了进一步研究不同专家在通道、空间与频率域的职责分工,对各专家模块进行消融并可可视化输出。在推理阶段,通过将某一专家分支的前向输出暂时置零,并将屏蔽前后的输出之差进行分位数归一化与 Top-K 叠加,最终得到各专家的贡献热图。在热图中,红/黄色高亮区域表征了图像复原对被屏蔽专家的高度依赖性。

CA 通道专家解码器的贡献热力图如图 6(a) 所示,其高亮区域主要集中在大块雪团以及高饱和度或强对比度区域。结果表明,通道专家更侧重于处理图像中具有显著色度或对比度的局部细节。

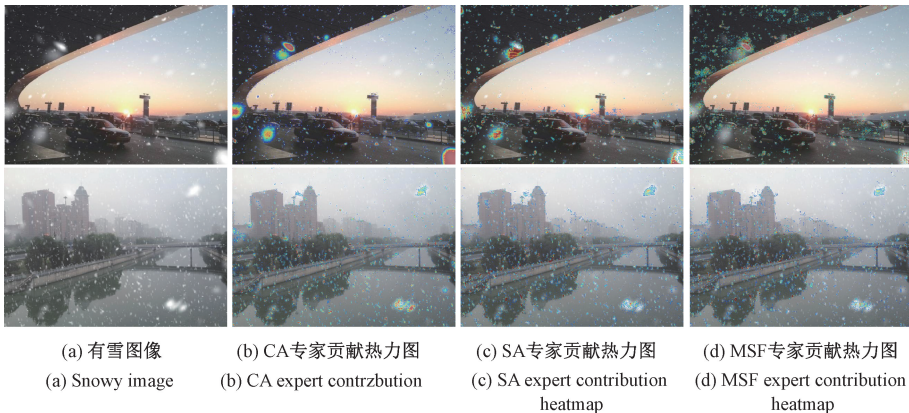


图 6 专家贡献热力图可视化

Fig. 6 Expert contribution heatmap visualization

SA 空间专家解码器的贡献热力图如图 6(b) 所示,与

通道专家不同,空间专家的响应更集中于零星散布的细小

雪粒。这得益于其内部 OCA 模块所具备的较大感受野,从而赋予其更强的空间结构感知能力。

MSF 专家解码器的贡献热力图如图 6(c)所示,其特征表现为覆盖范围广且响应细节密集。高亮区域不仅密集出现在较大雪粒与锐利边缘处,在背景区域也呈现大范围的中低强度响应。这说明该专家主要负责恢复全局高频细节并抑制噪声。

4 结 论

本研究提出了一种基于互补混合专家模型与一致性偏置子网路由机制的单幅图像去雪模型 SynergyRestorer。通过设计互补混合专家解码器,实现专精与协作专家的多维特征互补,显著增强了模型的建模能力。同时,引入了一致性偏置子网路由机制,通过融合多源特征并引入一致性信号偏置,实现特征间协调与冲突的动态平衡,提高了专家选择的准确性。实验结果表明,SynergyRestorer 在多个公开数据集上优于现有方法。但本文方法在部分边缘场景中仍存在计算开销优化空间,未来将进一步探索轻量化版本和部署优化等方向。

参考文献

- [1] ZHENG X H, LIAO Y H, GUO W, et al. Single-image-based rain and snow removal using multi-guided filter [C]. International Conference on Neural Information Processing, 2013, 8228: 258-265.
- [2] LIU Y F, JAW D W, HUANG S C, et al. DesnowNet: Context-aware deep network for snow removal[J]. IEEE Transactions on Image Processing, 2018, 27(6): 3064-3073.
- [3] CHEN W T, FANG H Y, DING J J, et al. JSTASR: Joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal[C]. European Conference on Computer Vision, 2020: 754-770.
- [4] CHEN W T, FANG H Y, HSIEH C L, et al. All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss[C]. 2021 IEEE/CVF International Conference on Computer Vision, 2021: 4176-4285.
- [5] LIN J H, JIANG N F, ZHANG ZH T, et al. LMQFormer: A laplace-prior-guided mask query transformer for lightweight snow removal[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(11): 6225-6235.
- [6] SHEN Z H, XUAN Y, YANG Q Y. Wavelet-enhanced desnowing: A novel single image restoration approach for traffic surveillance under adverse weather conditions [J]. IEEE Access, 2025, 13: 135407-135419.
- [7] JACOBS R A, JORDAN M I, NOWLAN S J, et al. Adaptive mixtures of local experts [J]. Neural Computation, 1991, 3: 79-87.
- [8] FEDUS W, ZOPH B, SHAZEER N. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity [J]. Journal of Machine Learning Research, 2022, 23 (1): 1532-4435.
- [9] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[J]. ArXiv preprint arXiv, 2010.11929, 2020.
- [10] 韩冉,曾广森,王荣杰. 基于残差块网络的图像去雨算法[J]. 仪器仪表学报, 2021, 41(8):175-182.
HAN R, ZENG G M, WANG R J. Image derain algorithm based on residual block network[J]. Chinese Journal of Scientific Instrument, 2021, 41(8): 175-182.
- [11] 周先春,吕梦楠,芮旸,等. 基于注意力机制的双卷积图像去噪网络[J]. 电子测量与仪器学报, 2025, 39(2):60-71.
ZHOU X CH, LYU M N, RUI Y, et al. Image denoising using dual convolutional neural network with attention mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39 (2): 60-71.
- [12] 武丽,俞俊,张征浩,等. 基于多残差和多重特征融合的去雾算法[J]. 国外电子测量技术, 2024, 43(6):12-21.
WU L, YU J, ZHANG ZH H, et al. Fogremoval algorithm based on multiple residuals and multiple feature fusion [J]. Foreign Electronic Measurement Technology, 2024, 43(6): 12-21.
- [13] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: Efficient transformer for high-resolution image restoration[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 5718-5729.
- [14] CHEN X Y, WANG X T, ZHOU J T, et al. Activating more pixels in image super-resolution transformer [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 22367-22377.
- [15] CHEN X Y, LI ZH Y, PU Y D, et al. A comparative study of image restoration networks for general backbone network design[C]. European Conference on Computer Vision, 2024: 74-91.

- [16] CHEN X, LI H, LI M Q, et al. Learning a sparse transformer network for effective image deraining[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 5896-5905.
- [17] LUO Y, ZHAO R, WEI X B, et al. WM-MoE: Weather-aware multi-scale mixture-of-experts for blind adverse weather removal[J]. ArXiv preprint arXiv, 2303.13739, 2023.
- [18] ZAMFIR E, WU Z W, MEHTA N, et al. Complexity experts are task-discriminative learners for any image restoration[C]. 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2025: 12753-12763.
- [19] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical vision transformer using shifted windows[C]. 2021 IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [20] LIU Y, QIN ZH Y, ANWAR S, et al. Invertible denoising network: A light solution for real noise removal [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13360-13369.
- [21] JAW D W, HUANG S C, KUO S Y. DesnowGAN: An efficient single image snow removal framework using cross-resolution lateral connection and GANs [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(4): 1342-1350.
- [22] ZHANG K H, LI R Q, YU Y J, et al. Deep dense multi-scale network for snow removal using semantic and depth priors[J]. IEEE Transactions on Image Processing, 2021, 30: 7419-7431.
- [23] WANG Y L, MA CH, LIU J ZH. SmartAssign: Learning a smart knowledge assignment strategy for deraining and desnowing [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 3677-3686.
- [24] 石明珠, 糟斌, 苏宇皓, 等. 联合图像通道与像素双注意力机制精细化单幅图像去雪[J]. 光学精密工程, 2024, 32(12): 1954-1964.
- SHI M ZH, ZAO B, SU Y H, et al. Dual attention refinement for single image desnowing[J]. Optics and Precision Engineering, 2024, 32(12): 1954-1964.
- [25] QUAN Y H, TAN X H, HUANG Y, et al. Image desnowing via deep invertible separation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(7): 3133-3144.
- [26] CHEN ER K, CHEN S X, YE T, et al. Degradation-adaptive neural network for jointly single image dehazing and desnowing [J]. Frontiers of Computer Science, 2024, 18(2): 182707.
- [27] CHEN S X, YE T, LIU Y, et al. MSP-Former: Multi-scale projection transformer for single image desnowing [C]. ICASSP 2023—IEEE International Conference on Acoustics, 2023: 1-5.

作者简介

尤逸晖, 硕士, 主要研究方向为计算机视觉、图像恢复。

E-mail: 2024124925792@nuist.edu.cn

张立新(通信作者), 硕士, 正高级工程师, 主要研究方向为图像处理与模式识别。

E-mail: zhanglx@cxwu.edu.cn

朱灵龙, 博士, 讲师, 主要研究方向为图像处理与分析。

E-mail: llzhu@cxwu.edu.cn