

航天器对小行星抵近探测过程中的视频超分辨率方法

陈宇涵^{1,2} 陈宇¹ 邓丽¹ 陈实¹

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 100049)

摘要 针对抵近探测中平台运动、抖动导致的动态图像序列模糊及分辨率低的问题,提出了一种基于 BasicVSR++ 的视频超分辨率方法。通过引入空间和通道注意力机制强化细节特征提取,结合共享投射权重、多组机制和采样点调制优化对齐模块,弥补正则卷积在长距离依赖与自适应空间聚集的不足。采用下采样与低通滤波器结合的方式,减少高频成分,提升抗图像抖动鲁棒性,同时引入新上采样模块,通过融合局部与全局特征生成自适应上采样核,进一步扩展感受野,以更好地恢复全局结构并重建细节。仿真实验结果显示,本文提出的方法在峰值信噪比 (PSNR) 和结构相似性 (SSIM) 指标上,分别比原始方法提高了 2.2% 和 2.1%,验证了本文方法在抵近探测图像序列超分辨率重建质量提升方面的有效性。

关键词 抵近探测, 视频超分辨率, 注意力机制, 采样点对齐

中图分类号 TP39

Video Super-resolution Method for Spacecraft Approaching and Detecting Asteroids

CHEN Yuhan^{1,2} CHEN Yu¹ DENG Li¹ CHEN Shi¹

1(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

2(University of Chinese Academy of Sciences, Beijing 100049)

Abstract In the imaging process of approach detection, dynamic image sequences often have problems such as image blur and insufficient resolution due to platform movement and jitter. This paper studies the super-resolution of image sequences in the process of approach detection and proposes a video super-resolution method based on BasicVSR++. By introducing spatial and channel attention mechanisms to enhance the model's ability to extract detail features, combined with shared projection weights, multi-group mechanisms and sampling point modulation, the effect of the alignment module is improved. While improving the network feature extraction capability, it makes up for the shortcomings of regular convolution in long-distance dependency and adaptive spatial aggregation. At the same time, downsam-

2025-01-06 收到原稿, 2025-05-22 收到修定稿

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

pling is combined with a low-pass filter to reduce the high-frequency components of the image, which improves the robustness of the model to slight image jitter. In addition, a new upsampling module is introduced to combine local and global features, generate an adaptive upsampling kernel to expand the receptive field, and better restore the global structure and reconstruct details. The simulation experimental results show that the proposed method improves the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) indicators by 2.2% and 2.1% respectively compared with the original method, which proves the effectiveness of the method proposed in this paper in improving the quality super-resolution reconstruction of the image sequence in close proximity.

Key words Approaching detection, Video super-resolution, Attention mechanism, Sampling point alignment

0 引言

成像探测是小行星探测的重要手段之一。航天器飞抵过程中获得的视频图像珍贵且独特。成像区域随抵近过程快速变化, 视场内目标有动态位移, 并且由于平台运动和抖动、太阳光照变化等因素影响, 图像存在噪声、运动模糊、分辨率不够高等问题。

近年来, 超分辨率技术取得了一定进展, 特别是在深度学习领域^[1]。图像超分辨率是指由一幅低分辨率图像或图像序列恢复出高分辨率图像的技术^[2,3]。高分辨率 (High Resolution, HR) 表示图像具有高像素密度, 可以提供更多的细节, 这些细节往往在应用中起到关键作用。根据分类准则的不同, 可以将图像超分辨率重建技术划分为不同的类别。从输入的低分辨率图像数量角度来看, 可以分为单帧图像的超分辨率重建和多帧图像 (视频) 的超分辨率重建^[4]; 从变换空间角度来看, 可以分为频域超分辨率重建和空域超分辨率重建^[5]; 从重建算法角度来看, 可以分为基于退化模型和基于学习的超分辨率重建^[6,7]。基于学习的超分辨率经典算法 (SRCNN^[1], EDSR^[8], SRGAN^[9] 等) 在静态图像的超分辨率重建方面表现优异, 但是在应对动态场景时, 常常存在信息丢失、图像模糊等问题^[4], 无法适应高速动态目标的细节重建。此外, 基于时间的可变形卷积、金字塔级联网络等技术, 例如 EDVR^[10] 和 TDAN^[11], 虽然对动态视频的处理有一定改善, 但在应对复杂光照变化和运动模糊时依然存在不足^[4,12]。

本文针对抵近探测过程中获得的图像序列开展超分辨研究, 基于 BasicVSR++^[13] 进行适配改动, 提出一种适用于航天器抵近探测任务的视频超分辨方法, 提高图像的分辨率, 并增强模型的鲁棒性。

1 抵近探测过程特征分析

航天器对小天体进行抵近探测时, 具有如下特点。

(1) 观测区域逐渐缩小: 随着运动平台逐渐接近目标, 观测区域 (视野) 会逐渐缩小。

(2) 视场快速变化: 运动平台的高速运动和姿态机动使得观测视场迅速变化, 运动距离在一分钟内可从几十千米变化到几百米。

(3) 分辨率递增梯度: 在获取的探测图像序列中, 通过图像叠加处理, 视场中心区域的像素分辨率逐渐提高, 分辨率增幅约为原始分辨率的两倍。

(4) 高帧频图像序列: 在高速运动的抵近探测过程中, 为避免图像模糊, 探测器采集的图像序列需具备高帧率, 通常帧率不低于 20 Hz。

(5) 帧频变化: 随着航天器速度与观测距离的动态变化, 采集图像的帧率也随之调整。航天器速度越快, 与目标距离越近, 帧率要求越高, 以确保成像质量。

(6) 系统误差: 探测器在成像过程中, 探测系统可能受到多种系统误差的影响^[14], 导致图像质量下降, 产生噪声增加、运动模糊以及分辨率不足等问题。

2 基于 BasicVSR++ 模型的适配构建

针对抵近过程的探测视频开展超分辨率研究, 提出了一种基于 BasicVSR++ 模型的适配方法。BasicVSR++ 模型大致可以分为聚合模块、对齐模块和上采样模块^[13], 这里通过对其模型中的部分模块进行相应调整, 增强其性能和鲁棒性, 使其更加适用于抵近探测中的超分辨率成像任务。

2.1 空间与通道注意力机制模块

在抵近探测的场景中, 由于观测区域的逐渐缩小以及视场的快速变化, 目标物体例如小行星等在图像中的相对位置和尺寸发生变化^[14]. 这一过程中, 尽管抵近成像可能产生的离焦问题可以通过焦面预置技术予以解决, 但是实际成像仍面临一些挑战: 一方面, 探测器系统在成像时会受到多种系统误差的影响, 导致图像出现噪声增加、运动模糊等问题; 另一方面, 由于平台运动、视场快速变化以及光照条件的动态改变, 使得图像特征在不同帧间的变化较为复杂. 基于上述分析, 本文引入注意力机制, 提升网络在动态环境中的适应性和图像处理的准确性.

引入注意力机制的目的是增强模型对核心特征的提取和对非核心特征的过滤^[15], 从而提升网络的整体表现. 该机制在通道和空间两个维度上展开^[16,17], 其中空间注意力引导模型聚焦于关键区域特征. 通道注意力提高模型对关键特征通道的关注度.

常用的SENet (Squeeze-and-Excitation Network) 等模块仅在通道上获取注意力^[17], 而未能体现空间维度的关注^[18]. 本研究选用 CBAM (Convolutional Block Attention Module) 模块, 在通道和空间上同时应用注意力^[19], 提升特征提取能力, 从而获得更优表现.

CBAM 模块结构如图 1 所示, CBAM 通过通道和空间两个注意力子模块优化特征处理^[19]: 通道注意

力利用全局平均池化和最大池化生成通道描述符, 通过全连接层和 Sigmoid 突出重要通道; 空间注意力则在此基础上通过池化操作和卷积生成空间注意力图, 突出关键区域.

通道模块首先对输入特征图执行全局最大池化和全局平均池化, 生成代表每个通道全局特征值的两个向量, 如图 2 所示. 这些特征向量输入到共享的全连接层, 学习不同通道的注意力权重, 使网络自适应地识别重要通道. 经过 Sigmoid 激活函数处理后的交叉结果产生最终的通道注意力权重, 这些权重与原始特征图的每个通道相乘, 突出对当前任务有帮助的特征, 并抑制无关通道.

通道注意力模块计算公式如下:

$$M_c(F) = \sigma(F_{\text{avg}}^{\text{pool}} + F_{\text{max}}^{\text{pool}})_{\text{MLP}} = \sigma\{W_i[W_o(F_{\text{avg}}^c)] + W_i[W_o(F_{\text{max}}^c)]\}. \quad (1)$$

可以看到, 通道注意力机制通过对特征图在空间维度的压缩, 实现了对所有像素信息的均衡考量以及对强响应区域的强调. 将聚合的空间信息传递到共享网络中进一步压缩, 最终生成通道注意力图, 提升了模型对关键特征的识别能力和响应效率.

空间注意力模块通过最大池化和平均池化捕捉上下文特征, 将池化结果拼接后经卷积和 Sigmoid 处理生成权重图^[19]. 这些权重与原始特征相乘, 实现对关键区域的增强和次要区域的抑制, 提升图像信息的

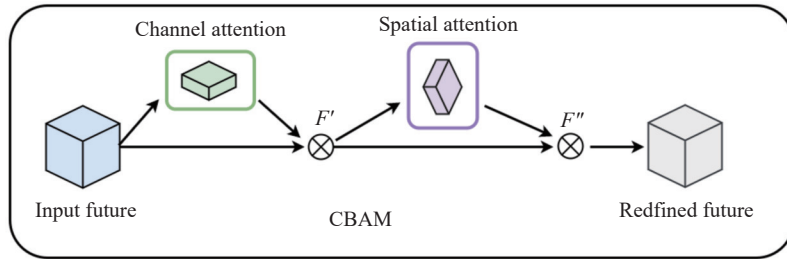


图 1 CBAM 模块结构

Fig. 1 Structure of CBAM module

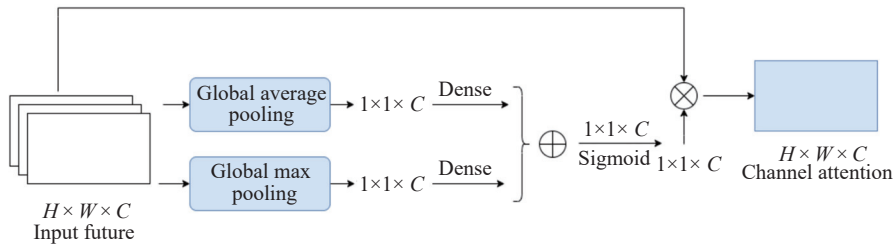


图 2 通道注意力模块

Fig. 2 Channel attention module

解析能力.

如图3所示, 空间注意力机制先通过最大池化操作从各通道中提取最大值, 再通过平均池化操作计算通道的平均值, 二者均以图像的高和宽为运算次数. 将两种池化得到的单通道特征图合并成一个包含两个通道的特征图^[19]. 通过一个 7×7 的卷积层处理合并后的特征图, 使用 Sigmoid 激活函数, 确保输出权重系数 M_s 的值位于 $0 \sim 1$ 内, 最终通过得到的空间注意力权重 M_s 与原特征图 F' 进行按位乘积得到融合后的特征图, 以提高特征表示的质量和相关性, 实现通道和空间维度上的自适应特征优化, 其计算公式为

$$M_s(F) = \sigma\{f^{7 \times 7}([F_{\text{avg}}^{\text{pool}}, F_{\text{max}}^{\text{pool}}])\} = \sigma\{f^{7 \times 7}([F_{\text{avg}}^s, F_{\text{max}}^s])\}. \quad (2)$$

通道注意力 (CA) 和空间注意力 (SA) 模块可以并行或顺序组合. 研究^[19]表明, 按顺序组合且先应用通道注意力效果更佳, 因此利用空间注意力特征对原始特征图进行加权相乘处理, 得到最终特征表示, 图4

给出了 CBAM 模块与 BasicVSR++ 中 ResNet 残差块的结合方式, 通过在残差连接的前后插入 CBAM, 增强了模型的特征提取能力.

CBAM 注意力模块通过自适应调整通道与空间特征权重^[19], 增强了关键特征的捕捉能力, 同时保持较低的计算开销, 其高效且轻量化的设计在抵近探测任务中展现了应用价值. 在动态视场和观测区域不断变化的抵近探测场景中, CBAM 的空间注意力机制能够引导网络聚焦于关键部位, 从而有效识别复杂地形中的细节特征, 并优化不同表面区域的关注度, 提升目标细节的重建质量. 同时, 通道注意力机制通过动态调整通道间的权重分布, 强化了帧间特征变化的提取能力, 为抵近探测过程中视频超分辨率任务的细节重建提供了支持.

2.2 采样点调制优化对齐策略

视频超分辨率算法需要解决帧间运动导致的特征不对齐问题^[20]. 为更好地学习运动物体的特征, 需

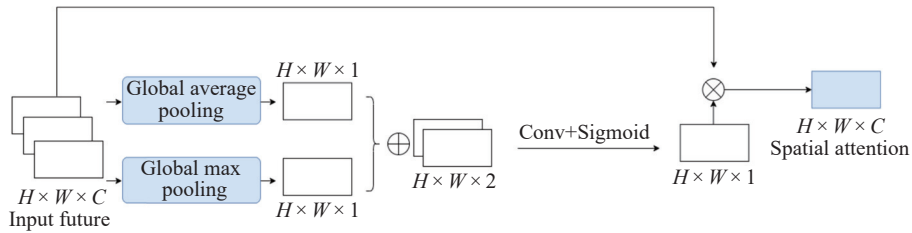


图3 空间注意力模块

Fig. 3 Spatial attention module

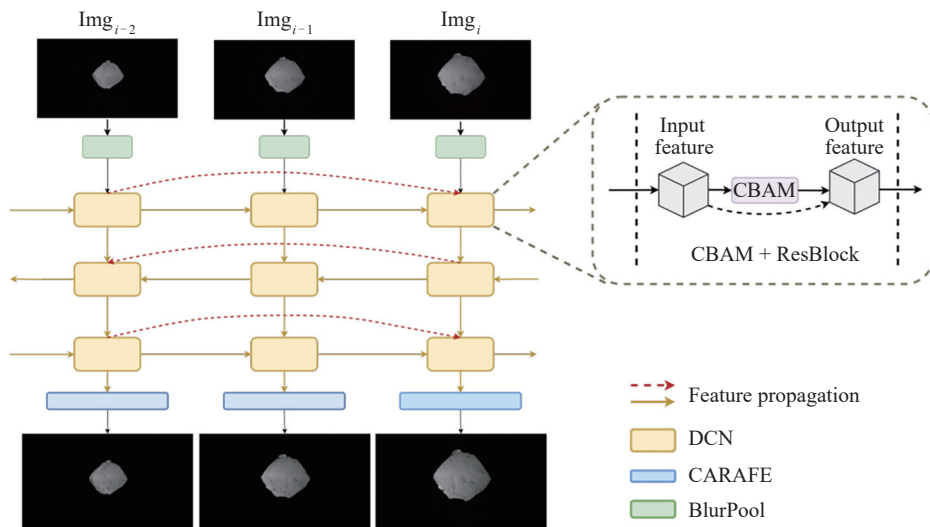


图4 CBAM 结合残差块

Fig. 4 CBAM combined with residual block

要将不同帧之间位置发生变化的同一物体对齐到相同位置, 确保物体特征在帧间保持一致, 即帧间对齐^[21].

在早期的对齐方法研究中, 运动估计和运动补偿技术 (MEMC) 是经典研究主题. 多数研究^[22-25]采用 MEMC (Motion Estimation and Motion Compensation) 捕获帧间运动. 虽然早期 VSRnet^[26]采用了传统 MEMC 方法, 但是后续 VESPCN^[22], TOFlow^[24]和 FRVSR^[25]等方法都发展了专门针对 MEMC 的子模块或子网, 进一步提升了性能. 然而 MEMC 方法在大运动或者亮度变化下表现不稳^[21], 难以应对抵近探测中视场快速变化和位置偏移导致的模糊与误差. 可变形卷积因其对环境变化的鲁棒性^[10], 成为应对此类高动态场景的有效解决方案.

相比多头自注意力机制 (MHSA), 传统卷积神经网络在处理图像数据时存在两大局限性^[27]: 一是仅聚合邻近像素信息, 难以有效建模全局特征; 二是卷积核静态参数无法动态调整空间聚合权重, 导致在自适应空间聚合方面表现不足, 缺乏对输入特征差异性的敏感度^[28]. Zhu 等^[29]的研究显示, 可变形卷积能够增强 CNN 对对象几何变化的建模能力. 在 VSR 领域, TDAN 方法^[11]首次使用可变形卷积实现帧间对齐, 随后 EDVR^[10], DNLN^[30]和 D3 Dnet^[31]等超分方法进一步利用此技术改善帧对齐效果.

BasicVSR++中的 DCN 算子 (BDCN) 通过引入可学习的偏移量实现局部与全局的交互^[13], 具有长程建模能力, 并能够动态调整调制标量和偏移量, 从而具备类似 MHSA 的自适应空间聚合能力^[29].

本文引入了改进的 DCN 算子^[32] (IDCN), 考虑到 BDCN 中每个点通过其独立的聚合窗口处理, 且无需 Key 概念^[32], 因此考虑去除 Sigmoid, 将介于 0~1 之间的调制标量转换为类似于卷积的无界动态权重^[32], 此改动继承了卷积的灵活性, 与输入独立窗口形状和权值的静态属性被适应性替代^[32], 赋予调制因子更大的自由度和更快的收敛速度, 增强其动态特性和表达能力.

为进一步提升性能, 对引入的 IDCN 算子^[33]在 BDCN 基础上进行改动. 首先, 采用共享投影权重和多组机制, 在 BDCN 中, 每个采样点配备独立的投射权重, 导致参数量与采样点总数线性相关, 增加了参数和内存复杂度^[33]. IDCN 通过借鉴可分离卷积的设

计理念, 采纳了位置无关的权重以替代原有的分组权重, 从而在不同采样点之间共享投影权重, 并成功保留了所有采样位置的依赖性, 这一创新提升了模型的参数效率和内存利用率, 同时降低了系统的复杂度^[33]. 其次, 引入多组机制, 在 IDCN 架构中引入了分组空间聚合的策略, 使每个组具备独立的采样偏移量, 并允许 IDCN 的每一层通过不同的空间聚合模式产生丰富的特征表达, 从而提升了模型处理复杂视觉任务的能力^[33].

IDCN 整合了偏移预测网络与可变形卷积, 利用偏移量确定每个输入位置的位移, 并基于此处理特征, 有效融合了传统卷积和注意力机制的特点: 一方面, 通过滑动窗口方式处理输入数据, 维持卷积的固有偏置^[33]; 另一方面, 其采样偏移量 Δp 和空间聚合权重 m 根据输入特征动态生成, 表现出动态特性, 增强了类似注意力机制的功能. 其数学表达式为

$$y(p_0) = \sum_{g=1}^G \sum_{k=1}^K w_g m_{gk} x_g(p_0 + p_k + \Delta p_{gk}). \quad (3)$$

其中, G 表示组数量. w_g 表示每组内共享投影权重, m_{gk} 表示第 g 组第 k 个采样点的调制因子, 经过归一化处理. IDCN 算子增强了视觉大模型的适用性, 通过补足传统卷积在长程依赖和动态空间信息整合的不足, 实施稀疏全局建模, 同时保留 CNN 的归纳偏置^[33], 实现计算效率与模型准确性的平衡. 其动态稀疏卷积窗口避免了大型密集卷积核的优化问题和高成本, 并通过可变形卷积实现感受野的自动学习, 支持长短距离交互.

引入的可变形对齐算子采用可调扩张率, 在保持计算效率的同时, 实现了更大感受野的特征提取. 该算子增强了模型对动态场景的适应性, 提高了其在高速动态环境中的泛化能力, 实现了更稳定的特征提取与帧间对齐, 能够有效捕捉快速运动目标及不规则天体的特征, 使其在小行星探测等场景中展现出应用价值.

2.3 动态上采样

在抵近探测的图像序列超分辨率重建过程中, 高精度图像重建对结果至关重要. 其中图像上采样是核心环节, 尤其是在从低分辨率图像恢复细节时. 传统的最近邻和双线性插值虽然简单, 但是仅依赖局部特征, 忽略了全局信息, 效果有限. 相比之下, 反卷积上采样通过统一核扩展特征图, 增强了细节, 但是由于

引入额外参数,增加了模型复杂度和计算量,在大规模数据或资源受限场景下效率低下.

BasicVSR++采用的 PixelShuffle^[34] 尽管具有较高的计算效率,但是其固定核结构缺乏内容感知能力,难以在复杂场景中充分恢复图像细节和结构,尤其是在处理小行星探测等图像内容复杂或细节变化较大的场景中.

CARAFE^[35] 上采样方法提供了更具灵活性和精度的解决方案. CARAFE 结合局部与全局特征,扩展感受野,使得模型能够在更大的区域内聚合上下文信息,提升了模型全局结构与细节的重建能力^[35]. 通过内容感知机制动态生成自适应采样核,灵活调整以适应不同输入和场景,实现高效图像重建. 其轻量级设计以较少的计算代价实现了性能提升,同样适用于需要处理大量数据的抵近探测任务. 基于上述分析,本文采用 CARAFE (Content-Aware Reassembly of Features) 轻量级上采样模块,其具体结构如图 5 所示,其中 W_σ 为上采样后特征图的宽; H_σ 为上采样后特征图的高; C_m 为降维后的特征通道数; k_e 为卷积核的卷积核大小.

CARAFE 模块将上采样视为特征重组过程. 每个输出像素 L' 通过与输入特征图中以 L 为中心的局部邻域和上采样核进行点积计算,从而重组特征. 其通过以下两个步骤计算: 通过核预测模块基于维度为 $H \times W \times C$ 的输入特征图,为每个目标像素位置生成

重组核,并根据上采样比率 U 扩展特征图尺寸至 $UH \times UW \times C$; 内容感知重组模块利用所生成重组核对特征进行重新组合,确保特征图的精确扩展和信息完整性,从而提升空间分辨率与特征表达质量.

从图 5 可以看出, CARAFE 由上采样核预测模块和特征重组模块组成. 特征图首先被送入上采样核预测模块. 在这一模块中,特征图的通道数 C 通过一个包含 1×1 卷积、批量归一化 (BN) 和 SiLU 激活函数的卷积层进行降维,降维后的通道数为 C_m ,可由下式计算:

$$C_m = \sigma^2 K_{up}^2. \quad (4)$$

其中, σ 为上采样倍数; K_{up} 为上采样核大小. 随后通道数被压缩后的输入特征图通过一个核大小为 $K_e \times K_e$ 的卷积层来预测上卷积核,该卷积层的输出通道数为 σ^2 , 然后采用 PixelShuffle 方法在空间维度上展开通道维,生成形状为 $H_\sigma \times W_\sigma$ 的上采样核,接下来使用 Softmax 激活函数归一化处理. 在特征重组模块中,使用上采样核预测模块生成的上采样核与通过常规方法上采样得到的特征图进行点乘,从而得到最终的输出特征图. 这一过程有效结合上采样核预测与特征重组,提高了上采样的精度和效率.

抵近过程中成像面临多项挑战,例如图像噪声、多视角观测引起的立体几何效应、相位不稳定性. CARAFE 通过内容感知的自适应核机制,能够在每

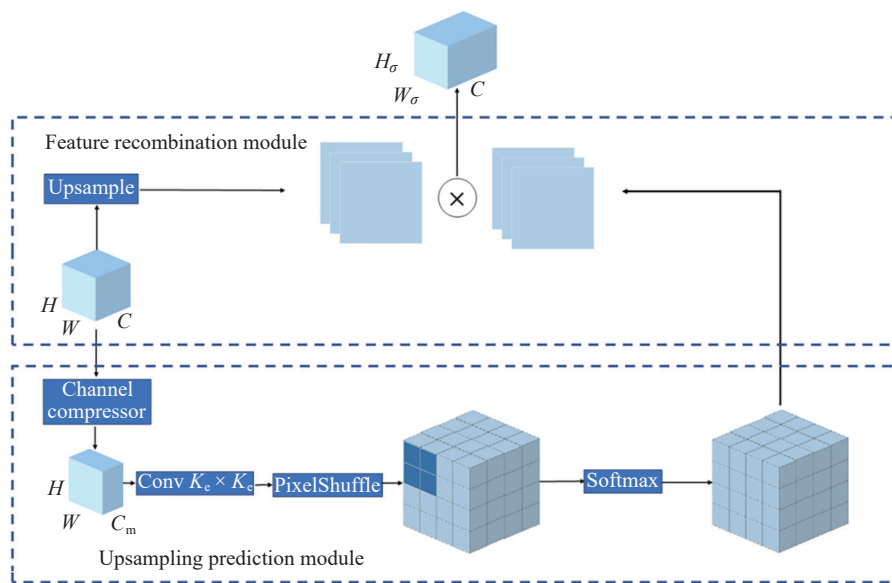


图 5 动态上采样模块结构

Fig. 5 Dynamic upsampling module structure

个像素位置动态调整上采样权重, 适合处理含噪声和运动模糊的复杂图像. 其大感受野特性有助于纠正多视角观测带来的几何扭曲. 此外, CARAFE 通过实时调整上采样核, 适应视野变化, 并细化图像细节, 特别适合抵近过程中分辨率逐步提高的场景. 其高效设计和较少的计算及参数量使其适合部署在资源受限的抵近探测系统中, 能够快速处理高帧频图像序列, 根据不同场景和帧频变化灵活调整上采样核, 实现高质量的图像重建.

图 6 给出了动态上采样模块在网络中的位置及作用. 该模块通过自适应生成上采样核重组特征, 增强网络的细节恢复能力和重建精度. 凭借大感受野和轻量化设计, CARAFE 可以提升模型处理抵近探测图像序列的能力, 并且在计算资源受限的空间科学任务中改善低分辨率图像的特征提取与目标探测性能.

2.4 平移不变性的下采样模块

平移不变性 (Translation Invariance) 是指在卷积神经网络中网络对于输入图像及其平移版本能够产生相似的输出. 这意味着无论图像中的对象是否发生位移, CNN 都能识别出相同的特征, 并输出相同的分类结果或标签.

在实际应用中, 这些性质往往只在输入平移量为采样步长 N 的整数倍时才成立, 这种情况被称为周期平移不变性/等变性^[36]. 这表明在某些特定的平移步长下, 函数才展示出预期的不变性或等变性特征.

在抵近探测中, 观测平台或目标的快速移动常导致图像捕捉出现抖动、位移或视角变化, 因此平移不变性尤为关键. 这一特性可确保模型对微小变动保持

鲁棒性, 有助于帧间特征对齐, 从而更好地学习运动物体的特征.

尽管滑动窗口卷积理论上应提供一定的平移不变性, 实际中 CNN 对输入图像的微小变动非常敏感, 轻微的像素改动或图像平移可能导致输出变化, 甚至预测错误^[36], 特别是在深层网络中, 这种问题更为明显. 研究表明, CNN 只能从大量数据中学习部分不变性^[37], 难以掌握变换不变性.

有研究分析了 CNN 平移不变性退化的原因, 主要在于下采样操作未遵循采样定理^[36]. 理论上, 平移的步幅应为二次采样因子的整数倍, 以保持特征一致性. 然而现代 CNN 的池化层采样因子通常过大, 导致平移不变性无法有效保证. 此外, 研究表明^[36], 无论是最大池化、平均池化, 还是步长大于 1 的卷积, 较大步长的下采样操作均可能削弱网络的平移不变性, 甚至仅平移一个像素, 最大池化的结果也会有显著差异.

传统池化提供的平移不变性有限, 尤其是在不考虑采样定理时, 可能导致混叠现象. 最大池化仅在感受野内最大值保持不变时, 平移不变性才有效, 而大步长会引起输出波动. 为了抗混叠, 常用模糊下采样 (例如平均池化) 作为低通滤波手段^[37]. 平均池化虽然抗锯齿效果较好, 性能却逊色于最大池化, 使得其在图像处理性能与抗锯齿间取舍, 难以兼顾.

本文通过在下采样前引入低通滤波器, 合理优化其位置和频率设置, 在保留最大值池化功能的同时有效降低混叠效应, 增强网络的平移不变性. 此外, 通过在卷积神经网络中插入低通滤波器, 防止下采样过程中的走样问题.

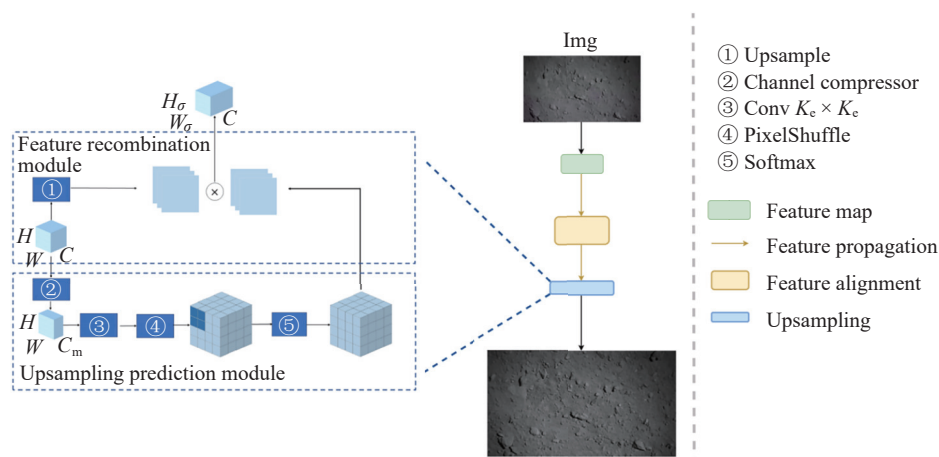


图 6 动态上采样模块在网络中的作用

Fig. 6 Role of the dynamic up-sampling module in the network

针对 CNN 在下采样过程中未遵循采样定理而易引入混叠效应、导致平移不变性退化的问题, 采用 BlurPool^[36] 下采样方法, 在最大值选取与子采样之间插入低通滤波器, 实现抗锯齿与平移不变性兼顾. 传统最大池化可以分解为先进行步长为 1 的最大池化操作, 再执行下采样; BlurPool 通过在最大池化和下采样之间引入低通滤波器^[36] (即最大池化、低通滤波与下采样的组合) 替代原始最大池化, 缓解平移不变性的丢失问题.

这一改动不仅优化了最大池化的两个阶段过程, 即先通过密集滑窗选取最大值以捕捉局部信息, 然后在低通滤波器辅助下完成子采样以减少混叠现象, 而且结合平均池化的抗锯齿特性与最大池化的局部特征提取优势, 从而提升了模型的抗锯齿能力与图像处理的稳定性. 实践证明, BlurPool 在提升平移不变性、减少混叠现象的同时, 可集成到现有神经网络中, 为模型实现抗锯齿和性能提升提供了方案^[36], BlurPool 的整体结构与应用方式如图 7 所示.

MaxPooling 可以结合 BlurPool 进行优化: MaxPooling 先以步长 1 进行最大池化, 再以步长 2 的 BlurPool 进行平滑下采样, 这样既保留了特征捕获能力, 又减少了大步长带来的混叠失真; AveragePooling 可直接用步长 2 的 BlurPool 替代, 利用其固有的平滑滤波特性提升图像质量, 其数学表达式为

$$M_s(F) = \sigma\{f^{7 \times 7}([F_{\text{avg}}^{\text{pool}}, F_{\text{max}}^{\text{pool}}])\} = \sigma\{f^{7 \times 7}([F_{\text{avg}}^s, F_{\text{max}}^s])\}. \quad (5)$$

但是这种改动也带来了新的挑战: 由于低通滤波后丢失了部分高频信息, 经过 BlurPool 后, 数据的方差变小, 特征之间的数值差异减少, 使得网络难以快速收敛, 网络提取特征的难度变大. 为解决这一问题, 采用 BatchNorm 加大特征间数值上的差异, 增强特征提取能力, 加快网络的学习过程. 加入 BatchNorm

后, 网络的学习速度得到改善.

该方法在动态环境下的抵近探测场景中展现出应用价值, 通过在下采样过程中引入低通滤波功能, 不仅可以有效平滑图像处理, 减轻因高速运动引起的相位不稳定、运动模糊及视场变化等不利影响, 提升帧间对齐精度, 还增强了系统抗干扰能力, 能够在恶劣天气下减弱信号干扰、滤除高频噪声, 并降低微陨石或尘埃对图像质量的影响, 为复杂场景下的抵近探测任务和超分辨率算法的优化提供基础.

基于上述分析, 本文引入调整的 BlurPool 模块, 在降低混叠效应的同时, 通过 BatchNorm 加速收敛, 解决了因高频信息丢失引起的网络训练变慢问题, 增强了模型对动态场景下目标平移抖动的鲁棒性, 为超分辨率算法的优化提供了新思路.

2.5 小结

针对航天器高速运动和姿态机动等带来的运动模糊问题, 通过引入 IDCN 算子, 利用动态卷积核和优化的注意力机制, 实现复杂运动场景不同帧间特征的高效对齐; 同时采用 BlurPool 下采样方法, 通过平滑图像削弱运动模糊, 也提升了帧间对齐精度.

考虑航天器运动产生的各类噪声, 这里引入 CARAFE 上采样模块, 通过内容感知的自适应核机制, 在每个像素位置动态调整上采样权重, 从而在处理含噪声图像时, 有效降低噪声对细节重建的干扰, 进而恢复图像细节和结构. 此外, 通过 CBAM 模块结合空间与通道注意力机制, 强化对关键特征的提取, 同时抑制噪声等非关键信息, 提升模型对噪声图像的适应性.

在抵近探测过程中, 探测距离逐渐逼近时, 探测目标越来越清晰. 通过改进的上采样和下采样模块协同工作, 其中 CARAFE 上采样模块根据图像内容动态调整上采样核, 适应分辨率变化, 更好地恢复图像

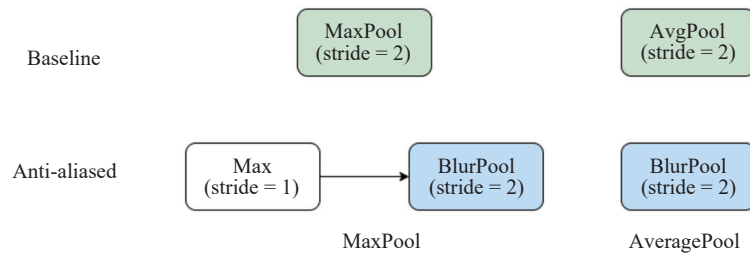


图 7 BlurPool 下采样结构及应用方式

Fig. 7 BlurPool downsampling structure and application methods

细节; BlurPool 下采样方法在降低混叠效应的同时, 保证在分辨率变化过程中模型对微小变动的鲁棒性, 使模型能够较为稳定地处理不同分辨率图像, 提升整体性能.

3 仿真分析

3.1 迁移学习

在超分辨率重建领域中, 训练样本集的丰富度及其规模对于提高模型的表现至关重要. 针对小行星图像的超分辨率重建任务, 理想情况下要求覆盖多样化的轨道位置、光照条件、表面特征等情况下的海量高分辨率样本. 但是现实限制了高分辨率小行星图像集的构建, 从而影响了基于深度学习的超分辨率算法在此类图像上的进展.

迁移学习为解决这一难题提供了一种有效途径, 通过借鉴在源域中获得的深层次知识, 迁移学习使得模型能够跨领域解决问题, 适用于样本匮乏的目标领域^[38]. 本研究参考自然图像的亮度通道与小行星图像在一定条件下的视觉相似性, 采取迁移学习策略将自然图像的超分辨率技术应用于小行星图像的超分辨率增强. 以较为广泛获取的自然图像作为训练基础, 应用卷积神经网络技术构建初始超分辨率模型; 以小批量精选的小行星图像对该模型进行微调训练, 优化模型对小行星特有特征的适应性, 实现针对小行星表面细节处理的优化.

具体地, 以 REDS 数据集作为源域对模型进行预训练, REDS 包含了 300 个不同场景的视频序列, 分辨率为 720×1280 pixel, 训练集、验证集和测试集分别包含 240, 30, 30 个视频. 随后, 将源域模型参数作为初始参数, 用小行星图像作为目标域数据集进行微调训练. 为进一步增强模型的泛化能力, 对收集的原始小行星图像进行傅里叶变换, 在此基础上进行包括水平翻转、仿射变换、图像倍乘、高斯模糊和图像亮度调整等扩展处理, 模拟不同的视觉变化. 通过这种方式, 获得了 3000 张经过增强的小行星图像, 包括不同类型的小行星和轨道环境, 用于迁移学习目标域训练.

所有图像均经过精确的标注处理, 确保训练数据的准确性和一致性. 目标域训练集和测试集按照 75% 和 25% 的比例划分为 2250 张训练图像和 750 张测试图像, 用以提供充足的训练数据, 并确保有足够的测

试样本验证模型在现实环境中的表现.

微调后, 模型在已有的泛化学习基础上, 提升了对小行星表面细节和轮廓的还原能力, 即使目标域数据有限, 模型依然能够在超分辨率重建任务中展现较高的精度, 满足小行星探测的需求.

3.2 实验结果

研究使用 REDS 和 Vimeo-90 K 数据集进行超分辨率模型的训练与测试. REDS 数据集中, REDS4 作为测试集, REDSval4 作为验证集, 其余片段用于训练, 测试集还包含 Vid4, UDM10 和 Vimeo90 K-T. 所有模型基于双三次插值和模糊下采样进行 4 倍下采样测试. 采用 Adam 优化器和余弦退火方案, 主网络和流量网络的初始学习率分别为 1×10^{-4} 和 2.5×10^{-5} , 总迭代次数为 6×10^5 , 前 5000 次迭代中流量网络的权重保持固定. 批量大小设置为 8, 输入的低分辨率帧大小为 64×64 pixel, 并使用 Charbonnier 损失函数. 在硬件方面, 实验使用了 Intel i7 12700 H CPU, NVIDIA GeForce RTX 3070 GPU (12 GB 显存) 和 32 GB 内存, 运行环境为 Windows10, CUDA 12.1, cudnn v8.6.0, 软件采用 Pytorch 1.7.1, 并在 Pycharm 中进行编译, 实验环境配置参数见表 1.

为全面评估所提方法的有效性, 分别与 SRCNN, EDVR, BasicVSR++ 前沿算法进行对比, 结果如图 8 所示. 从图 8 可以看出本文方法在小行星抵近探测视频超分辨率重建方面的性能.

对图 8 中结果进行比较分析可以看出, 采用本文方法后, 小行星表面的微小纹理和轮廓得到了更好的还原. 本文提出的方法能够更好地恢复小行星的轮廓、裂缝和陨石坑等细节, 并在应对快速运动带来的模糊和抖动时, 保持对这些细节的识别能力.

仿真场景中引入高斯噪声, 图像分辨率从 128×128 pixel 变化至 256×256 pixel, 帧频为 30 Hz. BasicVSR++ 与本文方法的超分辨结果如图 9 所示.

表 1 网络训练所用网络配置条件
Table 1 Configuration conditions used for network training

软硬件配置	型号/版本号
CPU	Intel Core i7-12700
GPU	NVIDIA RTX 3080
内存	12 GB
软件环境	Python3.9, CUDA12.1, Pytorch1.7.1

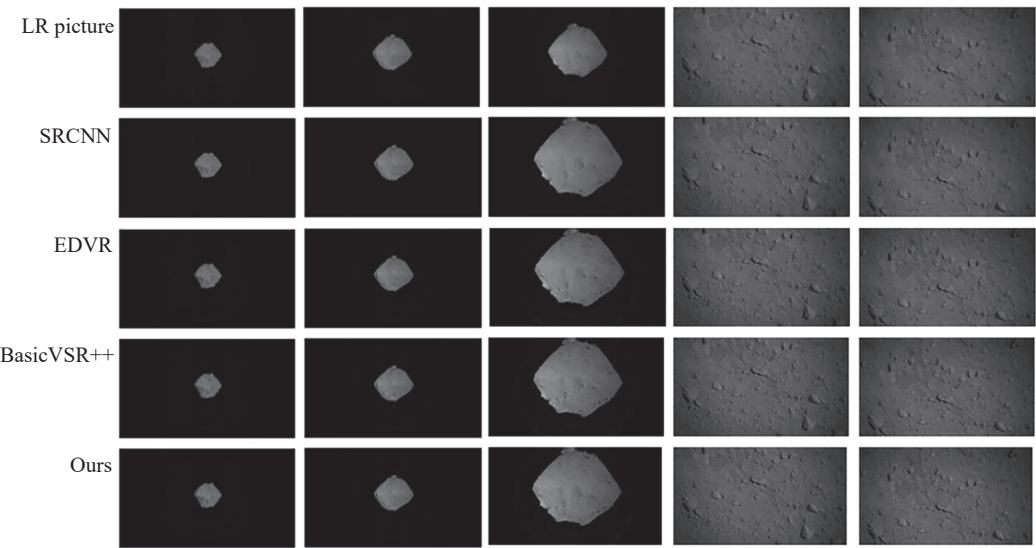


图 8 不同超分辨率方法重建图像对比

Fig. 8 Comparison of reconstructed images obtained by different super-resolution methods

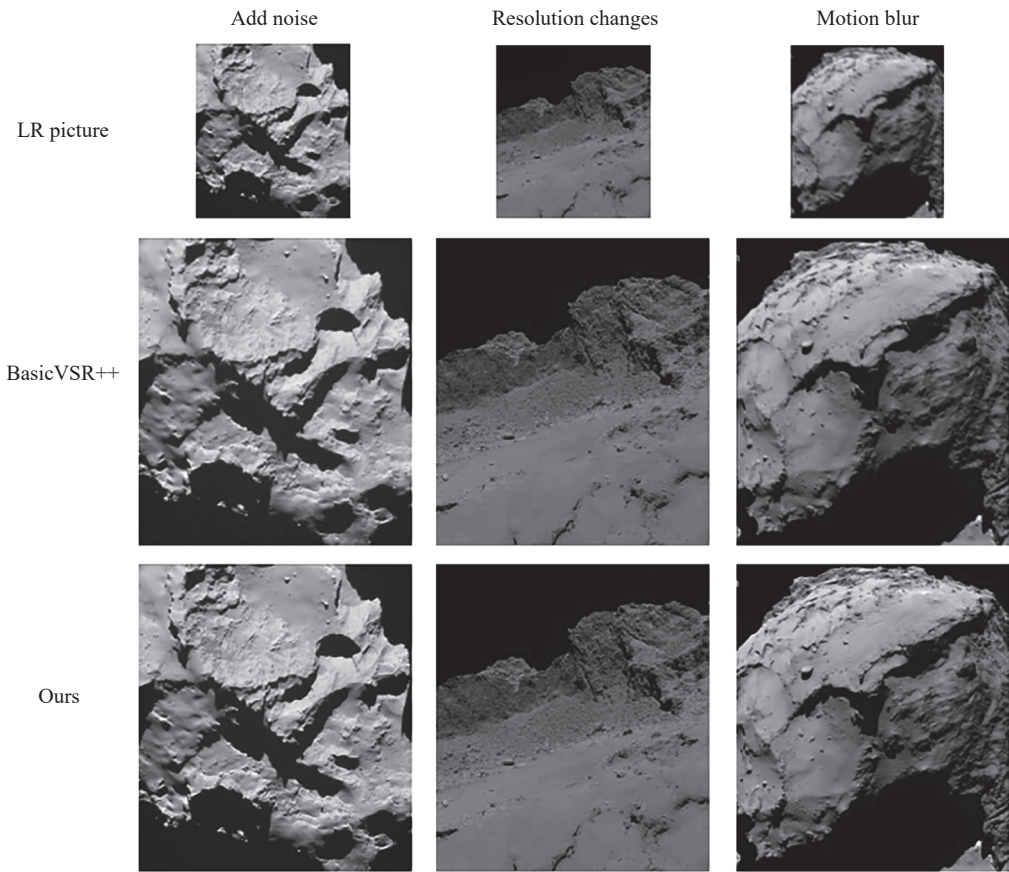


图 9 添加噪声、提升分辨率及运动模糊场景下模型效果对比

Fig. 9 Comparison of model effects under noise adding, resolution increasing and motion blur scenes

图 9 给出了在添加高斯噪声、图像分辨率逐帧提升 1 倍 (128×128 pixel 至 256×256 pixel) 和运动模糊场景下不同方法的重建效果. 由图 9 可见, 相比 BasicVSR++, 本文方法重建的图像噪点更少, 目标

轮廓更加清晰,同时不同分辨率区域之间的过渡更加平滑.

由表 2 可见,考虑噪声的情况下,本文方法的 SSIM 值为 0.9171, PSNR 值为 31.82, 均高于其他模型. 这表明即使存在一定噪声,本文方法能更好地保留图像结构相似性,提升图像的峰值信噪比,恢复出更高质量的图像. 在模拟运动模糊和图像分辨率变化场景中,本文方法的 SSIM 和 PSNR 指标同样优于其他算法,验证了其对于运动模糊的有效补偿能力及对分辨率变化的良好适应性.

综上可知,本文提出的改进方法能够适应存在噪声、运动模糊等多种影响的复杂场景,相比传统方法有效提升了图像质量,同时具有较高的鲁棒性.

3.3 消融实验

为验证模型改动部分的有效性,对 4 个模块进行了消融实验,结果列于表 3.

实验结果表明,CBAM 模块增强了网络对重要特征的聚焦能力,特别是在处理细节丰富的区域时,提升了图像质量和细节恢复能力. 从表 3 具体来看,第 1 组实验与第 2 组实验的结果对比显示,加入 CBAM

后, PSNR 提升了 0.12 dB, SSIM 提升了 0.73%, 这主要是由于 CBAM 通过通道和空间注意力机制有效增强了模型对图像中重要特征的感知能力.

CARAFE 动态上采样模块通过其自适应核生成机制,提高了上采样的精度,使得重建的图像不仅提高了分辨率,而且细节保真度也得到改善. 第 3 组实验替换原有上采样模块为 CARAFE 轻量级上采样模块后,相比第 2 组实验, PSNR 提升了 0.21 dB, SSIM 提升了 0.45%. 这主要是由于 CARAFE 模块能够根据图像内容动态调整上采样核,更好地恢复图像细节和纹理,尤其是在细节的重建方面显示出更优的表现.

引入的对齐模块使得网络在处理图像对齐时更为精确,特别是在小尺度目标和细节处理方面表现更加突出,通过获取相邻帧的高度相关信息,增强了特征的表达能力. 引入 IDCN 后,相比第 3 组实验,第 4 组实验的 PSNR 提升了 0.14 dB, SSIM 提升了 0.548%. IDCN 通过动态卷积核和优化的注意力机制,增强了模型对特征变化的适应能力以及对重要特征信息的捕捉和表达能力,特别是在对齐不同帧时的精确度,有效提升了整体图像的视觉质量和指标表现.

表 2 多干扰因素下模型成像指标对比
Table 2 Comparison of model imaging indicators under multiple interference factors

	Metrics	Bicubic	SRCNN	TOFlow	EDVR	RSDN	BasicVSR++	Ours
Add noise	SSIM	0.8684	0.8654	0.8748	0.8912	0.8954	0.9008	0.9171
	PSNR	28.61	29.54	29.67	30.69	30.95	31.28	31.82
Resolution changes	SSIM	0.8625	0.8751	0.8779	0.8852	0.8952	0.9010	0.9173
	PSNR	28.93	29.10	29.79	30.65	31.04	31.30	31.83
Motion blur	SSIM	0.8586	0.8693	0.8767	0.8792	0.8972	0.9011	0.9174
	PSNR	28.27	28.98	29.15	30.12	30.84	31.31	31.81

表 3 模型消融实验结果
Table 3 Model ablation test results

Test No.	CBAM	CARAFE	IDCN	BlurPool	PSNR/dB	SSIM
1	×	×	×	×	31.39	0.9019
2	✓	×	×	×	31.51	0.9085
3	✓	✓	×	×	31.72	0.9126
4	✓	✓	✓	×	31.86	0.9176
5	✓	✓	✓	✓	31.91	0.9181

注 符号✓表示该实验组包含该模块,符号×表示未包含该模块.

在上述实验基础上, 在第 5 组实验中进一步引入 BlurPool 模块后, 相比第 4 组, PSNR 提升了 0.05 dB, SSIM 值提升了 0.0005%. 结果表明, BlurPool 在提升图像清晰度和细节恢复方面起到积极作用, 验证了其在抗运动模糊中的作用. BlurPool 通过其低通滤波特性, 有效减少了动态场景下或视场快速变化时的混叠效应, 增强了平移不变性, 提高了模型的鲁棒性.

4 结论

针对复杂环境下的抵近探测场景, 提出了一种基于 BasicVSR++ 的超分辨率重建方法. 通过引入空间和通道注意力机制以及动态卷积, 结合共享投射权重与多组机制优化对齐算子, 进一步通过下采样方式增强了平移不变性, 提升了模型在复杂背景下的鲁棒性和细节重建能力. 在给定的测试集上进行测试, 实验结果表明, 本方法在 PSNR 和 SSIM 等指标方面均比传统方法有所提升, 在性能提高的同时, 计算量增长较低, 为小行星探测等观测任务提供了解决方案. 未来工作中, 将继续深入探索该方法在空间科学领域的应用, 特别是在小行星探测、空间目标跟踪等场景下的潜力, 并结合多光谱成像等先进技术, 进一步优化算法的适应性, 提升其在不同平台和不同光学条件下的性能.

利益冲突 所有作者均声明不存在利益冲突.

参考文献

- [1] DONG C, LOY C C, HE K M, *et al.* Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(2): 295-307
- [2] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 1646-1654
- [3] YANG J C, HUANG T. Image super-resolution: Historical overview and future challenges[M]//MILANFAR P. Super-Resolution Imaging. Boca Raton: CRC Press, 2017: 1-34
- [4] LIU H Y, RUAN Z B, ZHAO P, *et al.* Video super-resolution based on deep learning: a comprehensive survey[J]. *Artificial Intelligence Review*, 2022, **55**(8): 5981-6035
- [5] LEPCHA D C, GOYAL B, DOGRA A, *et al.* Image super-resolution: a comprehensive review, recent trends, challenges and applications[J]. *Information Fusion*, 2022, **91**: 230-260
- [6] HA V K, REN J C, XU X Y, *et al.* Deep learning based single image super-resolution: A survey[C]//9th Advances in Brain Inspired Cognitive Systems. Cham: Springer, 2018: 106-119
- [7] WANG Z H, CHEN J, HOI S C H. Deep learning for image super-resolution: A survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, **43**(10): 3365-3387
- [8] JENEFA A, KURIAKOSE B M, EDWARD N V, *et al.* EDSR: Empowering super-resolution algorithms with high-quality DIV2K images[J]. *Intelligent Decision Technologies*, 2023, **17**(4): 1249-1263
- [9] NAGANO Y, KIKUTA Y. SRGAN for super-resolving low-resolution food images[C]//Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management. New York: ACM, 2018: 33-37
- [10] WANG X T, CHAN K C K, YU K, *et al.* EDVR: Video restoration with enhanced deformable convolutional networks[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Long Beach: IEEE, 2019: 1954-1963
- [11] TIAN Y P, ZHANG Y L, FU Y, *et al.* TDAN: Temporally-Deformable Alignment Network for video super-resolution [C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 3357-3366
- [12] GOPALAKRISHNAN S, CHOUDHURY A. A “deep” review of video super-resolution[J]. *Signal Processing: Image Communication*, 2024, **129**: 117175
- [13] CHAN K C K, ZHOU S C, XU X Y, *et al.* BasicVSR++: Improving video super-resolution with enhanced propagation and alignment[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 5962-5971
- [14] RABINOWITZ D L. Detection of Earth-approaching asteroids in near real time[J]. *Astronomical Journal*, 1991, **101**(4): 1518-1529
- [15] NIU Z Y, ZHONG G A, YU H. A review on the attention mechanism of deep learning[J]. *Neurocomputing*, 2021, **452**: 48-62
- [16] GUO M H, XU T X, LIU J J, *et al.* Attention mechanisms in computer vision: a survey[J]. *Computational Visual Media*, 2022, **8**(3): 331-368
- [17] SOYDANER D. Attention mechanism in neural networks: where it comes and where it goes[J]. *Neural Computing and Applications*, 2022, **34**(16): 13371-13385
- [18] HU J, SHEN L, SUN G, *et al.* Squeeze-and-excitation networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132-7141
- [19] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional block attention module[C]//Proceedings of the 15th Euro-

- pean Conference. Cham: Springer, 2018: 3-19
- [20] WEN W L, REN W Q, SHI Y H, *et al.* Video super-resolution *via* a spatio-temporal alignment network[J]. *IEEE Transactions on Image Processing*, 2022, **31**: 1761-1773
- [21] SHI S W, GU J J, XIE L B, *et al.* Rethinking alignment in video super-resolution transformers[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2022: 2615
- [22] CABALLERO J, LEDIG C, AITKEN A, *et al.* Real-time video super-resolution with spatio-temporal networks and motion compensation[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2848-2857
- [23] WANG L G, GUO Y L, LIN Z P, *et al.* Learning for video super-resolution through HR optical flow estimation [C]//14th Asian Conference on Computer Vision. Cham: Springer, 2019: 514-529
- [24] XUE T F, CHEN B A, WU J J, *et al.* Video enhancement with task-oriented flow[J]. *International Journal of Computer Vision*, 2019, **127**(8): 1106-1125
- [25] SAJJADI M S M, VEMULAPALLI R, BROWN M. Frame-recurrent video super-resolution[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6626-6634
- [26] JO Y, OH S W, KANG J, *et al.* Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3224-3232
- [27] HASSANIN M, ANWAR S, RADWAN I, *et al.* Visual attention methods in deep learning: An in-depth survey[J]. *Information Fusion*, 2024, **108**: 102417
- [28] CORDONNIER J B, LOUKAS A, JAGGI M. On the relationship between self-attention and convolutional layers [C]//International Conference on Learning Representations. Addis Ababa: OpenReview. net, 2019
- [29] ZHU X Z, HU H, LIN S, *et al.* Deformable convNets v2: More deformable, better results[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 9300-9308
- [30] Wang H, Su D, Liu C, *et al.* Deformable Non-Local Network for VideoSuper-Resolution[J]. *IEEE Access*, 2019, **7**: 177734-177744
- [31] YING X Y, WANG L G, WANG Y Q, *et al.* Deformable 3D convolution for video super-resolution[J]. *IEEE Signal Processing Letters*, 2020, **27**: 1500-1504
- [32] WANG W, DAI J, CHEN Z, *et al.* InternImage: Exploring large-scale vision foundation models with deformable convolutions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 14408-14419.
- [33] XIONG Y W, LI Z Q, CHEN Y T, *et al.* Efficient deformable convNets: Rethinking dynamic and sparse operator for vision applications[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 5652-5661
- [34] SHI W Z, CABALLERO J, HUSZÁR F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1874-1883
- [35] WANG J Q, CHEN K, XU R, *et al.* CARAFE: Content-aware reassembly of features[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 3007-3016
- [36] ZHANG R. Making convolutional networks shift-invariant again[C]//Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019: 7324-7334
- [37] AZULAY A, WEISS Y. Why do deep convolutional networks generalize so poorly to small image transformations [OL]. arXiv preprint arXiv: 1805.12177, 2018. <https://arxiv.org/pdf/1805.12177>
- [38] WEISS K, KHOSHGOFTAAR T M, WANG D D. A survey of transfer learning[J]. *Journal of Big Data*, 2016, **3**(1): 1-40

作者简介



陈宇涵 男, 现就读于中国科学院国家空间科学中心, 硕士研究生, 专业为计算机技术。主要研究方向为图像处理技术。
E-mail: 1027690140@qq.com

邓丽 (通信作者) 女, 博士研究生导师, 中国科学院国家空间科学中心研究员, 主要研究方向为大数据及图像处理技术、目标智能检测技术、星间测量和相对定位技术等。
E-mail: dengli@nssc.ac.cn