

· 综 述 ·

基于参考图像的超分辨率重建算法综述

张东晓, 唐 妮

(集美大学理学院, 福建 厦门 361021)

[摘要] 单帧图像超分辨率重建是指从低分辨率图像重建出高分辨率图像的过程, 是一类典型的不适定性问题。近期出现一种新的重建思路: 将已有的高分辨率图像作为参考, 利用参考图中丰富的纹理来补偿低分辨率图像缺失的细节信息, 从而缓解单帧图像超分辨重建的不适定性。这种基于参考图的超分辨率重建, 与普通的单帧重建不同, 除输入低分辨率图像外, 还需额外输入一张或多张与低分辨率图像内容或纹理相似的高分辨率参考图像。尽管该重建思路的相关研究开展时间较短, 但也涌现出了很多先进的重建算法, 因此急需对这些方法做一个全面的梳理。在分析和借鉴相关研究成果的基础上, 从基于图像对齐的方法和基于图像块匹配的方法入手, 对近年来出现的新方法进行综述性介绍, 并探讨其未来的发展方向, 以期抛砖引玉, 为该领域的后续研究提供有益参考。

[关键词] 图像超分辨率重建; 参考图; 图像对齐; 图像块匹配

[中图分类号] TP 391

A Review of Reference-Based Image Super-Resolution Reconstruction Algorithms

ZHANG Dongxiao, TANG Ni

(School of Science, Jimei University, Xiamen 361021, China)

Abstract: Single image super-resolution reconstruction refers to the process of reconstructing high-resolution images from low-resolution images, which is a typical ill-posed problem. Recently, a new reconstruction idea has emerged: Taking the existing high-resolution image as a reference, the rich texture in the reference image is used to compensate for the missing detail information of the low-resolution image, so as to alleviate the ill-posedness of single image super-resolution reconstruction. This new method is called reference-based image super-resolution reconstruction. Different from ordinary single image super-resolution reconstruction, the input of this method also needs one or more high-resolution reference images with similar content or texture to the low-resolution image. Although the related research on this reconstruction method has been carried out for a short period of time, many advanced reconstruction ideas and algorithms have emerged, and a comprehensive review is urgently needed. Therefore, on the basis of analyzing and drawing on relevant research results about alignment-based and patch-matching-based methods, this paper gives a general introduction to the new methods emerging in recent years, and discusses the future development direction. It is hoped that this will be a useful reference for subsequent research in this field.

[收稿日期] 2022-05-25

[基金项目] 福建省自然科学基金项目 (2020J01710); 集美大学国家基金培育计划项目 (ZP2020063)

[作者简介] 张东晓 (1980—), 男, 博士, 副教授, 硕导, 主要从事视频与图像处理、机器学习方向研究。E-mail: zdx1980@jmu.edu.cn

Keywords: image super-resolution reconstruction; reference image; image alignment; image patch matching

0 引言

图像超分辨率重建 (super-resolution reconstruction, SR) 是指从一张或多张低分辨率 (low-resolution, LR) 图像重建出高分辨率 (high-resolution, HR) 图像的过程^[1], 是一类典型的不适定性问题, 也是计算机视觉和图像处理领域中重要的研究方向。图像 SR 在实际生活中有着丰富的应用, 如智能安防^[2]、医学影像^[3]、无人机侦察^[4]、大气遥感^[5]等。

图像 SR 可分为多帧图像超分辨率重建 (multiple image super-resolution reconstruction, MISR) 和单帧图像超分辨率重建 (single image super-resolution reconstruction, SISR) 两大类。MISR^[6-9]通常针对同一场景不同角度、不同方位的图像序列, 利用这些序列之间存在的互补信息进行重建。但在某些场景中 MISR 并不适用, 如在提升老旧照片分辨率时, 对同一场景往往无法再次成像, 所以无法获得存在互补信息的 LR 图像序列, 此时 MISR 通常会失效。针对这种只有单张 LR 图像的情形, SISR 通常通过学习的方法来重建出 HR 图像。本文主要关注 SISR 方面的进展。

在经典的 SISR 方法中, 基于图像自相似的方法^[10]和基于稀疏表示的方法^[11]均取得过优异的重建效果。近年来, 由于深度学习具有强大的特征表达能力, 基于深度学习的图像 SR 胜过了经典 SISR 方法, 成为了研究的热点。该方法通过神经网络直接学习 LR 图像到 HR 图像即端到端的映射, 从而重建出 HR 图像。

自从 Dong 等^[12]首次将卷积神经网络 (convolutional neural networks, CNN) 用于图像超分辨率重建以来, 基于深度学习的图像 SR 得到快速发展, 不断涌现出新的网络结构, 重建效果也是逐步提升。如基于 SRCNN (super-resolution CNN)^[12]的改进算法^[13-14]、基于残差网络的重建方法^[15-17]、基于密集连接的重建方法^[18-20]等。

近期, 在 SISR 的研究中, 一部分学者继续关注设计更好的网络结构、更优秀的损失函数; 另一部分学者则开始关注 SISR 的新思路: 将已有的 HR 图像作为参考, 利用参考图中丰富的纹理来补偿 LR 图像缺失的细节信息, 从而缓解图像 SR 的不适定性。这种基于参考图的超分辨率重建 (reference-based super-resolution reconstruction, RefSR), 其重建过程如图 1 所示。与普通的 SISR 不同, RefSR 的输入除 LR 图像外, 还需额外输入一张或多张与 LR 图像内容或纹理相似的 HR 参考图像。众多实验结果表明, 在 Ref 图像的帮助下, RefSR 能够重建出更多真实的细节和纹理信息, 其结果优于 SISR。

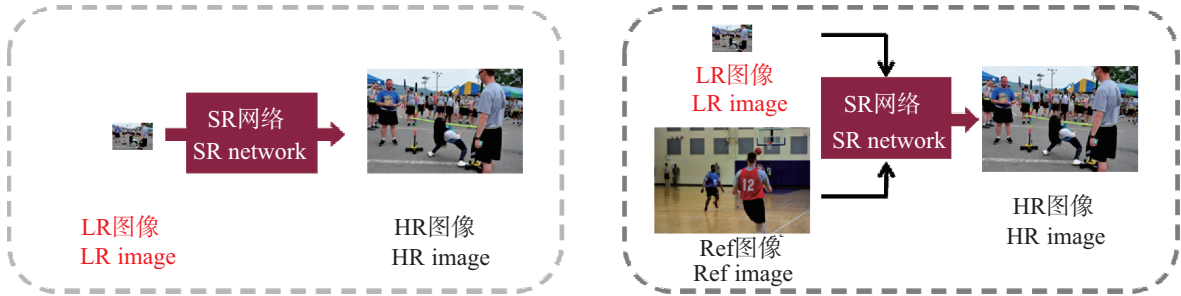


图 1 普通SISR方法与RefSR方法的对比

Fig.1 Comparison of common SISR method and RefSR method

尽管 RefSR 的相关研究开展时间较短, 但也涌现出了很多先进的重建思想和算法, 因此急需对此做一个全面的梳理。在现有 SISR 的综述性论文中, 鲜有专门介绍 RefSR。如 Wang 等^[21]从有监督、无监督和典型应用领域三个方面, 全面阐述了基于深度学习的 SISR 技术; 唐艳秋等^[22]从模型类型、网络结构、信息传递方式等方面对 SISR 方法中各种算法进行了详细评述; 曲延云等^[23]将基于深度学

习的 SISR 方法划分为有监督 SISR 方法和不成对 SISR 方法，并进行了深入分析等。这些关于超分辨率重建的综述性论文均未曾提及 RefSR 方法，更未对其进行详细阐述。因此，本文在分析和借鉴了若干图像 RefSR 研究成果的基础上，从基于图像对齐的方法和基于图像块匹配的方法入手，对近年来的 RefSR 方法进行综述性的介绍，以期抛砖引玉，为该领域的后续研究提供有益参考。

1 RefSR 的主要方法

在 RefSR 算法设计过程中，需要考虑如何更好地使用 Ref 图像。通用的做法是在 Ref 图像中寻找 LR 图像对应的高频信息，尽可能利用 Ref 图像的纹理和细节信息来帮助 LR 图像重建。如图 2 所示，依据 Ref 图像与 LR 图像的对应方式，可以将 RefSR 分为两大类：基于图像对齐的方法和基于图像块匹配的方法。基于图像对齐的 RefSR 方法是利用光流、可变形卷积等模型将输入的 LR 图像与 Ref 图像进行全局配准，再将对齐后的 Ref 图像的纹理用于 LR 图像重建，其代表模型有 SSEN、Cross-Net。而基于图像块匹配的 RefSR 方法则是将输入图像分割为若干个块，对每个块进行相似度匹配，再利用匹配后的 LR/Ref 图像进行重建。该方法代表模型有 SRNTT、TTSR、MASA 等。

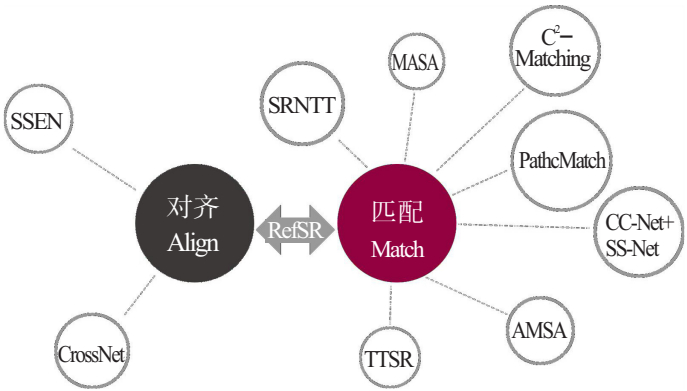


图 2 RefSR方法分类
Fig.2 Classification of RefSR methods

在 Ref 图像和 LR 图像是同一个场景的情况下，基于图像对齐的方法能够取得优异的重建效果，但是当它们的场景不相同，基于图像对齐方法的重建效果则不理想。相较而言，基于图像块匹配的方法对参考图像要求不是很高，因此成为了目前该领域的研究热点。

1.1 基于图像对齐的 RefSR 方法

由于 Ref 图像与 LR 图像存在差异，若直接将 Ref 图像中的纹理信息融入到 LR 图像中，则重建出的 HR 图像的效果不佳，所以将两幅图像对齐显得尤为重要。这种基于图像对齐的 RefSR 方法的核心是图像对齐操作，其框架如图 3 所示。

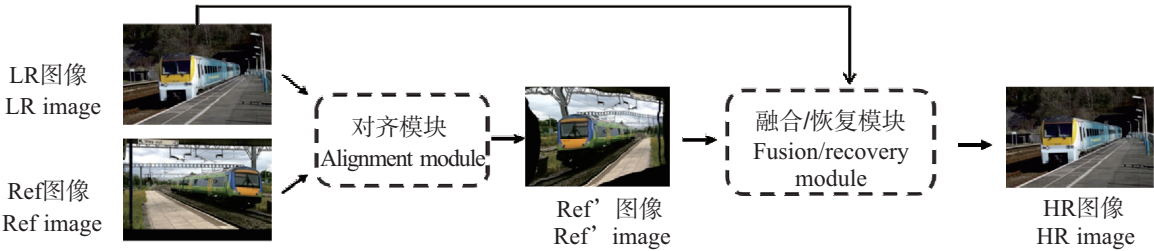


图 3 基于图像对齐的RefSR方法框架图
Fig.3 Framework of RefSR method based on image alignment

1.1.1 CrossNet

基于图像对齐方法的代表是 Zheng 等^[24]的研究，他们提出了一个跨尺度端到端的图像 SR 网络（CrossNet），其中图像对齐使用了光流法。该网络利用 SISR 方法对 LR 图像进行上采样，得到与 Ref 图像相同大小的上采样图，再分别提取出上采样图和 Ref 图像的多尺度特征；然后，利用改进的 FlowNetS 模型^[25]学习不同尺度 LR/Ref 图像特征的光流信息来更新 Ref 图像，从而实现图像对齐操作；最后，通过融合操作重建出 HR 图像。该方法假设两幅图像具有较强的相似性，所以当 LR/Ref 图像相关性不强时，效果会有所下降。

1.1.2 SSEN

由于 CrossNet 中光流对齐的成本高, 且使用其他预训练好的网络进行光流估计是不准确的, 所以 Shim 等^[26]提出了基于可变形卷积的网络结构 SSEN (similarity search and extraction network, SSEN)。SSEN 利用可变形卷积来寻找 LR/Ref 图像的对应关系, 且用动态偏移估计器对可变形卷积的偏移量进行估计。同时, 为了捕获特征内部和特征之间的全局相关性, Shim 等^[26]在动态偏移估计器中加入了非局部块。该方法能够处理非刚性变换的图像, 计算量较小, 但是无法解决长距离对应问题。

1.2 基于图像块匹配的 RefSR 方法

由于对图像进行全局配准要求图像具有较强的相似性, 当 Ref 图像与 LR 图像只是纹理相似或者内容相似时, 基于对齐的 RefSR 方法的效果将直线下降。针对以上问题, 一些学者开始研究用图像块匹配的方式来寻找 Ref 图像与 LR 图像的对应关系。这种基于图像块匹配的 RefSR 方法的核心是图像块匹配模型的构建以及如何处理好匹配后的不对齐问题, 其框架如图 4 所示, 其中图 4a 为 LR/Ref 图像块匹配示意图, 图 4b 为基于图像块匹配的 RefSR 方法的流程图。

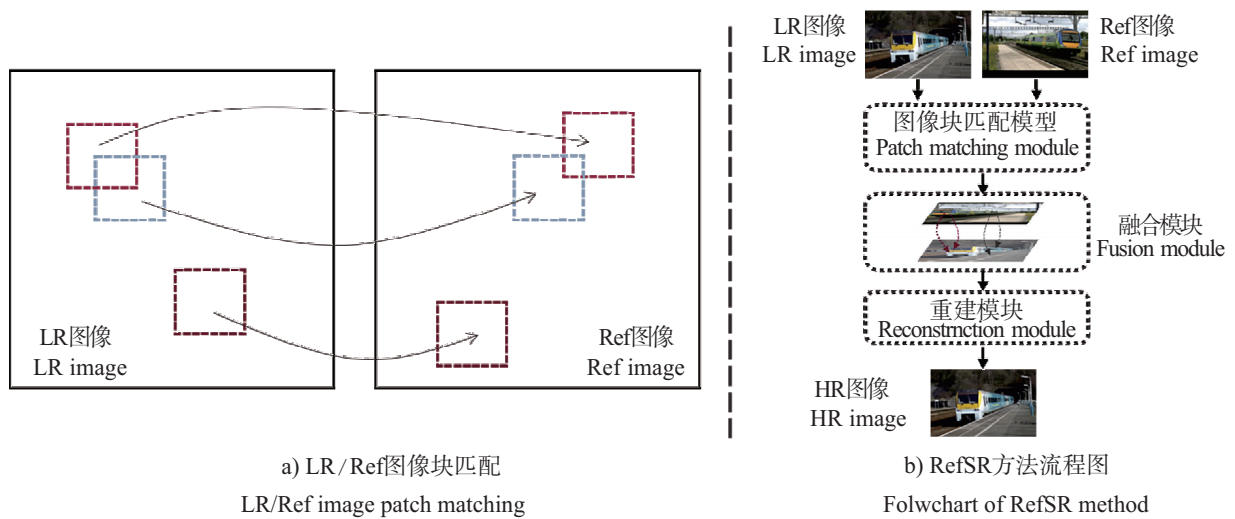


图 4 基于图像块匹配的RefSR方法框架图

Fig.4 Framework of RefSR method based on image patch matching

1.2.1 图像块匹配

基于图像块匹配的方法可以追溯到 Boominathan 等^[27]的研究。他们在提高光场 (light field, LF) 成像的分辨率和景深时, 将单反数码相机捕获的 HR 图像作为 Ref 图像, 指导 LF 图像的恢复。其算法核心是 Ref 图像的使用方式: Ref 图像下采样后, 计算其一阶、二阶梯度, 利用欧式距离在梯度特征图中寻找与 LR 图像最相邻的 9 个块, 然后加权平均得到融合结果。在后续研究中, 该方法常被称作块匹配 (patch match), 它的匹配过程是在原始像素空间上进行的, 操作简单, 但匹配过程中没有充分利用高频信息, 在融合阶段也只是简单地加权平均, 导致高频信息没有得到很好地融合。

1.2.2 CC-Net + SS-Net

针对图像块匹配方法的问题, Zheng 等^[28]使用 CC-Net (cross-scale correspondence network) 模块, 通过卷积神经网络提取 LR/Ref 图像特征, 利用内积计算相似度, 选择相似度最高的图像块作为匹配对, 然后, 将匹配好的图像块送入到 SS-Net (super-resolution synthesis network) 模块进行多尺度融合, 最后得到 SR 图像。相对于图像块匹配方法中的梯度特征, 该方法使用 CNN 提取的特征进行匹配, 其准确度更高; 但是在融合阶段是对每个图像块单独进行重建, 所以该方法容易产生块状效应。

1.2.3 SRNTT

受图像风格迁移的启发, Zhang 等^[29]将 RefSR 作为纹理迁移问题, 设计了一个端到端的 SRNTT (super-resolution by neural texture transfer) 模型: 首先, 对提取的图像特征进行密集块匹配, 然后将

匹配后的 Ref 图像进行纹理迁移，从而使得重建的 SR 图像拥有丰富的纹理信息。与图像块匹配方法不同，SRNTT 方法是在特征空间进行匹配，虽然该方法促进了多尺度的纹理迁移，允许模型从语义相关的角度实现图像块匹配，但是其计算量大，不利于实际应用。

1.2.4 TTSR

Yang 等^[30]指出 SRNTT 忽略了全局信息，这使得现有方法对某类 LR 图像重建效果不佳。针对此问题，他们提出了一个包含注意力机制的 SR 网络架构 TTSR（texture transformer network for image super-resolution）。该网络通过使用注意力机制挖掘深层次的特征对应关系，再将匹配好的特征送入跨尺度融合模块，最后得到 SR 图像。TTSR 的特征匹配由 4 个模块组成，包括纹理特征提取模块、相关性嵌入模块、用于纹理迁移的硬注意力模块和用于纹理合成的软注意力模块。而在相关性嵌入模块，他们将 LR 图像中提取的特征作为转换器中的查询（query），将 Ref 图像中提取的特征作为转换器中的键（key），以获得硬注意力矩阵和软注意力矩阵。由于作为键的 Ref 图像是经过下采样后再上采样的图像，该操作无法避免信息丢失，从而导致匹配精确度下降，特别是在细节区域下降得较为明显。

1.2.5 MASA

目前，虽然密集图像块匹配方法的效果最好，但其计算量巨大且耗费存储空间；所以，Lu 等^[31]提出了一个粗-细对应匹配方案，称为 MASA（matching acceleration and spatial adaptation for RefSR）：首先对 LR/Ref 图像进行一个大尺度的图像块匹配，再利用图像的局部相似性，对大尺度图形成块分块进行小尺度匹配。该方法可以大幅减少计算量。同时，虽然 Ref 图像与 LR 图像具有相似的内容或纹理，但其颜色和亮度可能不一样，若直接进行融合操作则效果不佳。所以他们在进行特征融合之前，加入了空间自适应模块，对得到的 Ref 特征图重新映射，使得模型对颜色和亮度更具有鲁棒性。

1.2.6 AMSA

虽然 MASA 方法的计算量有所下降，但是其计算复杂度仍是图像尺寸的平方级；所以，Xia 等^[32]对 2009 年 Barnes 等^[33]提出的图像块匹配方法进行了改进，提出了 CFE-PatchMatch（coarse-to-fine embedded patch match）快速匹配方法，并将其应用于 RefSR 中，使 LR/Ref 图像的匹配计算量从平方级降到了接近线性级。为了避免尺度不对齐问题，Xia 等^[32]提出了 AMSA（accelerated multi-scale aggregation network）模型。其中，针对小规模的不对齐问题，提出了动态融合模块；针对大尺度不对齐问题，提出了多规模融合模块。这两个融合模块相互配合，产生了很好的融合效果。

1.2.7 C²-Matching

Jiang 等^[34]认为由于在输入 LR 图像与 Ref 图像之间存在两个差距——变换差距（如旋转、缩放等）和分辨率差距，使得执行局部迁移变得较为困难。针对这些问题，他们提出了 C²-Matching（cross transformation and cross resolution matching）技术，用于跨变换和跨分辨率的关系匹配。对于变换差距，利用对比网络拉近匹配对之间的距离，疏远不匹配对之间的距离；对于分辨率差距，提出教师-学生关系蒸馏网络，与传统的知识蒸馏网络不同，此网络用 HR-HR 匹配来指导相对困难的 LR-HR 匹配。然后，通过设计的动态融合模块来解决潜在的错位问题。该技术在目前常用的数据集中，都显现出了极强的泛化能力，以及对大尺度和旋转变换的鲁棒性。

2 损失函数、数据集与评价标准

2.1 损失函数

损失函数在提升模型性能方面起着关键作用，常用于计算模型生成的 HR 图像与原始基准图像的差异，从而引导 RefSR 网络重建出更好的 HR 图像。损失函数的选择也尤为重要，一个好的损失函数能够帮助模型更优和更快地收敛，同时得到具有高精度，且与真实 HR 图像接近的 HR 图像。在 RefSR 方法中常用的损失函数主要有以下几种：

1) 重建损失（reconstruction loss） 重建损失是图像 SR 中最常用的损失函数，一般可分为 L_1 损

失和 L_2 损失,其表达式分别为:

$$L_{\text{rec_l1}}(I^{\text{SR}}, I) = \frac{1}{hwc} \sum_{i,j,k} |I_{i,j,k}^{\text{SR}} - I_{i,j,k}|,$$

$$L_{\text{rec_l2}}(I^{\text{SR}}, I) = \frac{1}{hwc} \sum_{i,j,k} (I_{i,j,k}^{\text{SR}} - I_{i,j,k})^2.$$

其中: h , w 和 c 分别表示图像的长、宽和通道数; I^{SR} 表示生成的 HR 图像; I 表示真实图像。

2) 感知损失 (perceptual loss) 感知损失已被证明可用于提升视觉质量,并在图像 SR 中广泛应用。感知损失的核心思想是增强预测图像和目标图像在特征空间中的相似性,其公式为:

$$L_{\text{per}} = ||\varphi_i(I^{\text{SR}}) - \varphi_i(I)||_2.$$

其中: φ_i 表示某种特征提取网络的第 i 层。

3) 对抗损失 (adversarial loss) 在图像 SR 领域中,只需将 SR 模型作为一个生成器,再定义一个鉴别器来判断输入图像是否由生成器产生,就可以构造出一个对抗学习网络。而对抗损失能够有效地使模型产生更加真实的图像,其具体表达式为:

$$L_{\text{gan_ce_g}} = -\lg D(I^{\text{SR}}),$$

$$L_{\text{gan_ce_d}} = -\lg D(I) - \lg(1 - D(I^{\text{SR}})).$$

其中: $D(\cdot)$ 表示判别器; $L_{\text{gan_ce_d}}$ 是鉴别器的对抗损失; $L_{\text{gan_ce_g}}$ 是 SR 模型生成器的对抗损失。

4) 纹理损失 (texture loss) Zhang 等^[29]将纹理损失引入到 RefSR 方法中,来减少参考图像与生成图像的纹理差异,其表达式为:

$$L_{\text{tex}} = \sum_l \lambda_l ||G_r(\varphi_l(I^{\text{SR}}) \odot S_l^*) - G_r(\varphi_l(I^{\text{Ref}}) \odot S_l^*)||_F.$$

其中: $G_r(\cdot)$ 表示格莱姆矩阵; λ_l 表示第 l 层特征图对应的正规范化因子; S_l^* 表示所有 LR 图像块计算得到的最佳匹配分数矩阵; \odot 为 Hadamard 乘积,表示矩阵对应元素相乘; $||\cdot||_F$ 表示 Frobenius 范数; φ_l 表示第 l 层特征图。

2.2 数据集

在 RefSR 研究中,最常用的数据集为 CUFED5,该数据集是 Zhang 等^[24]创建的,包含 1 个训练集 (11 871 个 160 px × 160 px 的图像对) 和 1 个测试集 (126 组图像)。在测试集中,每组图像包含 1 张 HR 图像和 5 个不同相似度的参考图像,如图 5 所示。



图 5 CUFED5 测试集中的两组图

Fig.5 Two-group images in the CUFED5 test set

由于 CUFED5 数据集中占比最重的为室内、人、室外类型的图像,其他类型的图像较少,最近 Jiang 等^[34]给出了一个新的数据集 WR-SR,该数据集包含了 80 个图像对。相比于 CUFED5 数据集,WR-SR 数据集还包含了动物、地标、建筑等图像类,其内容更加丰富,如图 6 所示。除此之外,常用的测试集还有 Sun80^[35] (80 张自然图像且有一些参考图像)、Urban100^[36] (100 张建筑图像,无参考图像但图像有自相似性)、Manga109^[37] (109 张漫画,无参考图像)。



图 6 WR-SR数据集的示例图
Fig.6 Example of WR-SR datasets

2.3 评价标准

图像质量评价主要包括主观评价和客观评价两类。在 RefSR 研究中，常用的客观评价指标有峰值信噪比（peak signal to noise ratio, PSNR）和结构相似性（structural similarity index, SSIM）^[38]；主观评价一般采用主观质量评分法（mean opinion score, MOS）^[39]。在图像 SR 中，使用对抗损失、感知损失等会获得更好的视觉效果，但是客观指标 PSNR 和 SSIM 会有所下降；所以，在使用 PSNR/SSIM 评价 RefSR 方法时，不仅要考虑基于所有损失（感知损失、对抗损失、重建损失等）的模型评价，还要单独评价只由重建损失训练的模型。

3 几种 SR 方法对比

自从卷积神经网络被引入图像 SR 以来，各种基于深度学习的图像 SR 方法不断涌现。近年来，RefSR 方向吸引了大批学者，他们基于深度神经网络开展了很多卓有成效的研究。本文通过比较总结上述具有代表性的 RefSR 方法，概括了不同 RefSR 方法的网络特点及算法优点，详情如表 1 所示。

表 1 RefSR 方法对比

Tab.1 Comparison of RefSR methods

方法 Methods	网络特点 Network Features	算法优点 Algorithm Advantages
CrossNet	SISR 上采样 + 光流对齐	端到端,全局对齐,速度较快
	SISR upsampling + optical flow alignment	End-to-end, global alignment, faster
SSEN	可变形卷积 + 动态偏移估计 + 非局部块	能够处理非刚性变换的图像,计算量较小
	Deformable convolution + dynamic offset estimation + non-local blocks	Capable of handling non-rigidly transformed images, less computationally intensive
PatchMatch	一二阶梯度 + 最小 L_2 距离	方法简单,操作性强
	First and second order gradient + minimum L_2 distance	The method is simple and operable
CC-Net + SS-Net	CNN 特征提取 + 内积	匹配更准确
	CNN feature extraction + inner product	Match more accurately
SRNTT	多层次密集图像块匹配 + 纹理损失	不受 Ref 图像全局结构的限制
	Multi-level dense image patch matching + texture loss	Not limited by the global structure of the Ref image
TTSR	注意力机制 + 多尺度交叉特征融合	挖掘深层次的特征对应关系
	Attention mechanism + multi-scale cross feature fusion	Mining deep-level feature correspondences
MASA	粗-细对应匹配 + 空间自适应模块	节省计算时间,对颜色和亮度更具鲁棒性
	Coarse-fine correspondence matching + space adaptive module	Saves computation time and is more robust to color and brightness
AMSA	粗-细嵌入式图像块匹配 + 多尺度动态聚合	接近线性级的计算复杂度
	Coarse-fine embedded image patch matchin + multi-scale dynamic aggregation	Computational complexity close to linear level

续表

方法 Methods	网络特点 Network Features	算法优点 Algorithm Advantages
C ² -Matching	对比学习网络 + 教师-学生关系蒸馏网络 + 可变形卷积	对图像尺寸和位置更具鲁棒性,提高细节 区域的视觉质量
	Contrastive learning network + teacher-student relational distillation network + deformable convolution	More robust to image size and position, improving visual quality in detail regions

如前所述,不管是基于图像对齐的 RefSR 方法,还是基于图像块匹配的方法,都关注着 LR 图像与 Ref 图像的对应问题;但是它们又各有侧重,如表 1 所示,这些方法各具特色,各有优点。具体归纳如下:

1) 在方法的选择上,基于图像块匹配的方法是 RefSR 中的主流方法,其对图像的相关性要求没有基于图像对齐的方法高。此外,现有部分学者也开始研究多参考图的图像 SR 方法,如 Yan 等^[40]提出的 CIMR-SR 模型等。

2) 在参考图像的选择上,参考图像与 LR 图像内容越相似,RefSR 方法重建出的 HR 图像效果越好。表 2 为不同方法在不同相似度的参考图像的帮助下,重建出的 HR 图像的 PSNR 和 SSIM。其中:L1 到 L4 表示输入相似度依次递减的参考图像,即 L1 是与 LR 图像相似度最高的 HR 图像,L4 为相似度最低的 HR 图像;LR 表示用 LR 图像作为输入的参考图像。由表 2 可以看到,算法的重建效果与参考图像的选择是有关系的,参考图像与 LR 图像相似度越高,重建效果越好。图 7 为图 5 展示的两幅 CUFED5 测试集图像基于不同相似度参考图像的细节重建结果,GT 表示真实 HR 图像,可以看到,参考图像与 LR 图像相似度越高,重建出的 HR 图像的细节恢复得越好,这也证明了参考图像对 Ref-SR 重建的结果有一定的影响。

表 2 不同相似度的参考图像对 PSNR 和 SSIM 的影响
Tab.2 Influence of reference images with different similarities on PSNR and SSIM

相似度水平 Similarity level	L1		L2		L3		L4		LR	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
CrossNet	25.48	0.764	25.48	0.764	25.47	0.763	25.46	0.763	25.46	0.763
SSEN	26.78	0.791	26.52	0.783	26.48	0.782	26.42	0.781	26.38	0.781
SRNTT-rec	26.15	0.781	25.98	0.776	25.98	0.775	25.95	0.774	25.91	0.776
TTSR-rec	26.99	0.800	26.74	0.791	26.64	0.788	26.58	0.787	26.43	0.782
MASA-rec	27.45	0.814	26.92	0.796	26.83	0.793	26.74	0.790	26.59	0.784
C ² -Matching-rec	28.24	0.841	27.39	0.813	27.17	0.806	26.94	0.799	26.52	0.784
AMSA-rec	28.58	0.849	27.52	0.816	27.25	0.809	27.04	0.803	26.63	0.789

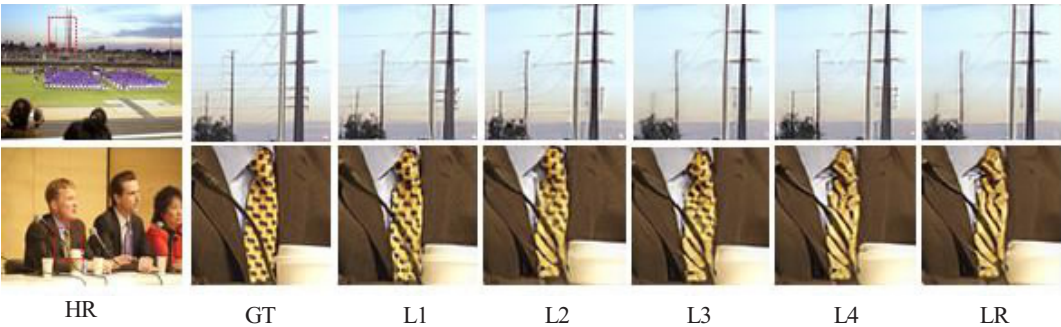


图 7 不同相似度的参考图像对结果的视觉影响
Fig.7 Visual effects of reference images with different similarities on the results

3) 从客观评价标准来看, RefSR 方法的效果要优于现有的 SISR 方法。表 3 为不同的图像 SR 方法在 CUFED5、Sun80、Urban100、Manga109 测试集上的实验结果, 实验结果均取自相关文献。这里除了比较 RefSR 不同算法的效果, 也比较了主流的 SISR 算法。RefSR 方法不仅比较了原始模型的 PSNR 和 SSIM, 还对比了只采用重建损失训练模型(用-rec 表示)的 PSNR 和 SSIM。由结果可见, 与现有的 SISR 方法相比, RefSR 方法的结果有大幅提升。

表 3 不同 SR 方法的 PSNR 和 SSIM
Tab.3 PSNR and SSIM of different SR methods

算法 Methods		CUFED5		Sun80		Urban100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
SISR	SRCNN ^[12]	25.33	0.745	28.26	0.781	24.41	0.738	27.12	0.850
	EDSR ^[17]	25.93	0.777	28.52	0.792	25.51	0.783	28.93	0.891
	RCAN ^[41]	26.06	0.702	29.86	0.810	25.42	0.768	29.38	0.895
	SRGAN ^[39]	24.40	0.702	26.76	0.725	24.07	0.729	25.12	0.802
	ENet ^[42]	24.24	0.695	26.24	0.702	23.63	0.711	25.25	0.802
	ESRGAN ^[43]	21.90	0.633	24.18	0.651	20.91	0.620	23.53	0.797
	RankSRGAN ^[44]	22.31	0.635	25.60	0.667	21.47	0.624	25.04	0.803
RefSR	CrossNet	25.48	0.764	28.52	0.793	25.11	0.765	23.36	0.741
	SSEN	25.35	0.742	—	—	—	—	—	—
	SSEN-rec	26.78	0.791	—	—	—	—	—	—
	SRNTT	25.61	0.764	27.59	0.756	25.09	0.774	27.54	0.862
	SRNTT-rec	26.24	0.784	28.54	0.793	25.50	0.783	28.95	0.885
	TTSR	25.53	0.765	28.59	0.774	24.62	0.747	28.70	0.886
	TTSR-rec	27.09	0.804	30.02	0.814	25.87	0.784	30.09	0.907
	MASA	24.92	0.729	27.12	0.708	23.78	0.712	—	—
	MASA-rec	27.54	0.814	30.15	0.815	26.09	0.786	—	—
	C ² -Matching	27.16	0.805	29.75	0.799	25.52	0.764	29.73	0.893
	C ² -Matching-rec	28.24	0.841	30.18	0.817	26.03	0.770	29.79	0.896
	AMSA	27.31	0.809	29.83	0.803	25.60	0.770	29.79	0.896
	AMSA-rec	28.50	0.849	30.29	0.819	26.18	0.789	30.57	0.914

4) 从视觉效果上看, 在 RefSR 方法中, Ref 图像为 LR 图像提供了更多的细节纹理信息。图 8 为目前效果最好的几种图像 SR 方法的效果(实验结果均取自相关文献), 并展示了每张图像的低分辨率图及其参考图像, 以及两种 SISR 方法(ESRGAN 和 BankSRGAN)和三种 RefSR 方法(MASA、C²-Matching、AMSA)对某个细节的重建效果。从图 8 可以看到, RefSR 方法在人脸、字母等细节信息的恢复上效果更好, 恢复出的细节信息更真实, 其中 AMSA 模型效果是 5 种方法中最好的。

5) 在 RefSR 方法中, 模型中加入注意力机制、非局部块等, 能大幅提升图像 SR 的效果。复杂的网络带来了好的重建效果, 但也出现了计算量大、存储空间不足等问题。针对这些问题, 研究者们对网络的设计不仅会考虑图像重建效果, 还会考虑模型的计算量、参数量, 使得模型更有利于实际应用。表 4 列出现有模型的参数量、计算量, 可以看到 AMSA 模型的计算量是最小的, 且其效果也是目前最好的。

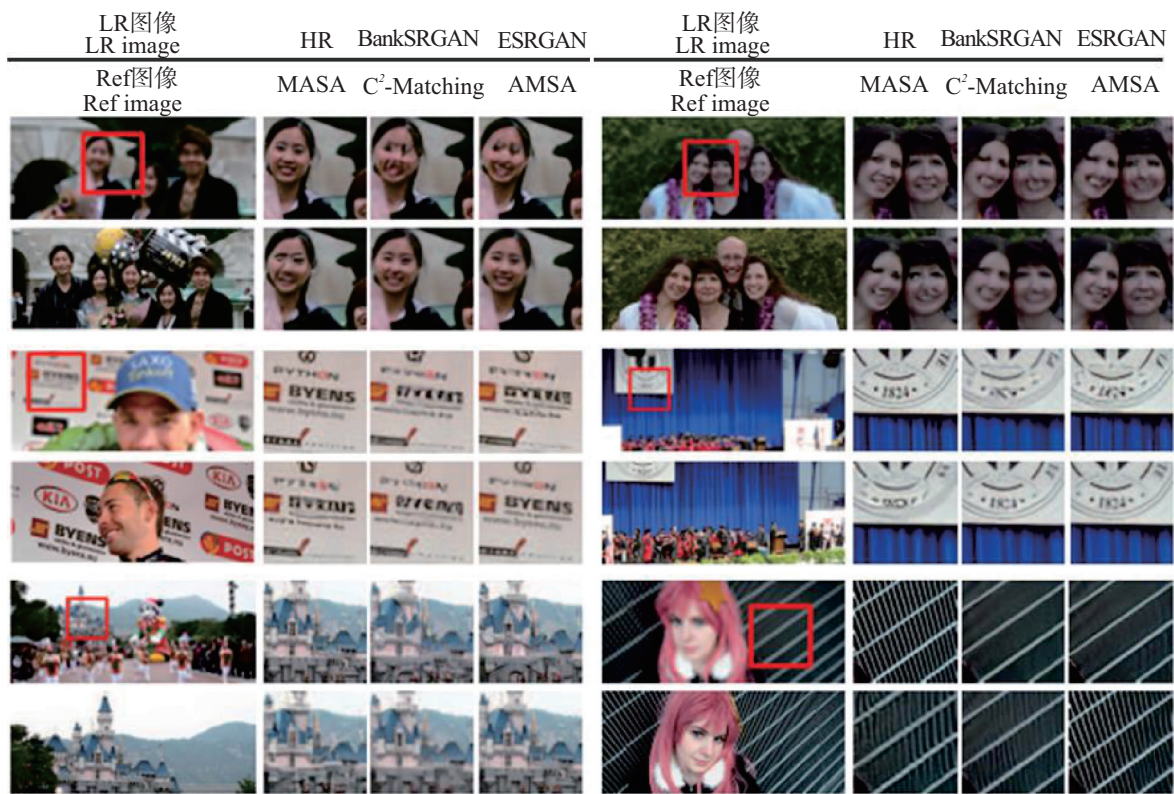


图 8 不同SR方法的视觉对比图

Fig.8 Visual comparison of different SR methods

表 4 RefSR 中各模型参数数量和计算量

Tab.4 Parameters and FLOPS of each model in RefSR

算法 Methods	CrossNet	SRNTT	TTSR	MASA	AMSA
参数量 Paramerers/ 10^6	35.18	5.75	6.99	4.03	4.12
计算量 Calculations/ 10^9	348.31	6500.70	1044.28	367.93	279.52

4 未来研究展望

图像 SR 作为计算机视觉领域底层视觉部分的经典问题，一直以来受到人们的广泛关注。目前，基于深度学习的 RefSR 方法取得了不错的效果。尽管如此，但仍存在一些问题有待解决：

1) 现有的 RefSR 方法中，LR 图像通常是由 HR 图像下采样得到的，以此模拟自然情况下图像退化过程，而在实际应用中图像的退化过程还包括噪声、模糊、压缩等。如何学习得到可以处理复杂退化图像的网络还有待探讨。

2) 轻量化的 RefSR 方法。受算力的限制，现有 RefSR 方法无法在手机、平板等移动设备上直接使用，所以，如何将现有的 RefSR 方法轻量化，在保证效果的同时，降低参数量和模型复杂度是 RefSR 研究的热点。

3) 在 SISR 方法中，现有的网络模块（如递归学习、密集连接等）和学习机制（如多路径学习、非局部相似性等）对 SISR 精度有很好的提升。在 RefSR 方法中，能否运用这些模块来进一步提升算法的性能，值得进一步去探讨。

[参考文献]

- [1] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(10): 3365-3387.
- [2] RASTI P, UIBOUPIN T, ESCALERA S, et al. Convolutional neural network super resolution for face recognition in surveillance monitoring[C]//International Conference on Articulated Motion and Deformable Objects. Cham: Springer, 2016: 175-184.
- [3] ARMANIOUS K, JIANG C, FISCHER M, et al. MedGAN: medical image translation using GANs[J]. Computerized Medical Imaging and Graphics, 2020, 79: 101684.
- [4] LU T, WANG J, ZHANG Y, et al. Satellite image super-resolution via multi-scale residual deep neural network[J]. Remote Sensing, 2019, 11(13): 1588.
- [5] 杨亚霖. 基于探墙雷达的图像超分辨率重建算法研究[D]. 南昌: 南昌大学, 2020.
- [6] 张东晓, 鲁林, 李翠华, 等. 基于亚像素位移的超分辨率图像重建算法[J]. 自动化学报, 2014, 40(12): 2851-2861.
- [7] ZHANG D, JODOIN P M, LI C, et al. Novel graph cuts method for multi-frame super-resolution[J]. IEEE Signal Processing Letters, 2015, 22(12): 2279-2283.
- [8] BHAT G, DANELLJAN M, VAN GOOL L, et al. Deep burst super-resolution[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 9209-9218.
- [9] BHAT G, DANELLJAN M, YU F, et al. Deep reparametrization of multi-frame super-resolution and denoising[C]// IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 2460-2470.
- [10] GLASNER D, BAGON S, IRANI M. Super-resolution from resolution from a single image[C]// IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2009: 349-356.
- [11] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [12] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(2): 295-307.
- [13] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]//European Conference on Computer Vision. Cham: Springer, 2016: 391-407.
- [14] SHI W, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 1874-1883.
- [15] GENG J Q, ZHANG D X. Large-factor single image super-resolution based on back projection and residual block[C]// Twelfth International Conference on Graphics and Image Processing (ICGIP 2020). Xi'an: ICGIP, 2021: 11720. DOI: 10.1117/12.2589365.
- [16] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 1646-1654.
- [17] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 136-144.
- [18] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network[C]// European Conference on Computer Vision. Cham: Springer, 2018: 252-268.
- [19] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 2472-2481.
- [20] TAI Y, YANG J, LIU X, et al. Memnet: a persistent memory network for image restoration[C]// IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 4539-4547.
- [21] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: a survey[J]. IEEE Transactions on pattern analysis and machine intelligence, 2020, 43(10): 3365-3387.
- [22] 唐艳秋, 潘泓, 朱亚平, 等. 图像超分辨率重建研究综述[J]. 电子学报, 2020, 48(7): 1407-1420.
- [23] 曲延云, 陈蓉, 李翠华, 等. 深度学习单帧图像超分辨率重建研究综述[J]. 厦门大学学报(自然科学版), 2021, 60

(3):555-570,641-642.

- [24] ZHENG H,JI M,WANG H,et al. Crossnet:an end-to-end reference-based super resolution network using cross-scale warping[C]// European Conference on Computer Vision. Cham:Springer,2018:88-104.
- [25] DOSOVITSKIY A,FISCHER P,ILG E,et al. FlowNet:learning optical flow with convolutional networks[C]// IEEE/CVF International Conference on Computer Vision. Piscataway,NJ:IEEE,2015:2758-2766.
- [26] SHIM G,PARK J,KWEON I S. Robust reference-based super-resolution with similarity-aware deformable convolution[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2020:8425-8434.
- [27] BOOMINATHAN V,MITRA K,VEERARAGHAVAN A. Improving resolution and depth-of-field of light field cameras using a hybrid imaging system[C]// IEEE International Conference on Computational Photography. Piscataway,NJ:IEEE,2014:1-10.
- [28] ZHENG H,JI M,HAN L,et al. Learning cross-scale correspondence and patch-based synthesis for reference-based super-resolution[C]// British Machine Vision Conference. London,UK:British Machine Vision Association,2017:2.
- [29] ZHANG Z,WANG Z,LIN Z,et al. Image super-resolution by neural texture transfer[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2019:7982-7991.
- [30] YANG F,YANG H,FU J,et al. Learning texture transformer network for image super-resolution[C]// IEEE/CVF conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2020:5791-5800.
- [31] LU L,LI W,TAO X,et al. Masa-sr:matching acceleration and spatial adaptation for reference-based image super-resolution[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2021:6368-6377.
- [32] XIA B,TIAN Y P,HANG Y C,et al. Coarse-to-fine embedded patchmatch and multi-scale dynamic aggregation for reference-based super-resolution[C]// AAAI Conference on Artificial Intelligence. Menlo Park,CA:AAAI,2022:2768-2776.
- [33] BARNES C,SHECHTMAN E,FINKELSTEIN A,et al. PatchMatch:a randomized correspondence algorithm for structural image editing[J]. ACM Trans Graph,2009,28(3):24.
- [34] JIANG Y,CHAN K C K,WANG X,et al. Robust reference-based super-resolution via C2-matching[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Pittsburgh,PA,USA:IEEE,2021:2103-2112.
- [35] SUN L,HAYS J. Super-resolution from internet-scale scene matching[C]// IEEE/CVF International Conference on Computer Vision. Piscataway,NJ:IEEE,2012:1-12.
- [36] HUANG J B,SINGH A,AHUJA N. Single image super-resolution from transformed self-exemplars[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2015:5197-5206.
- [37] MATSUI Y,ITO K,ARAMAKI Y,et al. Sketch-based manga retrieval using manga109 dataset[J]. Multimedia Tools and Applications,2017,76(20):21811-21838.
- [38] WANG Z,BOVIK A C,SHEIKH H R,et al. Image quality assessment:from error visibility to structural similarity[J]. IEEE Transactions on Image Processing,2004,13(4):600-612.
- [39] LEDIG C,THEIS L,HUSZAR F. Photo-realistic single image super-resolution using a generative adversarial network[C]// IEEE/CVF International Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE,2017:4681-4690.
- [40] YAN X,ZHAO W,YUAN K,et al. Towards content-independent multi-reference super-resolution:adaptive pattern matching and feature aggregation[C]//European Conference on Computer Vision. Cham:Springer,2020:52-68.
- [41] ZHANG Y,LI K,LI K,et al. Image super-resolution using very deep residual channel attention networks[C]// European Conference on Computer Vision. Cham:Springer,2018:286-301.
- [42] SAJJADI M S M,SCHOLKOPF B,HIRSCH M. Enhancenet:Single image super-resolution through automated texture synthesis[C]// IEEE/CVF International Conference on Computer Vision. Piscataway,NJ:IEEE,2017:4491-4500.
- [43] WANG X,YU K,WU S,et al. Esrgan:enhanced super-resolution generative adversarial networks[C]// European Conference on Computer Vision. Cham:Springer,2018:63-79.
- [44] ZHANG W,LIU Y,DONG C,et al. Ranksrgan:generative adversarial networks with ranker for image super-resolution[C]// IEEE/CVF International Conference on Computer Vision. Piscataway,NJ:IEEE,2019:3096-3105.

(责任编辑 朱雪莲 英文审校 黄振坤)