

图像超分辨率研究进展（2020–2025）

CNN 方法

近年来传统的卷积网络（CNN）仍然是超分辨率（SR）研究的基准，研究重点转向轻量化和更大感受野设计。一类代表性工作是BCRN（2024），该模型采用ConvNeXt残差块和BSConv深度可分离卷积，结合空间注意力和通道注意力模块，实现了极小参数量下的高性能^{1 2}。BCRN在Set14数据集（×2）上的PSNR达33.66 dB³，超过了同期方法；其代码已开源（GitHub: [kptx666/BCRN](#)）⁴。另一种轻量化结构是LrfSR（2025），通过扩张卷积构建大感受野的信息蒸馏模块，并引入高效通道和空间注意力（ECCA，SESA），在Set5/Set14等数据集上分别取得32.23 dB/28.65 dB的PSNR⁵；开源地址为[wanqin557/LrfSR](#)⁶。这些工作展示了在有限计算资源下改进PSNR/SSIM指标的效果。总体来看，CNN方法在指标（如PSNR、SSIM）上已有很高水平，但仍面临如何进一步提升感知质量、加速推理和拓展到真实模糊场景的挑战。

- **BCRN (2024)** – 使用ConvNeXt残差块、BSConv和注意力模块构建超轻量化SR网络，参数仅约0.29M，却达到了轻量级方法的最佳性能^{1 2}。在Set5/Set14/B100/Urban100/Manga109等基准（×2）上的PSNR分别为38.04/33.66/32.21/32.28/38.99⁷（超越其他轻量级模型），同时具备极低的计算量；代码开源⁴。
- **LrfSR (2025)** – 提出基于大感受野的信息蒸馏模块（利用扩张卷积）和高效注意力机制（ECCA通道注意力、SESA空间注意力）^{8 5}。在Set5/Set14/B100/Urban100/Manga109（×4）上PSNR为32.23/28.65/27.59/26.36/30.53 dB，略优于此前工作；模型参数量小，注重在轻量化同时捕获更丰富的上下文⁵。模型代码可在[wanqin557/LrfSR](#)上获得⁶。
- （可选）**其他轻量CNN** – 如IMDN、LapSRN等早期轻量网络通过多级信息蒸馏或残差设计减少参数。最新研究继续通过可分离卷积、Dense连接和注意力优化等手段提升性能。

对抗生成网络 (GAN) 方法

GAN方法强调感知质量，以真实感著称，但通常牺牲部分PSNR。典型代表是ESRGAN（2018）和其改进版。真实场景超分方面，Real-ESRGAN（2021）通过在训练合成数据中引入高阶复杂退化（多级模糊、噪声、JPEG等）和U-Net判别器，显著改善了真实图像的重建效果^{9 10}。实验证明其在人眼观感上优于先前方法^{9 10}（例如Real-ESRGAN模型的生成结果在真实场景中细节更丰富，噪点伪影更少），并提供了开源代码¹¹。另一类工作如BSRGAN（ICCV 2021）针对盲SR设计了更复杂的退化过程（随机shuffle模糊+下采样+噪声等），并训练“PSNR导向+感知导向”双分支模型。BSRGAN在多个合成退化数据集上获得了最佳的LPIPS评分（感知指标最低），虽然PSNR相对较低¹²，说明它在感知质量上优于其他方法。总的来说，GAN方法在近年仍不断改进对抗训练稳定性和泛化能力，但面临指标与视觉效果的权衡问题。

- **Real-ESRGAN (2021)** – 将ESRGAN扩展到真实退化SR，提出高阶退化模型（双阶段Blur/Down+Noise+JPEG等）和U-Net判别器^{9 10}。在多种真实世界图像上实现了优越的视觉效果（细节更丰富，伪影更少）^{9 10}；代码开源¹¹。该方法表明，通过模拟真实退化可显著提升SR模型的现实应用性能。
- **BSRGAN (2021)** – 设计了一个随机shuffle退化流程并训练两阶段模型（BSRNet和BSRGAN），BSRGAN分支优化感知损失。实验证明BSRGAN在感知指标上表现最好，整体LPIPS最低（最佳视觉质量），虽然PSNR指标略低于传统方法¹²。该工作强调了如何平衡通用性退化建模与生成图像质量。
- （可选）**多尺度GAN变体** – 如RankSRGAN、SFTGAN等继续在感知-失真权衡上探索，通过多级损失和辅助正则改善感知质量。开放问题是如何同时获得高PSNR和高感知质量，以及如何在真实模糊场景中训练稳定。

Transformer 架构方法

Transformer 及其变种在SR中崭露头角，优势是捕获长程依赖和全局信息。目前主要包括纯Transformer和混合CNN-Transformer结构。**SwinIR** (2021) 利用滑窗自注意力提出了基于Swin Transformer的图像恢复模型，在经典SR任务上相对于CNN基线实现了 $\sim 0.14\text{--}0.45\text{ dB}$ 的PSNR提升^[13]，同时参数量大幅减少（相较于RCAN等节省约67%）。其代码已开源^[14]【41】。**Restormer** (CVPR2022) 提出通道+空间混合注意力（MDTA模块）和门控卷积前馈（GDFN），在多项恢复任务（包括去雨、去噪和SR）上均达到了SOTA水平^[15]^[16]。Restormer证明了编码-解码Transformer可高效建模局部与全局特征。**HAT** (arXiv 2023/2024) 结合通道注意力与窗口注意力，引入交叠跨窗口注意（cross-window attention）以增强不同窗口间信息流通，并采用同任务预训练策略，全方位提升了SR性能^[17]；其分数在基准和真实场景数据集上均达到最新SOTA，代码公开^[18]【48】。此外，多项混合模型如**ESRT** (CVPRW2022) 和**IFIN** (SciRep2024) 也被提出：ESRT引入高效多头注意力(EMHA)，将Transformer内存需求从原始的16GB降至仅4GB^[19]；IFIN则通过结构感知块和Swin Transformer块的交互式特征模块整合CNN与Transformer优点，实现了极低算力下的高性能^[20]^[21]（代码：wwaannngllii/IFIN^[21]）。这些Transformer模型在PSNR/SSIM上保持竞争力的同时，更侧重全局上下文建模，近年来成为SR领域的研究热点。

- **SwinIR (Liang et al. 2021)** – 基于Swin Transformer的强基线，包含浅层特征提取、若干Residual Swin Transformer Blocks (RSTB) 和重构模块^[22]。在经典SR、轻量化SR和真实SR上均刷新SOTA：相对于之前最优方法，SwinIR在各类数据集上最多提升了0.14–0.45 dB PSNR^[13]，并显著减少了参数量（最高削减67%）。其官方实现已发布^[14]【41】。
- **Restormer (Zamir et al. CVPR2022)** – 提出Restoration Transformer架构，通过通道维度的MDTA注意力和门控卷积前馈（GDFN）模块实现高效的局部–全局信息融合。Restormer在图像去噪、去雨、反卷积等恢复任务上获得稳定的性能优势；论文称其在16个基准数据集上均实现了“显著性能提升”^[16]。例如，在多个合成去雨任务中，Restormer的PSNR普遍高于同期方法0.1–0.5 dB。
- **ESRT (Lu et al. 2022)** – 针对Transformer计算和显存瓶颈，设计了轻量级CNN骨干和专门的Efficient MHA。在将经典Transformer的显存占用从16GB缩减到约4GB的同时，ESRT仍获得了与原模型“可竞争”的重建效果^[19]（示例：ESRT占用GPU仅4191MB）。ESRT和相关高效注意力机制为资源受限环境的SR提供了解决方案^[19]。
- **HAT (Chen et al. 2023/24)** – “Hybrid Attention Transformer”结合通道注意力和窗口自注意力，并加入交叠跨窗口注意机制来增强不同窗口间的信息流动^[17]。通过额外的同任务预训练策略，HAT在标准和真实SR任务上均刷新了SOTA（量化指标和视觉效果均优异）^[17]。该工作验证了激活更多输入像素对SR性能的潜力；源码见XPixelGroup/HAT【48】。
- **IFIN (Wang et al. 2024)** – 结合CNN和Transformer优势，设计交互式特征模块（SAAB结构块、Swin Transformer块、增强空间自适应块），实现多层次特征提取和融合^[20]。实验表明IFIN在保持低模型复杂度的同时，在常见基准上达到或超越主流SR网络的重建质量^[20]。作者提供了开源代码^[21]。

扩散模型方法

扩散（Diffusion）模型近年成为SR生成的一种强大框架。**SR3** (Saharia et al. 2021) 将条件去噪扩散概率模型应用于图像SR：首先从高斯噪声开始迭代去噪，最终生成高分辨率图像^[23]。SR3在不同放大倍数的人脸与自然图像SR任务上展现优异性能：在CelebA-HQ 8×SR的用户研究中，其“欺骗率”接近50%，显著高于同期GAN方法的34%^[23]；该模型在ImageNet级链式生成中也取得了FID ≈ 11.3 的竞争性成绩。尽管多步扩散推理速度较慢，后续工作致力于加速和蒸馏。最新如**D3SR** (2024) 通过知识蒸馏和大型扩散判别器，将多步扩散蒸馏到单步模型，保持或超越了原方法的质量，同时显著加速推理（参数减少30%以上）^[24]。该类方法在定量（PSNR/SSIM等）和定性指标上与原多步扩散方法持平^[24]。总体而言，扩散模型为高质量SR提供了新的思路，但挑战在于效率优化和理论解释其结果稳定性。

- **SR3 (Saharia et al. 2021)** – 通过去噪扩散过程实现SR，生成过程从随机噪声出发依次去噪^[23]。该方法在不同倍率的人脸和自然图像上均表现出色：例如，在CelebA 8×SR任务中，SR3生成的图像“欺

骗”了近50%的人类评审（视为真人照片），远高于传统GAN方法的34%²³；这说明SR3输出效果极为真实自然。

- **单步扩散加速 (e.g. D3SR 2024)** – 为解决扩散模型的计算开销，一类方法利用知识蒸馏将多步模型简化为单步。Li等（2024）提出的大规模扩散判别器模型（D3SR）能在定量指标上与原多步模型相当甚至更优，同时推理速度快得多²⁴。该方法较先前需要数十甚至上百步的扩散，仅需一步即可获得相近的PSNR/视觉质量²⁴。此类研究表明扩散模型在保证效果的前提下，可通过蒸馏和新的损失设计实现效率优化。

自监督学习与无监督方法

自监督/无监督SR方法不依赖外部HR标签，而是利用图像自身或未配对数据学习。典型思路包括内部学习（如ZSSR自2018年提出，在给定LR图像上训练网络）、GAN-或循环一致性（CycleGAN）等。近年工作也在改进退化建模和结构先验。一项最新工作（Zhang等，2024）提出了“高分辨率感知”机制：在自监督SR框架中动态调整退化模型（由控制器根据当前SR质量调整建模参数），并引入特征对齐正则，直接约束生成图像特征与高分辨率分布一致²⁵。该方法无需HR标签即可针对真实LR域微调现成SR模型，其生成结果视觉效果自然且在感知质量上达到最新水平²⁵。总体来说，自监督SR目前多用于解决真实场景SR（未标注退化情况）问题，但仍面临提升细节清晰度和理论分析的挑战。

- **高分辨率感知自监督 (Zhang等, 2024)** – 针对真实场景缺乏HR标注的问题，该方法通过调节退化过程和高分辨率先验（特征对齐）来加强网络对高频细节的学习²⁵。实验证明，其在无需HR标签的前提下，对目标域图像进行微调后能生成具有“自然观感”的高质量SR图像²⁵。
- （可选）**内部示例学习、无监督循环等** – 如ZSSR、KernelGAN等经典方法利用单张图像内的自相似性和GAN估计退化核。但它们往往PSNR有限，适用于特定退化假设。目前更关注如何结合外部域信息和自监督策略来增强对真实退化的适应性，如LoSwap等领域适应方法。

多任务与多模态联合学习

多任务/多模态SR模型尝试同时利用额外信息或任务来提升SR性能。一类是多模态条件SR，例如最近的MMSR（2024）将自然语言、深度图、语义分割和边缘等多种指导信息联合用于SR：输入包括LR图像及其对应的图像说明（Caption）、深度图、分割掩码和边缘图等，通过条件扩散生成SR图像²⁶。实验表明，MMSR在视觉真实度和内容一致性上优于仅基于图像的SISR方法²⁶。另一类是多任务网络，如SEG-ESRGAN（2022）在遥感领域同时进行SR和语义分割：其双分支设计一边输出5倍超分图像，一边输出高分辨率地物分割图，从而充分利用了标签丰富的遥感影像信息²⁷。这种联合学习方式可以提高SR结果的可用性和鲁棒性，但也需要平衡不同任务目标。未来趋势可能包括将SR与其他视觉任务（如分类、检测或深度估计）联合优化，以及利用更多传感器数据（多光谱、红外、激光雷达等）来辅助SR。

- **MMSR (2024)** – 多模态SR模型，将图像Caption、深度图、语义分割图和边缘图作为条件输入，与LR图像一起引导SR生成。研究表明，丰富的上下文信息能够抑制不一致伪影并生成更逼真的HR图像，使MMSR在视觉质量上超越了传统SISR方法²⁶。
- **SEG-ESRGAN (2022)** – 多任务SR框架，同时输出超分图像（×5）和对应的高分辨率地物分割图。该网络在遥感图像（Sentinel-2）上联合利用分割任务信息，从而提升了SR重建效果和语义一致性²⁷。实验表明该多任务策略能够更有效地利用标注信息，获得更高质量的SR结果。

基于元学习/迁移学习的方法

迁移学习在SR中多用于将模型从合成退化域适应到真实域，或实现模型参数的快速更新。一种简单形式是预训练+微调：例如吴等（2022）在红外成像超分中采用元迁移学习策略，首先在大规模可用数据集上预训练网络，然后在目标内部数据上进行微调²⁸。这种做法增强了模型对新场景的适应性，提高了在目标域的表现²⁸。类似地，一些研究使用自适应层或可调整的网络结构，使模型可针对不同放大因子或不同退化类型快速更新。尽管“元学习”在SR领域的研究尚不多见，但其理念正在影响对少量标注或任务变化时模型快速收敛的探索。未

来趋势包括通过MAML等元学习算法实现跨任务的快速迁移，以及利用更大尺度预训练（例如扩散模型、大型生成模型）的知识迁移来提高SR泛化能力。

- **红外元迁移 (Wu等, 2022)** – 提出一种元迁移策略：先在外数据数据集上预训练SR模型，再在目标红外数据集上微调。该方法增强了模型的迁移能力，使其在特定场景下能够更好地恢复高分细节 ²⁸。
- **其他迁移学习方式** – 例如通过固定CNN骨干只微调最后几层，或使用小样本自适应（few-shot learning）快速适应新放大倍率。有研究尝试用元学习（如MAML）让模型学会在少量样本上优化降质参数。总体而言，迁移学习在现实世界SR中的挑战是如何利用少量目标域数据弥合合成与真实退化之间的差距，同时保证训练稳定性与效果提升。

主要挑战与发展趋势：当前SR研究虽然取得了PSNR/SSIM等指标的高度，但与真实世界低分辨率图像匹配仍困难。迁移学习和自监督方法应对这类挑战，将是未来重点；同时追求**更高感知质量**和**更快推理速度**的网络设计（如蒸馏的Transformer和Diffusion模型）仍在不断发展。多任务和多模态方法则为SR引入了新的信息来源（例如深度、语义、文本提示等），拓宽了技术视野。未来研究可能会更多地融合大规模生成模型的预训练能力，使单幅SR与图像生成等领域的进展交叉提升。在性能指标方面，除了传统的PSNR/SSIM外，感知指标（LPIPS、FID等）和用户主观评价也将成为衡量关键，促使算法在视觉效果上更贴近自然图像。

参考文献：以上各项方法及数据均来自近年相关论文和开源项目 ¹ ² ⁵ ⁹ ¹⁰ ¹² ¹³ ¹⁶ ¹⁹ ²⁵ ²⁶ ²⁷ ²⁸ 等。

¹ ² ⁴ A very lightweight image super-resolution network | Scientific Reports

https://www.nature.com/articles/s41598-024-64724-y?error=cookies_not_supported&code=236c9934-ad6a-4d49-9f9b-310b7c7d34b2

³ ⁷ Table 5 Quantitative comparison with state-of-the-art methods on benchmark datasets.

https://www.nature.com/articles/s41598-024-64724-y/tables/5?error=cookies_not_supported&code=8de6e6c3-2f90-41d8-ad24-9b24d29d3ea1

⁵ ⁶ ⁸ A lightweight large receptive field network LrfSR for image super-resolution | Scientific Reports

https://www.nature.com/articles/s41598-025-96796-9?error=cookies_not_supported&code=abddf4cb-fa9b-4382-9e96-15824cb871c8

⁹ ¹⁰ ¹¹ [2107.10833] Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data

<https://arxiv.org/abs/2107.10833>

¹² Designing a Practical Degradation Model for Deep Blind Image Super-Resolution

https://openaccess.thecvf.com/content/ICCV2021/papers/Zhang_Designing_a_Practical_Degradation_Model_for_Deep_Blind_Image_Super-Resolution_ICCV_2021_paper.pdf

¹³ ¹⁴ ²² [2108.10257] SwinIR: Image Restoration Using Swin Transformer

<https://arxiv.org/abs/2108.10257>

¹⁵ ¹⁶ Restormer: Efficient Transformer for High-Resolution Image Restoration

https://openaccess.thecvf.com/content/CVPR2022/papers/Zamir_Restormer_Efficient_Transformer_for_High-Resolution_Image_Restoration_CVPR_2022_paper.pdf

¹⁷ ¹⁸ [2309.05239] HAT: Hybrid Attention Transformer for Image Restoration

<https://arxiv.org/abs/2309.05239>

¹⁹ Transformer for Single Image Super-Resolution

https://openaccess.thecvf.com/content/CVPR2022W/NTIRE/papers/Lu_Transformer_for_Single_Image_Super-Resolution_CVPRW_2022_paper.pdf

20 21 **Lightweight interactive feature inference network for single-image super-resolution | Scientific Reports**

https://www.nature.com/articles/s41598-024-62633-8?error=cookies_not_supported&code=4874275e-a32c-4470-88f8-0fac738431f6

23 **[2104.07636] Image Super-Resolution via Iterative Refinement**

<https://arxiv.org/abs/2104.07636>

24 **Unleashing the Power of One-Step Diffusion based Image Super-Resolution via a Large-Scale Diffusion Discriminator**

<https://arxiv.org/html/2410.04224v3>

25 **High-Resolution Be Aware! Improving the Self-Supervised Real-World Super-Resolution**

<https://arxiv.org/html/2411.16175v1>

26 **The Power of Context: How Multimodality Improves Image Super-Resolution**

<https://arxiv.org/html/2503.14503v1>

27 **SEG-ESRGAN: A Multi-Task Network for Super-Resolution and Semantic Segmentation of Remote Sensing Images**

<https://www.mdpi.com/2072-4292/14/22/5862>

28 **Meta Transfer Learning-Based Super-Resolution Infrared Imaging | Request PDF**

https://www.researchgate.net/publication/358688635_Meta_Transfer_Learning-Based_Super-Resolution_Infrared_Imaging