



计算机科学

COMPUTER SCIENCE

基于稀疏编码非局部注意力对偶网络的病理图像超分辨率重建

梁美彦, 张宇, 梁建安, 陈庆辉, 王茹, 王琳

引用本文

梁美彦, 张宇, 梁建安, 陈庆辉, 王茹, 王琳. 基于稀疏编码非局部注意力对偶网络的病理图像超分辨率重建[J]. 计算机科学, 2023, 50(6A): 220700016-8.

LIANG Meiyuan, ZHANG Yu, LIANG Jianan, CHEN Qinghui, WANG Ru, WANG Lin. [Pathological Image Super-resolution Reconstruction Based on Sparse Coding Non-local AttentionDual Network](#) [J].

Computer Science, 2023, 50(6A): 220700016-8.

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)

[ATT&CK框架下基于事件序列关联的网络高级威胁检测系统](#)

Network Advanced Threat Detection System Based on Event Sequence Correlation Under ATT&CK Framework

计算机科学, 2023, 50(6A): 220600176-7. <https://doi.org/10.11896/jsjcx.220600176>

[结构化混合注意力网络的图像超分辨率重建](#)

Image Super-resolution Reconstruction Based on Structured Fusion Attention Network

计算机科学, 2023, 50(6A): 220600240-5. <https://doi.org/10.11896/jsjcx.220600240>

[基于深度多尺度卷积稀疏编码的图像去噪算法](#)

Image Denoising Algorithm Based on Deep Multi-scale Convolution Sparse Coding

计算机科学, 2023, 50(4): 133-140. <https://doi.org/10.11896/jsjcx.220100090>

[求解资源受限项目调度问题的分支定价算法](#)

Branch & Price Algorithm for Resource-constrained Project Scheduling Problem

计算机科学, 2022, 49(12): 274-282. <https://doi.org/10.11896/jsjcx.211100276>

[基于动态金字塔和子空间注意力的图像超分辨率重建网络](#)

Image Super-resolution Reconstruction Network Based on Dynamic Pyramid and Subspace Attention

计算机科学, 2022, 49(11A): 210900202-8. <https://doi.org/10.11896/jsjcx.210900202>

基于稀疏编码非局部注意力对偶网络的病理图像超分辨率重建

梁美彦¹ 张宇¹ 梁建安¹ 陈庆辉¹ 王茹¹ 王琳^{2,3}

1 山西大学物理电子工程学院 太原 030006

2 山西白求恩医院,山西医科大学第三医院 太原 030032

3 华中科技大学同济医学院附属同济医院 武汉 430030

摘要 高分辨率的病理学图像是疾病高精度诊断的客观依据,在精准医学领域具有重要意义。然而,受硬件设备分辨率和扫描时长的限制,实时获取高分辨率病理图像存在困难。经典的图像超分辨率重建算法由于模型的参数较难估计,导致重建后图像细节模糊且不够真实,不适用于病理学图像。为此,文中提出稀疏编码非局部注意力对偶网络,通过上采样和降采样对偶分支中的稀疏编码非局部注意力机制、高斯约束以及参数共享策略来实现病理学图像的超分辨率重建。重建后的病理图像峰值信噪比和结构相似性分别达到了 30.84 dB 和 0.914。研究表明,所提方法不但能够实现病理学图像中高频细节的精确重建,轻量化的稀疏编码非局部注意力机制也有效地提高了建模的效率,是病理学图像超分辨率重建的一种有效方法。

关键词: 稀疏编码;非局部注意力;对偶网络;病理图像;超分辨率

中图法分类号 TP391

Pathological Image Super-resolution Reconstruction Based on Sparse Coding Non-local Attention Dual Network

LIANG Meiyang¹, ZHANG Yu¹, LIANG Jianan¹, CHEN Qinghui¹, WANG Ru¹ and WANG Lin^{2,3}

1 College of Physics and Electronics Engineering, Shanxi University, Taiyuan 030006, China

2 Shanxi Bethune Hospital, Third Hospital of Shanxi Medical University, Taiyuan 030032, China

3 Tongji Hospital Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

Abstract High-resolution pathological images are the objective criteria for high-precision disease diagnosis, which have great significance in the field of precision medicine. However, it is difficult to obtain high-resolution pathological images in real time, due to the limited resolution and constrained scanning time of hardware devices. Classical image super-resolution reconstruction algorithm is not suitable for pathological images because the parameters of the model are difficult to estimate, resulting in blurred and unrealistic image details after super-resolution reconstruction. Therefore, sparse-coding non-local attention dual network (SNADN) is proposed, which uses Gaussian constraints, hash coding and parameter sharing strategy in the dual branches, to achieve high accuracy and high efficiency super-resolution reconstruction of pathological images. The PSNR and SSIM of the reconstructed pathological images can reach 30.84 dB and 0.914, respectively. The results show that SNADN can not only achieve accurate reconstruction of high-frequency details in pathological images, but also the lightweight sparse coding non-local attention mechanism can effectively improve the modeling efficiency. It is an effective method for super-resolution reconstruction of pathological images.

Keywords Sparse coding, Non-local attention, Dual network, Pathological image, Super resolution

1 引言

组织病理学检查是疾病诊断的“金标准”。病理专家通过高、低倍镜相结合的方式,来获取病理学图像的宏观结构及微观特征,再对这些特征进行综合分析而得出诊断结论。其中,高分辨率病理学图像能够提供病变部位的高频细节,是疾病

高精度诊断的客观依据。然而,受硬件设备分辨率和扫描时长的限制,导致实时获取高分辨率病理图像存在一定困难。若能找到一种方法对病理学图像进行高精度、高分辨率实时重建,对于病理学诊断具有重要意义。

图像超分辨率重构算法主要分为基于插值的方法^[1-3]、基于重构的方法和基于深度学习的方法。基于插值的超分辨率

基金项目:国家自然科学基金(11804209);山西省面上青年基金(201901D211173);山西省高校科技创新计划(2019L0064);山西省应用基础研究计划项目(201901D211172);山西省中央引导地方科技发展资金项目(YDZX20201400001547);山西省高等学校科技创新项目(2020L0048)

This work was supported by the National Natural Science Foundation of China(11804209), Shanxi Provincial Youth Fund(201901D211173), Shanxi Provincial University Science and Technology Innovation Program (2019L0064), Shanxi Province Applied Basic Research Program (201901D211172), Central Government of Shanxi Province Guides Local Science and Technology Development Fund Projects (YDZX20201400001547) and Science and Technology Innovation Project of Shanxi Provincial Colleges and Universities(2020L0048).

通信作者:梁美彦(meiyangliang@sxu.edu.cn)

算法又包括最近邻插值、双线性插值、双三次插值^[4]等。基于插值的算法虽然操作比较简单,但是不能增加图像实际的高频信息,从而导致生成的图像较为模糊。基于重构的超分辨率算法有迭代反投影法^[5]、凸集投影法^[6]、最大后验概率法^[7]等。这类算法主要通过模型对输入的一组高度相关的低分辨率图像进行参数估计,来实现高分辨率重建,其不仅计算复杂度高,而且无法自适应学习,当场景中无法获取足够的信息时,会导致重构的高分辨率图像细节模糊,或重建结果与真实图像结构不一致。

深度学习的方法通过自适应特征提取,不仅能够利用上下文语义信息对高分辨率图像进行参数估计,而且可以最大程度地保证重建结果与原始图像的结构相似性,因此逐渐被应用于超分辨率领域^[8-10]。目前,基于深度学习的图像超分辨率重建方法主要包括两大类:1)网络通过学习图像低分辨率空间到高分辨率空间的单向映射来实现图像超分辨率重建;2)网络通过学习图像低分辨率空间与高分辨率空间的双向映射,并利用降采样过程中损失的信息来重建高分辨率图像。基于学习单向映射的代表工作如下。2014年,Dong等^[11]首次提出超分辨率卷积神经网络 SRCNN,通过卷积操作获取图像的语义信息,实现了图像的超分辨率重建。以 SRCNN 为基础,FSRCNN^[12],VDSR^[13]和 ESPCN^[14]模型分别通过反卷积、减小卷积核尺寸和亚像素卷积操作来降低模型的计算复杂度。2017年,Christian首次将生成对抗网络^[15]用于图像超分辨率重建任务,提出了(Super-Resolution GAN,SRGAN)^[16]网络,通过在生成器中的卷积操作来获取图像语义信息,同时与判别器进行对抗训练,再分别引入感知损失和对抗损失,来引导生成器生成高分辨率的图像。后续研究都是以 VDSR 和 SRGAN 为基础模型,在模型提取低分辨率空间特征时加入跳层连接、注意力模块并设计轻量化的结构,再辅以感知损失对图像进行约束,这样不但能够提升图像的超分辨率重建效果,而且加快了模型的收敛速度^[17-38],其中代表性工作如下。2020年,Mei等引入跨尺度非局部注意力模块来融合不同尺度上提取到的特征信息,同时结合局部特征以及每个尺度下的非局部特征,提出了跨尺度非局部注意力网络(Cross-scale Non-local Attention Network, CNAN)^[39],利用低分辨率下不同尺度的特征对自然图像进行超分辨率重建,4倍超分辨率重建后 PSNR 和 SSIM 达到了 27.80dB 和 0.74。2021年,Liu等利用设计的轻量级多尺度模块(Lightweight Multi-scale Block, LMSB)^[40]提取低分辨率空间的多尺度特征和加速多个尺度的特征融合来获取更丰富的细节。2021年,Mei等提出了非局部稀疏注意力网络(Non-Local Sparse Attention Network, NLSAN)^[41],通过非局部稀疏注意力机制提取图像在低分辨率空间上特征图的关键信息,并利用这些信息对图像进行超分辨率重建。4倍超分辨率重建后图像的 PSNR 和 SSIM 分别为 27.78 dB 和 0.74,其性能不但优于使用跳层连接的网络,同时较低的参数量使建模时间缩短为原来的 1/3。

基于学习双向映射来实现超分辨率重建的代表性工作如下。2019年,Hu等将 U-net^[42]用于图像超分辨率重建并提出了改进的 RU-net^[43]模型,该模型通过学习自然图像高分辨率空间与低分辨率空间的双向映射,来获取降采样模型中

丢失的高频信息,并将该信息加入上采样模型中,用于图像的超分辨率重建。2020年,Xiao等提出了可逆神经网络(Invertible Image Rescaling, IRN)^[44],通过两个完全对称的上采样和降采样子模型,以及参数共享机制,不但减少了训练参数,而且提升了模型的重建效果。2020年,Guo等提出了对偶回归网络(Dual Regression Networks, DRN)^[45],通过设计降采样核重建低分辨率自然图像,并设计上采样核估算高分辨率图像,该模型利用形成的闭环结构来提供额外的监督。此外,利用参考图像也可以进行图像超分辨率重建^[46]。2020年,Yang等提出了文本迁移网络(Texture Transformer, TTSR)^[47],利用注意力机制提取参考图像与高分辨率图像中相似的高频信息,4倍超分辨率重建后图像的 PSNR 和 SSIM 分别为 25.87 dB 和 0.784。上述方法的特点的总结如表 1 所列。

表 1 超分辨率算法综述

Table 1 Summary of super-resolution algorithms

Method	Reference	Evaluation standard	Features
双三次插值算法	Davidon et al.		
最邻近插值算法	Cover et al.	视觉效果	图像平滑,产生振铃效应
迭代反投影法	Irani	视觉效果	算法复杂,不能自适应学习,丢失高频细节
凸集投影法	Patti et al.		
邻域嵌入法	Seung et al.		不能自适应学习
稀疏表达法	Yang et al.	视觉效果	可以自适应学习,计算耗时
SRCNN	Dong et al.	PSNR, SSIM, IFC, YCbCr 空间	
FSRCNN	Dong et al.		
Sub-pixel CNN	Shi et al.		
VDSR	Kim et al.		自适应学习,不能解决超分辨率重建解不唯一的问题,损失函数的选择决定重建图像的效果
LapSRN	Lai et al.		
SRDenseNet	Tong et al.	L ₁ 损失, 渐采样	
EDSR	Lim et al.		
WDSR	Yu et al.		
RDN, RCAN	Zhang et al.		
RFN	Liu et al.		
CSNL	Mei et al.		
SRGAN	Christian	生成对抗损失	改善图像的视觉效果,模型训练复杂
ESRGAN	Wanget al.		
RU-net	Hu et al.	PSNR, SSIM	可以解决超分辨率重建解不唯一的问题,但训练参数过多
IRN	Xiao et al.		
TTSR	Yang et al.		图像不真实
MCSC	Deng et al.		

目前,基于深度学习的图像超分辨率重建算法基本都是针对自然图像的。针对病理学图像高分辨率重建的研究未见报道。这是由于病理学图像的特征与自然图像不同,病理学图像的高频成份较多,而诊断依据也在很大程度上取决于高频分量。若病理学图像重建时存在误差或模糊,将严重影响病理医生的诊断结果。此外,图像的超分辨率重建算法使用不当将导致解不唯一。为此,本文提出基于稀疏编码的非局部注意力对偶网络来实现病理学图像的超分辨率重建,主要贡献有:

(1)通过降采样和上采样平行对偶分支学习图像低分辨率空间与高分辨率空间的双向映射,并利用高斯约束以及参数共享策略,来解决图像超分辨率重建时解不唯一的问题。

(2)为了在保证网络性能的同时减少模型参数,提出采用稀疏编码非局部注意力机制提供的轻量化模块来获取图像中

的关键信息,在有效高分辨重建的同时降低信息冗余。

2 模型

稀疏编码非局部注意力对偶网络框架如图 1 所示。该网络主要由两个平行对偶分支组成,即降采样分支(见图 1(a))和上采样分支(见图 1(b))。当网络输入为高分辨率病理学图像,输出为低分辨率病理学图像时,训练降采样分支,在

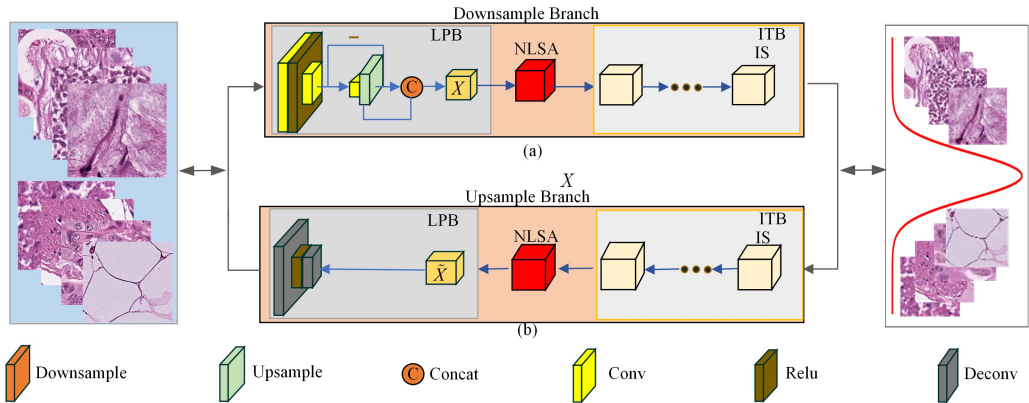


图 1 稀疏编码非局部注意力对偶网络

Fig. 1 Sparse coding non-local attention dual network

在稀疏编码非局部注意力对偶网络中,降采样分支和上采样互为对偶变换。在这两个平行分支中,主要由拉普拉斯金字塔模块(Laplacian Pyramid Block, LPB)、稀疏编码非局部注意力模块(NonLocal Sparse-coding Attention, NLSA)以及可逆变换模块(Invertible Transform Block, ITB)组成。拉普拉斯金字塔模块在降采样分支中,将输入图像通过卷积降采样、插值上采样,再用原始特征图与上采样的特征图相减分离出图像的高频信息,并与原来的低频特征图沿通道维拼接组成新的特征图 X ,因此特征图 X 前 64 个通道是低频特征,后 64 个通道是高频特征。拉普拉斯金字塔模块在上采样分支中的作用是将融合了低频特征和高频特征的特征图进行 2 倍上采样,使特征图恢复到原始大小;NLSA 模块通过哈希编码提取特征图中的关键信息,并最大程度避免信息冗余;两个对偶分支中的可逆变换模块都包含 8 个可逆子块(Invertible Subblock, IS),可逆变换模块通过可逆子块中的非线性变换的堆叠,进一步将特征图的高频信息和低频信息沿通道维进行分离和融合,即学习高分辨率图像空间与低分辨率图像空间双向映射。两个分支之间通过参数共享策略,可以将降采样分支中得到的符合高斯分布的高频信息嵌入上采样分支中,实现病理学图像的高分辨率精确重建。图 1 为 2 倍超分辨率对偶网络的示意图,4 倍超分辨率网络可以通过串联 2 个稀疏编码非局部注意力对偶网络组成。

2.1 稀疏编码非局部注意力模块

2.1.1 非局部注意力

非局部注意力可以提取特征图中的重要信息,通过特征图中待求位置与任一位置的相关性提取关键性特征。设输入特征图 X ,按通道划分转换为 $X = \{x_1, \dots, x_i, x_j, x_k, \dots\}$, $\forall x \in R^c$ 且 $X \in R^{n \times c}$, $f(\cdot, \cdot)$ 是特征间的相似性表示, $g(\cdot)$ 是特征转换函数,对于划分后的任意一个特征点,其注意力机制后特征点 y 表示为:

$$y = \sum_{j=1}^n \frac{f(x_i, x_j)}{\sum_{k=1}^n f(x_i, x_k)} g(x_j) \quad (1)$$

这个过程中模型学习高分辨率图像到低分辨率图像的特征变换映射,同时加入高斯约束,使变换后的高频信息符合高斯分布。当模型输入为低分辨率病理学图像,输出为高分辨率病理学图像时,训练上采样分支,即学习低分辨率图像到高分辨率图像的特征变换映射,同时通过参数共享策略将降采样分支中符合高斯约束的高频特征嵌入上采样分支,从而生成高分辨率病理学图像。

由上述分析可知,非局部注意力是通过计算图像空间全局范围内任意两点之间的相似性来得到每个特征点的注意力再进行分配,这将严重增加算力的消耗。

2.1.2 稀疏编码非局部注意力模块(NLSA)

局部敏感哈希编码是求解最近邻问题的一种随机化算法,在高维空间上运算速度更快。其基本思想是从局部敏感哈希函数族中随机选择哈希函数,将每个数据点映射到哈希表中。在查找邻居点时,只查找与查询点具有相近编码值的点。对于每个待查询的特征点,本文提出的局部敏感哈希编码方式在全局范围内快速查找与当前查询点相关性较大的特征点。

稀疏编码非局部注意力模块结构如图 2 所示。该模块由两个分支构成,分支 1 用于获取特征图 X 的哈希桶。具体过程为:

(1)对特征图 X 沿通道维度进行归一化。特征图上任意点 x 归一化后用 \hat{x} 表示,其计算式为:

$$\hat{x} = \frac{x}{\sum_{x \in X} \|x\|_2} \quad (2)$$

(2)对每个特征点进行哈希编码。过程如图 2(a)所示,特征点 \hat{x} 的编码值 F 的计算式如式(3)所示,其他特征点做同样的计算。

$$F = \max(A(\hat{x})) = \max(\{h_1(\hat{x}), h_2(\hat{x}), \dots, h_i(\hat{x}), \dots, h_n(\hat{x})\}) \quad (3)$$

其中, A 表示稀疏函数簇, $h_i(x)$ 表示第 i 个哈希函数,稀疏函数簇 A 由 n 个哈希函数 $h_i(x)$ 组成。

(3)根据每个特征点的编码值构建哈希桶,如图 2(b)所示。图 2(b)中,每个特征点通过哈希编码后,按照哈希值从大到小排序和分类,最终获得(I),(II),(III),(IV),(V),..., m 个哈希桶,每个桶内有 k 个最相关的特征点。哈希桶表示方式如下:

$$C_j = [\hat{x}^{j,1}, \hat{x}^{j,2}, \dots, \hat{x}^{j,k}], j \in 1, \dots, m \quad (4)$$

其中, $\hat{x}^{j,k}$ 表示编码到第 j 个桶的第 k 个特征点。每个桶内

特征点的哈希编码值相近。

(4) 分支 2 获取每个特征点经注意力机制后的结果。以特征点 \hat{x} 为例, 过程如图 2(c) 所示, 若 $A(\hat{x})$ 位于第 l 个桶, 其中 $l \in [1, m]$, 只需求得 \hat{x} 与该桶内 k 个特征点的相关性, 得到

注意力分数即可, 如式 (5)、式 (6) 所示:

$$\delta_l = \text{Index}(C_l), l \in [1, m] \quad (5)$$

$$y(\hat{x}_i, \delta_i) = \sum_{j \in \delta_i} \frac{f(\hat{x}_i, \hat{x}_j)}{\sum_{k \in \delta_i} f(\hat{x}_i, \hat{x}_k)} g(\hat{x}_j) \quad (6)$$

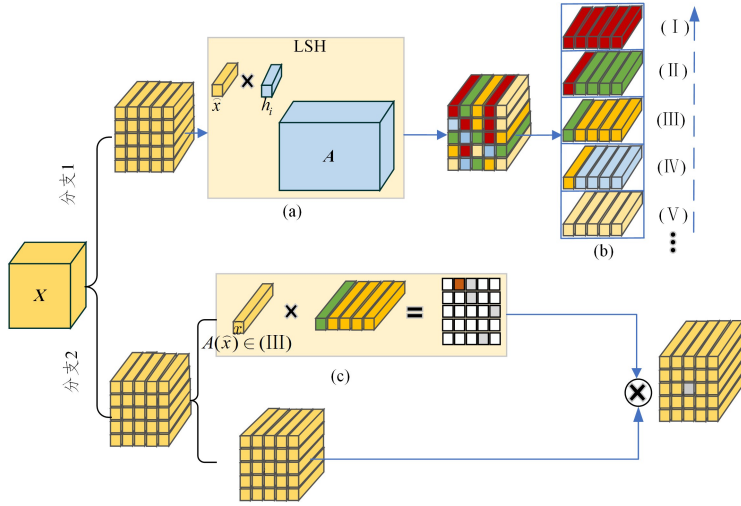


图 2 稀疏编码非局部注意力模块
Fig. 2 Non local Sparse-coding attention block

2.2 可逆变换模块

2.2.1 可逆理论

文献[48]提出了非线性独立分量估计网络 (Non-linear Independent Component Estimation, NICE), 用于建模复杂高维数据。然而, 该理论的前提是, 选择合适的数据特征表示使数据更易于建模。为此, 将源数据 x 通过非线性变换 $z = f(x)$ 映射到潜在空间, 并使变换后的数据 $p_z(z)$ 符合可分解分布, 即产生 d 维独立潜在变量 $p_{z_d}(z_d)$ 来表示源空间的特征, 最后, 可以将该潜在变量 z 通过 $\tilde{x} = f^{-1}(z)$ 来恢复源数据分布, 表达式如下:

$$p_z(z) = \prod_d p_{z_d}(z_d) \quad (7)$$

$$p(\tilde{x}) = \prod_d p(z_d) p(\tilde{x} | z_d) \quad (8)$$

其中, $p_z(z)$ 为潜在变量 z 的概率密度分布函数, d 为源数据空间的维度, $p_{z_d}(z_d)$ 表示潜在变量每个维度的概率密度分布函数, $p(\tilde{x})$ 表示源空间数据的概率密度函数, $p(\tilde{x} | z_d)$ 表示在特定维度 z_d 下, f^{-1} 需要拟合的条件概率密度分布函数。

在降采样分支中, 本文采用上述方法将高分辨率病理图像的高频和低频特征分别通过 f 变换映射成潜在分布, 即低分辨率图像和符合高斯分布的高频信息, 新的低频信息和高频信息之间彼此独立且高频各个通道维度之间也彼此独立。即在降采样分支中利用仿射变换 f 将高分辨率图像中的高频特征图 X_h 通过 $X_{h'} = f(X_h)$ 映射成高斯分布, 即潜在高频特征图 $X_{h'}$, 图像的低频特征图 X_l 通过 $X_{l'} = f(X_l)$ 映射成潜在低频特征图 $X_{l'}$ 。其中, $X_{h'}$ 与 $X_{l'}$ 是彼此独立的高频和低频特征, 如图 3(a) 所示。在上采样分支中, 重建高频信息时通过逆变换 $\tilde{X}_h = f^{-1}(X_{h'})$ 进行恢复, 同理, 低频信息通过逆变换 $\tilde{X}_l = f^{-1}(X_{l'})$ 进行恢复, 如图 3(b) 所示。其中, f 和 f^{-1} 由 8 个图 3(a) 和 3(b) 对应虚线框中的可逆子模块构成, 每个可逆子模块又包含 ϕ, ρ, η 这 3 种变换的组合。每个可逆子模块由可逆操作通过多尺度变换模块构成。X' 表示低分辨率图像和符合高斯分布的高频特征按通道拼接后的特征。以

4 倍降采样过程为例, 图中标注了特征图的大小。

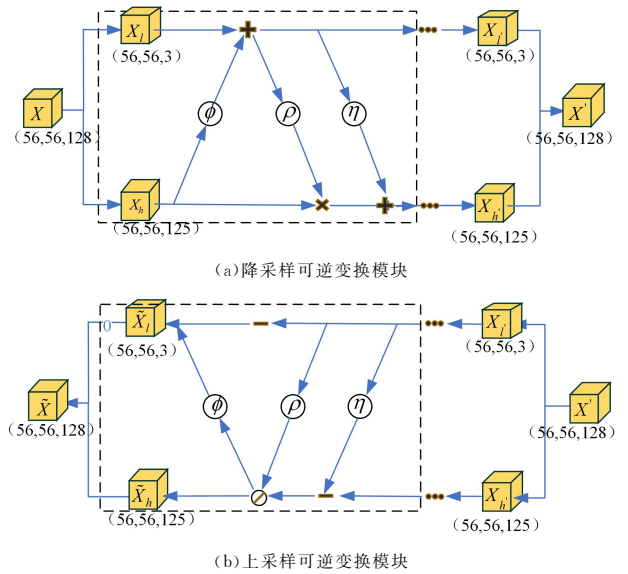


图 3 可逆变换模块

Fig. 3 Invertible transform block

2.2.2 可逆理论推导

概率分布函数 $F(x)$ 的定义为:

$$F(x) = P(\dot{X} \leq x)$$

由 $\tilde{x} = f^{-1}(z)$ 可得:

$$F(\tilde{x}) = P(f^{-1}(z) \leq x)$$

对 $P(\cdot)$ 括号内两边同时取 f 变换得 $F(x)$ 在潜在空间中的表示:

$$F(\tilde{x}) = P(z \leq f(x)) = F_z(f(x))$$

两端对 x 求导可得 $F(\tilde{x})$ 的概率密度函数 $p(\tilde{x})$:

$$p(\tilde{x}) = p_z(f(x)) \left| \det\left(\frac{\partial f(x)}{\partial x}\right) \right| \quad (9)$$

其中, $F(\cdot)$ 表示变量的分布函数, $P(\dot{X} \leq x)$ 表示变量 X 小于等于某值 x 的概率。 $p_z(f(x))$ 是 $F_z(f(x))$ 的导数,

$\det(\cdot)$ 表示行列式。

由式(9)可知,高分辨率图像高频信息和低频信息的表达式如下:

$$p(\tilde{X}_h) = p_Z(f(X_h)) \left| \det \frac{\partial f(X_h)}{\partial X_h} \right| \quad (10)$$

$$p(\tilde{X}_l) = p_Z(f(X_l)) \left| \det \frac{\partial f(X_l)}{\partial X_l} \right| \quad (11)$$

因此, f 必须满足以下两个条件,才能恢复高频和低频信息:

- (1) f 是可逆的;
- (2) f 的雅可比行列式容易计算。

为此,我们采用图3中堆叠的可逆结构^[49]来构建最优的 f 。

2.2.3 可逆操作

降采样分支中,可逆子模块低频特征和高频特征的计算式如下:

$$X_i^{+1} = X_i + \phi(X_i) \quad (12)$$

$$X_h^{+1} = X_h \odot \rho(X_i^{+1}) + \eta(X_i^{+1}) \quad (13)$$

同理,其对应上采样分支中可逆操作对应的计算方法如下:

$$X_i = (X_i^{+1} - \eta(X_i^{+1})) \oslash \rho(X_i^{+1}) \quad (14)$$

$$X_h = X_i^{+1} - \phi(X_i) \quad (15)$$

其中, X_i 和 X_h 表示第 i 个可逆子模块输入的低频特征和低频特征, ϕ 表示高频特征到低频特征的加性变换操作, η 表示低频特征向高频特征的加性变换操作, ρ 表示低频特征向高频特征的乘性变换操作, X_i^{+1} 和 X_h^{+1} 分别表示第 i 个可逆子模块输出的低频特征和低频特征。

2.2.4 多尺度密集模块

多尺度密集模块由两个分支组成,如图4所示。模块的两个分支分别采用膨胀率为1和2的卷积核,来获取不同感受野下图像的语义信息,卷积核大小都为 3×3 。为了避免梯度消失,在每两个层之间加入了跳跃连接,最后将不同尺度提取到的特征进行融合生成新的特征图。多尺度密集模块能够在可逆变换模块中利用多尺度信息,实现低频特征和低频特征的相互转换。图4为高频特征 X_h 向低频特征 X_i 变换操作的示意图,即 ϕ 所表示的变换,该模块分别作用于 ϕ, η, ρ 这3种变换中。

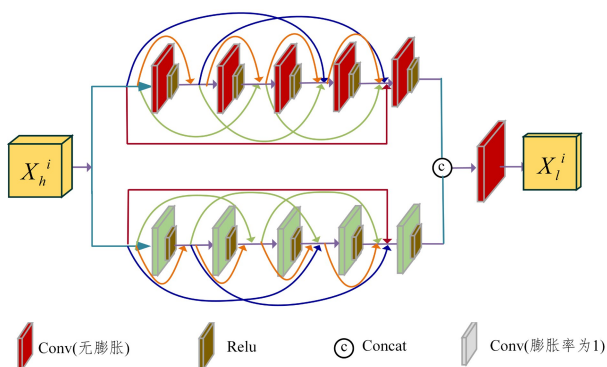


图4 多尺度模块
Fig. 4 Multi-scale module

2.3 算法

算法1给出了稀疏编码非局部注意力对偶网络的伪代码。在训练降采样分支时,真实的高分辨率病理学图像作为输入,低分辨率病理学图像和符合高斯分布的高频特征作为输出,建模过程中采用 L_2 损失函数,并对生成的高频信息进行高斯约束。在训练上采样分支时,真实的低分辨率病理学

图像和高斯约束的高频信息作为上采样网络输入,生成的高分辨率病理学图像作为上采样分支的输出。建模过程采用 L_1 损失函数。

算法1 稀疏编码非局部注意力对偶网络

1. 训练降采样分支
2. 输入:病理学图像训练集中高分辨率图像 $\{I_i^H\}$
3. 输出:符合高斯分布的高频信息 $X_{h'}$,低分辨率图像 $\{I_i^L\}$
4. 训练上采样分支
5. 输入:低分辨率图像 $\{I_i^L\}$,符合高斯分布的高频信息 $X_{h'}$
6. 输出:高分辨率图像 $\{\hat{I}_i^H\}$
7. Procedure:
8. For all i do:
9. $X_{l'}, X_{h'} \leftarrow$ Downsampling branch(I_i^H)
10. $\nabla_D L_2(X_{l'}, I_i^L) + \text{Gaussian Constraint}(X_{h'}) / *$ 更新降采样分支, L_2 代表均方差
11. $\hat{I}_i^H \leftarrow$ Upsampling branch($I_i^L, X_{h'}$)
12. $\nabla_u L_1(\hat{I}_i^H, I_i^H) / *$ 更新上采样分支, L_1 代表平均绝对误差,
13. End for
14. 测试 $\hat{I}_i^H \leftarrow$ Upsampling branch($I_i^L, X_{h'} \sim N(0, 1)$) / * 测试集低分辨率图像 I_i^L , 高分辨率图像 I_i^H

3 实验和结果

3.1 数据集

组织病理学图像是由德国海德堡国家癌症中心和海德堡大学医学中心获得的8种苏木精和伊红染色的人类结肠组织。数据集包含20000个互不交叠的结肠组织图像块,低分辨率病理学图像的大小为 56×56 ,高分辨率图像大小为 224×224 。数据集按照8:1:1的比例,划分为训练集、验证集和测试集。因此,训练集、验证集和测试集的图像数目分别为16000,2000和2000。训练过程分别通过 L_1 损失函数和 L_2 损失函数对上采样支路生成的高分辨率图像和降采样支路生成的低分辨率图像进行优化,本模型采用Adam优化器,初始学习率为 2×10^{-4} ,一阶矩衰减系数 $\beta_1 = 0.9$,二阶矩衰减系数 $\beta_2 = 0.999$ 。本模型采用 RTX 3090 来进行训练,批量大小为4,训练迭代100轮,建模时间约86h。

3.2 结果与分析

稀疏编码非局部注意力对偶网络对病理学图像超分辨率重建的结果如图5所示。

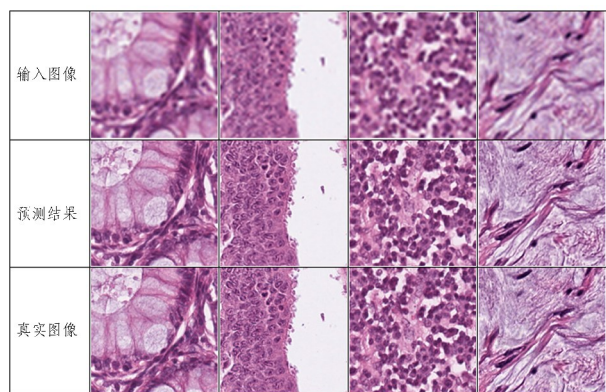


图5 稀疏编码非局部注意力对偶网络对病理学图像超分辨率重建的结果

Fig. 5 Super-resolution reconstruction of pathological images using sparse coding non-local attention dual network

结果表明,病理学图像在 4 倍超分辨率重建后,PSNR 和 SSIM 分别达到了 30.84 dB 和 0.914,重建后的图像在纹理、颜色和风格等方面与原始图像保持了较高的一致性,生成的高分辨率图像通过不同病理学家进行同行评审,就该方法的有效性达成了共识。此外,该模型通过稀疏编码非局部注意力模块,在不降低网络性能的前提下,实现了轻量化的设计,整个网络模型的参数量仅为 3.2×10^6 。

此外,本文通过对比稀疏编码非局部注意力与非局部注意力机制所关注的范围和重要程度,来说明稀疏编码非局部注意力对偶模型的有效性。图 6(a)和图 6(b)为非局部注意力与稀疏编码非局部注意力关注范围的对比。图 6 中的点表示与当前查询点的相关性程度,颜色越深表示相关性越大。对于每个待查询特征点,非局部注意力机制通过计算特征图上全局范围内每个像素点与查询特征点的相关性(注意力分数),来获取当前查询点的特征,用于图像超分辨率重建;而稀疏编码非局部注意力则通过哈希编码,仅需要计算与当前查询点相关性最大的特征像素点的注意力分数,来获取查询点的关键性特征,用于图像的超分辨率重建。图 6(b)中仅需要计算图 6(a)中颜色最深的特征,即关键性特征,就可以在不降低模型性能的前提下,大幅度降低模型的参数量,与可逆神经网络(IRN)相比,参数量降低了 25.6%。

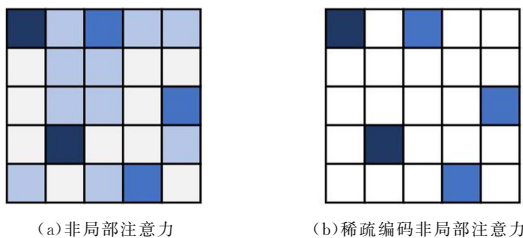


图 6 不同注意力机制关注的特征点示意图

Fig. 6 Schematic diagram of feature pixels in two attention mechanisms

为了进一步验证稀疏编码非局部注意力模块对网络性能的影响,对该模块进行了消融实验,5 次实验的平均结果如

表 2 所列。从结果可以看出,对稀疏编码非局部注意力模块进行消融后的模型,其 PSNR 和 SSIM 都存在一定的下降。这也证明了稀疏编码非局部注意力模块在提取特征图中关键信息的有效性。

表 2 消融实验

Table 2 Ablation study

	消融 NLSA 模块	Ours
PSNR/dB	30.56	30.84
SSIM	0.908	0.914

3.3 讨论

为了验证稀疏编码非局部注意力对偶网络超分辨率的效果,分别将其与双三次插值、SRGAN、CSNL、TTSR、IRN 模型的结果进行了对比。图 7 给出了上述方法对病理学图像进行 4 倍超分辨率重建的结果。由图 7 可知,双三次插值重建的图像细节模糊而且色彩失真,这是由于插值算法不能增加图像的高频信息,而病理图像色彩成分较多也导致了插值结果存在色彩失真。SRGAN 和 CSNL 虽然一定程度上提升了图像的分辨率,但无法重建图像的纹理信息,这是因为基于学习单向映射的超分辨率方法仅能学习到病理学图像低分辨率空间到高分辨率空间的统计规律,并通过建模基本实现两个空间的映射,但对于高频细节较多的病理学图像,这类方法无法准确得到与真实图像一致性较高的高分辨率图像。而 TTSR 在建模过程中虽然有可供参考的高频图像,但是生成的图像在细节方面也存在模糊的问题,这是由于 TTSR 模型是通过参考其他高分辨率图像信息来重建当前图像的,参考图像的与待恢复图像的高频信息分布的差异,也使图像在高分辨率重建过程中产生模糊。IRN 和稀疏编码非局部注意力对偶网络利用上采样和降采样两个分支来学习低分辨率空间与高分辨率空间之间的双向映射,并通过双向学习利用降采样产生的高频信息对低分辨率图像进行重建,该方法不但可以还原图像的细节和色彩等信息,而且与真实高分辨率图像保持了较高的一致性,其 4 倍超分辨率结果的 PSNR 和 SSIM 都在 30 dB 和 0.91 以上。

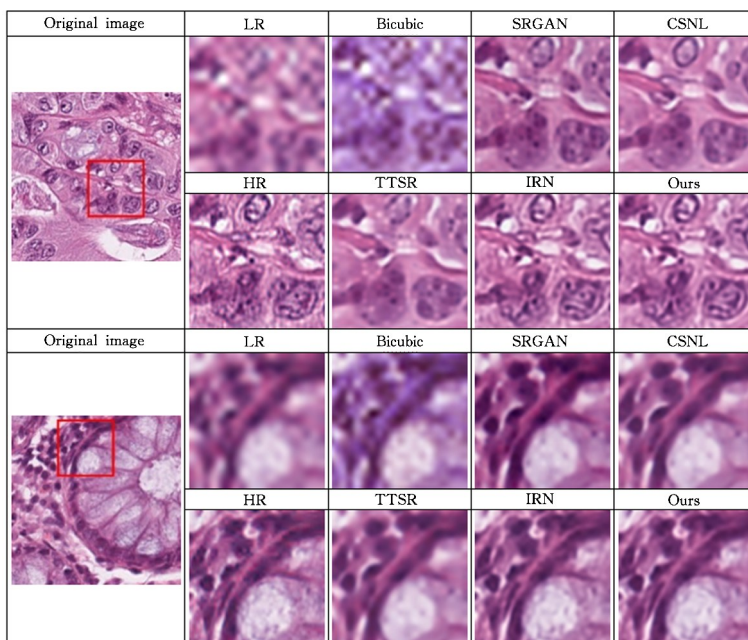


图 7 不同的超分辨率算法对病理组织图像重建 4 倍的结果

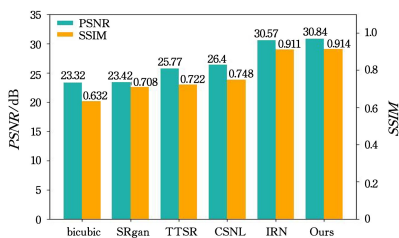
Fig. 7 4X reconstruction of pathological images using different super-resolution algorithms

表 3 列出了稀疏编码非局部注意力对偶网络与上述超分辨率算法在 4 倍超分辨率和 2 倍超分辨率下图像重建的评价结果。由表 3 和图 8 可知,在 4 倍和 2 倍的超分辨率图像重建时,稀疏编码非局部注意力对偶网络重建效果优于双三次插值、SRGAN、CSNL 和 TTSR 算法,这说明基于双向映射的超分辨率方法可以利用对偶分支中降采样过程产生的高频信息对图像进行高分辨率重建,与单向映射的方法相比,重建结果与真实图像具有较高的一致性。稀疏编码非局部注意力对偶网络与 IRN 算法相比,4 倍超分辨率图像的 PSNR 和 SSIM 都得到了一定的提高,同时参数减少了 1.1×10^6 ,这表明 SNADN 通过哈希编码可以在模型参数较少的情况下提取到关键特征,用于超分辨率重建。在 2 倍超分辨率实验中,虽然 PSNR 和 SSIM 指标比 IRN 略低,这是由于重建倍率较低,输入图片的本身已经包含较为明显的特征信息,但 IRN 模型也可以相对容易地提取到图像的特征信息并对病理学图像进行高分辨率重建。

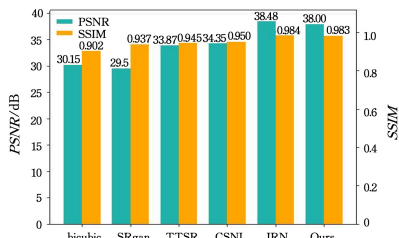
表 3 不同超分辨率算法的定量比较

Table 3 Quantitative comparison of different super-resolution algorithms

Method	Scale	Parameters	PSNR/dB	SSIM
Bicubic	$\times 4$		23.32	0.632
SRgan	$\times 4$	1.6×10^6	23.42	0.708
CSNL	$\times 4$	1.6×10^6	26.40	0.748
TTSR	$\times 4$	6.7×10^6	25.77	0.722
IRN	$\times 4$	4.3×10^6	30.57	0.911
OURS	$\times 4$	3.2×10^6	30.84	0.914
Bicubic	$\times 2$		30.15	0.902
SRgan	$\times 2$	1.6×10^6	29.50	0.937
CSNL	$\times 2$	0.7×10^6	34.35	0.950
TTSR	$\times 2$	4.4×10^6	33.87	0.945
IRN	$\times 2$	1.6×10^6	38.48	0.984
OURS	$\times 2$	1.0×10^6	38.00	0.983



(a) 4 倍超分辨率重建



(b) 2 倍超分辨率重建

图 8 不同超分辨率算法 PSNR 和 SSIM 的结果图

Fig. 8 PSNR and SSIM of different super-resolution algorithms

结束语 本文提出了基于稀疏编码非局部注意力的对偶模型,该模型通过降采样和上采样两个对偶分支,并结合高斯约束以及参数共享策略,来学习病理学图像低分辨率空间与高分辨率空间的双向映射,从而实现了组织病理学图像的超分辨率重建,重建结果的 PSNR 和 SSIM 分别为 30.84 dB 和 0.914。结果表明,基于稀疏编码非局部注意力的对偶模型不但解决了病理学图像超分辨率重建时解不唯一的问题,轻量化

的稀疏编码非局部注意力模块也在降低计算量的同时可以获取输入图像的关键信息,这是一种有效的超分辨率重建方法。

参 考 文 献

- [1] TSAI R, HUANG T. Multiframe image restoration and registration[J]. *Advances in Computer Vision and Image Processing*, 1984, 1(2): 317-339.
- [2] YUE L, SHEN H, LI J, et al. Image super-resolution: The techniques, applications, and future [J]. *Signal Processing*, 2016, 128: 389-408.
- [3] TSAI R, HUANG T. Moving image restoration and registration [C]// *IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1980: 418-421.
- [4] LIU L, LUO T. Application of Interpolation in Image Processing [J]. *Silicon Valley*, 2009, 9: 9-10.
- [5] IRANI M, PELEG S. Improving resolution by image registration [J]. *CVGIP: Graphical Models and Image Processing*, 1991, 53(3): 231-239.
- [6] PATTI A J, SEZAN M, TEKALP A. High resolution standards conversion of low resolution video [C]// *1995 International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1995, 4: 2197-2200.
- [7] LIU C, SUN D. On Bayesian adaptive video super resolution [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 36(2): 346-360.
- [8] SEUNG H, LEE D. The manifold ways of perception [J]. *Science*, 2000, 290(5500): 2268-2269.
- [9] YANG J, WRIGHT J, HUANG T, et al. Image super-resolution via sparse representation [J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
- [10] HINTON G, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [11] DONG C, LOY C, HE K, et al. Learning a deep convolutional network for image super-resolution [C]// *European Conference on Computer Vision*. Cham: Springer, 2014: 184-199.
- [12] KIM J, LEE J, LEE K. Accurate image super-resolution using very deep convolutional networks [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 1646-1654.
- [13] DONG C, LOY 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]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 1874-1883.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [16] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 4681-4690.
- [17] KIM J, LEE J, LEE K. Deeply-recursive convolutional network for image super-resolution [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 1637-1645.

- [18] IANDOLA, MOSKEWICZ M, KARAYEV S, et al. Densenet: Implementing efficient convnet descriptor pyramids[J]. arXiv: 1404.1869, 2014.
- [19] TONG T, LI G, LIU X, et al. Image super-resolution using dense skip connections[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017:4799-4807.
- [20] ZHANG K, SUN M, HAN T, et al. Residual networks of residual networks; Multilevel residual networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 28(6):1303-1314.
- [21] TAI Y, YANG J, LIU X, et al. Memnet: A persistent memory network for image restoration[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017:4539-4547.
- [22] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017:136-144.
- [23] YU J, FAN Y, YANG J, et al. Wide activation for efficient and accurate image super-resolution[J]. arXiv:1808.08718, 2018.
- [24] 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.
- [25] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]// Proceedings of the European Conference on Computer vision (ECCV). 2018:286-301.
- [26] LAI W, HUANG J, AHUJA N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:624-632.
- [27] CHEN R, ZHANG H, LIU J. Multi-attention augmented network for single image super-resolution[J]. Pattern Recognition, 2022, 122:108349.
- [28] WANG X T, YU K, DONG C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:606-615.
- [29] LIU J, ZHANG W, TANG Y, et al. Residual feature aggregation network for image super-resolution [C] // Proceedings of the IEEE/CVFC conference on Computer Vision and Pattern Recognition. 2020:2359-2368.
- [30] LIU Y, ZHANG S, XU J, et al. An accurate and lightweight method for human body image super-resolution [J]. IEEE Transactions on Image Processing, 2021, 30:2888-2897.
- [31] AN Z, ZHANG J, SHENG Z, et al. RBDN: Residual Bottleneck Dense Network for Image Super-Resolution[J]. IEEE Access, 2021, 9:103440-103451.
- [32] ANDONI A, INDYK P, LAARHOVEN T, et al. Practical and optimal LSH for angular distance[J]. Advances in Neural Information Processing Systems, 2015, 28:1225-1233.
- [33] TERASAWA K, TANAKA Y. Spherical LSH for approximate nearest neighbor search on unit hypersphere[C]// Workshop on Algorithms and Data Structures. Springer, Berlin, Heidelberg, 2007:27-38.
- [34] TIAN C, ZHANG X, LIN J, et al. Generative Adversarial Networks for Image Super-Resolution; A Survey[J]. arXiv:2204.13620, 2022.
- [35] WANG X, YU K, WU S, et al. Esrgan; Enhanced super-resolution generative adversarial networks[C]// Proceedings of the European Conference on Computer Vision (ECCV) workshops. 2018:8-14.
- [36] JIANG D, ZHANG S, DAI L, et al. Multi-scale generative adversarial network for image super-resolution[J]. Soft Computing, 2022, 26(8):3631-3641.
- [37] PARK S, MOON Y, CHO N. Flexible Style Image Super-Resolution using Conditional Objective[J]. IEEE Access, 2022, 10:9774-9792.
- [38] SONG J, YI H, XU W, et al. Dual Perceptual Loss for Single Image Super-Resolution Using ESRGAN [J]. arXiv: 2201.06383, 2022.
- [39] MEI Y, FAN Y, ZHOU Y, et al. Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:5690-5699.
- [40] PANG S, CHEN Z, YIN F. Lightweight multi-scale aggregated residual attention networks for image super-resolution[J]. Multimedia Tools and Applications, 2022, 81(4):4797-4819.
- [41] MEI Y, FAN Y, ZHOU Y. Image super-resolution with non-local sparse attention[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:3517-3526.
- [42] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Cham; Springer, 2015:234-241.
- [43] HU X, NAIEL M, WONG A, et al. RUNet: A robust UNet architecture for image super-resolution[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2019:505-507.
- [44] XIAO M, ZHENG S, LIU C, et al. Invertible image rescaling [C]// European Conference on Computer Vision. Cham; Springer, 2020:126-144.
- [45] GUO Y, CHEN J, WANG J, et al. Closed-loop matters: Dual regression networks for single image super-resolution[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:5407-5416.
- [46] DENG X, DRAGOTTI P. Deep convolutional neural network for multi-modal image restoration and fusion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(10):3333-3348.
- [47] YANG F, YANG H, FU J, et al. Learning texture transformer network for image super-resolution [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:5791-5800.
- [48] DINH L, KRUEGER D, BENGIO Y. NICE: Non-linear independent components estimation[C]// Workshop of the International Conference on Learning Representations. 2015.
- [49] DINH L, SOHI-DICKSTEIN J, BENGIO S. Density estimation using real nvp[J]. arXiv:1605.08803, 2016.



LIANG Meiyang, born in 1984, Ph.D., associate professor. Her main research interests include machine learning, deep learning and medical image processing.