

高效单图像超分辨率重建:深监督对称蒸馏网络

毛盼娣¹,徐道连²

(1.重庆城市科技学院 电气工程与智能制造学院,重庆 402167;

2.重庆大学 光电工程学院,重庆 400044)

摘要:过去几十年,卷积神经网络(Convolutional Neural Networks,CNNs)在单图像超分辨率(Single Image Super-Resolution,SISR)方面取得了明显的进展.现在大部分基于 CNNs 的方法都致力于构造新的架构去提升重建性能,这通常依赖大量计算和存储成本,难以应用于移动设备.提出了一种新颖的基于深监督对称蒸馏网络的高效单图像超分辨率重建方法(Deeply-Supervised Symmetry Distillation Network,DSSD),通过构造高频特征递归模块(High-frequency Feature Recursive Module,HFRM)和对称退化模块(Symmetry Degradation Module,SDM)缓解教师网络中提取高分辨率(High-Resolution,HR)高频信息不够准确这一问题.为了约束教师网络中提取的高频特征,采用深监督方法使教师网络蒸馏的知识与学生网络互补.在 DIV2K 数据集上的实验表明,DSSD 有效增强了单图像超分辨率(SISR)的性能,HFRM 和 SDM 的引入能够有效帮助 DSSD 提取更多图像高频细节.

关键词:深监督对称蒸馏网络;超分辨率;教师网络;高频特征递归模块;对称退化模块;特权信息

中图分类号:TP391.41

文献标志码:A

单图像超分辨率(SISR)旨在利用低分辨率(Low-Resolution,LR)图像重建出对应的高分辨率(HR)图像,近年来该方法在物体检测^[1]、信息取证^[2]等领域应用十分广泛.DONG 等^[3]提出 SRCNN 首次为 SISR 引入了卷积神经网络(CNNs).随后,基于 CNNs 的 SISR 方法致力于设计更深更广的网络架构以实现性能提高.然而这类方法往往参数量和计算量巨大,例如 EDSR^[4]和 RCAN^[5]等.在现实场景下真正需求的是轻量化网络模型,尤其是可以部署到移动设备的实时模型,因而如何更好地实现网络轻量化是一个亟待解决的问题.

SISR 模型轻量化首先想到的便是设计特殊的模型,减少网络的层数、使用小的卷积核、采用轻量化组件,例如分组卷积^[6]和可变形卷积^[7]等.但是特殊设计的模型依旧存在问题:一是轻量级网络与 EDSR 和 RCAN 等大网络相比具有很大的差距;二是轻量级模型中一些组件在实际硬件芯片上不能被很好地支持. SISR 领域的蒸馏方法应运而生,其使用性能好的大模型(教师网络)蒸馏信息提升小模型(学生网络),使学生网络也能获得接近教师网络的性能.本文基于第二种方法探索.

SISR 领域的蒸馏方法中,首次提出的 SRKD^[6]对教师网络和学生网络的不同阶段进行蒸馏,考虑到教师网络和学生网络通道数不相同,SRKD 对中间信息的统计进行监督.随后,FAKD 在 SRKD 的基础上引入了特征关联机制^[8],进一步提升了学生网络的性能.但这些方法仍有一定的局限性,于是 PISR 首次引入特权信息(HR 图像)^[9],用教师网络提取 HR 高频特征补充到学生网络,明显提升了重建性能.利用特权信息的方法最近得到了明显的关注,基于参考图像的超分辨率方法 C2-Matching^[10]利用特权信息大幅提升了性能(目前最好的 RefSR 模型).明显地,教师网络的性能决定了学生网络的上限,但是现有利用特权信息的方法存在很大的问题.

收稿日期:2022-06-12;修回日期:2022-07-26.

基金项目:重庆市教委科学技术研究项目(KJQN202002501);重庆市高等教育教学改革研究项目(213473).

作者简介:毛盼娣(1985—),女,安徽宿州人,重庆城市科技学院副教授,研究方向为电子技术、图像处理等.

通信作者:徐道连(1960—),男,四川广安人,重庆大学副教授,博士,研究方向为微型机械传感器、噪声及振动的理论与检测等,E-mail:kexuelove2000@163.com.

根据局部相似性,教师网络提取 HR 高频特征因没有很好的约束导致学生网络对输入互补性不够准确,如图 1 所示.针对这个问题,本文提出了高效单图像超分辨率重建^[11]:深监督对称蒸馏网络(DSSD).DSSD 针对教师网络提取 HR 的高频特征不够准确提出了一个新颖的教师网络.为了约束中间层构造了高频特征递归模块(HFRM),这种递归方式对中间特征约束更强.为了使教师网络学习到足够准确的高频特征,提出了一种对称退化结构(SDM),SDM 与 HFRM 一样进一步约束了教师网络中的 SR.特别地,整个教师网络使用深监督加强对教师网络的约束,同时解决了教师网络难以训练的问题,使转移到学生网络的蒸馏知识更加准确.同时,本文为了更好适用于移动设备,选择只有 13 k 参数量和 6 GFlops 计算量的 FSRCNN^[6]网络.本文主要贡献:1)提出了一种新颖的通用蒸馏教师网络架构 DSSD,该架构可以更准确地提取 HR 高频细节.2)其中构造了 HFRM 和 SDM 模块,目的是使教师网络更准确地提取 HR 高频特征.这两个模块采取递归方式并且结构简单.3)为了约束教师网络中提取的高频特征,采用深监督方法使教师网络蒸馏的知识可以与学生网络互补.

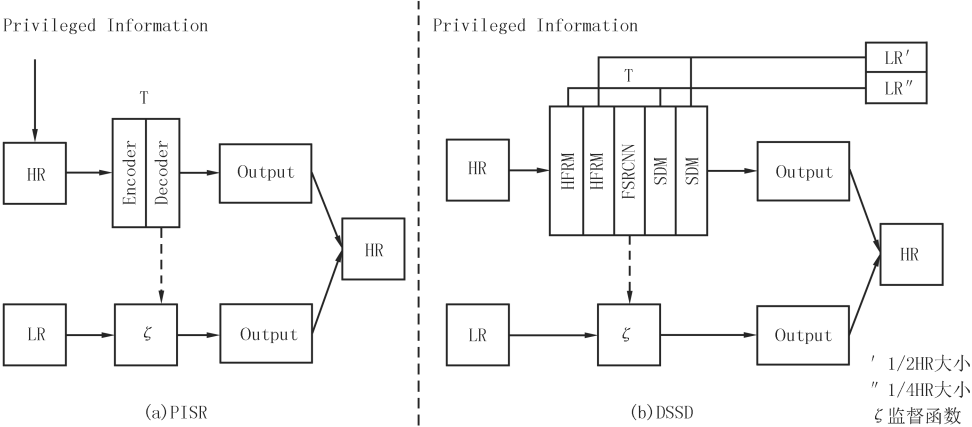


图1 PISR和DSSD模型
Fig.1 PISR and DSSD model

1 相关工作

1.1 单图像超分辨率

随着 CNNs 的发展,DONG 等^[3]首次提出基于 CNNs 的 SISR 方法 SRCNN 学习 LR 到 HR 的映射关系.之后的方法使用大量卷积层,利用密集和跳跃连接等保留更多的特征^[12],与之前方法相比有了显著的提升.为了减少显存和运行时间,高效 SISR 被提出.对于显存高效方法,其本质是利用特殊设计的架构减少参数量.对于高效运行时间方法,其本质是较低的计算量,使用级联、多分支架构或组卷积等方式.但是这些特殊设计的高效 SISR 方法不能适用于移动设备.

1.2 知识蒸馏

知识蒸馏的提出是为了转移大模型(教师网络)软标签到小模型(学生网络)以增强小模型性能,它被广泛运用在网络压缩上^[13].之后特征蒸馏被提出,可以从教师网络转移特征至学生网络,帮助学生网络更好地训练.利用特权信息的方法本质是使用额外信息,这在训练时需要额外成本但是在测试时不需要,被称之为广义蒸馏.广义蒸馏能够从教师网络转移到学生网络更多的信息以增强学生网络.特别地,变分信息蒸馏 VID 通过最大化教师和学生网络特征之间的互信息来转移知识,本文同 PISR 一样使用 VID.

本文方法相关方法包括 SRKD,FAKD 和 PISR.SRKD 对教师网络和学生网络不同阶段进行蒸馏,考虑到教师网络与学生网络通道数不同,SRKD 对中间特征的统计信息进行监督.FAKD 改进了 SRKD,在整个蒸馏的过程中 FAKD 与 SRKD 类似,区别在于 FAKD 在蒸馏损失方面提出了特征关联机制,取得了不错的效果.最近 PISR 被提出,PISR 首次将 HR 作为特权信息引入 SISR.PISR 与 SRKD、FAKD 有本质的不同,其将 HR 直接输入教师网络,允许教师网络提取更多的信息转移到学生网络.

1.3 深监督

深监督是在训练神经网络时,针对中间隐藏层特征透明度不高以及深层网络难以训练的问题,对隐藏层进行再监督,促进网络更快更好收敛.DSN 首次提出深监督概念,但 DSN 缺点在于使用的网络结构不够深,为了更好地训练深度网络,通过给神经网络的某些层添加了辅助监督来解决这个问题.此外,应用于分割的 Unet++ 也适用深监督增强医学图像分割的性能.由于教师网络中间层较多,为了更好约束中间层,使中间特征能更准确地蒸馏到学生网络,本文采用深监督的方式,明显提升了 DSSD 的性能.

2 方 法

本文提出了一个新颖的深监督对称蒸馏教师网络.该网络利用特权信息训练教师网络,训练好的教师网络会帮助学生网络更好地重建.由于篇幅有限,本文以倍率 4 为例:使用 X 和 Y 表示 LR 和 HR 图像.其中 ' $\prime$ ' 表示二分之一 HR 大小, ' $\prime\prime$ ' 表示四分之一 HR 大小, ζ 表示监督函数.

这种利用特权信息的对称结构类似于在课堂学习(训练阶段)时,不仅仅会有正确答案(HR),而且老师会解释其中的来龙去脉(HFRM),更进一步老师会告知错在哪里(SDM)以及该怎么修正(深监督).

2.1 教师网络

学生网络性能和教师网络性能直接相关,所以提高教师网络性能是关键.为此本文设计了一种对称教师网络结构如图 2 所示,该网络输入与 PISR 相同(HR).使用 HFRM 模块逐级下采样,每经过一个 HFRM 模块分辨率缩小二分之一,在递归 2 个 HFRM 模块后变成了 HR 分辨率的四分之一,与 LR 尺寸相同可以输入 FSRCNN.在经过 FSRCNN 之后生成 SR,将 SR 也采用与 HFRM 相同的下采样可以保证 SR 蕴含更多的高频特征,于是构造了 SDM 模块.特别地,SDM 与 HFRM 模块在结构上相同,形成一种对称网络结构.同时为了转移更准确的高频特征,教师网络采取深监督的方式对中间层进行约束,取得了不错的效果.

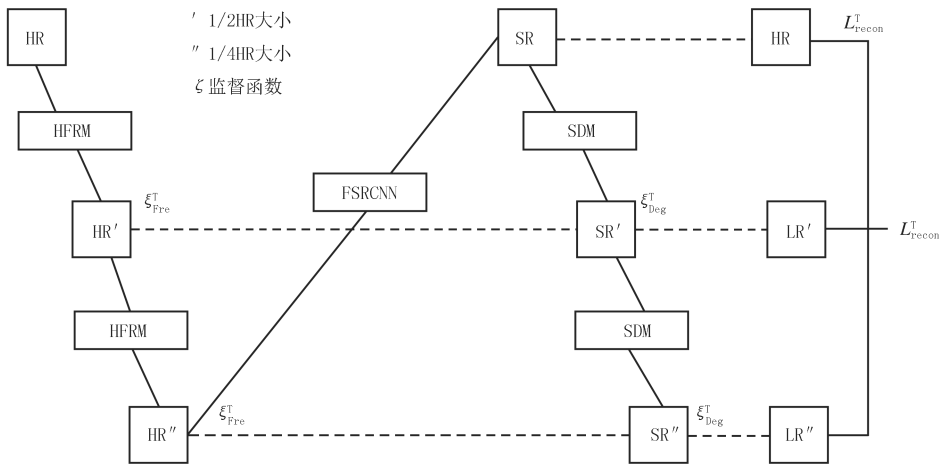


图2 教师网络整体图

Fig.2 Overall map of the teachers' network

2.1.1 高频特征递归模块

本文构造了 HFRM 模块将 HR 递归投影到低维特征空间生成紧凑的高频特征.与机械式双三次下采样不同,基于 CNNs 的 HFRM 能学习到更多与 LR 互补的高频特征,然后从紧凑特征重构 SR 图像.递归 HFRM 受到深监督的约束可以提取更准确的高频特征.这一过程表示为式(1)、(2):

$$\hat{X}'^T = G^T(Y), \quad (1)$$

$$\hat{X}''^T = G^T(\hat{X}'^T). \quad (2)$$

HFRM 模块详细的结构如图 3 所示,k3n56s1 为先经过一个 3×3 的卷积将 3 通道扩展到 56(为了与 FSRCNN 保持一致),步长大小为 1,激活函数为 PReLU.第 2 个 3×3 的卷积步长为 2 进行下采样,之后再有两个 3×3 的卷积,先将 56 通道降至 12 通道,再变为 3 通道.经过 2 个 HFRM 生成与 LR 大小一致的紧凑特征交给 FSRCNN 进行重构,其中递归能很好地受到深监督约束.

2.1.2 重建网络

本文目的是构造适用于移动设备的网络,既要很小的参数量也要很小的计算量,同时避免特殊设计的网络架构,所以 FSRCNN 是一个好的选择.教师网络中重建网络过程如式(3)所示:

$$\hat{Y}^T = F^T(\hat{X}^{T^T}), \quad (3)$$

其中: $\hat{Y}^T$ 是重建网络生成的输出, $F^T(\hat{X}^{T^T})$ 是一个函数 F 的应用于输入 $\hat{X}^{T^T}$ 的结果, $\hat{X}^{T^T}$ 是输入数据.

2.1.3 对称退化模块

为了保证经过 FSRCNN 重建的 SR 保留足够准确的高频特征,本文构造了 SDM 模块(图 4),之前大部分方法主要是最小化 SR 与 HR 之间的损失而忽略了解空间的问题.对于 LR 来说,LR 与 SR 是一种一对多的关系,利用 SDM 模块去约束其解空间的大小,保证 SR 具有更准确的高频特征.特别为了保持简单的结构,SDM 模块与 HFRM 模块有着相同的构造,同样采用递归下采样的方式并受到深监督的约束,如式(4)、(5):

$$\hat{D}'^T = G^T(\hat{Y}^T), \quad (4)$$

$$\hat{D}''^T = G^T(\hat{D}'^T), \quad (5)$$

其中: $\hat{D}'^T$ 表示高分辨率(HR)图像的特征, G^T 表示一种从输入到特征表示的映射, $\hat{Y}^T$ 是输入图像 $\hat{X}^T$ 的估计, $\hat{D}''^T$ 表示超分辨率(SR)图像的特征.

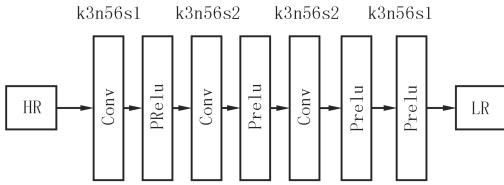


图3 HFRM模块

Fig.3 HFRM module

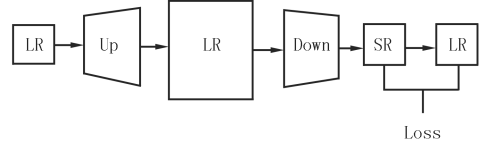


图4 SDM模块

Fig.4 SDM module

2.2 深监督

DSSD 除了新颖的对称架构以外,还有一个明显的特点:它由多个相同的监督构成深监督约束.如图 2 所示,每个 HFRM 模块可以下采样 2 倍,于是 4 倍率需要递归 2 次.HR 图像在经过 HFRM 模块之后会保留更多高频特征,这种特征会转移到学生网络.但是在保留的过程中,注意到学生网络的输入是双三次下采样得到的 LR 图像,尽管尽可能多地保留高频特征对学生网络有用,但是总体上不宜与 LR 偏差太大(局部相似性).为了解决这个问题,本文使用深监督策略,每一级分辨率都受到监督,保证在保留高频特征的同时保证合理性.

$$L_{\text{recon}}^T = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W |Y_{ij} - \hat{Y}_{ij}^T|, \quad (6)$$

$$\zeta'_{\text{Fre}}^T = \frac{1}{H'W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} |X'_{ij} - \hat{X}'_{ij}^T|, \quad (7)$$

$$\zeta''_{\text{Fre}}^T = \frac{1}{H''W''} \sum_{i=1}^{H''} \sum_{j=1}^{W''} |X''_{ij} - \hat{X}''_{ij}^T|, \quad (8)$$

$$L_{\text{Fre}}^T = \zeta'_{\text{Fre}}^T + \zeta''_{\text{Fre}}^T, \quad (9)$$

$$\zeta'_{\text{Deg}}^T = \frac{1}{H'W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} |X'_{ij} - \hat{D}'_{ij}^T|, \quad (10)$$

$$\zeta''_{\text{Deg}}^T = \frac{1}{H''W''} \sum_{i=1}^{H''} \sum_{j=1}^{W''} |X''_{ij} - \hat{D}''_{ij}^T|, \quad (11)$$

$$L_{\text{Deg}}^T = \zeta'_{\text{Deg}}^T + \zeta''_{\text{Deg}}^T, \quad (12)$$

$$L_{\text{total}}^T = L_{\text{recon}}^T + \lambda_{\text{Fre}}^T L_{\text{Fre}}^T + \lambda_{\text{Deg}}^T L_{\text{Deg}}^T, \quad (13)$$

L_{recon}^T 代表 FSRCNN 重建损失, L_{Fre}^T 代表经过 HFRM 模块的损失, X'_{ij} 为 1/2HR 大小, X''_{ij} 为 1/4HR 大小.同理, L_{Deg}^T 为 SDM 模块的损失.其中,式(13)中 λ_{Fre}^T 和 λ_{Deg}^T 均为 10^{-4} 大小.

2.3 学生网络

学生网络拥有与教师网络中重建网络(FSRCNN)相同的结构,但是输入不同.学生网络的输入为双三次

下采样的 LR,然后重建 SR,如式(14):

$$\hat{Y}^S = F^S(X), \quad (14)$$

初始化网络权重对学生网络的重建性能至关重要^[14].使用教师网络的蒸馏知识初始化学生网络,可以将教师网络训练好的重建能力转移给学生网络,为后期优化提供了很好的起点.

尽管教师网络的蒸馏知识可以很好地初始化学生网络,但是由于学生网络的输入不同,因此需要使用重建损失和蒸馏损失进一步优化学生网络.其中式(16)中 λ^S 为 10^{-6} .

$$L_{\text{recon}}^S = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W |Y_{ij} - \hat{Y}_{ij}^S|, \quad (15)$$

$$L_{\text{total}}^S = L_{\text{recon}}^S + \lambda^S L_{\text{distill}}^S. \quad (16)$$

蒸馏损失将教师网络的蒸馏知识转移到学生网络,本文同 PISR 使用 VID.目的是最大化教师网络和学生网络对应特征信息熵之间的互信息,

$$L_{\text{distill}}^S = \frac{1}{CHW} \sum_{i=1, j=1, k=1}^{CHW} \ln b_{ijk} + \frac{|f_{ijk}^T - \mu_{ijk}|}{b_{ijk}}. \quad (17)$$

由于教师网络和学生网络输入不同,当学生网络不能从蒸馏中获益时,尺度参数 b_{ijk} 能自适应地增大以减少蒸馏程度. $\ln b_{ijk}$ 相当于一个正则项,防止 b_{ijk} 造成 loss 的一个平凡解. μ 和 b 从学生网络的特征中估计,其中 μ 和 b 大小均为 $C \times H \times W$.估计方法是使用一个小网络,该网络有两个分支,分别用于估计 μ 和 b .两个分支使用相同的 1×1 卷积层,中间为 PReLU 激活函数.对于估计 b 的尺度分支会在最后一层添加 softplus 函数,作用是使尺度参数为正,估算模块是为蒸馏过程服务的,仅在训练时使用,

$$\zeta(x) = \ln(1 + e^x). \quad (18)$$

3 实 验

3.1 实验细节

本文使用 DIV2K 数据集训练,该数据集由 800 对 LR 和 HR 图像组成,其中 LR 是由 HR 双三次下采样得到.随机将 HR 图像剪切为 192×192 大小的块,同样的 LR 块取自对应比例因子的 LR 图像.同时使用数据增强技术,包括随机旋转和水平翻转.教师网络使用随机初始化训练,训练模型时,批处理大小为 16,一共训练 1 000 轮,其中 $\beta_1 = 0.900$, $\beta_2 = 0.999$.对于学习率大小,设置为 10^{-3} 并且使用余弦退火技术将其减少到 10^{-5} .使用标准评估集,用峰值信号比(PSNR)和结构相似程度(SSIM)在 Y 通道上评估.本文运行于 NVIDIA Titan RTX.

3.2 实验结果

附表 I 比较了 DSSD 学生模型与最先进模型的性能,特别是高效 SISR 方法(* 表示使用 DIV2K 重新训练,最好和最坏结果分别用下划实线和下划曲线标出).为了进行定量对比,评价指标为标准数据集在 2、3、4 倍率下的 PSNR 和 SSIM,也包含模型的参数量\计算量\运行时间,其中计算量是在 $1\,280 \times 720$ 大小的 HR 图像上测量.从附表 I 可以看到:DSSD 学生模型在所有比例因子上均优于 PISR,甚至 4 倍率 Set5 超 PISR 接近 0.12 dB,但是大数据集提升不如小数据集.因 FSRCNN 只有 13 k 参数量和 6 GFlops 计算量,其重建能力在大数据集上已经捉襟见肘.同理,DSSD 在较小的比例因子下的提升也有限,是因为较小的比例因子下丢失的信息较少,小参数量的模型重建能力有限.针对这个问题,本文后续在 3.3.4 讨论了较大模型的性能提升.

3.3 消融实验

消融实验对 DSSD 每个组件进行对比实验,以验证每个模块的作用,组合在一起之后会达到最好的效果.

3.3.1 高频递归模块实验

为了验证 HFRM 的有效性,设计了表 1 对比实验:这里只探究 HFRM 模块,在设计此对比实验时去掉 SDM 模块以保证公平性.可以看到直接双三次下采样效果最差,因为机械式下采样就等于直接输入 LR 图像,导致教师网络和学生网络的输入相同.FSRCNN 的 PSNR 为 30.89 dB,直接双三次下采样方法训练本质上是用教师网络训练好的特征初始化学生网络.PISR 的 Encoder 使用 CNNs 学到一部分 LR 没有的高频特征,但转移的高频特征不够准确(如图 5 对比所示).HFRM 模块采用递归的方式约束中间特征,目的是可以

学习到更准确更合理的高频特征.

表 1 HFRM 模块的结果比较

Tab. 1 Comparison results of HFRM

方法	PSNR/dB
双三次下采样	30.93
PISR 的 Encoder	30.96
HFRM	31.04

3.3.2 对称退化模块实验

为了验证 SDM 的有效性,本文设计了表 2 对比实验.这里保留 HFRM 只改变 SDM 模块以获得公平的对比环境.从表 2 可以看出,去掉 SDM 就是表 1 最后一个实验.重点在表 2 的后 3 个消融实验:将 SDM 使用双三次下采样结果反而变差,原因是这种机械方式破坏了 SR 的退化结果,深监督会使 SR 向更坏的方向发展.而基于 CNNs 下采样限制 SR 的高频特征提高了 SR 的重建性能.最后是本文构造的 SDM 模块,在深监督的配合下进一步约束了 SR 高频信息的合理性,获得了更好的重建效果.

表 2 SR 的退化方式对比实验

Tab. 2 Comparison results of the way of degradation to SR

方法	去掉 SDM	双三次下采样	基于 CNNs 下采样	SDM
PSNR/dB	31.04	31.03	31.05	31.07

3.3.3 深监督实验

为了验证深监督的有效性设计了表 3 对比实验.这里保留 HFRM 和 SDM 模块以探究深监督对 DSSD 教师网络的整体贡献大小:DSSD 中的深监督如图 2 所示,除了 L^T_{recon} 不变以外对每一级递归的监督进行叠加实验.如表 3 所示,在只有 L^T_{recon} 时性能下降的很大.保留 ζ'^T_{Fre} 和 ζ'^T_{Deg} 性能有了大幅增加.进一步增加第二级递归监督形成最终的深监督,使 SR 的重建性能达到最大,证明了深监督有效.

表 3 监督对 SR 性能的影响

Tab. 3 The effect of supervision on SR performance

方法	只保留 L^T_{recon}	保留 $L^T_{recon}, \zeta'^T_{Fre}$ 和 ζ'^T_{Deg}	深监督
PSNR/dB	30.99	31.05	31.07

3.3.4 教师网络通用性对比实验

为了探究 DSSD 教师网络是否具有通用性,本文挑选了几个参数量较大的轻量级 SISR 方法,见表 4.图 6 为标准测试集在 4 视觉上的比较为 665 k 和 591 k, * 表示用 DIV2K 训练不设置蒸馏,可以看见 DSSD 在 SISR 方法都有提升.特别地,作用在较小的数据集 Set5 上提升较大,而在较大数据集上 B100 提升有限,这是因为 FSRCNN 参数量只有 13 k 已经到了性能极限.在增大参数量的网络 VDSR 上 DSSD 增强了更多,接近 0.18 dB(Set5),而在 IDN 上的大数据集 B100 上增加了 0.1 dB.这些实验验证了 DSSD 教师网络具有通用性.

表 4 其他 SR 方法的定量结果

Tab. 4 Quantitative results of other SR methods

Methods	4Set5/B100	Methods	4Set5/B100	Methods	4Set5/B100
FSRCNN*	30.89/27.05	VDSR*	31.37/27.25	IDN*	32.03/27.49
FSRCNN*(PSNR)	30.95/27.08	VDSR*(PISR)	31.51/27.29	IDN*(PISR)	32.01/27.51
FSRCNN*(Ours)	31.07/27.11	VDSR*(Ours)	31.69/27.35	IDN*(Ours)	32.08/27.61

3.3.5 视觉效果

图 6 展示了使用学生网络在标准测试集上的重建结果.可以清楚地看到学生模型提供了比原始基线模

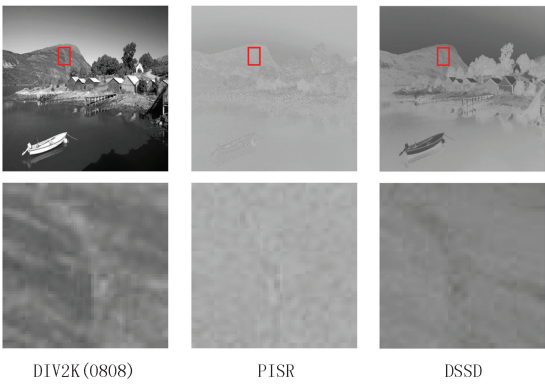


图5 PISR和DSSD提取高频信息能力对比

Fig.5 Comparison of the ability of PISR and DSSD to extract high frequency information

型(FSRCNN)和目前 SOTA 蒸馏方法 PISR 更好的性能.提出 DSSD 的目的在于探究如何更好地提取 HR 中包含的高频纹理细节,这些高频纹理细节可以使图像的边缘更清晰更明显,可以明显看见 B100 上数据集少了 PISR 的棋盘伪影,同时 Urban100 上数据集边缘明显变清晰锐利,Set14 中 barbara 的重建效果明显纹理细节更多.这些效果图证明了 DSSD 确实具有很好提取 HR 高频细节的能力.

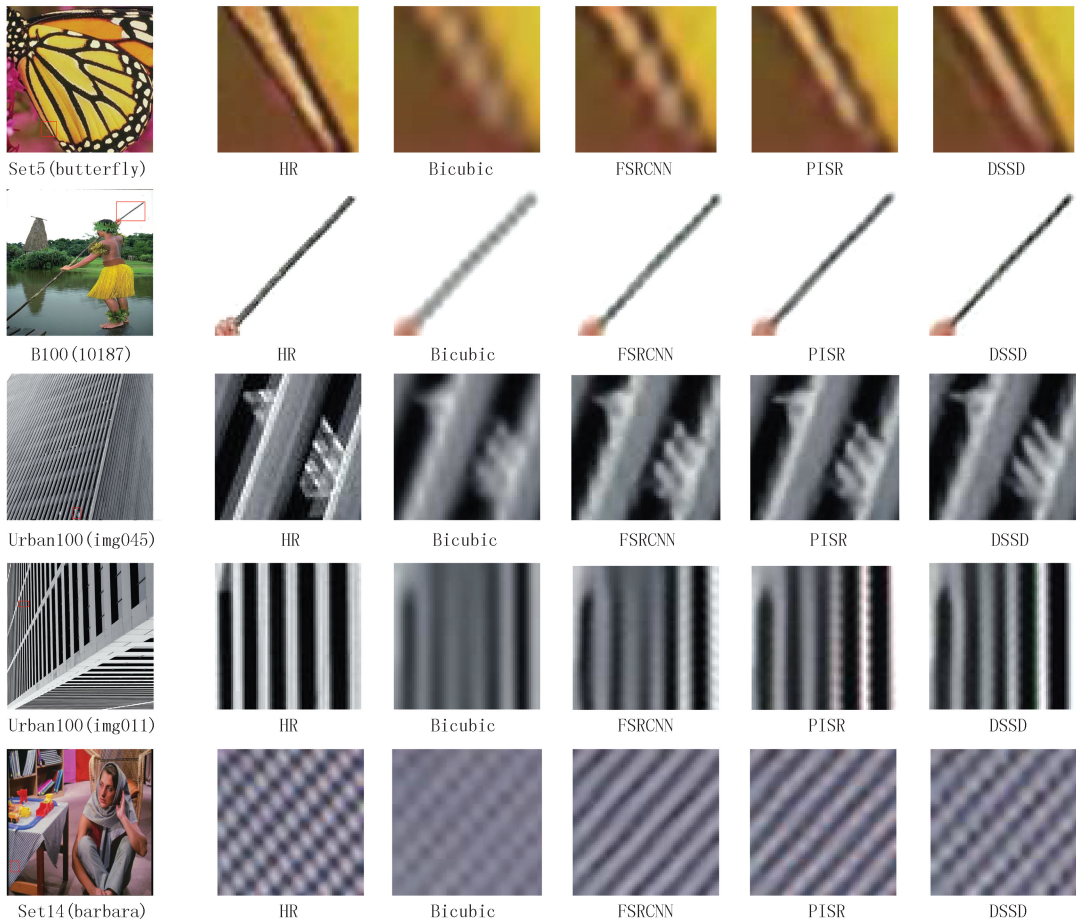


图6 标准测试集在4视觉上的比较

Fig.6 Comparison in 4× vision on the standard test set

4 结 论

本文提出了一种新颖的高效单图像超分辨率重建方法:深监督对称蒸馏网络.针对教师网络提取 HR 高频信息不够准确这一问题,构造了高频特征递归模块和对称退化模块,有效提高了重建性能.在整个蒸馏的教师网络中,利用深监督能更好地约束中间特征使提取的特征更准确,消融实验很好地证明了本文方法的有效性,将在未来的工作中继续探索蒸馏方法中特征转移部分以进一步提高蒸馏方法的性能.

附 录

附表 I 见电子版(DOI:10.16366/j.cnki.1000-2367.2023.06.007).

参 考 文 献

[1] BAI Y C,ZHANG Y Q,DING M L,et al.SOD-MTGAN:small object detection via multi-task generative adversarial network[C]//Computer Vision-ECCV 2018:15th European Conference.New York:ACM,2018:210-226.

[2] LIN W S,TJOA S K,ZHAO H V,et al.Digital image source coder forensics via intrinsic fingerprints[J].IEEE Transactions on Information Forensics and Security,2009,4(3):460-475.

[3] 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.

[4] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). [s.l.]: IEEE, 2017: 1132-1140.

[5] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//Computer Vision-ECV 2018: 15th European Conference. New York: ACM, 2018: 294-310.

[6] IOANNOU Y, ROBERTSON D, CIPOLLA R, et al. Deep roots: improving CNN efficiency with hierarchical filter groups[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). [s.l.]: IEEE, 2017: 5977-5986.

[7] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). [s.l.]: IEEE, 2017: 764-773.

[8] HE Z B, DAI T, LU J, et al. Fakd: feature-affinity based knowledge distillation for efficient image super-resolution[C]//2020 IEEE International Conference on Image Processing(ICIP). [s.l.]: IEEE, 2020: 518-522.

[9] LEE W, LEE J, KIM D, et al. Learning with privileged information for efficient image super-resolution[M]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 465-482.

[10] 刘嘉佩, 曹林, 杜康宁. 双层级神经网络的人脸超分辨率重建[J]. 计算机工程与应用, 2020, 56(13): 210-215.

LIU J P, CAO L, DU K N. Face image super-resolution based on two-layer cascade neural network[J]. Computer Engineering and Applications, 2020, 56(13): 210-215.

[11] 刘月峰, 杨涵晰, 蔡爽, 等. 基于改进卷积神经网络的单幅图像超分辨率重建方法[J]. 计算机应用, 2019, 39(5): 1440-1447.

LIU Y F, YANG H X, CAI S, et al. Single image super-resolution reconstruction method based on improved convolutional neural network [J]. Journal of Computer Applications, 2019, 39(5): 1440-1447.

[12] TONG T, LI G, LIU X J, et al. Image super-resolution using dense skip connections[C]//2017 IEEE International Conference on Computer Vision(ICCV). [s.l.]: IEEE, 2017: 4809-4817.

[13] 仇真, 奚雪峰, 崔志明, 等. 基于多分辨率自蒸馏网络的小样本图像分类[J]. 计算机工程, 2022, 48(12): 232-240.

QIU Z, XI X F, CUI Z M, et al. Few-shot image classification based on multi-resolution self-distillation network[J]. Computer Engineering, 2022, 48(12): 232-240.

[14] JIANG Y M, CHAN K C K, WANG X T, et al. Robust reference-based super-resolution via C2-matching[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). [s.l.]: IEEE, 2021: 2103-2112.

Efficient single-image super-resolution: deeply-supervised symmetric distillation network

Mao Pandi¹, Xu Daolian²

(1. School of Electrical Engineering and Intelligent Manufacturing, Chongqing Metropolitan College of Science and Technology, Chongqing 402167, China; 2. School of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China)

Abstract: Convolutional neural networks(CNNs) have made significant advances in single image super resolution(SISR) over the past few decades. Most CNNs-based approaches nowadays are devoted to constructing new architectures to improve reconstruction performance, which usually rely on large computation and storage costs and are difficult to apply to mobile devices. In this paper, we propose a novel efficient single-image super-resolution reconstruction method(DSSD) based on deeply supervised symmetric distillation networks, the problem that the high-frequency information of High-Resolution extracted in the teacher network is not accurate enough is alleviated by constructing a high-frequency feature recursion module(HFRM) and a symmetric degradation module(SDM). To constrain the high frequency features extracted from the teacher network, a deep supervision approach was used to make the knowledge distilled from the teacher network complementary to the student network. Experiments on the DIV2K dataset show that DSSD effectively enhances the performance of SISR, and the introduction of HFRM and SDM can effectively help DSSD to extract more high-frequency details of images.

Keywords: deeply-supervised symmetric distillation networks; super-resolution; teacher network; high-frequency feature recursion module; symmetric degeneration module; privileged information

附表 I 各种高效 SISR 方法在标准测试集上的平均 PSNR 和 SSIM 值

Attached tab. I Average PSNR and SSIM values of various efficient SISR methods on the standard test set								
Scale	Methods	Param./k	MultiAdds/G	Runtime/ms	Set5	Set14	B100	Urban100
2	FSRCNN	13	6.0	0.83	<u>37.05/0.956 0</u>	<u>32.66/0.909 0</u>	<u>31.53/0.892 0</u>	<u>29.88/0.902 0</u>
2	FSRCNN *	13	6.0	0.83	37.15/0.956 8	32.71/0.909 5	31.58/0.891 3	30.05/0.904 1
2	PSIR	13	6.0	0.83	37.33/0.957 6	32.79/0.910 5	31.65/0.892 6	30.24/0.907 1
2	DSSD(ours)	13	6.0	0.83	<u>37.34/0.958 0</u>	<u>32.81/0.910 8</u>	<u>31.67/0.892 8</u>	<u>30.28/0.907 9</u>
2	Bicubic	—	—	—	<u>33.66/0.929 9</u>	<u>30.24/0.868 8</u>	<u>29.56/0.843 1</u>	<u>26.88/0.840 3</u>
2	DRCN	1 774	17 974.3	239.93	<u>37.63/0.958 8</u>	<u>33.04/0.911 8</u>	<u>31.85/0.894 2</u>	<u>30.75/0.913 3</u>
2	DRRN	297	6 796.9	105.76	37.74/0.959 1	33.23/0.913 6	32.05/0.897 3	31.23/0.918 8
2	MemNet	677	2 662.4	21.06	37.78/0.959 7	33.28/0.914 2	32.08/0.897 8	31.31/0.919 5
2	CARN	1 592	222.8	8.43	37.76/0.959 0	33.52/0.916 6	32.09/0.897 8	31.92/0.925 6
2	IDN	591	136.5	7.01	37.83/0.960 0	33.30/0.914 8	32.08/0.898 5	31.27/0.919 6
2	SRFBN	3 631	1 126.7	108.52	<u>38.11/0.960 9</u>	<u>33.82/0.919 6</u>	<u>32.29/0.901 0</u>	<u>32.62/0.932 8</u>
2	IMDN	694	159.6	6.97	38.00/0.960 5	33.63/0.917 7	32.19/0.899 6	32.17/0.928 3
3	FSRCNN	13	6.0	0.83	33.18/ <u>0.914 0</u>	<u>29.37/0.824 0</u>	<u>28.53/0.791 0</u>	<u>26.43/0.808 0</u>
3	FSRCNN *	13	6.0	0.83	<u>33.15/0.915 7</u>	29.45/0.825 0	<u>28.52/0.789 5</u>	26.49/0.808 9
3	PSIR	13	6.0	0.83	33.31/0.917 9	<u>29.57/0.827 6</u>	28.61/0.791 9	26.67/0.815 3
3	DSSD(ours)	13	6.0	0.83	<u>33.36/0.918 9</u>	29.51/0.827 3	<u>28.63/0.792 4</u>	<u>26.68/0.815 4</u>
3	Bicubic	—	—	—	<u>30.39/0.868 2</u>	<u>27.55/0.774 2</u>	<u>27.21/0.738 5</u>	<u>24.46/0.734 9</u>
3	DRCN	1 774	17 974.3	239.93	33.82/0.922 6	29.76/0.831 1	28.80/0.796 3	27.15/0.827 6
3	DRRN	297	6 796.9	105.76	34.03/0.924 4	29.96/0.834 9	28.95/0.800 4	27.53/0.837 8
3	MemNet	677	2 662.4	21.06	34.09/0.924 8	30.00/0.835 0	28.96/0.800 1	27.56/0.837 6
3	CARN	1 592	222.8	8.43	34.29/0.925 5	30.29/0.840 7	29.06/0.803 4	28.06/0.849 3
3	IDN	591	136.5	7.01	34.11/0.925 3	29.99/0.835 4	28.95/0.801 3	27.42/0.835 9
3	SRFBN	3 631	1 126.7	108.52	<u>34.70/0.929 2</u>	<u>30.51/0.846 1</u>	<u>29.24/0.808 4</u>	<u>28.73/0.864 1</u>
3	IMDN	694	159.6	6.97	34.36/0.927 0	30.32/0.841 7	29.09/0.804 6	28.17/0.851 9
4	FSRCNN	13	6.0	0.83	<u>30.72/0.866 0</u>	<u>27.61/0.755 0</u>	<u>26.98/0.715 0</u>	<u>24.62/0.728 0</u>
4	FSRCNN *	13	6.0	0.83	30.89/0.874 8	27.72/0.759 9	27.05/0.717 6	24.76/0.735 8
4	PSIR	13	6.0	0.83	30.95/0.875 9	27.77/0.761 5	27.08/0.718 8	24.82/0.739 3
4	DSSD(ours)	13	6.0	0.83	<u>31.07/0.878 3</u>	<u>27.82/0.762 2</u>	<u>27.11/0.719 4</u>	<u>24.88/0.740 1</u>
4	Bicubic	—	—	—	<u>28.42/0.810 4</u>	<u>26.00/0.702 7</u>	<u>25.96/0.667 5</u>	<u>23.14/0.657 7</u>
4	DRCN	1 774	17 974.3	239.93	31.53/0.885 4	28.02/0.767 0	27.23/0.723 3	25.14/0.751 0
4	DRRN	297	6 796.9	105.76	31.68/0.888 8	28.21/0.772 1	27.38/0.728 4	25.44/0.763 8
4	MemNet	677	2 662.4	21.06	31.74/0.889 3	28.26/0.772 3	27.40/0.728 1	25.50/0.763 0
4	CARN	1 592	222.8	8.43	32.13/0.893 7	28.60/0.780 6	27.58/0.734 9	26.07/0.783 7
4	IDN	591	136.5	7.01	31.82/0.890 3	28.25/0.773 0	27.41/0.729 7	25.41/0.763 2
4	SRFBN	3 631	1 126.7	108.52	<u>32.47/0.898 3</u>	<u>28.81/0.786 8</u>	<u>27.72/0.740 9</u>	<u>26.60/0.801 5</u>
4	IMDN	694	159.6	6.97	32.21/0.894 8	28.58/0.781 1	27.56/0.735 3	26.04/0.783 8