

基于边缘感知的图像复原神经网络损失函数^①王昱昊^{②*} 李萌^{③**} 支天^{**} 张曦珊^{**} 周学海^{*}

(* 中国科学技术大学计算机科学与技术学院 合肥 230026)

(** 中国科学院计算技术研究所 北京 100086)

摘要 像素级损失是图像复原任务中最基本、最常用的损失函数。然而,像素级损失无法区分图像中的边缘部分是否模糊,对所有像素点的处理一致,这与人类视觉系统对边缘部分的感知不同。通过对图像边缘部分的分析,本研究发现当输出图像模糊时,其与真实图像对应像素的误差方向与真实图像中该像素的二阶梯度方向相同。基于此本文提出了一种基于边缘感知的结构保留损失函数(SPLoss)。本研究设计了模糊因子图来检测模糊像素,然后通过增加模糊位置像素误差的权重来惩罚模糊像素。SPLoss 可以检测到输出图像中的模糊像素,并在优化网络时对其进行惩罚。对于包括图像超分辨率和图像去模糊在内的图像复原任务,本文在多个网络结构上进行了实验。实验结果表明,使用本文提出的损失函数训练的网络可以恢复拥有更好的感知质量的图像,在感知相关客观指标和视觉效果方面均超越原有像素级损失。

关键词 图像复原;图像超分辨率;图像去模糊;神经网络(NN);损失函数

0 引言

图像复原是指从退化图像中恢复高质量图像的方法,包括图像超分辨率、图像去模糊等。近年来,基于神经网络(neural network, NN)的方法在图像复原任务上取得了巨大成功。除了网络结构^[1-3]在不断发展,作为优化目标的损失函数在基于神经网络的图像复原方法中也引起了很多关注^[4-5]。

基于神经网络的图像复原方法通常使用退化图像作为输入,并通过最小化输出图像和对应真实图像之间的误差来训练模型^[6],该误差由损失函数定义。

常用的损失函数可以分为逐像素损失和感知驱动损失。逐像素损失最小化输出图像像素值和真实图像像素值之间的绝对距离,例如均方误差(mean

squared error, MSE 或 L_2)^[7-8] 和平均绝对误差(mean absolute error, MAE 或 L_1)^[9-10]。感知驱动的损失比较了特征空间中输出图像和真实图像之间的差异,例如感知损失(perceptual loss)^[11-12] 和对抗损失(adversarial loss)^[4,13]。为了准确地生成高质量的真实输出,感知驱动的损失通常与逐像素损失一起使用。逐像素损失的优化目标是最大化峰值信噪比(peak signal-to-noise ratio, PSNR)。然而,一些工作表明,逐像素损失会产生一些模糊的结果,并且无法恢复图像的精微结构^[11,14]。

为了解决上述问题,本文为图像复原任务提出了一种基于边缘感知的结构保留损失函数(structure preserving loss function, SPLoss),它可以使经过训练的网络在保留精微图像结构的同时恢复准确的图像。为了保留图像结构,恢复图像的边缘是必要

① 国家自然科学基金(61925208,61906179,U20A20227),中国科学院战略性先导科技专项(XDB32050200)和中国科学院稳定支持基础研究领域青年团队计划(YSTR-029)资助项目。

② 男,1997年生,硕士生;研究方向:计算机系统结构;E-mail:ustcwyh@mail.ustc.edu.cn。

③ 通信作者,E-mail:limeng2020@ict.ac.cn。
(收稿日期:2021-11-30)

的。对于人类视觉系统,锐利的边缘拥有更好的感知质量。像素级损失(pixel-wise loss)无法区分图像边缘是模糊还是锐利,以 L_1 损失为例,输出图像的像素减去真实图像的像素在模糊边缘和锐利边缘可能具有相同的值但方向相反。因此,基于人类视觉系统对边缘的感知差异,本研究设计了一个方向敏感的逐像素损失函数。

本文提出的算法整个过程可以分为2个阶段。由于图像的二阶梯度图揭示了图像中边缘的方向和强度,首先使用真实图像的二阶梯度图结合恢复图像和真实图像之间的像素误差方向来检测模糊像素;然后通过增加这些像素中误差的权重来惩罚模糊像素。SPLoss将迫使网络对模糊像素进行反方向优化,最终得到精细结构。本研究对流行图像复原网络(包括图像超分辨率和图像去模糊)进行了广泛的实验来评估SPLoss的有效性。在本研究中,使用SPLoss重新训练原始网络的结果与原始网络的结果进行比较。实验结果表明,与原始损失函数相比,SPLoss成功地增强了恢复图像的清晰度和保真度。

1 相关工作

1.1 基于神经网络的图像复原方法

神经网络在高级语义任务中取得了巨大成功,并已应用于图像复原任务。早期的图像复原神经网络^[15-16]由多个卷积层组成,可以通过端到端的训练从损坏的图像中恢复原始图像,相对于传统算法具有优越的性能。后来,许多更深的网络架构被提出。一些新颖有效的架构也被用于图像复原任务,例如残差块^[4]、密集网络^[13]、递归学习^[17]、注意力机制^[10]和多级网络^[2,18]等。

1.2 图像复原任务的损失函数

损失函数在基于神经网络的图像复原方法中扮演了重要的角色,极大地影响了生成图像的质量。近年来,损失函数备受关注。图像复原的损失函数可以分为像素级损失和感知驱动损失。像素级损失计算图像对应像素的误差,感知驱动损失旨在提高感知质量。

像素级损失是最常用的损失函数,它约束了输出图像和真实图像像素之间的相似性。图像复原任务中最早使用的损失函数是MSE损失^[7-8]。MSE损失直接优化峰值信噪比(PSNR),这是图像复原任务的传统指标。近年来,由于使用 L_1 损失训练的模型达到了更高的PSNR,因此 L_1 损失已开始在许多工作^[9-10]中被使用。此外,一些工作^[2]使用了比 L_1 损失更鲁棒的沙博尼耶损失(Charbonnier loss),该损失函数的梯度在0值附近更加稳定。所有像素级损失都难以恢复图像的高频细节和纹理,这会导致图像模糊和感知质量差^[11,14]。

为了解决由像素级损失引起的问题,感知驱动损失被提出,它们通常使用预训练的神经网络模型来约束超出图像像素的特征。Johnson等人^[11]采用VGG网络作为特征提取器,误差被定义为特征域中的距离。Zhang等人^[19]提出了一种学习感知图像块相似性(learned perceptual image patch similarity, LPIPS)指标。他们在感知相似性数据集上校准了现有的分类网络。Rad等人^[12]使用生成的分割标签来估计合适的边界感知损失,同时考虑背景的纹理相似性。此外,还出现了一些基于生成对抗网络(generative adversarial network, GAN)的工作。对抗性损失使用鉴别器网络将解决方案推向自然图像流形。这些损失函数已经在图像超分辨率^[4,13]和图像去模糊^[20-21]任务中应用。Sajjadi等人^[22]还提出了纹理匹配损失能更好地产生图像细节。这些方法在重建细节的同时通常会产生不准确的纹理和不自然的伪像^[5],而且感知驱动的损失通常与像素级损失一起使用。

2 像素级损失在图像中边缘处的表现

图像模糊是图像复原任务中的一个重要问题,它会导致图像结构受到破坏。图像模糊倾向于减少当前像素和周围像素之间的差异,尤其是在边缘位置。本文观察到像素级损失无法区分图像边缘的像素是否模糊。

下面以 L_1 损失为例进行分析(其他像素损失类似)。如图1所示,清晰的图像比模糊的图像具有

更好的视觉质量,并且更好地保留了图像结构。然而,清晰图像和真实图像之间的 L_1 损失接近于模糊图像和真实图像之间的损失。在图1中,实线表示真实图像中边缘区域的像素值,虚线1和虚线2分别表示清晰图像和模糊图像中边缘区域的像素值。虚线1和实线之间的误差被定义为清晰误差,虚线2和实线之间的误差被定义为模糊误差。模糊误差和清晰误差的值相同但方向相反。但是,由于 L_1 损失不包含方向信息,无法区分这2种情况,模糊图像的 L_1 损失与清晰图像相同。大量研究表明, L_1 损失会导致模型输出模糊图像。

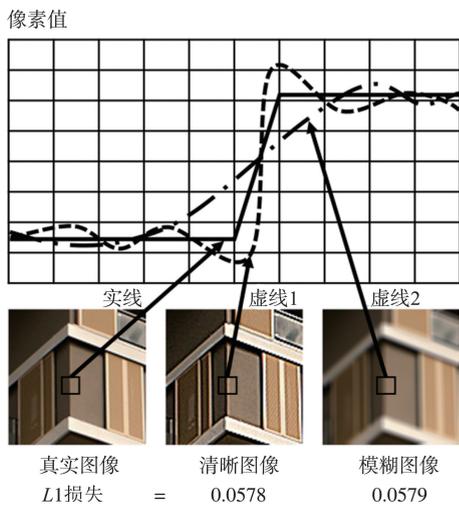


图1 L_1 损失在边缘处的表现

为了更好地保留图像中的锐利边缘和精细结构,损失函数应该区分边缘是否模糊,并且模糊像素应该得到比正常像素更大的惩罚。受此启发,本文提出了一种基于边缘感知的结构保留损失函数(SPLoss)。

3 结构保留损失函数

SPLoss的计算流程包括2个阶段:检测模糊像素和惩罚模糊像素。整个流程如图2所示。

3.1 检测模糊像素

本文设计了模糊因子图来检测模糊像素。输出图像中的像素是否模糊,与其和真实图像中对应像素的相对位置以及真实图像中相邻像素的变化方向

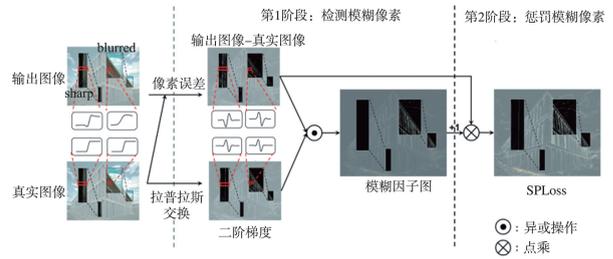


图2 SPLoss算法流程

有关。二阶梯度可以反映图像像素变化的方向和强度。由观察可知,当输出图像中的像素被模糊时,它与真实图像中对应像素之间的误差方向与真实图像中该位置的二阶梯度方向相同。为了解释这一点,首先明确以下定义。 $y_{i,j}^{out}$ 定义为输出图像中位置 (i,j) 处的像素, $y_{i,j}^{gt}$ 为真实图像中位置 (i,j) 处的像素, $\nabla^2 y_{i,j}^{gt}$ 为由Laplace算子在真实图像中位置 (i,j) 处计算的二阶梯度。输出图像与位置 (i,j) 处的真实图像之间的误差定义为

$$e_{i,j} = y_{i,j}^{out} - y_{i,j}^{gt} \quad (1)$$

如图2所示,如果该像素是模糊像素,则像素误差的方向与真实图像中相应位置的二阶梯度方向相同。接下来定义与图像大小相同的模糊因子图。如果一个像素是模糊像素,那么模糊因子图上的值就是真实图像中对应位置的二阶梯度的绝对值;否则为0。位置 (i,j) 处的模糊因子图的值定义为

$$m_{i,j} = \begin{cases} |\nabla^2 y_{i,j}^{gt}|, \nabla^2 y_{i,j}^{gt} > 0, e_{i,j} > 0 \\ 0, \nabla^2 y_{i,j}^{gt} > 0, e_{i,j} \leq 0 \\ 0, \nabla^2 y_{i,j}^{gt} \leq 0, e_{i,j} > 0 \\ |\nabla^2 y_{i,j}^{gt}|, \nabla^2 y_{i,j}^{gt} \leq 0, e_{i,j} \leq 0 \end{cases} \quad (2)$$

3.2 惩罚模糊像素

在计算损失时,SPLoss提高了模糊像素与真实图像中相应像素之间的误差权重以惩罚模糊像素。误差由模糊因子图加权。位置 (i,j) 处像素误差的权重定义为

$$w_{i,j} = 1 + \alpha m_{i,j} \quad (3)$$

其中 α 表示惩罚因子,即对模糊像素的惩罚强度。最后,SPLoss定义为

$$L_{SP} = \frac{1}{WH} \sum_{i=0}^W \sum_{j=0}^H w_{i,j} |e_{i,j}| \quad (4)$$

在 $SPLoss$ 中,具有相同误差的像素在不同位置具有不同的权重。输出图像最终会被优化到一个清晰的方向,因为模糊的像素会受到更多的惩罚。当惩罚因子 α 为 0 时, $SPLoss$ 等价于 L_1 损失。惩罚因子 α 的影响在后面章节的实验中展开讨论。

4 实验与结果分析

本节评估了本文所提方法在不同图像复原任务(包括图像超分辨率和图像去模糊)中的最先进网络结构上的有效性。此外,还研究了惩罚因子 α 的影响。

4.1 训练方法和评价指标

用 $SPLoss$ 替换原有方法中像素级损失来重新训练最先进的图像复原网络结构,并将结果与原始结果进行比较。每个实验的实现细节在下面每个小节中详细描述。 $SPLoss$ 的惩罚因子 α 为 2,所有实验的超参数与原实验相同。所有实验均基于 PyTorch 框架,在 NVIDIA Tesla A100 GPU 上实现。

本实验中选择了 4 个常用的指标:PSNR、结构相似性(structural similarity, SSIM)^[23]、LPIPS^[19] 和自然图像质量评估器(natural image quality evaluator, NIQE)^[24]。较低的 LPIPS 和 NIQE 值表示较高的感知质量。需要注意的是,对图像超分辨率任务

的结果评估在 YCbCr 空间的 Y 通道中进行评估。对于 NIQE 指标,所有任务结果的评估都在 Y 通道进行。

4.2 图像超分辨率任务中的 $SPLoss$ 实验

对于图像超分辨率任务,实验中将 $SPLoss$ 应用于 EDSR^[25]、ESRGAN^[13] 和 NLSN^[1], 缩放因子为 $\times 4$ 。EDSR 和 NLSN 是先进的 PSNR 驱动的图像超分辨率方法。它们的原始损失函数是 L_1 损失。ESRGAN 是先进的感知驱动方法。特别地,对于 ESRGAN 模型,分别在用 L_1 损失(RRDBNet)训练的版本和用 L_1 损失、感知损失和对抗性损失训练的版本上进行实验。为了与原始版本加以区分,本文将用 $SPLoss$ 训练的版本分别命名为 EDSR-SP、RRDBNet-SP、ESRGAN-SP 和 NLSN-SP。

本实验使用 DIV2K 作为 EDSR-SP 和 NLSN-SP 的训练数据集,使用 DIV2K 和 Flickr2K 作为 RRDBNet-SP 和 ESRGAN-SP 的训练数据集;使用 4 种常用的图像超分辨率任务基准测试集对训练结果进行评估:Set5、Set14、B100、Urban100。

图像超分辨率的实验结果如表 1 所示。可以发现所有使用 $SPLoss$ 的方法都比原始方法实现了更好的 LPIPS 和 NIQE。除了 Set5 测试数据集之外,使用 $SPLoss$ 的方法比原始方法实现了更好的 SSIM。由于 L_1 损失直接优化 PSNR,因此仅使用 L_1 损

表 1 图像超分辨率任务定量结果对比

数据集	指标	EDSR	EDSR-SP	RRDBNet	RRDBNet-SP	NLSN	NLSN-SP	ESRGAN	ESRGAN-SP
Set5	PSNR	32.48	32.11	32.73	32.39	32.70	32.35	30.46	30.68
	SSIM	0.8988	0.8982	0.9011	0.9003	0.9000	0.8999	0.8516	0.8538
	LPIPS	0.1861	0.1785	0.1843	0.1799	0.1684	0.1649	0.0806	0.0697
	NIQE	7.1441	6.8699	7.1230	6.9179	7.0472	6.9121	5.2280	4.7905
Set14	PSNR	28.81	28.33	29.00	28.58	28.95	28.56	26.28	26.65
	SSIM	0.7879	0.7879	0.7915	0.7926	0.7889	0.7904	0.6980	0.7103
	LPIPS	0.2782	0.2667	0.2757	0.2642	0.2716	0.2626	0.1374	0.1254
	NIQE	6.1305	6.0522	6.1375	5.9915	6.0856	5.8453	3.8774	3.7424
B100	PSNR	27.72	27.41	27.84	27.52	27.81	27.56	25.29	25.73
	SSIM	0.7420	0.7454	0.7453	0.7491	0.7416	0.7463	0.6495	0.6669
	LPIPS	0.3763	0.3520	0.3723	0.3502	0.3568	0.3377	0.1695	0.1641
	NIQE	6.3498	5.9326	6.3598	6.0498	6.2978	5.9825	3.6636	3.4045
Urban100	PSNR	26.65	26.22	27.03	26.53	27.11	26.77	24.35	24.96
	SSIM	0.8036	0.8044	0.8152	0.8157	0.8146	0.8166	0.7327	0.7526
	LPIPS	0.2083	0.1877	0.2005	0.1804	0.1949	0.1774	0.1261	0.1194
	NIQE	5.4342	5.4191	5.6235	5.6156	5.4475	5.4104	4.2066	3.8892

失的方法可以获得更高的 PSNR。特别是对于使用感知损失和对抗性损失的 ESRGAN 模型,ESRGAN-SP 所有指标均有提升。接下来比较基于 SPLoss 的方法和原始方法之间的视觉效果。图 3 表示 NLSN 与 NLSN-SP 的视觉比较结果。图 4 表示 ESRGAN 与 ESRGAN-SP 的视觉比较结果。结果表明,基于 SPLoss 的方法比相应的原始方法更清晰、更逼真,更好地保留了图像的结构。

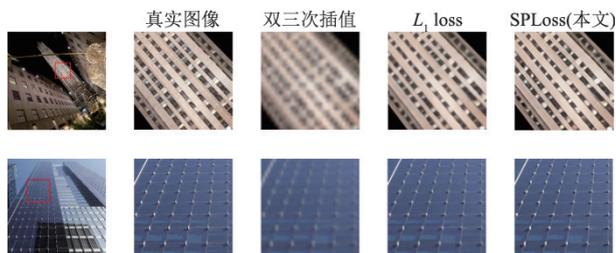


图 3 图像超分辨率任务 NLSN 网络视觉效果对比



图 4 图像超分辨率任务 ESRGAN 网络视觉效果对比

4.3 图像去模糊任务中的 SPLoss 实验

对于图像去模糊任务,在最先进的网络架构 MPRNet^[2]上试验 SPLoss。它的原始损失函数是 L_1

损失。用 SPLoss 训练的版本被命名为 MPRNet-SP。

本实验使用 GoPro 作为 MPRNet-SP 的训练及测试数据集,其中包含 2103 个用于训练的图像对和 1111 个用于验证的图像对。

图像去模糊的实验结果如表 2 所示。使用 SPLoss 训练的 MPRNet-SP 实现了更好的 SSIM、LPIPS 和 NIQE,比原始方法获得了可观的收益。图 5 显示了视觉比较的结果。

表 2 图像去模糊任务定量结果对比

数据集	指标	MPRNet	MPRNet-SP
GoPro	PSNR	32.66	32.44
	SSIM	0.9362	0.9371
	LPIPS	0.0908	0.0790
	NIQE	5.2189	4.7728

4.4 惩罚因子影响实验

在 SPLoss 中,惩罚因子 α 是影响生成图像质量和指标的重要参数。本文通过一系列实验探索不同 α 的影响。使用不同的 α (包括 1.0、2.0、3.0、4.0、6.0、8.0 和 10.0) 重新训练最先进的图像超分辨率网络结构 NLSN。训练数据集为 DIV2K,测试数据集为 Set5、Set14、B100、Urban100。对于每个指标,比较所有测试数据集的平均值。以 L_1 损失训练的原始结果作为本实验的参考对象。定量比较如表 3 所示。结果表明,当 α 增加时,PSNR 变低,但感知质量指标变好。当 α 为 2.0 时,SSIM 获得最佳值,其他指标平衡,图像的结构得到最好保留。

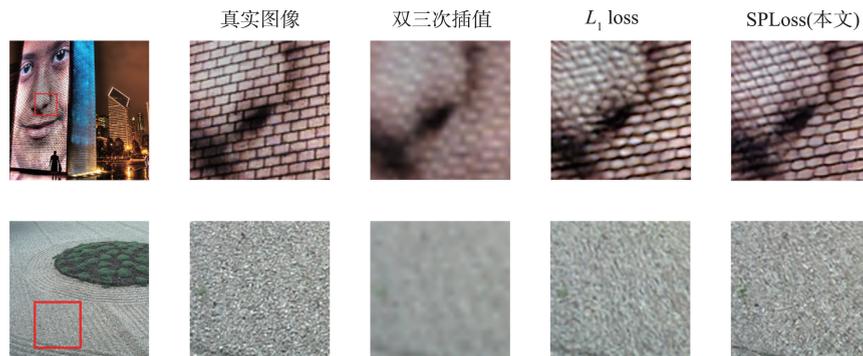


图 5 图像去模糊任务 MPRNet 网络视觉效果对比

表3 模糊因子的影响

指标	$L_1(\alpha=0.0)$	$\alpha=1.0$	$\alpha=2.0$	$\alpha=3.0$	$\alpha=4.0$	$\alpha=6.0$	$\alpha=8.0$	$\alpha=10.0$
PSNR	29.14	29.00	28.81	28.63	28.47	28.19	27.97	27.79
SSIM	0.8113	0.8133	0.8134	0.8128	0.8121	0.8104	0.8086	0.8071
LPIPS	0.2479	0.2397	0.2357	0.2335	0.2324	0.2315	0.2314	0.2315
NIQE	6.2195	6.1075	6.0376	5.9905	5.9531	5.9058	5.8582	5.8179

5 结论

本文基于人类视觉系统对边缘部分的感知差异,为基于神经网络的图像复原任务引入了一种基于边缘感知的结构保持损失函数(SPLoss)。本文所提出的损失函数可以检测并惩罚恢复图像中的模糊像素,这使经过训练的网络在恢复图像的同时保留精细图像结构。SPLoss 首先使用真实图像的二阶梯度图结合输出图像和真实图像之间像素误差的方向来检测模糊像素。然后通过增加这些像素中错误的权重来惩罚模糊像素。本研究在图像超分辨率和图像去模糊任务上进行了大量实验,并且探究了惩罚因子对结果的影响。一系列定量和定性实验结果表明,本文提出的方法可以在复原图像的同时保留精细结构,提升了图像感知质量,在感知相关客观指标和视觉效果方面均超越原有像素级损失。

参考文献

- [1] MEI Y, FAN Y, ZHOU Y. Image super-resolution with non-local sparse attention[C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 3516-3525.
- [2] ZAMIR S W, ARORA A, KHAN S, et al. Multi-stage progressive image restoration [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 14821-14831.
- [3] 肖进胜, 刘恩雨, 朱力, 等. 改进的基于卷积神经网络的图像超分辨率算法[J]. 光学学报, 2017, 37(3): 94-102.
- [4] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C] // Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 4681-4690.
- [5] SOH J W, PARK G Y, JO J, et al. Natural and realistic single image super-resolution with explicit natural manifold discrimination [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 8122-8131.
- [6] 李佳星, 赵勇先, 王京华. 基于深度学习的单幅图像超分辨率重建算法综述[J]. 自动化学报, 2021, 47(10): 2341-2363.
- [7] ZHANG K, ZUO W, CHEN Y, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising [J]. IEEE Transactions on Image Processing, 2017, 26(7): 3142-3155.
- [8] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1646-1654.
- [9] ANWAR S, BARNES N. Real image denoising with feature attention [C] // Proceedings of the IEEE International Conference on Computer Vision. Long Beach: IEEE, 2019: 3155-3164.
- [10] 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. Munich: ECCV, 2018: 286-301.
- [11] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution [C] // European Conference on Computer Vision. Amsterdam: ECCV, 2016: 694-711.
- [12] RAD M S, BOZORGTABAR B, MARTI U V, et al. SROBB: targeted perceptual loss for single image super-resolution [C] // Proceedings of the IEEE International Conference on Computer Vision. Seoul: IEEE, 2019: 2710-2719.
- [13] WANG X, YU K, WU S, et al. Esrgan: enhanced super-resolution generative adversarial networks [C] // Proceedings of the European Conference on Computer Vision Workshops. Munich: ECCV, 2018: 63-79.
- [14] MA C, RAO Y, CHENG Y, et al. Structure-preserving super resolution with gradient guidance [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 7766-7775.
- [15] JAIN V, SEUNG S. Natural image denoising with convolutional networks [J]. Advances in Neural Information Processing Systems, 2008, 21: 769-776.
- [16] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [C] // European Conference on Computer Vision. Zürich: ECCV, 2014: 184-199.
- [17] TAO X, GAO H, SHEN X, et al. Scale-recurrent network for deep image deblurring [C] // Proceedings of the

- IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8174-8182.
- [18] CHEN L, LU X, ZHANG J, et al. HINet: half instance normalization network for image restoration [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 182-192.
- [19] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 586-595.
- [20] KUPYN O, MARTYNIUK T, WU J, et al. DeblurGAN-v2: deblurring (orders-of-magnitude) faster and better [C] // Proceedings of the IEEE International Conference on Computer Vision. Seoul: IEEE, 2019: 8878-8887.
- [21] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. DeblurGAN: blind motion deblurring using conditional adversarial networks [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8183-8192.
- [22] SAJJADI M S M, SCHOLKOPF B, HIRSCH M. EnhanceNet: single image super-resolution through automated texture synthesis [C] // Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 4491-4500.
- [23] 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.
- [24] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a ‘completely blind’ image quality analyzer [J]. IEEE Signal Processing Letters, 2012, 20(3): 209-212.
- [25] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017: 136-144.

Edge perception based loss function for image restoration

WANG Yuhao^{*}, LI Meng^{**}, ZHI Tian^{**}, ZHANG Xishan^{**}, ZHOU Xuehai^{*}

(^{*} School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026)

(^{**} Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100086)

Abstract

The pixel-wise loss is the most basic and commonly used loss function for image restoration. However, pixel-wise loss cannot distinguish whether the edge in the image is blurred or not, and treats all pixels of equal importance, which is different from the perception of the edge part by the human visual system. It can be observed that when the pixel in the output image is blurred, the direction of the error between it and the corresponding pixel in the ground truth is the same as the direction of the second-order gradient of this position in the ground truth. Inspired by this, a novel structure preserving loss function (SPLoss) is proposed for image restoration. A blurred factor map is designed to detect blurred pixels, and then penalize blurred pixels by increasing the weight of the error. SPSLoss detects blurred pixels in restored images and penalizes them while optimizing networks. Experimental results show that the proposed method can restore images while preserving fine image structures on popular image restoration networks including image super-resolution and image deblurring.

Key words: image restoration, image super-resolution, image deblurring, neural network (NN), loss function