

用于单图像超分辨率的增强深度残差网络

潘林丰

摘要

近年来,随着深度卷积神经网络的发展,超分辨率的研究也取得了进展了。特别是残差学习表现出很好的性能,本复现论文开发了一种增强深度超分辨率网络(EDSR)^[1],性能超越了当前现有的超分辨率方法。EDSR之所以有显著的性能提升,是因为去除了传统残差网络中不必要的模块。另一个原因就是在训练过程保持稳定的情况下,扩展了模型的规模。另外,本复现论文该提出了多尺度深度超分辨率模型(MDSR),可以重建不同尺度的高分辨率图像。论文提出的方法在基准数据集上表现出优于现有方法的性能。

关键词: 超分辨率; 残差网络

1 引言

近年来,图像超分辨率问题受到越来越多的关注。超分辨率的目标是将低分辨率图像重建成高分辨率图像。随着深度卷积神经网络的发展,最近对超分辨率的研究取得了进展。特别是残差学习应用在超分辨率上,其中 SRResNet^[2]表现出很好的性能,但是 SRResNet 没有对 ResNet 架构^[3]进行修改。然而 ResNet 最初提出用于计算机视觉的 high level task,例如图像分类和检测。因此,将 ResNet 直接应用在超分辨率等 low level task 并不是最好的。所以,论文基于 SRResNet,通过分析和删除一些不必要的模块优化网络结构,得到 EDSR 模型。实验证明,改进后的 EDSR 模型有更好的性能。此外,因为多个尺度的超分辨率是相互关联的任务,论文还设计了一个多尺度融合的架构,称为 MDSR。从参数上看,与多个不同尺度的 EDSR 对比,MDSR 的参数更少,但有着和 EDSR 相当的性能。在标准基准数据集和 DIV2K 数据集上评估模型,提出的 EDSR 和 MDSR 在 PSNR 和 SSIM 有着优于现有方法的性能。

2 相关工作

为了解决超分辨率问题,早期的方法使用基于采样理论的插值技术。然而,这些方法在预测详细、逼真的纹理细节方法有着局限性。以前的研究对问题采用自然图像统计来重建更好的高分辨率图像。最近,深度神经网络的强大能力促进了超分辨率研究的发展。Dong 等人^[4]首先提出了一种基于深度学习的 SR 方法,针对 SR 研究了各种卷积神经网络。Kim 等人首先引入了残差网络来训练更深层的网络,称为 VDSR^[5],取得了优异的性能。他们特别指出,残差连接减轻了在超分辨率网络中携带身份信息负担,加速收敛。Mao 等人^[6]利用编码器-解码器网络和对称的残差连接解决了一般的图像恢复问题,并且加速了收敛。在许多基于深度学习的超分辨率算法中,输入图像在输入网络之前通过 bicubic 插值进行上采样。可以在网络末端训练上采样模块来代替使用插值图像作为输入的方式,这样做就可以在不减小模型容量的情况下减少大量计算。然而,这些方法有一个缺点:不能像 VDSR 在单一框架下处理多尺度问题。本论文中,解决了多尺度训练和计算效率的问题。不仅利用学习到的特征在每个尺度上的相互关系,而且提出了一个新的多尺度模型,可以重建不同尺度下的高分辨率图像。

此外，L2 损失函数是图像恢复最广泛使用的损失函数，而该论文使用 L1 损失函数来训练网络，故在复现工作中选择了使用 L2 损失函数来训练，并且与 L1 损失进行性能对比。

3 本文方法

3.1 残差块

EDSR 删除了残差块的 batch normalization 层，因为 BN 层对特征进行了归一化处理，破坏了图像的对比度信息。使输出图像的质量变差。实验证明，这个修改大大提高了性能。此外，GPU 内存的使用也得到充分减少，因为 BN 层消耗的内存量与卷积层相同，与 SRResNet 相比，节省了接近一半的内存。并且 EDSR 的残差块外面没有 ReLU 激活函数。残差块结构如图 1所示：

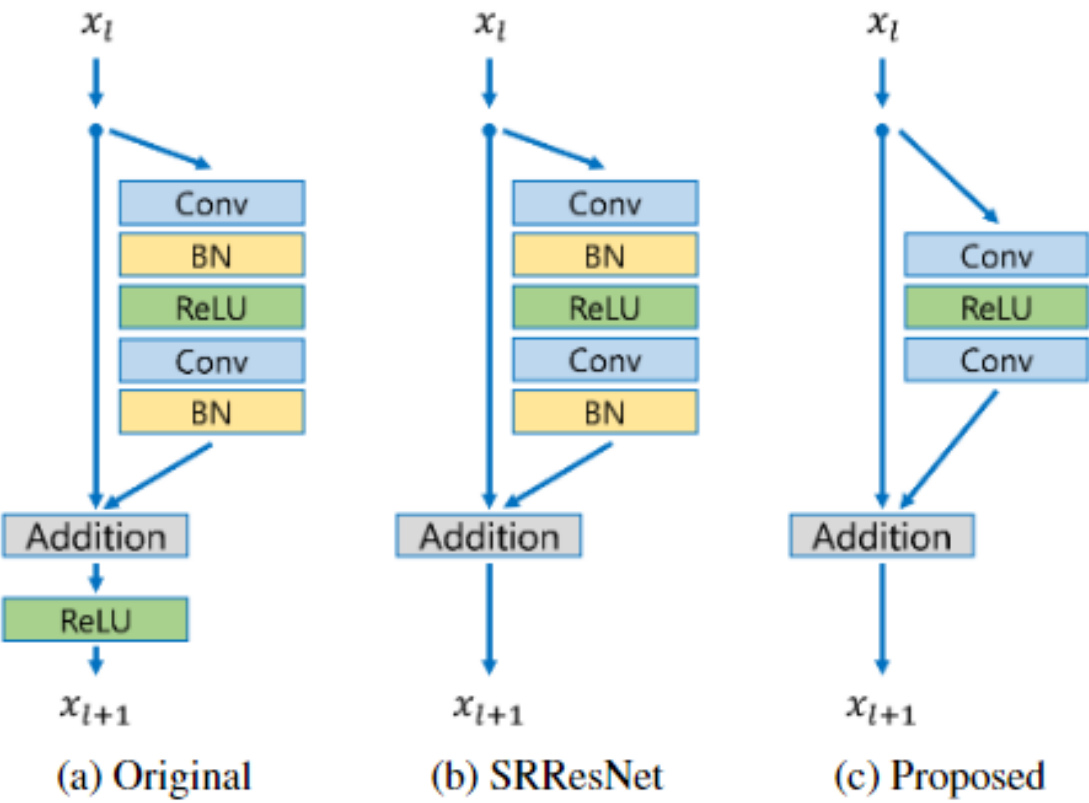


图 1: 残差块

3.2 EDSR

EDSR 结构与 SRResNet 相似，由卷积层、残差块、上采样模块组成，在残差块外面没有 ReLU 激活函数。EDSR 网络结构如图 2所示：

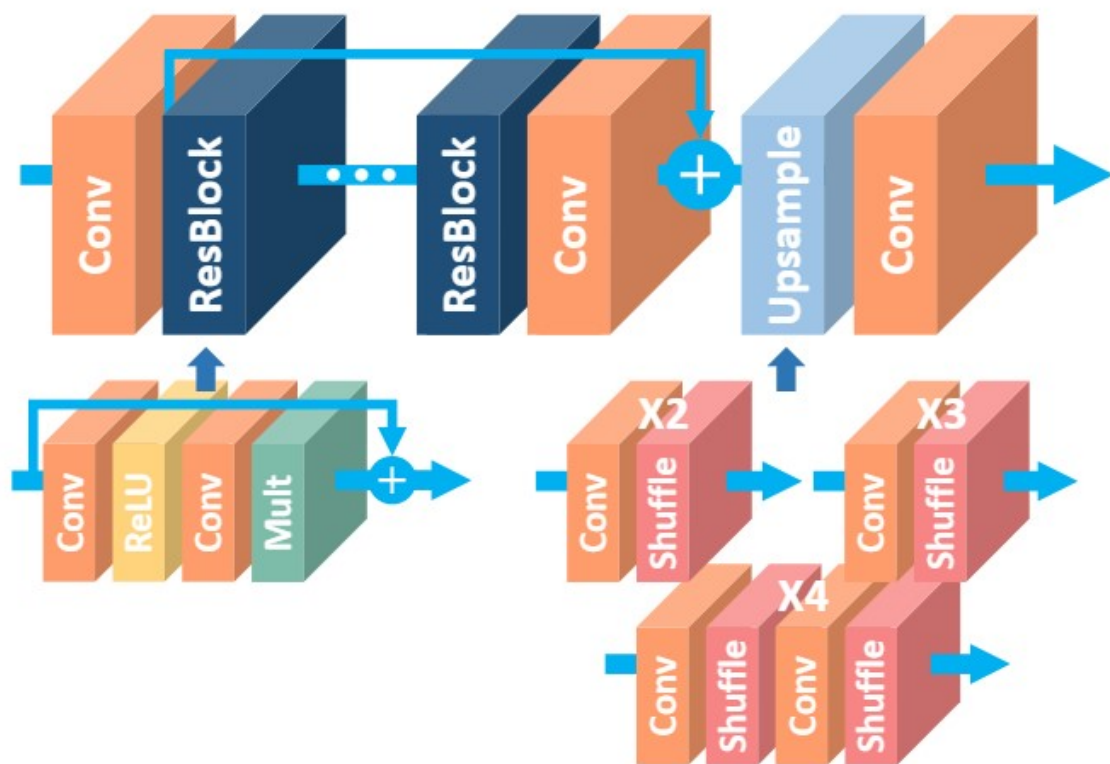


图 2: EDSR 网络结构

3.3 MDSR

MDSR 为不同尺度设置不同的残差块，以在多个尺度下处理超分辨率。在 MDSR 的尾部，平行放置不同尺度上的上采样模块进行多尺度重建。从参数上看，3 个不同尺度的 EDSR 一共 4.5M 个参数，而 MDSR 只有 3.2M 个参数，参数更少，而 MDSR 与多个不同尺度的 EDSR 的性能相当。MDSR 网络结构如图 3所示：

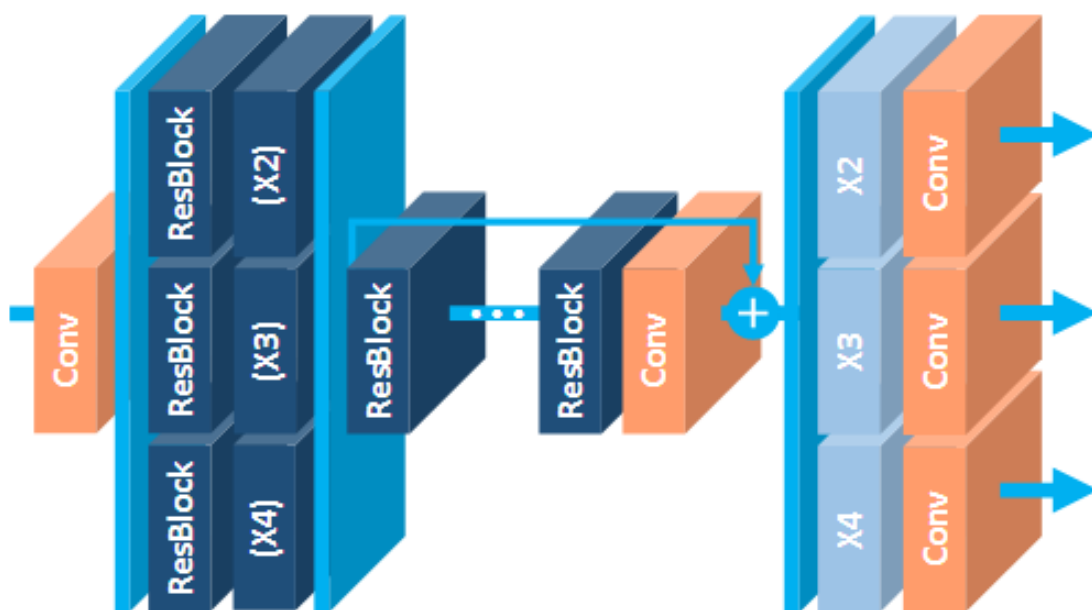


图 3: MDSR 网络结构

4 复现细节

4.1 与已有开源代码对比

论文使用 L1 损失函数训练模型，而 L2 损失函数是图像超分辨率问题最广泛使用的损失函数，并且最小化 L2 损失的同时可以最大化 PSNR，PSNR 是超分辨率问题的主要性能指标。论文发现 L1 损失能加速收敛，但是超分辨率图像的视觉效果不如使用 L2 损失。为了得到视觉效果更好的超分辨率图像，本复现工作使用 L2 损失函数来训练模型。

4.2 数据集

本复现工作使用的数据集的 DIV2K 数据集，这是用于图像恢复任务的高质量（2K 分辨率）图像数据集。DIV2K 数据集由 800 张训练图像、100 张验证图像和 100 张测试图像组成。由于测试数据集的 ground truth 没有发布，所以比较验证数据集的性能。

4.3 预处理

使用来自低分辨率图像的大小为 48x48 的 RGB 输入补丁，以及对应的高分辨率补丁。并且使用了随机水平翻转和 90 次旋转来扩充实验数据。通过减去 DIV2K 数据集的平均 RGB 值来预处理所有图像。

5 实验结果分析

本部分对实验所得结果进行分析，详细对实验内容进行说明，实验结果进行描述并分析。

5.1 定量分析

下图分别为使用 L1 loss 和 L2 loss 训练模型得到的 PSNR 图, 从图中可以看出，使用 L1 loss 和 L2 loss 训练模型的 PSNR 性能接近，在训练最后时 PSNR 达到最好，均有 34dB。

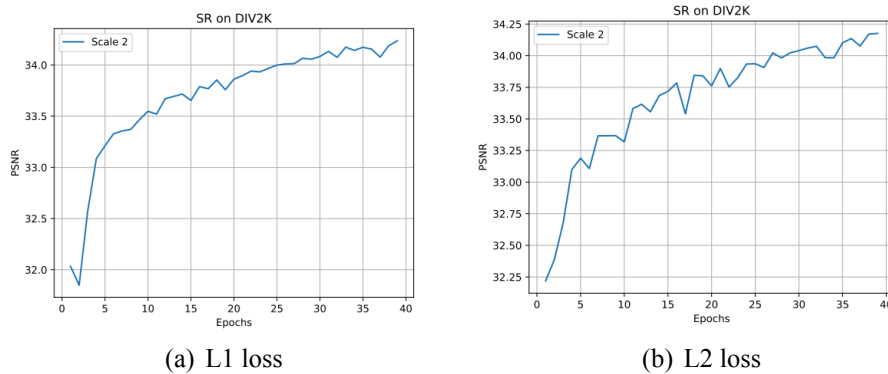


图 4: 使用 L1 loss 和 L2 loss 训练模型得到的 PSNR

5.2 定性分析

通过对使用 L1 loss 和 L2 loss 训练模型得到的超分辨率图像进行性能对比，从下图可以看出，L2 loss 训练模型得到的超分辨率图像在纹理细节上更加清晰，得到了更好的视觉效果。



图 5: 使用 L1 loss 和 L2 loss 训练的超分辨率图像

6 总结与展望

论文通过删除 ResNet 中的 batch normalization 层, 得到了 EDSR 模型, 提高了性能。此外, 还开发了一个多尺度超分辨率网络来减少模型大小和训练时间, 可以有效地处理不同尺度的超分辨率, 与多个不同尺度的 EDSR 对比, 性能相当, 并且参数更少。本复现工作使用 L2 损失函数训练 EDSR 模型, 与论文使用的 L1 损失函数对比, 图像质量指标 PSNR 性能接近, 而在感知上看, L2 损失函数在重建图像纹理细节上做得更好, 得到的超分辨率图像更加清晰, 有更好的视觉效果。

参考文献

- [1] 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.
- [2] LEDIG C, THEIS L, HUSZÁR 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.
- [3] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [4] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [C] // Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13. 2014: 184-199.
- [5] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1646-1654.
- [6] MAO X, SHEN C, YANG Y B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections[J]. Advances in neural information processing systems, 2016, 29.