

基于GAN的超分辨率重建算法研究

吴熹鹏^{1*}, 黄溪¹, 王先兵^{1#}, 何涛², 吴中鼎³, 柴婉秋³

¹武汉大学国家网络安全学院, 湖北 武汉

²中国科学院计算技术研究所, 北京

³贵阳铝镁设计研究院有限公司, 贵州 贵阳

Email: xipeng1994@gmail.com, #xbwang@whu.edu.cn

收稿日期: 2020年10月8日; 录用日期: 2020年10月23日; 发布日期: 2020年10月30日

摘要

图像超分辨率重建是图像复原的重要分支之一, 算法任务是进行图像或者视频的修复或重建, 目标是提高原始图像或视频帧的分辨率, 目前很多超分辨率重建算法使用生成对抗模型来提升分辨率。本文拟采用密集型卷积网络对生成对抗模型的生成器部分进行改进, 并提供一个将高层次特征与低层次特征联合起来的方法, 将每一层的特征图传播到后续卷积层, 从而在模型不同卷积层之间建立一条捷径, 使得梯度信息在层次较深的网络模块中实现回传。然后在此基础上采用生成对抗网络生成高频信息, 用来补充图像或视频帧细节, 获得更加真实的高分辨率图像, 达到提升重建结果的目标。

关键词

图像超分辨率, 生成对抗模型, 密集卷积网络, 深度学习

A Study of Super Resolution Algorithm Basing on Generative Adversarial Network

Xipeng Wu^{1*}, Xi Huang¹, Xianbing Wang^{1#}, Tao He², Zhongding Wu³, Wanqiu Chai³

¹School of Cyber Science and Engineering, Wuhan University, Wuhan Hubei

²Institute of Computing Technology, Chinese Academy of Sciences, Beijing

³Guiyang Aluminum Magnesium Design and Research Institute Co., Ltd., Guiyang Guizhou

Email: xipeng1994@gmail.com, #xbwang@whu.edu.cn

Received: Oct. 8th, 2020; accepted: Oct. 23rd, 2020; published: Oct. 30th, 2020

Abstract

Image super-resolution is one of the most important branches of image restoration, which is used

*第一作者。

#通讯作者。

文章引用: 吴熹鹏, 黄溪, 王先兵, 何涛, 吴中鼎, 柴婉秋. 基于 GAN 的超分辨率重建算法研究[J]. 计算机科学与应用, 2020, 10(10): 1908-1920. DOI: 10.12677/csa.2020.1010201

for restoring or reconstructing images or videos, thus enhancing the resolution of the original images or frames of video. At present, most image super-resolution algorithms are using the generative adversarial network. In this paper, DenseNet is used to improve the generator part of the GAN model, and a method of combining high-level features and low-level features is provided to spread the feature map of each layer to the subsequent convolution layer, so as to establish a shortcut between different convolution layers of the model, making gradient information in the deeper network module propagate back. On this basis, the high-frequency information is generated by the GAN, which is used to supplement the image or video frame details, and obtain more real high-resolution images, thus improving the reconstruction results.

Keywords

Super-Resolution, Generative Adversarial Network, DenseNet, Deep Learning

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像超分辨率(Super-resolution, SR)重建是图像复原的重要分支之一,通过算法对图像或视频帧进行修复,提升原始图像或视频帧的分辨率。目前,通过对单张低分辨率图像进行重建生成单张高分辨率图像的方法大致可以分为三类:基于插值的方法、基于重建的方法、基于学习的方法。

基于插值[1]的方法。插值是比较容易想到也是最先用来进行图像重建的方法。该方法将输入图像信息作为算法的先验知识,算法可以得到较满意的重建效果。不过,因简单插值方法有限的表达能力,很容易在分辨率提升过程中产生振荡,并可能会产生锯齿状纹理,导致最终的重建图像比较模糊。

基于重建[2]的方法。该方法则采取下采样,并强制约束平滑,使得原图像和重建图像一致。该方法有一个明显的缺陷,当放大倍数过大时,很多重要的高频细节可能会丢失,导致重建图像质量大幅度下降。

基于学习[3]的方法。该方法是目前的研究热点,通过深度学习网络来解决上述两种方法产生的问题。该方法对训练样本中的低分辨率图像和相对的高分辨率图像进行字典学习[4],得到图像之间统计学意义上的关联,然后利用关联来提升低分辨率图像的分辨率,最终得到高清重建图像。

SRGAN [5]将生成对抗网络(GAN)用于超分辨率 SR 重建问题,采用生成器重建低频图像内容,通过 GANs 补全图像高频细节内容,从而解决传统 MSE 损失函数缺少高频信息难题。虽然 SRGAN 的重建图像质量超过其他深度学习方法,但 SRGAN 更侧重于感知相似性,即从图像风格上相似,而不是像素相似性,在细节上可能会出现某些不一致的情况。

本文基于密集网络和残差网络相结合的密集残差单元来改进 SRGAN 中的生成器结构,可以在使用更深的网络层次进行重建图像低频信息时,不会导致梯度消失,在此基础上优化目标函数,从而提高训练效率。训练好的模型在 Set5、Set14、Urban100、BSD100 等公开数据集中进行了测试,实验结果表明本文的改进方法有效,具有较明显提升。

2. 相关理论

2.1. 生成对抗网络 GAN

GAN 是一种生成式神经网络[6],它把训练过程比作两个参与者竞争的过程。其中,生成器通过学习

与训练数据生成新样本；新样本则由判别器来判断是真实样本还是假样本。在此基础上不断进行改进和完善。

2.2. 残差网络 ResNet

残差网络(ResNet) [7]的提出是为了解决深度神经网络的梯度弥散问题。残差块通过在网络中加入短路连接实现梯度回传,如图1所示,假定想要的隐层输出为 $H(x)$,通过加入一个短路连接 identity mapping,将经过卷积层的输出 $F(x)$,与输入 x 进行求和,即得到输出 $H(x) = F(x) + x$ 。

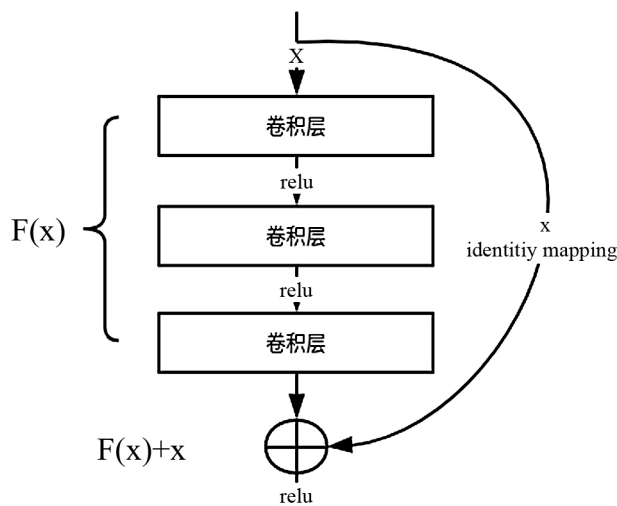


Figure 1. ResNet structure
图 1. 残差结构

加入 identity mapping 之后,即使在深层网络,梯度也可以通过 identity mapping 回传,不像以前的网络通过单纯的权值连乘进行梯度回传,从而避免反向传播导致的梯度弥散。也可以理解为残差网络在做反向传播时通过短路缩短了所经过的网络层数。

2.3. SRGAN

SRGAN 将 ResNet 应用于超分辨率重建,生成器 SRResNet 中每个残差块中都由两个 3×3 大小的卷积核以及 64 个卷积核的卷积层组成,每个卷积层之后接一个批规范化层,并使用 PReLU 作为激活函数。

SRGAN 提出了感知损失用于解决重建结果感知质量不佳问题,采用基于预训练好的 VGG19 网络 VGG loss。由于 SRGAN 使用 GAN 进行高分辨率重建,添加了一个判别器,因此不仅需要考虑内容损失,还需考虑对抗损失。

2.4. 密集卷积网络 DenseNet

如图 2 所示,密集卷积网络[8]是一种更加简便的连接模式,通过连接具有相同大小的特征图,确保网络各层间信息流通。每一层卷积层都会接受之前所有卷积层的输出,同时会把自身层的输出传递给之后所有卷积层。与残差网络通过相加将特征整合不同,密集卷积网络则是通过堆叠相同大小的特征图。

由于密集网络要堆叠整个网络的特征图,而且要保存卷积层输出作为后续输入,资源消耗巨大。因此密集思想仅在小个块(DenseBlock)中使用,并在网络中减少特征图数量。

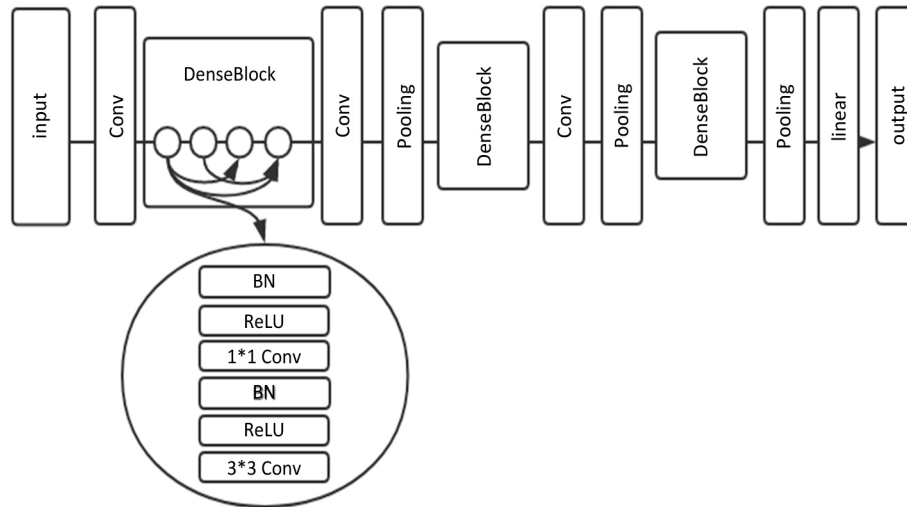


Figure 2. DenseNet structure
图 2. 密集卷积网络结构

2.5. 残差密集块

残差密集块[9] (Residual dense block, RDB)吸收 DenseNet 的优点, 包括密集连接层、局部特征融合(LFF)和局部残差学习(LRL)。残差密集块能确保每一层卷积层能提取更多的特征图, 并通过特征融合在全局范围保留分层特征, 从而产生隐式深层监督。本文使用 RDB 替代残差网络来改进 SRGAN 的生成器, 使用密集残差块来提取特征, 与 DenseNet 相比可使用更多卷积核。

3. 训练框架

3.1. 生成器架构

本文采用 RDB 改进 SRGAN 的生成器架构, 主要是利用 DenseNet 的局部密集连接, 通过堆叠实现充分利用不同卷积层级之间的特征。改进的残差密集网络(RDN)生成器通过局部特征融合 LFF 来调整特征图数量, 通过局部残差学习 LRL 整合上层 RDB 与本层 RDB 的信息。最后加入全局特征融合(GFF), 整合所有 RDB 提取的特征。

生成器网络(如图 3 所示)包括四个模块, 1) RDB 浅层特征提取的模块; 2) 将 RDB 堆叠起来的分层特征提取模块; 3) GFF 全局特征整合模块; 4) 亚元素卷积模块, 通过图像放大得到最终重建图像。

本文将网络输入记作 I_{LR} , 输出记作 I_{SR} 。首先采用一个 3×3 卷积层完成第一次浅层特征提取, 卷积核数量为 64 个, 浅层特征提取记作:

$$F_{-1} = \text{Conv}(I_{LR}) \quad (1)$$

F_{-1} 不仅会用于下一步的分层特征提取, 还将用于全局残差学习。第二次的浅层特征提取与第一层相同, 也是 3×3 卷积层, 64 个卷积核。第二层浅层特征提取的输出 F_0 将作为后续一系列 RDB 结构的输入。

$$F_0 = \text{Conv}(I_{LR}) \quad (2)$$

假设生成器网络一共采用 K 个 RDB, 则第 d 个 RDB 的输出 F_d 表示为:

$$F_d = H_{RDB,d}(F_{d-1}) = H_{RDB,d}\left(H_{RDB,d-1}\left(\cdots\left(H_{RDB,1}(F_0)\right)\cdots\right)\right) \quad (3)$$

其中 $H_{RDB,d}$ 表示第 d 个 RDB。

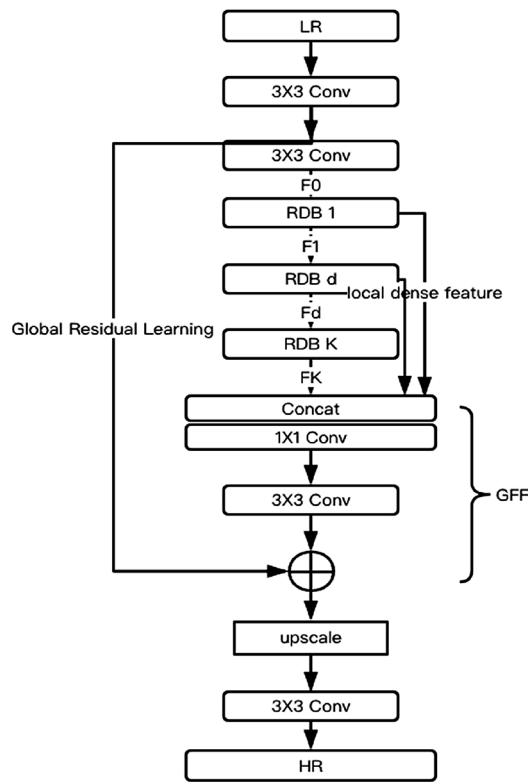


Figure 3. Generator model

图 3. 生成器模型

生成器网络一共使用了 12 个 RDB 进行堆叠。每个 RDB 结构如图 4 所示，由 6 个卷积层组成，每层包含 64 个 3×3 卷积核，采用 PReLU 作为激活函数。每层卷积层的输入来自于前面所有卷积层的输出堆叠形成，最后堆叠所有层的输出，并通过一个由 64 个大小为 1×1 的卷积核组成的卷积层调整特征图数量，并与上一个 RDB 的输出一起进行局部残差学习，作为这个 RDB 的输出。

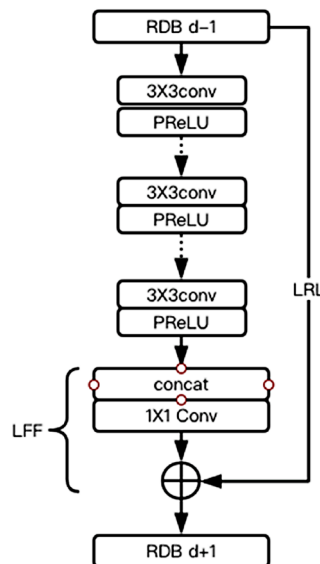


Figure 4. Residual dense block structure

图 4. 残差密集块结构

生成器网络然后将 12 个 RDB 的输出堆叠在了一起,即将不同层的局部特征进行全局特征整合 GFF。GFF 由一个堆叠层和一个 64 个 1×1 卷积核组成的卷积层组成,紧接着采用一个 3×3 卷积层进一步提取特征,最后再与最开始产生的 F_{-1} 进行全局残差学习。

生成器网络接下来用一个卷积层将形状提升到 $H \cdot W \cdot C \cdot r^2$,然后用亚元素卷积层将放大图像到 $rH \cdot rW \cdot C$,最后采用一个 3×3 卷积层把图像压缩到 RGB3 通道,输出高分辨率图像。

3.2. 判别器架构

判别器用于判别图像样本的真伪,判别器模型如图 5 所示。

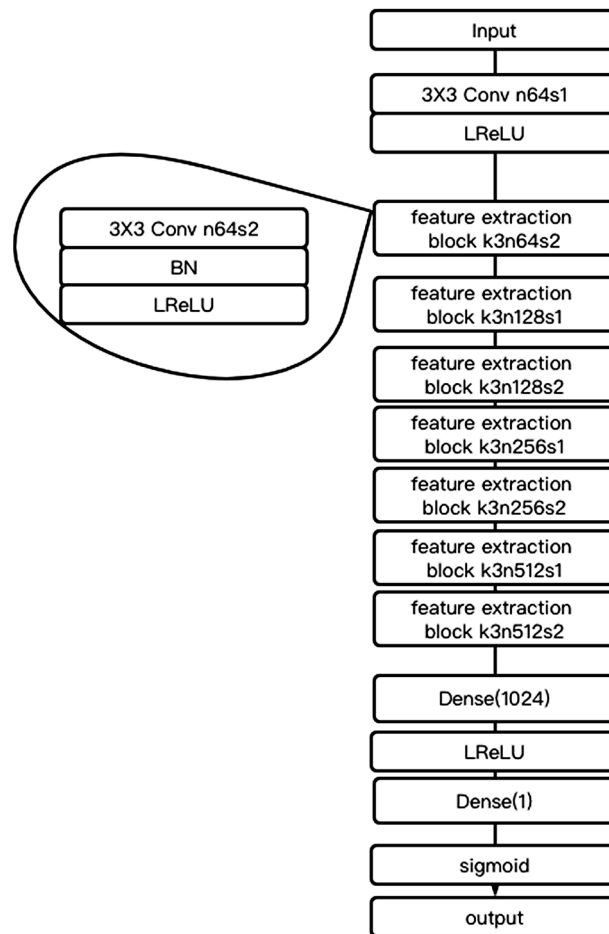


Figure 5. Discriminator model

图 5. 判别器模型

判别器网络首先使用一个 64 个 3×3 卷积核组成的卷积层提取特征,并使用 Leaky ReLU 作为激活函数,紧接着是一连串特征提取模块。每一个特征提取模块都由一层 3×3 卷积层,紧跟一层批规范化层 (Batch-normalization, BN) 层,后面接 LReLU 作为激活函数。第一个模块卷积层有 64 个卷积核,步长为 2。后面两个模块卷积层的卷积核数量均为 128,步长分别为 1 与 2。再后面两个模块卷积层的卷积核数量为 256,步长一样分别为 1 与 2。最后两个模块卷积层的卷积核数量为 512,步长也分别为 1 与 2。然后通过一个 1024 个神经元隐层平铺特征矩阵,再通过一个隐层得到最终输出,使用 sigmoid 激活函数获得真实图像概率,即可给出判别结果。

3.3. 模型训练

模型训练流程如图 6 所示。模型训练使用 ImageNet2014 [10]数据集，随机选取 5 万张高清图像，通过使用 bicubic 插值进行缩小 4 倍压缩，得到对应的低分辨率图像，然后在图像中随机截取一个 32×32 大小区域作为训练样本，固定训练样本大小的原因是因为判别器有全连接层。为了增加数据的多样性，对每张图像随机决定是否进行上下或者水平翻转，从而防止模型过拟合。

验证集则使用 DIV2K 高质量数据集[11]，该数据集有 800 张 2 K 分辨率图像，也通过 bicubic 插值等操作得到对应的低分辨率图像。训练过程中每训练 1000 次就使用验证集对模型效果进行评估，并将效果最好的参数保存下来。由于 GAN 在刚开始训练阶段，容易因达不到纳什平衡而得到无法预料的结果，因此在训练刚开始阶段需要随时观察实际重建结果，如果不正常，需要调参重新开始训练。

训练过程中先针对生成器单独训练，然后针对生成器使用 GAN，将损失函数中内容损失设置为 MSE，并将对抗损失系数设置为 10^{-3} ，再进行训练。最后再将损失函数替换为感知损失，并将 MSE 损失也加入到内容损失中，再进行训练，得到最终模型。

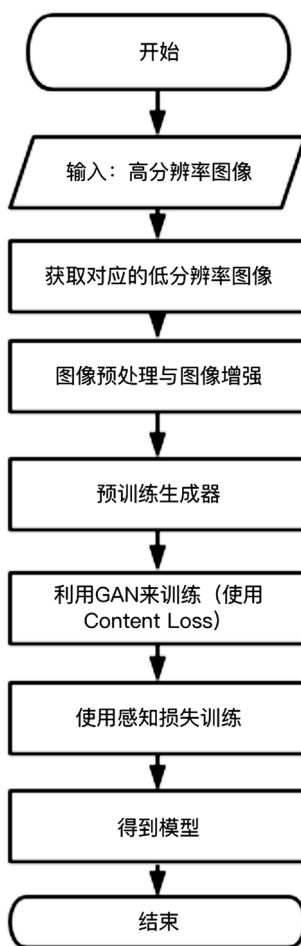


Figure 6. Training process

图 6. 训练流程

模型基于开源深度学习框架 Keras2 实现，深度学习框架采用 tensorflow，python 的版本为 3.4，操作系统为 Ubuntu16.04，显卡 GTX1060 (6 G)。训练优化方法采用 Adam [12]，初始学习率设置为 10^{-4} ，每

训练 1000 轮将其缩减一半，训练耗时一天。

模型最终使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(SSIM)对重建结果进行评估，但需要先将图像从 RGB 空间转换到 YCbCr 空间中，在 Y 通道上计算 SSIM 和 PSNR 值。

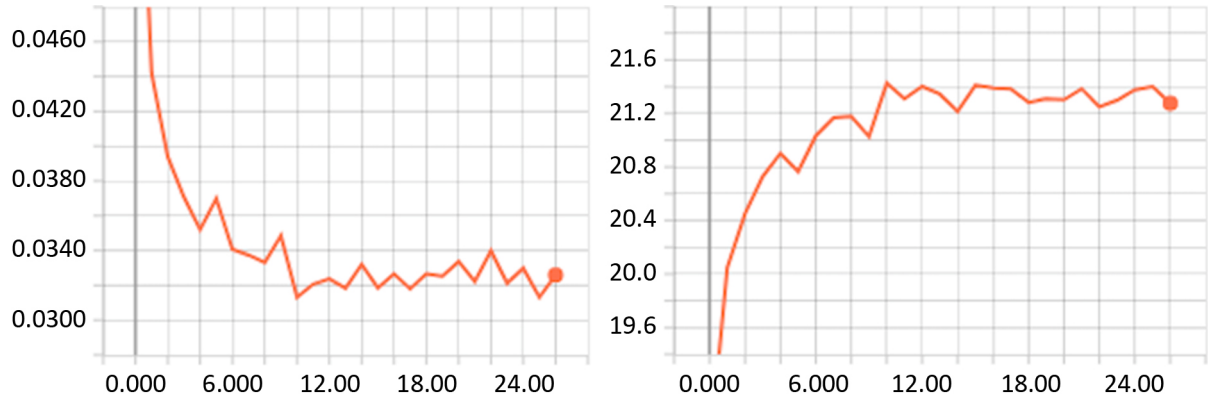


Figure 7. Pre-training generator left: loss right: PSNR

图 7. 预训练生成器 左图为 loss 右图为 PSNR

图 7 左边图像是预训练生成器过程中 loss 的趋势图，图中可以看出预训练过程中 loss 震荡较严重，主要原因是 GPU 显存只有 6G 太小，每次样本数量不够多，导致 loss 发生比较明显的震荡，但总体仍呈下降趋势。横坐标代表训练轮数，每一轮包含 1000 次迭代。从图中可以看出前十几轮迭代 loss 下降非常明显，后面的训练，loss 下降非常有限，而且产生剧烈震荡。

通过调整学习率模型可以进一步收敛，并期望最终收敛到局部最优值。为实现这一目标，我们采用 Adam 作为模型训练优化器，根据训练迭代数通过自适应的方式来调整学习率，其优化更新方法为：

$$m = \beta_1 \cdot m + (1 - \beta_1) \cdot dx \tag{4}$$

$$mt = \frac{m}{(1 - \beta_1^t)} \tag{5}$$

$$v = \beta_2 \cdot v + (1 - \beta_2) \cdot dx^2 \tag{6}$$

$$vt = \frac{v}{(1 - \beta_2^t)} \tag{7}$$

$$x+ = - \frac{lr \cdot mt}{np.sqrt(vt) + eps} \tag{8}$$

其中 lr 为学习率， eps 是一个超参数用来防止分母为 0， β_1 和 β_2 为优化器中的超参数，其中 $\beta_1 = 0.9$ ， $\beta_2 = 0.999$ ， $eps = 1e-8$ ， t 为迭代轮数。

PSNR 表示信号最大可能功率和影响其表示精度的破坏性噪声功率的比值。PSNR 的值越大越好，会随着训练时间越变越大。从图 7 右边可以看出，PSNR 图像的走向趋势与 MSE 走向相反，在前十几轮的迭代中迅速增加，后面有震荡趋于稳定。而且，PSNR 值不高，最后稳定在 21.3 左右，究其原因应该是 ImageNet 样本图像分辨率不高，压缩后高频细节丢失严重。

图 8 是验证集训练结果，十几轮后验证集损失降低到一个较低水平，说明模型未产生过拟合现象。验证集使用 DIV2K 数据集，样本图像分辨率比 ImageNet 高，PSNR 值提升较明显，因此在训练过程中 PSNR 值不高与样本图像原始分辨率较低存在关联关系，很明显过于模糊的图像会导致超分辨率重建变得困难。

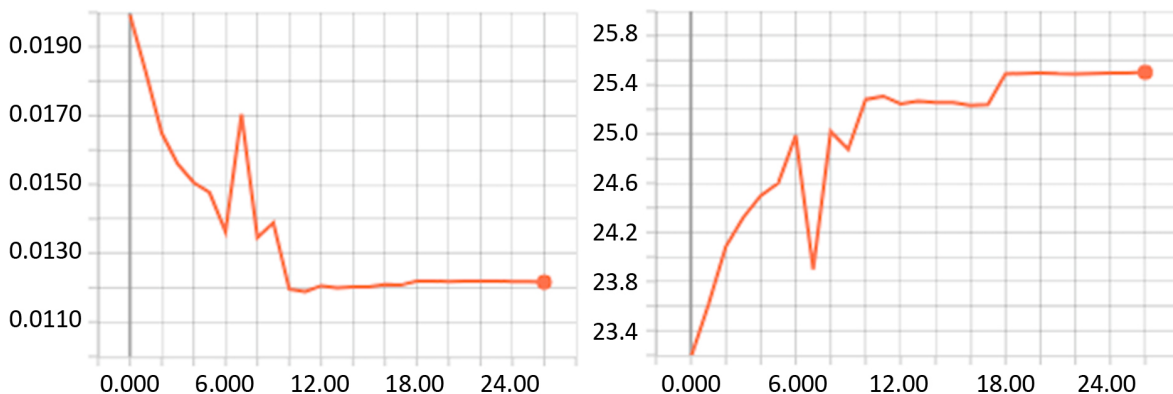


Figure 8. Pre-training validator left:loss right:PSNR

图 8. 预训练验证器 左图为 loss 右图为 PSNR

图 9 显示的是 GANs 训练过程中对抗损失与内容损失走向。由于先进行了生成预训练，内容损失刚开始下降，后面一直震荡。由于判别器刚开始还未训练，依靠随机猜测输出类别，对抗损失较低。后面随着训练的进行，对抗损失越来越高，后面随着生成器生成的样本越来越接近真实样本，对抗损失再度降低，后面震荡减小，模型逐渐收敛到局部最优。

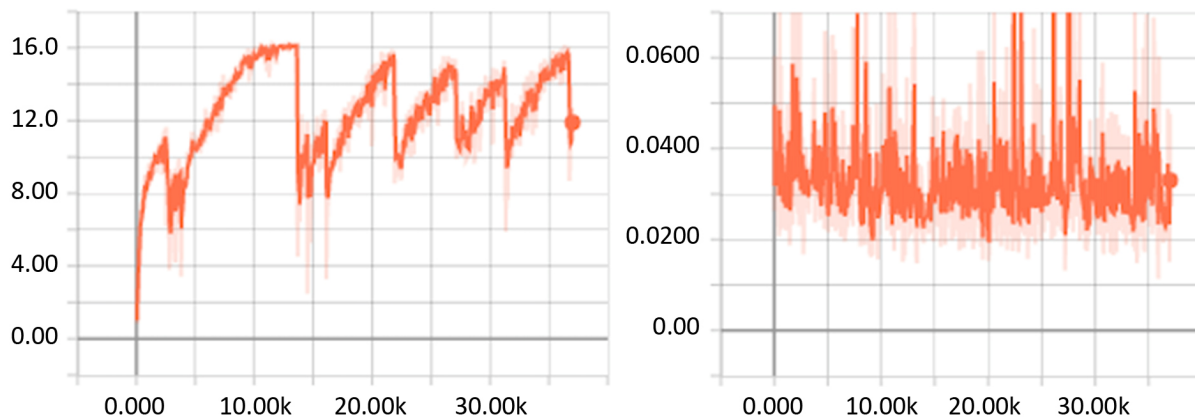


Figure9. GANs training process left:discriminator loss right:content loss

图 9. 生成对抗网络训练过程,左图为对抗损失, 右图为内容损失

这个阶段 PSNR 震荡非常明显，应该是模型生成的高频细节并不是原始图像的真实细节，比如会多出一些假的纹理和细节。从而导致重建图像与真实图像间的 MSE 相差较大，MSE 是 PSNR 值计算的衡量标准，最终导致明显的震荡。

4. 实验过程

4.1. 实验数据集

实验数据集选择 Set5 [13], Set14 [14], BSD100 [15], Urban100 [16]这些经常被图像超分辨率重建用于作为结果对比的公开数据集。四个数据集分别包含 5 张、14 张、100 张、100 张高分辨率图片及对应的低分辨率图片，图片格式均为 png。实验采用 python 中的 matplotlib 包将这些图像读到内存，将图像像素为[0,255]的值域进行转换，低分辨率图像像素值域映射到[0,1]，高分辨率图像像素值域映射到[-1,1]。

4.2. 实验设计

本文的对比实验包括：改进生成器之后的 SRGAN 与原文使用 SRResNe 生成器的 SRGAN 对比；与其它公开结果的模型进行对比实验，包括与最近邻插值法、双立方插值法[17]、SRCNN [18]、SCSR [19]等进行对比实验；进行转置卷积与亚元素卷积的对比实验，分析两者对重建结果与训练速度的影响。

实验使用 keras 复现 SRGAN 的 SRResNet 生成器网络，设置的参数与 RDB 训练的网络一致，使用 ImageNet 数据集进行训练。

实验使用的最近邻插值法和双立方插值法属于传统插值超分辨率重建算法，优点体现在速度快，缺点也非常明显，因为丢失高频细节导致重建结果会出现模糊的情况。实验使用的 SRCNN 是最早几种采用卷积网络进行超分辨率重建的模型。模型使用的卷积层数较少，仅堆叠了三层卷积层，模型首先通过双立方插值进行图像放大，然后采用卷积网络进行特征提取和筛选，最后一层卷积输出作为重建结果。

本文为了验证模块数量对重建结果的影响，以及残差密集块对梯度弥散的减轻程度，采用堆叠不同数量 RDB 模块进行实验。因为显卡显存容量大小限制，本文分别将 RDB 堆叠数量设置为 8, 10, 12 进行对比实验，每个 RDB 模块采用 6 个卷积核数量为 32 的卷积层进行特征提取。模型使用的其它卷积层，包括最开始的浅层特征提取、LFF、GFF，均采用 64 个卷积核。实验结果显示在堆叠较多的 RDB 模块后，仍能起到促进作用。

本文针对 DenseNet 改进主要是在 RDB 结构中加了更多 1×1 卷积，因此每层卷积层可以使用更多的卷积核进行特征提取。本文针对卷积核数量进行了对比实验，分别将 RDB 中卷积层的卷积核数量设为 16、32、64 进行对比实验，分析卷积核数量对模型的影响。

4.3. 实验结果

4.3.1. 堆叠 RDB 块数量对比实验

Table 1. The number of RDB's influence on the network

表 1. 堆叠 RDB 数量对网络的影响

	Set5	Set14	训练时间	参数数量
K = 8	29.437/0.848	26.809/0.816	161 ms	2,502,019
K = 10	29.536/0.883	26.855/0.816	193 ms	3,041,155
K = 12	29.717/0.895	27.005/0.821	223 ms	3,580,291

表 1 中的 K 表示堆叠 RDB 块数量，训练时间是模型训练完一个 batch 所消耗的时间。表 1 中第 2 列与第 3 列数据显示，模型增加堆叠 RDB 块，重建图像的 PSNR 值与 SSIM 值会增加，K = 12 的值比前两者的值都高，说明堆叠更多的 RDB 模块能够提升重建结果，原因是每堆叠一个 RDB，网络会增加 6 个卷积层，从而提升网络表示能力，最终提升重建结果，但提升的精度细微。从表 1 的第 4 列与第 5 列可以看到，随着堆叠的 RDB 块增加，训练时间大大增加，带来细微精度的提升；模型参数数量也随着堆叠的 RDB 块增加而提高了不少，如果 GPU 的显存不够大，参数数量过多会导致显存不够训练无法继续，堆叠的 RDB 块不能太多，在对重建精度要求不高的使用场景，可以堆叠较少 RDB 提升训练效率和重建速度。

4.3.2. 卷积核数量对比实验

Table 2. The number of kernel's influence on the network
表 2. 卷积核数量对网络表现的影响

	Set5	Set14	参数数量
G = 16	29.491/0.880	26.955/0.816	1,597,699
G = 32	29.717/0.895	27.005/0.821	3,580,291
G = 64	29.889/0.888	27.38/0.827	10,033,795

表 2 中 G 表示卷积层包含的卷积核数量，实验数据表明，增加卷积核数量可以提升模型能力，但提升不是很明显。与 RDB 堆叠数量相似，当特征图数量过多时参数数量增大，最终会导致显存不够无法继续训练。

4.3.3. 卷积方法对比实验

Table 3. Compare transposed convolution with sub-pixel convolution
表 3. 转置卷积与亚元素卷积对比

	Set5	Set14	训练时间
转置卷积	29.887/0.885	27.39/0.832	381 ms
亚元素卷积	29.889/0.888	27.38/0.827	319 ms

本文采用转置卷积与亚元素卷积两种不同的卷积方案进行了对比实验，实验结果如表 3 所示，两种不同的卷积方式的结果精度并没有太大的差别。但在训练时间上，亚元素卷积比转置卷积要快一些。

4.3.4. 不同生成器对比实验

Table 4. Compare different generators
表 4. 不同生成器对比

	Set5	Set14	Urban100	BSD100
SRResNet	23.303/0.635	21.024/0.535	19.381/0.567	20.983/0.545
RDB	26.245/0.814	24.717/0.756	22.753/0.729	24.348/0.732



Figure 10. left:SRGAN (SRResNet) right:SRGAN (RDB)
图 10. 左图为 SRGAN (SRResNet)右图为 SRGAN (RDB)

本文的核心工作是改进 SRGAN 的生成器，并对基于 RDB 改进生成器的 SRGAN 模型与原来使用 SRResNet 作为生成器的 SRGAN 模型在四个测试集上分别进行了超分辨率重建对比实验，表 4 显示的是实验结果，结果显示基于本文改进的 RDB 生成器模型，在每个数据集上，PSNR 与 SSIM 均有较明显的

提升, 见图 10。

实验结果显示将密集块与残差块结合起来, 的确可以提升 SRGAN 的重建结果。由于密集网络充分利用了不同层级之间的特征, 低频信息重建与高频细节重建都起到了帮助。虽然 SRGAN 主要依靠 GAN 补足高频细节, 但是生成器的改进对重建也带来了帮助。

4.3.5. 不同算法对比实验

本文最后将本文改进后的 SRGAN 模型与其它经典的超分辨率算法在各个公开测试集进行了对比实验, 结果如表 5 所示。

Table 5. Compare to other algorithms

表 5. 与其他算法结果对比

	Set5	Set14	Urban100	BSD100
最近邻插值	24.607/0.731	22.982/0.646	20.733/0.619	23.728/0.629
双立方线性插值	26.694/0.789	24.247/0.686	21.701/0.651	24.654/0.661
SCSR	27.237/0.806	24.631/0.706	22.510/0.695	25.258/0.694
SRCNN	28.089/0.829	25.161/0.720	22.626/0.696	25.345/0.698
SRGAN(RDB)	26.245/0.814	24.717/0.756	22.753/0.729	24.348/0.732

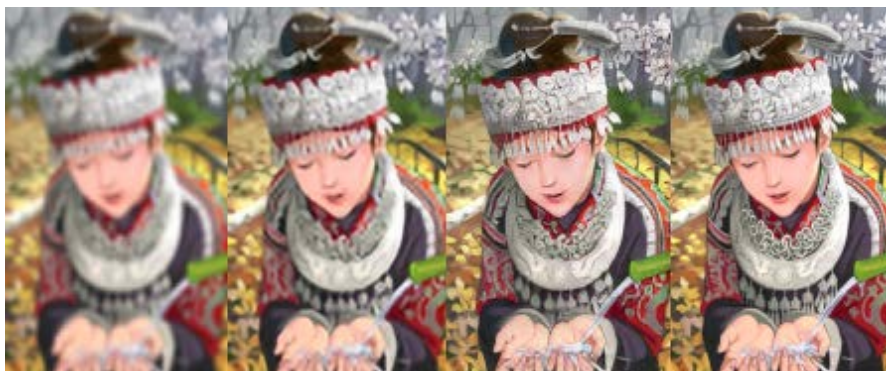


Figure 11. From left to right: LR SRCNN SRGAN(RDB) HR

图 11. 左一 LR 左二 SRCNN 右二 SRGAN(RDB)右一 HR

从表 5 的结果可以看出 GAN 模型的重构图像由于生成了许多假细节, 在 PSNR 指标上比不过一些方法, 但重构图像在真实度上有较大提升, 更符合人眼观察到的高清图像, 见图 11。

因此, 对于重构图像真实度要求不是特别高的场景, 可以采用 SRGAN 进行超分辨率重建, 比如复原哪些压缩严重的风景画或者照片, 但对于那些对图像真实要求很高的医疗、遥感等图像的重构就不太适合。

5. 结束语

本文采用密集型卷积网络对生成对抗模型的生成器进行改进, 提升了 SRGAN 模型进行超分辨率重建能力。虽然本文研究达到了预期目的, 但仍有很多改进的空间。因为堆叠 RDB 块造成网络深度过深, 虽然密集网络能减轻梯度弥散, 但训练和测试时间增长非常快。另外, 利用 GANs 重建图像的真实感会有较大提升, 但 PSNR 值却稍低, 说明 PSNR 可能不是特别适合用来作为指标来衡量重建效果。由于 GANs 的训练是基于两个模型的对抗竞争, 超参数设置比较敏感, 易出现难以预料的结果。如何提升 GANs 训练的鲁棒性, 尽量避免出现不收敛情况是未来继续研究的方向。

基金项目

黔科合重大专项字[2016] 3012。

参考文献

- [1] Zhou, F., Yang, W.M. and Liao, Q.M. (2012) Interpolation-Based Image Super-Resolution Using Multi-Surface Fitting. *IEEE Transactions on Image Processing*, **21**, 3312-3318. <https://doi.org/10.1109/TIP.2012.2189576>
- [2] Lin, Z.C. and Shum, H.Y. (2004) Fundamental Limits of Reconstruction Based Super Resolution Algorithms under Local Translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**, 83-97. <https://doi.org/10.1109/TPAMI.2004.1261081>
- [3] Pan, Z.-X., Yu, J., Hu, S.-X. and Sun, W.-D. (2014) Single Image Super Resolution Based on Multi-Scale Structural Self-Similarity. *Acta Automatica Sinica*, **40**, 594-603.
- [4] Lian, Q.-S., Shi, B.-S. and Chen, S.Z. (2015) Research Advances on Dictionary Learning Models, Algorithms and Applications. *Acta Automatica Sinica*, **41**, 240-260.
- [5] Ledig, C., Theis, L., Huszar, F., et al. (2016) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 4680-4690. <https://doi.org/10.1109/CVPR.2017.19>
- [6] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014) Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, Montreal, 8-13 December 2014, 2672-2680.
- [7] Dong, C., Loy, C.C., He, K., et al. (2014) Learning a Deep Convolutional Network for Image Super-Resolution. In: *European Conference on Computer Vision*, Springer, Cham, 184-199. https://doi.org/10.1007/978-3-319-10593-2_13
- [8] Huang, G., Liu, Z., Weinberger, K.Q., et al. (2017) Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, Vol. 1, 3. <https://doi.org/10.1109/CVPR.2017.243>
- [9] Zhang, Y., Tian, Y., Kong, Y., et al. (2018) Residual Dense Network for Image Super-Resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 2472-2481. <https://doi.org/10.1109/CVPR.2018.00262>
- [10] Deng, J., Dong, W., Socher, R., et al. (2009) Imagenet: A Large-Scale Hierarchical Image Database. *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 248-255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [11] Agustsson, E. and Timofte, R. (2017) Ntire 2017 Challenge on Single Image Super-Resolution: Dataset and Study. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, Honolulu, 21-26 July 2017, Vol. 3, 2. <https://doi.org/10.1109/CVPRW.2017.150>
- [12] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization.
- [13] Bevilacqua, M., Roumy, A., Guillemot, C., et al. (2012) Low-Complexity Single-Image Super-Resolution Based on Nonnegative Neighbor Embedding. *Proceedings British Machine Vision Conference 2012*, Surrey, 3-7 September 2012, 135.1-135.10. <https://doi.org/10.5244/C.26.135>
- [14] Zeyde, R., Elad, M. and Protter, M. (2010) On Single Image Scale-Up Using Sparse-Representations. In: *International Conference on Curves and Surfaces*, Springer, Berlin, Heidelberg, 711-730. https://doi.org/10.1007/978-3-642-27413-8_47
- [15] Martin, D., Fowlkes, C., Tal, D., et al. (2001) A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. *Eighth IEEE International Conference on Computer Vision*, Vol. 2, 416-423.
- [16] Huang, J.B., Singh, A. and Ahuja, N. (2015) Single Image Super-Resolution from Transformed Self-Exemplars. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 5197-5206. <https://doi.org/10.1109/CVPR.2015.7299156>
- [17] De Boor, C. (1962) Bicubic Spline Interpolation. *Journal of Mathematics & Physics*, **41**, 212-218. <https://doi.org/10.1002/sapm1962411212>
- [18] Dong, C., Chen, C.L., He, K., et al. (2016) Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **38**, 295-307. <https://doi.org/10.1109/TPAMI.2015.2439281>
- [19] Yang, J.C., Wright, J., Huang, T.S. and Ma, Y. (2010) Image Superresolution via Sparse Representation. *IEEE Transactions on Image Processing*, **19**, 2861-2873. <https://doi.org/10.1109/TIP.2010.2050625>