

文章编号: 1671-6833(2021)05-0001-06

## 基于无配对生成对抗网络的图像超分辨率重建

李学相, 曹 淇, 刘成明

(郑州大学 软件学院, 河南 郑州 450002)

**摘 要:** 针对基于生成对抗网络的图像超分辨率重建方法依赖配对数据集训练且结果不稳定的问题, 提出了一个新的基于无配对图像的模型 NM-SRGAN。首先, 通过使用循环生成对抗网络作预处理模块, 使模型可以不依赖配对数据集进行训练且获得更好的输入图像, 同时该模型取消了 BN 层的使用, 解决了结果不稳定的问题。然后, 使用了协方差矩阵捕捉图像的二阶信息, 增加了二阶损失函数, 更加注重于捕捉图像细节区域部分的变化。最后, 通过使用新的 VGG 损失函数提升了图像的边缘纹理细节。对提出的 NM-SRGAN 模型在 4 个标准数据集上进行测试评估, 使用客观评价标准对结果图进行评价, NM-SRGAN 模型较目前若干先进模型中的最佳峰值信噪比分别提升了 0.19、0.03、0.13、0.02 dB, 在 4 个数据集上的评价价值均达到最高。实验结果表明, 该模型在稳定性、图像质量和细节方面较经典算法均有较好的提升。

**关键词:** 超分辨率; 深度学习; 生成对抗网络; 无配对; 二阶统计量

**中图分类号:** TP399

**文献标志码:** A

**doi:** 10.13705/j.issn.1671-6833.2021.05.018

### 0 引言

将低分辨率(low resolution, LR)图像经过一系列变换得到高分辨率(high resolution, HR)图像的任务称为图像超分辨率<sup>[1]</sup>。图像超分辨率作为比较基础的视觉问题, 越来越受到计算机视觉界的广泛关注, 并且在军事、医学、安全等领域拥有广阔的应用前景<sup>[2]</sup>。

基于深度学习的超分辨率(super resolution, SR)任务是近年来研究的热门问题。Dong 等<sup>[3]</sup>提出的基于 3 层卷积的端到端的模型 SRCNN 是图像超分辨率在深度学习方面的开篇之作, 为此后各种基于深度学习的工作指明了方向。为了提升 SR 的性能, 各种深度学习的模型被提出。随着网络的加深, 超分辨率在 PSNR(峰值信噪比)和 SSIM(结构相似性)指标<sup>[4]</sup>方面取得了不错的成果, 但所得到的图像缺乏高频纹理细节, 在视觉质量方面达不到满意的效果。针对此问题, Ledig 等<sup>[5]</sup>提出利用残差块建立的基础模型 SRGAN, 构建了基于生成对抗网络(generative adversarial network, GAN)超分辨率重建的经典模型, 在图像感

知质量方面取得了较好的效果。

但是, SRGAN 在其结果方面展现了极度的不稳定性, 部分图像出现伪影现象<sup>[6-7]</sup>, 其结果在 PSNR 与 SSIM 方面产生了较大偏差, 如图 1 所示。在真实场景中, SRGAN 表现出了极高的不适应性, 包括噪声以及下采样未知等问题。同时, 在传统的卷积神经网络工作中, 通常使用最大化池或平均池和全连接层来获取一阶统计量, 而二阶统计量被认为是比一阶统计量更好的区域描述符<sup>[8]</sup>。

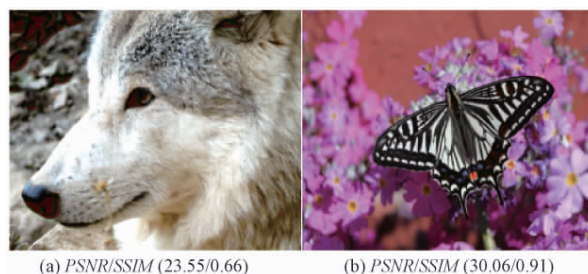


图 1 SRGAN 生成的图像对比图

Figure 1 Image comparison graph generated by SRGAN

基于以上问题, 提出一种改进的生成对抗网络模型(image super-resolution based on no match

收稿日期: 2021-01-19; 修订日期: 2021-03-01

基金项目: 国家重点研发计划项目(2020YFB1712401)

作者简介: 李学相(1965—), 男, 河南周口人, 郑州大学教授, 主要从事人工智能、物联网安全研究, E-mail: lxx@zzu.edu.cn.

generative adversarial network, NM-SRGAN), 其中 NM 代表无配对输入, SRGAN 代表基于生成对抗网络的超分辨率重建。首先, 该模型取消了 BN 层的使用, 解决了伪影的问题, 并提升了结果的稳定性。之后考虑到现实生活中高分辨图像不一定容易获得以及现实中下采样未知及噪声等问题, 受到循环生成对抗网络(cycle-gan)的启发, 使用 cycle-gan 作为预处理模块置于网络的前端, 改善了图像超分辨率在现实世界的适应性问题并达到无配对输入的目的<sup>[9]</sup>。最后通过在感知损失中增加二阶特征损失及修改 VGG 损失, 使输出的图像细节部分改善更加明显。基于以上改进, 该模型获得了更强的稳定性以及更高的图像质量。

## 1 相关工作

### 1.1 循环生成对抗网络

循环生成对抗网络(cycle-gan)的出现解决了在视觉问题(例如超分辨率领域)上难以找到匹配的高质量图片的问题, 因为很多模型训练时都依赖匹配的图像。cycle-gan 旨在学习数据域之间的风格变换, 从而减少对匹配数据的依赖性<sup>[10]</sup>。cycle-gan 在图像超分辨率、风格变换、图像增强等方面具有较强的适应性。

### 1.2 协方差矩阵

协方差矩阵作为二阶统计量, 被认为比一阶统计量能更好地捕捉图像的区域特征。给定一组特征, 可以使用协方差矩阵来统计输出图像的二阶信息<sup>[11]</sup>, 协方差矩阵可以表示为

$$C = \frac{1}{n-1} \sum_{i=1}^n (f_i - \bar{f})(f_i - \bar{f})^T. \quad (1)$$

式中:  $f_1, f_2, \dots, f_n \in R^d$  为特征集合且  $\bar{f} = \sum_{i=1}^n \frac{1}{n} f_i$ 。

## 2 NM-SRGAN 模型

### 2.1 网络结构

新的生成对抗网络模型如图 2 所示。预处理模块在图 2 的 Pretreatment 框架内, 预处理部分共包括生成器 G1 和 G2 与鉴别器 D 3 个部分, 具体结构如图 3 所示。图 2、3 中 Conv 代表卷积层, BasicBlock 代表模块中的基本块, 括号内的数字代表滤波器的数量。对于图 3 中的生成器 G1 与 G2, 在头部与尾部各使用 3 层  $3 \times 3$  的卷积层, 在中间使用 6 个基本块, 基本块的结构与图 2 中生成器的基本块的结构相同, 步长始终为 1。鉴别

器 D 设置了 5 层卷积层, 随着网络的加深特征个数不断增加。在图 2 中, Real 代表无噪声的 LR 图像, Sample 代表 G1 生成的图像。给定一个输入图像  $x$ , G1 的作用是学习并生成与 Real 相似的图像 Sample, 来骗过鉴别器 D, 鉴别器 D 则负责将实际样本 Real 与生成样本 Sample 区分开来, 因为训练过程中输入图像  $x$  与 Real 并不是配对的数据集, 所以使用 G2 来保证输入  $x$  与输出 Sample 的一致性。基于 G1、G2 与 D 的作用, 整个预处理模块也就达到了无配对训练和生成更好的 LR 输入图像的目的。然后将预处理部分的结果输入到主生成对抗网络中。图 2 中包含整个生成器结构和主鉴别器结构。在主生成对抗网络部分, 最前端有 1 层  $3 \times 3$  的卷积层, 卷积层之后包含 16 个基本块。在 SR 任务中, 去除 BN 层已经被证明可以提升性能并减少计算复杂度<sup>[6-7]</sup>, 因此取消了基本块中 BN 层的使用。每个基本块中包含 2 层  $3 \times 3$  的卷积及 1 层激活函数层, 在基本块之后包含 2 个增加特征尺寸的反卷积层及 1 层卷积层。主生成网络将预处理模块输入的 LR 图像生成超分辨率图像。在主鉴别网络部分有 8 层卷积层, 与生成网络中的基本块不同, 该模型没有删除鉴别器中的 BN 层, 因为鉴别器网络中的 BN 层并不会影响最终性能。主鉴别器网络的基本块中包含 1 层卷积层、1 层 BN 层和 1 层激活函数层, 最后使用 2 层全连接层与 1 层 Sigmoid 函数来获得样本分类的概率<sup>[5]</sup>, 并通过损失函数来优化网络参数。

整个生成对抗网络的训练过程中, 生成器与鉴别器发生持续的博弈, 直至生成网络可以骗过鉴别器<sup>[12]</sup>, 模型收敛, 得到想要的超分辨率图像。

### 2.2 损失函数

#### 2.2.1 循环网络损失

通过学习 cycle-gan 来设置损失函数。cycle-gan 的损失分为普通生成网络损失、循环一致性损失以及容易忽略的 identity 损失<sup>[10]</sup>。通过 G1 来生成样本, 用来生成一个更好的 LR 图像, 通过 D 用来区分 G1 生成的样本与真实样本, 普通 GAN 的损失为

$$l_{GAN} = \frac{1}{S} \sum_i^S \|f_D(f_{G1}(x_i) - 1)\|_2. \quad (2)$$

式中:  $S$  表示训练的样本数;  $x$  表示输入。

为了保证输入  $x$  与输出  $y$  的一致性设计了一个循环一致性损失, 循环一致性损失为

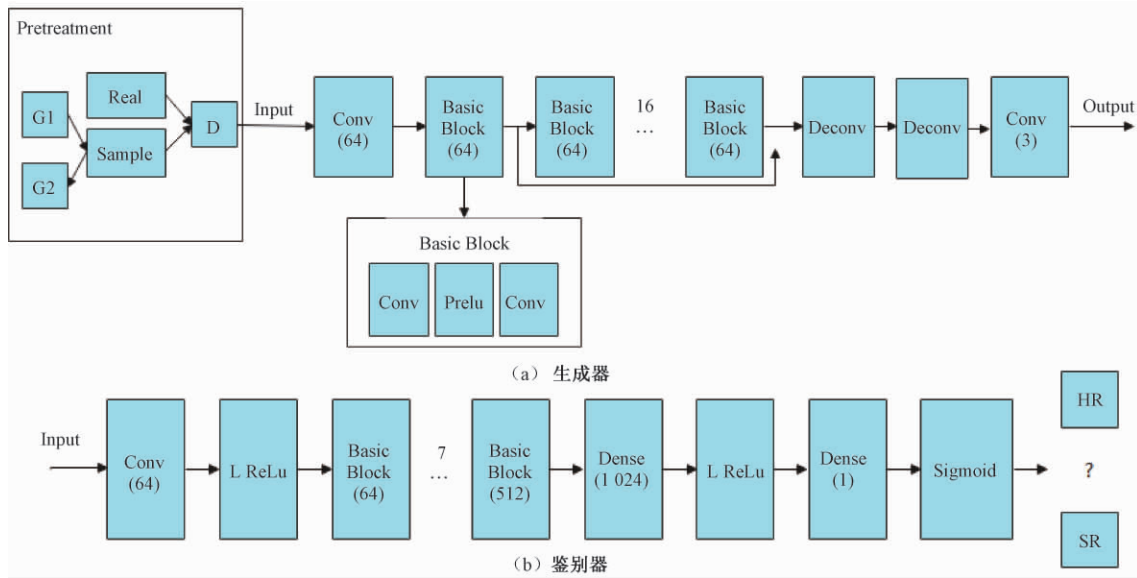


图2 生成器与鉴别器结构图

Figure 2 Generator and discriminator structure diagram

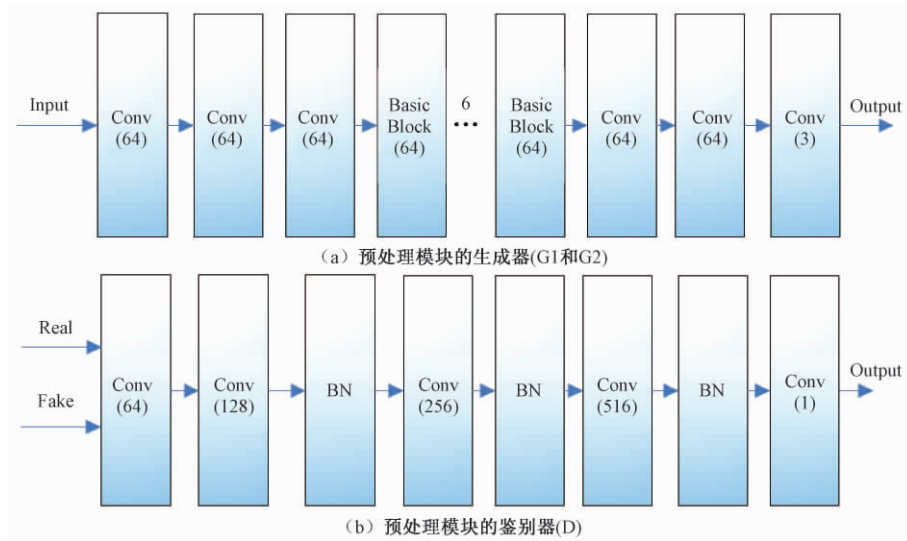


图3 预处理模块中的生成器网络与鉴别器网络结构图

Figure 3 Structure of the generating network and discriminating network in the preprocessing module

$$l_{\text{cycle}} = \frac{1}{S} \sum_i \|f_{G1}(x_i) - x_i\|_2. \quad (3)$$

identity 损失可以使生成的图像更加稳定,如果不加该损失,可能会使图像的整体颜色发生改变,identity 损失为

$$l_{\text{id}} = \frac{1}{S} \sum_i \|f_{G1}(y_i) - y_i\|_2. \quad (4)$$

因此,单循环网络总损失为

$$l_{\text{total}} = l_{\text{GAN}} + w_1 l_{\text{cycle}} + w_2 l_{\text{id}}. \quad (5)$$

式中:  $w_1$  和  $w_2$  为不同损失的权重。

### 2.2.2 感知损失

在本小节中,对原感知损失主要做出了2点改进:①增加了二阶损失函数;②修改了原有的VGG损失。

将协方差矩阵应用于图像的流程如图4所示。首先将卷积网络最后输出的特征图(三维矩阵)扁平化为二维矩阵,设  $\mathbf{X} \in \mathbf{R}^{w \times h \times c}$  为经过几个卷积层得到的输出,其中  $w$  为宽度,  $h$  为高度,  $c$  为通道数。 $\mathbf{X}$  通过扁平化得到  $\mathbf{X}_0 \in \mathbf{R}^{D \times c}$ , 其中  $D = w \times h$ , 如果  $f_1, f_2, \dots, f_n \in \mathbf{R}^d$  是  $\mathbf{X}_0$  的列, 那就可以通过式(1)计算协方差矩阵。这样得到的协方差矩阵通常位于对称正定(SPD)矩阵的黎曼流形上。但是直接扁平化会导致几何信息的缺失,如果使用标准方法采用对数运算扁平化黎曼流形结构,会导致得到的协方差矩阵过大,因此需要在保持几何结构的同时进行降维。

通过学习文献[13]中的算法设计了特殊降维层,如图5所示。图5中 BiRe 为 Bimap 与

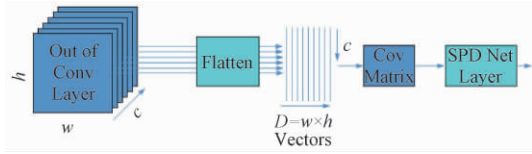


图4 图像扁平化流程图

Figure 4 Flowchart of image flattening



图5 降维流程图

Figure 5 Dimension reduction flowchart

ReEig 的合体。其中 Bimap 为双线性映射层,双线性映射层可以解决在降维的过程中保持几何结构的问题。 $k$  次双线性映射可以表示为

$$\mathbf{X}_k = f_b^k(\mathbf{X}_{k-1}; \mathbf{W}_k) = \mathbf{W}_k \mathbf{X}_{k-1} \mathbf{W}_k^T \quad (6)$$

式中:  $\mathbf{X}_{k-1}$  为 SPD 矩阵;  $\mathbf{W}_k \in \mathbf{R}^{d_k \times d_{k-1}}$  为权矩阵;  $\mathbf{X}_k \in \mathbf{R}^{d_k \times d_k}$  为输出矩阵。

ReEig 层为特征修正层,像 Relu 层一样被用于引入非线性, $k$  次 ReEig 层可以定义为

$$\mathbf{X}_k = f_r^k(\mathbf{X}_{k-1}) = \mathbf{U}_{k-1} \max(\epsilon \mathbf{I}, \sigma_{k-1}) \mathbf{U}_{k-1}^T \quad (7)$$

式中:  $\mathbf{X}_{k-1}$  为输入的 SPD 矩阵;  $\mathbf{X}_k$  为输出;  $\mathbf{U}_{k-1}$  通过特征值分解  $\mathbf{X}_{j-1} = \mathbf{U}_{k-1} \sum_{k-1} \mathbf{U}_{k-1}^T$  定义; max 操作为逐元素矩阵运算。

最后的 LogEig 层为对数特征层,用于赋予黎曼流形中的元素一个李群结构,使矩阵能够扁平化,并可以应用于标准的欧式运算<sup>[12]</sup>。第  $k$  层应用的对数特征层可以被定义为

$$\mathbf{X}_k = f_l^k(\mathbf{X}_{k-1}) = \log(\mathbf{X}_{k-1}) = \mathbf{U}_{k-1} \log(\sum_{k-1}) \mathbf{U}_{k-1}^T \quad (8)$$

式中:  $\mathbf{X}_{k-1}$  为输入矩阵;  $\mathbf{X}_k$  为输出矩阵,  $\mathbf{X}_k = \mathbf{U}_{k-1} \sum_{k-1} \mathbf{U}_{k-1}^T$  为特征值分解; log 操作为逐元素矩阵运算,且 log 运算对任意符合意义的底数都成立。

将 HR 图像及超分辨率图像经过以上流程扁平化及降维之后得出的协方差矩阵分别设为  $\mathbf{C}^{\text{HR}}$  和  $\mathbf{C}^{\text{SR}}$ ,则二阶损失函数为

$$l_{\text{cov}}^{\text{SR}} = \frac{1}{N} \sum_i^n (\mathbf{C}^{\text{HR}}(x_i) - \mathbf{C}^{\text{SR}}(x_i))^2 \quad (9)$$

式中:  $N$  为矩阵内元素个数。

在 Ledig 等<sup>[5]</sup>提出的原感知损失函数中包含内容损失与生成对抗损失。其中,内容损失又包括 MSE 损失与 VGG 损失。本节对 VGG 损失进行了修改,虽然依然使用了预训练的 19 层的 VGG 网络<sup>[14]</sup>的 ReLU 激活层来定义了 VGG 损

失,用来求出欧式距离,其中的输入不再是 HR 图像与 SR 图像的原图,而是 HR 与 SR 图像的一阶梯度图,记做  $I_g^{\text{HR}}$  与  $G_{\theta G}(I_g^{\text{SR}})$ 。VGG 损失为

$$l_{\text{VGG}_{ij}}^{\text{SR}} = \frac{1}{W_{ij} H_{ij}} \sum_{x=1}^{W_{ij}} \sum_{y=1}^{H_{ij}} (\Phi_{ij}(I_g^{\text{HR}})_{x,y} - \Phi_{ij}(G_{\theta G}(I_g^{\text{SR}})_{x,y}))^2 \quad (10)$$

式中:  $W_{ij}$  与  $H_{ij}$  表示各个特征图的尺寸;  $\Phi_{ij}$  表示在 VGG 网络中第  $i$  个最大化层之前通过第  $j$  个卷积获得的特征映射。

感知损失中的 MSE 损失与对抗损失一致,没有变化,因此不过多介绍。基于以上流程得出二阶损失及修改的 VGG 损失,则总感知损失为

$$l^{\text{SR}} = l_{\text{MSE}}^{\text{SR}} + l_{\text{VGG}}^{\text{SR}} + 10^{-3} l_{\text{cov}}^{\text{SR}} + 10^{-3} l_{\text{Gen}}^{\text{SR}} \quad (11)$$

式中:  $l_{\text{MSE}}^{\text{SR}}$  为 MSE 损失也是目前基于卷积的任务普遍使用的优化目标;  $l_{\text{Gen}}^{\text{SR}}$  为生成对抗损失,目的是生成让鉴别器无法区分的数据分布。

### 3 实验

#### 3.1 数据集

使用 NTIRE 2018 Super-Resolution Challenge 中的 DIV2K 作为训练数据集。DIV2K 的每种降采样数据集都包含 800 张训练图片,100 张验证图片以及 100 张测试图片。本文使用了其中的未知下采样的数据集进行训练,该数据集具有随机的模糊内核及像素偏移量。将 LR 数据集的前一半作为输入图像,同时将 HR 数据集的后一半设置为需要的 HR 图像并进行下采样,得到无噪的 LR 图像。训练过程并不依赖成对的数据集,实验在 4 倍比例因子下进行,同时采用数据集中 100 张照片进行验证。

#### 3.2 训练细节

训练共分为 2 步:第 1 步将输入的 HR 图像通过 cycle-gan 作预处理得到 1 个比双三次 Bicubic 算法更好的 LR 图像;第 2 步对得到的 LR 图像做提升分辨率的工作。实验使用 Anaconda3 搭建的 python3.7 的环境,基于 Pytorch 来构建模型,并使用 GeForce RTX 2080Ti (×4) GPU 进行训练。

在第 1 步中,将式(5)中的  $w_1$  与  $w_2$  分别设置为 10 与 5,使用 Adam 优化器,其中  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,学习率初始设置为  $2 \times 10^{-4}$ ,batchsize 设置为 16,随着迭代次数的增加,逐步减少学习率与 batchsize,直至函数收敛。在第 2 步中将第 1 步中获得的更好的 LR 图片作为输入,输入到修改后的生成对抗模型进行训练以获得 SR 图像,将



学习率设置为  $10^{-4}$ , 学习率的其他参数与第 1 步一样, 并设置学习率的衰减, *batchsize* 设置为 16, 迭代次数为 2 000 次。

3.3 实验结果

对第 1 步的结果进行定量研究, 并与双三次 Bicubic 方法进行比较, 结果表明, NM-SRGAN 模型的预处理方法可以无配对训练并获得一个更好的输入的 LR 图像, 如图 6、表 1 所示。将实验结果与深度学习方法 SRCNN、VDSR、DRCN 及 SRGAN 等在目前公开的数据集 Set5、Set14、BSD100、Urban100<sup>[15]</sup> 上进行测试, 通过 *PSNR* 与 *SSIM* 值的比较, 如表 2 所示, 该模型在 4 个数据集上均有提升, 较其他模型的最佳 *PSNR*, 分别提升了 0.19 dB、0.03 dB、0.13 dB、0.02 dB。最后通过结果图的部分截取对比发现, 该模型的结果

可以获得更高质量的细节( 如对比图 1 帽子的毛线、对比图 2 蝴蝶翅膀上的纹理), 如图 7、8 所示。

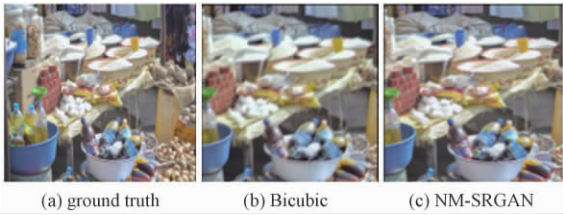


图 6 预处理方法对比图

Figure 6 Comparison diagram of preprocessing methods

表 1 预处理方法评估表

Table 1 Preprocessing method evaluation sheet

方法	<i>PSNR</i> /dB	<i>SSIM</i>
Bicubic	22.85	0.63
NM-SRGAN	<b>23.54</b>	<b>0.65</b>

表 2 不同模型重建图像评估表

Table 2 Evaluation table of images reconstructed by different models

数据集	Bicubic		SRCNN		VDSR		DRCN		SRGAN		NM-SRGAN	
	<i>PSNR</i> /dB	<i>SSIM</i>	<i>PSNR</i> /dB	<i>SSIM</i>	<i>PSNR</i> /dB	<i>SSIM</i>	<i>PSNR</i> /dB	<i>SSIM</i>	<i>PSNR</i> /dB	<i>SSIM</i>	<i>PSNR</i> /dB	<i>SSIM</i>
Set5	28.42	0.81	30.07	0.86	31.35	0.88	31.52	0.89	29.40	0.84	<b>31.71</b>	0.89
Set14	26.00	0.70	27.18	0.78	28.01	0.76	28.02	0.80	26.02	0.73	<b>28.05</b>	<b>0.81</b>
BSD100	25.96	0.66	26.68	0.72	27.21	0.72	27.21	0.74	25.16	0.66	<b>27.34</b>	<b>0.75</b>
Urban100	23.14	0.65	24.52	0.72	25.18	0.75	25.20	0.75	23.20	0.65	<b>25.22</b>	0.75

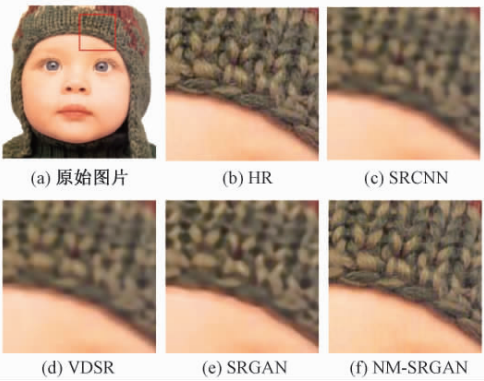


图 7 在不同数据集下的图像截取对比图 1

Figure 7 Comparison 1 of image captures in different data sets

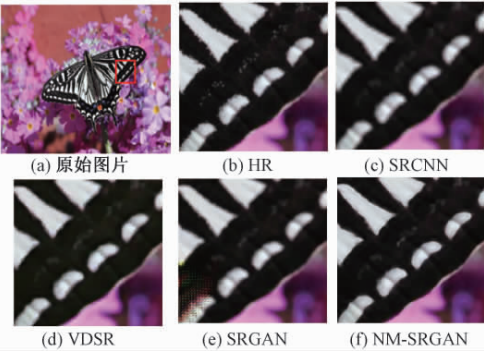


图 8 在不同数据集下的图像截取对比图 2

Figure 8 Comparison 2 of image captures in different data sets

4 结论

本文在对 SRGAN 的算法及模型进行学习之后对 SRGAN 模型进行改进。通过添加预处理模块达到无配对训练的目的并获得一个更好的输入图像, 同时通过删除 BN 层来稳定训练结果。此外, 还重新设置了感知损失函数, 来更好地恢复图像细节。通过以上改进提出了一个全新的 NM-SRGAN 模型, 将结果在公开数据集上与经典方法进行比较, 并用客观评价标准 *PSNR* 及 *SSIM* 进行评价与比较, 结果显示, 该模型在 4 个标准数据集上的评价价值均有较好提升, 且图像细节部分重建良好。但是该模型计算略显复杂, 下一步将对网络进行简化以提升训练速度。

参考文献:

[1] NASROLLAHI K, MOESLUND T B. Super-resolution: a comprehensive survey [J]. Machine vision and applications, 2014, 25( 6) : 1423–1468.

[2] YUAN Y, LIU S Y, ZHANG J W, et al. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops

- ( CVPRW ).Piscataway: IEEE,2018: 814–823.
- [3] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution [J]. Lecture notes in computer science, 2014, 8692: 184–199.
- [4] YANG C Y, MA C, YANG M H. Single-image super-resolution: a benchmark [J]. Lecture notes in computer science, 2014, 8692: 372–386.
- [5] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition ( CVPR ). Piscataway: IEEE, 2017: 105–114.
- [6] WANG X T, YU K, WU S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C] // 15th European Conference on Computer Vision, ECCV 2018. Cham: Springer, 2018: 1–16.
- [7] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. ( 2015–02–11 ) [2020–12–21]. <https://arxiv.org/abs/1502.03167>.
- [8] ACHARYA D, HUANG Z W, PAUDEL D P, et al. Co-variance pooling for facial expression recognition [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops ( CVPRW ). Piscataway: IEEE, 2018: 480–486.
- [9] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望 [J]. 自动化学报, 2017, 43( 3 ): 321–332.
- [10] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] // 2017 IEEE International Conference on Computer Vision ( ICCV ). Piscataway: IEEE, 2017: 2242–2251.
- [11] 苏咪咪, 叶中行. 协方差矩阵奇异情况下的最优投资组合 [J]. 应用概率统计, 2005( 3 ): 244–248.
- [12] 李方彪, 何昕, 魏仲慧, 等. 生成式对抗神经网络的多帧红外图像超分辨率重建 [J]. 红外与激光工程, 2018, 47( 2 ): 26–33.
- [13] HUANG Z W, GOOL L V. A Riemannian network for SPD matrix learning [EB/OL]. ( 2016–12–22 ) [2020–12–21]. <https://arxiv.org/abs/1608.04233>.
- [14] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [EB/OL]. ( 2016–08–29 ) [2020–12–21]. <https://arxiv.org/abs/1608.08155>.
- [15] 毛晓波, 张志超. 基于二维经验模态分解的单幅图像超分辨率重建 [J]. 郑州大学学报(工学版), 2014, 35( 5 ): 15–18.

## Image Super-resolution Based on No Match Generative Adversarial Network

LI Xuexiang, CAO Qi, LIU Chengming

( School of Software, Zhengzhou University, Zhengzhou 450002, China )

**Abstract:** Image super-resolution reconstruction based on generative adversarial networks ( GAN ) is subject to the dataset training with an unstable result. To solve this problem, a new NM-SRGAN model is established. The cycle-gan is firstly used as the preprocess module to make the model free from the dataset for training with better input of the image, and the model cancels BN layer to solve the unstable results. Besides, covariance matrix is adopted to capture the second-order information of the image, and second-order loss function is added with a focus on the changes of the image details. The new VGG loss function is used to improve the marginal texture of the image. The proposed NM-SRGAN model is verified by four standard datasets, and the resulting images are assessed by the objective evaluation indices. Compared with the existing models, NM-SRGAN model has an improved evaluation value of 0.19, 0.03, 0.13, and 0.02 dB, respectively, reaching up to the maximum among the four datasets. Results show that the proposed method, compared with traditional algorithms, has achieved better improvements in stability and image quality with better details.

**Key words:** super-resolution; deep learning; generative adversarial network; no matching; second-order statistic