

引文格式:

伍新洲, 袁宁徽, 沈立. 一种高效的多模型图像超分辨率框架 [J]. 集成技术, 2019, 8(5): 49-57.

Wu XZ, Yuan NH, Shen L. An efficient multi-model super resolution framework [J]. Journal of Integration Technology, 2019, 8(5): 49-57.

一种高效的多模型图像超分辨率框架

伍新洲 袁宁徽 沈立

(国防科技大学计算机学院 长沙 410073)

摘要 图像超分辨率 (Super Resolution, SR) 技术能够从低分辨率图像中恢复出高分辨率图像, 已被广泛应用于遥感、医学影像、目标跟踪与识别等多个领域。随着深度学习研究的深入, 该技术也被成功应用于 SR 相关研究中, 但现有工作往往只关注输出图像的质量, 而忽略了训练和重构效率。该文基于对图像特征和训练效率的观察, 提出了一种基于多模型的 SR 框架——MMSR, 能够根据不同的图像特征选择合适的网络模型, 从而在不影响输出图像质量的情况下有效缩短训练时间。面向 DIV_2K 图像集的测试结果表明, 该框架能够实现平均 66.7% 的性能提升, 同时具有良好的可扩展性。

关键词 超分辨率; 深度学习; 图像分类; 多模型

中图分类号 TP 391.4 文献标志码 A doi: 10.12146/j.issn.2095-3135.20190810001

An Efficient Multi-Model Super Resolution Framework

WU Xinzhou YUAN Ninghui SHEN Li

(School of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract Super resolution (SR) technique is an important means for image resolution improvement, which has been widely used in remote sensing, medicine image processing, target recognition and tracking etc. In recent years, the deep learning techniques also have been applied in the SR domain successfully. However, researchers pay most of their attentions on the quality of the output images, but ignore the training or reconstruction efficiency. In this paper, we found that for images with different texture features, the most appropriate models are usually different. Based on this observation, a multi-model super resolution framework (MMSR) is proposed, which can choose a suitable network model for each image for training. Experimental results with the DIV_2K image set indicate that, efficiency can be improved 66.7% without the loss of image quality. Moreover, MMSR exhibits good scalability.

Keywords super resolution; deep learning; image classification; multi-model

收稿日期: 2019-08-10 修回日期: 2019-08-27

基金项目: 国家自然科学基金项目 (61972407)

作者简介: 伍新洲, 硕士研究生, 研究方向为处理器体系结构和运行时技术; 袁宁徽, 硕士研究生, 研究方向为处理器体系结构和运行时技术; 沈立 (通讯作者), 博士, 教授, 研究方向为处理器体系结构和运行时技术, E-mail: lishen@nudt.edu.cn.

1 引 言

图像超分辨率(Super Resolution, SR)重构技术能够从一张或多张低分辨率图像恢复出高分辨率图像,已被成功地应用于监控、卫星图像、医学影像、目标识别与跟踪等诸多与图像处理密切相关的领域中。2014年,Dong等^[1]提出的SRCNN(Super Resolution Convolution Neural Network)首次将深度学习技术应用于SR研究中。随着深度学习技术的不断发展与广泛应用,越来越多的SR研究工作都在深度学习的基础上展开。目前基于深度学习的SR技术主要使用单张低分辨率图像进行重构,因此也被称作SISR(Single Image Super Resolution)。VDSR(Very Deep Convolutional Network based SR)^[2]也是SR技术发展历程中的一个具有重要意义的工作,它使用更深的网络并利用残差和很大的学习率来加快收敛速度。Leding等^[3]提出的SRGAN则是2017年以来十分流行的网络模型,它使用生成对抗网络(Generative Adversarial Network, GAN)模型来恢复图像中的纹理细节,大大提高了所得到高分辨率图像的真实感。

现有关于SR的研究基本都集中在如何提升输出高分辨率图像的质量上。研究者们设计了不同的网络模型或对特定的网络模型进行参数调优以获得更高质量的输出结果,但这也使SR所用的网络模型变得越来越复杂。复杂的网络模型增加了SR的训练时间和输出图像的重建时间,因此如何提高训练和重建效率成为SR研究不可回避的关键问题。实际上,由于不同图片具有不同的纹理特征,训练和重构这些图片所使用的网络模型以及所产生的时间开销也有很大的不同。一般来说,对于纹理特征较为简洁的图片,使用简单的网络模型就可以得到质量较高的重建结果,而纹理相对复杂的图片则需要使用相对复杂的网络模型才能得到满意的结果。

基于以上观察,本文提出了一种多模型SR框架——MMSR(Multi-Model Super Resolution Framework)。该框架中集成了多个不同的网络模型,在训练或重建时能够根据图像的纹理特征为每幅图像选择合适的网络模型,因此可以显著地缩短训练时间而不影响重构图像的质量。最终,本文在SRGAN^[3]的基础上实现了MMSR框架并面向DIV_2K进行了性能测试。测试结果表明,MMSR可以减少平均40%的训练时间,并且具有很好的可扩展性。本文的主要创新在于:

(1)提出并实现了一种多模型SR框架——MMSR。该框架能够根据输入图像的纹理特征为它们指定合适的网络模型,因而能够在不影响输出图像质量的同时提高训练效率。同时,MMSR框架可以与其他SR优化技术结合在一起使用,获得叠加的性能加速比。

(2)提出了一种图像分类方法——超过阈值的总方差(Total Variance above the Threshold, TVAT),能够根据图像像素的总方差将其分为不同类型。与传统的图像分类方法相比,TVAT的计算非常简洁,可以定量描述图像纹理的复杂度而不引入过多的额外计算开销。

2 相关工作

近年来,深度学习技术已被广泛应用在许多图像处理和分析领域,也包括图像超分辨率技术^[1-11]。Dong等^[1]最早将深度学习技术应用于SR领域,提出了一个简单的三层卷积神经网络模型SRCNN。该三个网络层分别实现了特征提取、非线性映射以及图像重构。低分辨率输入图像在进入SRCNN之前,先由一个三次插值模块转换为高分辨率图像。但这种方法缺少足够的高频信息,且会引入额外的计算。作为对SRCNN工作的改进,FSRCNN^[4]去掉了插值模块并使用一个反卷积层完成图像重构。VDSR^[2]是SR技术

发展过程中另一个标志性工作, 它使用了 20 层的网络并首次使用残差来减少计算量。在这之后, 几乎所有的 SR 研究工作都使用了残差。后来, 研究者们又尝试将生成对抗网络^[3]、递归神经网络^[5]等已有的网络模型应用于 SR 中, 或设计出由对称的卷积层-反卷积层构成的网络结构^[6], 或采用更深的网络结构^[7], 或对已有的网络结构进行一定的改进^[8], 均取得了一定的效果。在 Ledig 等^[3]首次将 GAN 网络用于 SR 之后, 研究者们又提出了 SFTGAN^[9]和 ESRGAN^[10]等方案。前者将语义分割概率图作为语义类别先验条件, 即确定图像中属于天空、水或草地等的区域, 以便生成更加丰富真实的纹理; 后者则为了去除 SRGAN 结果中的人工伪影, 增强结果的视觉质量, 从生成网络、判别网络和感知损失三个方面进行了提升。国内也有学者在 SR 领域进行研究^[11], 但多数工作是对已有方法或模型进行改进。

在 GAN 网络被用于解决 SR 问题之前, 均方误差被用作训练的损失函数。尽管使用这种方法可以得到较高的峰值信噪比, 但重构得到的图像会损失不少高频细节, 这会影响人们的视觉体验。图 1 为 SRGAN 的工作流程, 包括一个使用

生成网络的生成模块和一个使用判别网络的对抗模块。在判别网络的最后一层, SGRAN 使用感知损失 (Perceptual Loss) 确保输出图像的质量。其中, 感知损失描述了生成的图像与参考的高分辨率图像之间的差别, 若感知误差过大 (如超过了某个阈值), 将进行重新生成。

目前几乎所有基于深度学习的 SR 技术都在关注如何通过提出新的网络模型或对某个网络模型的参数进行调优来提高输出高分辨率图像的精度, 而忽略了训练的效率, 这会造成计算资源的浪费。因此, 本文提出 MMSR 框架, 以期提高 SR 的训练效率。因为 SRGAN 是目前被广泛应用于 SR 研究中且效果最好的模型, 所以本文以 SRGAN 为基础实现了所提出的 MMSR 框架。

研究者们关注的另一个问题是如何评价输出图像的质量。用来描述图像精度的参数有很多, 如均方误差、峰值信噪比、自然图像评价质量 (Natural Image Quality Evaluator, NIQE)、知觉指数 (Perceptual Index, PI)、结构相似度、信息保真度等。一般来说, 这些指标大致可分为两类。第一类通过像素特征来描述图像质量, 包括均方误差、峰值信噪比、结构相似度等。但是由于这些指标忽略了图像的纹理特征, 因而输出图

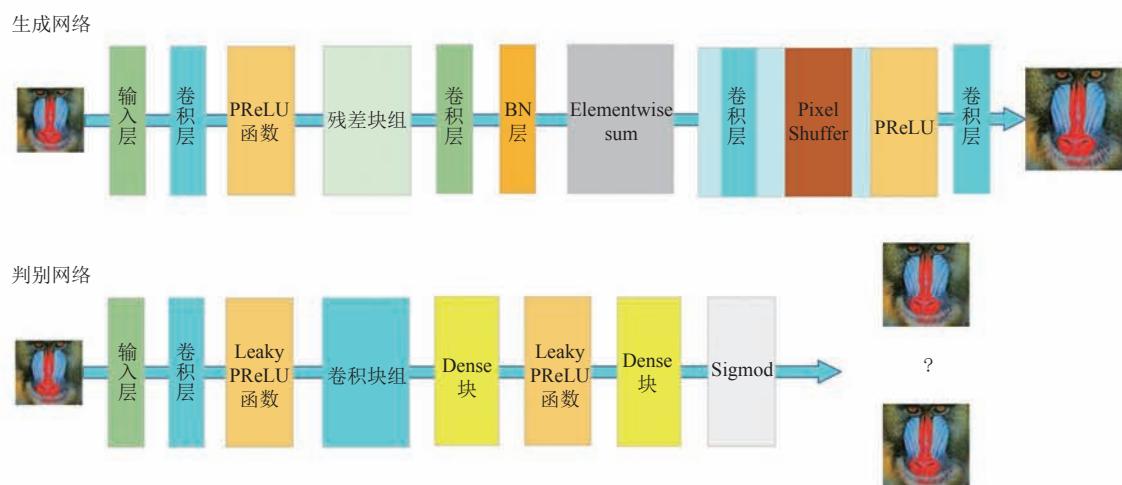


图 1 SRGAN 的结构

Fig. 1 Structure of SRGAN

像往往过于平滑或模糊。另一类指标则以人类的视觉感受为基础，如 NIQE 和 PI，其中 PI 的值在 NIQE 的基础上计算得到^[12]。有时甚至直接以人眼的视觉感受作为指标。显然，图像锐度越高、越自然，其 NIQE 值或 PI 值就越大。在近来的 SR 研究中，第二类指标逐渐成为主流的选择，而 PI 更是其中最主要的选择。因此，本文也选择 PI 值来描述输出图像的质量。

3 MMSR 框架

本节将首先介绍 MMSR 框架，它由一个训练模块和一个重构模块组成。其中，训练模块使用训练集对网络模型进行优化，重构模块则将输入的低分辨率图像重构为高分辨率图像。MMSR 具有很强的通用性，可以将不同的深度网络模型集成在该框架中。除了 MMSR 框架外，本节还将介绍本文提出的图像分类方法 TVAT。

3.1 框架概述

MMSR 的第一部分是训练模块，如图 2 所示。它分为两个阶段：分类阶段和多模型训练阶段。在分类阶段，图像按照其纹理特征被放入不同的类别。而在多模型训练阶段，系统会为不同的类别选择合适的网络模型，每一类图像会进入不同的模块进行训练。将图像分类会使后面的训练更有针对性，也会提高训练效率。各网络模型的主要区别在于其参数不同，如生成网络中残差块的数量以及判别器网络的层数。

MMSR 的第二个部分是重构模块，具体如图 3 所示。它由 4 部分组成，即分块模块、分类模块、多模型训练模块和重构模块，其中分类模块和多模型训练模块的结构与图 2 相同。该模块的流程如下：首先，分块模块将输入的低分辨率图像分为多块；然后，每一块进入分类模块并按照其纹理特征进行分类，同时为其指定网络模型，随后进入相应的模块进行重构；最后，重构出的各

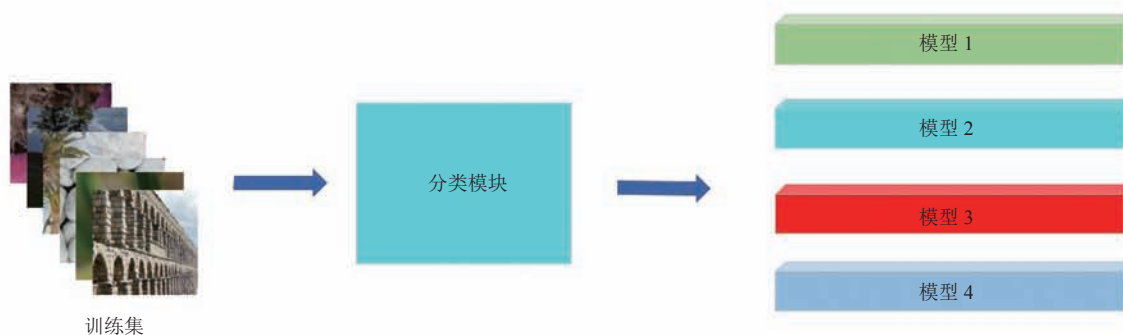


图 2 MMSR 训练模块的结构

Fig. 2 Structure of the training module in MMSR

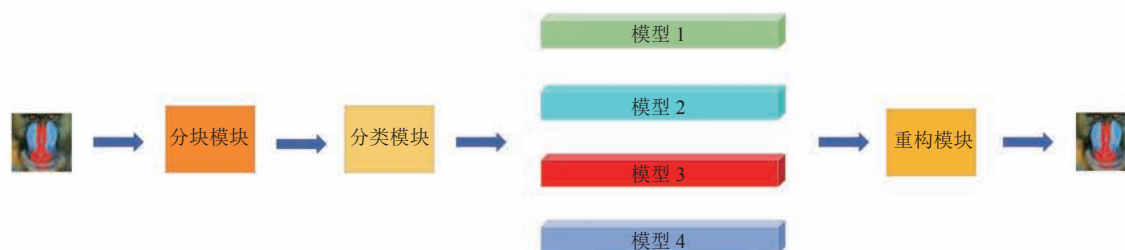


图 3 MMSR 重构模块的结构

Fig. 3 Structure of the reconstruction module in MMSR

图像块将被拼接在一起, 得到一个完整的图像。

将图像块分类后, 每一类会被送入不同的模块进行训练, 具体如图 4 所示。这些模块可以采用完全不同的网络模型, 也可以使用相同的、但具有不同参数的网络模型。本文采用第二类方式, 使用不同参数的 SRGAN 模型, 这是因为 SRGAN 模型输出的图像质量高于其他模型。此外, 在计算资源丰富的系统(如有多个计算节点的系统)中, 这些模块可被部署在不同的计算节点上。

如图 5 所示, 重构层将一组图像块合并为一幅输出高分辨率图像。因为分块之后, 不同类别的图像块会进入不同的模型训练, 训练完成时间的差异会导致图像的恢复时间有所不同。只有在所有图像块都训练结束后, 重构模块才开始生成输出图像。

有时, 图像重构过程中会出现边界效应, 影

响输出图像的质量, 如图 6 中的输出图像中有明显的图像块边界痕迹。一组边界重叠的图像块拼接在一起会消除边界效应。因此, 本文提出的框架在分块时会使相邻块的边界有一定的重叠, 并且可以调整重叠部分的宽度, 以确保尽可能完全消除边界效应。

3.2 图像分类

一般来说, 不同的图像具有不同的纹理特征, 但现有的绝大多数基于深度学习的 SR 方法都忽略了图像特征对训练或重构效率的影响。因此, 本文使用了一种能够根据图像纹理特征进行图像分类的低开销方法——TVAT。与现有的基于机器学习的分类方法相比, TVAT 的计算更为简洁, 开销也更低。

3.2.1 基于 TVAT 值的分类

为了寻找适合不同纹理特征的图像的网络模型, 本文对一些图像进行了测试。结果发现, 在



图 4 根据图像块的纹理特征将它们送入不同的模块

Fig. 4 Image segments are sent to different modules according to their texture features

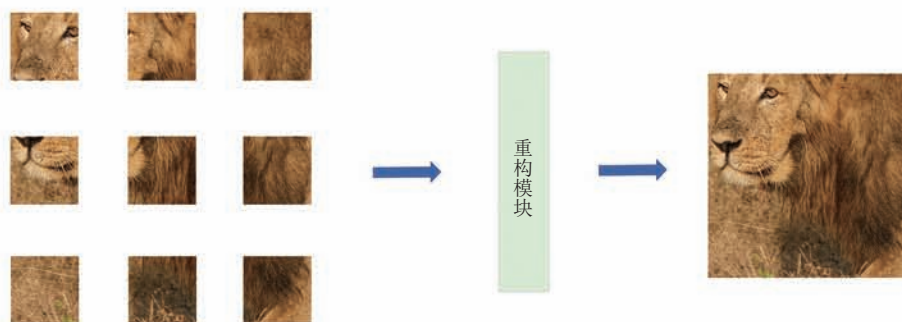


图 5 恢复模块将一组图像拼接为一幅完整的高分辨率图像

Fig. 5 The reconstruction module put a group of image segments together into a complete high resolution image

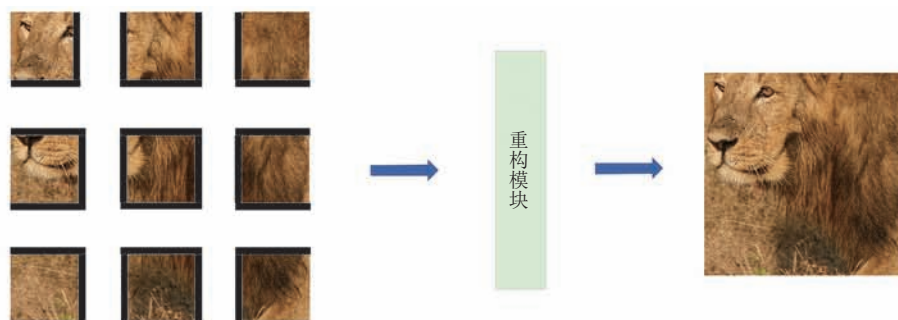


图 6 将一组边界重叠图像拼接为一幅完整的高分辨率图像

Fig. 6 A group of boundary overlapped images are put together into a complete high resolution image

使用同样的网络模型时，图像越复杂，所需的训练时间往往也越长。因此，本文试图寻找一种描述图像纹理特征的方法。描述图像纹理特征的最常用方法一般都是基于方差的。但这类方法的效果并非总是很理想，因为有些纹理均匀的图像的方差较大，但其每个像素的方差往往很小。因此，本文使用一种新的方法来描述纹理：首先计算出一个 3×3 块的中心点的方差，称作单像素方差 (Variance of Single Pixel, VSP)；其次，将一幅图像中所有大于某个阈值的 VSP 相加 (这里的阈值可通过实验确定)，得到这幅图像的 TVAT 值。通过测试以 $0 \sim 25$ 间的所有整数作为阈值时的分类效果发现，当阈值为 5 时效果最好，因此本文选择 5 作为计算 TVAT 值时的阈值。

图像中第 i 个像素点的 VSP 值可用公式 (1) 计算：

$$VSP_i = \sum_{j=1}^8 [(R_i, G_i, B_i) - (R_j, G_j, B_j)]^2 \quad (1)$$

图像的 TVAT 值可用公式 (2) 计算：

$$TVAT = \left(\sum_{i=1}^n VSP_i \times judge_{index_i} \right) / (cols \times rows) \quad (2)$$

其中， $cols$ 和 $rows$ 分别表示图像的行数和列数；

$judge_{index_i}$ 为阶跃函数，可用公式 (3) 计算：

$$judge_{index_i} = \begin{cases} 0, & VSP_i < \text{阈值} \\ 1, & VSP_i \geq \text{阈值} \end{cases} \quad (3)$$

3.2.2 TVAT 值和输出结果

本节讨论 TVAT 值与输出结果质量之间的关

系，如表 1 所示。表 1 中， x - y 表示一个生成网络使用了 x 个残差块、判别网络有 y 层的 GAN 模型。例如，16-8 表示 GAN 模型的生成网络有 16 个残差块，判别网络有 8 层。从表 1 可以看出，对于大多数图像，其 TVAT 值越大，图像的纹理越复杂。需要注意的是，并不是网络层数越深，输出图像的质量就越好。

从 DIV_2K 图像集中随机选择了 80 幅图像测试它们在不同模型下的重构质量，并对每一幅图像的 TVAT 值以及输出图像的质量进行统计，结果如图 7 所示。通过分析得到以下结论：对于 TVAT 值较小 ($0 \sim 2$) 的图像，使用 4-6 模型足以得到最优的结果；当 TVAT 值为 $2 \sim 4$ 时，2-2 模型或 16-8 模型均可以得到较好的结果；而当 TVAT 值大于 4 时，4-2 模型和 16-8 模型的结果

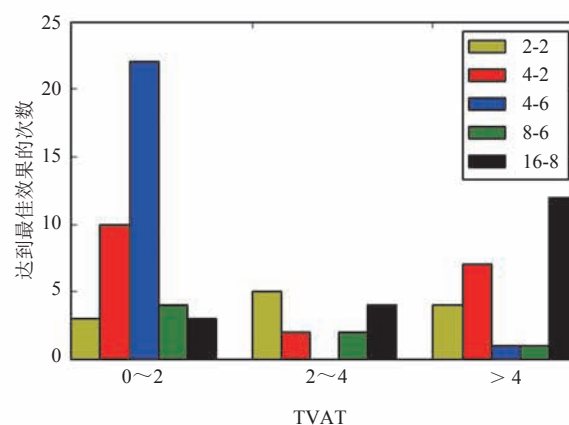


图 7 不同特征 (TVAT 值) 的图像在不同网络上的训练效果

Fig. 7 The training results of images with different features (TVAT values) on different networks

表 1 图像的 TVAT 值与不同的 GAN 网络模型之间的关系

Table 1 Relationship between the TVAT values and different GAN network models

| 测试图像编号 | TVAT 值 | 知觉指数(PI) | | | | | 最佳网络模型 |
|--------|----------|----------|----------|----------|----------|----------|--------|
| | | 2-2 | 4-2 | 4-6 | 8-6 | 16-8 | |
| 1 | 0.223 16 | 10.770 9 | 10.709 3 | 8.947 4 | 10.677 2 | 13.453 1 | 4-6 |
| 2 | 0.141 25 | 10.728 6 | 10.241 2 | 9.801 5 | 10.626 5 | 11.030 4 | 4-6 |
| 3 | 5.840 19 | 6.845 6 | 6.784 8 | 6.901 9 | 6.847 8 | 6.481 0 | 16-8 |
| 4 | 0.363 91 | 14.958 5 | 10.445 1 | 14.367 7 | 12.960 2 | 11.126 8 | 4-2 |
| 5 | 1.462 80 | 7.046 6 | 6.912 8 | 7.655 5 | 7.474 5 | 7.762 6 | 4-2 |
| 6 | 4.545 11 | 5.823 2 | 5.717 0 | 6.066 1 | 5.476 3 | 5.357 8 | 16-8 |
| 7 | 4.158 10 | 6.405 3 | 6.427 0 | 6.826 3 | 6.431 7 | 6.696 9 | 2-2 |
| 8 | 2.500 22 | 6.840 8 | 6.917 2 | 6.870 3 | 6.720 1 | 7.001 5 | 8-6 |
| 9 | 0.055 13 | 12.887 8 | 9.068 2 | 14.955 9 | 12.136 2 | 14.219 2 | 4-2 |
| 10 | 3.307 24 | 6.621 1 | 6.218 2 | 6.956 8 | 6.158 1 | 6.191 4 | 8-6 |
| 11 | 3.013 82 | 7.277 8 | 6.991 4 | 6.658 8 | 6.694 7 | 6.531 8 | 16-8 |
| 12 | 0.064 23 | 11.714 4 | 10.397 7 | 10.388 3 | 11.199 8 | 11.494 6 | 4-6 |
| 13 | 4.395 49 | 6.655 9 | 6.576 5 | 6.788 1 | 6.661 9 | 5.849 8 | 16-8 |
| 14 | 0.274 29 | 11.217 2 | 10.588 5 | 9.865 7 | 12.955 8 | 10.693 3 | 4-6 |
| 15 | 0.323 17 | 9.661 1 | 9.091 4 | 8.961 0 | 9.673 0 | 9.531 3 | 4-6 |
| 16 | 5.959 09 | 7.742 3 | 7.560 6 | 7.674 5 | 7.712 7 | 7.199 4 | 16-8 |

最好。因此, 本文使用图像的 TVAT 值对其进行分类。

4 测试与分析

4.1 实验环境

我们使用一个 4 个 CPU+GPU 节点的小规模异构集群测试 MMSR 框架的性能和可扩展性。该集群每个节点的基本参数如表 2 所示。

表 2 实验所用计算节点的主要参数

Table 2 Main parameters of the computation nodes

| 硬件/软件模块 | 参数 |
|---------|--------------------------------------|
| CPU | Intel® Xeon® E5-2660 v3 @2.6 GHz x 2 |
| GPU | NVIDIA Tesla K80 x 2 |
| 内存容量 | 64 GB |
| 操作系统 | Linux CentOS 7.4 |
| 开发环境 | Anaconda 3, Pytorch 1.0 |

本文选择 DIV_2K 图像集作为训练集和测试集。为对 MMSR 和其他 SR 框架输出图像的质量进行衡量, 本文采用 PI 作为评价指标。PI 是一种

被广泛使用的指标, 可以使用公式(4)^[12]计算:

$$PI = \frac{1}{2} [(10 - Ma) + NIQE] \quad (4)$$

其中, $NIQE$ 利用待评价图像特征模型参数与预先建立的模型参数之间的距离来确定图像质量; Ma 是一种基于人的视觉感知的超分辨图像的评价指标, 它反映了图像的一些空间域和频率域的统计学特征^[12]。从图 8 可以看出, PI 值较小的图

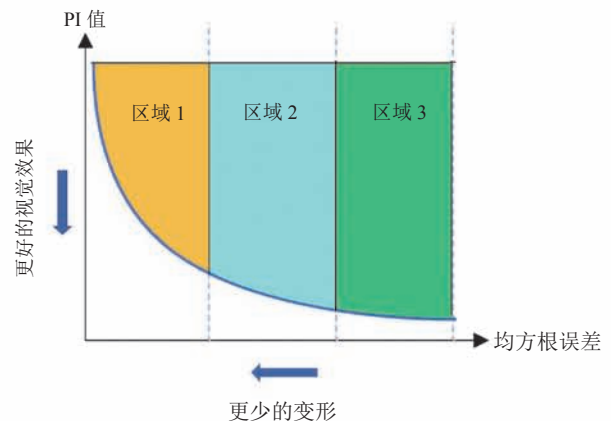


图 8 图像感知质量与图像失真之间的关系

Fig. 8 The relationship between PI and the loss of image

像具有较低的失真，即图像质量也比较高^[13-14]。

使用 PI 值描述图像质量的好处在于它更符合人眼的视觉感受，这是它与其他指标最大的不同。而且，SRGAN 模型能够更好地还原出图像的纹理细节，以获取更好的输出图像质量。

本文选择 Python 语言实现 MMSR 框架，使用 PyTorch^[15-17] 框架实现了训练模块。其中，PyTorch 是一个开源的 Python 机器学习库(基于 Torch 实现)，正在被越来越广泛地应用在当前的 AI 领域。

4.2 实验结果

实验中，根据图像的 TVAT 值将它们分为 3 类，并送往不同的 GPU 节点进行训练。从表 3 可以看到，与 16-8 SRGAN 网络模型^[9](这也是标准的 SRGAN 模型)相比，MMSR 的训练时间大大缩短。当使用单个 GPU 节点时，这 3 类图像只能串行地进行训练，MMSR 带来的性能加速比为 1.62；而在使用 3 个 GPU 节点时，这 3 类图像可以并行地进行训练，因此 MMSR 带来的性能加速比进一步提升，约为 2.9。

图 9 比较了使用 MMSR 和其他方法(SRGAN、Bicubic)所获得输出图像的质量。PI 值越小，表明输出图像的视觉效果越好，质量也

表 3 不同方法的训练时间

| SR 框架 | 训练时间(s) | 平均时间(s) |
|----------------------|------------|------------|
| SRGAN (单个 GPU 节点) | 16 415.600 | 17 171.338 |
| | 17 680.502 | |
| | 17 417.912 | |
| MMSR (单个 GPU 节点) | 10 683.279 | 10 624.901 |
| | 10 592.480 | |
| | 10 598.944 | |
| MMSR (三个 GPU 节点) | 5 921.761 | 5 894.845 |
| | 5 872.322 | |
| | 5 890.451 | |

就越高。从图 9 可以看出，Bicubic 方法的输出图像质量相对较差，MMSR 与 SRGAN 较接近，但对于某些图像(如 2、7、12、15 号图像)，MMSR 的效果要好于另外两种方法。

5 结束语

为具有不同纹理特征的图片选择合适的模型来提高 SR 的训练效率，本文提出了一种通用的多模型超分辨率图像重构框架——MMSR。为了定量地描述图像的纹理特征同时不引入过多的额外计算，本文设计了一种阈值控制的、根据总方

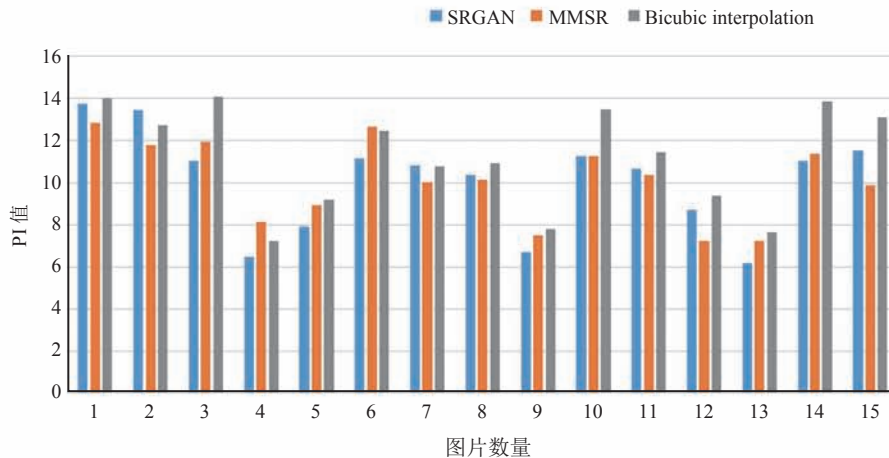


图 9 不同方法输出图像的质量

Fig. 9 The quality of images generated by different approaches

差进行分类的方法——TVAT。实验结果表明, 本文所提出的 MMSR 框架能够有效地缩短 SR 的训练时间, 并且基本不会影响输出图像的质量。MMSR 也表现出较好的可扩展性。此外, 不同的 SR 网络模型和 SR 优化技术均可以很容易地集成在 MMSR 框架中。当然, 本文所提出的 TVAT 只是一种简单的图像特征提取方法, 但通过这种方法证明了 MMSR 框架的有效性。今后将进一步研究更适合 MMSR 框架的图像特征提取和分类方法。

参 考 文 献

- [1] Dong C, Loy CC, He KM, et al. Learning a deep convolutional network for image super-resolution [C] // 2014 European Conference on Computer Vision (ECCV 2014), 2014: 184-199.
- [2] Kim JW, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), 2016: 1646-1654.
- [3] 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 2017), 2017: 105-114.
- [4] Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network [C] // 2016 European Conference on Computer Vision (ECCV 2016), 2016: 391-407.
- [5] Kim JW, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), 2016: 1637-1645.
- [6] Mao XJ, Shen CH, Yang YB. Image restoration using convolutional auto-encoders with symmetric skip connections [J]. arXiv: 1606.08921, 2016.
- [7] Tai Y, Yang J, Liu XM. Image super-resolution via deep recursive residual network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), 2017: 2790-2798.
- [8] Tong T, Li G, Liu XJ, et al. Image super-resolution using dense skip connections [C] // 2017 IEEE International Conference on Computer Vision (ICCV 2017), 2017: 4809-4817.
- [9] Wang XT, Yu K, Dong C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018), 2018: 606-615.
- [10] Wang XT, Yu K, Wu SX, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C] // ECCV 2018 Workshops, 2018: 63-79.
- [11] 张静, 石益强, 纪雯. 基于 CNN 与 ELM 的二次超分辨率重构方法研究 [J]. 计算机学报, 2018, 41(11): 2581-2597.
- [12] PIRM2018. The PIRM challenge on perceptual super resolution [EB/OL]. [2019-08-21]. <https://www.pirm2018.org/PIRM-SR.html>.
- [13] Blau YC, Michaeli T. Rethinking lossy compression: the rate-distortion-perception tradeoff [C] // International Conference on Machine Learning (ICML) 2019, 2019: 675-685.
- [14] Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [15] Yegulalp S. Facebook brings GPU-powered machine learning to Python [EB/OL]. InfoWorld, 2017-01-19[2019-08-21]. <https://www.infoworld.com/article/3159120/facebook-brings-gpu-powered-machine-learning-to-python.html>.
- [16] Ben L. Why AI and machine learning researchers are beginning to embrace PyTorch [Z]. O'Reilly Media, 2017.
- [17] Nikhil K. Deep Learning with Python [M]. Berkeley: Apress, 2017: 195-208.