

文章编号: 2095-2163(2020)10-0032-06

中图分类号: TP391.41

文献标志码: A

# 基于生成对抗网络的织物图案生成方法

李 锋, 邵 健

(东华大学 计算机科学与技术学院, 上海 201620)

**摘要:** 织物图案设计作为一种传统工艺,在当代艺术的环境下需要更加多元化的设计思路。本文基于深度卷积生成对抗网络(DCGAN)进行织物图案设计,为了消除 DCGAN 中的反卷积操作产生的棋盘效应,使用缩放卷积对原模型生成器中的反卷积操作进行优化,并在缩放卷积的上采样过程中分别采用最近邻插值和双线性插值,同时与原始 DCGAN 模型的生成样本进行实验对比。实验结果表明,使用缩放卷积改进的 DCGAN 可以更有效的提升生成图像的质量。

**关键词:** 织物图案设计; 深度卷积生成对抗网络; 棋盘效应; 缩放卷积

## Fabric Pattern Generation Based on Generative Adversarial Network

LI Feng, SHAO Jian

(College of Computer Science and Technology, Donghua University, Shanghai 201620, China)

**[Abstract]** As a traditional craft, fabric pattern design requires more diversified design ideas in the context of modern and contemporary art. This paper designs fabric patterns based on the Deep Convolution Generated Adversarial Network (DCGAN). In order to eliminate the checkerboard artifacts produced by deconvolution operation in DCGAN, this paper uses resize convolution to optimize the deconvolution in the generator of DCGAN. In the up-sampling process of resize convolution, nearest neighbor interpolation and bilinear interpolation are used respectively, and experimental comparison with the generated images of the original DCGAN is performed. Experiments show that the improved DCGAN using resize convolution can improve the quality of generated images more effectively.

**[Key words]** fabric pattern design; DCGAN; checkerboard artifacts; resize convolution

### 0 引言

纺织业作为中国的支柱性产业,在国民经济中占有重要的地位,并且具有很强的国际性优势<sup>[1]</sup>。目前纺织品的艺术设计作为纺织品行业的一个重要竞争因素,时尚创新逐渐成为新的市场要求。

图像的自动生成一直是计算机视觉领域的一个重点研究方向<sup>[2]</sup>。如今,深度学习等技术的发展带来了具有重要研究意义的图像生成技术。生成对抗网络(Generative Adversarial Nets, GAN)算法是近几年图像处理领域比较热门的技术,GAN 的提出给更多的领域提供了新的解决方案,已有众多研究成果如恢复残缺图像、生成逼真人脸、提升图像分辨率等;深度卷积生成对抗网络(Deep Convolutional GAN, DCGAN)作为 GAN 的典型衍生模型已成为近几年来备受关注的热点技术。然而在 DCGAN 生成图片时会产生类似棋盘格的干扰信息,这一问题普遍存在于众多模型中。在深度学习中,常常会使用从低分辨率向高分辨率逐步转换的方式来生成图

片,在 DCGAN 中亦是如此,具体表现为生成器中的反卷积操作。由于反卷积操作在多数情况下会导致不均匀的像素重叠,使得生成图像中的部分区域产生类似棋盘格的色彩信息,即棋盘效应。

针对该现象本文使用缩放卷积即上采样与正向卷积的方式代替 DCGAN 中的反卷积操作,对 DCGAN 进行改进。将该改进 DCGAN 模型应用于织物图案设计,实现了由人工智能模拟人类思维过程,完成创造性任务。

### 1 生成对抗网络

由 Goodfellow 等人于 2014 年提出的生成对抗网络(GAN)在深度学习领域迅速掀起热潮,成为近几年来备受关注的热点技术。GAN 在图像合成、超分辨率、风格迁移以及其他图像生成任务上表现出突出的性能。GAN 包括一个生成器 G(Generator)和一个判别器 D(Discriminator),其结构如图 1 所示。其中,生成器 G 通过输入噪声  $z$  去学习真实样本的分布,并尝试生成判别器 D 无法将其与真实样

**基金项目:** 国家重点研发计划项目(2017YFB0309800)。

**作者简介:** 李 锋(1969-),男,博士,教授,主要研究方向:模式识别、深度学习;邵 健(1995-),男,硕士研究生,主要研究方向:计算机视觉、图像处理。

**通讯作者:** 邵 健 Email: jee-shao@163.com

**收稿日期:** 2020-05-18

本区分开的样本, 而判别器  $D$  的目标是尽力区分开这两种样本, 可以将其视为二分类器, 用 0 和 1 分别代表伪样本和真实样本<sup>[3]</sup>。随着生成器  $G$  与判别器  $D$  的持续迭代更新, 两者的性能都会此起彼伏的增长。通过不断的博弈, 直到生成器  $G$  生成的样本可以以假乱真欺骗判别器  $D$  时, 判别器  $D$  的判别准确率约为 50%, 模型处于纳什均衡, 表示生成器  $G$  成功的生成了符合真实样本分布的样本。

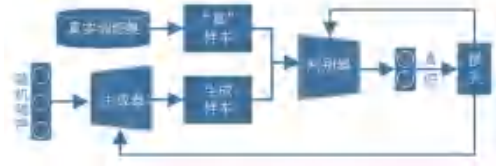


图 1 GAN 结构图

Fig. 1 Structure diagram of GAN

GAN 中判别器  $D$  的损失值来自两种情况: 一是将真实样本判别为生成器  $G$  生成的样本, 二是将生成器  $G$  生成的样本判别为真实样本<sup>[4]</sup>。由于生成器  $G$  和判别器  $D$  是零和博弈的过程, 所以其损失值即为判别器  $D$  损失值的负值, 可由式(1)表示:

$$\begin{aligned} J(D) &= -\frac{1}{2} E_{x \sim p_{data}(x)} [\log D(x)] - \\ &\quad \frac{1}{2} E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \\ J(G) &= -J(D) \end{aligned} \quad (1)$$

其中,  $D(x)$  表示判别器  $D$  判定输入样本为真实样本的概率,  $p_{data}(x)$  和  $p_z(z)$  分别表示真实样本分布和生成样本分布。根据式(1), GAN 的优化目标函数表示为式(2):

$$\begin{aligned} \min_G \max_D V(D, G) &= \\ E_{x \sim p_{data}(x)} [\log D(x)] &+ E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \end{aligned} \quad (2)$$

判别器  $D$  希望判别输入的真实样本的值趋于 1, 同时对生成样本的判别值尽可能的趋于 0, 即让  $D(x)$  无限趋近于 1,  $D(G(z))$  无限趋近于 0。相反, 生成器  $G$  的目标是学习到真实样本的分布  $p_{data}(x)$  并成功欺骗判别器  $D$ , 即使得  $D(G(z))$  尽可能的趋近于 1。

在实际训练的过程中, 训练初期生成器  $G$  几乎不可能生成出与真实样本相似的图像, 即无法很好的模拟真实样本的分布, 判别器  $D$  就会很轻易的分辨出样本的真假, 导致判别器  $D$  很快收敛。同时, 反馈的损失值也非常小, 导致生成器  $G$  在迭代更新

的过程中产生较小的梯度, 更新梯度便会下降甚至停止, 导致生成器  $G$  极早的停止学习。因此, 在训练生成器  $G$  时, 将其原损失中的  $\log(1 - D(G(z)))$  替换为  $\log G(z)$ , 这使得生成器  $G$  的弱梯度引起的模型发散问题得以解决。

上述零和博弈中, 生成器  $G$  与判别器  $D$  的损失值之和严格等于 0, 即饱和式零和博弈, 同时, 还有一种可以解决因为生成器  $G$  梯度消失导致 GAN 无法收敛的问题的方式, 称作非饱和式零和博弈<sup>[5]</sup>, 其定义如式(3)。由于式(2)是整体围绕判别器  $D$  而定义的损失, 导致判别器  $D$  一旦收敛就无法更新生成器  $G$ 。而式(3)通过重新定义生成器  $G$  的损失来避免生成器严重约束于判别器。

$$\begin{aligned} J(D) &= -\frac{1}{2} E_{x \sim p_{data}(x)} [\log D(x)] - \\ &\quad \frac{1}{2} E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \\ J(G) &= -\frac{1}{2} E_{z \sim p_z(z)} [\log D(G(z))] \end{aligned} \quad (3)$$

## 2 改进的 DCGAN 模型

针对原始 GAN 模型难以收敛、生成样本缺乏多样性以及生成样本质量较差等问题, Alec Radford 等人对原始 GAN 模型进行改进, 于 2015 年提出了深度卷积生成对抗网络 (DCGAN)。相对于原始 GAN 模型, DCGAN 增加了更多的隐藏层和参数。DCGAN 与 GAN 理论具有一致的标准结构, 主要区别在于 DCGAN 将卷积神经网络与 GAN 结合, 通过两个卷积神经网络实现生成器  $G$  和判别器  $D$ 。具体来说, DCGAN 存在如下的改变:

(1) 取消池化层, 取而代之使用跨步卷积, 在生成器  $G$  中使用反卷积进行上采样;

(2) 取消全连接层, 使用卷积层连接生成器  $G$  和判别器  $D$  的输入输出;

(3) 生成器  $G$  和判别器  $D$  中均使用批量归一化 (Batch Normalization, BN), 可以有效缓解初始化不理想的问题;

(4) 生成器  $G$  中除了输出层使用 Tanh, 其他各层使用 ReLU 作为激活函数。该改进方式可以有效防止梯度消失, 同时加速网络收敛。激活函数 ReLU 与 Tanh 表达式分别如式(4)和式(5):

$$ReLU(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

$$Tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5)$$

(5) 判别器 D 使用 LeakyReLU 作为激活函数, 其表达式(6)如下:

$$\text{LeakyReLU}(x) = \begin{cases} x, & x > 0; \\ \alpha x, & x \leq 0, \alpha \text{ 为负值梯度值.} \end{cases} \quad (6)$$

在 DCGAN 的判别器 D 中, 通过使用卷积神经网络模型对输入样本进行不断的卷积与下采样运算, 最后完成真假样本的分类。而在生成器 G 中, 则使用反卷积来完成上采样运算。反卷积的具体运算过程如图 2 所示。首先填充输入图像, 然后进行卷积运算, 裁剪后的卷积运算结果即为反卷积的运算结果。

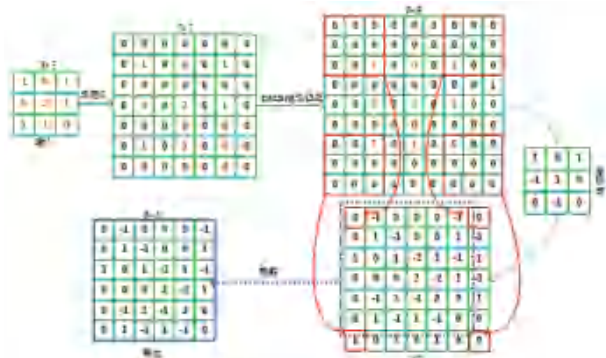


图 2 反卷积运算

Fig. 2 Deconvolution operation

反卷积操作的一个弊端是会造成不均匀的像素重叠, 导致生成图像产生明显的棋盘效应<sup>[6]</sup>, 在卷积核大小非步长的整数倍时尤为突出。这种重叠的样式体现在两个维度上, 两个坐标方向上不均匀的重叠相乘产生了类似棋盘格的样式特性。在一维情况下, 一个  $\text{size} = 3, \text{stride} = 2$  的反卷积操作所展示的棋盘效应如图 3 所示。在实际生成图像中棋盘效应的表现如图 4 所示, 其为使用未改进的 DCGAN 循环训练 5000 次生成的格子织物图案。



图 3 反卷积重叠

Fig. 3 Deconvolution overlap

为了有效抑制生成图像的棋盘效应, 本文中使用了缩放卷积 (resize convolution) 代替原 DCGAN 模型生成器中的反卷积操作。缩放卷积首先对图像上采样, 将图像大小变为原图的 4 倍, 经过一层正向卷积操作。上采样操作分别使用最近邻插值和双线性插值两种方案进行实验。设原图像 A 的大小为  $m \times n$ , 对图像 A 上采样后的图像 B 的大小为  $a \times b$ 。现

取图像 B 中的一点  $(x, y)$ , 可计算对应应在图像 A 中的理论位置为  $(\frac{x \times m}{a}, \frac{y \times n}{b})$ 。而该理论位置可能为小数, 即在图像 A 中无实际对应的点, 对于最近邻插值, 将对该点四舍五入, 即取图像 A 中距离该理论点最近的点作为图像 B 中的点  $(x, y)$ ; 对于双线性插值, 如图 5 所示。



图 4 棋盘效应

Fig. 4 Checkerboard Artifacts

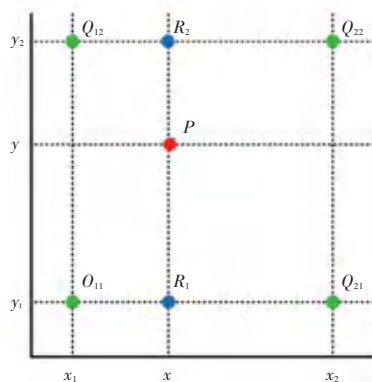


图 5 双线性插值

Fig. 5 Bilinear interpolation

图 5 中, P 为理论点,  $Q_{11}, Q_{12}, Q_{21}, Q_{22}$  为 P 点最近的 4 个点, 将在图像 A 中找到该理论点最近的 4 个点。其计算方法为首先通过线性插值得点  $R_1$  与  $R_2$ , 再基于点  $R_1$  与点  $R_2$  线性插值, 求得点 P, 式(7)、式(8)和式(9):

$$f(x, y_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}), \quad (7)$$

$$f(x, y_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}), \quad (8)$$

$$f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2). \quad (9)$$

$f(x, y)$  即为通过双线性插值计算后图像 B 中点  $(x, y)$  的值。

同时, 在训练网络的过程中, 生成器 G 的损失值一直处于递增的趋势。该现象是由于在训练过程中判别器 D 达到过拟合, 其损失值在初期迅速的向 0 逼近, 以致于 G 无法继续学习。本文通过在判别器 D 的输入中添加高斯噪声来抑制判别器 D 达到过拟合状态, 提高生成器 G 的学习能力。

本文基于 DCGAN 改进后的模型结构,如图 6 所示。将原模型生成器 G 中 3 个反卷积层替换为 3 个缩放卷积层,即上采样和正向卷积操作。判别器 D 首先对输入样本添加高斯噪声,经过多层卷积运算后映射 1 个标量结果。

在深度学习模型中,网络的超参数对网络性能起着决定性的作用。超参数的设置和更新标准为尽可能的使模型的损失值变小,即模型输出的真实值与理论值尽可能的接近,生成的样本质量越高。假设给定  $N$  个样本,经过模型训练输出结果为  $y = (y_1, y_2, \dots, y_n)$ , 其对应的理论输出结果为  $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ , 对此可将训练样本的损失定义为  $L(y, \hat{y})$ , 超参数将向使得  $L(y, \hat{y})$  最小化的方向更新。本文采用交叉熵计算损失值,式(10)表示每批训练样本的损失函数:

$$L(y_j, \hat{y}_j) = -\hat{y}_j \log_2 y_j - (1 - \hat{y}_j) \log_2 (1 - y_j). \quad (10)$$

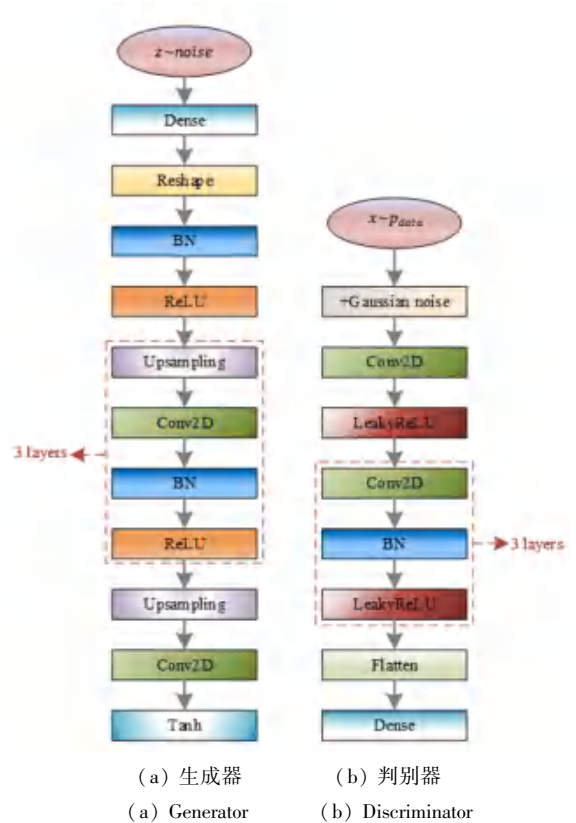
其中,  $\hat{y}_j$  代表第  $j$  批样本的真实理论值,  $y_j$  代表模型实际输出值。所有批次样本的损失值的均值即为模型的整体损失值,如式(11)所示:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{j=1}^N [\hat{y}_j \log_2 y_j + (1 - \hat{y}_j) \log_2 (1 - y_j)]. \quad (11)$$

为了使得 DCGAN 模型收敛并达到稳定的效果,需要求得可以最小化  $L(y, \hat{y})$  值的权重参数  $\omega$  和偏置  $b$ 。实验中采用梯度下降法更新  $\omega$  和  $b$  的值,并在生成器 G 和判别器 D 中均使用 Adam 优化器。式(12)和式(13)表示  $\omega$  和  $b$  的更新方法。其中,  $\eta$  代表学习率。

$$\omega_l \rightarrow \omega_l' = \omega_l - \eta \frac{\partial L}{\partial \omega_l}, \quad (12)$$

$$b_l \rightarrow b_l' = b_l - \eta \frac{\partial L}{\partial b_l}. \quad (13)$$



(a) 生成器 (b) 判别器  
(a) Generator (b) Discriminator

图 6 改进 DCGAN 网络结构

Fig. 6 Improved DCGAN structure

改进 DCGAN 的网络结构如图 6 所示,生成器 G 和判别器 D 的网络参数分别见表 1、表 2。其中,  $m@k \times k \times n$  表示当前层的卷积核设置为  $k \times k \times n$ , 并运算出  $m$  个特征图。

表 1 生成器 G 网络参数

Tab. 1 Generator network parameters

层号	描述	结构	输出
1	全连接层	[8192, 1] → [4, 4, 512] → ReLU	[4, 4, 512]
2	缩放卷积层 1	256@5×5×512, strides:2×2 → ReLU	[8, 8, 256]
3	缩放卷积层 2	128@5×5×256, strides:2×2 → ReLU	[16, 16, 128]
4	缩放卷积层 3	64@5×5×128, strides:2×2 → ReLU	[32, 32, 64]
5	缩放卷积层 4	3@5×5×64, strides:2×2 → tanh	[64, 64, 3]

表 2 判别器 D 网络参数

Tab. 2 Discriminator network parameters

层号	描述	结构	输出
1	卷积层 1	64@5×5×3, strides:2×2 → Leaky ReLU	[32, 32, 64]
2	卷积层 2	128@5×5×64, strides:2×2 → Leaky ReLU	[16, 16, 128]
3	卷积层 3	256@5×5×128, strides:2×2 → Leaky ReLU	[8, 8, 256]
4	卷积层 4	512@5×5×256, strides:2×2 → Leaky ReLU	[4, 4, 512]
5	全连接层	[4, 4, 512] → [8192, 1]	[1]

### 3 实验分析

本文从互联网中收集了 2 420 张格子图案的织物图像作为原始数据集,并对这些图像添加随机噪声,翻转变换以及亮度更改等处理,将数据集样本数量扩充至 8 000,所有样本数据均已进行中心裁剪和归一化等预处理操作,样本图像分辨率为  $64 \times 64$ 。数据集示例如图 7 所示。



图 7 格子织物图案样本示例图

Fig. 7 Sample images of plaid fabric pattern

本文中实验的硬件环境为 Intel(R) Xeon(R) Bronze3204 CPU, 16G RAM 和 NVIDIA GeForce RTX 2080Ti 11G 显卡;软件环境为 TensorFlow 1.14 及 Python3.5。实验中采用小批量随机梯度下降的方式进行训练,以提升模型训练的速度和准确性。设置卷积核大小  $5 \times 5$ ,步长为 2,学习速率 0.001,共 5 000 次循环,单次循环覆盖所有数据集样本。本文共进行 3 组实验,分别为原始 DCGAN 模型、缩放卷积中采用最近邻插值的改进 DCGAN 模型 NN-DCGAN 以及缩放卷积中采用双线性插值的改进 DCGAN 模型 Bilinear-DCGAN。在相同实验环境下分别进行训练,3 组实验在不同训练阶段的生成图像效果如图 8 所示。

在评价模型效果和图像质量时,主要分为主观评估和客观评估两种方式。针对主观视觉方面,通过实验结果可以发现,随着网络的不断迭代更新,生成图案质量有明显的提升。从 epoch = 5k 的生成图像可以发现 Bilinear-DCGAN 和 NN-DCGAN 都比原始 DCGAN 生成的样本纹理更加清晰,具有更少的噪声,可以明显的抑制棋盘效应。

为了更加有效的验证模型效果,本文对对比实验中的不同模型进行了客观评估和分析。模型的训练损失值一方面表征了模型的收敛效果,同时也体现了生成图像和真实图像的相似度,即生成图像的真实

性。对比实验中训练总损失变化曲线如图 9 所示。

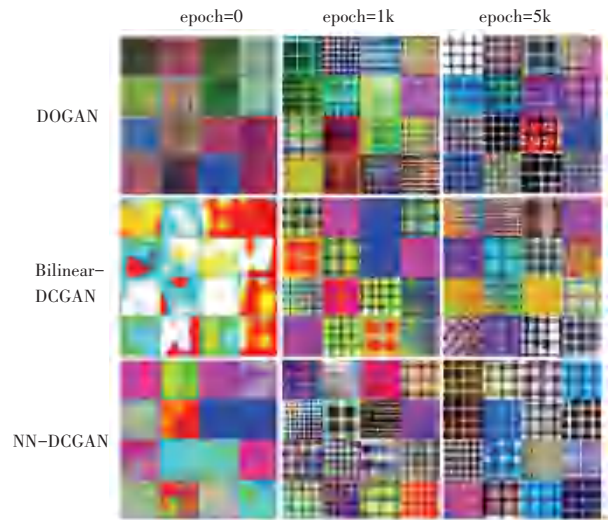


图 8 不同模型实验效果

Fig. 8 Experimental results of different models

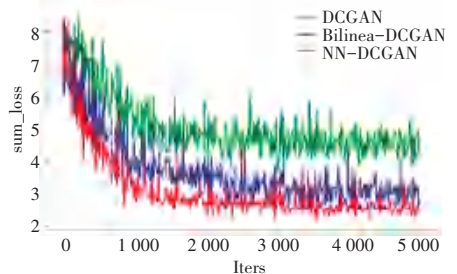


图 9 不同模型的损失曲线

Fig. 9 Loss curves of different models

同时,本文使用 Inception Score (IS) 和 Fréchet Inception Distance (FID) 两种客观评估方法对模型生成图像的真实性和多样性进行评估,IS 值越大代表生成样本质量越高,FID 值越小代表模型效果越好。最终评估结果见表 3。

表 3 模型评估结果

Tab. 3 Model evaluation results

	IS 值	FID 值
真实样本	36.06	2.68
DCGAN	18.44	54.59
Bilinear-DCGAN	20.50	51.02
NN-DCGAN	22.16	49.55

从图 9 和表 3 的评估结果可以得出,相对原始 DCGAN 而言,Bilinear-DCGAN 和 NN-DCGAN 在训练过程以及生成图像质量方面都具有更佳的效果,证明该改进方法对抑制棋盘效应并提高图像质量起到了积极作用。同时也发现在本文针对织物图案的实验中 NN-DCGAN 具有最佳的实验效果和评估结果,即在缩放卷积的上采样方式中,使用最近邻插值上采样相比使用双线性插值具有更好的效果。

(下转第 43 页)