



基于双鉴别器 GAN 的包装类产品外观设计法

doi:10.3969/j.issn.1674-7100.2020.02.011

李佳昕 孙刘杰 王文举

上海理工大学
出版印刷与艺术设计学院
上海 200093

摘要: 针对传统包装类产品外观设计工作中设计周期较长、设计难度较大以及设计成本相对较高等问题,将双鉴别器 GAN 应用于包装类产品外观设计工作中。在 BicycleGAN 的基础上,添加了一个局部鉴别器,并重新构建了相应的损失函数以及目标函数。实验时,输入对象为“边缘图像+原图”的图像对,输出对象为真实性概率排前 10 的图像。从多样性、PSNR 值以及 SSIM 值 3 个方面对实验结果进行评估,结果表明:双鉴别器 GAN 生成的图像不仅拥有较好的多样性,其图像细节质量也有所提高。因此,将双鉴别器 GAN 应用于包装类产品的外观设计工作是可行的,一方面能够为设计师提供设计灵感,另一方面也节省了人力与物力成本,提高了工作效率。

关键词: 双鉴别器 GAN; 外观设计; 图像转换

中图分类号: TP317.4 **文献标志码:** A

文章编号: 1674-7100(2020)02-0077-07

引文格式: 李佳昕,孙刘杰,王文举.基于双鉴别器 GAN 的包装类产品外观设计法[J].包装学报,2020,12(2):77-83.

1 研究背景

深度学习是机器学习中一个新的研究方向,其概念源于对神经网络的研究。在图像处理领域,利用深度学习可以解决不少问题,如图像分割、图像检测、图像翻译等。2014年,I. J. Goodfellow 等提出了生成式对抗网络(generative adversarial network, GAN)^[1-3],并将其应用于图像翻译任务中,研究结果表明 GAN 取得了比传统卷积神经网络(convolutional neural networks, CNN)更好的效果。大部分 GAN 是基于成对或非成对的输入图像,学习并生成与目标图像相对应的输出图像,使得生成的

图像具有输入图像的纹理和目标图像的内容。其中,较为经典的有 P. Isola、Zhu J. Y. 等提出的有监督的 Pix2Pix 模型^[4]和无监督的 CycleGAN 模型^[5]。随后,Zhu J. Y. 等^[6]又提出了有监督的 BicycleGAN 模型,实现了图像翻译工作从单一模态到多模态的升级。S. Iizuka 等^[7]在利用 GAN 解决图像修复问题时,提出了双鉴别器 GAN 模型,该模型可以更好地提升图像质量。

外观设计是以包装为用途的产品设计的重点之一。在设计时,需要考虑包装的材料与结构、品牌的文化内涵、色彩的运用以及契合节日主题等因素。这使得包装类产品外观的设计工作变得复杂,设计难

收稿日期:2020-01-26

基金项目:上海市科学技术委员会科研计划基金资助项目(18060502500)

作者简介:李佳昕(1995-),女,江苏泰州人,上海理工大学硕士生,主要研究方向为数字图文信息处理,
E-mail:1258482473@qq.com。

通信作者:孙刘杰(1965-),男,安徽六安人,上海理工大学教授,博士,主要从事印刷机测量与控制技术、数字印刷防伪技术、光信息处理技术方面的教学与研究,E-mail:liujiesunx@163.cpm

度加大，时间成本与人力成本增加。

包装类产品的外观设计可以看作是图像到图像间的风格转换问题，但目前将深度学习与包装外观设计相结合的研究鲜有报道。深度学习可以根据输入草图，生成外观设计的效果图，这既为包装类产品的外观设计工作提供了配色和纹理选择的灵感，也为后续的个性化修改工作提供了参考。因此，本课题组提出了基于双鉴别器 GAN 的包装类产品外观设计方法，以期包装外观设计提供一种新模式。

2 双鉴别器 GAN 介绍

2.1 双鉴别器 GAN 原理

双鉴别器 GAN 包含一个生成器 G、一个编码器 E 以及两个并联的鉴别器即全局鉴别器 GD 与局部鉴别器 LD。输入图像是成对的“边缘图像 + 原图”，边缘图像是通过 Canny 算法从原图提取边缘信息而获得的，标记为 A；原图即目标图像，标记为 B，通过编码器 E 编码其颜色、纹理等信息。

双鉴别器 GAN 的算法流程如图 1 所示。

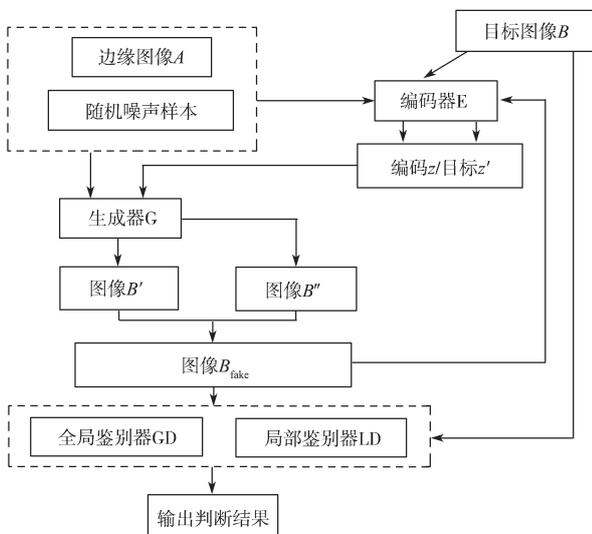


图 1 算法流程图

Fig. 1 Algorithm flow chart

在生成网络部分，生成器 G 根据输入的边缘图像 A 结合随机噪声 z_1 生成图像 B' ，即

$$B' = G(A, z_1) \quad (1)$$

将目标图像 B 输入编码器 E 进行编码，从而获得编码 z，即

$$z = E(B) \quad (2)$$

随后，将边缘图像 A 与编码 z 输入生成器 G，得到新的图像 B'' ，即

$$B'' = G(A, z) \quad (3)$$

最后通过 cat 函数将图像 B' 和图像 B'' 拼接为图像 B_{fake} ，即

$$B_{fake} = \text{cat}(B', B'') \quad (4)$$

B_{fake} 作为生成器 G 的最终输出结果，被输入到两个鉴别器进行判断。其中，GD 的输入为完整的 B_{fake} ，LD 的输入为 1/4 个 B_{fake} ，鉴别器组的判断结果将作前向反馈，以优化算法。

2.2 生成器 G

生成器 G 采用 U-Net 网络架构^[8]。U-Net 网络是一种典型的基于 CNN 的端到端的网络结构，主要用于图像分割领域。一般 U-Net 网络由卷积层、下采样层（最大池化层）、上采样层（反卷积层）和非线性激活函数（ReLU 等）组成，其结构示意图如图 2 所示。

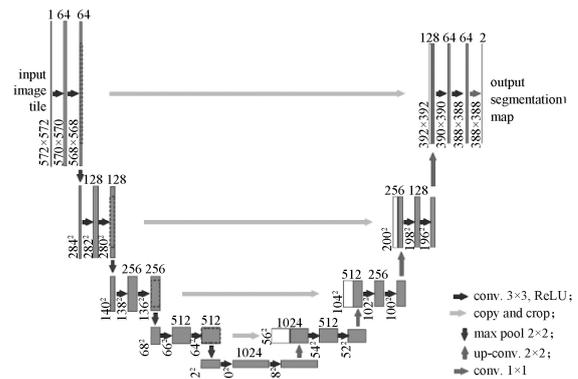


图 2 U-Net 网络结构图

Fig. 2 U-Net structural diagram

U-Net 网络的优点是没有全连接层，是对称的跳跃连接对。当进行上采样操作时，利用对应的下采样层信息使深层卷积具有浅层的简单特征，最后得到的结果能更大程度地还原原图信息。

2.3 编码器 E

编码器 E 采用 Resnet^[9] 网络结构。Resnet 网络又称为残差网络，其核心思想是在网络中增加了直连通道，这样即可以使原始的输入信息能够直接传入后面的层中，又使后面的层无需再学习前面一层所输出的所有信息。Resnet 网络的原理如图 3 所示。图中，X 是浅层的输出，F(X) 是对 X 的补充，两者相加即可得到深层的输出。

Resnet 网络可以较好地解决传统 CNN 网络在学习过程中存在的信息丢失、梯度消失、梯度爆炸等问题。因此，与传统 CNN 网络相比，Resnet 网络可以训练更深层次的网络，学习到更多的信息。

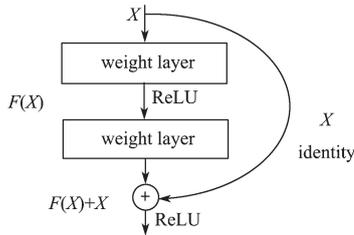


图3 Resnet 网络原理图

Fig. 3 Resnet network schematic diagram

2.4 全局鉴别器与局部鉴别器

鉴别器组有两个部分, 分别是用于判断整幅图像的全局鉴别器 GD 以及用于判断图像局部的鉴别器 LD^[9]。

经过多层卷积层后传统判别器输出的是 0~1 之间的标量值, 而 PatchGAN 输出的是 $n \times n$ 的矩阵, 其可以学习更多的细节信息, 综合整幅图像不同部分的影响, 从而使判断更加精准。因此, 本文的两个鉴别器均采用 PatchGAN^[1-3,7]。PatchGAN 的核心思想是将输入图像裁成多个大小为 70×70 的 patch, 每个 patch 卷积之后输出一个 0~1 之间的值 X_{ij} , 即当前 patch 中样本为真实的概率, 当所有的 patch 完成卷积之后, 计算 X_{ij} 的均值, 得到最终的判断结果。

根据式 (4) 可知, 鉴别器组的输入图像为 B_{fake} 。其中, 全局鉴别器 GD 的输入是完整的 B_{fake} , 全局鉴别器 GD 的损失函数 l_{GD} 为

$$l_{\text{GD}} = E_{A,B} [\log(\text{GD}(A, B_{\text{fake}}))] + E_{A,B,z} [\log(1 - \text{GD}(A, G(A, z)))] \quad (5)$$

局部鉴别器 LD 的输入是 1/4 个 B_{fake} , 记为 B_{fc} 。 B_{fc} 是以 crop 函数随机检测到的边缘点为中心, 剪裁大小为 128×128 的图像。对 A 和 B 进行同理剪裁, 得到的图像分别记为 A_{crop} 和 B_{crop} 。局部鉴别器 LD 的损失函数 l_{LD} 为

$$l_{\text{LD}} = E_{A_{\text{crop}}, B_{\text{crop}}} [\log(\text{LD}(A_{\text{crop}}, B_{\text{crop}}))] + E_{A_{\text{crop}}, B_{\text{fc}}} [\log(1 - \text{LD}(A_{\text{crop}}, B_{\text{fc}}))] \quad (6)$$

因此, 鉴别器组的损失函数 l_{D} 为

$$l_{\text{D}} = \alpha_{\text{GD}} l_{\text{GD}} + \beta_{\text{LD}} l_{\text{LD}}, \quad (7)$$

式中 α_{GD} 、 β_{LD} 为权重, $\alpha_{\text{GD}}=0.75$, $\beta_{\text{LD}}=0.25$ 。

由于局部鉴别器 LD 的输入是以检测到的边缘点作为中心的 1/4 图像, 因此局部鉴别器 LD 可以获取生成图像的边缘特征、局部特征, 从而更好地配合全局鉴别器 GD 进行鉴别工作, 改善生成图像的局部细节。采用双鉴别器的好处在于, 既可在保证多样性的前提

下, 使得生成图像的局部可以产生更加逼真的纹理, 特征更加明显, 又可以使得生成的图像在局部和全局上保持较好的一致性, 不会显得突兀。

2.5 损失函数

由图 2 可知, 双鉴别器 GAN 是一个循环结构, 理论上实现了从目标图像 B 到编码 z , 再到图像 B_{fake} , 最后到目标 z' 的循环。

1) 从目标图像 B 到编码 z , 再到图像 B_{fake} ^[10] 的部分, 编码器 E 直接将目标图像 B 映射到潜在 z 空间, 得到潜在编码 z 的分布 $Q(z|B)$, 且 $Q(z|B) \hat{=} E(B)$, 再结合边缘图像 A 经生成器 G 生成图像 B_{fake} , 最后由鉴别器组判断 B_{fake} 的真实性。此过程的损失函数 $l_{B \rightarrow B_{\text{fake}}}$ 为:

$$l_{B \rightarrow B_{\text{fake}}} = E_{A, B \sim p(A, B)} [\log(D(A, B))] + E_{A, B \sim p(A, B), z \sim E(B)} [\log(1 - D(A, G(A, z)))] \quad (8)$$

此外, 为了使生成器 G 的输入与输出相匹配和训练更加稳定, 采用函数 $l' = E_{A, B \sim p(A, B), z \sim E(B)} \|B - G(A, z_1)\|$ 平衡目标图像 B 和图像 B_{fake} 之间的损失。当 B 未知时, 添加 KL 散度损失^[3,6] 使编码 z 的潜在分布接近随机高斯分布, 即

$$l_{\text{KL}} = E_{B \sim p(B)} [D_{\text{KL}}(E(B) \| N(0, 1))] \quad (9)$$

式中 $D_{\text{KL}}(p \| q) = - \int p(z) \log \frac{p(z)}{q(z)} dz$ 。

从目标图像 B 到图像 B_{fake} 的目标函数为

$$\min_{G, E} \max_D V(G, E) = l_{B \rightarrow B_{\text{fake}}}(G, D, E) + \lambda l'(G, E) + \lambda_{\text{KL}} l_{\text{KL}}(E) \quad (10)$$

2) 从编码 z 到图像 B_{fake} , 再到目标 z' ^[11-13] 的部分, 生成器 G 根据边缘图像 A 与随机噪声 z_1 生成图像 B' , 编码器 E 则利用随机抽取的噪声 z_1 和边缘图像 A 获得 z' 的目标分布, 即 $z' = E(G(A, z_1)) = E(B')$ 。此时, 随机噪声 z_1 与目标 z' 之间的损失函数 $l_{z_1 \rightarrow z'}$ 为

$$l_{z_1 \rightarrow z'} = E_{A \sim p(A), z_1 \sim p(z)} \|z - E(G(A, z_1))\|; \quad (11)$$

编码 z 与目标 z' 的损失函数 $l_{z \rightarrow z'}$ 为

$$l_{z \rightarrow z'} = E_{A, B \sim p(A, B)} [\log(D(A, B))] + E_{A \sim p(A), z_1 \sim p(z)} [\log(1 - D(G(A, z_1)))] \quad (12)$$

目标函数为

$$\min_{G, E} \max_D V(G, E) = l_{z \rightarrow z'}(G, D) + \lambda_{z_1 \rightarrow z'} l_{z_1 \rightarrow z'}(G, E) \quad (13)$$

完成整个循环后, 双鉴别器 GAN 的目标函数为

$$\min_{G,E} \max_D V(G,E) = I_{B \rightarrow B_{fake}}(G,D,E) + \lambda I'(G,E) + \lambda_{KL} I_{KL}(E) + I_{z \rightarrow z'}(G,D) + \lambda_{z_1 \rightarrow z_2} I_{z_1 \rightarrow z_2'}(G,E) \quad (14)$$

3 实验结果与分析

本文将双鉴别器 GAN 应用于包装类产品外观设计设计中, 以产品的边缘图像和目标图像作为成对的输

入, 完成从一张边缘图像映射到多张实物图像的任务。实验条件为 Linux 18.04 系统 + Pytorch 1.0 平台, 训练采用的数据集是来自 iGAN-project 的 137 000 张不同材质的手提包和拉杆箱图像。输入图像是大小为 256×256 的目标图像 B 和与其对应的边缘图像 A , 输出图像是与目标图像相似但不完全相同的图像。本算法最后输出 10 张真实性概率最高的图像。部分测试数据的单号实验结果如表 1 所示。

表 1 双鉴别器 GAN 辅助产品外观设计的实验结果

Table 1 Experimental results of GAN assisted packaging appearance design with double discriminator



彩图

产品名称	边缘图像 A	目标图像 B	输出图像序号				
			1	3	5	7	9
布制手提袋							
塑料环保袋							
纸质手提礼品盒							
纸质手提袋							
PC 拉杆箱							

由表 1 可以看出: 利用本文提出的双鉴别器 GAN 辅助包装类产品外观设计的效果较好, 生成图像不仅多样且边缘清晰、色彩丰富; 相较于目标图像, 布制手提袋的输出图像拥有更多的颜色以及不同的花纹; 塑料环保袋的输出图像不仅较完整地保留了目标图像的一些细节信息, 并在此基础上, 尝试给出了不同的配色参考; 纸质手提礼品盒的输出图像提供了更加丰富的纸张颜色以及不同纹样的搭配方案; 纸质手提袋的输出图像中“LISA”字样的品牌信息依旧清晰可见, 并且给出了不同的色彩和花纹设计;

拉杆箱的箱面材质为 PC, 其输出图像色彩更加丰富并且色彩搭配也十分合理, 不显突兀。

本文分析了双鉴别器 GAN、BicycleGAN 以及 Pix2Pix 分别用于包装类产品外观设计的效果。由于 BicycleGAN 与双鉴别器 GAN 的输出图像无法一一对应, 本文选择了两组测试图像的单号输出图像进行对比。以拉杆箱为测试对象, 3 种算法的单号输出图像如表 2 所示。以纸质手提礼品盒为测试对象, 3 种算法的单号输出图像如表 3 所示。表中圆圈标识部分存在局部边缘特征不明显、阴影、模糊等情况。

由表 2 和表 3 可以看出, BicycleGAN 的输出图像存在局部边缘特征不明显、阴影、模糊情况, 图像效果较差; Pix2Pix 的输出图像只有一张, 且存在色彩模糊、着色不均匀的情况; 双鉴别器 GAN 的输出图像, 同样存在一定程度的颜色缺失以及边缘模糊情

况, 但是相较于 BicycleGAN 和 Pix2Pix 的输出结果, 双鉴别器 GAN 的输出图像的边缘更加清晰, 着色更均匀, 且一定程度上改善了局部阴影问题, 视觉效果得到提升。

表 2 拉杆箱的实验结果对比表

Table 2 Comparison table of experimental results of draw bar box



彩图

算 法	输出图像序号				
	1	3	5	7	9
双鉴别器 GAN					
BicycleGAN					
Pix2Pix					

表 3 纸质手提礼品盒实验结果对比表

Table 3 Comparison table of experimental results of paper portable box



彩图

算 法	输出图像序号				
	1	3	5	7	9
双鉴别器 GAN					
BicycleGAN					
Pix2Pix					

表4和表5是以拉杆箱为测试对象的3种算法输出图像的峰值信噪比(peak signal to noise ratio, PSNR)和结构相似性(structural similarity index, SSIM)对比。较高的PSNR值和SSIM值表示输出图像的质量较好。

表4 拉杆箱实验结果的PSNR值对比

Table 4 Comparison of PSNR values of draw bar box test results

算 法	输出图像序号					平均值
	1	3	5	7	9	
双鉴别器 GAN	12.1557	11.5171	13.8530	14.5744	12.1253	12.8451
BicycleGAN	15.4217	11.1981	12.8637	12.8038	11.6372	12.7849
Pix2Pix			13.3513			13.3513

表5 拉杆箱实验结果的SSIM值对比

Table 5 Comparison of SSIM values of draw bar box test results

算 法	输出图像序号					平均值
	1	3	5	7	9	
双鉴别器 GAN	0.5654	0.6580	0.5891	0.6638	0.6610	0.6275
BicycleGAN	0.7772	0.6493	0.5811	0.6020	0.5919	0.6403
Pix2Pix			0.6310			0.6310

由表4和表5可知: BicycleGAN的输出图像1的PSNR值和SSIM值最高, 追溯至表1与表2则不难发现, 输出图像1与原图B十分相像; 而双鉴别器GAN的其他输出图像的PSNR值和SSIM值略高于BicycleGAN; Pix2Pix的输出图像过于单一。

表6和表7是以纸质手提礼品盒为测试对象的3种算法输出图像的PSNR值以及SSIM值对比。

表6 纸质手提礼品盒实验结果PSNR值对比

Table 6 Comparison of PSNR values of paper portable gift box

算 法	输出图像序号					平均值
	1	3	5	7	9	
双鉴别器 GAN	14.6266	14.4120	10.5327	12.7210	14.5207	13.3626
BicycleGAN	11.3350	14.2970	10.7594	11.5579	11.7656	11.9430
Pix2Pix			12.4842			12.4842

表7 纸质手提礼品盒实验结果SSIM值对比

Table 7 Comparison of SSIM values of paper portable gift box

算 法	输出图像序号					平均值
	1	3	5	7	9	
双鉴别器 GAN	0.5735	0.8206	0.6657	0.5296	0.6313	0.6441
BicycleGAN	0.5437	0.6837	0.5714	0.3668	0.6939	0.5719
Pix2Pix			0.3192			0.3192

由表6和表7可知, 相比BicycleGAN和Pix2Pix,

双鉴别器GAN的输出图像的PSNR值和SSIM值略高一些。

综合上述结论可得, 将双鉴别器GAN应用于包装类产品的外观设计是可行的。

4 结语

针对包装类产品外观设计, 本文在BicycleGAN的基础上, 添加了一个局部鉴别器, 并重新构建了相应的损失函数以及目标函数, 提出了双鉴别器GAN。实验结果表明: 相较于BicycleGAN、Pix2Pix, 双鉴别器GAN提高了输出图像的边缘清晰度以及局部细节, 但是同样存在颜色缺失等问题; 将其应用于包装类产品的外观设计中是可行的。一方面, 多样性的输出结果可以作为灵感来源, 为设计工作提供更加丰富的参考; 另一方面, 由于输出结果拥有了更明显的局部细节特征, 因此也可以简化相对较为繁琐的包装外观设计工作, 甚至可直接进行后续的个性化设计或者修改。这样不仅大大缩短了设计周期, 还能进一步节省人力物力, 降低成本, 使得设计工作更加轻松高效。

由于实验条件有限, 本文缺少一定的主观评价, 这需要继续深入研究; 再者, 双鉴别器GAN是有监督的学习, 其对于数据集的要求相对较高, 若要辅助完成较为复杂的外观设计工作, 目前尚存在困难, 这也是今后本课题组需要解决的难点之一。

参考文献:

- [1] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks[C]//Advances in Neural Information Processing Systems NIPS. Montreal: MIT Press, 2014: 2672-2680.
- [2] 王硕诚. GAN在图像翻译上的研究与应用[D]. 贵阳: 贵州大学, 2019.
WANG Shuocheng. Research and Application of GAN in Image Translation[D]. Guiyang: Guizhou University, 2019.
- [3] 侯学东. 基于生成对抗网络的图像翻译[D]. 南京: 南京邮电大学, 2018.
HOU Xuedong. Image Translation Based Generative Adversarial Networks[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2018.
- [4] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-Image Translation with Conditional Adversarial

- Networks[C]//2017 IEEE International Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 5967-5976.
- [5] 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. Venice: IEEE, 2017: 2242-2251. DOI: 10.1109/ICCV.2017.244.
- [6] ZHU J Y, ZHANG R, PATHAK D, et al. Toward Multimodal Image-to-Image Translation[EB/OL]. [2020-01-11]. https://www.researchgate.net/publication/321417717_Toward_Multimodal_Image-to-Image_Translation.
- [7] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and Locally Consistent Image Completion[J]. ACM Transactions on Graphics, 2017, 36(4): 1-14.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: MICCAI, 2015: 234-241.
- [9] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. DOI:10.1109/CVPR.2016.90.
- [10] HINTON G E. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786): 504-507.
- [11] CHEN X, DUAN Y, HOUTHOOFT R, et al. InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets[EB/OL]. [2020-01-05]. <https://arxiv.org/abs/1606.03657>.
- [12] DONAHUE J, KRÄHENBÜHL P, DARRELL T. Adversarial Feature Learning[EB/OL]. [2020-01-05]. <https://arxiv.org/abs/1605.09782v6>.
- [13] DUMOULIN V, BELGHAZI I, POOLE B, et al. Adversarially Learned Inference[EB/OL]. [2019-04-06]. https://www.researchgate.net/publication/303755744_Adversarially_Learned_Inference.

(责任编辑: 邓 彬)

Appearance Design Method of Packaging Product Based on Dual Discriminator GAN

LI Jiaxin, SUN Liujie, WANG Wenju

(College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai 200093)

Abstract: To solve the problems of long design cycle, great design difficulty and relatively high design cost in the traditional packaging product design, the dual discriminator GAN was applied to the packaging product design. On the basis of BicycleGAN, a local discriminator was added, and the corresponding loss function and objective function were reconstructed. In the experiment, the input object was the image pair of "edge image + original image", and the output object was the image with the top 10 authenticity probability. The experimental results were evaluated from three aspects: diversity, PSNR value and SSIM value. The results showed that the image generated by dual discriminator Gan not only had better diversity, but also improved the image detail quality. Therefore, it was feasible to apply the dual discriminator Gan to the appearance design of packaging products. On one hand, it could provide designers with design inspiration, on the other hand, it could save human and material resources, and improve working efficiency.

Keywords: dual discriminator GAN; appearance design; image translation