

# 融合特征增强的频谱知识蒸馏<sup>①</sup>

李睿, 龚啸威

(兰州理工大学 计算机与通信学院, 兰州 730050)

通信作者: 龚啸威, E-mail: [xwxw0505@163.com](mailto:xwxw0505@163.com)



**摘要:** 现有生成对抗网络压缩方法通常更侧重于网络架构和空间域的优化, 忽视了频谱域优化对蒸馏效果和模型性能的影响. 这种局限性导致轻量化模型在生成图像的高频细节时, 与教师模型之间存在一定的差异; 同时, 在图像转换任务中, 现有的特征提取方法通常会导致图像细节丢失. 针对这些问题, 提出了一种融合特征增强的频谱知识蒸馏 (FESD-CycleGAN) 方案. 在 FESD-CycleGAN 中, 首先通过对特征图的部分特征通道进行偏移, 扩大感受野, 增强特征多样性, 从而提升了生成图像的细节和整体表现. 其次基于对频谱域进行知识蒸馏能够使生成器捕捉图像的高频细节这一特性, 提出在对特征图进行特征增强的基础上, 将空间域与频谱域的知识蒸馏相结合, 从而增强了模型对生成图像细节的把控能力. 实验结果表明, 在 horse2zebra、summer2winter 和 edges2shoes 数据集上, FESD-CycleGAN 与基线模型 DCD 相比, *FID* 值分别降低了 2.19、0.68 和 0.76, 达到 54.98、73.41 和 27.45, 有效提升轻量化模型的生成性能.

**关键词:** 特征增强; 频谱知识蒸馏; CycleGAN; 模型压缩; 图像增强

引用格式: 李睿, 龚啸威. 融合特征增强的频谱知识蒸馏. 计算机系统应用, 2026, 35(1): 76-87. <http://www.c-s-a.org.cn/1003-3254/10040.html>

## Spectral Knowledge Distillation with Integrated Feature Enhancement

LI Rui, GONG Xiao-Wei

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** Existing generative adversarial network (GAN) compression methods focus more on optimizing network architecture and the spatial domain, while neglecting the impact of spectral-domain optimization on distillation effectiveness and model performance. This limitation results in discrepancies between lightweight models and teacher models in generating high-frequency image details. In addition, conventional feature extraction methods in image translation often cause detail loss. To address these issues, this study proposes a spectral knowledge distillation scheme with integrated feature enhancement (FESD-CycleGAN). In FESD-CycleGAN, by shifting certain feature channels in the feature map, the receptive field is expanded and feature diversity is enhanced, thus improving both the details and the overall quality of generated images. Moreover, since spectral-domain knowledge distillation enables the generator to capture high-frequency details of images, knowledge distillation in both the spatial and spectral domains is integrated on top of feature enhancement in the feature map. This approach enhances the model's ability to preserve fine details in generated images. Experimental results show that on the horse2zebra, summer2winter, and edges2shoes datasets, FESD-CycleGAN reduces the *FID* by 2.19, 0.68, and 0.76 compared to the baseline DCD model, achieving scores of 54.98, 73.41, and 27.45, respectively. The generative performance of lightweight models is effectively improved by FESD-CycleGAN.

**Key words:** feature enhancement; spectral knowledge distillation; CycleGAN; model compression; image enhancement

<sup>①</sup> 收稿时间: 2025-05-28; 修改时间: 2025-06-20, 2025-07-10; 采用时间: 2025-07-29; csa 在线出版时间: 2025-10-29  
CNKI 网络首发时间: 2025-10-30

生成对抗网络 (generative adversarial network, GAN)<sup>[1]</sup>模型由两个核心组件构成: 判别器和生成器. 生成器的目的是生成逼真的假样本, 尽量使其看起来像真实数据, 目的是去骗过判别器. 而判别器的目的是去区分真实的样本和生成的样本, 最终达到一个纳什均衡的状态. 随着训练的进行, 生成器生成的样本会越来越像真实数据, 而判别器则越来越难以分辨出真实样本和生成的样本. 这个对抗性的训练过程最终使生成器学会生成逼真的样本. 目前图像到图像转换的 GAN 已经成功地应用于图像合成<sup>[2]</sup>、风格迁移<sup>[3]</sup>、图像编辑<sup>[4]</sup>和图像增强<sup>[5]</sup>. 然而, 由于 GAN 的超高计算和内存开销使得在实际部署中, 尤其是对于移动设备或嵌入式系统等资源有限的场景, 使用如此庞大的模型变得不现实. 为了解决这一问题, 人们已经研究了一系列的模型压缩方法, 如剪枝、量化、低秩分解、知识蒸馏和神经网络架构搜索 (NAS) 等, 来减少模型的计算量和存储空间. 其中, 知识蒸馏, 特别是特征图的蒸馏, 被认为是提高轻量化生成器性能的有效方法. 在知识蒸馏过程中, 将教师模型的中间激活信息作为一种知识来源, 通过提取并传递这些特征给学生模型, 使学生模型学习到教师模型的中间表示, 这已经被证明可以提高学生模型的性能和泛化能力.

近年来, 随着 GAN 的快速发展, 出现了许多用于 GAN 压缩方法的研究工作并取得了显著成果, 特别是应用在图像到图像的转换模型中. 例如, Wang 等人<sup>[6]</sup>提出了 GAN-Slimming, 在极小极大优化框架下, 通过集成模型蒸馏、剪枝和量化这 3 种主流的压缩技术, 用于 GAN 压缩并获得了先进的压缩结果. Li 等人<sup>[7]</sup>提出了 GCC, 考虑到了鉴别器在压缩中所能起到的作用, 设计了一种生成器与判别器协同压缩的方法并证明了其有效性. Gao 等人<sup>[8]</sup>提出了 ACKD, 一种从教师生成器的各个阶段积累知识, 然后将其转移到学生生成器的基于积累知识蒸馏的压缩方法, 这种压缩方法表现出了卓越的性能. 尽管这些方法在 GAN 压缩领域取得了重要进展, 但是现有压缩方法多局限于网络架构和空间域的优化, 而忽略了频谱域优化对生成图像质量和蒸馏效果的潜在影响. 此外, 如何有效地对不同层次的特征进行增强并将其转移到学生模型, 也是影响知识蒸馏效果的重要因素. 基于此, 本文从特征增强和频谱域的角度, 研究了知识蒸馏在图像转换应用中效果的影响因素.

在图像转换任务中, Lee 等人<sup>[9]</sup>表明, 简单地进行特征提取往往会导致细节丢失, 尤其是在边缘和纹理等区域, 并且模型的泛化能力较差. 因此, 对特征进行增强可以有效提升特征表示, 从而改善生成图像的质量. 此外, Frank 等人<sup>[10]</sup>的研究表明, 图像生成任务中的频谱差异主要集中在高频区域, 而这些高频数据对细节的准确性起着重要作用, 因此忽视这些差异可能会导致生成图像的质量下降, 尤其是在注重细节的生成任务中. Durall 等人<sup>[11]</sup>也表明, 通过分析高频数据, 可以在频谱域中有效区分生成的图像和真实图像. 相比真实图像, 生成图像通常在高频信息上存在较大的差异, 因此, 缩小这种频谱差异将有助于提升 GAN 在图像转换任务中的性能.

本文结合部分通道偏移和频谱知识蒸馏方法在 CycleGAN 模型的图像转换任务中, 设计了一个方案 (FESD-CycleGAN), 用于增强特征图蒸馏的水平, 提升轻量化学生生成器生成高质量图像的性能. FESD-CycleGAN 的主要贡献如下.

(1) 针对特征增强在改善生成图像质量中的作用, 本文引入了一种简洁而又高效的特征增强方法——部分通道偏移. 该方法能够扩展有效感受野并显著提升特征表示, 同时避免增加额外的网络参数和计算成本, 从而有效提升生成器的性能.

(2) 对于 CycleGAN 的轻量化任务, 在知识蒸馏过程中, 本文基于特征增强, 将频谱域蒸馏与空间域蒸馏相结合, 进一步缩小教师模型与学生模型之间的输出差异.

## 1 相关工作

### 1.1 GAN 和 GAN 压缩

由于 GAN 能够在不同数据域中生成高质量和可扩展的图像, 导致其在图像到图像转换的任务中取得了显著的成功. 具体来说, Isola 等人<sup>[12]</sup>提出的 Pix2Pix 模型用于训练配对数据的图像到图像的转换任务. 如将线条图的鞋转换为实物鞋的图像. CycleGAN<sup>[13]</sup>作为一种无配对图像到图像转换的模型, 通过循环一致性损失和对抗训练实现了高效的图像生成. 尽管 GAN 的各种方法取得了显著进展, 但其高内存和计算开销限制了在资源受限设备上的应用. 为此, 近年来研究者们提出了多种面向 GAN 的压缩技术, 旨在降低 GAN 的模型大小和计算复杂度, 同时保持性能和效率. 例如,

对于 CycleGAN 模型的压缩, Shu 等人<sup>[14]</sup>提出了一种基于循环一致性损失的协同进化算法来对不重要的权值进行修剪. Li 等人<sup>[15]</sup>采用神经架构搜索识别高效网络, 并通过一次性网络进行权值共享来压缩 GAN. Kumar 等人<sup>[16]</sup>提出了一种创新的 GAN 压缩方法, 利用可微剪枝与超网络的结合, 实现了对条件 GAN 的压缩.

### 1.2 知识蒸馏

知识蒸馏是一种模型压缩方法, 旨在通过让较小的学生网络模仿较大的教师网络的输出行为, 将知识从教师网络有效传递到学生网络中. 这一思想最早是由 Buciluă 等人<sup>[17]</sup>提出的, 然后由 Hinton 等人<sup>[18]</sup>进一步提出了知识蒸馏的概念, 并在 Softmax 中添加了一个温度超参数, 用于平滑教师网络的 logits 分布. Zhou 等人<sup>[19]</sup>提出特征图的每个通道对应一个视觉模式, 因此通过通道注意力将教师网络的通道信息传递到学生网络中, 并且只让学生网络学习教师网络的正确输出. Ren 等人<sup>[20]</sup>提出的是一种在线多粒度蒸馏方案, 该方法中的学生生成器是通过逐步优化的教师生成器进行蒸馏, 不需要使用单独的判别器. 同时, 在蒸馏过程中, 学生生成器从教师模型的多个网络层和语义层次中学习多维度的知识.

### 1.3 特征偏移

在图像转换领域, 如何设计参数更少、计算复杂度更低的模型是一个关键的研究方向. 相比于常用的卷积操作会带来一定的参数和计算成本, 特征偏移则旨在彻底消除这些额外的计算成本和参数负担. 因此, 在效率与性能的权衡中, 特征偏移比门控机制<sup>[21]</sup>等复杂设计表现得更为出色. 特征偏移已经被广泛应用于多个领域. 例如, 在视频分析中, 特征偏移被用于时序偏移模块 (TSM)<sup>[22]</sup>等模型, 以提升视频分类和动作识别的性能. 此外, 在目标检测和图像分类任务中, AddressNet<sup>[23]</sup>通过动态空间偏移增强了特征信息的利用, 从而在复杂背景中能更准确地识别目标. 然而, 在图像到图像转换任务中, 特征偏移的潜力尚未得到充分挖掘. 在本文中将特征偏移应用于 CycleGAN 模型, 以提升生成图像的质量.

### 1.4 CNN 的频率分析

随着 GAN 技术的进步, 图像生成的表现也在不断提升. 近年来的研究发现, 尽管生成图像与真实图像在视觉上越来越相似, 但它们在某些情况下仍然存在一些不可忽视的差异. 此外, 频谱分析在图像处理上也起

着一定的关键作用, 但是当前 GAN 模型生成的图像在频谱分布上始终难以达到与真实数据一致的效果. Zhang 等人<sup>[24]</sup>首次揭示了 GAN 生成管道中的上采样模块会引发特定的频谱伪影. 此外, Rahaman 等人<sup>[25]</sup>的研究表明, 神经网络对低频成分的偏向导致了频谱上的选择性偏差, 这种偏向也解释了其泛化能力的来源. 然而, 高频数据对于注重细节的图像生成任务尤为重要. 为了应对这一问题, Cai 等人<sup>[26]</sup>提出了一种频域图像转换框架 FDIT, 该方法通过分离图像中的低频和高频信息, 引入频率特征来优化图像生成, 使生成的图像更具真实性. 近年来, 虽然频谱域已得到广泛应用, 但是在图像生成领域中仍然需要进一步的研究.

## 2 方法

FESD-CycleGAN 模型框架如图 1 所示. FESD-CycleGAN 的框架结构与 CycleGAN 类似, 主要由教师生成器 (teacher generator,  $G^T$ )、学生生成器 (student generator,  $G^S$ ) 和教师判别器 (teacher discriminator,  $D^T$ ) 组成. 与 CycleGAN 框架的主要区别在于, FESD-CycleGAN 框架中增加了特征增强模块 (feature enhancement). 具体来说, 该模块对学生生成器和教师生成器生成的特征图进行了特征增强, 并在判别器中引入了频谱分类器 (spectral classifier,  $SC$ ), 以用来测量学生生成器和教师生成器生成的特征图之间的频谱差异, 从而实现频谱域的知识蒸馏. 该模型能够通过扩大特征图的有效感受野来提高生成器的性能, 同时也进一步缩小教师模型与学生模型的输出差距, 从而提高生成图像的质量.

FESD-CycleGAN 模型的工作流程为: 对于来自域  $X$  的图像, 目标是将其转换为对应的域  $Y$  图像 (例如, 马到斑马的转换: 马为域  $X$  的图像, 斑马为域  $Y$  的图像). 首先, 将域  $X$  中的图像分别输入  $G^S$  和  $G^T$ , 生成中间特征图. 接着, 将来自学生和教师生成器的中间特征图进行下采样以对齐维度, 然后对特征图进行特征增强. 随后, 再将增强后的特征图输入到  $D^T$  和  $SC$  中, 分别进行空间域以及频谱域的知识蒸馏, 以实现在空间域和频谱域上生成感知逼真的图像. 对于将域  $Y$  图像转换为域  $X$  图像的过程, 与上述过程类似. 通过上述过程的充分训练, FESD-CycleGAN 模型可以生成更高质量的图像.

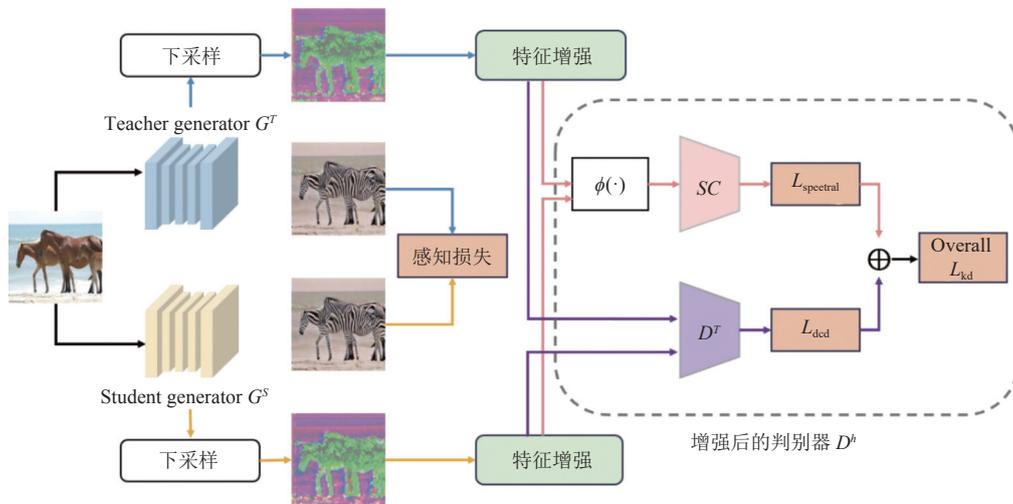


图1 FESD-CycleGAN 模型框架

### 2.1 特征偏移

简单地进行特征提取往往难以捕捉细节信息,尤其是在边缘和纹理等区域,容易导致生成图像细节丢失.在图像转换任务中,提升特征表示的有效性能显著改进模型的性能和生成图像的质量.基于此,Zhang等人<sup>[27]</sup>提出了部分通道偏移(PCS)方法,具体而言,PCS方法通过在两个空间维度  $H$  和  $W$  上对部分特征通道进行移动,打破原本整齐对齐的特征分布,使其在空间上产生一定的偏移.移位后的特征网格中,移除的元素会被直接舍弃,而空出的网格位置则用零来填补,以保持特征网格的结构完整性.通过打破原有对齐结构,PCS增加了特征图的空间变换,使模型能够更好地捕捉广泛的空间信息和多样的特征表达,进而有效扩大感受野并增强特征多样性.此外,PCS方法可以在不增加额外计算成本的情况下,提升生成图像的细节和整体表现.因此,为了提高特征表示能力,将PCS方法引入到模型结构中,通过扩大感受野,使模型能够更准确地获取局部区域的特征关系,从而提升压缩后模型的性能,生成高质量的图像.如图2所示,对部分特征通道进行偏移可以扩大有效感受野,从而增强对特征信息的提取.其中,左图是由多个特征通道组成的原始特征图.对于每个通道,通过  $3 \times 3$  的卷积核在特征图上捕获一个  $3 \times 3$  的感受野;右图是通过引入PCS操作,扩展了有效感受野范围.

为了从理论上分析PCS对感受野的影响,假设输入特征图为  $F$ ,卷积核尺寸为  $k \times k$ ,卷积核半径为  $r$ .对于特征图中某个像素位置  $(i, j)$ ,未偏移通道对应的感

受野为以  $(i, j)$  为中心的  $k \times k$  邻域范围,如式(1)所示:

$$R_{\text{normal}} = [i-r, i+r] \times [j-r, j+r], r = \lfloor k/2 \rfloor \quad (1)$$

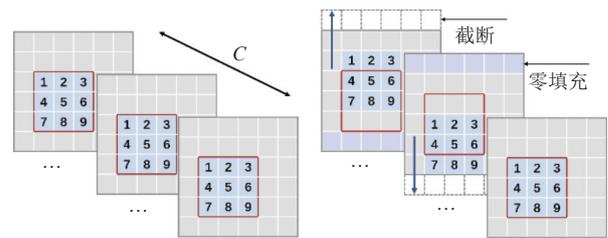


图2 特征偏移

当对部分通道施加空间偏移量  $(\Delta h, \Delta w)$  时,这些通道对应的感受野如式(2)所示:

$$R_{\text{shifted}} = [i-r+\Delta h, i+r+\Delta h] \times [j-r+\Delta w, j+r+\Delta w] \quad (2)$$

结合未偏移通道与偏移通道后的有效感受野为两者的并集,如式(3)所示:

$$R_{\text{effective}} = R_{\text{normal}} \cup R_{\text{shifted}} \quad (3)$$

这一操作能够显著扩展模型的感受野范围,但需要注意的是,当偏移幅度过小时(如  $\Delta h=0.5$ ),有效感受野扩展有限;当偏移幅度过大时(如  $|\Delta h| \gg r$ ),会导致  $R_{\text{normal}}$  与  $R_{\text{shifted}}$  几乎无重叠,破坏空间局部性.

因此,通过移动部分通道的方法,可以有效扩大感受野,从而提升生成图像的细节和整体质量,但这一方法也存在着一些问题.例如,为了保持特征网格的结构完整性,直接将移位后的特征元素舍弃,并用零填充空出的位置,可能会导致偏移通道中的边界信息丢失;如果特征偏移的幅度过大,可能会导致特征的空间局部

性遭到破坏;虽然偏移只作用于部分通道,但仍然会产生大量的偏移状态,而完全枚举这些状态是不现实的;此外,偏移通道在所有通道中的排列方式也是一个问题.为此,针对这些问题需要对偏移的幅度、方向、偏移通道的数量以及偏移通道在所有通道中的排列方式等进行合理的约束,从而使偏移操作更加简洁高效.

### 2.1.1 偏移幅度

对于一个形状为 $H \times W \times C$ 的中间特征 $x$ , $H$ 、 $W$ 、 $C$ 分别表示特征图的高度、宽度以及通道数.定义了一种部分通道偏移操作,操作结果为 $\bar{x}$ ,计算公式如式(4)所示.

$$\bar{x} = Tc(x, \theta) \tag{4}$$

其中, $Tc(\cdot)$ 表示通道偏移操作, $\bar{x}$ 表示经过舍弃移位特征元素和零填充后的增强特征,并且其形状与原特征 $x$ 的形状相同. $\theta$ 用于定义特征 $x$ 中 $C$ 个通道的偏移模式集合.例如, $x_i (1 \leq i \leq C)$ 代表特征图的第 $i$ 个通道,那么对应的偏移模式为 $\theta_i$ .

偏移模式 $\theta_i$ 包括在特征通道的高度方向的偏移 $h_i (1 \leq i \leq C)$ 以及宽度方向的偏移 $w_i (1 \leq i \leq C)$ ,并且满足式(5)中的条件.

$$-H \leq h_i \leq H, -W \leq w_i \leq W \tag{5}$$

根据上述分析,对于单个通道的特征图,如果不对偏移方式进行任何限制,那么所有可能的偏移情况的数量会达到 $(2H+1) \times (2W+1)$ .为了减少可能的偏移情况,为偏移幅度设置了限制阈值 $L_m$ ,其中 $0 < L_m < H$ 且 $0 < L_m < W$ .此时,偏移约束条件为 $|h_i| < L_m$ 以及 $|w_i| <$

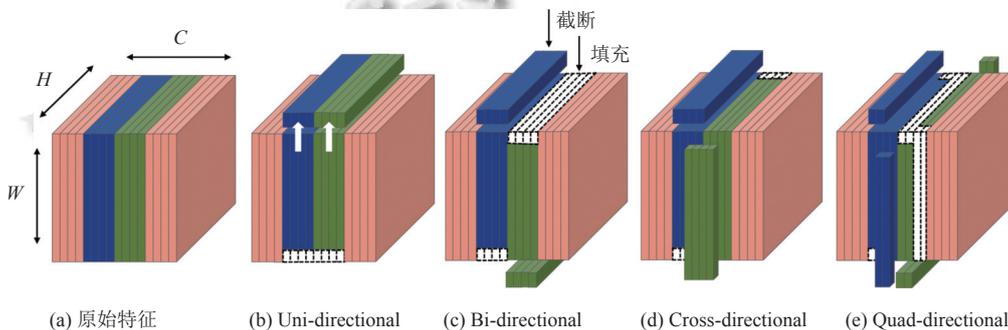


图4 应用 PCS 到中间特征的 4 种偏移方向

## 2.2 频谱知识蒸馏

为了解决生成对抗网络中的判别器对高频信息处理不足的问题,本文鉴于 SSD-GAN<sup>[28]</sup>对判别器结构的

$L_m$ ,从而将可能的偏移情况减少至 $(2L_m+1)^2$ .并且在 $L_m \ll H$ 和 $L_m \ll W$ 的条件下,这种方式能够显著减少单个通道的偏移情况.

### 2.1.2 偏移数量

偏移过多的通道数可能会导致边界信息丢失,所以为偏移数量也设置了限制阈值 $L_c$ ,其中, $0 < L_c < C$ .并且定义了通道偏移率 $\alpha = N/C$ ,其中 $N \leq L_c$ 表示在具有 $C$ 个通道的特征图中被偏移的通道数量,以更清晰地描述当通道数量 $C$ 发生变化时的部分通道偏移.

### 2.1.3 排列方式

为了保持特征的空间局部性并减少额外计算开销,采用了一种简化策略来解决偏移通道的排列问题.具体来说,将偏移通道集中处理,并在最终排列中以连续的形式插入到特定位置,而非与未偏移通道交替排列.本文的策略是将偏移通道统一插入中间位置,形成一个连续的偏移通道组,如图3所示.实验证明,这种简化的排列方式在不增加额外计算开销的前提下,有效提升了模型性能.

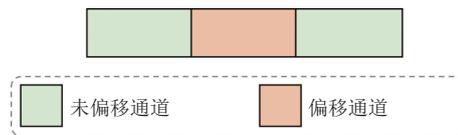


图3 排列方式

### 2.1.4 偏移方向

本文将偏移方向简化成了4种模式,如图4所示.图中展示的不同特征移位增强了有效感受野,同时提高了特征的多样性.

对判别器的这种改进,可以激励生成器减少生成图像和真实图像的频谱差异,从而提高生成图像的质量,同时不会显著增加计算成本.此外,将源自 $G^S$ 和 $G^T$ 的增强特征图输入到 $SC$ 中,计算特征图的频谱蒸馏损失 $L_{\text{spectral}}$ ,并将 $L_{\text{spectral}}$ 整合到总蒸馏损失 $L_{\text{kd}}$ 中,进一步增强了生成效果.因此,本文通过提出频谱知识蒸馏的机制,解决了判别器中高频信息缺失问题.具体来说,通过在原始判别器 $D^T$ 中加入 $SC$ ,测量来自 $G^T$ 和 $G^S$ 的特征图之间的频谱差异.该方法在增强特征图的基础上,将空间域的知识蒸馏与频谱域的知识蒸馏相结合,鼓励生成器更好地学习高频数据的分布,从而生成质量更好、细节更丰富的图像.

首先将输入的图像转换为灰度图像 $f(x, y)$ ,然后计算其二维傅里叶变换.计算公式如式(6)所示,将图像从空间域映射到频域,得到对应的频域 $F(u, v)$ .

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (6)$$

接着计算频谱幅值,其计算公式如式(7)所示:

$$|F(u, v)| = \sqrt{F_{\text{real}}^2 + F_{\text{imag}}^2} \quad (7)$$

其中,实部 $F_{\text{real}}$ 和虚部 $F_{\text{imag}}$ 共同描述图像频域特性.

然后对 $|F(u, v)|$ 进行对数压缩,用于缩小频谱幅值的动态范围.再使用频谱平移将低频部分移至中心位置,并通过径向平均将二维频谱特征转换为一维频谱特征.最终,通过对频谱特征进行归一化处理,可以提升数据分布的一致性,提高数据稳定性.

增强后的判别器 $D^h$ 包含两部分,一个是用于评估特征图在空间域中真实度的常规判别器 $D^T$ ,另一个是用于衡量特征图在频谱域中真实度的频谱分类器 $SC$ .因此, $D^h$ 结合了对输入数据在空间域和频谱域上的真实性评估.首先将来自 $G^T$ 和 $G^S$ 的特征图输入到 $D^T$ 中,计算它们在空间域中的差异,以实现空间域的蒸馏,该损失被称为判别器协作蒸馏损失 $L_{\text{dcd}}$ .其次,通过灰度分量函数 $\phi(\cdot)$ 提取这两个特征图的频谱信息,并将其输入 $SC$ 中,计算它们在频谱域中的差异,从而实现高频信息的知识蒸馏,该损失被称为频谱蒸馏损失 $L_{\text{spectral}}$ .最终,将这两个损失结合起来作为总蒸馏损失 $L_{\text{kd}}$ ,具体损失函数在第2.3节给出.

### 2.3 损失函数

本文提出的FESD-CycleGAN模型的损失函数由

对抗损失 $L_{\text{GAN}}$ 、感知损失 $L_{\text{per}}$ 、判别器协作蒸馏损失 $L_{\text{dcd}}$ 和频谱蒸馏损失 $L_{\text{spectral}}$ 构成.接下来将分别介绍这4个损失函数.

#### 2.3.1 对抗损失

GAN的训练目标通常是最小化生成器的损失,最大化判别器的损失.然而,与传统GAN的最小-最大优化方法不同的是,本文没有引入一个独立的学生判别器,而是将教师判别器从零开始训练,以确保其与教师生成器之间的良好匹配.此外,教师判别器还充当协作判别器,用来判断学生生成器的输入是否是真实数据.因此,对抗损失 $L_{\text{GAN}}$ 的计算公式如式(8)所示:

$$\begin{aligned} \min_{G^T, G^S} \max_{D^T} L_{\text{GAN}} = & \mathbb{E}_{y \in Y} [\log D^T(y)] \\ & + \mathbb{E}_{x \in X} [\log(1 - D^T(G^T(x)))] \\ & + \lambda_s \times \mathbb{E}_{x \in X} [\log(1 - D^T(G^S(x)))] \quad (8) \end{aligned}$$

其中, $Y$ 表示真实数据分布, $X$ 表示输入数据分布. $\lambda_s$ 是一个权衡参数,当 $\lambda_s = 0$ 时,学生生成器的对抗损失项消失,意味着学生生成器不再直接与教师判别器进行对抗性训练,从而使得模型可以不再引入独立的学生判别器.

#### 2.3.2 感知损失

感知损失 $L_{\text{per}}$ 用于确保在图像生成任务中能够产生更具视觉吸引力的结果.具体来说, $L_{\text{per}}$ 包括风格重建损失 $L_{\text{sty}}$ 和特征重建损失 $L_{\text{feat}}$ .其计算公式如式(9)所示:

$$L_{\text{per}} = \lambda_{\text{sty}} \cdot L_{\text{sty}} + \lambda_{\text{feat}} \cdot L_{\text{feat}} \quad (9)$$

$L_{\text{sty}}$ 通过计算生成图像与目标图像的Gram矩阵来提取风格特征,从而确保生成图像与目标图像在风格上相似.其计算公式如式(10)所示:

$$L_{\text{sty}} = \sum_{k \in J} \|G(\Phi_k(G^T(x))) - G(\Phi_k(G^S(x)))\|_1 \quad (10)$$

其中,集合 $J$ 包含希望提取的中间层索引, $G(\cdot)$ 为Gram矩阵的简写.

$L_{\text{feat}}$ 使用预训练的VGG网络来使得教师生成器和学生生成器的输出在特征空间中更加接近.其计算公式如式(11)所示:

$$L_{\text{feat}} = \sum_{k \in J} \frac{1}{C_k H_k W_k} \|\Phi_k(G^T(x)) - \Phi_k(G^S(x))\|_1 \quad (11)$$

其中, $C_k$ 、 $H_k$ 和 $W_k$ 分别表示第 $k$ 层特征图的通道数、高度和宽度.函数 $\Phi_k(\cdot)$ 返回从预训练VGG网络

中提取的第  $k$  层的激活输出。

### 2.3.3 判别器协作蒸馏损失

在基于 GAN 的生成任务中, 由于判别器本身具有丰富的信息特征, 因此将其引入了蒸馏过程. 具体来说, 将生成器的输出结果作为判别器的输入, 通过最小化真实图像和生成图像在判别器中间层输出上的差异, 使生成图像在不同层次上更接近真实图像的特征表达. 其计算公式如式 (12) 所示:

$$L_{\text{dcd}} = \sum_{k \in I_D} \sum_{i \in I_G} l(D_k^T(h(f(G_i^T(x))))), D_k^T(h(f(G_i^S(x)))))) \quad (12)$$

其中,  $G_i(x)$  表示生成器第  $i$  层的输出;  $f(\cdot)$  是仿射变换函数, 用于调整教师和学生特征图的通道数, 以匹配判别器输入的通道维度;  $h(\cdot)$  代表特征增强操作;  $D_k^T$  表示教师判别器第  $k$  层的输出;  $l(\cdot)$  是距离度量函数;  $I_G$  表示生成器中选取的中间特征层索引集合, 这些层是教师与学生生成器中需要对齐的部分, 用以控制知识传递的粒度;  $I_D$  表示教师判别器中选取的中间特征层索引集合, 这些中间特征包含图像真实性判别特征, 作为指导信号以引导学生生成器生成更高质量的图像。

### 2.3.4 频谱蒸馏损失

频谱蒸馏损失  $L_{\text{spectral}}$  用于表示教师生成器生成的特征图和学生生成器生成的特征图在频谱域信息上的差距. 具体来说, 通过使用灰度组件函数  $\phi(\cdot)$  提取教师与学生生成器特征图中的频谱信息, 然后使用  $SC$  衡量其频谱真实性. 通过这种方式, 频谱蒸馏损失在训练过程中能够减少学生生成器与教师生成器之间的频谱差异, 从而在频谱域中生成更逼真的图像. 其计算公式如式 (13) 所示:

$$L_{\text{spectral}} = \sum_{k \in I_D} \sum_{i \in I_G} l(SC(\phi(h(f(G_i^T(x))))), SC(\phi(h(f(G_i^S(x)))))) \quad (13)$$

其中,  $\phi(\cdot)$  表示灰度组件, 用于从特征图中提取频谱信息;  $SC(\cdot)$  表示频谱分类器, 用于衡量频谱真实性。

### 2.3.5 总损失

综合上述 4 种损失, 本文的整体损失函数如式 (14) 所示:

$$\min_{G^T, G^S} \max_{D^T} (L_{\text{GAN}} + L_{\text{per}} + \lambda_{\text{dcd}} \cdot L_{\text{dcd}} + \lambda_{\text{spe}} \cdot L_{\text{spectral}}) \quad (14)$$

其中,  $\lambda_{\text{dcd}}$  和  $\lambda_{\text{spe}}$  为平衡损失项. 本文中涉及的 5 个超参数分别为  $\lambda_{\text{dcd}}$ 、 $\lambda_{\text{spe}}$ 、 $\lambda_{\text{fea}}$ 、 $\lambda_{\text{sty}}$  以及  $\lambda_{\text{stu}}$ , 其具体设置均在实验部分详细描述。

## 3 实验

### 3.1 设置

#### 3.1.1 数据集

CycleGAN 在图像转换任务中无需源域和目标域之间存在一一对应的映射, 而是可以在未配对的样本集合上学习两个域之间的映射关系. 因此, 本文在 horse2zebra 和 summer2winter 两个未配对图像转换数据集上对 FESD-CycleGAN 模型进行实验. 而对于 Pix2-Pix, 本文在配对数据集 edges2shoes 上进行实验. 其中, horse2zebra 数据集包含马和斑马两类图像, 旨在通过模型学习马和斑马两个域之间的映射, 能够实现马与斑马图像的相互转换. summer2winter 数据集包含夏季和冬季的风景图像, 模型需要学习夏季和冬季两个域之间的映射, 目标是实现夏季场景和冬季场景的转换. edges2shoes 数据集则包含鞋子的轮廓图与对应的鞋子图像, 模型需学习从轮廓到鞋子图像的映射, 能够将鞋子的轮廓图转换为具有真实纹理、颜色和细节的鞋子图像。

#### 3.1.2 实现

本文在 CycleGAN 模型上进行了 100 个 epoch 的训练. 初始学习率设置为  $2 \times 10^{-4}$ , batch size 设置为 1. 根据实验训练经验, 将超参数设置为:  $\lambda_{\text{dcd}}=1$ 、 $\lambda_{\text{spe}}=0.2$ 、 $\lambda_{\text{fea}}=1 \times 10^1$ 、 $\lambda_{\text{sty}}=1 \times 10^4$ 、 $\lambda_{\text{stu}}=1$ . 为确保实验结果的公平性与可比性, 所有实验均在单张 RTX 3080 Ti GPU 上进行, 并在相同环境下对 DCD 方法<sup>[29]</sup>进行了重新训练。

#### 3.1.3 评估指标

本文采用 Fréchet inception distance (FID)<sup>[30]</sup> 指标对 CycleGAN 模型的性能进行评估. FID 通过预训练的 Inception-v3 网络提取真实图像与生成图像的特征向量, 然后计算特征向量之间的 Fréchet 距离, 得到 FID 值. FID 的值越低, 表示生成图像更接近于真实图像. 其计算公式如式 (15) 所示:

$$FID = \|\mu_r - \mu_g\|^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}) \quad (15)$$

其中,  $\mu_r$  和  $\mu_g$  表示真实图像和生成图像的特征均值,  $\Sigma_r$  和  $\Sigma_g$  表示真实特征向量和生成特征向量的协方差。

### 3.2 定量比较

在对定量结果进行分析时, 为了突出实验结果的优越性, 最优的实验数据用粗体来表示. 在表 1 中将

模型 FESD-CycleGAN 与现有的 GAN 压缩方法在 horse2zebra、summer2winter 和 edges2shoes 数据集上进行比较. 从表 1 中可以看出, 在 horse2zebra 数据集上, 大多数现有的 GAN 压缩方法表现较为不足, 导致  $FID$  值都有所上升. 然而, OMGD、DCD 以及本文提出的

FESD-CycleGAN 方法能够显著提升模型性能, 其中 FESD-CycleGAN 的  $FID$  表现优于现有的 GAN 压缩方法. 在 summer2winter 和 edges2shoes 数据集上, FESD-CycleGAN 的  $FID$  值同样得到比现有的 GAN 压缩方法更好的结果.

表 1 压缩 CycleGAN 应用于 horse2zebra 和 summer2winter 数据集、压缩 Pix2Pix 应用于 edges2shoes 数据集时的性能比较

模型	数据集	方法	MACs (G)	参数量 (M)	$FID$ ( $\downarrow$ )
CycleGAN	horse2zebra	Original <sup>[13]</sup>	56.80	11.30	61.53 (—)
		Co-evolutionary <sup>[14]</sup>	13.40	—	96.15 (−34.62)
		DMAD <sup>[31]</sup>	2.41	0.28	62.96 (−1.43)
		Wavelet KD <sup>[32]</sup>	1.68	0.72	77.04 (−15.51)
		GAN compression <sup>[15]</sup>	2.67	0.34	64.95 (−3.42)
		CAT <sup>[33]</sup>	2.56	—	60.18 (+1.35)
		OMGD <sup>[20]</sup>	1.408	0.137	59.81 (+1.72)
		DCD <sup>[29]</sup>	1.408	0.137	57.17 (+4.36)
		FESD-CycleGAN	1.408	0.137	<b>54.98 (+6.55)</b>
	summer2winter	Original <sup>[13]</sup>	56.80	11.30	79.12 (—)
		Co-evolutionary <sup>[14]</sup>	11.10	—	78.58 (+0.54)
		AutoGAN-Distiller <sup>[34]</sup>	4.34	—	78.33 (+0.79)
		DMAD <sup>[31]</sup>	3.18	0.30	78.24 (+0.88)
		CAT <sup>[33]</sup>	2.60	—	78.29 (+0.83)
		OMGD <sup>[20]</sup>	1.408	0.137	75.20 (+3.92)
		DCD <sup>[29]</sup>	1.408	0.137	74.09 (+5.03)
		FESD-CycleGAN	1.408	0.137	<b>73.41 (+5.71)</b>
		Pix2Pix	edges2shoes	Original <sup>[13]</sup>	18.60
DMAD <sup>[31]</sup>	2.99			2.13	46.95 (−12.64)
OMGD <sup>[20]</sup>	1.219			3.404	29.39 (+4.92)
DCD <sup>[29]</sup>	1.219			3.404	28.21 (+6.10)
FESD-CycleGAN	1.219			3.404	<b>27.45 (+6.86)</b>

具体而言, 本文提出的 FESD-CycleGAN 相比于原始的 CycleGAN 计算量降低了 97.5%、参数量降低了 98.8%, 表现出比当前最优方法 DCD 更优异的性能. 在 horse2zebra 数据集上, 相比于 DCD, FESD-CycleGAN 的  $FID$  值从 57.17 降至 54.98, 同时相比于原始 CycleGAN 的性能提升 6.55. 值得注意的是, DCD 的  $FID$  值为 57.17, 已经表现出较为优越的性能, 而 FESD-CycleGAN 方法仍可以进一步提升性能. 此外, 在 summer2winter 数据集上, FESD-CycleGAN 的  $FID$  值 73.41 比 DCD 的  $FID$  值 74.09 下降了 0.68; 在 edges2shoes 数据集上 FESD-CycleGAN 的  $FID$  值 27.45 比 DCD 的  $FID$  值 28.21 下降了 0.76, 性能同样得到了提升. 因此, FESD-CycleGAN 在提升轻量化生成器性能方面展示出了巨大的潜力.

为验证上述  $FID$  改进的可靠性, 进一步在 edges2-

shoes 数据集上评估了 LPIPS 与 PSNR 两个常用图像质量指标. 实验结果如表 2 所示, 相比 Original 模型, FESD-CycleGAN 的  $FID$  值下降了 6.86, LPIPS 降低了 0.008, PSNR 提升了 0.09 dB. 3 项指标均呈现出一致的优化趋势, 进一步佐证了以  $FID$  为核心评估标准的合理性与可靠性.

表 2 FESD-CycleGAN 与 Original 方法在 edges2shoes 数据集上的多指标评估

方法	$FID$ ( $\downarrow$ )	LPIPS ( $\downarrow$ )	PSNR ( $\uparrow$ ) (dB)
Original	34.31	0.191	21.08
FESD-CycleGAN	<b>27.45</b>	<b>0.183</b>	<b>21.17</b>

### 3.3 可视化

图 5、图 6 中进一步呈现了原始 CycleGAN 模型、OMGD、DCD 以及本文的 FESD-CycleGAN 模型所生成图像的可视化结果. 图 7 中呈现了高频细节的局部

放大对比. 从结果可以看出, 在 horse2zebra 数据集上, FESD-CycleGAN 不仅生成了视觉效果更好的图像, 还较好地保留了图像的背景细节. 相比其他压缩方法, FESD-CycleGAN 所生成的斑马条纹更清晰锐利. 在

summer2winter 数据集上, FESD-CycleGAN 能够在保持前景关键元素的同时, 平滑地转移背景风格. 在 edges2shoes 数据集上, FESD-CycleGAN 有助于捕捉鞋子的纹理细节.

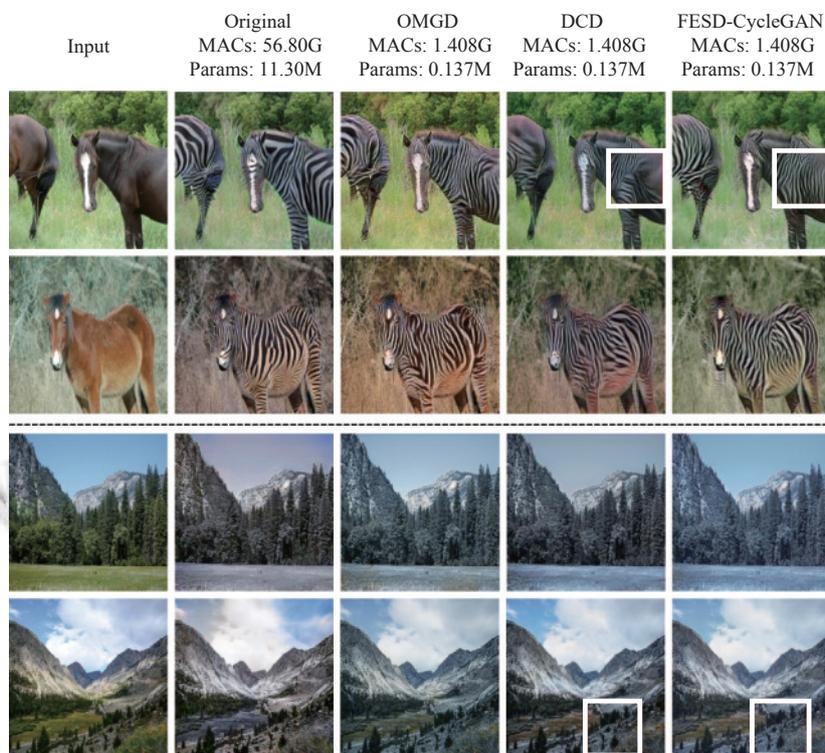


图 5 在 horse2zebra、summer2winter 数据集上与以往 CycleGAN 的可视化比较



图 6 在 edges2shoes 数据集上与以往 CycleGAN 的可视化比较

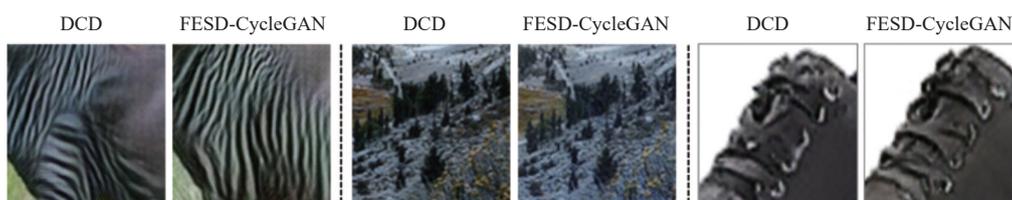


图 7 DCD 与 FESD-CycleGAN 方法的高频细节对比 (对应图 5、图 6 中框出的内容)

### 3.4 消融研究

为提升 PCS 方法的有效性并避免其可能引发的问题,需要对偏移方向、偏移幅度与通道偏移数量进行合理约束.为此,在 Baseline (BL) 上进行消融实验,从这 3 个维度评估其对性能的影响. BL 指原始 CycleGAN 压缩模型.

表 3 中展示了在通道偏移率  $\alpha = 1/8$  和偏移幅度  $|h_i|=2$  的条件下,4 种偏移方向的性能对比.实验结果表明,采取 Bi-directional 模式能够获得最优的 FID 值 73.96,从而使生成器表现出更好的性能.

表 3 4 种偏移方向应用于 BL 模型在 summer2winter 数据集上的性能比较

通道偏移率 $\alpha$	偏移幅度 $ h_i $	偏移方向	FID ( $\downarrow$ )
—	—	—	74.09
1/8	2	Uni-directional	74.05
1/8	2	Bi-directional	<b>73.96</b>
1/8	2	Cross-directional	74.26
1/8	2	Quad-directional	74.24

在表 4 中继续分析了在偏移方向为 Bi-directional 和通道偏移率  $\alpha = 1/8$  的条件下,不同偏移幅度对模型性能的影响.可以看到,较小的偏移幅度 ( $|h_i|=2$ ) 能够有效提升模型的性能.当偏移幅度较大时 ( $|h_i|=8$ ),图像的空间局部性可能会遭到破坏导致性能降低.因此,将偏移幅度的限制阈值  $L_m$  设置为 3.

表 4 不同偏移幅度应用于 BL 模型在 summer2winter 数据集上的性能比较

偏移方向	通道偏移率 $\alpha$	偏移幅度 $ h_i $	FID ( $\downarrow$ )
—	—	—	74.09
Bi-directional	1/8	<b>2</b>	<b>73.96</b>
Bi-directional	1/8	4	74.23
Bi-directional	1/8	8	75.46

此外,在偏移方向为 Bi-directional 和偏移幅度  $|h_i|=2$  的条件下,评估了不同通道偏移率  $\alpha$  对性能的影响.在表 5 中可以看到  $\alpha = 1/8$  的生成质量优于更小 (1/16) 或更大 (1/4) 的偏移比例.

表 5 不同通道偏移率应用于 BL 模型在 summer2winter 数据集上的性能比较

偏移方向	偏移幅度 $ h_i $	通道偏移率 $\alpha$	FID ( $\downarrow$ )
—	—	—	74.09
Bi-directional	2	1/16	74.18
Bi-directional	2	<b>1/8</b>	<b>73.96</b>
Bi-directional	2	1/4	74.25

为了分析部分通道偏移 (PCS) 方法和频谱蒸馏损失  $L_{\text{spectral}}$  对本文提出的 FESD-CycleGAN 模型的性能

影响,进行了消融研究,其中加粗的数据表示更优的结果,结果如表 6 所示.

表 6 消融实验 FID 结果 ( $\downarrow$ )

方法	horse2zebra	summer2winter
BL	57.17	74.09
BL+PCS	56.06	73.96
BL+PCS+ $L_{\text{spectral}}$	<b>54.98</b>	<b>73.41</b>

从表 6 中可以看出,在 horse2zebra 数据集中,本文的方法在 FID 值上均有下降.将 PCS 方法加入 BL 之后,FID 值从 57.17 降低到 56.06,降低了 1.11.这表明,偏移部分通道特征的特征增强方法有效促进了生成图像质量的提升.当将 PCS 方法与频谱蒸馏损失  $L_{\text{spectral}}$  结合起来时,也就是 FESD-CycleGAN 模型,得到最佳的 FID 值 54.98.这说明,改进后的判别器在区分真实图像和生成图像方面的能力得到了增强,从而有效激励生成器生成质量更高的图像.同样地,在 summer2winter 数据集中,将 PCS 方法加入 BL 之后,FID 值从 74.09 降低到 73.96,降低了 0.13.而结合频谱蒸馏损失  $L_{\text{spectral}}$  后,最终取得了最佳 FID 值 73.41,模型在生成图像质量上取得了显著提升.

### 3.5 推理效率分析

为评估 FESD-CycleGAN 在实际推理过程中的计算效率,在 RTX 3080 Ti GPU 上模拟交互式部署场景,对 horse2zebra 和 summer2winter 数据集进行了推理时间测试.实验设置为 batch size=1,推理前进行 100 次预热,随后遍历整个数据集,记录每张图像的推理耗时并计算平均值.结果如表 7 所示,FESD-CycleGAN 分别将 Original 模型的平均推理时间从 8.87 ms 和 8.65 ms 降低至 6.49 ms 和 6.13 ms,推理速度提升约 26.8% 和 29.1%.结合 MACs 和参数数量的大幅压缩,FESD-CycleGAN 在提升图像质量的同时显著优化了计算效率,具备良好的实时部署潜力.

表 7 推理效率结果

数据集	方法	MACs (G)	参数数量 (M)	延迟 (ms)
horse2zebra	Original	56.80	11.30	8.87
	FESD-CycleGAN	1.408	0.137	<b>6.49</b>
summer2winter	Original	56.80	11.30	8.65
	FESD-CycleGAN	1.408	0.137	<b>6.13</b>

## 4 结论

在本文中,提出了一种基于特征增强的频谱知识蒸馏方法——FESD-CycleGAN.该方法通过利用增强

后的特征图,将空间域和频谱域相结合,以实现更高效的知识蒸馏,从而能够在生成图像中表现出更精细的细节以及更高的整体质量.实验结果证明了 FESD-CycleGAN 在定量和定性性能上均取得了显著的提升,充分验证了 FESD-CycleGAN 模型的有效性和优越性.

尽管模型的整体性能有所提升,但仍存在改进空间.在 horse2zebra 数据集的图像转换任务中,生成斑马的条纹是关键.然而,生成的斑马条纹容易趋于某一固定模式,缺乏足够的多样性.未来的研究可以通过结合多尺度生成器方法以及模型压缩技术,进一步优化模型,从而提升生成图像的质量,同时减少计算和存储需求.

### 参考文献

- 1 Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks. *Communications of the ACM*, 2020, 63(11): 139–144. [doi: [10.1145/3422622](https://doi.org/10.1145/3422622)]
- 2 李帅先, 谭桂梅, 刘汝璇, 等. 利用自相似性实现医学图像合成的生成对抗网络. *中南民族大学学报(自然科学版)*, 2024, 43(1): 78–89.
- 3 Chen Y, Lai YK, Liu YJ. CartoonGAN: Generative adversarial networks for photo cartoonization. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 9465–9474. [doi: [10.1109/CVPR.2018.00986](https://doi.org/10.1109/CVPR.2018.00986)]
- 4 杜佳俊, 兰红. 基于多模态特征融合的图像编辑模型. *计算机科学与应用*, 2024, 14(6): 164–176. [doi: [10.12677/csa.2024.146153](https://doi.org/10.12677/csa.2024.146153)]
- 5 Jiang YF, Gong XY, Liu D, *et al.* EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 2021, 30: 2340–2349. [doi: [10.1109/TIP.2021.3051462](https://doi.org/10.1109/TIP.2021.3051462)]
- 6 Wang HT, Gui SP, Yang HC, *et al.* GAN slimming: All-in-one GAN compression by a unified optimization framework. *Proceedings of the 16th European Conference on Computer Vision*. Glasgow: Springer, 2020. 54–73.
- 7 Li SJ, Wu J, Xiao XF, *et al.* Revisiting discriminator in GAN compression: a generator-discriminator cooperative compression scheme. *Proceedings of the 35th International Conference on Neural Information Processing System*. Curran Associates Inc., 2021. 28560–28572.
- 8 Gao TW, Long RJ. Accumulation knowledge distillation for conditional GAN compression. *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris: IEEE, 2023. 1294–1303.
- 9 Lee D, Lee JY, Kim D, *et al.* Fix the noise: Disentangling source feature for transfer learning of StyleGAN. *arXiv: 2204.14079*, 2022.
- 10 Frank J, Eisenhofer T, Schönherr L, *et al.* Leveraging frequency analysis for deep fake image recognition. *Proceedings of the 37th International Conference on Machine Learning*. JMLR.org, 2020. 3247–3258.
- 11 Durall R, Keuper M, Keuper J. Watch your up-convolution: CNN based generative deep neural networks are failing to reproduce spectral distributions. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 7887–7896.
- 12 Isola P, Zhu JY, Zhou TH, *et al.* Image-to-image translation with conditional adversarial networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 5967–5976. [doi: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632)]
- 13 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017. 2242–2251. [doi: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244)]
- 14 Shu H, Wang YH, Jia X, *et al.* Co-evolutionary compression for unpaired image translation. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 3234–3243.
- 15 Li MY, Lin J, Ding YY, *et al.* GAN compression: Efficient architectures for interactive conditional GANs. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 5283–5293.
- 16 Kumar A, Anand K, Mandloi S, *et al.* CoroNetGAN: Controlled pruning of GANs via hypernetworks. *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris: IEEE, 2023. 1254–1263. [doi: [10.1109/ICCVW60793.2023.00136](https://doi.org/10.1109/ICCVW60793.2023.00136)]
- 17 Bucilua C, Caruana R, Niculescu-Mizil A. Model compression. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Philadelphia: ACM, 2006. 535–541.
- 18 Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. *arXiv:1503.02531v1*, 2015.
- 19 Zhou ZD, Zhuge CR, Guan XW, *et al.* Channel distillation: Channel-wise attention for knowledge distillation. *arXiv: 2006.01683v1*, 2020.
- 20 Ren YX, Wu J, Xiao XF, *et al.* Online multi-granularity

- distillation for GAN compression. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 6773–6783. [doi: [10.1109/ICCV48922.2021.00672](https://doi.org/10.1109/ICCV48922.2021.00672)]
- 21 Zhao XL, Hu XF, Liao Y, *et al.* Accurate MR image super-resolution via lightweight lateral inhibition network. Computer Vision and Image Understanding, 2020, 201: 103075. [doi: [10.1016/j.cviu.2020.103075](https://doi.org/10.1016/j.cviu.2020.103075)]
- 22 Lin J, Gan C, Han S. TSM: Temporal shift module for efficient video understanding. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 7082–7092. [doi: [10.1109/ICCV.2019.00718](https://doi.org/10.1109/ICCV.2019.00718)]
- 23 He YH, Liu XG, Zhong HS, *et al.* AddressNet: Shift-based primitives for efficient convolutional neural networks. Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2019. 1213–1222. [doi: [10.1109/WACV.2019.00134](https://doi.org/10.1109/WACV.2019.00134)]
- 24 Zhang X, Karaman S, Chang SF. Detecting and simulating artifacts in GAN fake images. Proceedings of the 2019 IEEE International Workshop on Information Forensics and Security (WIFS). Delft: IEEE, 2019. 1–6.
- 25 Rahaman N, Baratin A, Arpit D, *et al.* On the spectral bias of neural networks. Proceedings of the 36th International Conference on Machine Learning. Long Beach: Curran Associates Inc., 2019. 5301–5310.
- 26 Cai M, Zhang H, Huang HJ, *et al.* Frequency domain image translation: More photo-realistic, better identity-preserving. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 13910–13920.
- 27 Zhang XM, Li TR, Zhao XL. Boosting single image super-resolution via partial channel shifting. Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision. Paris: IEEE, 2023. 13177–13186. [doi: [10.1109/ICCV51070.2023.01216](https://doi.org/10.1109/ICCV51070.2023.01216)]
- 28 Chen YQ, Li G, Jin CC, *et al.* SSD-GAN: Measuring the realness in the spatial and spectral domains. Proceedings of the 2021 AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021. 1105–1112. [doi: [10.1609/aaai.v35i2.16196](https://doi.org/10.1609/aaai.v35i2.16196)]
- 29 Hu T, Lin MB, You LZ, *et al.* Discriminator-cooperated feature map distillation for GAN compression. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 20351–20360. [doi: [10.1109/CVPR52729.2023.01949](https://doi.org/10.1109/CVPR52729.2023.01949)]
- 30 Heusel M, Ramsauer H, Unterthiner T, *et al.* GANs trained by a two time-scale update rule converge to a local nash equilibrium. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6629–6640.
- 31 Li SJ, Lin MB, Wang Y, *et al.* Learning efficient GANs for image translation via differentiable masks and co-attention distillation. IEEE Transactions on Multimedia, 2023, 25: 3180–3189. [doi: [10.1109/TMM.2022.3156699](https://doi.org/10.1109/TMM.2022.3156699)]
- 32 Zhang LF, Chen X, Tu XB, *et al.* Wavelet knowledge distillation: Towards efficient image-to-image translation. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 12454–12464.
- 33 Jin Q, Ren J, Woodford OJ, *et al.* Teachers do more than teach: Compressing image-to-image models. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13595–13606.
- 34 Fu YG, Chen WY, Wang HT, *et al.* AutoGAN-Distiller: Searching to compress generative adversarial networks. Proceedings of the 37th International Conference on Machine Learning. JMLR.org, 2020. 3292–3303.

(校对责编: 李慧鑫)