

# 人脸转正GAN模型的高效压缩

魏 雷, 邱卫根, 张立臣

广东工业大学, 广东 广州  
Email: weilei@mail2.gdut.edu.cn

收稿日期: 2021年2月23日; 录用日期: 2021年3月18日; 发布日期: 2021年3月26日

## 摘 要

生成对抗网络(Generative Adversarial Network)在正面人脸图像生成方面大放异彩, 生成的正面人脸极其逼真受到研究人缘的青睐。但其强大的图像生成能力源自于其训练和使用过程中巨大的计算量, GAN结构越复杂, 其计算量需求, 这极大地限制了其交互式部署。为增强其部署的便利性, 减少GAN的计算量需求, 本文提出了一种通用的压缩算法, 该算法对人脸转正GAN进行了压缩, 减少了GAN中生成器的推理时间和模型大小。本文的实验证明了本文算法在相较于原网络大幅减少了计算量的情况下, 压缩后的GAN网络仍然保持了较好的图片质量。

## 关键词

GAN, 图像生成, 人脸正面化, 模型压缩

# Efficient Compression of Face Frontalization GAN Model

Lei Wei, Weigen Qiu, Lichen Zhang

Guangdong University of Technology, Guangzhou Guangdong  
Email: weilei@mail2.gdut.edu.cn

Received: Feb. 23<sup>rd</sup>, 2021; accepted: Mar. 18<sup>th</sup>, 2021; published: Mar. 26<sup>th</sup>, 2021

## Abstract

GAN performed extremely well in the frontal face image generation, the generated frontal face is very realistic favored by most researches. However, its powerful image generation ability comes from the huge calculation power and storage space required, and the more complex the GAN structure, the greater the demand for computation, which greatly limits its interactive deployment applications. To enhance the convenience of its deployment and reduce the computational requirements of the GAN, the paper proposes a general compression algorithm. The algorithm compresses the GAN model size of face frontalization and reduces the inference time. The experiments in this paper show that the compressed GAN network still obtains better image quality under the condition that the

computation and storage load are greatly reduced compared with the original networks.

## Keywords

GAN, Image Generation, Face Frontalization, Model Compression

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

人脸正面化是一项旨在将各种视角下的人脸对齐到规范位置(即正面)的计算机视觉处理任务。2D/3D 纹理映射[1] [2] [3], 统计建模[4] [5]和基于深度学习的方法[6] [7]都在这一领域取得了一定进展。

人脸正面化生成成为人脸数据扩充提供了一种有效的方式, 在数据量小的情况下可以通过这一方法扩增数据量。还有一些研究人员用其提高非受限环境下人脸识别的正确率, 也取得了很好的效果。TP-GAN (Twopathway GAN) [7]是第一个用双通路结构来进行人脸生成的, 这使得它能同时关注整体结构和局部细节。CR-GAN [8]的双通路架构被用来学习完整的嵌入空间。FF-GAN [9]是基于 3D 人脸模型的 GAN。PIM [10]对 TP-GAN 做了一些改进。具体地说, 这种改进是使用了一种域自适应策略, 可以提高具有极端姿势变化的面部的识别性能。

使用 GAN 进行正面人脸生成比以往基于 CNN (Convolutional Neural Networks)的方法需要的计算强度多很多倍, 甚至是高 1~2 个数量级。例如 TP-GAN 有很好的生成效果, 但是模型复杂, 训练耗时, 与 MobileNet-v3 的 0.44G MAC (multiply-accumulate)相比, TP-GAN 生成一个图像消耗 281G MAC, 不仅需要训练全局网络, 还要训练多个补丁网络, 另外其数据处理也有很高的延迟, 这极大地限制了它的使用场景。如下图 1, 清晰、逼真的生成结果的代价就是大量的计算需求, 在模型大小相差不多的情况下, 人脸生成 GAN 也比用于识别的 CNN 所需计算量多很多。

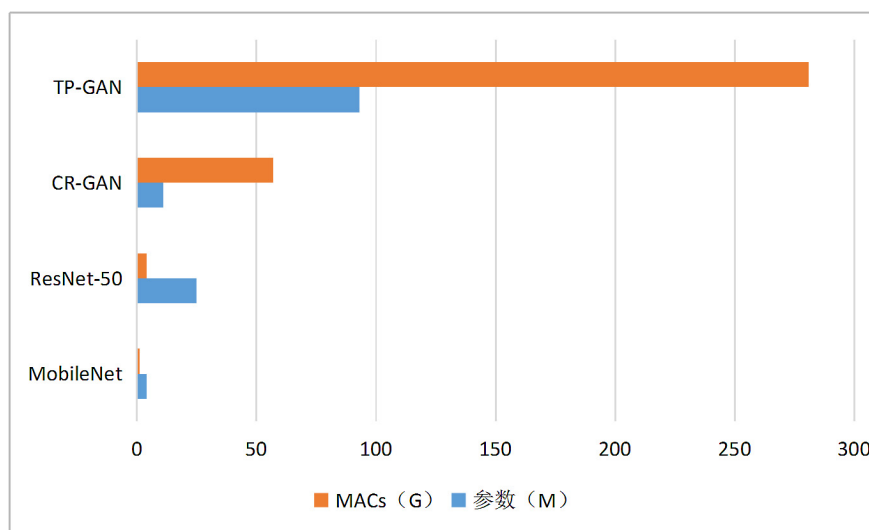


Figure 1. Image generation GAN requires much more calculations than image classification CNN

图 1. 图像生成 GAN 需要的计算量比图片分类 CNN 多得多

近日, 来自 MIT、Adobe、上海交通大学的研究者提出了一种用于压缩条件 GAN 的通用方法[11]。这一新技术在保持视觉保真度的同时, 将 pix2pix, CycleGAN 和 GauGAN 等广泛使用的条件 GAN 模型的计算量减少到 1/9~1/21。本文提出了一种算法, 将此思想应用到人脸转正的 GAN 模型上来。实验证明, 改进后的模型显著地减少了进行人脸转正工作的计算量, 压缩后 GAN 所需计算量减少了 7~19 倍, 生成图片的质量仍然得到了很好地保持。

## 2. 相关工作

### 2.1. GAN

GAN 使用对抗的学习方案, 通过生成器和判别器之间的零一博弈, 使双方都在训练上有所改进。生成对抗网络的目标函数可以表示为:

$$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \quad (1)$$

其中,  $x$  取自真实的样本分布,  $z$  取自噪声的随机分布。 $G$  和  $D$  分别表示生成网络和对抗网络。通过这种机制, GAN 可以很好地发现从源数据域到目的数据域的映射。基于它的这一特性, 研究人员挖掘出了其在图像生成方面的巨大潜力。随着 GAN 模型不断发展, 人脸正面化取得了显著成就。Twopathway GAN (TP-GAN) [12]率先提出了一种用于正面人脸合成的两通道方法, 该方法能够捕获局部细节并同时理解全局结构。此后不久, [13]开发了一个基于 GAN 的框架, 该框架重新组合了不同的身份和属性, 以在开放域中合成人脸时保留身份。之后的 PIM [10]以相互促进的方式共同学习姿势不变特征提取和人脸合成。

### 2.2. 模型加速

为了减少模型权重中的冗余, 加速深度学习模型的训练进程, 研究人员提出了一种对模型中各个层之间的连接进行剪枝的方法, 并以某些特定的硬件来加速其运行。一些后续的研究工作指出可以修剪所有的卷积滤波器来提高计算的规律性。用于模型压缩的 AutoML (AMC) [14]利用强化学习来自动确定每层的修剪率。后来, Liu 等人[15]用进化搜索算法代替了强化学习。

### 2.3. 知识蒸馏

Hinton 等[16]首次提出了知识蒸馏的概念, 用于将较大的教师网络中的知识转移到较小的学生网络中, 学生网络被训练来模仿教师网络的行为。有几种方法利用知识蒸馏来压缩识别模型[17]。最近, Aguinardo 等人[18]采用这种方法来加速无条件 GAN。与他们不同的是, 韩松等人[11]专注于条件 GAN。他们对条件 GAN 进行了几种蒸馏方法的实验, 但是仅观察到了边际改进, 不足以进行交互式应用。所以要达到压缩人脸生成 GAN 的目的, 还需要结合 NAS 技术。

### 2.4. NAS

神经网络结构搜索(NAS)已成功设计出了在大型图像分类任务中胜过手工设计的神经网络体系结构 [19] [20] [21]。为了有效地降低搜索成本, 研究人员最近提出了一步法(One-Shot)神经网络模型搜索框架 [22] [23] [24] [25], 其中不同的候选子网可以共享同一组权重。尽管所有这些方法都专注于图像分类模型, 但我们仍可以使用这项技术来寻找高效的人脸生成 GAN 架构。

## 3. 方法

对人脸生成 GAN 进行压缩有两个难点。首先, GANs 的训练动态本质上是高度不稳定的。其次, 识别模型和生成模型之间的巨大架构差异使得现有 CNN 压缩算法难以直接被应用。为解决上述问题, 韩松

等[11]提出了针对有效生成模型量身定制的训练协议，并通过神经网络结构搜索(NAS)进一步提高了压缩率。整体架构如图 2 所示，使用 ResNet 生成器作为示例。①使用预训练教师生成器  $G$ ，我们蒸馏了一个更小的“one-for-all”学生生成器，它包含所有可能的通道数量，我们在每个训练步骤选择不同通道数量的学生生成器。②然后我们可以从“one-for-all”生成器中提取很多的子生成器，并评估它们的表现。提取出的子生成器不需要再次进行训练，这就是使用“one-for-all”生成器的好处。③最终，我们选择在给定压缩比率和指标下表现最好的子生成器，进行微调，就获得了最终的压缩模型。

本文将这一思想运用于人脸生成 GAN 上，实验证明了其对人脸生成 GAN 的有效性。

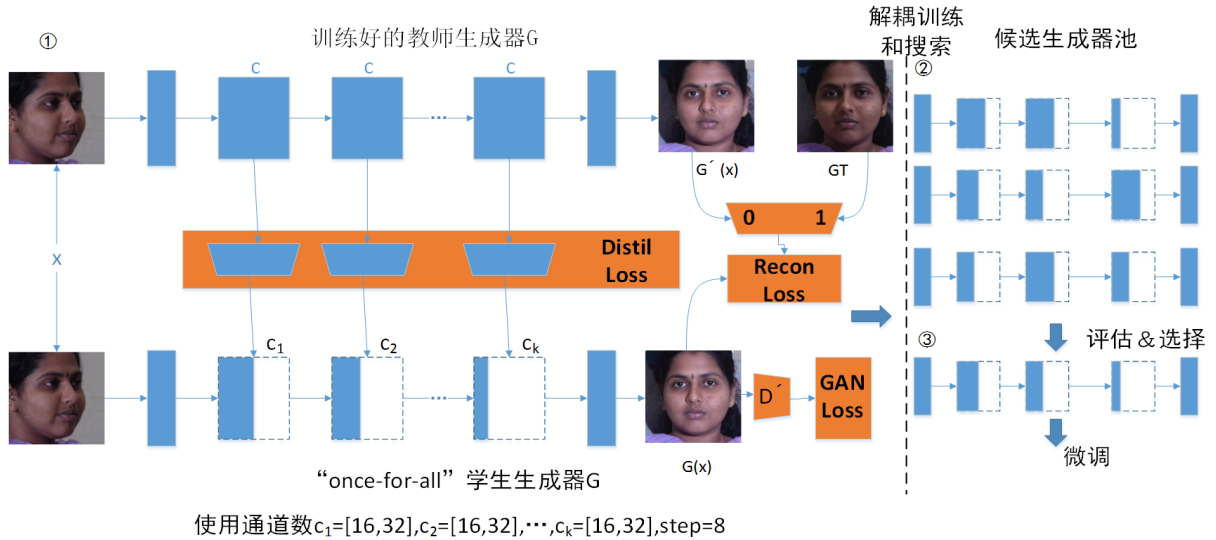


Figure 2. GAN Compression framework  
图 2. GAN 压缩框架

### 3.1. 训练目标

人脸正面化 GAN 的目标是学习从侧面人脸  $X$  到正面人脸  $Y$  的映射函数函数  $G$ ，可使用成对的数据进行训练 $(\{x_i, y_i\}_{i=1}^N, x_i \in X \text{ 且 } y_i \in Y)$ 。  $N$  表示训练图片的数目。模型学习目标形式化如下，其中  $\mathbb{E}_{\mathbf{x},\mathbf{y}} \triangleq \mathbb{E}_{\mathbf{x},\mathbf{y} \sim p_{\text{data}}}(\mathbf{x},\mathbf{y})$ ，  $\|\cdot\|_1$  表示 L1 正则模。

$$\mathcal{L}_{\text{recon}} = \mathbb{E}_{\mathbf{x},\mathbf{y}} \|G(\mathbf{x}) - \mathbf{y}\|_1 \quad (2)$$

#### 3.1.1. 继承鉴别器

尽管旨在压缩发生器，但鉴别器  $D$  也会存储训练好的 GAN 中的知识，用来发现当前生成器的弱点 [26]。因此，本文采用相同的鉴别器架构，使用来自教师的预先训练的权重，并让压缩好的生成器和鉴别器一起进行微调。在本文实验中，可以观察到预训练的鉴别器可以指导学生生成器的训练。使用随机初始化的鉴别器通常会导致严重的训练不稳定和图像质量下降。GAN 目标的形式化为：

$$\mathcal{L}_{\text{cGAN}} = \mathbb{E}_{\mathbf{x},\mathbf{y}} [\log D(\mathbf{x},\mathbf{y})] + \mathbb{E}_{\mathbf{x}} [\log(1 - D(\mathbf{x}, G(\mathbf{x})))] \quad (3)$$

本文模型中，使用教师鉴别器  $D'$  的权重初始化学生鉴别器  $D$ 。  $G$  和  $D$  使用标准的 minimax 优化进行训练。

#### 3.1.2. 中间特征蒸馏

CNN 模型压缩的一种广泛使用的方法是知识蒸馏[17] [27]。通过匹配输出层的逻辑分布，可以将暗

知识从教师模型转移到学生模型，从而提高学生模型的表现。但是，人脸生成 GAN 会输出确定性图像，而不是概率分布。因此，很难从教师模型的输出中提取暗知识。

为了解决上述问题，本文模型改为从教师生成器的中间表示形式提取暗知识，如先前的工作[17]所述。模型中间层包含更多通道，提供更丰富的信息，并允许学生模型除输出外获取更多信息。可以将蒸馏目标形式化为

$$\mathcal{L}_{\text{distill}} = \sum_{t=1}^T \left\| G_t(\mathbf{x}) - f_t(G'_t(\mathbf{x})) \right\|_2 \quad (4)$$

其中  $G_t(\mathbf{x})$  和  $G'_t(\mathbf{x})$  是在学生和教师模型中所选择的第  $t$  层的中间特征激活函数， $T$  表示层数。 $1 \times 1$  卷积层  $f_t$  将学生模型的特征映射到教师模型的特征中相同数量的通道，以共同优化  $G_t$  和  $f_t$ ，以最大程度地减少蒸馏损失  $l_{\text{distill}}$ 。

### 3.1.3. 中间特征蒸馏

本文最终目标函数如下：

$$\mathcal{L} = \mathcal{L}_{\text{cGAN}} + \lambda_{\text{recon}} \mathcal{L}_{\text{recon}} + \lambda_{\text{distill}} \mathcal{L}_{\text{distill}} \quad (5)$$

其中，超参数  $\lambda_{\text{recon}}$  和  $\lambda_{\text{distill}}$  控制每个项的重要程度。

## 3.2. 高效的生成器设计空间

选择一个设计良好的学生模型架构对于最终进行知识蒸馏至关重要。一般地，单纯地缩小教师模型的通道数并不能简单地生成一个紧凑的学生模型。经常地，在减少了 4 倍的计算量之后，性能开始显著下降。可能的原因之一是现有的生成器架构经常使用来自图像识别模型的[28] [29] [30]，这也许不是图像合成任务的最佳选择。下面，本文展示了如何从现有的 GAN 生成器中获得更好的架构设计空间，并在该空间内执行神经架构搜索(NAS)。

### 3.2.1. 卷积分解和层敏感性

通常地，有效的 CNN 设计广泛采用了分解版本的卷积(depthwise + pointwise) [31]，在表现与计算量之间的取得了很好的平衡。本文发现使用分解卷积也适用于 GAN 中的生成器设计。但是，将分解卷积直接应用于所有卷积层(如在分类器中)将大大降低图像质量。在某些层使用分解卷积会造成性能下降，而其他层则不会。此外，GAN 的层敏感性与识别模型的不同。例如，在 ResNet 生成器[32]中，resBlock 层消耗了大部分模型参数和计算成本，几乎不受分解的影响。相反，上采样层参数要少得多，却对模型压缩相当敏感。

### 3.2.2. 使用 NAS 自动减少通道

现有的发生器通常在所有层之间使用人工设置的通道数，其中存在冗余并且远非最佳。为了进一步提高压缩率，本文使用通道修剪[12] [14] [33] [34] [35]在生成器中自动选择通道宽度，以消除冗余，再次减少计算量。本文模型支持对通道数的细粒度选择，对于每个卷积层，可以从 8 的倍数中选择通道数，这可以平衡 MACs 和硬件并行性。

给定可能的通道配置  $\{c_1, c_2, \dots, c_K\}$ ，其中  $K$  是要修剪的层数，本文试图使用神经体系结构搜索找到最佳的通道配置  $\{c_1^*, c_2^*, \dots, c_K^*\} = \arg \min_{c_1, c_2, \dots, c_K} l$ ，s.t.  $\text{MACs} < F_t$ ，其中  $F_t$  是计算约束。通常，直接的方法是遍历所有可能的通道配置，训练它们直至收敛，然后从中评估并选择性能最佳的生成器。但是，随着  $K$  的增加，可能配置的数量呈指数增长，并且每种配置可能需要在每个阶段设置学习率和权重的不同超参数，非常耗时。

### 3.3. 解耦训练和搜索

为了解决这个问题, 本文模型遵循一步法神经体系结构搜索方法的最新研究[23] [25] [36], 将模型训练与体系结构搜索分离。首先训练一个支持不同通道数的“once-for-all”网络[25]。具有不同数目通道的每个子网都经过相同的训练过程, 可以独立运行。子网与“once-for-all”网络共享权重。图 2 说明了整个框架。假设原始教师生成器具有  $\{c_k^0\}_{k=1}^K$  个通道。对于给定的信道编号配置  $\{c_k\}_{k=1}^K$ ,  $c_k \leq c_k^0$ , 通过从“once-for-all”的相应权重张量中提取开始的  $\{c_k\}_{k=1}^K$  通道来获得子网的权重。遵循 Guo 等人的理论[36], 在每个训练步骤中, 使用一定的通道数配置对子网进行随机抽样, 计算输出和梯度, 并使用学习目标来更新提取的权重(等式 5)。由于前几个通道的权重被更频繁地更新, 因此它们在所有权重中扮演着更为关键的角色。

在训练了“once-for-all”网络之后, 通过直接在验证集上评估每个候选子网的性能来找到最佳子网。由于“once-for-all”网络经过权重共享的方式直接训练, 因此无需进行微调。通过这种方式, 可以将训练和搜索生成器体系结构分离开来: 只需要训练一次, 然后就可以在无需进一步训练的情况下评估所有可能的通道配置, 并选择最佳的作为搜索结果。然后将所选择的网络进行微调来进一步提升它的表现。

## 4. 实验

### 4.1. 实验设置

#### 4.1.1. 模型

本文在以下两种模型上进行实验以论证方法的有效性。

TP-GAN [12]是一个能够考虑整体和局部信息的双路径对抗生成网络, 可以用于从单一侧面照片合成高清晰的正面人脸图像。通过该方法合成的人脸图像能够很好地保留身份特征, 并且可以处理大量不同姿势和光照的图片。

CR-GAN [8]也是一个双路径生成对抗网络, 除了一个重建路径外, 还引入了一条路径以保持学习到的嵌入空间的完整性。这两条路径以参数共享的方式进行协作和竞争, 从而极大地提高了能力。

#### 4.1.2. 数据集

MultiPIE 数据集[37]是在受控环境中用于人脸合成和识别的最大公共数据库。MultiPIE 数据库包含 337 个身份个体, 每个个体下含有 20 种光照、15 种姿态以及两种标签, 分成四个时间段拍摄而成。并且, 数据库中还有少量的高像素度的的人脸图片供研究使用。整个 MultiPIE 数据库共有超过 750,000 张人脸图像, 是一个大规模的能用于多姿态人脸识别的数据库。本文实验使用前 200 名受试者的图像进行训练, 其中包括在  $\pm 90^\circ$ 内具有 13 个姿势和 20 个照明水平的样本。其余 137 个身份的样本构成了测试集, 而表达和照明中性的样本构成了画廊。请注意, 训练和测试集之间没有重叠的科目。

CAS\_PEAL\_R1 数据集[38]是一个公开发布的大规模中国人人脸数据库, 其姿态, 表情, 配饰和灯光变化都受到控制。它包括 1040 人的 30,863 张灰度图片, 其中男性 595 名, 女性 445 名。

#### 4.1.3. 实验细节

为了训练本文模型, 需要成对的人脸图像, 其中是侧面人脸图像, 是正面人脸图像。首先, 需要将所有图片修剪到 128,128 [16]。所使用的身份保留网络在 MS-Celeb-1M 上进行预训练, 并在 Multi-PIE 的训练集上进行微调。对于 CAS-PEAL-R1, 所有图片都被设置为灰度图片。用于训练 CAS-PEAL-R1 的身份保留网络已在 MS-Celeb-1M 的灰度图片上进行了预训练。

## 4.2. 结果

### 4.2.1. 身份保留能力

为了定量地论证压缩后模型的身份保留能力, 本文评估了在合成出的人脸图片上进行人脸识别的准确率。本文使用一个训练好的 29 层 Light-CNN 做为人脸识别模型来提取特征, 然后用余弦距离度量来计算这些特征的相似度。大姿态下的人脸图片提供的信息太少, 使得模型很难在生成的正面人脸图片中保留身份信息。表 1 比较了在 MultiPIE 数据集上不同角度下各个模型的表现, 表 2 比较了在 CAS\_PEAL\_R1 数据集上不同角度下各个模型的表现, 结果都以 rank-1 识别率表示。如表所示, 各个角度下的表现, 原模型与压缩模型都相当接近。

**Table 1.** Rank-1 recognition rate from various angles on the MultiPIE dataset (%)

**表 1.** MultiPIE 数据集上各个角度下的 rank-1 识别率(%)

	90°	75°	60°	45°	30°	15°	Avg
TP-GAN	64.64	77.43	87.72	95.38	98.06	98.68	86.99
CR-GAN	71.60	92.50	97.00	98.60	99.30	99.40	93.07
TP-GAN(c)	60.73	72.32	83.65	92.19	96.01	97.24	83.69
CR-GAN(c)	67.84	88.21	93.37	95.28	97.15	98.75	90.10

**Table 2.** Rank-1 recognition rate on the CAS\_PEAL\_R1 dataset (%)

**表 2.** CAS\_PEAL\_R1 数据集上的 rank-1 识别率(%)

Yaw	Pitch(-15°)					Pitch(0°)					Pitch(+15°)				
	0°	15°	30°	45°	Avg_1	15°	30°	45°	Avg_2	0°	15°	30°	45°	Avg_3	
TP-GAN	98.86	98.94	98.89	97.62	98.58	100	99.94	98.71	99.55	97.68	97.73	97.45	95.83	97.17	
CR-GAN	83.98	83.91	83.71	80.38	82.86	97.61	95.80	89.73	94.38	89.74	89.44	87.95	83.90	87.76	
TP-GAN(c)	97.73	96.94	97.47	96.52	97.16	99.57	99.43	97.18	98.93	95.93	95.42	94.17	93.28	94.70	
CR-GAN(c)	80.73	80.26	80.43	79.92	80.34	95.25	93.74	87.10	92.03	87.32	87.07	85.28	80.15	84.96	

### 4.2.2. 模型压缩效果

在表 3 中, 记录了在 MultiPIE 数据集上对 TP-GAN, PIM 和 CR-GAN 进行压缩的定量结果。通过使用“once-for-all”网络中性能最佳的子网, 这一 GAN 压缩方法可实现较大的压缩率。它可以将人脸生成 GAN 的计算量减少 7~25 倍, 并将模型的参数量缩小 5~31 倍, 从上一节可以看到, 在这么大的压缩比下, 模型的表现下降极小。这说明了该方法的有效性。

**Table 3.** Quantitative evaluation of GAN compression

**表 3.** 定量评估 GAN 压缩

模型	方法	参数	MACs
TP-GAN	原始	14.3 M	78.3 G
	压缩	0.52 M (27.3×)	4.03 G (19.4×)
PIM	原始	12.7 M	57.9 G
	压缩	0.84 M (15.1×)	5.17 G (11.2×)
CR-GAN	原始	10.5 M	32.4 G
	压缩	1.81 M (5.8×)	4.43 G (7.3×)

### 4.2.3. 生成图片效果

在这一部分，我们在视觉上比较了不同模型与其压缩模型的正面人脸生成效果。图 3 展示了在 MultiPIE 数据集上的定性比较结果。图 4 显示了在 CAS\_PEAL\_R1 各个模型的生成效果。我们提供输入，其真实正脸图片，不同角度下原始模型的输出以及压缩模型的输出。从图片可以看出压缩后的模型仍然可以生成相当接近原模型的逼真人脸图像。

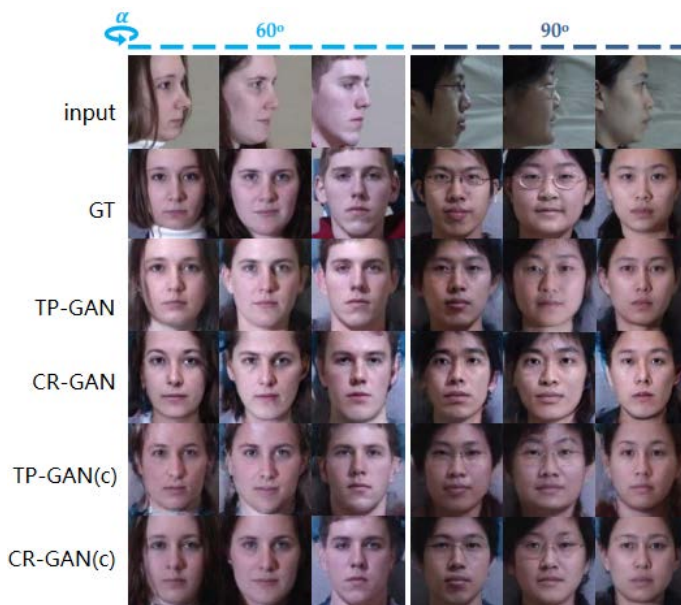


Figure 3. The performance of each model on the MultiPIE dataset  
图 3. MultiPIE 数据集下各模型表现

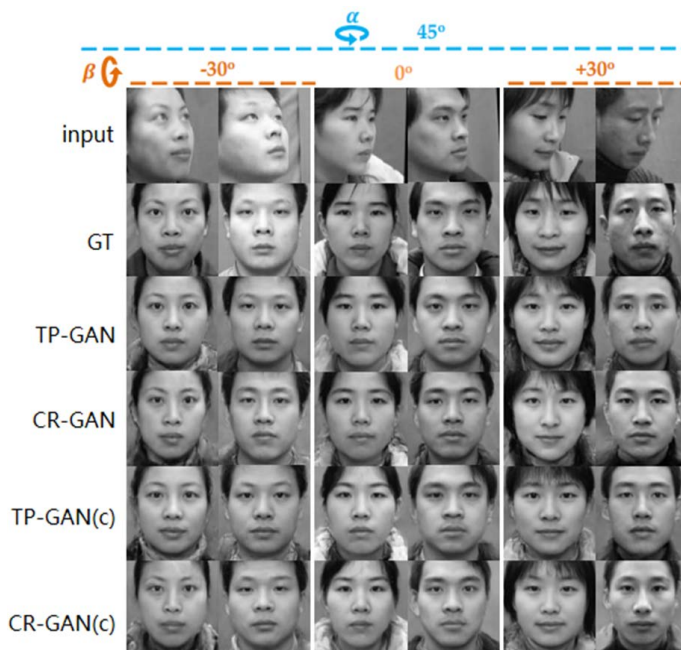


Figure 4. The performance of each model under the CAS\_PEAL\_R1 dataset  
图 4. CAS\_PEAL\_R1 数据集上各模型的表现



#### 4.2.4. 加快硬件推断

对于现实世界的交互式应用程序，硬件上的计算加速比减少计算量更为关键。为了验证本文方法的实际有效性，本实验在几种具有不同计算能力的设备上测量压缩模型的推理速度。为了模拟交互式应用程序，我们使用 1 的批量大小。首先执行 100 次热身运行，并测量接下来 100 次运行的平均时间。结果如表 4 所示。注意到，与 CPU 相比，GPU 上的加速效果不那么明显，这主要是由于 GPU 高度的并行性。不过，专注于在可能无法使用强大 GPU 的设备上使模型更易于访问，以便更多人可以使用交互式人脸生成 GAN 应用程序。

**Table 4.** Measured memory reduction and latency speedup

**表 4.** 测量内存减少和模型加速

模型	TP-GAN	CR-GAN
MAC 减少	17.8×	11.7×
内存减少	2.3×	1.4×
3060TI 加速	0.007 s (2.7×)	0.010 s (1.5×)
15 10400 CPU 加速	0.13 s (3.2×)	0.67 s (2.5×)

## 5. 结论

本文借鉴韩松团队提出的通用压缩框架思想，提出了一种人脸生成 GAN 压缩算法，使用知识蒸馏和神经体系结构搜索，来减轻训练的不稳定性并提高模型效率。本文实验表明，本文方法在保留视觉质量的同时，可以有效压缩多个人脸生成 GAN 模型。

## 参考文献

- [1] Ferrari, C., Lisanti, G., Berretti, S. and Del Bimbo, A. (2016) Effective 3D Based Frontalization for Unconstrained Face Recognition. 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, 1047-1052. <https://doi.org/10.1109/ICPR.2016.7899774>
- [2] Hassner, T., Harel, S., Paz, E. and Enbar, R. (2015) Effective Face Frontalization in Unconstrained Images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 4295-4304.
- [3] Jeni, L.A. and Cohn, J.F. (2016) Person-Independent 3d Gaze Estimation Using Face Frontalization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, Las Vegas, 26 June-1 July 2016, 87-95.
- [4] Booth, J., Roussos, A., Ponniah, A., et al. (2018) Large Scale 3D Morphable Models. *International Journal of Computer Vision*, **126**, 233-254. <https://doi.org/10.1007/s11263-017-1009-7>
- [5] Booth, J., Roussos, A., Zafeiriou, S., Ponniah, A. and Dunaway, D. (2016) A 3d Morphable Model Learnt from 10,000 Faces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 26 June-1 July 2016, 5543-5552.
- [6] Cao, J., Hu, Y., Zhang, H., et al. (2018) Learning a High Fidelity Pose Invariant Model for High-Resolution Face Frontalization.
- [7] Huang, R., Zhang, S., Li, T., et al. (2017) Beyond Face Rotation: Global and Local Perception Gan for Photorealistic and Identity Preserving Frontal View Synthesis. *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 2439-2448. vision. 2439-2448. <https://doi.org/10.1109/ICCV.2017.267>
- [8] Tian, Y., Peng, X., Zhao, L., et al. (2018) CR-GAN: Learning Complete Representations for Multi-View Generation. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence Main Track*, Stockholm, 13-19 July 2018, 942-948. <https://doi.org/10.24963/ijcai.2018/131>
- [9] Yin, X., Yu, X., Sohn, K., Liu, X.M. and Chandraker, M. (2017) Towards Large-Pose Face Frontalization in the

- Wild. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 3990-3999. <https://doi.org/10.1109/ICCV.2017.430>
- [10] Zhao, J., Cheng, Y., Xu, Y., Xiong, L., Li, J., Zhao, F., Jayashree, K., Pranata, S., Shen, S., Xing, J., *et al.* (2018) Towards Pose Invariant Face Recognition in the Wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-23 June 2018, 2207-2216. <https://doi.org/10.1109/CVPR.2018.00235>
- [11] Li, M., Lin, J., Ding, Y., *et al.* (2020) GAN Compression: Efficient Architectures for Interactive Conditional Gans. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 14-19 June 2020, 5284-5294. <https://doi.org/10.1109/CVPR42600.2020.00533>
- [12] He, Y.H., Zhang, X.Y. and Sun, J. (2017) Channel Pruning for Accelerating Very Deep Neural Networks. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 1389-1397.
- [13] Shen, Y., Luo, P., Yan, J., Wang, X. and Tang, X. (2018) Faceid-GAN: Learning a Symmetry Three-Player GAN for Identity-Preserving Face Synthesis. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-23 June 2018, 821-830. <https://doi.org/10.1109/CVPR.2018.00092>
- [14] He, Y.H., Lin, J., Liu, Z.J., Wang, H.R., Li, L.-J. and Han, S. (2018) AMC: AutoML for Model Compression and Acceleration on Mobile Devices. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 784-800. [https://doi.org/10.1007/978-3-030-01234-2\\_48](https://doi.org/10.1007/978-3-030-01234-2_48)
- [15] Liu, Z.C., Mu, H.Y., Zhang, X.Y., Guo, Z.C., Yang, X., Kwang-Ting Cheng, T. and Sun, J. (2019) Metapruning: Meta Learning for Automatic Neural Network Channel Pruning. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 3296-3305. <https://doi.org/10.1109/ICCV.2019.00339>
- [16] Hinton, G., Vinyals, O. and Dean, J. (2015) Distilling the Knowledge in a Neural Network.
- [17] Chen, G., Choi, W., Yu, X., *et al.* (2017) Learning Efficient Object Detection Models with Knowledge Distillation. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 742-751.
- [18] Aguinaldo, A., Chiang, P.-Y., Gain, A., Patil, A., Pearson, K. and Feizi, S. (2019) Compressing Gans Using Knowledge Distillation.
- [19] Zoph, B. and Le, Q.V. (2016) Neural Architecture Search with Reinforcement Learning.
- [20] Liu, C., Zoph, B., Neumann, M., *et al.* (2018) Progressive Neural Architecture Search. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 19-34. [https://doi.org/10.1007/978-3-030-01246-5\\_2](https://doi.org/10.1007/978-3-030-01246-5_2)
- [21] Liu, H., Simonyan, K., Vinyals, O., *et al.* (2017) Hierarchical Representations for Efficient Architecture Search.
- [22] Liu, H., Simonyan, K. and Yang, Y. (2018) Darts: Differentiable Architecture Search.
- [23] Cai, H., Zhu, L. and Han, S. (2018) Proxylessnas: Direct Neural Architecture Search on Target Task and Hardware.
- [24] Wu, B.C., *et al.* (2019) Fbnet: Hardware-Aware Efficient ConvNet Design via Differentiable Neural Architecture Search. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seoul, 27 October-2 November 2019, 10734-10742. <https://doi.org/10.1109/CVPR.2019.01099>
- [25] Cai, H., Gan, C., Wang, T., *et al.* (2019) Once-for-All: Train One Network and Specialize It for Efficient Deployment.
- [26] Azadi, S., Olsson, C., Darrell, T., *et al.* (2018) Discriminator Rejection Sampling.
- [27] Chen, Y., Wang, N. and Zhang, Z. (2018) DarkRank: Accelerating Deep Metric Learning via Cross Sample Similarities Transfer. *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**, 2852-2859. <https://ojs.aaai.org/index.php/AAAI/article/view/11783>
- [28] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *CVPR*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [29] He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 26 June-1 July 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [30] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-net: Convolutional Networks for Biomedical Image Segmentation. *MICCAI*, Munich, 5-9 October 2015, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [31] Howard, A.G., Zhu, M.L., Chen, B., Kalenichenko, D., *et al.* (2017) Efficient Convolutional Neural Networks for Mobile Vision Applications.
- [32] Johnson, J., *et al.* (2016) Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In: *European Conference on Computer Vision*, Springer, Cham, 694-711. [https://doi.org/10.1007/978-3-319-46475-6\\_43](https://doi.org/10.1007/978-3-319-46475-6_43)
- [33] Liu, Z., Li, J.G., Shen, Z.Q., Huang, G., Yan, S.M. and Zhang, C.S. (2017) Learning Efficient Convolutional Networks through Network Slimming. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice,

---

22-29 October 2017, 2736-2744.

- [34] Zhuang, Z.W., Tan, M.K., Zhuang, B.H., Liu, J., Guo, Y., Wu, Q.Y., *et al.* (2018) Discrimination-Aware Channel Pruning for Deep Neural Networks. *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, 3-8 December 2018, 875-886.
- [35] Luo, J.H., Wu, J. and Lin, W. (2017) ThiNet: A Filter Level Pruning Method for Deep Neural Network Compression. *Proceedings of the IEEE International Conference on Computer Vision, Venice, 22-29 October 2017*, 5058-5066. <https://doi.org/10.1109/ICCV.2017.541>
- [36] Guo, Z.C., Zhang, X.Y., Mu, H.Y., Heng, W., Liu, Z.C., Wei, Y.C. and Sun, J. (2019) Single Path Oneshot Neural Architecture Search with Uniform Sampling.
- [37] Gross, R., Matthews, I., Cohn, J., *et al.* (2010) Multi-Pie. *Image and Vision Computing*, **28**, 807-813. <https://doi.org/10.1016/j.imavis.2009.08.002>
- [38] Gao, W., Cao, B., Shan, S., *et al.* (2007) The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, **38**, 149-161. <https://doi.org/10.1109/TSMCA.2007.909557>