

引用格式:魏子涵,王慧,王晶,等. 基于特征增强的人脸属性转换[J]. 微电子学与计算机,2020,37(10):38-41. [WEI Z H, WANG H, WANG J, et al. Face attribute transfer based on feature enhancement[J]. Microelectronics & Computer, 2020, 37(10):38-41.]

基于特征增强的人脸属性转换

魏子涵^{1,2}, 王慧^{1,3}, 王晶^{1,3}, 高岚^{3,4}, 张伟功^{1,3}

(1 首都师范大学 北京成像理论与技术高精尖创新中心, 北京 100048;

2 西安电子科技大学 电子工程学院, 陕西 西安 710126;

3 首都师范大学 信息工程学院, 北京 100048;

4 电子系统可靠性技术北京市重点实验室, 北京 100048)

摘 要: 采用生成对抗网络及其扩展模型可以实现人脸属性转换, 但生成图像存在分辨率低及特征损失的问题. 本文基于 GeneGAN 网络, 提出一种基于特征增强的人脸属性转换算法, 在生成网络中增加网络深度, 建立高低维特征耦合通道, 通过对特定属性深度特征提取比较, 优化生成图像的身份及属性特征. 在 CelebA 数据集上的实验表明, 本文算法得到的高分辨率图像不仅在保留图像身份信息上有较好的改善, 同时在视觉效果和客观指标上都有较大的提高.

关键词: 生成对抗网络; 人脸属性转换; 特征增强; 深度特征; 反向传递

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2020)10—0038—04

Face attribute transfer based on feature enhancement

WEI Zi-han^{1,2}, WANG Hui^{1,3}, WANG Jing^{1,3}, GAO Lan^{3,4}, ZHANG Wei-gong^{1,3}

(1 Beijing Advanced Innovation Center for Imaging Theory and Technology, Capital Normal University,

Beijing, 100048, China; 2 College of Electronical and Information Engineering, Xidian University,

Xian, 710126, China; 3 College of Information Engineering, Capital Normal University,

Beijing 100048, China; 4 Beijing Key Laboratory of Electronic System Reliability and Prognostics,

Capital Normal University, Beijing 100048, China)

Abstract: The current facial attribute transformation based on generative adversarial networks (GAN) and GAN-based method has problems of low resolution and feature loss of the generated image. Based on the GeneGAN network, this paper proposes a face attribute conversion algorithm based on feature enhancement. By increasing the network depth in the generation network, establishing a high- and low-dimensional feature coupling channel, and extracting and comparing specific attribute depth features to optimize the identity and features of the generated image. Experiment based on CelebA dataset shows that the high-resolution images generated by the proposed algorithm have a better improvement in maintaining image identity information, and have a greater improvement in visual effects and objective indicators.

Key words: Generative adversarial network; Facial attributes transformation; Feature enhancement; Deep features; Backpropagation

收稿日期: 2020—04—20; 修回日期: 2020—05—22

基金项目: 国家自然科学基金(61772350); 北京市高水平教师队伍建设计划(CIT&TCD201704082); 北京市科技新星计划项目(Z181100006218093); 北京未来芯片技术高精尖创新中心科研基金资助项目(KYJJ2018008).

1 前言

基于生成对抗网络(GAN)实现人脸属性转换可以改善生成图像的质量^[1-2],为了进一步提高生成图像质量,条件生成对抗网络(cGAN)^[3]利用噪声分布和条件属性,在GAN模型的基础上增加条件变量 y 作为输入,通过条件变量 y 控制对应条件下图像数据的生成。

GeneGAN^[4]对GAN网络进一步扩展,其编码器部分可以将输入图像的背景和属性进行分离,分别提取图像背景和不同对象特征向量,使用不同的输入图像和提取的特征向量、图像背景信息进行组合,融合生成具有各种不同属性的图像。

为了生成具有多种属性的人脸图像,并使得生成图像能够更好地保留指定属性,本文提出一种具有特征保护的人脸属性转换方法。

2 相关工作

2.1 生成对抗网络

GAN由生成器和判别器两部分组成,其网络框架如图1所示。给定一个随机噪声向量 z ,生成器将 z 映射到生成图像 $G(z)$,把 $G(z)$ 和真实图像 x 作为输入图像输入到判别器,判别图像 $G(z)$ 的真假。

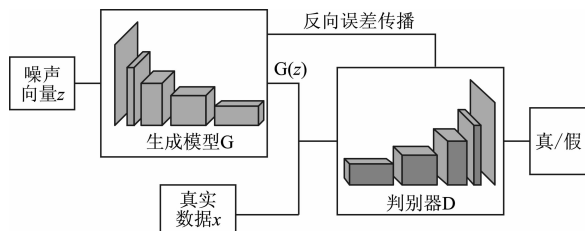


图1 GAN模型图

GAN采用零和博弈的思想,其优化过程是一个极小极大博弈的问题,优化的目标是达到纳什均衡^[5],即判别器无法区分真实图像与生成图像,其优化过程采用目标优化函数。

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

式中, $p_{\text{data}}(x)$ 和 $p_z(z)$ 分别是真实数据和随机噪声的分布, G 、 D 分别代表生成器、判别器的输出。 G 生成的图像越接近真实, $D(G(z))$ 就越大,这时 $V(D, G)$ 会变小,这时会得到式中的 $\min_G$; D 的判别能力越强, $D(x)$ 越大 $D(G(z))$ 就越小,这时 $V(D, G)$ 会变大,即得到 $\max_D$ 。 G 、 D 对抗训练,使最终得到的

生成图像接近于真实图像。

2.2 GeneGAN

GeneGAN^[4]主要是通过对偶学习,在不对称图像集进行目标迁移,比如把一个图像上的眼镜加入到另一幅图像上。其编码器将图像编码成内容(如人脸ID和图像背景)和属性(如微笑、胡须等)的特征向量,解码器则将属性特征向量和新的内容组合形成图像^[6]。通过改变解码器的输入属性,可以实现图像属性的转换。

假设 A_u 、 B_0 分别为具有属性 u (如刘海)和不具有属性 u 的两副不同的人脸图像。生成器将 A_u 、 B_0 分解为人脸 A 、 B 与属性 u ,解码器可使用人脸 B 和属性 u 合成图像 B_u 。为了获得更好地生成图像 B_u ,GeneGAN的解码器还会合成图像 A_u 、 A_0 、 B_0 ,并利用标准 G 损失、重建损失、零位损失、平行四边形损失等多种损失函数约束图像合成。

GeneGAN网络能够有效的分离内容和属性信息,解决了图像生成属性单一的问题。

3 基于特征增强的人脸属性转换模型

本文以GeneGAN为基础,提出一种如图2所示的可进行特征增强的多属性人脸图像生成模型,以源图像和参照属性图像作为输入,通过交换编码的属性部分,解码生成具有参照属性的图像,并通过特征优化模块对生成图像进行特征优化,以保留图像身份特征。这不仅可以改善生成图像的质量,还可以用于对图像集进行数据扩展。

深度神经网络由于其在特征提取方面良好的表现而被广泛应用,在一定范围内,随着网络深度的增加其提取特征能力逐渐增强。U-net网络^[7]不仅在图像分割方面具有很好的表现,在图像恢复精度方面同样具有很好的表现。本文借鉴U-net网络,在图2所示的网络模型中,将GeneGAN的网络深度由3层增加到5层,通过加大网络深度,提高生成图像质量。同时,将图中相互对称的编码器、解码器的中间层相互连接起来,实现编码器部分得到的低维特征与解码器得到的高维特征相结合,作为下一层的输入,这样可以在人脸属性转换的同时更好地保留图像的原始特征,生成更自然逼真的图像。

为提取图像的深度特征,在生成器后添加一个深度神经网络,该模块的主要作用是将数据从图像空间映射到深度特征空间,通过对训练样本和生成图像提取的深度特征进行优化,辅助找到优化后的对象特征,以保证生成更加真实的图像。

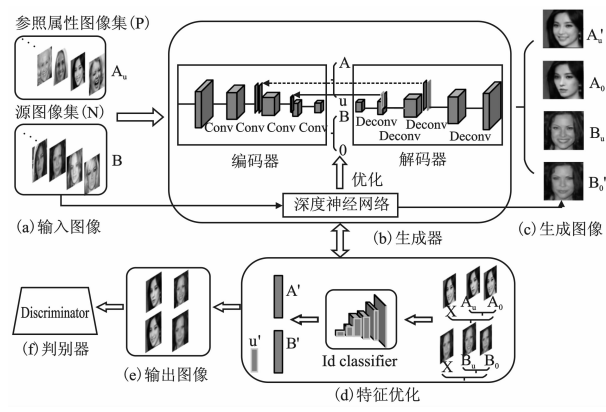


图 2 网络模型整体框架

3.1 生成网络

本文模型中的生成网络由编码器和解码器两部分构成. 输入图像先通过编码器被分离出背景特征和属性特征两部分, 然后通过解码器组合背景特征和对象特征, 生成具有新属性的图片样本. 如图 2 (a), 生成网络接收源图像集(N)和参照属性图像集(P)两组训练图像作为输入, 来自 P 集的图像 A_u 具有微笑属性, 而来自 N 集的图像 B_0 没有该属性, A_u 与 B_0 不要求是同一个人.

与 GeneGAN 相比, 图 2(b) 所示的生成网络中, 将编码器部分得到的低维特征与解码器得到的高维特征相结合, 作为解码器下一层的输入, 这样可以在人脸属性转换的过程中, 更好地保留图像的原始特征, 生成图像结果如图 2(c). 基于生成对抗网络生成的图像都会存在特征损失的情况, 本文在图 2 模型中增加图 2(d) 所示的特征优化模块, 利用深度神经网络特征提出能力将其作为特征优化模型, 通过提取原始图像和生成图像身份特征, 优化特征向量来保留原始身份信息. 最后得到经特征优化后的输出图像 2(e), 再输入到判别器 2(f) 中, 判别输入图像真假像.

判别器是一个普通的卷积神经网络, 在鉴别源图像和生成图像的过程, 会对自身的参数进行更新, 并利用判别器的训练结果指导生成器的训练, 这样可以提高生成图像的质量. 同时, 判别器要求相同属性的 A_u 和 B_u , A_0 和 B_0 无法区分, 从而确保分离出来的 A 中尽量不包含属性 u 的信息, 从而实现背景信息与属性信息的良好分离.

生成网络的编码器和解码器信息参数如表 1 和表 2 所示, 其中 Conv 表示下采样的卷积层, Deconv 表示上采样的反卷积层, 卷积核的大小均为 3×3 . 除网络的第一层不加批量规范化(Batch Normaliza-

tion, BN)^[8] 处理外, 其他层都加入 BN; 同时, 下采样使用 LeakyReLU 激活函数; 上采样层除最后一层使用 Tanh 激活函数, 其他层均使用 ReLU 激活函数.

表 1 编码器信息表

输入大小	网络层	输出大小
[256, 256, 3]	Conv(3 * 3, S=2)	[128, 128, 64]
[128, 128, 64]	Conv(3 * 3, S=2)	[64, 64, 128]
[64, 64, 128]	Conv(3 * 3, S=2)	[32, 32, 256]
[32, 32, 256]	Conv(3 * 3, S=2)	[16, 16, 512]
[16, 16, 512]	Conv(3 * 3, S=2)	[8, 8, 512]
concat [8, 8, 1024]		

表 2 解码器信息表

输入大小	网络层	输出大小
[8, 8, 1024]	Deconv(3 * 3, S=2)	[16, 16, 512]
[16, 16, 512]	Deconv(3 * 3, S=2)	[32, 32, 256]
[32, 32, 256]	Deconv(3 * 3, S=2)	[64, 64, 128]
[64, 64, 128]	Deconv(3 * 3, S=2)	[128, 128, 64]
[128, 128, 64]	Deconv(3 * 3, S=2)	输出图像

3.2 特征优化

为了对图像身份特征、属性特征进行优化, 本文在生成器后, 如图 2(d) 所示, 采用预训练好的深度神经网络 FaceNet^[9], 根据训练数据样本和生成图像提取的深度特征, 求生成图像的特征及属性损失, 获取更优的身份信息.

在深度神经网络中, 较浅的网络层有助于保留图像的对象特征, 而较深的网络层有助于保留图像的属性特征^[10]. 本文使用预训练 FaceNet 的 conv5 来映射身份特征 $f(*)$, 而使用 fc7 来映射属性特征 $h(*)$, 通过对应欧氏空间里的距离来直接关联相似度.

特征优化过程如图 2(d) 所示, 给定一个有或没有某属性的人脸输入图像 x , 通过编码器生成图像 $A_u(B_0)$ 和 $A_0(B_u)$, 如将 x, A_u 和 A_0 输入到预训练模块, 分别得到对应的特征映射, 并计算不同特征映射之间的欧氏距离. 为了优化身份特征, 使 $f(x)$ 与 $f(A_0)$ 的欧氏空间距离尽可能减小; 为了使增加的属性更加真实, 应当使 $f(x)$ 与 $f(A_u)$ 的欧氏空间尽可能减小.

3.3 损失函数

训练过程采用迭代训练策略, 每次通过反馈具有相反属性的一对图像来训练关于特定属性的模型并重复遍历所有属性. 例如: 收集一组微笑图像和另

一组非微笑图像作为输入. 生成的图像样本和原始图像都经过 FaceNet 提取出深度特征, 预训练神经网络模型下提取的特征为 $f(x)$ 、 $f(G(x))$. 重构人脸图像 $G(x)$ 不包含属性的变化, 直接求输入 x 与生成图像 $G(x)$ 在欧氏空间中的均方误差 L_{id} , 具体计算为:

$$L_{id} = \sum_{x \in R} \| f(x) - f(G(x)) \|^2$$

$f(x)$ 、 $f(G(x))$ 是从预训练模型中提取到的输入图像特征和重构图像特征. 感知损失能使生成图像接近输入图像特征, 因此属性优化计算为:

$$L_{attr} = \min_L \| f(x) - f(G(x/Attr)) \|_{L2}$$

$f(G(x/Attr))$ 是从生成图像提取的特征, $G(x/Attr)$ 包含胡须、微笑、刘海等属性的变化. 通过最小化 L_{id} 和 L_{Attr} 达到保护图像身份特征和优化属性特征的目的.

4 实验与分析

本文模型采用 CelebA^[11] 数据集进行了训练, 该数据集是一个公开的大规模的人脸属性数据集, 包括了 10257 个人的 202599 个面部图像, 每个图像都有 40 个属性注释.

4.1 数据集的预处理

在做人脸属性转换时, 只需要人脸区域, 因此需要对 CelebA 数据集进行预处理. 首先通过人脸特征检测定位一幅图像中的人脸并对齐, 然后将检测到的人脸区域进行裁剪, 最后通过仿射变换, 得到尺寸为 256×256 的图像. 将处理后的 200000 张左右的人脸图像作为训练集来训练网络, 实验验证时选取处理后任意一张图像.

4.2 实验结果分析

为了对生成图像与原图像的相似性进行客观评价, 可以使用峰值信噪比 (PSNR) 作为图像质量的衡量标准. PSNR 不考虑人眼的视觉特性, 而是基于对应像素点间的误差进行量化评估的一种评价方法, PSNR 值越大, 图像失真就越小.

使用本文方法与 GeneGAN 网络, 各取 20 组图像与原图像对比, 分别计算 PSNR, 结果如图 5 所示. PSNR_ours、PSNR_Gene 分别是本文算法模型、GeneGAN 网络生成图像的 PSNR 值. 从图 4 中可以看出, 本文算法模型的生成图像质量优于 GeneGAN 网络的生成图像, 增加的属性特征也比较真实.

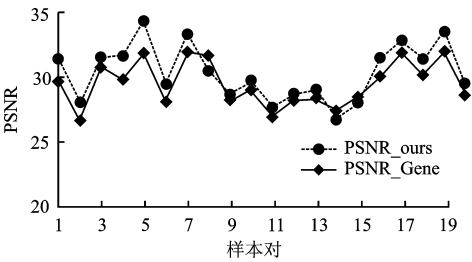


图 3 PSNR 值对比

将本文模型生成的图像、GeneGAN 网络生成图像及源图像输入到深度神经网络中, 求取对应的欧氏空间深度特征差, 差值越小, 说明生成图像与真实图像 ID 特征越接近, 实验结果如图 4 所示. sour_ours、sour_gene 分别是本文算法模型、GeneGAN 网络生成图像的深度特征差. 由图 4 可知, 本文算法模型生成的图像在身份特征保持方面优于 GeneGAN 网络.

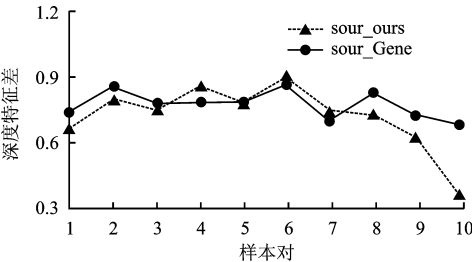


图 4 深度特征差

5 结束语

本文针对 GAN 网络在人脸属性转换时存在的生成图像分辨率低及特征损失问题, 基于 GeneGAN 网络, 提出了一种具有特征保护的人脸属性转换方法. 实验结果表明, 该方法在生成人脸图像时, 通过属性特征优化增强, 能够更好地保持图像身份特征和属性特征.

参考文献:

[1] TRAN L, YIN X, LIU X. Disentangled Representation Learning GAN for Pose-Invariant Face Recognition[C]. 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2017). Honolulu, HI, USA, 2017:1283-1292.

[2] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]. 28th Conference on Neural Information Processing Systems (NIPS2014). Montreal, CANADA, 2014:2672-2680.

输出电压特性,温度系数与 PSRR(电源抑制比)都比同类电路较好.整体电路设计基于传统的带隙基准电路,增加了两级运放,进行参数上的调整,以及加入一个负反馈的环路,有效的保证了输出参考基准电压的稳定性,实现了低功耗和高性能的设计要求.目前,已经成功运用于 BUCK 型 DC-DC 电压芯片中.

参考文献:

[1] XIE F, DAI Q Y. Design of a new CMOS integrated drop-out voltage-regulator IP block[J]. Chinese Journal of Electron Devices, 2009,6(9):1027-1030.

[2] BEHAZD RAZZVI. Design of analog CMOS integrated circuits[M]. CHEN G C. Xi'an:Xi'an Jiaotong University Press, 2002;309-315.

[3] WU J, YAO J N, CHANG C Y. CMOS analog IP linear integrated circuit[M]. Nan Jing: Southeast University Press, 2007;107-112.

[4] HE Z W, JIANG Y F. Design of CMOS bandgap voltage reference circuit with high PSR[J]. Modern Elec-

tronics Technique, 2014,37(13):153-158.

[5] ZHOU Z X, LAI Q T, JIANG Y, et al. A high PSRR bandgap reference[J]. Microelectronics & Computer, 2019,36(5):1-9.

[6] WANG Y X, ZHU B. A CMOS bandgap reference for PWM control buck DC-DC converters [J]. Chinese Journal of Electron Devices, 2013,36(2):252-255.

[7] ZHANG X Z, ZHANG T. A bandgap voltage reference with third - order curvature compensation[J]. Journal of Wuhan University of Science and Technology, 2015,38(1):67-71.

作者简介:

王 晖 男,(1995-),硕士研究生.研究方向为模拟集成电路设计.
E-mail:wanghui_wh@foxmail.com.

张 涛 男,(1967-),博士,教授.研究方向为数模混合集成电路设计.

刘 劲 男,(1981-),博士,副教授.研究方向为 Kalman 滤波器.

(上接第 41 页)

[3] MIRZA M, OSINDERO S. Conditional Generative Adversarial Nets [DB/OL]. arXiv preprint arXiv:1411.1784.

[4] ZHOU S, XIAO T, YANG Y, et al. GeneGAN: Learning Object Transfiguration and Attribute Subspace from Unpaired Data[C]. 28th British Machine Vision Conference (BMVC2017). London, UK, 2017;111. 1-111. 13.

[5] RATLIFF L J, BURDEN S A, SASTRY S S. Characterization and computation of local Nash equilibria in continuous games[C]. 51st IEEE Annual Allerton Conference on Communication, Control and Computing. Monticello, IL, USA, 2013;917-924.

[6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18(7): 1527-1554.

[7] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C]. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAD). Munich, GERMANY, 2015; 234-241.

[8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pat-

tern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.

[9] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A Unified Embedding for Face Recognition and Clustering[C]. 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2015). Boston, MA, USA, 2015;815-823.

[10] JOHNSON J, ALAHI A, LI F. Perceptual Losses for Real-Time Style Transfer and Super-Resolution[C]. 14th European Conference on Computer Vision. Amsterdam, NETHERLANDS, 2016;694-711.

[11] LIU Z, LUO P, WANG X, et al. Deep Learning Face Attributes in the Wild[C]. IEEE International Conference on Computer Vision (ICCV2015). Santiago, CHILE, 2015: 3730-3738.

作者简介:

魏子涵 男,(1999-).研究方向为嵌入式系统设计.

王 慧 女,(1993-),硕士研究生.研究方向为图像处理.

王 晶 女,(1982-),博士,副研究员.研究方向为计算机体系结构、智能芯片设计.

高 岚 女,(1987-),博士,讲师.研究方向为计算机体系结构、GPGPU 体系结构及编程优化、并行编程模型.

张伟功(通讯作者) 男,(1967-),博士,研究员.研究方向为计算机体系结构、系统总线技术.

Email:zwg771@cnu.edu.cn