

基于生成对抗网络的人脸姿态转正方法

王宏志¹, 祖东成¹, 康祺儿²

(1. 长春工业大学 计算机科学与工程学院, 长春 130012; 2. 长春工业大学 电气与电子工程学院, 长春 130012)

摘要: 针对人脸图像姿态转正过程中存在忽略细节特征、转正后的人脸图像纹理模糊、人脸特征与原图像身份特征差距过大的问题, 提出一种基于双判别器双注意力机制的生成对抗网络人脸转正方法. 首先, 在利用双判别器判断人脸身份和人脸姿态的同时, 设计一种人脸表征注意力模块, 增强人脸的整体特征, 防止人脸相关信息缺失, 提高模型完善人脸图像的能力; 其次, 设计一种自适应边缘增强注意力模块, 利用自适应注意力机制和 Sobel 滤波器, 增强人脸边缘细节特征和人脸关键特征, 生成五官和轮廓逼真的正面人脸; 最后, 采用新的归一化层 CrossNorm 提高分布变化下的鲁棒性. 在数据集 Multi-PIE 和 CFP 上进行测试实验的结果表明, 该模型相较于对比模型, 生成的正脸图像人脸转正效果更好.

关键词: 生成对抗网络; 注意力机制; 人脸转正; 人脸姿态

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1671-5489(2025)06-1646-09

Face Posture Correction Method Based on Generative Adversarial Networks

WANG Hongzhi¹, ZU Dongcheng¹, KANG Qi'er²

(1. College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China;
2. College of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: Aiming at the problems existing in the process of face image posture correction, such as ignoring detailed features, blurring the texture of the face image after correction, and excessive gap between the face features and the identity features of the original image, we proposed a face correction method based on the dual discriminator and dual attention mechanism of the generative adversarial network. Firstly, while using the dual discriminator to determine the face identity and face posture, we designed a face representation attention module to enhance the overall features of the face, prevented the loss of face-related information, and improved the model's ability to perfect face images. Secondly, we designed an adaptive edge-enhanced attention module. By using the adaptive attention mechanism and Sobel filter, the edge detail features and key features of the face were enhanced to generate a frontal face with realistic facial features and contours. Finally, a new normalization layer CrossNorm was adopted to improve the robustness under distribution changes, and the results of testing experiments on the Multi-PIE dataset and the CFP dataset show that the proposed model generates frontal face images with better face correction effects compared to the

收稿日期: 2024-06-11.

第一作者简介: 王宏志(1961—), 男, 汉族, 博士, 教授, 博士生导师, 从事数字信号处理与应用、图像处理和电机控制的研究, E-mail: wanghongzhi@ccut.edu.cn. **通信作者简介:** 祖东成(1999—), 男, 汉族, 硕士研究生, 从事图像处理和嵌入式的研究, E-mail: 3145004457@qq.com.

基金项目: 吉林省科技发展计划技术攻关项目(批准号: 20210201051GX).

comparison model.

Keywords: generative adversarial network; attention mechanism; face correction; face posture

目前, 基于深度学习的人脸识别算法在约束场景下已展现出超越人类视觉系统的识别精度. 然而, 现有的研究存在显著局限性, 主流的人脸识别算法训练普遍依赖标准正面化人脸数据库, 但当识别算法目标为极端姿态下的人脸时, 即使是性能最优的正面人脸识别算法, 识别率也会极大下降.

目前解决大姿态人脸识别难题的方法主要分为以下 3 类: 基于姿态鲁棒的特征提取方法、基于正面人脸生成的方法和基于子空间分析的方法. 其中基于特征提取的方法和基于子空间分析的方法都无法有效处理人脸姿态偏转较大的情形, 因此本文采用生成正面人脸的思路解决该问题. 早期的研究通过建立三维人脸模型得到对应的正面人脸, 大多数先从二维面推测三维面, 再从三维面构建正面视图, 但对特征点的要求较高; 当人脸偏转角度稍大时, 二维人脸图片中有些特征点不可见, 因此该方法存在局限性.

近年来, 利用生成对抗网络(generative adversarial network, GAN)^[1]进行人脸转正的方法备受关注. Tran 等^[2]在条件生成对抗网络(CGAN)等网络的基础上提出了 DR-GAN (disentangled representation learning GAN)网络, 生成器在姿态编码的引导下生成所需的侧面人脸图像, 通过判别器对生成图像的真假以及姿态进行分类, 最终得到对应姿态的正面图像. Huang 等^[3]提出了分别在人脸轮廓和人脸细节进行生成的双路径生成对抗网络结构(two-pathway GAN, TP-GAN), 在映射空间将人脸转变为高级特征编码后, 利用多个特征标定网络对不同方面的特征进行处理, 形成逼真的正面人脸. Zhao 等^[4]提出了在姿态鲁棒模型(pose invariant model, PIM)中采用双通道的模型结构, 为人脸重构提供详细的先验引导. Yin 等^[5]提出了一种双注意力机制生成对抗网络(dual-attention GAN, DA-GAN)模型, 其为解码器增添了自注意力机制, 使人脸的纹理细节得到加强, 但忽略了编码时的特征提取也是生成人脸的重要步骤. 文献^[6]提出了双代理姿态引导生成对抗网络(couple-agent pose-guided generative adversarial network, CAPG-GAN), 使用 5 个人脸关键点作为结构先验, 将输入图像与姿态图像进行拼接, 得到正面的生成图像. Duan 等^[7]提出了一种由粗到精的 BoostGAN 架构, BoostGAN 将多遮挡轮廓面作为输入, 然后为每个轮廓生成正面, 最后对多正面进行正面提升.

上述方法更侧重高维度的抽象信息^[8], 而忽略了各阶段特征提取得到的人脸全局信息. 人脸全局信息包含丰富的人脸相关信息和整体网络的架构关联内容, 可用来辅助人脸生成, 但却在编码过程中丢失, 在解码过程中又被忽视. 该问题不仅使大姿态下生成的正面人脸图像丢失关键的特征信息, 细节模糊, 还会导致在非同源的测试数据上得到的生成图像效果太差.

为进一步改善生成的正面人脸图像细节丢失、纹理模糊等问题, 本文设计一种新的 DD-GAN (dual discriminator dual attention mechanism for generative adversarial networks)模型, 将重点关注在各阶段的特征提取部分, 通过对比实验和模块消融实验, 验证本文方法有效性. 本文主要贡献如下:

1) 针对已有结果中忽略图像的人脸全局特征信息, 只关注人脸外观关键信息, 从而导致生成的正面人脸图像信息缺失, 整体特征表现模糊等问题, 设计了人脸表征注意力模块(face representation attention, FRA), 从而扩大感受野、注重空间和通道信息, 尽可能专注人脸的全局相关特征信息, 分配不同的权重, 处理整体人脸特征, 保证生成图像与真实图像的一致性;

2) 针对人脸转正图像重建困难, 视觉效果和准确率检测效果冲突等问题, 设计一种自适应边缘增强注意力模块(adaptive edge-enhanced attention, AEEA), 增强网络捕捉人脸边缘轮廓信息的能力, 自适应地保留最有用的自注意力值进行特征聚合, 找到人脸特征之间的聚类联系, 提高对人脸关键特征的关注, 生成更符合人眼感知的正面图像;

3) 针对模型训练过程中过拟合, 或训练良好的模型在新环境中性能不佳等问题, 增添了新的归一化层 CrossNorm^[9]以扩大训练分布, 增强网络的泛化鲁棒性.

1 DD-GAN 模型

一个仅有卷积、池化和全连接等操作构成的网络模型通常无法精确找到关注的区域, 因此本文将

生成对抗网络的基础模型与注意力机制相结合^[10], 使用人脸表征注意力模块和自适应边缘增强注意力模块, 提高网络对人脸全局特征和人脸细节特征的关注度, 从而提高生成图像的质量. 并从归一化层方面改进网络训练结果.

1.1 网络模型结构

图 1 为 DD-GAN 模型的主体框架. 本文的生成对抗网络由生成器 G 、判别器 D_i 和 D_p 组成, 受多任务学习机制的启发, 生成器 G 由编码器、人脸表征注意力模块、自适应边缘增强注意力模块和解码器组成. 编码器由 10 个卷积块组成, 除第一个卷积块外都嵌入了人脸表征注意力模块. 解码器由 10 个反卷积块构成, 都嵌入了自适应边缘增强注意力模块. 判别器 D_i 和 D_p 则由 14 个卷积块组成, 且在除第一个卷积块前嵌入了人脸表征注意力模块^[11].

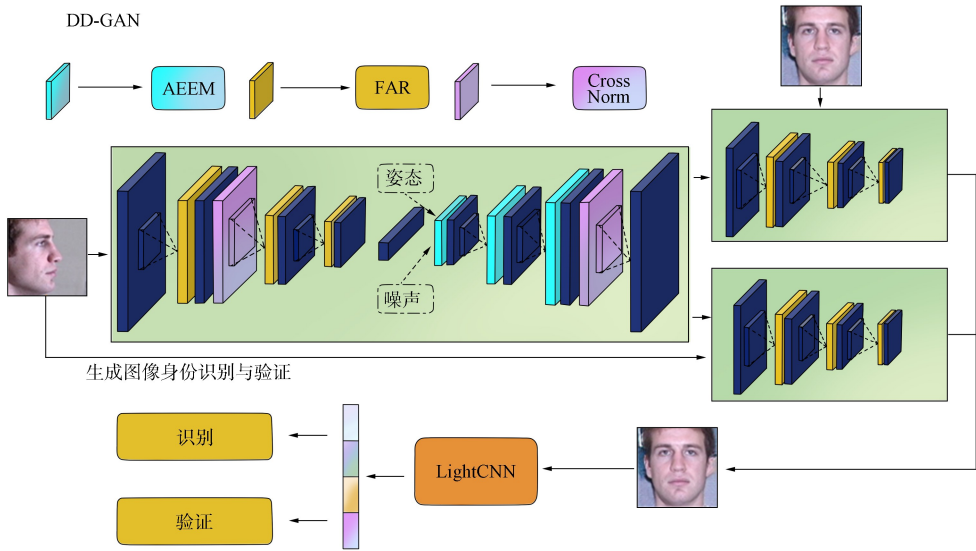


图 1 DD-GAN 模型主体框架

Fig. 1 Main framework of DD-GAN model

生成对抗网络中的编码器先对经过预处理后的图像进行卷积, 图像经过裁剪后的大小为 96×96 , 图像经过一次卷积后, 会将得到的特征向量输入到设计的人脸表征注意力模块 FRA, 将模型的注意力专注在感兴趣的人脸特征区域. 使用激活函数 ReLU, 该激活函数在映射和保留特征时, 会导致一些神经元的输出变化为 0, 从而形成网络的稀疏性, 可有效防止梯度消失和过拟合的形成. 同时采用归一化层 CrossNorm 增加规范化特征和数据的组合, 对卷积后得到的通道施加风格增强, 从而扩大训练分布, 加快收敛速度. 编码器的作用是从输入的人脸图像 x 学习人脸身份表征, 表示为

$$y(x) = G_{enc}(x).$$

解码器的输入是经过编码器和模块 FRA 共同处理的 $y(x)$, 输入为 320 维的特征向量、独热姿态编码^[12] c 和固定噪声 z . 通过第一个反卷积块后的特征向量经过嵌入的自适应边缘增强注意力模块 AEEA, 聚拢人脸轮廓的边缘细节, 通过 Top-K 的贡献分数, 留存细化特征后的重要部分. 解码器的归一化层同样用 CrossNorm 解决分布移位问题. 除最后一层的激活函数使用 tanh, 输出映射到 $[-1, 1]$, 丰富生成人脸细节, 其余卷积块都使用 ReLU. 解码器最后生成的正面人脸图像大小为 $96 \times 96 \times 3$. 双判别器的输入则是真实的人脸及生成人脸, 判别器会分别判断生成人脸的身份和姿态是否符合生成要求, 最后身份判别器与姿态判别器利用 LightCNN^[13] 进行联合概率输出, 判断出符合要求的转正人脸.

1.2 人脸表征注意力模块

观察整张人脸图像, 每个区域实际上都与其他区域有隐性的特征联系^[14], 但在人脸生成的大部分研究中, 都只关注眼睛、嘴和鼻子等局部关键区域, 忽略了其他部分和人脸的关系, 导致一些非关键

区域, 如额头、下巴的特征遗失. 因此, 本文改进注意力模块 CBAM^[15], 设计了人脸表征注意力模块 FRA. 在输入的人脸图像经过卷积处理得到更高维度的抽象特征后, 再使用 FRA 模块, 从而扩大感受野, 对维度特征进行关联聚合, 忽略图片背景信息, 聚拢人脸整体的特征信息, 防止人脸细节丢失. FRA 模块结构如图 2 所示.

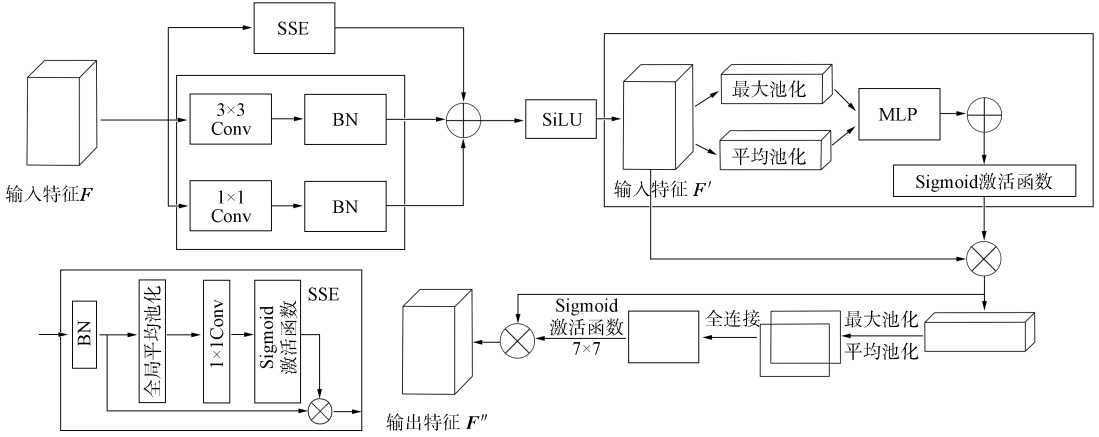


图 2 人脸表征注意力模块结构

Fig. 2 Structure of face representation attention module

本文通过 Conv1×1 将空间维度 $H \times W$ 压缩到 1×1 提取通道信息, Conv3×3 提取局部特征, 增加感受野, 捕捉局部的上下文细节. 利用通道挤压和空间激励(squeeze and spatial excitation, SSE)模块, 通过全局平均池化捕获全局上下文信息, 1×1 卷积调整通道的权重, 达到增强全局特征的效果^[16]. 扩大感受野后, 将处理后的特征图分别经过最大池化层和平均池化层抽取映射空间的全局信息, 再基于多层感知机(MLP)在空间维度上进行特征整合和降维压缩, 由于 FRA 模块的目标是提取捕捉人脸的全局特征^[17], 忽略背景信息干扰, 突出显示空间映射特征哪些与人脸有关, 因此生成一维的通道注意力特征图进行计算, 计算方法如下:

$$F_f(F') = \sigma(\text{MLP1}(F'_{\text{avg}}) + \text{MLP1}(F'_{\text{max}})), \tag{1}$$

$$F_{\text{inter}} = F' \otimes F_f, \tag{2}$$

其中 σ 为非线性激活函数, F_{inter} 为经过乘积得到的特征图, F'_{avg} 为经过平均池化的特征图, F'_{max} 为经过最大池化得到的特征图.

得到人脸表征在通道上的注意力特征图后, 将特征图经过平均池化层和最大池化层操作处理, 关注空间域中的信息区域, 强调人脸特征位置, 抑制背景影响, 并将处理结果连接生成 1×1 维的空间注意力特征图, 得到网络在扩大感受野、收集人脸特征后的面部特征:

$$F'' = F_{\text{inter}} \otimes F_{\text{sig}}, \tag{3}$$

其中 F_{sig} 为经过 7×7 卷积核运算后的二维空间注意力特征图.

1.3 自适应边缘增强注意力模块

初步聚合人脸特征后, 网络的目标从粗略的人脸全局转到细微的人脸局部, 旨在生成真实的人脸图像. 自适应边缘增强注意力 AEEA 模块利用边缘增强^[18], 注重人脸轮廓细节特征, 采用 Top-K 贡献分数^[19], 保留最重要的部分, 细化关键特征, 帮助生成网络生成正面人脸. AEEA 模块网络结构如图 3 所示.

经处理过后的人脸特征图已突出了一些重要的视觉信息, 为进一步细化人脸关键表征, 将特征图输入 Sobel 滤波器, 进行水平和垂直方向的边缘检测, 结合两个方向计算的边缘强度, 找到像素变化剧烈的区域, 即人脸轮廓区域与外部交界处的边缘点. 滤波后, 通过 BN 层、ReLU 激活函数和 3×3 卷积以渐进的方式增强轮廓的表达. 最后, 将全局特征图与经过边缘增强的轮廓特征图残差连接、加权求和, 再利用 1×1 卷积使特征压缩为单通道, 进行后续的处理.

融合计算后的特征图将提供更丰富的信息和重要的人脸轮廓细节输入下一层网络结构, 推断更具

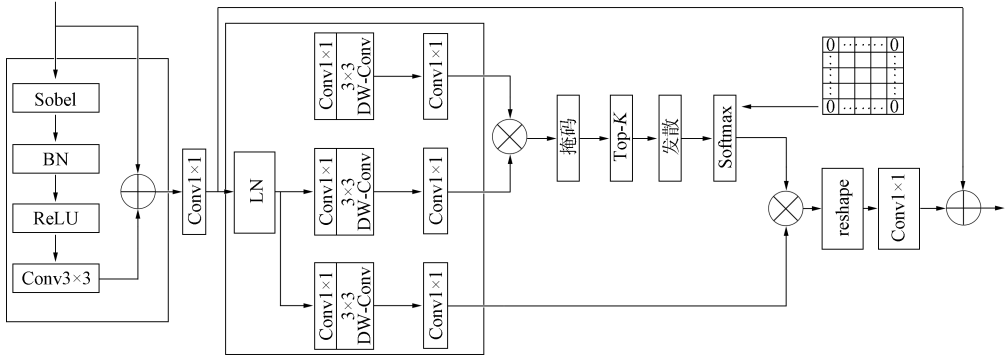


图 3 AEEA 模块网络结构

Fig. 3 Network structure of AEEA modules

细节的人脸特征. 输入特征经过 1×1 卷积, 得到注意力的查询 Q 、键 K 和值 V 的混合表示后, 对混合表示进行深度可分离. 即将 Q, K 和 V 的混合表示沿通道维度分割成 3 份, 分别表示为新的 Q, K 和 V 变量. 为利用多头注意力的计算表达, 将赋值后的变量重新排列, 每个注意力头的 Q, K 将进行归一化, 以便更好地计算注意力分数. 再将多个注意力头创建对注意力分数不会造成影响的全零张量作为掩码, 控制注意力中的关注分布^[20]. 掩码的创建为注意力分数特征图的计算提供了前置帮助, 将 Q 和 V 的转置矩阵进行矩阵相乘, 得到每个位置上的分数, 即衡量 Q 和 K 的相似度, 查询向量应该对每个键向量分配多少注意力. 然后使用 Top-K 方法, 分别获取每个位置上前 K 个元素、返回值和对应的索引. 其中, K 是根据训练过程中输入数据动态计算, 即根据输入的维度自适应地选择前 K 个重要关系.

凭借这些索引对注意力分数^[21]进行调整, 使用 Scatter 发散操作在每个注意力头的掩码张量中将相应位置的值设为 1, 其他位置不变. 这些设置好的掩码张量, 将保留注意力分数中对应位置 1 的元素, 其他位置元素设置为一 ∞ , 最后连接所有的输出, 经过线性投影得到想要的特征^[22]. 完成屏蔽较低关注权重分配的不必要特征, 可突出人脸的细节特征, 从而实现 Top-K 的自适应分配权重的目的, 计算公式如下:

$$\text{Att}(Q, K, V) = \text{Softmax}\left(T_k\left(\frac{QK^T}{\omega}\right)\right)V, \tag{4}$$

其中 ω 为动态控制稀疏度大小的可调参数, T_k 为自适应分配的 Top-K 算子, 表示为

$$(T_k(S))_{ij} = \begin{cases} S_{ij}, & S_{ij} \geq t_i, \\ 0, & \text{其他}, \end{cases} \tag{5}$$

t_i 为 QK^T 第 j 行中第 k 个最大值.

1.4 损失函数

面对不同种类的 GAN 网络, 损失函数在网络的训练和拟真效果上具有关键的制约作用, 能防止在训练中发生过拟合、训练时间过长和生成样本数据质量差等问题. 本文根据人脸表征的各种特性, 嵌入双判别器损失函数、生成器损失函数、感知损失函数、身份保持损失函数和对称损失函数, 使生成的图像在肉眼真实度和特征距离验证方面更接近真实的正面人脸图像.

1.4.1 双判别器损失函数和生成器损失函数

由于生成网络的特征提取不能使特征明确地分类为身份特征和姿态特征, 会导致身份丢失等问题, 所以本文设计双判别器, 联合 D_i 和 D_p 的概率输出, 使用以下双判别器损失函数进行制约优化:

$$L_D = \min_G \max_D V(D, G) = E_{x \sim P_d(x), z \sim P_z(z)} [\log(D_{i(N^i+1)}(\tilde{x})) + \log(D_{i(N^i+1)}(G(x_1 \cdots x_n, z)))] + E_{x \sim P_d(x), z \sim P_z(z), c \sim p_c(c)} [\log(D_{p(N^i+1)}(\tilde{x})) + \log(D_{p(N^i+1)}(G(x_1 \cdots x_n, z, c)))] \tag{6}$$

G 和 D 在迭代训练中交替优化, 双判别器损失函数使 G 生成虚假图像的能力和 D 判别出虚假图像的能力达到平衡. 所以, 为生成器可以制衡判别器, 其损失函数为

$$L_G = \max_G V_G(\mathbf{D}, \mathbf{G}) = E_{x \sim P_d(x), z \sim P_z(z)} [\log(\mathbf{D}_{i(N^i)}(\mathbf{G}(x_1 \cdots x_n))) + (\mathbf{D}_{p(N^i)}(\mathbf{G}(x_1 \cdots x_n, z)))]. \quad (7)$$

1.4.2 感知损失函数

仅使用生成器和双判别器之间的对抗损失会导致生成图像的纹理失真, 为改善该问题, 本文采用基于 LightCNN 网络 ReLU 的感知损失, 以确保在生成图像过程中尽量少丢失图像的纹理细节, 其公式为

$$L_{pl} = \sum_n \frac{1}{C_n H_n W_n} \sum_{i=1}^{H_n} \sum_{j=1}^{W_n} \|\Phi_n(x)_{ij} - \Phi_n(\mathbf{G}(x))_{ij}\|, \quad (8)$$

其中 n 为 LightCNN 在第 n 个最大池化层的特征表达, Φ_n 为特征大小.

1.4.3 身份保留损失函数

身份保留损失函数由 LightCNN 网络的最后两层计算实现, 借助网络特征提取的功能, 凸显人脸的结构特征, 保持身份的稳定性的, 表达式为

$$L_{id} = \sum_{k=1}^2 \frac{1}{W_k \times H_k \times C_k} \sum_{x=1}^{W_k} \sum_{y=1}^{H_k} \sum_{z=1}^{C_k} |\mathbf{F}(I^g)_{x,y,z}^k - \mathbf{F}(I^t)_{x,y,z}^k|, \quad (9)$$

其中 I^g 和 I^t 分别表示生成图像和输入的真实图像, \mathbf{F} 为 LightCNN 的中间层部分.

1.4.4 对称损失函数

对称损失函数主要利用人脸左右部分之间的差异度量, 衡量生成图像人脸对称区域之间的像素差距. 以对称损失函数作为训练指标, 可使生成图像的质量得到提升, 减少遮挡等外部影响, 计算公式如下:

$$L_s = \frac{1}{W/2 \times H \times C} \sum_{x=1}^{W/2} \sum_{y=1}^H \sum_{z=1}^C |I_{x,y,z}^g - I_{x,y,z}^t|, \quad (10)$$

1.4.5 总体损失函数

模型的总体损失函数为

$$L = L_G + L_D + \lambda_1 L_s + \lambda_2 L_{id} + \lambda_3 L_{pl}, \quad (11)$$

其中 $\lambda_1, \lambda_2, \lambda_3$ 表示不同损失的权重参数.

1.5 CrossNorm 归一化层

为增强训练的分布, 生成更好的图像效果, 本文引用 CrossNorm 归一化层, 嵌入设计网络, 观察所用训练数据集的总体风格, 即对输入图像的每个通道进行风格观察, 从而在 C 通道维度之间或者核心之间交换统计信息. 通过这种方法, CrossNorm 能有效填补训练过程中因风格变化而导致的空白, 从而提升正面人脸图像生成的多样性和质量.

2 实验

2.1 实验环境及数据集选择

本文在 Windows10 系统和 3090 显卡进行模型训练, 使用 Pytorch1.13.1 深度学习框架, 所用数据集为 Multi-PIE 和 CFP. 其中数据集 Multi-PIE 是在不同姿态下光照和表情同时受限情况下发布的最大的多视图人脸识别数据集, 它包含了 337 个对象的 754 204 张人脸图像. 为准确评估本文模型性能, 将采用数据集内 seting1 的内容进行训练验证, seting1 有 250 个身份对象, 选择姿态角度为 $-90^\circ \sim 90^\circ$ 内的 13 种偏转角度进行数据集训练. 数据集 CFP 则是一个专用于多角度人脸验证的数据集, 由 500 个不同对象共 7 000 张人脸组成. 每个对象提供 10 张正脸照片和 4 张侧脸照片, 其中正脸图像之间的姿态变化小于 10° , 而正脸与侧脸之间的角度变化接近 90° .

2.2 对比实验

2.2.1 基于数据集 Multi-PIE 的评估

为确保 DD-GAN 模型结构的有效性, 将本文模型与现有的基于生成对抗网络算法和基础特征提

取网络,如 DR-GAN,TP-GAN,FNM,LightCNN 在不同数据集上进行人脸识别结果对比,验证模型性能.表 1 列出了不同模型在数据集 Multi-PIE 上的识别精度对比结果.

表 1 不同模型在数据集 Multi-PIE 上的识别精度对比结果

Table 1 Comparison results of recognition accuracy of different models on Multi-PIE dataset

模型	姿态角度						%
	±15°	±30°	±45°	±60°	±75°	±90°	
TP-GAN	97.93	96.34	94.71	84.46	70.11	59.82	
DR-GAN	95.30	92.62	81.76	78.21	67.33	33.56	
FNM	98.02	96.53	95.20	83.61	62.31	41.20	
LightCNN	99.76	94.82	70.94	31.58	9.37	2.43	
DD-GAN	98.12	97.01	95.38	85.19	70.90	60.54	

模型利用 LightCNN 提取生成图像和真实图像的特征,计算两者之间的余弦距离,从而得出具体的人脸识别精度.由表 1 可见,正面人脸与侧面人脸之间的旋转角度越大,模型的转正效果越差,本文 DD-GAN 模型的人脸转正效果明显优于其他模型,能在微小旋转角度保持相对优势.在人脸大姿态环境下的识别,如 60°和 75°的极端旋转角度,DD-GAN 模型的识别精度也优于其他模型,表明本文模型能在大姿态场景下更好地保留身份信息,生成质量优异的人脸正面图像.

2.2.2 基于数据集 CFP 的评估

为测试本文模型在不同数据集上的泛化能力,采用数据集 CFP 对 DD-GAN 模型进行评估.在数据集中采用 200 个相同对象和 200 个不同对象的测试对进行验证.重复上述操作 10 次,取 10 次结果的平均值与其他模型进行对比.表 2 列出了不同模型在数据集 CFP 上的人脸识别精度对比结果.由表 2 可见,DD-GAN 模型在不同数据集上对数据的泛化能力和生成效果的鲁棒性都优于其他模型.在从正面到正面这一生成图像结果中,DD-GAN 模型利用优化后且有针对性的特征提取,生成了人脸识别精度较高的正面人脸,超过了 TP-GAN 等模型的准确率,略低于 LightCNN 模型.但在正面到侧面的人脸验证中,DD-GAN 模型的人脸识别精度超过了其他对比模型,表明得到了更逼真的正面人脸图像,从而证明了该算法的有效性.

表 2 不同模型在数据集 CFP 上的人脸识别精度对比结果

Table 2 Comparison results of face recognition accuracy of different models on CFP dataset

模型	正面-正面		模型	正面-侧面		%
	正面-正面	正面-侧面		正面-正面	正面-侧面	
TP-GAN	97.85	92.41	LightCNN	98.29	87.22	
DR-GAN	97.08	91.02	DD-GAN	98.01	93.23	
FNM	97.26	92.95				

2.2.3 人脸转正可视化结果

图 4 为 DD-GAN 模型生成的正面人脸图像.由图 4 可见,各角度的侧面人脸图像经过 DD-GAN 模型处理后生成的正面人脸图像质量优异,小姿态变化人脸生成的转正图像视觉效果逼真.大姿态偏转图像经过 DD-GAN 模型聚合人脸全局信息,增强轮廓信息,找寻关键特征后的生成效果也十分优异,能保留大部分人脸特征,清晰辨别身份信息.

2.3 消融实验

为证明本文设计的人脸表征注意力模块、自适应边缘增强注意力模块和替换 CrossNorm 归一化层的有效性,使用特征向量的余弦距离作为评价指标,在数据集 CFP 上进行训练,以验证人脸偏转嵌入不同模块后的性能.表 3 列出了各模块消融实验结果.由表 3 可见,采用双判别器生成对抗网络作为基线,将不同的模块嵌入模型中,注意力机制在模型生成图像的质量上具有关键作用.在数据集 CFP 侧面到正面的人脸验证中,基础模型的侧面人脸识别精度为 81.73%,增加人脸表征注意力模块 FAR,将模型关注区域转到整体人脸,弥补了基础网络提取人脸全局特征不足的问题,此时识别精度为 84.18%.在基础模型中加入自适应边缘增强注意力模块 AEEA,注重人脸轮廓和人脸特征的内部关系,此时识别精度为 84.54%,模型性能得到增强.最后替换卷积层后的批量归一化层,换为

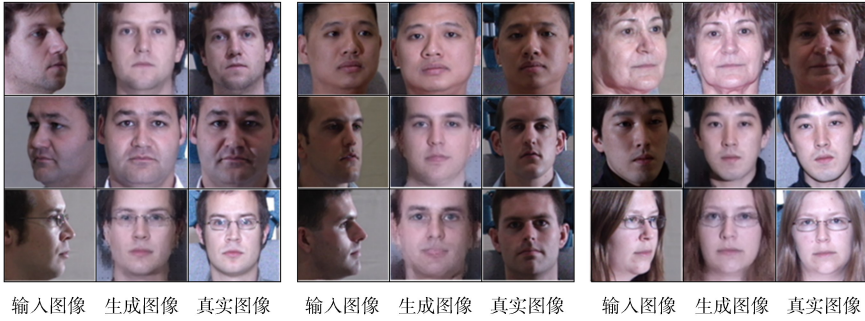


图 4 DD-GAN 模型生成的正面人脸图像

Fig. 4 Frontal face images generated by DD-GAN model

CrossNorm, 替换后的模型和基础模型相比, 效果提升了 0.48 个百分点.

表 3 各模块消融实验结果

Table 3 Results of ablation experiments for each module

模型	正面-正面	正面-侧面	模型	正面-正面	正面-侧面
基础	84.21	81.73	+AEEA	88.39	84.54
+FAR	87.05	84.18	+CrossNorm	85.36	82.21

综上所述, 针对人脸识别训练过程中人脸全局身份特征丢失、关键特征捕捉能力低、视觉效果差的问题, 本文设计了一种用于人脸转正的 DD-GAN 模型, 通过人脸表征注意力模块 FRA 和自适应边缘增强模块 AEEA 增强模型生成正面人脸的能力, 并在提高人脸转正性能的同时引入了 CrossNorm 归一化层, 增强模型的泛化性. 实验结果表明, 本文 DD-GAN 模型生成效果, 无论是在人眼主观判断的视觉效果还是客观的人脸识别精度上都优于其他的生成正面人脸算法.

参 考 文 献

[1] KRICHEN M. Generative Adversarial Networks [C]//2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). Piscataway, NJ: IEEE, 2023: 1-7.

[2] TRAN L, YIN X, LIU X M. Disentangled Representation Learning GAN for Pose-Invariant Face Recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 1415-1424.

[3] HUANG R, ZHANG S, LI T Y, et al. Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 2439-2448.

[4] ZHAO J, CHENG Y, XU Y, et al. Towards Pose Invariant Face Recognition in the Wild [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 2207-2216.

[5] YIN Y, JIANG S Y, ROBINSON J P, et al. Dual-Attention GAN for Large-Pose Face Frontalization [C]//2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020). Piscataway, NJ: IEEE, 2020: 249-256.

[6] VALLE R, BUENAPOSADA J M, BAUMELA L. Reducing Head Pose Estimation Data Set Bias with Synthetic Data [J]. IEEE Access, 2025, 13: 73530-73539.

[7] DUAN Q Y, ZHANG L. Look More Into Occlusion: Realistic Face Frontalization and Recognition with Boostgan [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 214-228.

[8] TANG Z, GAO Y, ZHU Y, et al. Crossnorm and Selfnorm for Generalization under Distribution Shifts [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 52-56.

[9] 张曦, 库少平. 基于生成对抗网络的人脸超分辨率重建方法 [J]. 吉林大学学报(工学版), 2025, 55(1): 333-338. (ZHANG X, KU S P. Face Super-resolution Reconstruction Method Based on Generative Adversarial Network [J]. Journal of Jilin University (Engineering and Technology Edition), 2025, 55(1): 333-338.)

- [10] HUANG Y S, ALHLFFEE M H B. Improving Face Recognition by Integrating Decision Forest Into GAN [J]. *Applied Artificial Intelligence*, 2023, 37(1): 531-565.
- [11] 解奕鹏, 秦品乐, 曾建潮, 等. 基于姿态注意力的特定角度人脸正面化网络 [J]. *计算机技术与发展*, 2023, 33(7): 47-54. (XIE Y P, QIN P L, ZENG J C, et al. Point-Specific Face Frontal Network Based on Postural Attention [J]. *Computer Technology and Development*, 2023, 33(7): 47-54.)
- [12] 朱立忠, 李永壹. 基于生成对抗网络的人脸转正方法研究 [J]. *通信与信息技术*, 2023(6): 22-25. (ZHU L Z, LI Y Y. Research on Face-Positive Method Based on Generated Adversarial Network [J]. *Communication and Information Technology*, 2023(6): 22-25.)
- [13] WU X, HE R, SUN Z, et al. A Light CNN for Deep Face Representation with Noisy Labels [J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(11): 2884-2896.
- [14] MOSTOFA M, SAADABADI M S E, MALAKSHAN S R, et al. Pose Attention-Guided Profile-to-Frontal Face Recognition [C]//2022 IEEE International Joint Conference on Biometrics (IJCB). Piscataway, NJ: IEEE, 2022: 1-10.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.l.]: ECCV, 2018: 3-19.
- [16] GOYAL A, BOCHKOVSKIY A, DENG J, et al. Non-deep Networks [J]. *Advances in Neural Information Processing Systems*, 2022, 35: 6789-6801.
- [17] ZHANG Z Z, LAN C L, ZENG W J, et al. Relation-Aware Global Attention for Person Re-identification [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 3186-3195.
- [18] 陈婷, 王通, 张冀武, 等. 基于边缘先验融合动态门控特征的人脸图像修复 [J]. *计算机应用研究*, 2023, 40(11): 3478-3484. (CHEN T, WANG T, ZHANG J W, et al. Face Image Repair Based on Edge Prior Fusion of Dynamic Gated Features [J]. *Research on Computer Application*, 2023, 40(11): 3478-3484.)
- [19] CHEN X, LI H, LI M Q, et al. Learning a Sparse Transformer Network for Effective Image Deraining [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2023: 5896-5905.
- [20] 王威威, 罗杨, 李熹. 融合注意力机制和多层次特征的语义分割方法 [J]. *工业控制计算机*, 2025, 38(4): 93-95. (WANG W W, LUO Y, LI X. Semantic Segmentation Method Integrating Attention Mechanism and Multi-level Features [J]. *Industrial Control Computer*, 2025, 38(4): 93-95.)
- [21] 杨青, 于明, 阎刚. 基于深度频率特征注意力机制的图像去雨方法 [J]. *吉林大学学报(工学版)*, 2024, 54(10): 2952-2962. (YANG Q, YU M, YAN G. Image Rain Removal Method Based on Depth Frequency Feature Attention Mechanism [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2024, 54(10): 2952-2962.)
- [22] 王进, 王猛旗, 张昕跃, 等. 基于多头注意力机制字词联合的中文命名实体识别 [J]. *江苏大学学报(自然科学版)*, 2024, 45(1): 77-84. (WANG J, WANG M Q, ZHANG X Y, et al. Chinese Named Entity Recognition Based on Multi-head Attention Mechanism Word Union [J]. *Journal of Jiangsu University (Natural Science Edition)*, 2024, 45(1): 77-84.)

(责任编辑: 韩 啸)