

基于改进 CycleGAN 的人脸剪纸自动生成^①

牟 玮, 童 晶, 韦 剑, 张明懿

(河海大学 信息科学与工程学院, 常州 213022)

通信作者: 童 晶, E-mail: tongjing.cn@gmail.com



摘 要: 针对人脸剪纸手工设计难度大, 制作周期长等问题, 本文首次利用生成对抗网络生成高质量人脸剪纸. 面向人脸剪纸艺术特点, 提出了一种基于 CycleGAN 的改进网络: 1) 在原始 CycleGAN 生成器中引入 CBAM 注意力模块, 增强网络特征提取能力; 2) 引入针对鼻、眼、唇等关键面部区域的局部鉴别器, 提升人脸剪纸中以上区域的生成效果; 3) 设计基于图像边缘信息与 SSIM 的损失函数, 取代 CycleGAN 的前向循环一致损失, 消除所得人脸剪纸中的阴影. 相较于其他人脸剪纸自动生成方法, 本文方法可快速生成与原始人脸相似度高、线条连续流畅、具有艺术美感的人脸剪纸. 此外, 本文还提出了一种人脸剪纸连通性后处理方法, 使所得结果更符合中国传统剪纸整体连通的特点.

关键词: 人脸剪纸; 剪纸生成; 生成对抗网络; 连通性

引用格式: 牟玮, 童晶, 韦剑, 张明懿. 基于改进 CycleGAN 的人脸剪纸自动生成. 计算机系统应用, 2023, 32(12): 1-11. <http://www.c-s-a.org.cn/1003-3254/9318.html>

Automatic Portrait Paper-cut Generation Based on Improved CycleGAN

MU Wei, TONG Jing, WEI Jian, ZHANG Ming-Yi

(College of Information Science and Engineering, Hohai University, Changzhou 213022, China)

Abstract: Aiming at the difficulty in manually designing portrait paper-cuts, this study employs the generative adversarial network (GAN) to generate high-quality portrait paper-cuts for the first time. Based on the artistic characteristics of portrait paper-cuts, an improved network based on CycleGAN is proposed. 1) The CBAM attention module is introduced into the CycleGAN generator to enhance the feature extraction of the network. 2) The local discriminator for key facial regions such as nose, eyes, and lips is introduced to improve the generation effect of the above areas in generated portrait paper-cuts. 3) A new loss function is designed based on image edge information and SSIM, which will be adopted to replace the original forward cycle-consistency loss of CycleGAN and eliminate the shadows in the portrait paper-cuts. Compared with other automatic generation methods of portrait paper-cuts, the proposed method can quickly generate paper-cuts featuring high similarity to the original human face, continuous and smooth lines, and aesthetic beauty. Additionally, this study also puts forward a post-processing method of portrait paper-cut connectivity to make the obtained results more consistent with the overall connectivity of traditional Chinese paper-cuts.

Key words: portrait paper-cut; paper-cut generation; generative adversarial network (GAN); connectivity

剪纸是我国宝贵的非物质文化遗产, 人脸剪纸是其中深受欢迎的一类. 人脸剪纸以干净流畅的线条、

镂空的艺术手法、简洁的二值逻辑刻画原始人脸特征^[1]. 保持与原始人脸的相似性且兼具艺术美感是人脸

① 基金项目: 国家自然科学基金 (61772172)

收稿时间: 2023-05-17; 修改时间: 2023-06-26; 采用时间: 2023-07-07; csa 在线出版时间: 2023-10-25

CNKI 网络首发时间: 2023-10-26

剪纸的基本要求,这极为考验剪纸艺术家的专业能力,故纯手工制作人脸剪纸通常难度大、周期长。因此,许多学者从事剪纸自动生成研究^[1-6],其中不乏面向人脸剪纸的自动生成方法^[1,3,4]。

现有剪纸自动生成方法通常将剪纸生成视为图像二值化问题,具体方法可被划分为两类^[3]: 1) 基于阈值的二分类方法^[2,4,5]; 2) 基于图像分割与能量优化的方法^[6]。前者对原始图像整体或分区计算阈值,依据阈值将其二值化,得到类似剪纸的效果;后者将原始图像分割为一系列区域,设计与这些区域特性相关的能量函数,通过优化能量函数为各区域分配二值标签。而对于人脸剪纸生成,另有学者提出了基于人脸五官剪纸模板库的方法^[1,3]。该方法首先对预收集的大量人脸剪纸进行五官分割提取,创建五官模板库。由人脸照片生成人脸剪纸时,分别从五官模板库中选取与当前人脸五官最相近的五官剪纸模板,将其组合得到完整人脸剪纸。另外,可通过对五官剪纸模板进行变形^[1],使其更符合原始人脸五官形状。以上3种方法各有不足: 1) 基于阈值的方法易受原始人脸照片分辨率及光照影响,所得结果多有噪声,且存在边缘粗糙,线条不流畅,不易于剪刻等问题。此外,往往需人工对阈值进行调节方可取得最佳结果; 2) 基于图像分割与能量优化的方法以分割出的小区域作为二值化基本单元,若将人脸分割为若干区域后进行二值化,会破坏人脸语义特征,所得结果往往无法满足人脸剪纸相似性与艺术性要求; 3) 基于五官剪纸模板库的方法需预先建库,工作量大,且人脸剪纸生成效果有赖于五官模板库内剪纸模板的质量与数量,无法保证稳定的优质结果。另外,基于模板库的方法需对人脸五官之外的区域做额外处理,例如头发区域或眼镜等配饰。

针对现有人脸剪纸自动生成方法存在的以上不足,本文将人脸剪纸生成视为图像风格化任务,将人脸图像由真实照片风格转换为剪纸风格。就图像风格化这一任务而言,现有方法可分为非真实感渲染(non-photo-realistic rendering, NPR)、神经风格迁移(neural style transfer)、图像翻译(image-to-image translation)等3个方向。NPR领域对人物肖像进行风格化的方法包括基于笔触的方法^[7-9]、基于区域的方法^[3,10-12]、基于纹理迁移的方法^[13,14]。基于笔触的方法基本思想是模拟艺术作品中的笔画放置方式,而剪纸并非基于笔触的艺术形式,故不能使用此类方法。基于区域的方法首先

将输入图片分解为不同区域,然后要么分别将这些区域与艺术家作品模板进行匹配^[3,10,11],要么对不同区域分别采取不同的绘制算法^[12],最后将所有区域组合得到完整肖像作品。该方法流程较复杂,且方法可迁移性不强。基于纹理迁移的方法将样本图像中的纹理迁移至目标图像,对于油画等纹理信息丰富的艺术风格可取得较好结果,不适用于剪纸等纹理信息较少的艺术作品。神经风格迁移^[15]利用神经网络分别从内容图像 A 和风格图像 B 中提取内容特征与风格特征,生成内容与 A 一致,风格与 B 一致的目标图像。同样地,此类方法^[16-18]对于纹理信息丰富的艺术风格能取得较好结果,但剪纸纹理信息较少,以其作为目标风格往往存在风格迁移不彻底,图像内容发生形变等问题。且基于样例的神经风格迁移方法从单张图捕获风格特征,所得风格准确度往往不及使用大量数据的基于学习的方法。图像翻译^[19]旨在学习两个图像域间的映射关系,进而可将源域图像转换至目标域。相较于神经风格迁移,图像翻译适用于更广范围的风格转换任务,且其转换效果往往更优,故受到了大量关注。Isola等^[19]提出的Pix2Pix是基于“源域-目标域”成对训练数据进行图像翻译的代表,而CycleGAN^[20]是在缺少成对数据时进行图像翻译的典型方法,许多学者面向不同的应用场景对CycleGAN进行了扩展与改进^[21,22]。由于大量“人脸照片-对应剪纸”成对数据获取不易,本文采取基于CycleGAN的图像翻译方法。

CycleGAN作为通用图像翻译框架,并不专门面向人脸图像翻译任务,故对于具有高度语义约束的人脸图像无法取得好的转换效果。若将原始CycleGAN用于人脸照片与人脸剪纸转换,所得人脸剪纸会出现部分五官缺失或形变、面部存在阴影和模糊、面部轮廓不完整、线条不够连续流畅等问题。因此,本文对原始CycleGAN进行改进,使其可生成与原始人脸相似度高、面部线条整洁流畅、易于剪刻、具有艺术美感的人脸剪纸。

此外,整体连通是传统剪纸的一大重要特点,现有剪纸自动生成方法要么忽略了该问题^[1,4,6],要么用固定的连接线生成单一死板的连通结果^[3]。也有人将此任务交由用户手动处理^[2],或采用会破坏人脸特征与剪纸美观性的连通性处理方法^[2,5]。以上方式在自动化程度及处理效果方面各有不足,因此,本文提出一种新的人脸剪纸连通性后处理方法。

本文的贡献包括: 1) 将人脸剪纸自动生成视为图像风格化任务, 首次利用生成对抗网络生成高质量人脸剪纸. 2) 提出改进的 CycleGAN, 设计基于注意力机制的生成器, 减少面部五官残缺及形变现象; 引入针对鼻、眼、唇区域的局部鉴别器, 提升人脸剪纸中以上区域的生成效果; 设计基于结构相似性指标 (structural similarity, SSIM) 及图像边缘信息的损失函数, 替换 CycleGAN 的前向循环一致损失, 消除所得人脸剪纸中的阴影与模糊. 3) 提出了一种对人脸剪纸面部特征破坏较少、符合人工设计原则的人脸剪纸连通性后处理方法.

1 改进的 CycleGAN

CycleGAN 是由 Zhu 等^[20]提出的图像翻译模型, 它用于学习两个图像域 X, Y 间的映射关系, 从而可将两个域中的图像进行转换, 其工作原理如图 1 所示.

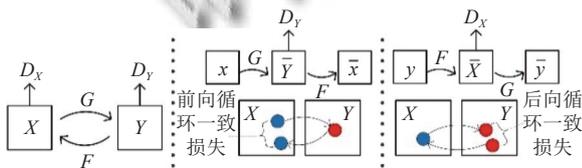


图 1 CycleGAN 原理结构

CycleGAN 具有两个结构相同的生成器 G, F , 前者用于学习从 X 域到 Y 域的映射, 后者学习从 Y 域到 X 域的映射. 同时 CycleGAN 具有两个结构相同的鉴别器 D_x, D_y . D_x 用于区分输入图像是 X 域真实图像还是生成器 F 合成的图像, D_y 用于区分 Y 域真实图像和生成器 G 合成的图像. 此外 CycleGAN 引入了循环一致性损失^[20], 分为前向循环一致损失与后向循环一致损失, 前者要求 X 域中任一图像 x 均能通过图像转换循环被还原为原始图像, 即 $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, 后向循环一致损失与此类似, 要求 $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$.

本文将人脸照片域记为 X , 将人脸剪纸域记为 Y , 理论上可利用 CycleGAN 的生成器 G 将任意人脸照片 x 转换为对应剪纸 $G(x)$. 而人脸图像具有高度语义约束, 人脸剪纸尤其注重面部特征的表达, 但原始 CycleGAN 对面部特征的提取与表达能力有限, 无法得到令人满意的结果. 具体表现有: 生成器对人脸图像的特征提取能力有限, 生成图像时易受冗余特征影响, 对部分重要面部特征表达不足; 鉴别器的鉴别结果易受次要区域影响, 对关键面部特征关注不足. 生成器与鉴别器的上

述不足会导致所得人脸剪纸 $G(x)$ 中部分五官残缺或形变. 此外, CycleGAN 的前向循环一致损失期望原始人脸照片 x 与重建人脸照片 $F(G(x))$ 具有像素级别的一致性. 而由于人脸剪纸域信息远少于人脸照片域, 故由 $G(x)$ 重建人脸照片时很难得到与 x 像素级别一致的重建结果. 该矛盾会使得生成器 G 合成人脸剪纸时在其中嵌入冗余信息, 而这会导致 $G(x)$ 中存在阴影与模糊, 也会导致部分区域发生形变.

对于 CycleGAN 的上述局限性, 本文对其进行改进. 首先在生成器中引入 CBAM 注意力模块^[23], 设计 CBAM-res 模块取代生成器中原始残差块, 提高生成器网络的特征提取能力, 增强有用特征并抑制无用特征响应. 引入针对鼻、眼、唇区域的局部鉴别器, 用于加强网络对这 3 个关键面部区域的生成能力, 缓解部分五官缺失或形变的问题. 此外, 结合 SSIM^[24]、图像边缘信息设计新的损失函数 $L_{\text{forw_new}}$, 取代原始前向循环一致损失. 改进的 CycleGAN 原理结构图如图 2 所示, 图中仅展示了由人脸照片生成人脸剪纸时的原理.

1.1 生成器

CycleGAN 的两个生成器采用相同的编码-解码结构. 编码器主要由卷积层组成, 解码器主要由反卷积层组成, 两者之间是数个残差块. 原始生成器对人脸图像的特征提取能力有限, 生成图像时易受冗余特征影响, 对部分重要面部特征表达不足, 所得人脸剪纸中面部轮廓及五官会存在缺失或形变.

对于此问题, 本文在生成器中引入 CBAM 注意力模块, 使网络更加关注重要特征层和重要空间区域. 该模块是一种由空间注意力机制和通道注意力机制混合而成的注意力模型, 其示意图如图 3 所示. 本文在残差块中引入 CBAM, 记为 CBAM-res 块, 其结构见图 4. CBAM-res 块中的跳跃连接增加了数据传输支路, 也确保了梯度信息能够有效地在深层网络中进行传递. 而引入的 CBAM 模块利用通道注意力机制为各通道特征添加权重, 使网络更加关注重要程度高的特征层, 降低次要通道的冗余信息对生成图像的影响. 同时空间注意力机制给人脸图像中眼睛、鼻子、嘴巴等重要区域添加更大的权重, 给脸颊、头发等次要区域添加较小权重. 通过将两种注意力模块串联加入残差块中, 生成器网络可以更加有效地提取和表达重要特征, 抑制无用特征, 从而提升人脸剪纸的生成质量.

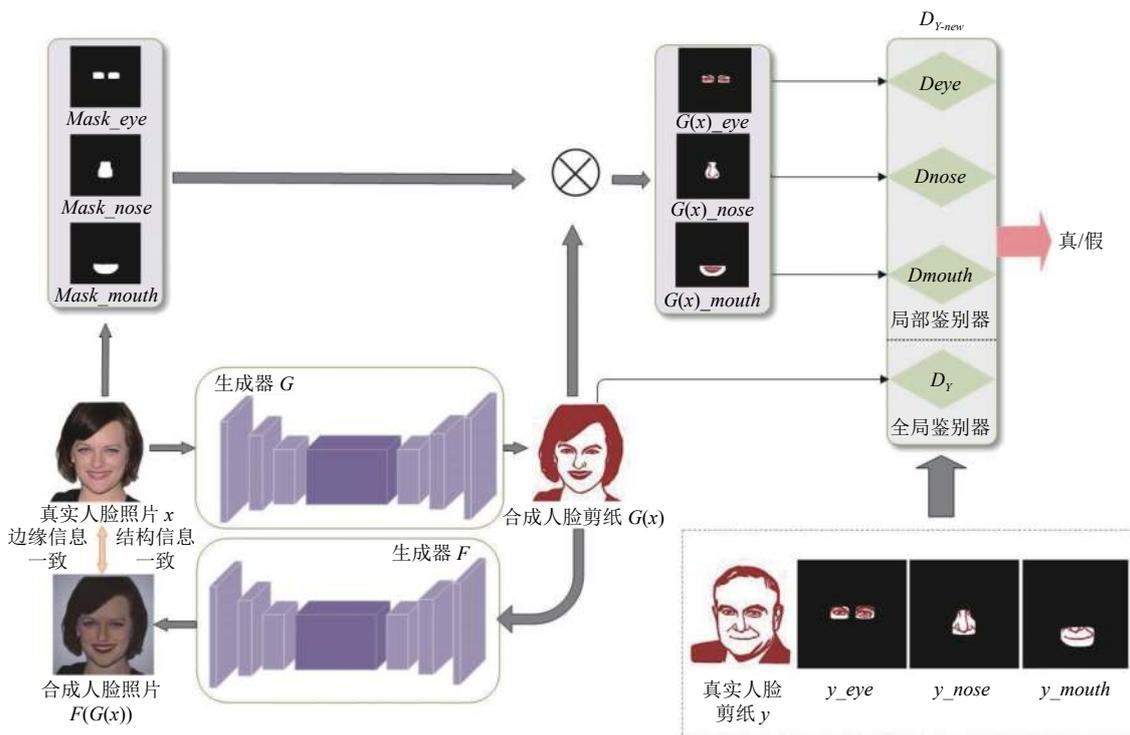


图2 改进 CycleGAN 原理结构

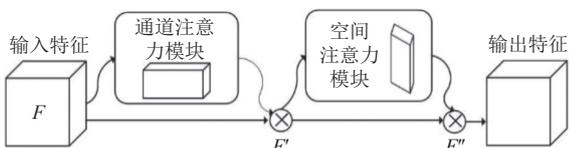


图3 CBAM 注意力机制

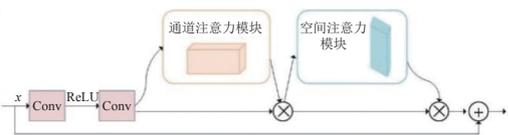


图4 CBAM-res 模块

1.2 鉴别器

CycleGAN 中鉴别器 D_X 与 D_Y 具有相同的结构, 均采用 70×70 的 PatchGAN 网络结构, 判断感受野为 70×70 的局部图像块是否为真. 直接将原始 CycleGAN 的鉴别器应用于人脸剪纸生成任务时, 次要区域的鉴别结果会对完整图像鉴别结果产生影响. 有时即使关键面部特征生成效果不佳, 完整人脸图像依旧会被鉴别为真. 因此, 仅采用鉴别器 D_X, D_Y 会使合成的人脸剪纸 $G(x)$ 存在面部区域生成效果不佳的问题, 例如部分结果中眼睛、鼻子、嘴巴等区域存在残缺或丢失现象. 为加强上述区域的生成效果, 本文在保留 CycleGAN 原始鉴别器 D_X, D_Y 的前提下, 引入分别对人脸剪纸中

眼睛、鼻子、嘴巴等重要区域进行鉴别的局部鉴别器 $D_{eye}, D_{nose}, D_{mouth}$. 将本文对人脸剪纸进行鉴别的鉴别器记为 D_{Y-new} , 它包括原始 CycleGAN 中的全局鉴别器 D_Y 及本文引入的 3 个局部鉴别器.

局部鉴别器也采用了 PatchGAN 结构, 但其输入不同于全局鉴别器. 全局鉴别器 D_Y 的输入为真实人脸照片 x , 而局部鉴别器的输入是对人脸照片进行遮罩处理后的图像. 以局部鉴别器 D_{eye} 为例, 若欲利用 D_{eye} 对真实人脸照片 x 进行局部鉴别, 则首先利用人脸解析网络^[25]提取 x 的眼睛区域, 并将该区域适度扩大. 然后将眼睛区域像素值设置为 255, 其他区域像素值设置为 0, 得到掩膜图像 $Mask_{eye}$. 接着, 通过式 (1) 得到局部鉴别器 D_{eye} 的输入图像 y_{eye} :

$$y_{eye} = Mask_{eye} / 255 \times y \quad (1)$$

若要利用 D_{eye} 对合成人脸照片 $G(x)$ 进行鉴别, 则应对人脸照片 x 进行面部解析, 得到掩膜图像 $Mask_{eye}$, 采用相同方式得到 D_{eye} 的输入图像 $G(x)_{eye}$.

1.3 损失函数

CycleGAN 的损失函数如式 (2), 其中, $L_{GAN}(G, D_Y)$ 、 $L_{GAN}(F, D_X)$ 为对抗损失, L_{cyc_forw} 为前向循环一致损失, L_{cyc_back} 为后向循环一致损失. 对于 X 域的每一张

图像 x , CycleGAN 的前向循环一致性要求图像 x 依次经过生成器 G 与 F 后能被还原回 x . 式 (3) 为 CycleGAN 的前向循环一致损失.

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y) + L_{GAN}(F, D_X) + \lambda L_{cyc_forw} + \lambda L_{cyc_back} \quad (2)$$

$$L_{cyc_forw}(G, F) = \mathbb{E}_{x \sim X} [\|F(G(x)) - x\|_1] \quad (3)$$

人脸照片域与人脸剪纸域信息丰富度不均衡, 后者所含信息更少, 因此, 合成的人脸剪纸 $G(x)$ 所含信息不足以生成与 x 像素级相似的重建照片 $F(G(x))$. 而 CycleGAN 前向循环一致损失在 x 与 $F(G(x))$ 间加入了像素级一致的强约束, 使生成器 G 生成人脸剪纸 $G(x)$ 时嵌入多余信息, 进而使 $F(G(x))$ 像素级接近 x . 而多余信息的嵌入会使 $G(x)$ 中存在阴影与模糊, 且将导致面部特征发生形变. 同时, CycleGAN 在原始人脸照片与重建人脸照片间使用的 $L1$ 损失逐像素比较二者差异, 未考虑人类视觉感知, 因此也会降低 $G(x)$ 的艺术效果. 此外, 观察人脸剪纸作品可发现, 当人脸剪纸与原始人脸保持边缘信息一致, 整体结构信息一致即可满足可辨识性及相似性需求.

综上所述, 本文基于以下 3 个想法设计了新的前向循环一致损失 L_{forw_new} : 1) L_{forw_new} 要对原始人脸照片 x 与重建人脸照片 $F(G(x))$ 间的相似性加以约束, 如此方可驱动生成器 G 生成与原始人脸相似, 可辨识的人脸剪纸; 2) 不要求 x 和 $F(G(x))$ 像素级别一致, 仅要求二者边缘信息一致、结构信息一致; 3) 衡量 x 与 $F(G(x))$ 差异时应考虑人类视觉感知与人脸剪纸艺术特点. L_{forw_new} 具体形式如式 (4) 所示:

$$L_{forw_new}(G, F) = \mathbb{E}_{x \sim X} [k_1 \|Hed(F(G(x))) - Hed(x)\|_1] + \mathbb{E}_{x \sim X} [k_2 (1 - SSIM(F(G(x)), x))] \quad (4)$$

其中, Hed 为边缘检测模块^[26], 可提取出图像边缘信息, L_{forw_new} 中第 1 项使 x 与 $F(G(x))$ 的边缘信息一致. 第 2 项中的 $SSIM$ 为衡量图像相似性的结构相似性指标, 该指标利用均值、方差和协方差衡量图像局部亮度、对比度及结构相似性, 进而度量图像整体的相似性, 其评价结果与人主观感知较为一致.

2 人脸剪纸连通性处理

2.1 现有连通性处理方法

整体连通是传统剪纸的重要艺术特点, 现有剪纸自动生成方法通常将其作为后处理选项. Meng 等^[3]

预定义了一组用来加强人脸剪纸连通性的连接线, 该方法需根据人脸结构比例的不同调整连接线形态, 且处理效果单一. Lei 等^[2]为用户提供了处理剪纸连通性的接口, 由用户选定需要连通的任意两个区域 P, Q , 并给定一点 M , 系统会自动绘制一条经过 M 点的直线用于连接 P 和 Q . 该方法自动化程度低, 且用直线作为连接线有损所得剪纸的美观性. Xu 等^[5]利用 Dijkstra 最短路径搜索算法确定能将孤立区域进行连接的最短路径. 路径搜索时以图像像素作为搜索结点, 边权由像素灰度值确定. Xu 等^[5]所提方法在自动化程度及处理效果方面均优于上述两种方法. 但人脸剪纸图案简约且面部语义约束强, 且人类视觉系统对包括多余线条在内的人脸特征的错误极为敏感, 故其方法不适用于人脸剪纸的连通性加强. 具体存在如下不足: 1) 以给定区域的所有边界像素点依次作为出发点进行路径搜索, 且对搜索范围不加限制, 时常会导致确定的连接线形态或位置不合理, 破坏面部语义特征. 例如, 利用 Xu 等^[5]的方法对图 5(a) 进行连通性后处理, 所得结果图 5(b) 中下眼皮与鼻子连接在一起. 2) 仅由像素灰度值确定 Dijkstra 算法中的边权, 由此得到的连接线往往会破坏人脸剪纸美感, 且所得的连接线不符合人工设计原则.



图 5 利用 Xu 等的方法对非连通人脸剪纸进行后处理

2.2 人脸剪纸连通性后处理

本文提出了一种新的人脸剪纸连通性后处理方法, 主要工作在于: 1) 首先利用面部轮廓线初步处理连通性, 降低后续计算复杂度; 2) 不同于 Xu 等^[5]直接在全局范围内处理孤立区域, 本文首先在各五官小区域中处理局部连通性, 确保面部五官各自的完整性及语义特征, 然后再处理全局连通性; 3) 利用 Dijkstra 算法搜索路径时设计了考虑图像梯度信息的新的边权, 得到更加符合人工设计原则、不破坏面部美观性的连接线. 整体流程可见图 6, 具体分析及说明见下文.

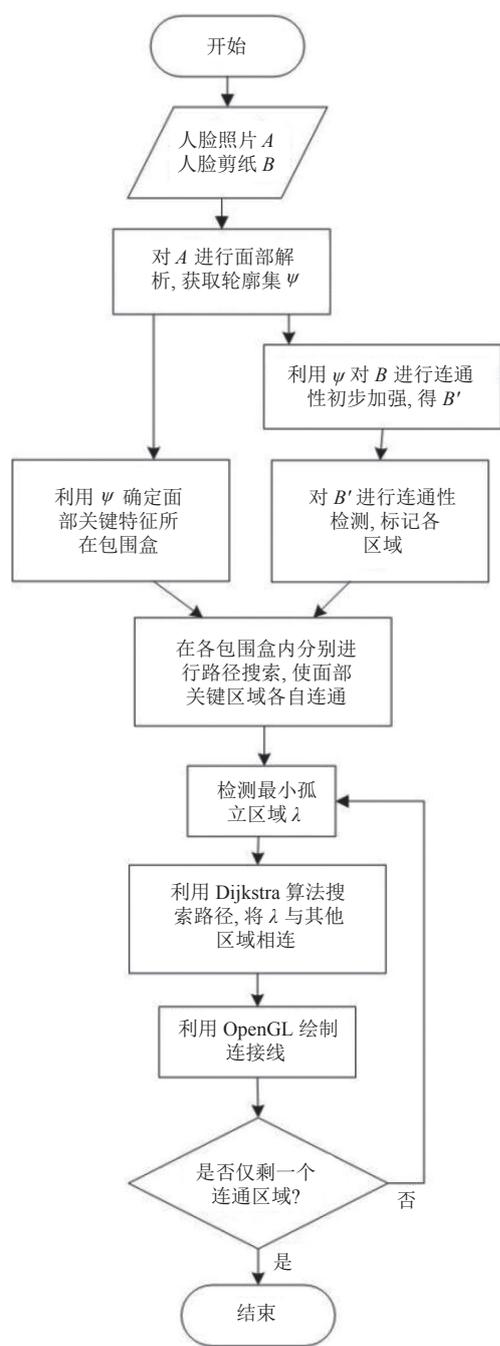


图6 人脸剪纸连通性后处理方法

首先, 利用人脸解析网络^[25]确定人脸剪纸中眼睛、鼻子、嘴巴、眉毛、头发等区域的轮廓线, 通过所得轮廓线对人脸剪纸图像进行初步连通性后处理. 即在原始人脸剪纸中附加上述轮廓线, 利用轮廓线将部分离散区域连接起来. 然后, 对剪纸图像进行八连通扫描线种子填充标记, 对同一区域内的像素点取相同标签, 进而确定各区域.

接下来, 借助 Dijkstra 最短路径搜索算法将孤立区域和其他区域连接起来. 不同于 Xu 等^[5]直接选取全局最小孤立区域, 并确定其边界点到达其他区域的最短路径, 本文首先在各局部区域内执行此过程. 具体地, 首先利用面部区域轮廓线分别确定眉毛、眼睛、鼻子、嘴巴区域的包围盒, 并对其适度扩大. 然后, 分别在各包围盒内执行 Dijkstra 最短路径算法, 使每个包围盒内的区域连通. 在合理小范围内进行搜索一方面保证了面部特征语义完整, 另一方面提高了整体搜索效率. 在面部五官各自连通后, 检测最小的孤立区域 λ , 并在全局范围内搜索路径, 将其与其他区域连接.

由于剪纸艺术家手工制作人脸剪纸时往往会沿着面部线条添加辅助连接线, 例如眼角的皱纹, 鼻翼与嘴角间的法令纹等. 因此, 本文希望利用 Dijkstra 算法搜索出的路径除了满足长度较短、经过面部较暗区域两个条件外, 还尽可能沿着面部原有线条. 故设计了将梯度信息考虑在内的边权重, 使搜索确定的连接线符合人工设计原则. 该权重形式如式 (5) 所示:

$$w(p_i, p_j) = v(p_i) + v(p_j) + \frac{k}{G(p_i) + G(p_j)} \quad (5)$$

其中, p_i 、 p_j 为人脸灰度图中的一对邻接像素点, 也是路径搜索算法中的搜索结点. $w(p_i, p_j)$ 为结点 p_i 、 p_j 间的权重, $v(p_i)$ 为 p_i 的灰度值, $v(p_j)$ 同理. $G(p_i)$ 、 $G(p_j)$ 分别为 p_i 、 p_j 的梯度值.

确定了区域 λ 引出的连接线后, 利用 OpenGL 的线条绘制函数将其绘制进人脸剪纸图像中, 并更新剪纸连通状态. 重复以上步骤, 直至人脸剪纸整体连通.

3 实验分析

本文实验主要包括两部分: 由改进的 CycleGAN 生成非连通人脸剪纸、人脸剪纸连通性后处理.

3.1 非连通人脸剪纸自动生成实验

3.1.1 数据准备

本文实验数据包括人脸照片与人脸剪纸. 人脸照片来自公开数据集 CelebAMask-HQ, 选取其中 492 张用于训练, 200 张用于测试. 人脸剪纸来源于网络, 共挑选出风格统一的 441 张作为训练数据, 测试集无需提供人脸剪纸. 首先借助 U²-Net^[27]网络提取原始人脸照片及人脸剪纸中的人像区域, 例如对于图 7(a) 中的人脸照片, 可利用 U²-Net 得到如图 7(b) 所示的人像 mask

图, 然后借助 OpenCV 得到去除背景的纯人像图 7(c). 接着对纯人像图片进行人脸对齐、尺寸调整、裁剪, 得到尺寸为 256×256 的人脸照片与人脸剪纸, 以此作为最终的训练及测试数据.

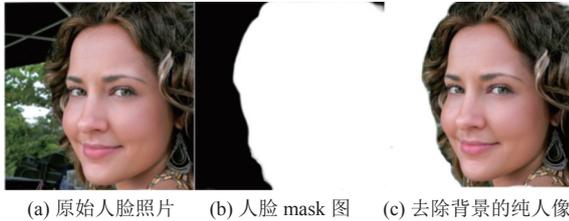


图 7 获取去除背景的人脸照片

3.1.2 实验设置与训练

本实验所使用的的硬件配置为 24 GB 的 NVIDIA GeForce RTX 3090 显卡, 深度学习框架为 PyTorch 1.7.1, Python 版本为 3.8.5, CUDA 版本为 11.0, 操作系统为 Ubuntu 18.04. 批处理大小设为 1, 式 (4) 中 k_1 为 10, k_2 为 5.

在训练过程中检查生成器的生成结果, 发现训练至第 200 个 epoch 时生成器 G 即可生成令人满意的人脸剪纸, 故一共训练 200 个 epoch. 对关键损失函数 $L_{\text{forw_new}}$ 及 $L_{\text{cyc_back}}$ 绘制损失曲线图, 结果分别如图 8、图 9 所示. 对于 CycleGAN 而言, 前向循环一致损失及后向循环一致损失整体呈下降趋势则说明模型训练情况良好, 本文训练结果符合该要求. 而对于 GAN 而言, 损失函数的收敛情况能够说明的问题有限, 训练的最终目的应是生成器能够生成好的结果. 故可通过观察比较生成器 G 生成的人脸剪纸, 确定模型训练效果, 后文设置了相关对比试验.

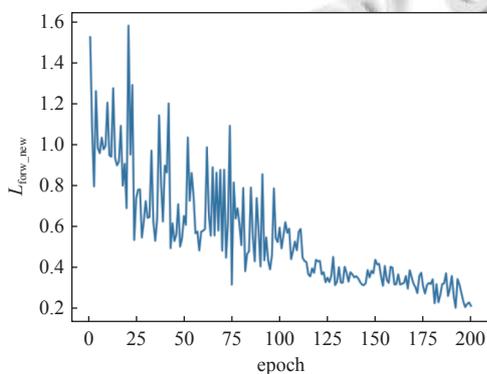


图 8 本文新的前向循环一致损失 $L_{\text{forw_new}}$ 的收敛图

对于 GAN 训练速度慢、训练不稳定的问题, 本文编码实现时用最小二乘损失函数代替交叉熵损失函数,

该损失在训练时更加稳定. 此外, 为了减少模型的振荡, 训练时遵循 Shrivastava 等^[28]的策略, 更新鉴别器时使用历史生成的图像, 而非最新的生成器所得图像. 对于鉴别器学习慢的问题, 采取 TTUR^[29]策略, 让鉴别器收敛更快.

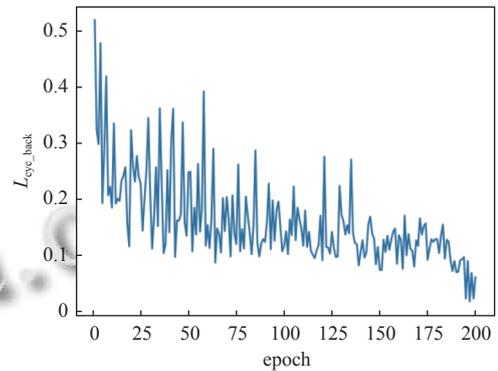


图 9 后向循环一致损失 $L_{\text{cyc_back}}$ 的收敛图

3.1.3 实验结果展示

1) 与阈值法及 CycleGAN 对比

本节首先将改进的 CycleGAN 所得人脸剪纸与阈值处理法^[5]及原始 CycleGAN 所得结果进行对比, 对比结果如图 10 所示. 不难发现, 阈值法所得结果通常无法很好地体现人物面部细节, 往往噪声过多、线条不流畅、不易于实际剪裁. 而 CycleGAN 所得人脸剪纸部分面部区域出现缺失或形变, 有时面部存在阴影或模糊. 改进的 CycleGAN 所得结果干净简约、线条清晰流畅、兼具相似性与艺术美感.

2) 与剪纸模板库法对比

Meng 等^[3]自行制作了五官剪纸模板库, 并由此生成人脸剪纸, 图 11(a)、图 11(b) 分别为文献 [3] 中 Meng 等提供的人脸照片与基于模板库生成的对应剪纸. 本文利用改进的 CycleGAN 生成图 11(a) 对应的人脸剪纸, 结果可见图 11(c).

可以发现, 改进 CycleGAN 所得结果能够满足人脸相似度需求. 此外, 正如前文所言, 基于模板库的方法难以取得稳定的优质结果, 其结果有赖于模板的丰富度, 对于和五官模板相似度较低的人脸通常难取得好的结果. 同时, 基于五官模板库的方法需对人脸以外的区域做单独处理, 例如头发、颈部等, 整体流程复杂. 而对于图 12 中存在面部配饰(眼镜等), 或闭着眼, 或非正脸的特殊人脸照片, 基于五官模板库的方法便不再适用. 相较于基于模板库的方法, 本文方法可生成稳

定结果,且不存在预先建立模板库的大工作量,对于特殊人脸照片也可得到较好的生成效果。

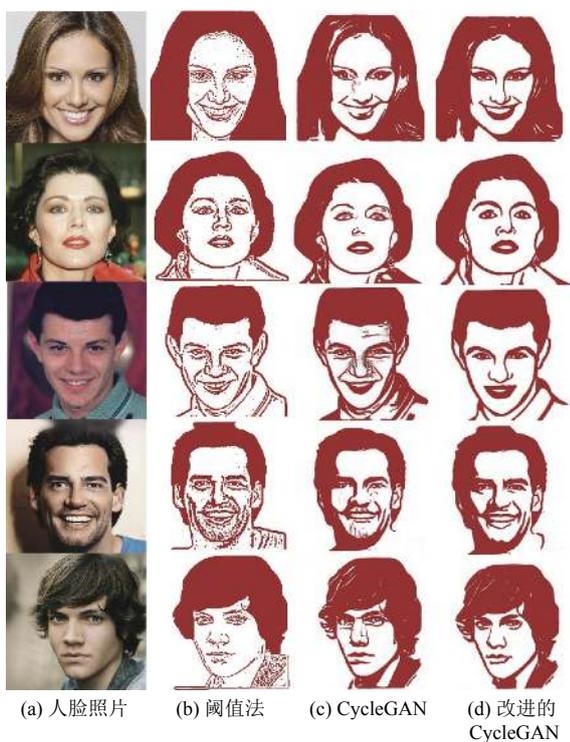


图 10 与阈值法及原始 CycleGAN 所得结果进行对比

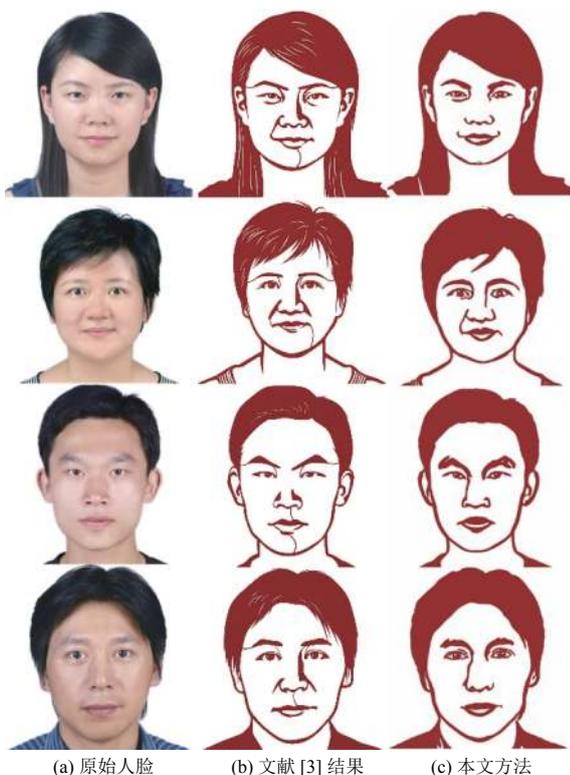


图 11 与基于剪纸模板库的方法进行对比



图 12 本文方法对特殊人脸可取得较好结果

3.1.4 消融实验

设置消融实验探究本文新增及改进模块对整体模型的提升效果,分别测试 CBAM-res 模块、局部鉴别器、改进的前向循环一致损失 L_{forw_new} 对模型的有效贡献.从改进 CycleGAN 中分别去除 CBAM-res 块(替换为原始 CycleGAN 的残差块)、去除局部鉴别器、去除 L_{forw_new} (替换为原始 CycleGAN 的前向循环一致损失),所得结果如图 13 所示.图 13 中“w/o CBAM-res”“w/o D_{local} ”“w/o L_{forw_new} ”分别与上述 3 种情况对应.可以发现,去除 CBAM-res 块后人物面部轮廓及五官出现缺失,去除局部鉴别器后面部特征存在不同程度的形变和缺失,去除 L_{forw_new} 后所得人脸剪纸中存在面部阴影及模糊,五官也略有变形.这些结果表明本文做出的改进及新加的模块都是必要且有效的.

3.1.5 定量分析

本文采用 FID 作为合成图像的评价指标,通过计算真实图像和生成图像的特征向量之间的距离,评价二者之间的相似度,分数越低表示合成图像越趋近于真实图像. FID 公式如下:

$$FID(x, g) = \|\mu_x - \mu_g\|_2^2 + Tr(\lambda_x + \lambda_g - 2\sqrt{\lambda_x \lambda_g}) \quad (6)$$

其中, x 和 g 分别为真实图像和生成图像; μ_x 、 μ_g 为图像特征的均值; λ_x 、 λ_g 为图像特征的协方差矩阵; $Tr(\cdot)$ 为矩阵的迹. 分别利用阈值法、CycleGAN、改进 CycleGAN 生成若干人脸剪纸, 并分别计算 3 种合成人脸剪纸集与真实人脸剪纸集间的 FID 值. 根据表 1 可以发现本文所提改进 CycleGAN 的 FID 值最低, 说明本文方法生成的人脸剪纸与真实人脸剪纸最相近.

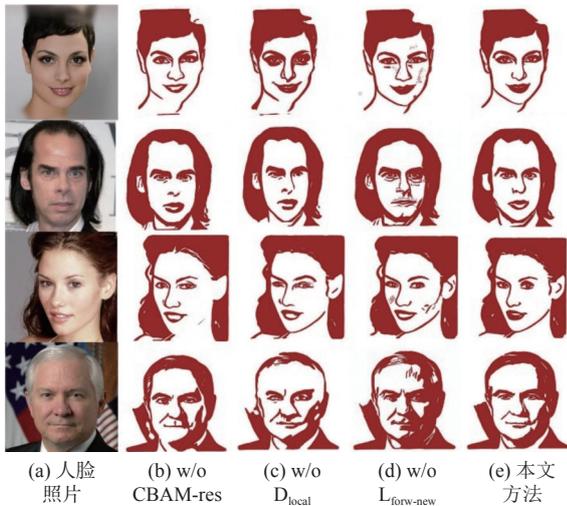


图 13 消融实验结果

表 1 3 种方法对应的 FID

方法	阈值法	CycleGAN	改进的CycleGAN
FID	291.7	217.6	173.9

3.2 连通性处理实验

将本文提出的人脸剪纸连通性后处理方法与 Xu 等^[5]所提连通性处理方法进行对比, 图 14 为对比结果图. 观察图 14 可发现, 文献 [5] 确定的连接线通常计算机处理痕迹明显, 会对人脸面部特征及美观性造成一定程度的破坏. 而本文方法所得连接线或沿着面部原有线条, 或位于对面部特征不造成明显破坏的位置, 符合人工设计习惯, 且能满足人脸剪纸美观性需求.

3.3 实际剪刻效果

由于本文方法所得人脸剪纸线条连续流畅, 面部干净简洁, 所以易于手工或使用自动雕刻机对其进行剪刻. 利用 xTool M1 自动切割机对本文所得人脸剪纸进行雕刻, 图 15 为雕刻机设备图, 实际雕刻效果可见图 16, 图 17 为更多实验结果. 不难发现, 无论是改进

CycleGAN 所得非连通结果, 或是进行连通处理后的连通人脸剪纸, 均可取得较好的雕刻效果.



图 14 人脸剪纸连通性后处理实验结果

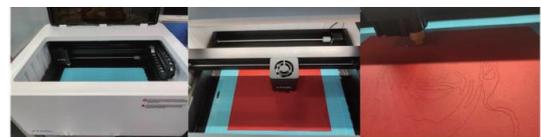


图 15 xTool M1 切割机



图 16 利用自动切割对所得人脸剪纸进行雕刻



图 17 对本文方法所得人脸剪纸进行实际雕刻

4 结论与展望

对于人脸剪纸人工设计难度大,制作周期长,现有自动生成方法效果不佳等问题,本文将人脸剪纸生成视为人脸风格化任务,首次利用生成对抗网络实现高质量人脸剪纸自动生成。面向人脸剪纸艺术特点,分别对 CycleGAN 的生成器、鉴别器、损失函数做了适应性改进。此外,本文提出了一种人脸剪纸连通性后处理方法。实验结果表明本文提出的改进 CycleGAN 能够生成与原始人脸相似度高,具有艺术美感,易于雕刻的人脸剪纸。本文所提人脸剪纸连通性处理方法能生成符合人脸剪纸人工设计习惯、不破坏剪纸美观性的连通结果。但由于剪纸数据集限制,本文所提剪纸自动生成方法目前仅适用于人脸,无法生成其他非人像主题的剪纸,在通用性上逊于阈值处理法等现有方法。

参考文献

- 1 王晓宇,赵妍,刘泽云,等.基于五官特征的人脸剪纸生成方法研究.计算机工程与应用,2018,54(17):208–213.[doi:10.3778/j.issn.1002-8331.1704-0334]
- 2 Lei HY, Qiu WL. The conversion from line drawings to paintings with paper-cutting designs. Proceedings of the 2011 International Conference on Computer Science and Network Technology. Harbin: IEEE, 2011. 2783–2786. [doi:10.1109/ICCSNT.2011.6182541]

10.1109/ICCSNT.2011.6182541]

- 3 Meng M, Zhao MT, Zhu SC. Artistic paper-cut of human portraits. Proceedings of the 18th ACM International Conference on Multimedia. Firenze: ACM, 2010. 931–934. [doi:10.1145/1873951.1874116]
- 4 Dang K, Song GL, Ming AL. Chinese paper-cut generated from human portrait. Proceedings of the 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW). Nantou: IEEE, 2016. 1–2. [doi:10.1109/ICCE-TW.2016.7520943]
- 5 Xu J, Kaplan CS, Mi XF. Computer-generated papercutting. Proceedings of the 15th Pacific Conference on Computer Graphics and Applications (PG 2007). Maui: IEEE, 2007. 343–350. [doi:10.1109/PG.2007.10]
- 6 Xu J, Kaplan CS. Artistic thresholding. Proceedings of the 6th International Symposium on Non-photorealistic Animation and Rendering. Annecy: ACM, 2008. 39–47. [doi:10.1145/1377980.1377990]
- 7 Wang TH, Collomosse JP, Hunter A, et al. Learnable stroke models for example-based portrait painting. Proceedings of the 2013 British Machine Vision Conference. Bristol: BMVA Press, 2013.
- 8 Berger I, Shamir A, Mahler M, et al. Style and abstraction in portrait sketching. ACM Transactions on Graphics, 2013, 32(4): 1–12. [doi:10.1145/2461912.2461964]

- 9 Zhao MT, Zhu SC. Portrait painting using active templates. Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Non-photorealistic Animation and Rendering. Vancouver: ACM, 2011. 117–124. [doi: [10.1145/2024676.2024696](https://doi.org/10.1145/2024676.2024696)]
- 10 Zhang Y, Dong WM, Ma CY, *et al.* Data-driven synthesis of cartoon faces using different styles. IEEE Transactions on Image Processing, 2017, 26(1): 464–478. [doi: [10.1109/TIP.2016.2628581](https://doi.org/10.1109/TIP.2016.2628581)]
- 11 Chen H, Liu ZQ, Rose C, *et al.* Example-based composite sketching of human portraits. Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering. Annecy: ACM, 2004. 95–153. [doi: [10.1145/987657.987673](https://doi.org/10.1145/987657.987673)]
- 12 Rosin PL, Lai YK. Non-photorealistic rendering of portraits. Proceedings of the 2015 Workshop on Computational Aesthetics. Istanbul: Eurographics Association, 2015. 159–170. [doi: [10.2312/EXP.20151189](https://doi.org/10.2312/EXP.20151189)]
- 13 Fišer J, Jamriška O, Simons D, *et al.* Example-based synthesis of stylized facial animations. ACM Transactions on Graphics, 2017, 36(4): 1–11. [doi: [10.1145/3072959.3073660](https://doi.org/10.1145/3072959.3073660)]
- 14 Selim A, Elgharib M, Doyle L. Painting style transfer for head portraits using convolutional neural networks. ACM Transactions on Graphics, 2016, 35(4): 129. [doi: [10.1145/2897824.2925968](https://doi.org/10.1145/2897824.2925968)]
- 15 Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2414–2423. [doi: [10.1109/CVPR.2016.265](https://doi.org/10.1109/CVPR.2016.265)]
- 16 Kaur P, Zhang H, Dana K. Photo-realistic facial texture transfer. Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2019. 2097–2105. [doi: [10.1109/WACV.2019.00227](https://doi.org/10.1109/WACV.2019.00227)]
- 17 Zhang YB, Li MH, Li RH, *et al.* Exact feature distribution matching for arbitrary style transfer and domain generalization. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 8025–8035. [doi: [10.1109/CVPR52688.2022.00787](https://doi.org/10.1109/CVPR52688.2022.00787)]
- 18 Yang S, Jiang LM, Liu ZW, *et al.* Pastiche master: Exemplar-based high-resolution portrait style transfer. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 7683–7692. [doi: [10.1109/CVPR52688.2022.00754](https://doi.org/10.1109/CVPR52688.2022.00754)]
- 19 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)]
- 20 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)]
- 21 Shi YC, Deb D, Jain AK. WarpGAN: Automatic caricature generation. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 10754–10763. [doi: [10.1109/CVPR.2019.01102](https://doi.org/10.1109/CVPR.2019.01102)]
- 22 Choi Y, Uh Y, Yoo J, *et al.* StarGANv2: Diverse image synthesis for multiple domains. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 8185–8194. [doi: [10.1109/CVPR42600.2020.00821](https://doi.org/10.1109/CVPR42600.2020.00821)]
- 23 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19. [doi: [10.1007/978-3-030-01234-2_1](https://doi.org/10.1007/978-3-030-01234-2_1)]
- 24 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600–612. [doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861)]
- 25 Luo L, Xue DY, Feng XL. Ehanet: An effective hierarchical aggregation network for face parsing. Applied Sciences, 2020, 10(9): 3135. [doi: [10.3390/app10093135](https://doi.org/10.3390/app10093135)]
- 26 Xie SN, Tu ZW. Holistically-nested edge detection. Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015. 1395–1403. [doi: [10.1109/ICCV.2015.164](https://doi.org/10.1109/ICCV.2015.164)]
- 27 Qin XB, Zhang ZC, Huang CY, *et al.* U²-Net: Going deeper with nested U-structure for salient object detection. Pattern Recognition, 2020, 106: 107404. [doi: [10.1016/j.patcog.2020.107404](https://doi.org/10.1016/j.patcog.2020.107404)]
- 28 Shrivastava A, Pfister T, Tuzel O, *et al.* Learning from simulated and unsupervised images through adversarial training. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2242–2251. [doi: [10.1109/CVPR.2017.241](https://doi.org/10.1109/CVPR.2017.241)]
- 29 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.

(校对责编: 牛欣悦)