

# 杭州电子科技大学

## 毕业设计（论文）开题报告

题    目     基于对抗式生成网络的角色形象生成

学    院     人文艺术与数字媒体学院

专    业     数字媒体技术

姓    名     徐嘉宁

班    级     18222513

学    号     18221787

指导教师     王毅刚

## 一、综述本课题国内外研究动态，说明选题的依据和意义

动漫是人们广为喜爱的一种艺术表现形式，一些经典的动漫人物形象可以唤起人们心中的美好回忆。在各种社交平台上人们常喜欢使用动漫人物作为账号的头像。然而大部分用户不一定具备动漫图像的自创能力，所以特定要求生成动漫风格的目标图像就成为了他们的现实需求。同时，动漫这种艺术形式在广告、游戏、影视作品、摄影等多个方面有着广泛的应用，但由于漫画的绘制和生成大多采用手工绘图，再经电脑渲染完成制作，并且一个新动漫角色的诞生需要在这基础上加入更多的步骤，花费的时间与人力都相对较多，因此使用电脑生成满足人为调控的漫画风格的形象成为了可期望的解决方案。在当前的技术条件下，利用生成式对抗网络解决这一问题成为了一个行得通的方案。

生成式对抗网络（GAN, Generative Adversarial Networks）是一种深度学习模型，是近年来复杂分布上无监督学习较具有前景的方法之一。GAN 常常应用于图像生成方面，例如超分辨率重建、图像去雾、语义分割以及风格迁移。

自 2014 年以来，生成式对抗网络的应用范围大幅增加。在广告领域，GAN 被用于创建虚拟时装模特照片，这使广告公司省下聘请模特、摄影师、化妆师的开支；在天文领域，GAN 被用于改善天文图像；在游戏影像领域，GAN 常常被用来改造早期游戏的画质，以 4K 或更高的分辨率重新创建低分辨率的 2D 纹理。

2014 年 10 月，Ian J. Goodfellow 为 GAN 发表了开山之作——《Generative Adversarial Networks》。在这篇论文之中，Goodfellow 等人首次提出了通过对抗过程估计生成模型的框架，并形象地解释了 GAN 的思想原理。他们在新框架中同时训练两个模型：一个是用于捕获数据分布的生成模型 G（Generative model），另一个是用来估计样本来自训练数据而不是 G 的概率的判别模型 D（Discriminative model）。这个框架相当于让两种神经网络模型进行持续不断的“博弈”。

这里以生成一张假钞图像为实例进行原理讲解，同时这也是原作者在论文中的形象比喻。

我们的最终目标是通过生成模型 G 得到一张足够优秀的假钞图片。

首先，生成模型 G 的任务是随机生成一张假钞图片。这张图片可能由一串随机的特征向量控制，例如我们可以人为设定特征向量的第一维控制图片中钞票的颜色，倒数第二维控制钞票的纹理。在输入随机的特征向量之后，G 按照特征向量随机生成一张图片。

其次，判别模型 D 的任务是判断 G 生成的假钞是否足够“真”。因此，D 的输入是 G 生成的图片，输出是一个标量值，通常它被限制在 0 到 1 的范围中，越小表示假钞越“假”；越大表示假钞越“真”。它的判断标准来自于我们提供的真钞图片的数据集，即对于 D 来说，这些数据集中的真钞的评价值都为 1，而 G 生成的假钞都为 0。

最后，G 为了能骗过 D 展开博弈。例如最开始 G 生成的假钞都是黑白的，D 发现真钞需要是有颜色的，输出 0 分；于是 G 再次生成假钞时加上了颜色，但纹理都很模糊，D 发现真钞的纹理都很清晰，输出 0 分；于是 G 再次生成了纹理清晰的假钞……直到 G 生成的假钞让 D 无法分辨真假，我们的任务便完成了。

GAN 在基于这个思想原理之上经过了 8 年的发展，至今演化出了若干种变体，例如 DC-GAN、Condition GAN、Cycle GAN、Info GAN、Big-GAN、Style GAN、Stack GAN 等等。

其中由 Radford 等人提出的 DC-GAN 标准生成对抗网络框架利用深度卷积层设计了生成器和判别器、少量使用全连接层，大大提升了 GAN 的生成性能，同时使用了批归一化和恰当的激活函数使得模型训练稳定并且可以生成高质量的图像，后续的 GANs 变体，包括各种生成三次元人物或者二次元人物的 GANs，基本上是采用了 DC-GAN 作为基础框架。

通常来说，不同的 GAN 变体着重于图像生成的不同方面。例如由 Tero Karras 等人提出的 Style GAN 通过逐层在生成器和判别器中添加卷积操作，在生成图像的过程中同时控制图像的低维和高维特征从而实现了图像的风格迁移；同样由 Tero Karras 等人提出的 PG-GAN，它通过先从低分辨率的生成器和判别器开始训练，逐渐向两种网络中添加层的方法增加了生成图片的空间分辨率，最终可以实现超高分辨率的人脸图像。

国内也有许多研究者研究 GAN 技术并进行改进。例如由来自清华大学的 Yang Chen 等人发明的 Cartoon GAN 和来自武汉大学的 Jie Chen 等人在前者基础上发明的 Anime GAN，都是建立在 Style GAN 之上将风格迁移技术引入 GAN 之中，使得生成图片能以漫画风格表现出来，而且不仅仅拘泥于人物形象。来自 Han Zhang 等人提出的 Stack GAN 可以做到通过文字描述符生成对应的高质量图像，其将 GAN 分为两个 stage：stage-I 根据给定的文本描述绘制对象的原始形状和基本颜色，并从随机噪声向量绘制背景布局，从而产生低分辨率图像。stage-II 纠正了 Stage-I 中低分辨率图像中的缺陷，并通过再次读取文本描述来完成对象的细节，从而生成高分辨率的照片般逼真的图像；Yanghua Jin 等人在论文中提到了一种基于 DARGAN (Dynamic Attention Recursive GAN) 的生成动漫角色头像的方法，在台湾大学李宏毅的生成对抗网络课程之中也包含了这部分内容，作者将论文中的内容在网页端实现，是我们可以方便体验到生成动漫头像的趣味。

从 GAN 的多态和广泛的应用中，不难看出，一个新的方法的引入很可能可以对生成图像产生截然不同的效果。我的项目也将以这篇论文为切入点，探究 GAN 技术的思想以及如何在代码中落实，最后提出改进点。

## 二、研究的基本内容，拟解决的主要问题：

项目研究的基本内容主要包括有：1.理解 GAN 中生成模型和判别模型两个神经网络内部的工作原理。2.熟悉如何使用需要用到的第三方框架，如 pytorch、tensorflow 和 keras。3.参考已有的、开源的关于 GAN 的项目代码，并在本地运行，分析图片生成的效果，总结其优缺点。4.结合其他 GANs 中的提升生成图像质量的思想，从而提升生成动漫形象的质量。5.扩展原有 GAN 算法功能，使得生成的图片不仅仅局限于头像，也包括肢体和服饰。

拟解决的主要问题主要包括：1. 在本地复现生成动漫人物头像。需要对神经网络、深度学习以及 tensorflow、pytorch 或者 keras 有深入的学习和了解。2. 在各种 GANs 中寻找适合的算法与现有算法相结合从而提升生成图片的质量。在 Yanghua Jin 等人的论文中，他们把利用 GAN 生成动漫人物头像的技术制作成了一个叫做“MakeGirlsMoe”的网站，人人都可以一键生成头像。但是在尝试过后，我发现网站上生成图片的质量普遍较低，比较模糊而且五官不够协调。在一些博客中作者们也提到了对图片生成的不满意，于是他们自己复现并重新训练模型以求达到更好的效果。我认为可以尝试加入其他 GANs 的思想来提升图片的质量，例如 PG-GAN，其通过从低到高分辨率的逐步训练过程而得到充满细节的人脸图像，对于生成动漫头像而言也许同样有效。3.扩展原有功能，使实现角色肢体与服饰的生成，例如生成人设图。在 Yanghua Jin 中，作者将大量人物的脸部截取保存，并整理成数据集提供给判别器。因此可以保留角色的肢体与服饰信息，更改数据集中的图片，也许能够重新训练完整的角色形象的图片。但过程中可能会遇到不可预测的问题，例如肢体的特征信息表示等，仍有待解决。大量的数据可以由爬虫在大型动漫图片网站上爬取得到。

## 三、研究步骤、方法及措施：

1.需要在本地克隆一些开源的项目进行研究，例如论文中已经实现并且开源在 github 上的项目，尤其是 DRAGAN 论文与 PG-GAN 论文中发表的项目。研究论文中的方法如何具体在代码中实现。

2.在本地复现基本的动漫头像的图片生成，对图片的质量进行多方面的分析，例如图片是否清晰、五官是否协调、颜色是否和谐、细节是否丰富……在过程中可以设定指标，使用 python 或者 matlab 等工具对生成的图片进行评价，并且绘制统计图便于展示。分析结果后，考虑并总结改进算法的方向。

3.参考网络中他人的改进思想和训练方法、参考其他 GANs 论文中的算法，尝试进行结合，例如 PG-GAN 和 Stack GAN。改进之后可以邀请身边同学对不同算法产生的结果进行质量打分，通过统计得分分析生成图像的质量，并确定算法的适配性。

4.对训练的数据集进行重组。使用爬虫爬取需要的图片，提取图片中的所需内容，对生成器和判别器重新训练，最后实现带肢体和服饰的角色形象图片。

5.重新评价改进算法对于肢体和服饰的生成效果以及对于图片整体的影响。同样可以采用邀请同学进行打分评价的形式。最后进行总结。

#### 四、研究工作进度：

序号	时间	内容
1	2022. 1. 1-2022. 1. 5	学生在教务系统中申请选题，指导老师确认
2	2022. 1. 5-2022. 3. 8	搜集资料，准备开题答辩
3	2022. 3. 9	开题报告会
4	2022. 3. 10-2022. 3. 24	研究所需基础知识查漏补缺及学习
5	2022. 3. 25-2022. 4. 6	开源项目研究、本地克隆运行
6	2022. 4. 7-2022. 4. 12	统计、分析产出结果
7	2022. 4. 13-2022. 5. 4	编写改进算法、扩展算法功能
8	2022. 5. 5-2022. 5. 10	测试改进后的结果、统计并分析
9	2022. 5. 11-2022. 5. 24	毕业论文初稿撰写、修改
10	2022. 5. 25-2022. 6. 3	论文查重、答辩、验收、评分、上报成绩
11	2022. 6. 4-学期末	资料整理及归档自查与交流

#### 五、主要参考文献：（所列出的参考文献不得少于 10 篇，其中外文文献不得少于 2 篇，发表在期刊上的学术论文不得少于 4 篇。）

- [1] Goodfellow I J , Pouget-Abadie J , Mirza M , et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.
- [2] Karras T , Aila T , Laine S , et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation[J]. 2017.
- [3] Karras T , Laine S , Aila T . A Style-Based Generator Architecture for Generative Adversarial Networks[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [4] Radford A , Metz L , Chintala S . Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. Computer ence, 2015.
- [5] Li A , Zheng C , Peng R , et al. Dynamic Attention Based Generative Adversarial Network with Phase Post-Processing for Speech Enhancement[J]. 2020.
- [6] Zhu J Y , Park T , Isola P , et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[J]. IEEE, 2017.
- [7] Chen Y , Lai Y K , Liu Y J . CartoonGAN: Generative Adversarial Networks for Photo Cartoonization[C]// IEEE/CVF Conference on Computer Vision & Pattern Recognition. IEEE, 2018.

- [8] Wu H , Zheng S , Zhang J , et al. GP-GAN: Towards Realistic High-Resolution Image Blending[J]. 2017.
- [9] Zhang H , Xu T , Li H , et al. StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks[J]. IEEE, 2017.
- [10]董虎胜, 刘诚志, 朱晶,等. 基于 StyleGAN 的动漫图像生成[J]. 甘肃科技纵横.
- [11] 朱嘉琪. 基于生成对抗网络的人脸图像合成 [D]. 合肥工业大学,2021.DOI:10.27101/d.cnki.ghfgu.2021.001094.
- [12] Jin Y , Zhang J , Li M , et al. Towards the Automatic Anime Characters Creation with Generative Adversarial Networks[J]. 2017.
- [13]孙天鹏,周宁宁,黄国方.新的基于 GAN 的局部写实感漫画图像风格迁移[J/OL].计算机工程与应用:1-14[2022-03-04].<http://kns.cnki.net/kcms/detail/11.2127.TP.20210706.1407.014.html>.

六、开题答辩小组评审意见：

考核点	背景及 意义阐述情况	研究内容 与任务书的 匹配程度	研究方案 合理性	进度安排 情况	答辩情 况	总分
满分	20	30	30	10	10	100
评分	18	25	24	9	9	85

开题答辩小组负责人签字：



2022年3月9日