[研究的背景及意义： 1](#_Toc6853)

[研究现状： 1](#_Toc27065)

[图像生成： 2](#_Toc1270)

[智能图像编辑： 2](#_Toc96)

[表情制作生成： 2](#_Toc13936)

[研究目标和内容： 2](#_Toc11330)

[研究技术路线以及可行性： 3](#_Toc4089)

[研究工作计划： 3](#_Toc11712)

[文献综述： 5](#_Toc28754)

# 课题学术与实际意义：

在深度学习等技术的推动下，图像生成和智能编辑技术越来越成熟。智能图像编辑就是用户通过简单的配置属性或者绘制草图，通过算法就能生成逼真的图像。就像语音识别可以理解人的声音信息一样，智能图像编辑某种程度上是对人想要表达的图像信息的理解，可以让我们不用掌握绘画技巧，就可以轻易地将想要的图像表达出来。

目前，智能图像编辑在很多领域取得了一定的成效，比如人脸，风景。通过研究这样的算法，可以在多方面取得应用，在社交娱乐领域，人们可以利用这类智能编辑方法，使用简单方式，让自己的照片变得具有艺术感，也可以直接创作令人印象深刻的图像艺术作品，和传统基于绘画的方式不一样的。在美妆美化领域，可以帮助用户尝试造型，妆容等；在设计领域可以帮助快速绘制出人或物的原型等。

# 研究现状：

图像生成：

近年来，伴随着深度学习，卷积神经网络，生成对抗网络（GAN）等技术的发展，图像生成领域有了长足的发展。现在的图像生成模型能够生成十分逼真的图像，并且在生成图像的同时让图像满足一定的条件。比如ProgressGAN[]，BigGAN[]，StyleGAN[]等能生成1024\*1024的高分辨率的逼真的图像。

图像风格迁移：

图像风格迁移就是将图像从一个域迁移到另一个域，比如将真实图像迁移变成某种风格的艺术照，CycleGAN[]可以实现利用编码器结合GAN组成循环结构，将一个领域的图像风格迁移到另一个领域，并且不依赖成对图像。比如从真实风景照片迁移到绘画图像。Adain通过添加自适应实例归一化层，实现了任意风格领域之间的图像迁移。

智能图像编辑：

随随着深度学习和GAN等技术的发展，尤其在StyleGAN提出之后，使得人们可以通过调节图像属性智能的生成十分逼真的图像，比如调节人脸图像的年龄等。Image2StyleGAN探究了的指定图像嵌入到StyleGAN前空间的方法，使得GAN图像生成的可解释性大大提高，促进了相关应用的发展。比如 https://www.gwern.net/Faces实现动漫人脸的编辑。

# 研究目标：

 实现一个基于GAN技术的智能图像编辑系统，能够比现有算法更加轻量和快速。使人们能在手机等终端通过简单的编辑操作，创作出令人印象深刻图像作品。

# 研究内容：

具体研究内容有：

1、收集不同方面的图像数据集以供训练，比如动画，风景等，并按算法需要对数据集做一定的预处理。

1. 这也是本文主要研究内容，包含使用风格迁移等算法将迁移图像风格，实现用户使用属性控制等方法编辑图像的算法模型，并且保证模型的轻量快速。训练模型，根据结果一步步改进优化算法。

3、编写用户交互应用，并部署算法模型到应用上，实现用户制作艺术图像的目的。优化模型的体积，大小等，将模型导出到移动设备上使用。编写用户交互界面，让用户通过简单方式实现具有艺术感的图像的创作编辑。

技术路线：

对于目标一，目前网络上可以获取大量的相关领域的图像，可以通过网页爬虫，手机端下载等方式获取到。另外，还有从相关的专们数据集下载。然后根据算法需求对图像做一定的处理，比如包含人脸的图像中裁剪人脸，缩放图像到需要的尺寸，图像标签对齐等。

对于目标二，拟采用基于GAN的方法设计模型，借鉴StyleGAN的思想。模型主体由编码器+解码器构成的生成器generator，和判别器discriminator构成。生成器，判别器拟采用CNN搭建而成，或者更新的transformer结构。借鉴StyleGAN的结构，将图像映射到潜空间，从而可以进行属性编辑。

对于目标三，将训练好的模型导出，然后开发移动端的APP，编写交互界面，用户可以通过简单的图像合成，属性控制等方式操纵图像。需要注意的是，基于实时交互的需求和手机等设备运算能力限制，模型需要满足轻量，快速的需求。 拟采用尽量降低生成图像尺寸，模型压缩等技术降低模型大小，并提升速度。另外还可以利用移动端专用的框架，比如MNN，运行模型，进一步提升速度。

# 可行性论证：

目前，生成对抗网络（GAN）等技术的发展，图像生成领域有了长足的发展。现在的图像生成模型能够生成十分逼真的图像，并且在生成图像的同时让图像满足一定的条件。StyleGAN[]等能生成1024\*1024的高清人脸的图像。基于StyleGAN的进一步研究能够在潜空间通过控制潜变量控制图像的生成。最终体现在用户端就是通过控制图片属性控制图像的生成。本文可以基于上述技术实现用户创作编辑一些特定领域图像，并且编辑结果十分逼真的的效果。

目前，对于深度学习的模型，研究者们提出了很多的方法进行模型压缩，比如剪枝，知识蒸馏，神经架构搜索等方法。通过这些方法，能够很大程度降低模型计算复杂度和模型大小等。Muyang Li等人提出的GAN Compression: Efficient Architectures for Interactive Conditional GANs方法，作为一种通用的GAN模型压缩方法，可以将GAN的模型复杂度降低数倍。本文可采用上述技术降低模型复杂度，达到实时交互的效果。

当前很多深度学习开源框架提供了移动端的支持。本文利用相应框架结合导出移动端模型，移动端的编程实现交互的编辑前端界面。

# 创新之处：

将基于GAN的智能图像编辑应用到更多领域，目前的研究多集中于人脸领域，使用模型压缩等算法让模型快速轻量，能够交互式运行。

# 拟解决的关键问题：

本课题拟解决的关键问题在于实现智能编辑的算法模型，以及对模型进行压缩让模型能够实时交互式的运行。

研究工作计划：

确定研究课题

研究文献，准备资料，确定研究思路

准备数据集：收集数据，数据清洗，数据标注

实现代码

训练模型

根据输出结果调整思路，代码，再进行训练调整

论文工作的总体时间安排：   
1、2020年11月-2021年3月，开题准备工作，查找资料，了解国内外的发展动态，对课题进行可行性分析，确定课题的最终研究方向及课题内容，学习相关知识技术。   
2、2021-04-01到2021-07-31 进行深入分析，认真学习与本研究课题相关的理论知识，形成进行相关系统的概要设计  
3、2021-08-01到2021-12-31 形成课题完整解决方案，实施方案，收集清洗数据，编写算法代码，进行模型训练等-  
4、2022-01-01到2022-02-28 完成对相关算法的改进，完成交互界面编写，并进行试验，验证有效性。  
5、2022-03-01到2022-05-31，总结和整理相关资料，进行论文的撰写，准备答辩。

文献：

PSGAN: Pose and Expression Robust Spatial-Aware GAN for Customizable Makeup Transfer

文献综述：

研究背景

研究现状

评述

参考文献

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 , 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.

[3] Huang X , Belongie S . Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization[J]. IEEE, 2017.

[4] Li B , Zhu Y , Wang Y , et al. AniGAN: Style-Guided Generative Adversarial Networks for Unsupervised Anime Face Generation[J]. 2021.

[5] Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space?[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2019.

[6] Zhu J Y , Park T , Isola P , et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[J]. IEEE, 2017.

[7] Abdal R , Qin Y , Wonka P . Image2StyleGAN++: How to Edit the Embedded Images?[J]. 2019.

[8] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets[J], 2014. 1, 2

[9] Li M , Lin J , Ding Y , et al. GAN Compression: Efficient Architectures for Interactive Conditional GANs[J]. 2020.

[10] Aguinaldo A , Chiang P Y , Gain A , et al. Compressing GANs using Knowledge Distillation[J]. 2019.

[11] Donahue J , Simonyan K . Large Scale Adversarial Representation Learning. IEEE, 2019.

**[12] Viazovetskyi Y , Ivashkin V , Kashin E . StyleGAN2 Distillation for Feed-forward Image Manipulation. 2020.**

## 论证记录：

问题一：本课题研究内容是什么

答：现在基于GAN的技术可以生成令人印象深刻的艺术图像，本课题主要研究内容是实现基于GAN技术智能生成和编辑艺术图像的算法，这有别于传统的图像创作方法，比如直接调节人脸照片的年龄生成一个人不同年龄的照片。本课题还研究使算法模型轻量化的技巧，实现交互式编辑的系统。

问题二：课题主要创新点在于

答：本文主要创新点有，1、将智能编辑创作算法应用到不同的图像领域 2、算法模型轻量化技术，实现模型的交互式运行

问题三：是否有足够文献支撑，你都查阅了哪些文献

答：本文在研究过程中查阅了各种相关资料，报考阅读相关领域最新的论文，查阅各种网络资料，体验和使用了现有的相关应用。

问题四：简述本课题的实现方式，技术路径

答：基于生成对抗网络GAN实现逼真的图像生成，借助StyleGAN的结构，将图像嵌入到潜空间，潜空间可以看做图像属性的一个解耦的表示，此时可以通过调整潜空间的变量实现对图像的编辑。然后利用现有的模型蒸馏，移动模型等技术可以实现对算法模型的轻量化。

GAN

STYLEGAN

Adain

Anigan

Image2stylegan

Cyclegan

Image2stylegan++

CGAN

GAN压缩1 压缩2

疑问：

很偏应用 并且应用领域也是一个很小的领域 这个怎么

还是那个创新性的评估，有些疑惑。

论文看多了，感觉自己的思路似乎都有人实现过了，创新性