**2022年矿大-华为智能基座竞赛**

**作品报告**

**作品名称： 用DCAGN生成柯南图片**

**姓名： 郑耀凯 专业班级： 人工智能21-2班**

**学号： 08213166 电子邮箱： yukuiyue@gmail.com**

**作品类别： AI作品赛 联系电话： 19516132795**

填写说明

1. 所有参赛项目必须为一个基本完整的设计与实现。作品报告书旨在能够清晰准确地阐述（或图示）参赛项目（或方案）。

2. 作品报告采用A4纸撰写。除标题外，所有内容必需为宋体、小四号字、1.5倍行距。

3. 作品报告中各项目说明文字部分仅供参考，作品报告书撰写完毕后，请删除所有说明文字。(本页不删除)

4. 作品报告模板里已经列的内容仅供参考，作者可以在此基础上增加内容或对文档结构进行微调。

5. 作品类别，请填写openEuler作品赛、AI作品赛或者鸿蒙作品赛。

**目 录**

[摘要 ……………………………………………………………………………………1](#_Toc33965493)

[第一章 作品概述 …………………………………………………………2](#_Toc33965494)

1.1 背景及相关工作……………………………………………………………………2

1.2 运行环境……………………………………………………………………………2

1.3 作品特色……………………………………………………………………………2

1.4 应用前景……………………………………………………………………………2

[第二章 作品设计与实现 …………………………………………………………3](#_Toc33965495)

2.1 系统方案……………………………………………………………………………3

2.2 实现原理……………………………………………………………………………3

2.3 硬件框图与软件流程………………………………………………………………4

[第三章 作品测试与分析 …………………………………………………………6](#_Toc33965496)

3.1 测试方案、环境、设备……………………………………………………………6

3.2 测试数据及结果分析………………………………………………………………6

[第四章 创新性说明…………………………………………………………………9](#_Toc33965497)

4.1数据集………………………………………………………………………………9

4.2 DCGAN ………………………………………………………………………………9

[第五章 总结 1](#_Toc33965498)0

[参考文献 1](#_Toc33965499)1

# 摘要

作为学习人工智能的一名同学，想参加2022年中国矿业大学华为智能基座软件与人工智能开发竞赛来加深对人工智能专业的理解，同时为了更熟练地使用Mindspore开发AI应用，完成了本次作品。

本作品名称为用DCAGN生成柯南图片（柯南是动漫《名侦探柯南》中的主角江户川柯南），主要功能为应用DCGAN通过输入预先准备好的训练集，来生成模拟柯南的图片。

本作品的数据集经过人为筛选后，主要画面均为柯南，所以得到的假数据集也会模拟柯南，数据较集中，便于帮助判断生成器的优劣。

GAN是一种深度学习模型，是近年来复杂分布上无监督学习最具前景的方法之一。

DCGAN作为GAN的一种，可以根据输入的数据来生成对输入数据模拟后的数据。所以在实用方面，可以自定义数据集输入，来得到假数据集。

GAN的图像生成可以用于许多领域，在原数据量不足时，可通过GAN来补充数据达到扩充数据量的效果。

关键词：智能基座 GAN DCGAN 柯南图片

# 第一章 作品概述

1.1 背景及相关工作

21世纪以来，在大数据、云计算和物联网等新兴技术不断成熟的大背景下，人工智能得到了第3次井喷式发展。其中，以深度学习为基础的图像处理、语音识别等技术的发展尤为迅猛。而生成式对抗网络（GAN）是如今深度学习领域中最为前沿也是最令人着迷的领域之一，它自2014年10月被Ian Goodfellow等提出以后，就一直受到人们的广泛关注与追捧。

DCGAN，全称是 Deep Convolution Generative Adversarial Networks（深度卷积生成对抗网络），是Alec Radfor等人于2015年提出的一种模型。该模型在 Original GAN的理论基础上，开创性地将CNN和GAN相结合以实现对图像的处理，并提出了一系列对网络结构的限制以提高网络的稳定性。

1.2 运行坏境

在华为ModelArts的notebook中，选择mindspore1.7.0-cuda10.1-py3.7-ubuntu18.04。

进入notebook后首先需要安装GPU CUDA10.1对应的mindspore1.8.1版本（参照[MindSpore安装指南](https://www.mindspore.cn/install)）。本作品程序在GPU上训练运行。

1.3 作品特色

选用了GAN中的DCGAN；输入输出都为我们较为熟悉的动漫角色柯南，便于判断网络的情况；代码中还使用了mindspore中的数据增强、激活函数等操作。

1.4 应用前景

在医学领域，有时会遇到数据量不足的情况，这时通过GAN网络生成与原图像相似的假图像来作为补充数据，可以扩充数据量。

例如在异常检测方面，当正常数据很多，而异常数据不多，可以通过生成器来学习到正常数据的分布，然后将需要检测的异常图通过前面的生成器得到其正常图的样子，然后对比找出异常。

# 第二章 作品设计与实现

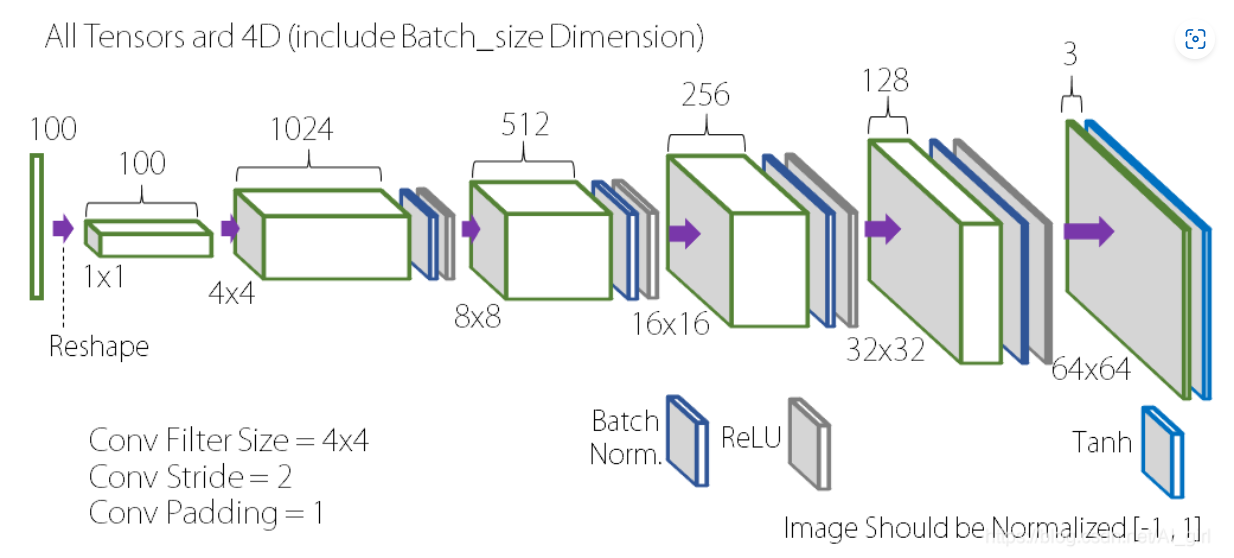
2.1 系统方案

1. 准备并上传准备好数据集，加载数据集、进行数据增强。
2. 定义转置卷积层、卷积层、生成器、损失函数等。
3. 训练并打印训练结果。

2.2 实现原理

DCGAN即使用卷积网络的对抗网络，其原理和GAN一样，只是把CNN卷积技术用于GAN模式的网络里，G（生成器）网在生成数据时，使用反卷积的重构技术来重构原始图片。D（判别器）网用卷积技术来识别图片特征，进而做出判别。同时，CDGAN中的卷积神经网络也做了一些结构的改变，以提高样本的质量和收敛速度。

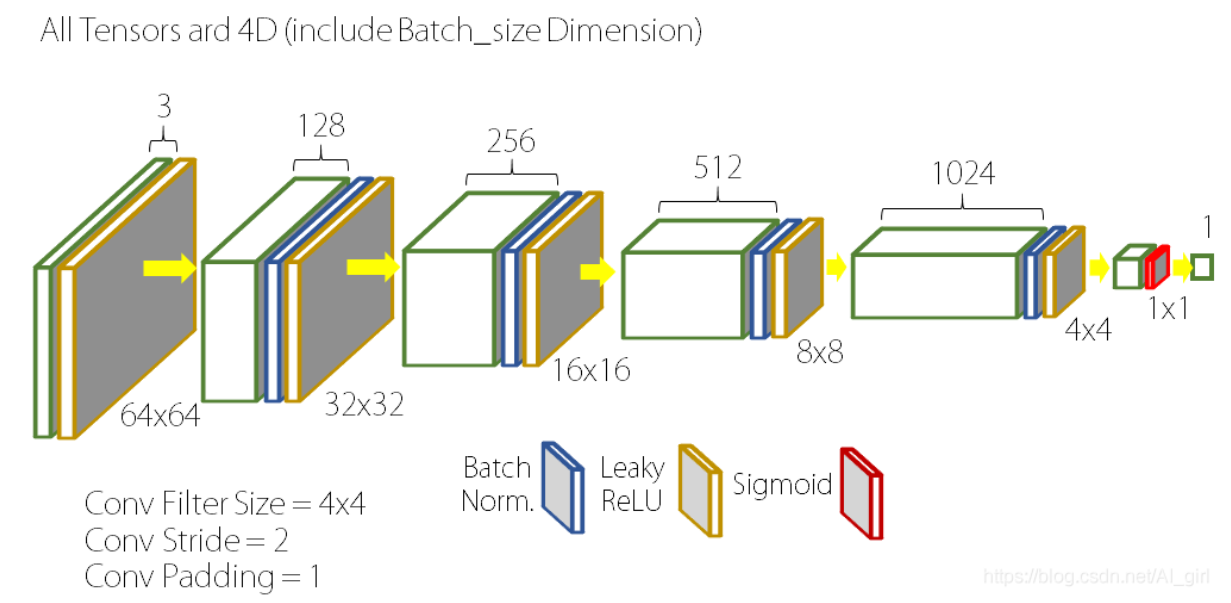
DCGAN的generator网络结构图如下：



1.G网中使用ReLU作为激活函数，最后一层使用Tanh作为激活函数。

2.去掉了FC层，使网络变为全卷积网络。

DCGAN的discriminator网络结构图如下：



1.D中取消所有的池化层，使用转置卷积（transposed convolutional layer）并且步长大于等于2进行上采样。

2.D网中也加入stride的卷积代替pooling。

3.在D网和G网中均使用批量归一化（batch normalization），而在最后一层时通常不会使用batch normalization，这是为了保证模型能够学习到数据的正确均值和方差。

4.D网络中使用LeakyReLU作为激活函数。

5.DCGAN中换成了两个卷积神经网络（CNN）的G和D，可以刚好的学习对输入图像层次化的表示，尤其在生成器部分会有更好的模拟效果。DCGAN在训练过程中会使用Adam优化算法。

2.3 硬件框图与软件流程

2.3.1硬件

配置为：GPU: 1\*V100(32GB)|CPU: 8核 64GB

2.3.2 软件流程

1. 先在华为ModelArts开发平台上上传准备好的数据集。
2. 加载数据集并进行数据增强操作，包括放缩、随机裁剪等。再进行数据映射，展现部分经过前述处理后的数据集。
3. 定义转置卷积层、BatchNorm2d层。
4. 定义网络生成器、卷积层、网络判别器。
5. 定义损失函数、连接生成器和损失函数、连接判别器和损失函数。
6. 创建隐向量、设置优化器。
7. 定义DCGAN网络、实例化。
8. 开始训练。
9. 打印训练结果，展示生成器生成的图片。
10. 检查训练结果，若结果不佳则考虑重新设定参数再次训练。

# 第三章 作品测试与分析

3.1 测试方案、环境、设备

测试方案主要为：（1）打印生成器和判别器的损失与训练迭代的关系图。

（2）重新生成隐向量，通过生成器输出图片并观察。

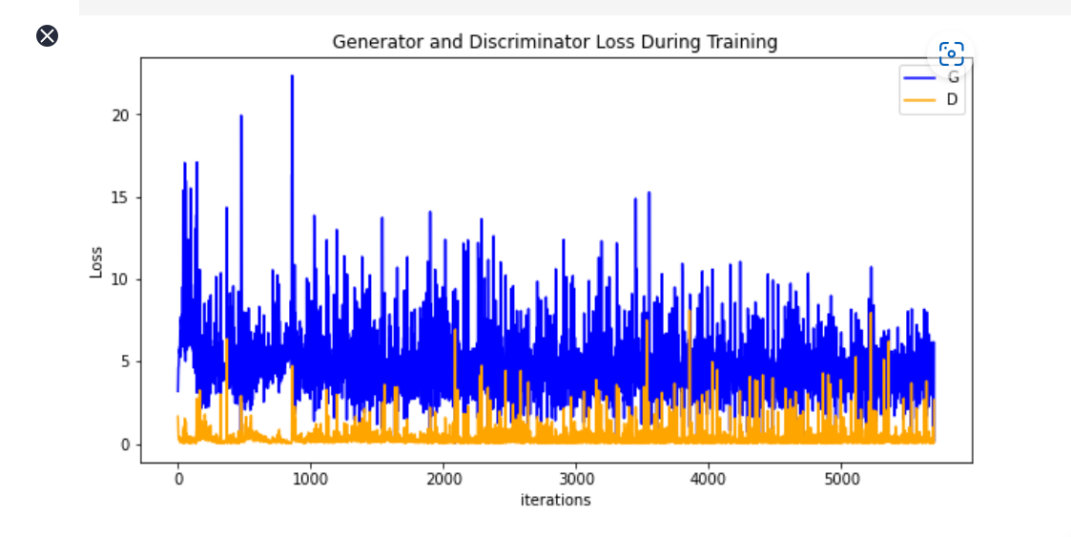
测试环境为ModelArts的notebook，选择其中的mindspore1.7.0-cuda10.1-py3.7-ubuntu18.04

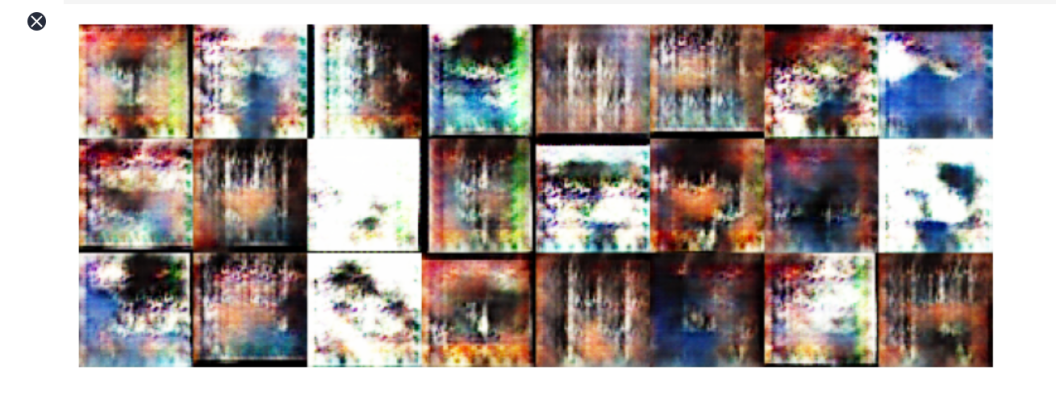
测试设备为联想r7000p

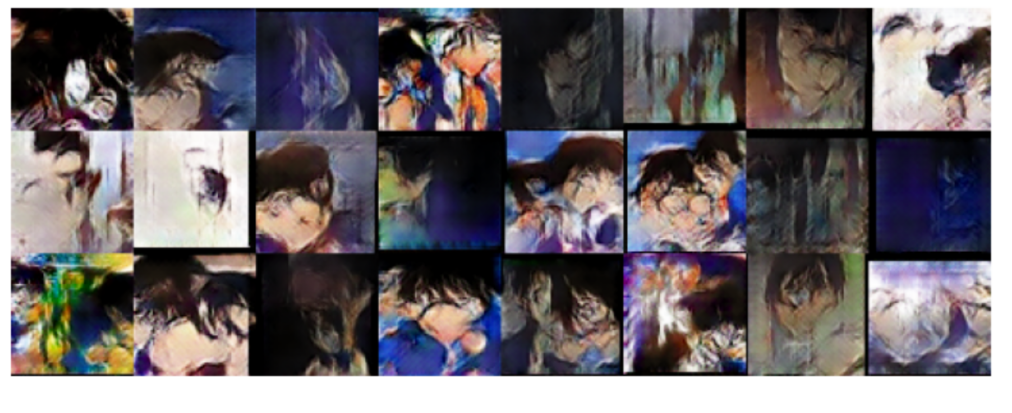
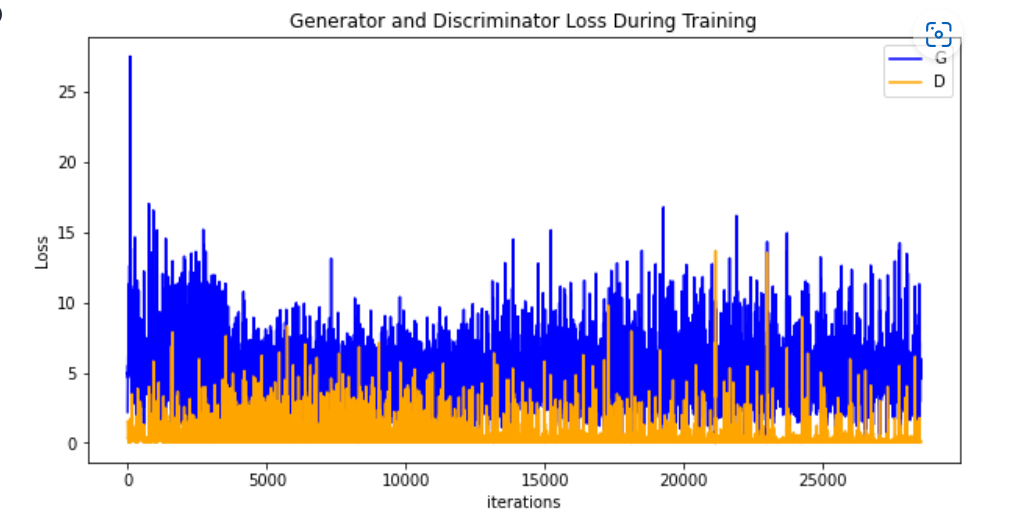
3.2 测试数据及结果分析

3.2.1测试数据

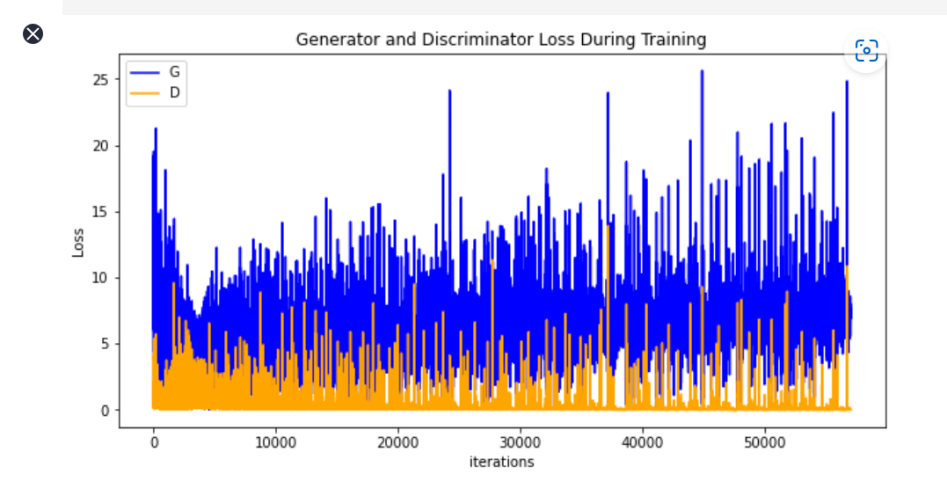
（1）num\_epochs = 10



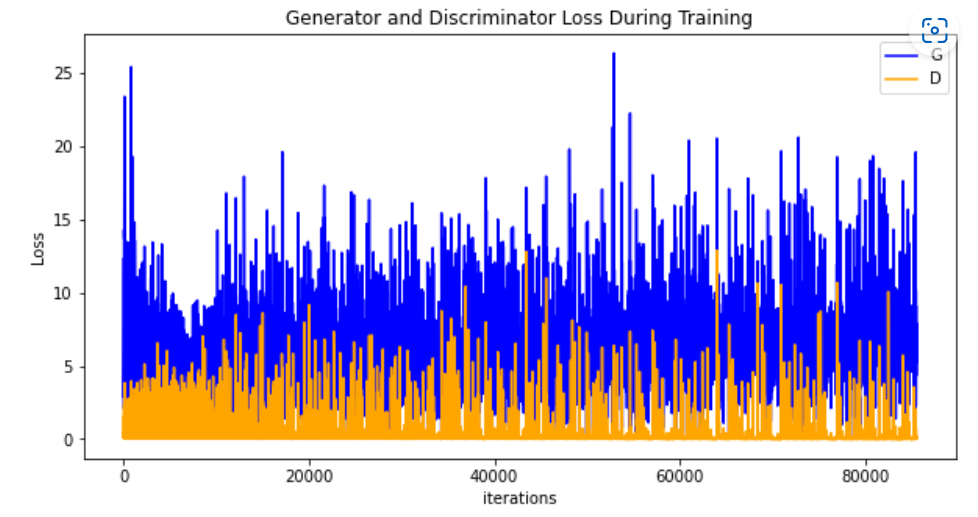


1. num\_epochs = 50 

（3）num\_epochs = 100



（4）num\_epochs = 150

0

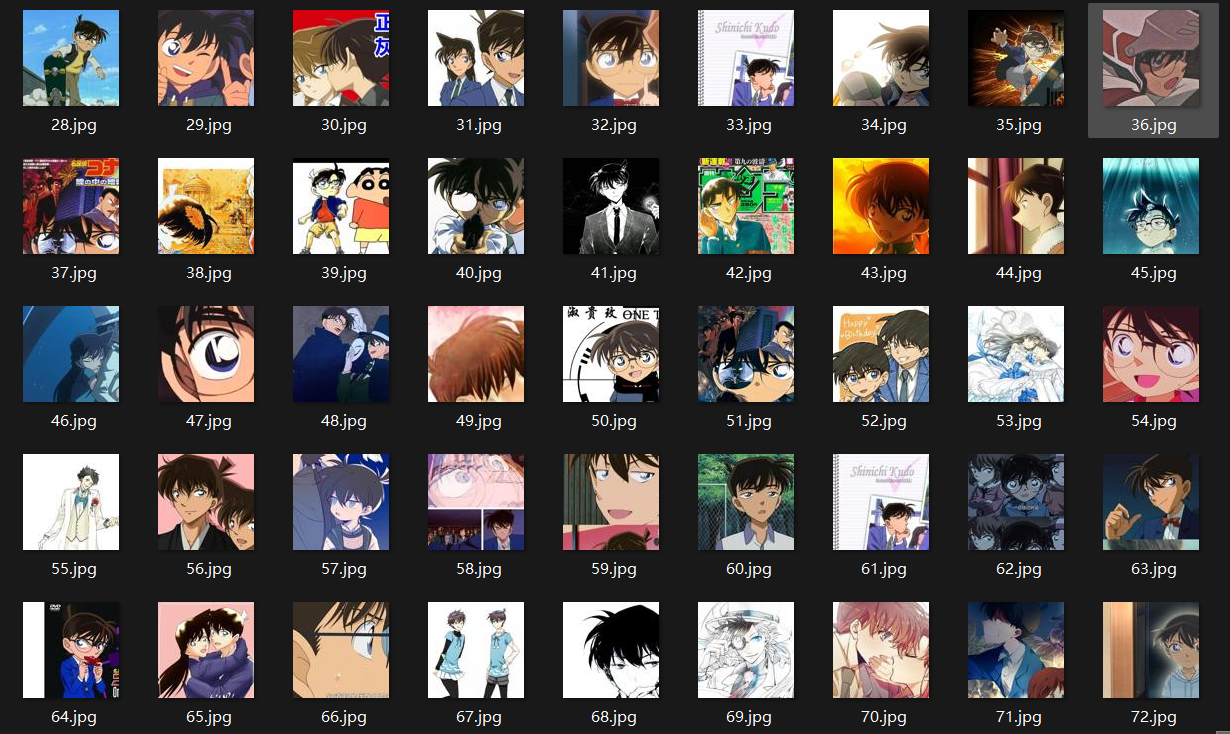
3.2.2 结果分析

随着num\_epochs的增大，生成器生成的图像效果越好。

# 第四章 创新性说明

4.1数据集

本作品的数据集一共为907张柯南图像，均经过仔细筛选。



4.2 DCGAN

（1）DCGAN所需要的各项参数中，学习率lr = 0.0001较为合适。

（2）在数据增强时加入了随机裁剪和随机翻转操作。另外，之前几次训练并因为数据集小导致效果不理想，后来使用repeat使每个data的数据量由原先的19变为570。

（3）DCGAN网络判别器中，使用了LeakyReLU激活函数，比ReLU函数效果好。

（4）使用Adam优化器，具有实现简单，计算高效，对内存需求少， 参数的更新不受梯度的伸缩变换影响，更新的步长能够被限制在大致的范围内（初始学习率）， 适用于不稳定目标函数，学习率自适应修正，不用手动调整的优势。

# 总结

本次的作品经过几个星期的学习、尝试和练习后得以完成。在此期间，对代码进行了多次修改，尝试使用华为的ModelArts来完成作品。曾发生过训练周期num\_epochs=200时出现梯度爆炸或梯度消失、学习率不合适导致梯度下降过慢等问题。

优化器尝试过Momentum和Adam，发现Adam具有学习率自适应修正，不用手动调整的优势，最终决定使用Adam优化器。

这次的作品从学习mindspore到编写代码，花费了不少精力，编写代码时不懂的地方很多，于是不断查找学习，对变量的命名也更加规范化，同时也加深了对代码的理解。

# 参考文献

[1]https://blog.csdn.net/qq\_24224067/article/details/103371910

[2]https://blog.csdn.net/AI\_girl/article/details/88424768

[3]https://www.mindspore.cn/docs/zh-CN/r1.8/index.html