****毕业论文（设计）

|  |  |
| --- | --- |
| **设计(论文)题目** | 基于DCGAN的语义图像修复研究与实现 |

**2019年 5月 10日**

目 录

[摘 要 I](#_Toc9156226)

[ABSTRACT II](#_Toc9156227)

[一、 绪 论 1](#_Toc9156228)

[(一) 选题背景及研究意义 1](#_Toc9156229)

[(二) 国内外研究现状 1](#_Toc9156230)

[1. 以下是国内外现在主流的图像修复的技术的方法： 1](#_Toc9156231)

[2. 关于基于深度学习的图像修复的研究现状 1](#_Toc9156232)

[(三) 论文主要研究内容 2](#_Toc9156233)

[(四) 论文组织结构 2](#_Toc9156234)

[二、 深度学习在图像修复上的应用 3](#_Toc9156235)

[(一) 将图像解释为概率分布 3](#_Toc9156236)

[(二) 卷积神经网络与图像修复 3](#_Toc9156237)

[1. 卷积神经网络简介 3](#_Toc9156238)

[2. 卷积网络的操作 4](#_Toc9156239)

[3. 卷积网络在图像修复上的应用 5](#_Toc9156240)

[(三) 生成对抗网络与图像修复 6](#_Toc9156241)

[1. 生成对抗网络简介 6](#_Toc9156242)

[2. 生成对抗网络在图像修复上的应用 6](#_Toc9156243)

[(四) 深度卷积对抗网络与图像修复 7](#_Toc9156244)

[1. 深度卷积对抗网络简介 7](#_Toc9156245)

[2. 深度卷积对抗网络与图像修复 7](#_Toc9156246)

[(五) 三种网络在图像修复上的比较 8](#_Toc9156247)

[三、 图像生成模型的构建 9](#_Toc9156248)

[(一) 环境配置 9](#_Toc9156249)

[1. 本地环境配置 9](#_Toc9156250)

[2. 网络环境配置方法1(无GPU电脑推荐) 9](#_Toc9156251)

[3. 网络环境配置方法2(无GPU电脑推荐) 10](#_Toc9156252)

[(二) 深度卷积对抗网络的激活函数 10](#_Toc9156253)

[(三) 深度卷积对抗网络的代价函数 11](#_Toc9156254)

[(四) 深度卷积对抗网络的图像生成训练步骤 11](#_Toc9156255)

[(五) 基于深度卷积对抗网络的人脸数据集图像生成模型 12](#_Toc9156256)

[1. 人脸图片数据集简介 12](#_Toc9156257)

[2. 基于人脸图片数据集的图像生成模型的构建 12](#_Toc9156258)

[3. 图像生成模型的运行以及分析改进 16](#_Toc9156259)

[4. 深度卷积对抗网络在图像生成模型上的优势 18](#_Toc9156260)

[四、 图像修复 19](#_Toc9156261)

[(一) 图像修复原理 19](#_Toc9156262)

[(二) 建立投影方程 19](#_Toc9156263)

[(三) 图像修复与图像生成 20](#_Toc9156264)

[1. 图像修复与图像生成的关系 20](#_Toc9156265)

[2. 图像修复与图像生成的区别 21](#_Toc9156266)

[(四) 图像修复模型的构建与运行 22](#_Toc9156267)

[1. 图像修复模型构建 22](#_Toc9156268)

[2. 图像修复参数的确定 22](#_Toc9156269)

[3. 图像修复模型的运行 23](#_Toc9156270)

[五、 总结与展望 24](#_Toc9156271)

[(一) 本论文所做工作总结 24](#_Toc9156272)

[(二) 进一步研究设想与展望 24](#_Toc9156273)

[参考文献 25](#_Toc9156274)

[附 录 27](#_Toc9156275)

[致 谢 29](#_Toc9156276)

摘 要

随着深度学习的兴起以及图片大数据集的生成，人们开始尝试应用深度学习的方法对图像进行研究与应用。在这些应用之中，图像修复成为应用热点。传统的图像修复一般采用PS工具对图像进行感知修复，这种方法虽然一定程度上修复率高，但存在修复过程复杂,修复时间较长的问题。应用深度学习的半监督训练的图像修复现已在图像自动化方法比赛取得优胜，逐渐成为自动化图像修复的主流方法。

本文主要研究如何利用深度卷积对抗生成网络进行基于图像语义的图像修复。首先输入完整图片不断训练得到一个切实可行的假图片生成网络模型，其次输入随机分布，应用感知损失以及缺失部分的边缘信息，不断优化输入分布达到逼近缺失图片分布的目的，最后把此分布送入生成器生成最佳假图片来补全缺失部分，以达到图像修复的目的。

关 键 词

深度学习；图像修复；深度卷积对抗生成网络；图片生成；边缘信息

ABSTRACT

With the rise of deep learning and the generation of large data sets, people began to study and apply images using deep learning methods. Among these applications, image inpainting becomes a hotspot for image applications. Traditional image inpainting generally uses PS tools to perceptually repair images. Although this method has a high repair rate to some extent, it has the problems of low efficiency and a long time. The image inpainting of semi-supervised training using deep learning has now won the competition in the image automation method competition, and has gradually become the mainstream method of automated image inpainting. Among them, the deep convolutional anti-generation network has become the mainstream network for image inpainting in deep learning applications.

This paper mainly studies image semantic inpainting using deep convolution to generate networks. First, input a large number of complete and unmissed images, and use the accelerator to continuously train to obtain a feasible false image automatic generation network model. When the fake picture trained by the model is realistic enough, the model is formally applied to repair the image before application. The image of the missing fixed part is enough, and the picture of the missing part is used as the input of the trained deep network convolutional neural network model, which uses the edge information of the missing part to the edge information and the perceptual information. Completing the missing part, and finally adjusting the parameters, optimizer, network structure parameters and other hyperparameters to achieve the purpose of image inpainting network optimization.

KEY WORDS

Deep-learning; Image-inpainting; Generated-fake-picture; DCGAN; Edge-information

绪 论

* 1. 选题背景及研究意义

传统的基于物理与数学的研究方法曾在视觉与图像领域占据一席之地。然而随着近几年大家共同见证了深度学习在视觉与图像应用方面取得的显著成果，再加上其应用效果远远超越传统的物理与数学方法，所以选择深度学习成为视觉领域研究的首选逐渐成为趋势。幸运的是深度学习在图像分类的领域上已取得卓越的效果，这之后一些研究者就把研究方向放在更上一层的应用上，比如图像编辑和图像生成等应用。在图像编辑与图像生成之间的一个应用图像修复同样也吸引了当下不止步于图像分类的研究者研究与探索。

图像修复是指利用复杂的算法重建图形中丢失或损坏的部分的过程。在现实生活中，这项工作仍然由经验丰富的图像修复师来完成。图像修复技术主要用来修复日常生活中被噪声污染或者人为破坏的破损图像，其次也可应用于替换图像中的小区域或者瑕疵。所以让图像修复借助深度学习的算法和框架自动化是一个值得深入研究的课题。

* 1. 国内外研究现状
     1. 以下是国内外现在主流的图像修复的技术的方法：

(1) 基于偏微分方法的图像修复[1],[2]：主要在建立了图像的先验和数据模型的前提上，采用数学上的偏分和变分的方法，利用待补全区域的周边局部信息来确定扩散信息和方向，此方法只适用于缺失部分面积较小，少纹理的图片破损修复。

(2) 基于样本采样的图形修复[1],[3]：采样一些与缺失部分较相似的图像块来填充缺失的图形图像块，此种修复追求整体一致性，但会造成图像拼接效果不佳，产生与原图差别较大的结果。

(3) 基于深度学习的图像修复[1],[4]：基于深度学习的图像修复是近来提出的一种新兴的方法。主要原理是利用深度学习构建学习网络学习图像的特征，得到训练样本间的非线性关系,然后利用学到的特征以及非线性关系输入缺失图片进行图像补全。

* + 1. 关于基于深度学习的图像修复的研究现状

在对深度学习与图像应用相结合的不断探索中，由Lecun Y[5]提出的卷积神经网络(CNN)因为其能利用卷积图像特征形成局部特征的优点而成为深度学习图像应用的一个重大突破，同时在在纹理研究中，发现一个用CNN训练的网络所提取的特征作为隐藏层输入之后生成的图片与原图片具有相似的语义。这之后由Goodfellow[6]提出的生成对抗网络(GAN)被应用于图像生成，被发现对于图像的生成具有视觉增强效果（提高分辨率）。这些网络的不断创新，更加坚定人们把深度学习应用于图像修复中的信心。

* 1. 论文主要研究内容

本论文以深度卷积对抗网络(DCGAN[7])在图像修复上的应用为研究的课题，展开实验与改进，主要研究内容包括探讨DCGAN在图像修复与CNN，GAN网络比较的优势，然后用DCGAN网络进行图像修复，重点阐述图像修复的原理，图像修复实验的步骤，效果评价，改进方法。

* 1. 论文组织结构

(1) 第一章（绪论）主要介绍所研究的课题的研究背景，并以此背景介绍了本课题研究的目的意义，证明了本课题值得深入研究。

(2) 第二章（深度学习在图像修复上的应用）主要围绕三大网络展开，首先是卷积神经网络，展开介绍了卷积神经网络的原理，关于构建卷积网络所用的卷积操作，转置卷积操作，池化操作以及利用卷积网络形成的自编码器进行图像修复，接下来介绍了GAN网络的原理，GAN网络的生成器与分类器的简介，GAN网络的训练过程和应用GAN来进行图像修复的原理。最后介绍了DCGAN网络，写明此网络与前述网络的区别以及本课题采用此网络进行网络修复的原因。

(3) 第三章（图像生成模型的构建），在该章节重点阐述了进行模型初构造的原理与过程，首先利用Celeba数据集加上构建好的DCGAN模型进行Celeba图像生成，根据生成效果验证DCGAN在图像生成的优势。

(4) 第四章（图像修复），在该章节将采用celeba数据集进行图像修复，复用上文构建好的Celeba图像生成模型，并基于该图像生成模型构造图像修复模型并对模型的原理进行阐明以及模型运行效果的评价。

(5) 第五章(总结)对研究的课题进行由上至下的总结，并提出本课题今后的修改方向。

深度学习在图像修复上的应用

* 1. 将图像解释为概率分布

在统计学上，常常将图片解释为概率分布的样本，所谓的概率分布其实就是对图像的像素值进行采样得到的分布。对于图像修复来说，我们可以利用待修复的原图的概率分布去估算破损块的像素值。如图2.1所示，同一张图片具有相同的概率分布，当输入固定的像素值（图示x=1），然后利用分布来找到所有可能的输出像素值(y),找到其中概率最大即是最可能的图片。

|  |
| --- |
|  |
| **图2.1 图像的概率分布图(图像二维分布固定x,找到最密集的y即为输出的图像分布)** |

* 1. 卷积神经网络与图像修复
     1. 卷积神经网络简介

卷积神经网络现已经在深度学习关于图像分类上的问题遥遥领先，对于卷积神经网络而言，主要原理有两个，一个是稀疏交互，另一个权值共享。其中稀疏交互是指在神经网络中输入与输出并非是全连接，即并非每个输出神经元都与输入神经元产生连接与交互，而是每个输出神经元都只与部分输入交互，这样能达到一种减少计算量的目的，同时当我们只需要研究图像的一些小特征的话，网络稀疏交互意味着我们可以通过提取部分输入的像素点而非全部像素点所得到这样的小特征(比如边缘特征)。卷积的另一原理即权值共享是指对于网络的多个输入神经元可以共享一个权值矩阵，这种权值共享保证我们不需要对每个输入与每个输出之间做一次非线性的映射，只需要学习出一个权值集合然后进行共享操作。

* + 1. 卷积网络的操作

卷积:把上面的权值集合称作滤波器，利用这个滤波器对图像矩阵进行内积的操作就称为卷积操作。内积操作如图2.2(卷积操作的示意图)所示是指逐个元素相乘再相加。

|  |
| --- |
| https://img-blog.csdn.net/20160702215705128 |
| **图2.2 卷积操作示意图(原7\*7图像矩阵，取出3\*3的部分矩阵与3\*3的卷积核做内积之后的得到卷积结果)** |

转置卷积，它的操作就像卷积的逆过程一样，但又在内容上存在差异，因为转置卷积是利用卷积核对输入进行上采样操作，在上采样期间，为了使输出的大小大于输入大小，对输出采用填充0的方法，不断扩充输出，所以这个输出不是卷积的输入。

|  |
| --- |
|  |
| **图2.3 转置卷积操作示意图 (从像素点经过3\*3的过滤器进行**  **转置卷积填充0不断扩充输入)** |

池化，池化的目的还是为了减少计算量，池化分为最大池化和平均池化。最大池化是指是对卷积得到的小特征再提取最能代表图片的特征，即特征数值最大的。同理平均池化是对卷积得到的小特征进行特征值相加再求平均得到相对能表示图像特征的数值。

* + 1. 卷积网络在图像修复上的应用

现行的卷积网络应用于图像修复主要是与自编码器结合，刚开始Pathank[8]等人提出一种输入损失小部分的图像，然后利用上下文信息进行卷积自编码器训练，最后输出针对破损部分的修复图，可是很快它们发现生成的图片过于模糊，然后它们增加了判别修复图是来自原图还是自编码器解码生成的对抗损失，虽然这个确实得到很高的清晰度，但是这样不能保证生成修补区域与原图未破损部分的一致性，好在之后提出了全局对抗损失，将修复图拼接到原图再对图像真伪性进行判别，成功解决了这个问题。

卷积自编码器主要包含两个编码器，两个编码器都是输入待修复的图像，不同的是一个是输出缺失小块的修复图，并把修复图与缺失小块的原图作对比生成局部对抗损失，另一个编码器是输出修补后的完整图与原图作全局对抗损失，努力使这样两个损失之和最低是卷积自编码器训练的目标，图2.4是卷积自编码器图像修复原理图。

|  |
| --- |
|  |
| **图2.4 卷积自编码器进行图像修复的原理图（待修复图输入编码-解码网络**  **生成了修补小块，对修补小块与原缺失小块以及完整原图做损失更新）** |

* 1. 生成对抗网络与图像修复
     1. 生成对抗网络简介

生成对抗网络其实两个网络互相博弈，最终达到纳什均衡。这两个网络一个是生成器网络，它的目标接受随机噪声，不断训练生成假图片，为了方便叙述把生成假图片的函数记作G(z)。另一个网络是判别器网络，该网络负责判断输入的图片是由G(z)生成的假图片还是原图，如果是原图的会输出1，如果是假图片输出0.这两个网络的博弈过程具体说就是生成器网络输入随机噪声后不断生成足够接近真实图片的假图片去欺骗判别器网络，而判别器网络就是要把生成器网络生成的假图片判断出来，这样两者互相博弈，最终到达纳什均衡。图2.5是GAN网络的对抗生成原理图。

|  |
| --- |
| https://segmentfault.com/img/remote/1460000016861246 |
| **图2.5 GAN网络生成图片原理图 （输入随机噪声生成器生成的假图片以及真图片又作为判别器的输入，判别器判别错误惩罚判别器，判别正确惩罚生成器）** |

* + 1. 生成对抗网络在图像修复上的应用

把GAN网络应用于图像修复[9]时，首先不考虑的是使用破损的不完整图像进行修复，而是采用完整的原图进行GAN网络训练图像生成模型。首先随机输入一些噪声，使用GAN模型的生成器与判别器互相博弈，使生成器能有产生接近原图的假图片的能力。但是此时的模型还不足够生成最佳的假图片进行图像修复，因为该假图片的输入的概率分布不一定与待修复图片的概率分布一致，所以对假图片的输入噪声进行迭代更新，直到训练后输入的概率分布与待修复图片一致，即可提取由生成器生成的假图片和破损原图一样大一样位置的破损块，利用破损块与待修复图做简单的通道拼接之后即可完成图像修复工作。图2.6即为使用GAN网络进行图像修复的原理图。

|  |
| --- |
|  |
| **图2.6 GAN网络进行图像修复的原理 (随机噪声经过生成器多次迭代后提取生成后假图片的部分填充到待修复图)**   * 1. 深度卷积对抗网络与图像修复      1. 深度卷积对抗网络简介   DCGAN网络简单来理解就是把GAN网络框架的生成器和判别器用卷积神经网络实现。但是在进行实验时，为了提高训练收敛的速度，DCGAN对卷积网络的结构进行了一些修正，这些修正如下：  (1) 取消卷积后的池化层，在判别器网络中取消的池化层改用卷积层，在生成器网络中取消的池化层改用转置卷积层。  (2) 对判别器网络训练时除了输出层外都采用relu函数进行激活。输出层为保证输出为两个输出，采用sigmoid函数。  (3) 对生成器网络训练时除了输出层外都采用修正的ReLU函数LeakyReLU函数进行激活。输出层为保证输出在-1到1之间，采用tanh激活。图2.7是DCGAN网络的生成器。   |  | | --- | |  | | **图2.7 DCGAN网络的生成器 (100维噪声经过三层转置卷积层图像**  **由窄深变为宽浅)** |   (4) 卷积后每一层在激活前都使用batch\_normalization归一化，避免参数过大，出现过拟合现象。  (5) 取消密集（dense）层。   * + 1. 深度卷积对抗网络与图像修复   DCGAN网络实际是GAN网络的升级，所以使用DCGAN网络进行图像修复的原理与GAN网络相似。总的来说就是训练生成器具备生成假图片的能力，然后不断调整输入，使之能找到最佳假图片。关于原理细节和实验细节将在第三，四章详细介绍。   * 1. 三种网络在图像修复上的比较   实验证明CNN，GAN,DCGAN都能应用于图像修复，但是三种网络进行图像修复时都存在优缺点，在综合表2.1对三种网络的优缺点比较之后，下面的实验选用DCGAN网络进行图像修复。  **表2.1 CNN，GAN，DCGAN在图像修复上的比较**   |  |  |  | | --- | --- | --- | | 修复图像方法 | 优点 | 缺点 | | 卷积自编码器 | 图像的任意破损区域都可修复 | 只能对生成低分辨的修复图像；图像修复的拼接效果不佳 | | 使用GAN网络 | 可以生成高分辨率修复的图像；GAN比CNN学到的特征更能代表原图 | 训练时不易收敛；训练不稳定，造成生成器的生成效果不理想；G的生成优化来自于D的分辨结果的反馈，若D的分辨结果过佳，则会造成G的梯度消失；若G生成假图片的效果过好，则会造成D无法分辨真假图片，导致模式崩溃。 | | 使用DCGAN网络 | 结合CNN和GAN的优势，降低GAN网络训练的不稳定性 | 对GAN训练的不稳定性虽然有-所降低，还是会存在不稳定 | |

图像生成模型的构建

在进行图像修复之前，首先必须做的是环境配置与图像生成的模型构建。

* 1. 环境配置

下面将讲解本实验的环境配置，本地配置和网络配置可以任选其一配置。

* + 1. 本地环境配置

(1)Anaconda的下载与安装：百度搜索Anaconda找到官网，运行安装包即可。（注意：在运行Anaconda安装程序最后行要勾选添加环境变量和设置默认python版本）。

1. 安装TensorFlow：
2. 在cmd中激活环境：activate Anaconda3
3. 接着使用pip安装: pip install tensorflow-gpu

(5) 安装cuda：去NVIDIA官网下载运行GPU所需的CUDA，根据自己的操作系统和显卡型号选择合适的CUDA版本

(6) 运行JupyTerNotebook输入测试代码：import tensorflow as tf

|  |
| --- |
|  |
| **图3.1 本地成功配置tensorflow图** |

* + 1. 网络环境配置方法1(无GPU电脑推荐)

1. 拥有一个google账号，打开google\_drive；
2. 在google\_drive添加关联应用Colaboratory；
3. 在google\_drive新建一个ipynb文件；
4. 打开该文件，在文件选项框，选择修改-》笔记本设置-》设置为GPU；
5. 网络环境不用安装tensorflow，可以直接运行。

|  |
| --- |
|  |
| **图3.2 Colaboratory提供的GPU云的配置详情图** |

* + 1. 网络环境配置方法2(无GPU电脑推荐)

1. 登录kaggle官网，打开kernel控制台；
2. 打开GPU开关即可运行。

|  |
| --- |
|  |
| **图3.3 Kaggle提供的GPU云的配置详情图** |

* 1. 深度卷积对抗网络的激活函数

**表3.1 DCGAN网络的激活函数**

|  |  |  |
| --- | --- | --- |
| **激活函数** | **函数图像** | **应用** |
| Tanh |  | 适用输出需要为-1到1的层 |
| Sigmoid |  | sigmoid函数当输出为0和1时曲线趋向平缓，所以适合二分类问题 |
| Relu |  | Relu函数比sigmoid收敛的快，但是只适用输出为正数的问题 |
| LeakyReLU |  | 修正版的Relu函数，可以适用输入含有负数的问题 |

* 1. 深度卷积对抗网络的代价函数

DCGAN模型的建立代价函数，让我们回顾一下前面所提到的，生成器要做的是生成足够可以迷惑判别器的假图片，让判别器输出1，而生成器的目标是判断输入的图片是否使真图片，如果是真图片的话，输出1，如果是假图片，输出0，所以代价函数必须要同时考虑生成器和判别器两者的目标。设由生成器生成的图片记为G(Z),X为原图,判别器判断的结果记作D(X),那么转化成符号语言就是G的目标就是把让D(G(Z))逼近1，即1-D(G(Z))逼近0，而D的目标就是让D(G(Z))逼近0，并且让D(X)逼近1，这样就得到代价函数。

**表3.2 代价函数**

|  |  |
| --- | --- |
| **D的代价函数** | **G的代价函数** |
|  |  |

* 1. 深度卷积对抗网络的图像生成训练步骤

1. 处理图片数据集：
2. 使用scipy库函数切割图片;
3. 划分图片数据集为训练集与验证集;
4. 对图片数据集进行规范化，使图片数据集的各像素点的值在[-1,1];
5. 对图片数据集进行随机打乱，并把打乱的数据集切分多个batch，方便训练（以后训练以batch为单位）。
6. 构建生成器：确定生成器有几层结构，确定输入噪声维度，转置卷积层的层数以及每一层卷积核大小，输入通道数，输出通道数，转置卷积后是否采用统一规范化和激活函数选择。
7. 构建判别器：确定判别有几层结构，卷积层的层数以及每一层卷积核大小，输入通道数，输出通道数，卷积后是否采用统一规范化和激活函数选择。
8. 构造代价函数：使用交叉熵代价函数构造如表3.1的代价函数。
9. 构造优化器：确定Adam优化器的学习率(learning\_rate)以及梯度加总的衰减率（beta1）。
10. 创建tensorflow会话，确定epoch，运行各张量和模型。
11. 运用测试集或者选定的样本集运行模型，测试生成的假图片。
12. 根据假图片生成情况，对模型进行优化，参数调整等操作，直至生成的假图片的效果达到预期。
    1. 基于深度卷积对抗网络的人脸数据集图像生成模型

在进行Celeba数据集应用DCGAN网络进行图像生成模型之前，本文已经采用过数据集mnist或者cifar10数据集进行了模型的初建，但是因为考虑到mnist数据集构成简单和cifar10类别数太多，而DCGAN网络只能对特定类别的图片进行修复，如果使用DCGAN网络对其他数据集缺失固定部分的图像进行图像修复训练，可能效果不佳，所以本文采用Celeba图像集进行修复训练，本章重点讲解Celeba图像生成构建的原理与过程。

* + 1. 人脸图片数据集简介

Celeba数据集最初是由The Chinese University of Hong Kong的MMLAB收集的用来训练的人脸图片集，该人脸图片集包含202,599张各种名人的脸部图像，并对人脸的5个标志：眼睛的左边与右边，鼻子，嘴巴的左边与右边进行标记，本文采用的数据文件img\_align\_celeba.zip为所有脸部图像进行了裁剪和对齐的数据集文件。

|  |  |  |
| --- | --- | --- |
| |  | | --- | |  | | **图3.4 Celeba数据集部分图片展示** | |
|  |

* + 1. 基于人脸图片数据集的图像生成模型的构建

下面按照本章描写的步骤对Celeba图片进行图像生成模型的构建。本部分将会详细介绍模型的形成。

1. Celeba图片数据的处理
2. 使用scipy的库函数对Celeba图片集进行裁剪[shape=64\*64\*3]。
3. 首先从数据集中提取样本集。
4. 对数据集进行处理。

对数据具体的数据处理步骤和方法如表3.3所示。

**表3.3 数据处理步骤图**

|  |  |
| --- | --- |
| **处理步骤** | **处理方法** |
| 对训练集进行剪切 | 使用scipy的crop方法 |
| 对每张图像进行规范化，范围在[-1,1] | 因为像素值在[0,255],所以对每个像素点除以255乘以2再减1 |
| 把图片打乱 | 使用tf的shuffle方法 |
| 把图片组合为batch  batch\_size=128 | 设定循环变量，以batch\_size为单位循环切割图片 |

1. 构建判别器

图3.5展示的Celeba判别器的模型的隐藏层由四层卷积层组成。为了方便叙述，下面的模型解释以单张图片为例，此四层卷积层都采用卷积步长为2的5\*5的过滤器（卷积核），并使用same方式填充，且都在卷积后进行规范统一化，都采用relu作为激活函数。进行上述操作的神经层实际输出了缩小两倍的图像和增大的通道数。构建判别器实际上进行四次上述操作，使图像缩小了十六倍（由64\*64变为4\*4），输出通道数不断增大(由3变为64再变为128再变为256最后变为512),最后输出层先用一次展平操作（即把4\*4\*512拉成一维向量），再使用全连接层使输出为1\*1，并使用sigmoid函数激活。

|  |
| --- |
|  |
| **图3.5 Celeba判别器构造模型 (输入的64\*64 RGB三通道的图片**  **经过四层卷积层输出分类)** |

1. 构建生成器

图3.6展示的Celeba的生成器模型的隐藏层由四层转置卷积层组成，同样为了方便叙述，下面的模型解释以单张图片为例，首先第一层使用全连接层把100维的输入噪声变为输出维度为8192\*1的图像矩阵，然后使用reshape把图像矩阵变为4\*4\*512，之后四层转置卷积层都采用卷积步长为2的5\*5的过滤器（卷积核），并使用same方式填充，且都在转置卷积后进行规范统一化，都采用LeakyReLU作为激活函数。进行上述操作的神经层实际输出了增大两倍的图像和减少的通道数。构建生成器实际上进行三次上述操作，使图像增加了八倍（由4\*4变为64\*64），输出通道数不断减少(由512变为256再变为128再变为64最后变为3),最后输出层为了保证输入在[-1,1]，使用tanh函数激活。

|  |
| --- |
|  |
| **图3.6 Celeba生成器构造模型 (输入的100维噪声经过四层转置卷积层**  **生成64\*64 RGB三通道的假图片)** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. 代价函数与优化器的构建   表3.4详细展示了图像生成的代价函数与优化器的创建的方法和代码。  **表3.4 图像生成代价函数与优化器的构建**   |  |  |  | | --- | --- | --- | | 判别器代价函数 | 方法 | 因为判别器想要输入真图片输出1，想要输入假图片输出0，所以需要两个代价函数，一个利用交叉熵函数判断输入真图片输出与1的差别，另一个利用交叉熵函数判断输入假图片输出与0的差别，但是为了防止判别效果太好,导致梯度消失,所以添加很小的参数smooth,输入真图片时判断真图片与1个略比1小的差别。当两个差别过大，就惩罚判别器，以此得到好的判别效果。 | | 代码 | #真图片输出与略比1小的标签的差距  d\_loss\_real = tf.reduce\_mean(tf.nn.sigmoid\_cross\_entropy\_with\_logits(  logits=d\_logits\_real, labels=tf.ones\_like(d\_logits\_real) \* (1 - smooth)))  #假图片输出与标签0的差距  d\_loss\_fake = tf.reduce\_mean(tf.nn.sigmoid\_cross\_entropy\_with\_logits(  logits=d\_logits\_fake, labels=tf.zeros\_like(d\_model\_fake)))  #两个差距之和构成判别器代价函数  d\_loss = d\_loss\_real + d\_loss\_fake | | 生成器代价函数 | 方法 | 因为生成器想要自己构造的假图片让判别器输出1，所以生成器的构造利用交叉熵函数判断判别器输入假图片输出与1的差别。 | | 代码 | #假图片输出与标签1的差距  g\_loss = tf.reduce\_mean(tf.nn.sigmoid\_cross\_entropy\_with\_logits(logits=  d\_logits\_fake, labels=tf.ones\_like(d\_model\_fake)) | | 判别器优化器 | 方法 | 输入学习率learning\_rate，权重衰减参数beta1，使用Adam优化器 | | 代码 | #判别器优化器  discriminator\_optimizer = tf.train.AdamOptimizer(learning\_rate=  learning\_rate, beta1=beta1) | | 生成器优化器 | 方法 | 输入学习率learning\_rate，权重衰减参数beta1，使用Adam优化器 | | 代码 | #生成器优化器  generator\_optimizer = tf.train.AdamOptimizer(learning\_rate=learning\_rate, beta1=beta1) |  1. 参数确定   此参数表3.5是经过不断优化实验和参考文献[10]所得。  **表3.5 Celeba训练参数表**   |  |  | | --- | --- | | **参数** | **参数值** | | 输入噪声维度(z\_size) | 100 | | 学习率(learning\_rate) | 0.0002 | | 权重衰减率(beta1) | 0.5 | | 增加对抗攻击(smooth) | 0.1 | | LeakyReLU的斜率(alpha) | 0.2 | | 训练的batch的大小（batch\_size） | 128 | | 卷积核大小（ksize） | 5\*5 | | 卷积步长(stride) | 2 | | 训练次数(epoch) | 10 | |

* + 1. 图像生成模型的运行以及分析改进

1. 图像生成模型的运行

此次模型采用优化一次判别器，再优化一次生成器的方法进行图像生成，图3.7展示了图像生成的过程。(输入为100维随机噪声向量，经过迭代不断生成假图片的过程)

|  |
| --- |
|  |
| **图3.7 图像生成过程图 (不断优化生成器,输入最左边的随机噪声一步步优化生成器得到不断清晰的人脸生成图像)** |

1. 图像生成模型的改进

虽然使用一次判别器迭代，一次生成器迭代达到图像生成的目的，但是通过观察图3.8损失变化图(红线是生成器生成损失，蓝线是判别器判别损失)，容易发现判别器判别损失有趋于0的现象，此种现象叫做梯度消失。

|  |
| --- |
|  |
| **图3.8 未优化的损失变化图 (未优化时出现了判别器损失接近0的现象)** |

为了解决这种梯度消失现象，本文对图像生成模型进行了改进，在原有基础上增加了一次生成器损失，也就是优化后的模型变为一次判别器优化迭代，两次生成器迭代优化。图3.9展示了优化后的损失变化图。很明显从图中能看出梯度消失得到了解决。

|  |
| --- |
|  |
| **图3.9 优化后的损失变化图 （优化后生成器损失与判别器损失的对抗达到新的平衡）** |

图3.10展示了进行优化后的图像生成过程，可以从图中看出图像生成的迭代效果也比之前有了提高。

|  |
| --- |
|  |
| **图3.10 优化后的图像生成过程图 (不断优化生成器,输入最左边的随机噪声一步步优化生成器比优化前同样的迭代次数更快得到不断清晰的人脸生成图像)** |

虽然进行了增加一次生成器优化的方法来缓解梯度消失，但是在实验过程中，随着迭代次数增多，还是出现了梯度消失，所以梯度消失只能缓解不能彻底解决。

* + 1. 深度卷积对抗网络在图像生成模型上的优势

通过对图像生成模型进行优化之后，该图像生成模型不仅缓解了梯度消失的问题，而且比起优化前和celeba数据集的原图来说，图像的分辨率都得到了提高，这也证明DCGAN网络在图像生成的优势。

图像修复

经过上一章节，我们已经搭建好图像生成模型，那么本章节将基于图像生成进行图像修复。

* 1. 图像修复原理

当构建好图像生成模型，就可以正式开始图像修复。首先输入随机噪声，随机噪声经过图像生成模型能生成假图片。然后我们利用假图片去修补原图像的缺失部分（破损部分）,换句话来说就是利用原图像的缺失部分在假图片上的投影替换原图像的缺失部分。图4.1展示了图像修复的原理。

|  |
| --- |
|  |
| **图4.1 图像修复原理图 (提取生成的假图的对应部分与破损原图进行拼接修复)** |

* 1. 建立投影方程

为了使获取破损部分在假图片的投影，需要用掩膜对图像进行处理。掩膜的作用是对图像特定部分的遮挡。在实际应用中，一般使用掩膜运算处理图像数据。

掩膜矩阵是一个只包含0和1的矩阵，图矩阵与掩膜矩阵相乘后，图矩阵对于掩膜矩阵为1的位置原样保留，对于掩膜矩阵为0的位置进行置0操作(不显示该位置)。图4.2展示了一个2\*2的图矩阵与一个2\*2掩膜运算的过程图。

|  |
| --- |
|  |
| **图4.2 掩膜运算图 (图矩阵经过掩膜矩阵,图矩阵在掩膜矩阵为1的投影原样保留，为0的地方为0)** |

使用一个与破损部分一样大的全0掩膜矩阵与原图矩阵进行运算，可以得到遮挡破损部分的原图，之后使用一个与破损部分一样大的全0掩膜矩阵与假图片矩阵进行运算，可以得到破损部分在假图片的投影。

使用掩膜运算之后，就可以得到修复图像的数学化公式（LC表示修复图，G(z(i))表示生成的假图片，x表示原图,MASK表示与原图同样大小的只有破损部分对应部分为全1，其它部分为全0的掩膜矩阵）：

 (4-1)

因为训练的时候不可能得到大量同一位置缺失的破损图像,所以使用与原图同样大小的只有破损部分对应部分为全0，其它部分为全1的掩膜矩阵生成缺失相同位置的图片。

* 1. 图像修复与图像生成
     1. 图像修复与图像生成的关系

为了使投影切近原图，需要通过DCGAN模型训练好的生成器不断生成假图片，不断迭代求解可替换的最优假图片。所以在进行图像修复模型时，必须复用之前的图像生成的模型。但是在进行图像修复时我们所优化的不再是生成器，而是优化生成器的输入，因为之前生成器的输入一直是随机输入的分布，而进行进行图像修复时。必须迭代求解找到所修复图片的分布，这就需要迭代优化生成器的输入z,图4.3展示的正是DCGAN网络迭代求解假图片的过程。

|  |
| --- |
|  |
| **图4.3 DCGAN网络迭代求解假图片的效果图 (为了使生成器输入分布贴近修复图分布需要不断迭代优化输入分布直至分布贴近生成器产生最佳假图片)** |

* + 1. 图像修复与图像生成的区别

#### 定义的代价函数

文献[11]提到图像修复主要依赖图像的两个信息，一个是上下文信息，一个是感知信息。

上下文信息主要指缺失图像的周边信息。

感知信息主要指缺失部分在用户的感觉中是否是能很好填补原图。

有了这两种信息就可以定义图像修复的代价函数。表4.1很好的解释了代价函数的构成。

**表4.1 图像修复代价函数的构成**

|  |  |
| --- | --- |
| **代价** | **原理** |
| 上下文信息 | 利用缺失部分的原图与缺失同样部分的假图片的差距来构成代价  (4-2) |
| 感知  信息 | 因为用户的感觉的作用和判别器的判别作用相似，所以该部分代价沿用假图片生成的代价(假图片生成被判别器误判的可能性)  (4-3) |
| 图像  修复 | 上下文信息代价加合感知信息代价  公式表示：(lam为调整的感知代价对优化的比例) (4-4) |

#### 重定义的优化器

与图像生成类似，图像修复同样使用Adam优化器对图像修复代价来进行最小优化。不过图像修复只会对生成器地输入进行迭代优化。

|  |
| --- |
|  |
| **图4.4 图像修复迭代优化z过程图 (橙线表示当假图片与原图片上下文信息差异过大就优化随机噪声z，蓝线表示当假图片与原图片的感知差异过大就优化随机噪声z)** |

* 1. 图像修复模型的构建与运行
     1. 图像修复模型构建

在原来的图像生成模型相应的增加一个图像修复代价优化器即可完成图像修复模型的构建。在训练时，应该在进行图像修复之前应该先迭代训练图像生成模型，找寻到能生成假图片的权重矩阵，再利用该权重矩阵进行图像修复模型迭代优化，具体来说是先完成图像生成迭代，然后停止图像生成优化，对图像生成的输入噪声迭代优化，找到与修复图相同的输入分布之后生成最佳假图片，提取假图片投影进行拼接即完成图像修复，图4.5展示这种图像修复模型的运行过程。

|  |
| --- |
|  |
| **图4.5 图像修复模型的运行过程图 (蓝色线为输入噪声z经过生成器进行**  **假图片生成的过程，橙色线为寻找最佳假图片不断调整输入分布z，得到**  **最佳假图片进行图像修复拼接的过程)** |

* + 1. 图像修复参数的确定

对于图像修复模型来说，需要保持原来图像生成模型的参数，然后设置新参数，表4.2显示了新参数的设置，其中Adam优化器的参数值参考了文献[12]

**表4.2 图像修复训练参数表**

|  |  |
| --- | --- |
| **参数** | **参数值** |
| 图像的迭代次数(nIndex) | 300 |
| 感知损失比重(lam) | 0.1 |
| 一阶矩估计的衰减率(beta1) | 0.9 |
| 二阶矩估计的衰减率(beta2) | 0.9999 |
| 防止除以0的很小的数(eps) | 0.2 |
| 学习速率（lr） | 0.001 |

* + 1. 图像修复模型的运行

因为每个修复图像的概率分布不同，所以此次的图像修复模型采用了对每张图片进行输入分布迭代优化的方法。图4.6展示图像修复的过程图，左边的图片是生成的最佳假图片，右边的图片是修复图。

|  |
| --- |
|  |
|  |
|  |
| **图4.6 迭代次数不同的图像修复效果图，左边是最佳假图片，右边是修复图，a)为迭代100次时的效果图，b)为迭代200次的效果图，c)为迭代300次的效果图** |

总结与展望

* 1. 本论文所做工作总结

现今图像修复主要还是依靠专业修复师来完成，人们因此开始思考自动化图像修复的实现，在这个思考过程中，深度学习领域与计算机视觉领域的结合给了人们新的思考方向。卷积网络，生成对抗网络的兴起，使得图像的自动化修复成为可能，并且未来可能达到较高的修复效果。回看本论文的研究内容，本论文主要围绕以下内容展开:

1. 对时下最热门的三种基于深度学习的图像修复技术进行网络的讲解和修复原理的简述。并对这三种方法进行对比之后，阐明了选择DCGAN网络修复图像的原因。
2. 在选用DCGAN网络进行图像修复之后，首先阐述了图像修复的最重要一环节——图像生成的环境配置，原理，代价函数，激活函数，并介绍了利用Celeba数据集进行图像生成的步骤以及实验结果分析，并在实验中对梯度消失现象提出了增加一次生成器优化的解决方案，并实践证明该方案的可行性。
3. 通过前面章节图像生成模型的构造，在此基础上讲了图像修复的原理，图像修复与图像生成的区别与联系，并再次用Celeba数据集进行图像修复的模型的构建与运行，最终完成了图像修复。
   1. 进一步研究设想与展望

使用本论文的方法对图像修复的效果仍然与专业的修复师差距较大，所以需要今后不断改进，今后的改进工作如下：

1. 生成对抗网络会出现梯度消失和模式崩溃的问题，没有一个特定的训练次数要求或者较好的解决方案，希望在之后的研究中能找到更好的解决方法达到更好的图像修复效果。
2. 对于只把上下文信息的损失作为图像修复的代价也对图像修复效果造成了影响，希望之后能发现其他能作为图像修复代价的损失函数使图像修复更上一层楼。
3. 部分图像修复并不能保证完美的修复贴合，期待继续研究图像修复问题，并解决这个贴合问题。

参考文献

1. 陈欢.图像修复技术[J].科技风, 2014(14):5+7
2. Bertalmio M, Bertozzi A L, Sapiro G.Navier-stokes, fluid dynamics, and image and video inpainting[C]//Proceedings of2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.Kauai, HI, USA:IEEE, 2001:355-362.[DOI:10.1109/CVPR.2001.990497]
3. Barnes C, Shechtman E, Finkelstein A, et al.PatchMatch:a randomized correspondence algorithm for structural image editing[J].ACM Transactions on Graphics, 2009, 28 (3) :#24.[DOI:10.1145/1531326.1531330]
4. Yang C, Lu X, Lin Z, et al.High-resolution image inpainting using multi-scale neural patch synthesis[C]//Proceedings of2017 IEEE Conference on Computer Vision and Pattern Recognition.Honolulu, HI, USA:IEEE, 2017:4076-4084.[DOI:10.1109/CVPR.2017.434]
5. Lecun Y, Bottou L, Bengio Y, et al.Gradient-based learning applied to document recognition[J].Proceedings of the IEEE, 1998, 86 (11) :2278-2324.[DOI:10.1109/5.726791]
6. Goodfellow I J, Pouget-Abadie J, Mirza M, et al.Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems.Montréal, Canada:ACM, 2014:2672-2680.
7. 洪洋. 深度卷积对抗生成网络综述[A]. 中国自动化学会系统仿真专业委员会、中国系统仿真学会仿真技术应用专业委员会.第18届中国系统仿真技术及其应用学术年会论文集（18th CCSSTA 2017）[C].中国自动化学会系统仿真专业委员会、中国系统仿真学会仿真技术应用专业委员会:中国自动化学会系统仿真专业委员会,2017:5.
8. Pathak D, Krhenbühl P, Donahue J, et al.Context encoders:feature learning by inpainting[C]//Proceedings of 2016 IEEEConference on Computer Vision and Pattern Recognition.Las Vegas, NV, USA:IEEE, 2016:2536-2544.[DOI:10.1109/CVPR.2016.278]
9. 强振平,何丽波,陈旭,徐丹.深度学习图像修复方法综述[J].中国图象图形学报,2019,24(03):447-463.
10. 曹志义,牛少彰,张继威.基于半监督学习生成对抗网络的人脸还原算法研究[J].电子与信息学报,2018,40(02):323-330.
11. R Yeh, C Chen, TY Lim, et al. MN Do - arXiv preprint arXiv:1607.07539, 2016
12. Alec Radford，Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks，[[1511.06434] Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks](https://link.zhihu.com/?target=https%3A//arxiv.org/abs/1511.06434)

附 录

致 谢

时光匆匆，转眼又到了属于我的毕业季。大学四年所学到的知识和在母校的点点滴滴都会是我最宝贵的财富。

感谢我的导师金长龙老师，首先感谢老师，因为老师的深度学习课程上所教授的知识给我毕设打下的坚实基础，其次感谢老师在毕设整个阶段对我的敦敦教导，无论是课程的选取，研究的方向，还是之后的图像修复知识的学习以及实验指导还有论文编写，老师总是帮助我理清思路，努力引导我自主解决问题，提高了我的学习能力。

感谢好友林晓婷和陈茜，感谢你们在论文整个阶段对我的鼓励和帮助。

感谢我的父母和家人，是你们的支持与激励，使我不断进步。

最后衷心感谢各位专家，能在百忙之中抽出宝贵时间阅读我的论文，敬请各位老师批评与指正论文的不足。