|  |
| --- |
| 吉 林 大 学  开放性创新实验报告  **（基于卷积神经网络的图像风格迁移）**  **姓 名： 史昊林**  **专 业：材料科学与工程学院**  **年 级：2019级**  **小 组：第四组**  **指导教师： 于美铭**  **年 月** |

吉林大学开放性创新实验报告

|  |  |  |  |
| --- | --- | --- | --- |
| **姓名** | 史昊林 | **项目名称** | 基于卷积神经网络的图像风格迁移 |
| **性别** | 男 | **项目编号** | [**966OEP83**](https://oep.jlu.edu.cn/search.php?s=966OEP83) |
| **学院** | 材料科学与工程学院 | **电话** | 15202963290 |
| **学号** | 1619026 | **邮箱** | 1665105642@qq.com |
| **实验目的：通过上课老师所讲和智能计算系统中深度学习知识学习卷积神经网络和学习“A neural algorithm of artistic style”论文，实现基于卷积神经网络的图像风格迁移实验** | | | |
| **实验准备：1.python3.7**  **2.pycharm2021.3**  **3.tensorflow**  **4.卷积神经网络算法** | | | |
| 总体设计： 以下展示的结果是基于论文的VGG网络生成的，用于训练执行物体识别和定位，更多细节如论文所示。使用标准的19层VGG网络包含16个卷积层和5个池化层提供的特征空间。按比例改变权重规范化网络，这样每层卷积滤波器在图像和位置上平均激活值就等于1.这种针对VGG网络的按比例缩放不会改变它的输出，因为它只包含修正线性激活函数，在特征地图上没有归一化层和池化层。我们也不使用任何全连接层。模型是公开的，可以在caffe架构中找到。对于图像合成，我们发现用平均池化取代最大池化操作，生成的结果更好，这就是图像使用平均池化层来生成。内容表示 **通常网络中每层定义了一个非线性滤波，它的复杂度随着每层在网络中的位置而增加。给定一个图像 ，卷积神经网络每层使用滤波对图像进行编码。Nl宽的滤波器的一个层大小为Ml的每个有Nl个特征地图，其中Ml是特征地图的高乘以宽。所以，l层的响应可以存储在一个矩阵中 ，其中 是l层的位置j上的第i个滤波器的激活值。为了将图像信息可视化在层级结构的不同层上进行编码，可以在一个带有白噪声的图像上执行梯度下降算法寻找可以匹配原始图像特征响应的另外一个图像（参考图1的内容重建），如论文24所述。令 和 分别为原始图像和生成的图像， 和 分别是l层的特征表示。那么，就可以定义两类特征表示之间的误差平方损失函数为：**  **损失函数的偏导数对应的l层的激活函数为：**  **图像的梯度可以用标准误差反馈传播计算（如图2右侧）。因此，我们可以改变初始随机图像直到卷积神经网络的某一层可以生成与原始图像相同的响应。**  **图1. 卷积神经网络中的图像表示。在卷积神经网络的每个处理阶段，一个给定的输入图像表示为滤波过的图像。滤波器的数量沿着处理的层级增加，滤波后的图像用某种降采样机制减少（比如，最大池化），可以减少网络中每层的总数量。**内容表示**：可以从特定的一个网络层上，在只知道网络响应的情况下重建输入图像，就能够在卷积神经网络的不同处理阶段将信息可视化。在原来的VGG网络中的‘conv1 2’ (a), ‘conv2 2’ (b), ‘conv32’ (c), ‘conv4 2’ (d) ‘conv5 2’ (e) 重建输入图像。可以发现从网络中的低层重建接近完美（a-c）。在网络的高层，细节像素信息会丢失，但高层的图像内容会被保留下来（d，e）。**风格表示**：在卷积神经网络顶层激活时，使用一个特征空间获取输入图像的纹理信息。风格表示计算了卷积神经网络不同层不同特征间的关系。从卷积神经网络层（‘conv1 1’ (a), ‘conv1 1’ 和‘conv2 1’(b),‘conv1 1’, ‘conv2 1’ and ‘conv3 1’ (c), ‘conv1 1’, ‘conv2 1’, ‘conv3 1’ and‘conv4 1’ (d), ‘conv1 1’, ‘conv2 1’, ‘conv3 1’, ‘conv4 1’and ‘conv5 1’ (e)）的不同子集上构建的风格表示重建输入图像的风格。这样创建的图像在一个逐步增加的规模上会匹配给定图像的风格，时会丢掉场景全局结构信息。**    **卷积神经网络在物体识别上训练，它们可以生成一个对图像的表示，在网络层级的处理过程中使得物体信息越来越明显，如论文10所示。因此，沿着网络层级处理过程，输入图像会发生转变，对图像的真正内容越来越敏感，但会对精细的外观变得相对不变。网络的高层会捕捉输入图像的高阶内容，比如物体和结构，但不会限制重构过程中的确切像素值（如图1中，内容重建 d，e）。相反，低层重建只是简单地复制原始图像中的确定的像素值（如图1，内容重建 a-c）。** | | | |
| 项目开发：使用经典CNN模型 VGG19. 图片风格算法原理。实验环境：  * python 3.7和pycharm2021.3. VGG网络非常深，通常有 16/19层，称作 VGG16 和 VGG19 ，卷积核大小为 3 x 3 ，16和19层的区别主要在于后面三个卷积部分卷积层的数量。通过加载vgg19的模型搭建可视化vgg19.所以我们需要一个训练好的神经网络 VGG * 一张风格图像，用来计算它的风格representation * 一张内容图像，用来计算它的内容representation， * 一张噪声图像，用来迭代优化   给定一张风格图像 **a** 和一张普通图像 **p**，风格图像经过 VGG 的时候在每个卷积层会得到很多 feature maps, 这些 feature maps 组成一个集合 A，同样的，普通图像 **p** 通过 VGG 的时候也会得到很多 feature maps，这些 feature maps 组成一个集合 P，然后生成一张随机噪声图像 **x**（在后面的实验中，其实就是普通图像 **p**）, 随机噪声图像 **x** 通过 VGG 的时候也会生成很多 feature maps，这些 feature maps 构成集合 G 和 F 分别对应集合 A 和 P, 最终的优化函数是希望调整 **x** , 让随机噪声图像 **x** 最后看起来既保持普通图像 **p** 的内容, 又有一定的风格图像 **a** 的风格。 | | | |
| **遇到的难点：1.搭建tensorflow遇到困难**  **2.图片风格差异过大，图片大小不一样导致出现失误**  **3.算法中损失函数和程序中的bug** | | | |
| **针对难点的解决方案：1.通过csdn找到方法下载并安装tensorflow1.15.0**  **2.找到大小一样的图像**  **3.分步调试，改变参数** | | | |
| **成果展示**（程序关键代码截图，包括硬件实验的实物照片）**：**  风格图像：  目标图像：  经过5次迭代之后得到的结果图：    目标图像2：南岭标志物    结果图：    关键代码段： | | | |
| **总结**(用到的主要知识点，学习心得等) **：**通过模型和论文，我们学习   * 经典神经网络模型 VGG * 图像风格迁移的算法原理基础 * 并动手实现任意图片的风格转换   通过学习智能计算系统，我学习  1.适合图像处理的卷积神经网络  包括卷积层，池化层，归一化层，全连接层，softmax计算原理，层的排布规律。  2.基于CNN的图像分类神经网络算法  主要介绍了AlexNet（LRN，dropout），VGG，Inception系列（1\*1卷积，BatchNorm），ResNet经典的神经网络分类算法。  3.基于CNN的图像检测神经网络算法  主要介绍了检测算法评价指标（IoU，mAP），R-CNN系列（Region Proposal，NMS，ROIPooling，RPN，anchor box），YOLO，SSD。  4.循环神经网络  主要介绍了循环神经网络的应用，网络结构，正/反向过程，以及梯度消失产生原因。  5.长短期记忆模型LSTM/GRU 6.生成对抗网络GAN 包括原始GAN结构，模式崩溃，DCGAN，以及conditionalGAN。 7.Example-Image Transfer | | | |