

華南師範大學

《人工智能导论》课程项目

课 程 项 目 报 告

项 目 题 目：图像风格迁移

所 在 学 院：计算机学院

项 目 组 长：麦境轩

小 组 成 员：孙明辉，梁家灏

开 题 时 间：2023 年 3 月 30 日

一、引言

图像风格迁移是一种计算机视觉和图形图像处理领域中的技术，它可以将一张图片的风格应用到另一张图片上，从而创造出新的艺术作品。风格迁移技术最初由 Gatys 等人在 2015 年提出，通过基于神经网络的方法实现。该技术利用预先训练好的深度卷积神经网络（DCNN）对图像进行重建，将一张参考图像中的艺术风格特征应用到目标图像中，同时保留目标图像的内容信息。这样，就可以很容易地生成一个新的图像，其中包含来自参考图像的风格，同时保留了目标图像的内容。风格迁移技术的核心思想是通过一张参考图像的风格来改变目标图像的外观。因此，风格迁移技术并不仅仅是一种简单的滤镜效果，它更具有艺术性和创造性，可以用于许多应用，例如艺术创作、图像编辑、动态视频生成等。

除了基于神经网络的方法之外，还有许多其他方法可用于实现图像风格迁移，例如基于统计方法的方法，基于插值和优化的方法等。然而，基于神经网络的方法是最先进的技术，在风格迁移领域中已经获得了广泛的应用，并产生了许多有趣的应用和变体，例如多风格图像风格迁移、视频风格迁移等。

二、国内外研究现状

图像风格迁移是近年来计算机视觉领域的热门研究方向之一，国内外学者们在该领域展开了广泛的研究。

国外方面，最初提出图像风格迁移技术的是 Gatys 等人，他们提出了基于神经网络的方法，并在 2016 年发表了题为《Image Style Transfer Using Convolutional Neural Networks》的文章。随后，该领域的研究呈现出快速发展的趋势，不断涌现出新的技术和算法，例如 Johnson 等人提出的快速图像风格迁移算法，Ulyanov 等人提出的神经风格迁移算法，Dumoulin 等人提出的多风格图像风格迁移算法等等。

国内方面，近年来也有不少学者在图像风格迁移领域进行了深入研究。例如，南京大学的毛立群教授及其合作者提出了自适应子空间协同优化算法来解决图像风格迁移中的优化问题；北京大学的郭恩明教授及其合作者提出了基于生成对抗网络（GAN）的风格迁移方法，可以更加灵活地控制风格迁移的结果；中山大学的胡姗姗副教授及其合作者则提出了一种基于条件生成对抗网络（cGAN）的图像风格迁移方法，使得可以施加语义信息来驱动目标图像的风格转换。

总的来说，图像风格迁移领域的研究在国内外都非常活跃，涌现出了许多有意义的成果，使得该技术在实际应用中具有了更多的可能性。

三、模型和算法

图像风格迁移技术主要涉及两个模型：内容模型和风格模型。

1. 内容模型：

内容模型用于表示一张图片的内容特征。它通常采用卷积神经网络（CNN）来提取输入图片的特征表示。CNN 是一种能够学习到图像中特征的模型，通过多层卷积和池化操作来对图片进行特征提取，最终得到一组高维抽象的特征表示。

在图像风格迁移中，可以使用预训练好的 CNN 模型（VGG19）来提取输入图片的内容特征。将输入图片通过 CNN 模型传递后，选择一个或多个特定的卷积层，将这些层的特征作为内容模型。

```
use_gpu = torch.cuda.is_available() #定义需要优化的参数
cnn = models.vgg19(pretrained=True).features
if use_gpu:
    cnn = cnn.cuda()

print(cnn)
```

在程序中导入 VGG19，并输出 cnn

```
Sequential(
  (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (1): ReLU(inplace=True)
  (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (3): ReLU(inplace=True)
  (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (6): ReLU(inplace=True)
  (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (8): ReLU(inplace=True)
  (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (11): ReLU(inplace=True)
  (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (13): ReLU(inplace=True)
  (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (15): ReLU(inplace=True)
  (16): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (17): ReLU(inplace=True)
  (18): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (19): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (20): ReLU(inplace=True)
  (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (22): ReLU(inplace=True)
  (23): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (24): ReLU(inplace=True)
  (25): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
)
```

输出 cnn 的部分层内容

2. 风格模型：

风格模型用于表示一张图片的风格特征。和内容模型不同，风格模型不是直接使用 CNN 模型来提取特征，而是使用 Gram 矩阵来描述图片的风格信息。

Gram 矩阵是一种衡量不同特征之间相关性的方式。它是将特定层的特征图转化为矩阵后进行乘法操作得到的。Gram 矩阵反映了特定层中不同特征之间的相关性，因此可以用来描述图像的风格特征。

在图像风格迁移中，可以将输入图片通过 CNN 模型传递后，在选择的卷积层上计算 Gram 矩阵，将其作为风格模型。

```
class gram_matrix(torch.nn.Module):  
    def forward(self, input):  
        a,b,c,d = input.size()  
        feature = input.view(a*b, c*d)  
        gram = torch.mm(feature, feature.t())  
        return gram.div(a*b*c*d)
```

定义 Gram 矩阵

目前，图像风格迁移的算法可以分为两大类：非深度学习方法和基于深度学习的方法。

非深度学习方法常用的有：

1. 基于特征匹配的方法。此类方法利用经典的计算机视觉技术，如 SIFT 等，在目标图像和参考风格图像中提取特征点和特征描述子，并通过特征点间的匹配实现风格迁移。代表性算法有 Texture Transfer 算法和 Fast Texture Transfer 算法。

2. 基于图像分解的方法。此类方法将目标图像分解成多个部分，从而实现对不同部位的风格迁移。代表性算法有光流算法和基于多尺度分解的算法。

3. 基于最小化风格差异和内容差异的方法。此类方法利用统计学方法将风格和-content 解耦，并通过最小化目标图像与参考风格图像之间的差异实现风格迁移。代表性算法有 Markov Random Fields 算法和 Bilateral Filtering 算法。

我们选择在 VGG19 中选择几层作为原始图像损失层和风格图像的损失层实现 LBFGS 算法，利用 pytorch 搭配 VGG19 和英伟达 CUDA 调用 GPU 进行迁移。

```

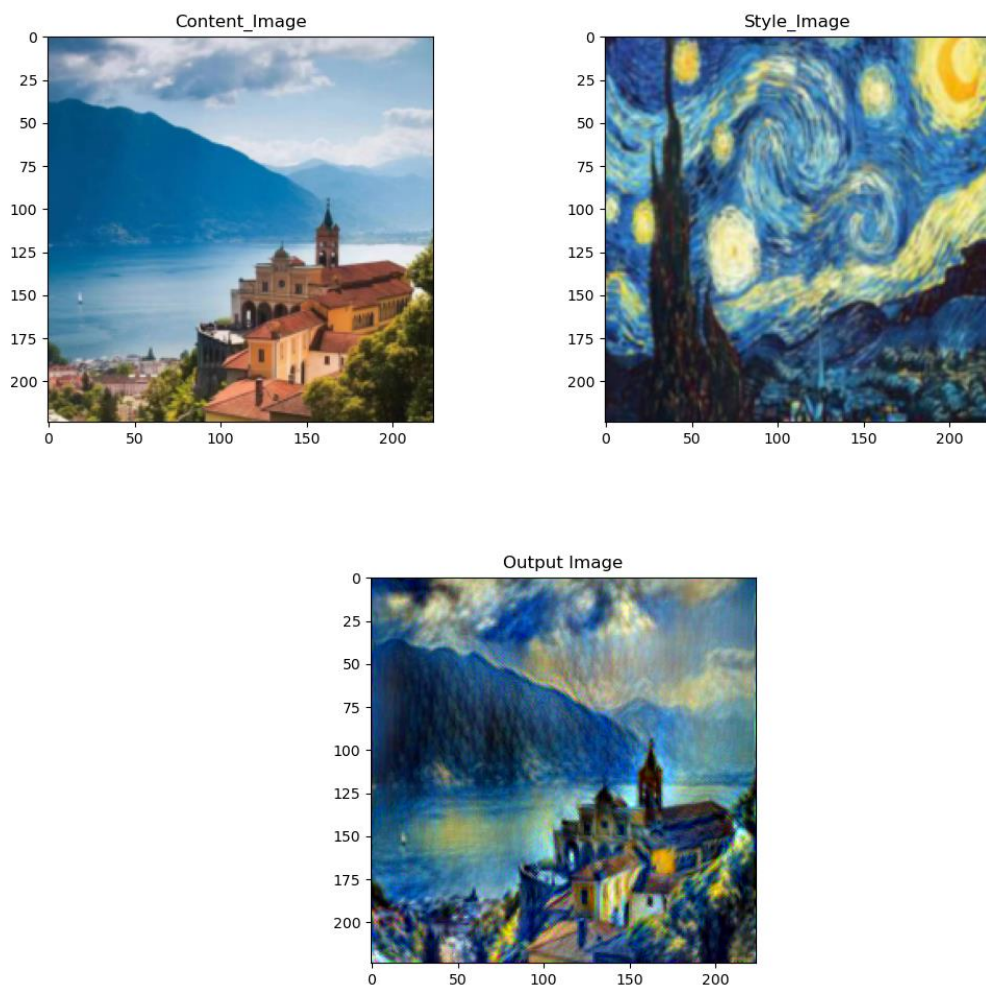
import torch
import torchvision
from torchvision import transforms, models
from PIL import Image
import matplotlib.pyplot as plt
from torch.autograd import Variable
import copy
import pylab #为了输出图片

```

导入相关包

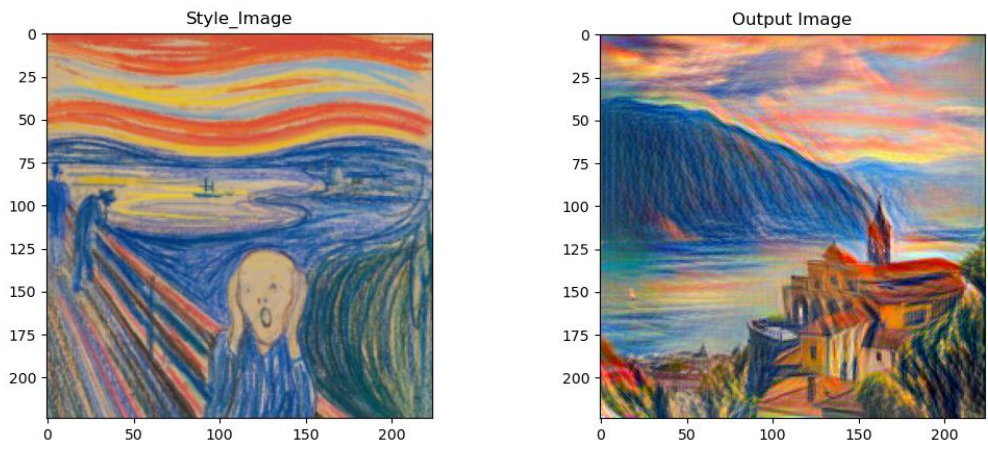
四、实验结果分析

风景：

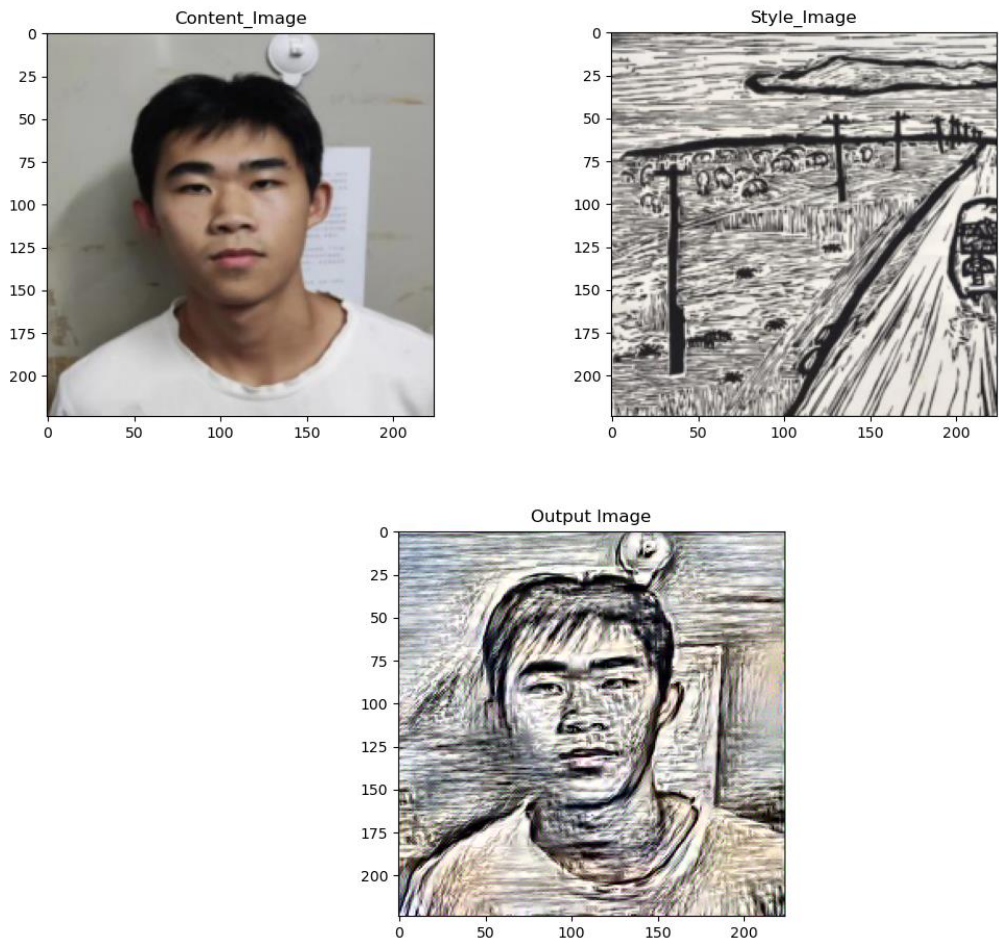


由于两张图片色调相差不大，因此迁移出来的风格并不明显，为了排除这方面的影响，我选择

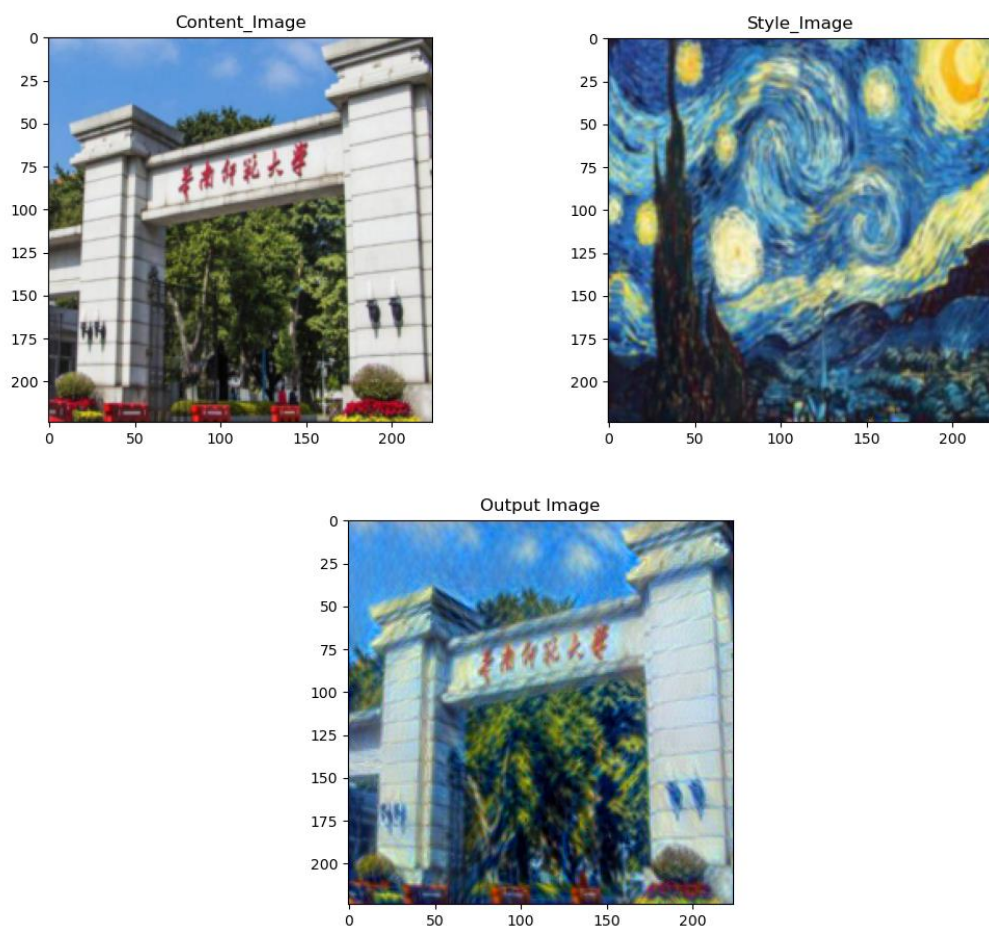
了另一种风格进行迁移，并且取得了不错的效果。



人像：



校门：



五、结论

图像风格迁移技术是一种基于计算机视觉和深度学习技术的艺术生成技术。针对不同的图像来源和风格参考，有多种算法可以实现风格迁移，包括基于特征匹配的方法、基于图像分解的方法、基于最小化风格差异和内容差异的方法、基于卷积神经网络的方法和基于生成对抗网络的方法。其中，基于深度学习的方法在实际应用中表现更好。

实验结果显示，图像风格迁移技术可以生成高质量的艺术图像。针对生成的图像质量和风格传递效果，评估时重要指标包括与参考图像的相似度、与源图像的分离度和处理时间等因素。虽然该技术在生成高质量艺术图片方面具有很大的潜力和发展空间，但也需要考虑图像内容的可解释性和保留问题，避免图像过度风格化或失真等问题。因此，在应用图像风格迁移技术时需要注意平衡图像内容和风格迁移的效果，并在保留图像潜在信息的基础上，提高艺术性和可理解度。

六、小组分工

组长	麦境轩	(1) 撰写开题报告、中期报告 (2) 制作汇报 PPT (3) 后期根据小组反思，共同提出问题，修正实验结果
组员	孙明辉	(1) 搭建代码运行环境，学习代码，训练模型 (2) 收集相关文献和参考资料，学习代码含义，逐行解读，写代码注释 (3) 对图像进行风格迁移，得到实验结果和输出图像 (4) 后期根据小组反思，共同提出问题，修正实验结果
组员	梁家灝	(1)撰写开题报告、中期报告 (2)制作汇报视频 (3)后期根据小组反思，共同提出问题，修正实验结果

七、参考文献

[1]Image Style Transfer Using Convolutional Neural Networks

[2]<https://arxiv.org/abs/1603.08155>

[3]<https://www.zhihu.com/question/49805962/answer/199427278>

[4]<https://arxiv.org/abs/1603.08155>

[5]<https://github.com/hzy46/fast-neural-style-tensorflow>

[6]<https://github.com/JaimeTang/PyTorch-and-Neural-style-transfer/blob/master/PyTorch-and-Neural-style-transfer.ipynb>