

Gatys-style和Lapstyle风格迁移的比较

作者：2100013132 张浩卓，2100013131 淦林川

引言

神经风格迁移是一种图像处理的技术，它可以将一幅图像的内容和另一幅图像的风格结合起来，创造出新的艺术图像。神经风格迁移的原理是利用卷积神经网络（CNN）提取图像的内容和风格特征，并通过优化算法最小化内容和风格之间的距离。神经风格迁移的一个经典方法是Gatys-style，它使用了VGG网络作为特征提取器，并使用L-BFGS作为优化器。Gatys-style方法可以产生高质量的风格迁移图像，但是也有一些缺点，比如计算量大，速度慢，容易产生失真和伪影等。为了解决这些问题，Lapstyle方法提出了一种基于拉普拉斯金字塔的神经风格迁移框架，它可以在不同的尺度上进行风格迁移。Lapstyle方法可以保留更多的内容细节，减少失真和伪影，同时提高了风格迁移的速度和效率。本任务的目的是比较Gatys-style和Lapstyle两种神经风格迁移的方法，并分析它们的优缺点和适用场景，并进一步测试更复杂的模型（如专门的生成式模型）完成此任务的效果。

问题描述

本任务关注的问题是比较Gatys-style和Lapstyle两种神经风格迁移的方法，具体来说，有以下几个方面：

- 数据集：使用自行收集的图片作为内容图像，使用WikiArt数据集作为风格图像。
- 预期结果：使用两种方法分别对同一对内容和风格图像进行风格迁移，并比较生成的风格迁移图像的质量和效果。
- 评估指标：使用以下几个指标来评估风格迁移的质量和效果：
 - 风格迁移度：使用风格图像和风格迁移图像之间的格拉姆矩阵（Gram matrix）的距离来衡量风格迁移的程度，距离越小，表示风格迁移越好。
 - 结构保留度：使用结构相似性指数（SSIM）来衡量风格迁移图像的结构保留的程度，SSIM越大，表示结构保留越好。
 - 主观评价：采用调查统计的方法让人类观察者对Style transfer的结果进行打分或排序，根据他们的偏好和感受来评价风格迁移的质量和多样性。

技术方法

本报告使用了两种神经风格迁移的方法，分别是Gatys-style和Lapstyle，下面分别介绍它们的技术细节。

Gatys-style方法

Gatys-style方法是一种基于CNN的神经风格迁移方法，它使用了VGG网络作为特征提取器，具体来说，它使用了VGG的前五个卷积块，分别记为`conv1_1`, `conv2_1`, `conv3_1`, `conv4_1`, `conv5_1`。Gatys-style方法的流程如下：

- 输入一幅内容图像 C 和一幅风格图像 S ，初始化一幅噪声图像 I 作为风格迁移图像。
- 将 C, S, I 分别输入VGG网络，得到它们在不同层的特征图。
- 定义内容损失函数 L_C ，它是 C 和 I 在 `conv4_2` 层特征图之间的MSE。

- 定义风格损失函数 L_S , 它是 S 和 I 在 conv1_1, conv2_1, conv3_1, conv4_1, conv5_1 层特征图的格拉姆矩阵之间的距离的加权和。
- 定义总损失函数 L , 它是 L_C 和 L_S 的加权和, 即 $L = \alpha L_C + \beta L_S$, 其中 α 和 β 是内容和风格的权重系数。
- 使用Adam优化器对 I 进行优化, 使得 L 达到最小值, 得到最终的风格迁移图像。

Lapstyle方法

Lapstyle方法是一种基于拉普拉斯金字塔的神经风格迁移方法, 它也使用了VGG网络作为特征提取器, 但是它在不同的尺度上进行风格迁移, 具体来说, 它使用了四个尺度, 分别记为coarse, medium, fine, finest。Lapstyle方法的流程如下:

- 输入一幅内容图像 C 和一幅风格图像 S , 对它们分别进行降采样, 得到四个尺度的图像, 记为 C_i 和 S_i , 其中 $i = 1, 2, 3, 4$ 分别对应 coarse, medium, fine, finest。
- 初始化四个尺度的风格迁移图像 I_i , 其中 I_1 是 C_1 加上随机噪声, I_2, I_3, I_4 是 C_2, C_3, C_4 加上双线性插值后的 I_1, I_2, I_3 。
- 对于每个尺度 i , 将 C_i, S_i, I_i 分别输入VGG网络, 得到它们在不同层的特征图。
- 对于每个尺度 i , 定义内容损失函数 L_{C_i} , 它是 C_i 和 I_i 在 conv4_2 层特征图之间的MSE。
- 对于每个尺度 i , 定义风格损失函数 L_{S_i} , 它是 S_i 和 I_i 在 conv1_1, conv2_1, conv3_1, conv4_1, conv5_1 层特征图的Gram矩阵之间的距离的加权和。
- 对于每个尺度 i , 定义结构损失函数 L_{T_i} , 它是 C_i 和 I_i 之间的SSIM。
- 对于每个尺度 i , 定义总损失函数 L_i , 它是 $L_{C_i}, L_{S_i}, L_{T_i}$ 的加权和, 即 $L_i = \alpha_i L_{C_i} + \beta_i L_{S_i} + \gamma_i L_{T_i}$, 其中 $\alpha_i, \beta_i, \gamma_i$ 是内容、风格和结构的权重系数。
- 对于每个尺度 i , 使用优化器对 I_i 进行优化, 使得 L_i 达到最小值, 得到最终的风格迁移图像。

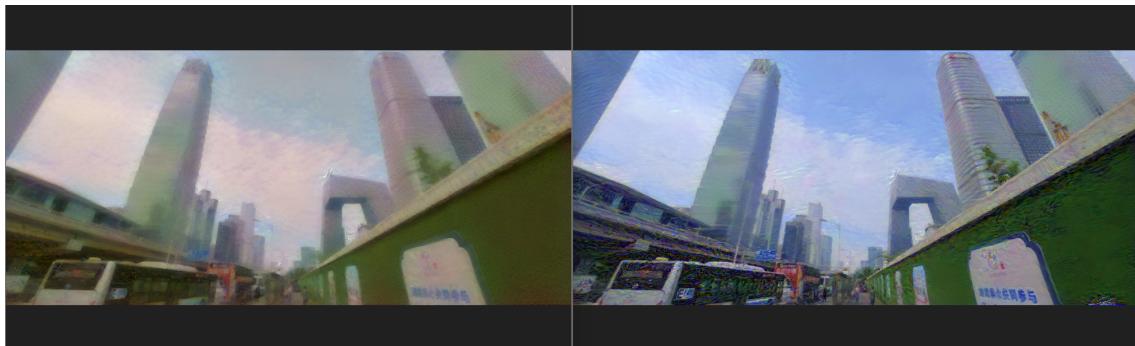
中期结果

目前我们实现了Gatys-style方法, 并进行了以下实验:

- 对于style transfer问题的提出: 给图片加“滤镜”的问题, 在生成图片初始化时将噪声改为content并调大 L_c 即可实现, 如:

原图:

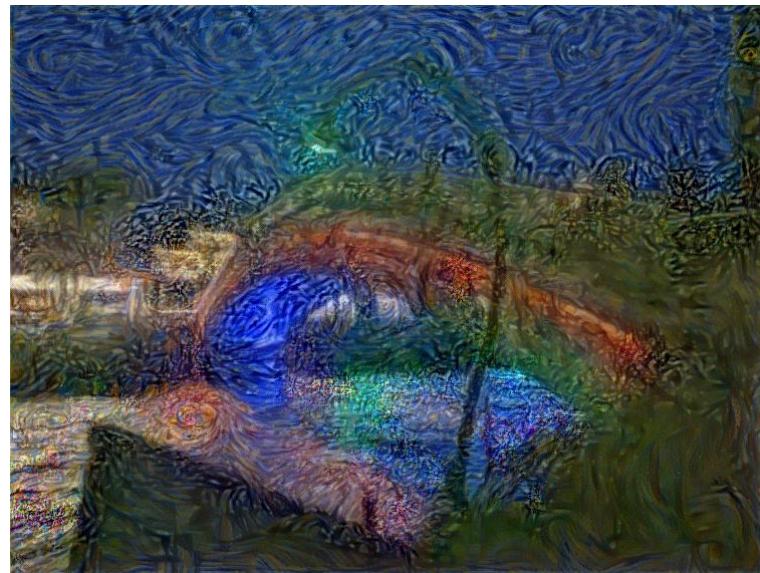




左图为原图 + Shipwreck of the Minotaur William Turner, 右图为原图 + Starry Night

但从图片中可以明显看到类似“油膜”的效果，主观评价细节实现较差。

- 同一个style image，在一定范围内，分辨率大的style image效果不如分辨率小的。
- 使用不同图片时，我们发现有的测试样本会出现严重的局部扭曲与鬼影，这是Gatys-style的不足之一。



未来计划

- 进一步优化Gatys-style
- 实现Lapstyle的风格迁移，在coco数据集上进行客观指标评价和主观评价并进行进一步分析
- 阅读更多采用CNN方法进行图像风格迁移优化的论文，并进行测试与比较

- 探索专门化的生成式模型在风格迁移方面的实现，并将其结果与Gatys-style和Lapstyle进行分析比较