

计算机科学与技术学院神经网络与深度学习课程实验报告

实验题目: style transfer		学号: 201600181073
日期: 2019. 4. 10	班级: 智能 16	姓名: 唐超
实验目的: <ul style="list-style-type: none">• 学习通过 VGG 提取图像特征;• 明白 neural style transfer 的基本原理以及流程。		
实验软件和硬件环境: Spyder&python3.6		
实验原理和方法: <p>风格迁移的工作是将原图的上下文内容与参考图的风格进行融合, 这种融合使得输出的图片在内容上接近 <code>content_image</code>, 在风格上接近 <code>style_image</code>。要完成以上两点, 需要定义输出图在内容上和风格上与输入图的 Loss。</p> <h3>1. content loss</h3> <p>内容差距我们用两张图片像素点的差的平方来衡量:</p> $L_{content} = mean((F_{INPUT} - F_{OUTPUT})^2)$ <h3>2. style loss</h3> <p>风格差距通过 Gram 矩阵来定义:</p> $G_{ij} = \sum_K^{C \times C} F_{ik} F_{jk}$ <p>Gram 矩阵就是在这个特征图上面定义出来的。每个特征图的大小一般是 $M \times N \times C$ 或者是 $C \times M \times N$ 这种大小, 这里 C 表示的时候厚度。Gram 矩阵的大小是有特征图的厚度决定的, 等于 $C \times C$。把特征图中第 i 层和第 j 层取出来, 这样就得到了两个 $M \times N$ 的矩阵, 然后将这两个矩阵对应元素相乘然后求和就得到了 $Gram(i, j)$, 同理 Gram 的所有元素都可以通过这个方式得到。这样 Gram 中每个元素都可以表示两层特征图的一种组合, 就可以定义为它的风格。</p> <p>Style loss 计算方式与 content loss 基本相同:</p> $L_{STYLE} = mean((G - F)^2)$		

实验步骤：（不要求罗列完整源代码）

1. 将 content.jpg 和 style.jpg resize 成同样大小、转化为 4D 的 tensor、做归一化处理；
2. 将上一步 content 的处理结果复制一份作为 target，并将 content、style、target 作为 vgg19 的输入，并分别提取各自第 1、6、11、20、29 层卷积特征，接着按实验原理中的公式计算 content_loss 和 style_loss；
3. 对 loss 分别反向求导并调整 target，每 100 个 step 输出一张合成图片。

结论分析与体会：

Style transfer 的结果如下：



figure 1 content



figure 2 style

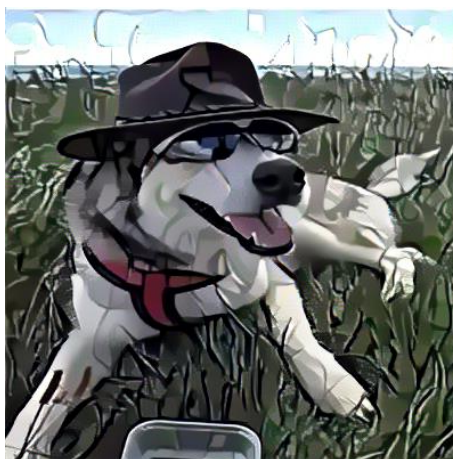


figure 3 target(1400step)

体会：通过这次实验，学习了用 vgg 网络提取图像特征的方法和 style transfer 的原理，加深了对 vgg 网络的认识。

就实验过程中遇到和出现的问题，你是如何解决和处理的，自拟 1—3 道问答题：

1. 开始时将 feature map 错误理解为图像的 rgb 通道，没有明白 style loss 的计算方法，后来再详细看 vgg19 的网络结构认识到 feature map 是卷积层的“厚度”；
2. Image 需要进行预处理，即 resize、转化为 tensor 并做归一化。