

Image Style Transfer Using Convolutional Neural Networks

此论文发表前风格迁移工作发展情况：基本没发展，比较接近的是纹理迁移，但纹理迁移保留内容却无法提取语义内容（以前提取到的语义也只局限到字体、数字等），导致无法对图像大改，达到风格迁移的程度

此论文的工作：

1. 使用vgg19的卷积层提取图像的信息
2. 对各层卷积的结果进行重建，发现浅层主要记录细节，深层损失细节记录内容
3. 构建内容损失的计算方法（其实就是用对应的层卷积结果计算各个位置平方差的和）并依据这个方法计算出损失的导数以优化生成的图像中每个像素，使内容不再产生变化（导数=F-P, F, P分别是生成图与原图被提取出的特征图上的点的值，且此时F为正，若F为负，导数为0）□

$$\frac{\partial \mathcal{L}_{\text{content}}}{\partial F_{ij}^l} = \begin{cases} (F^l - P^l)_{ij} & \text{if } F_{ij}^l > 0 \\ 0 & \text{if } F_{ij}^l < 0, \end{cases} \quad (2)$$

4. 构建风格的表示与损失计算方法：多个卷积核卷积会得到多张特征图，我们可以使用这些特征图计算一个GRAM矩阵来表示图像的风格（相当于把多个向量集合到了一起，对角线上的值表示特征图强度，其他点表示不同特征图间的相关性，我们就可以知道这个图在每一种卷积核下强度如何，其空间排布和其他特征的排布间的关系，可以很综合的表现一个图的特征），并定义了风格损失——不同GRAM矩阵间各个位置平方差的和再求和（要进行归一化的，也就是每个GRAM矩阵间的差值还要除以滤波器数量的平方 N^2 与图大小的平方 M^2 ），特征图中每个点的

导数如下□

$$\frac{\partial E_l}{\partial F_{ij}^l} = \begin{cases} \frac{1}{N_l^2 M_l^2} ((F^l)^T (G^l - A^l))_{ji} & \text{if } F_{ij}^l > 0 \\ 0 & \text{if } F_{ij}^l < 0. \end{cases} \quad (6)$$

5. 于是作者结合这两种损失作为整体损失，使用优化算法对随机噪声（如果不随机的话会得到相同的结果，无法得到多张风格迁移的图片）进行最优化以进行风格迁移，并对两种损失之间的权重进行了探究，此外作者还发现浅层卷积核会更多的注重图像细节，使得风格迁移结果更像是纹理附在照片上，不够自然。