《大数据与机器智能》大作业报告

第16组

自91 王文新 2019011354

选题及理由

我组选择了“风格迁移”作为本次《大数据与机器智能》课程的大作业课题。

风格迁移是一种基于深度学习的图像处理技术，是深度学习领域的一颗新星。利用这一技术可以完成提取某幅图像的风格，并利用这一风格重绘其他图像的任务。

这一技术有着很高的应用价值。从娱乐角度，这一技术给了人们一个将自己的摄影绘画作品处理成具有大师风格的艺术品的机会，如不久前大火的“P图软件”Prisma，本质就是通过风格迁移技术进行图像处理的，较高的效率与优秀的模型使得这一软件即使在手机上运行也能具有十分优异的照片处理表现，进而斩获人气。而研究角度上，风格迁移技术中伴随出现的诸多精妙的图像处理技术，如Gram矩阵、图像编码&解码网络等，对于图像处理领域的研究来说，有着很大的借鉴意义。从应用角度，风格迁移技术也不只局限于模仿画作，例如，可以通过学习多种夜景、日景等不同光照条件下拍摄的图片的风格，将一张人像进行风格迁移，进而有效地扩大人脸识别的训练数据集，提高人脸识别网络的鲁棒性。

实验过程

出于Tensorflow lite并未支持风格迁移的实例，我组从第三方搜集了介绍文案、博客、论文、代码等资源对两世代风格迁移技术的原理分别进行了系统的学习，并对两代风格迁移技术分别进行了实例尝试，进行了问题的总结和两代技术的对比。最终出于风格迁移这一项目具有静态的特殊性，我组选择了制作视频进行更好的效果展示。

两代网络结构及比较

第一代风格迁移技术中，仅使用了VGG-16卷积网络，用于图像特征的提取并利用特征进行比较与反向传播。抽取风格图像的从低到高多层特征图计算Gram矩阵，与输出图像的同样运算所得的Gram矩阵进行比较，得到风格损失函数；抽取内容图像的高层特征图，与输出图像的高层特征图计算L2距离得到内容损失函数，二者的线性加权求和作为总损失函数用于反向传播。

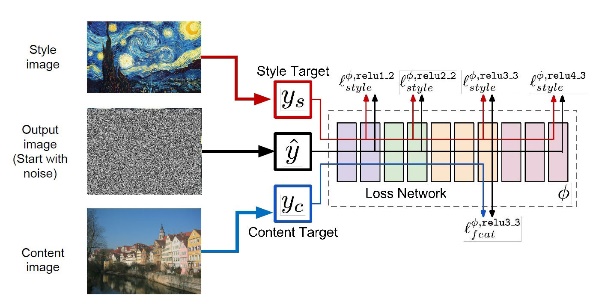


图1第一代风格迁移网络结构示意图

由于基于反向传播进行迁移，且每一次反向传播都将调节输出图像的每一个像素，这一网络的效率较为低下，在有限的时间内收敛状况不佳，结果中常出现不自然的色块。

第二代风格迁移技术中，通过训练编码-残差-解码网络，使之具备对任何内容图像进行风格迁移的能力，进而达到了快速风格迁移的水平。对该网络的训练仍为利用VGG-16网络，以同样的损失函数计算方式进行。这一代迁移技术中，训练可能耗费大量时间，但在训练数据集（即大量不同的内容图像）足量，训练所得模型较好的情况下，只需要编解码网络本身就可以前馈进行任意图像的风格迁移，速度与图像效果都相当可观。（效果展示可见../PPT中展示PPT或../VIDEO\_PRE中视频展示）。

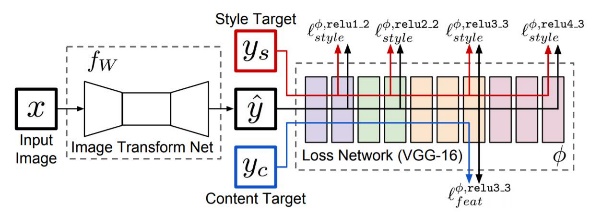


图2第二代风格迁移网络结构示意图

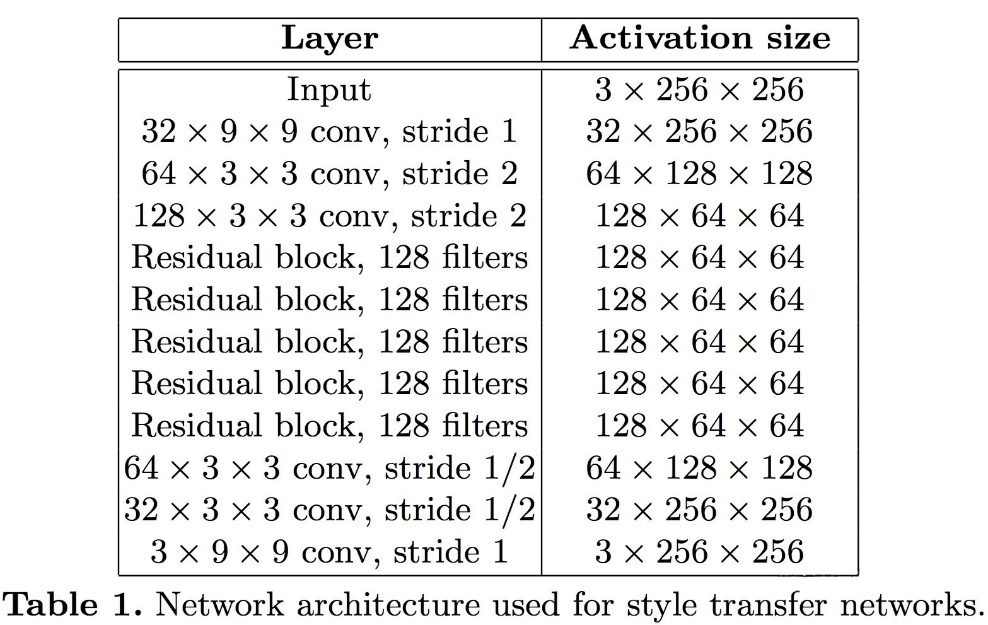


图3 Image Transform Net编解码网络结构示意图

总结

本次大作业的完成过程让我收获良多，我更加深入地理解了深度学习技术，更是在风格迁移方面获取了不少知识，在研究和实践的过程中，欣赏艺术之美的同时也感到了收获的喜悦。