

图片风格融合及快速迁移

陈 良

(江西理工大学信息工程学院, 江西 赣州 341000)

✉m13125397685@126.com



摘 要: 通过卷积神经网络对不同图片的内容和风格进行融合, 可生成风格多样化的图片。这不但为影视制作提供了丰富的素材, 更有利于图像修复和图像增强。针对这类问题, 前人曾提出一些算法, 但很难在时间和空间方面都达到很好的效果。这里提出一种基于TensorFlow(将复杂的数据结构传输至人工智能神经网络中进行分析 and 处理过程的系统)的条件归一化网络来支持多风格融合及图片快速迁移, 多风格可共用一个网络模型, 这大大减少了算法耗时, 并缓解了模型存储耗费空间大的问题, 节省了计算机资源。时间上优于传统迁移算法三个数量级, 空间上25种风格可共用一个模型。同时, 更大程度地保留了内容图的语义特征, 具有更好的视觉效果。

关键词: 条件归一化网络; 风格融合; 快速迁移; 共享模型

中图分类号: TP312 **文献标识码:** A

Picture Style Fusion and Rapid Migration

CHEN Liang

(School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China)

✉m13125397685@126.com

Abstract: Content and style of different pictures can be merged by convolutional neural network to generate pictures with diversified styles. This not only provides rich materials for film and television production, but also facilitates image restoration and enhancement. Aiming at this kind of problem, previous researchers have proposed some algorithms, which turned out to be difficult to achieve good results in both time and space. This paper proposes a conditional normalization network based on TensorFlow (a system that transmits complex data structures to artificial intelligence neural networks for analysis and processing) to support multi-style fusion and rapid image migration. Multiple styles share a network model, which greatly reduces time-consuming algorithm and relieves large space consumption for model storage, saving computer resources. It is three orders of magnitude better in time than traditional migration algorithms, and 25 styles in space share one model. At the same time, semantic features of the content map are retained to a greater extent, and the visual effect is better.

Keywords: conditional normalization network; style fusion; rapid migration; sharing model

1 引言(Introduction)

图像风格迁移不仅能改善电影动画制作效果, 降低制作成本, 更重要的是它有利于图像修复。图像风格迁移算法最初从研究纹理合成技术开始, 大多数纹理合成算法采用非参数方法^[1], 并没有提取图像的高级特征, 对纹理较为复杂的图像很难做到理想的风格迁移。Gatys^[2]等人提出了一种基于CNN的传统迁移算法, 增强了迁移效果, 但耗时很大。Johnson^[3]等人提出快速迁移算法即训练一种前馈生成网络, 生成速度提高了三个数量级, 但仍然无法同时获取多种风格。

本文提出一种条件归一化网络算法, 对每种风格归一化后再对其风格参数进行特定的调节, 令最终模型具有多风格

同步迁移且支持风格融合的能力。这样多风格可以共享一个模型, 不仅缩短了风格融合时间, 还大大缓解了模型存储的空间压力, 为后期部署节约了很大的成本。

2 算法设计(Algorithm design)

2.1 构思背景

考虑到快速风格迁移算法虽然加快了图像迁移的速度, 但每次迁移时的风格图像是固定的, 也就是每当更换风格时都需要加载另一个模型, 这不利于多风格迁移且多个模型很不利于存储, 更会浪费计算机资源。更重要的是, 这为后期模型部署带来了很大的压力, 降低了图片风格迁移的应用价值。于是设计一种网络可以满足对多个风格实现同步迁移,

并最终利用该网络训练得到针对多风格的单一模型就显得尤为重要。

显然,这里的训练数据和测试数据属于同一集合,因此不会存在过拟合问题。进行图像风格迁移时,不同风格图片的参数计算有一定的规律,那么不同的风格图片与同一内容图片进行风格迁移时就会存在一种共享模式,因此在实际的多风格迁移时多风格可以共享模型的参数和结构,这会大大减少模型占用空间。凭直觉来讲,不同的绘画有着相似的笔触,这代表着相似的视觉元素,但调色、用笔力度等方面存在差异,那么这多幅作品其实存在很多共通点。显然,在进行N风格迁移时仍然构建N个网络,这种共享模式就失去了意义,也很不利于模型的训练和存储。从艺术的角度来讲,绘画风格的多样性代表了一种丰富的视觉信息,有利于增强人们对绘画更高层次特征的理解。捕获到不同的图像风格,训练得到的模型允许任意方式组合的全新风格来探索新的艺术风格,为学习艺术风格的特征提供了一条独特的途径。

2.2 网络架构

条件归一化网络架构从输入层到输出层依次为三层卷积层、五层残差块和两层上采样,最后一层为卷积层,如图1所示;再结合损失函数网络便构成了条件归一化网络流程的整体网络架构(图2)。与快速风格迁移网络架构的主要区别为,条件归一化网络使用了条件归一化操作,并且该操作贯穿了整个网络的每一层,也就是图像的多风格迁移和风格融合是同时进行的。



图1 条件归一化网络

Fig.1 Conditional normalization network

图1的卷积层包含三种操作:卷积、归一化、relu。残差块使用的是Gross and Wilber的残差网络,每个残差块共五层结构,包含两个 3×3 的卷积层,每个卷积层有128个滤波器。残差网络的核心操作为跨卷积层做加法操作,目的为避免梯度消失。与其等效的卷积块相比,该残差块最后缺少一个relu非线性,在分类方面有更好的性能,同时加快了损失函数收敛的速度,更易于训练深层的网络^[4]。上采样同样包含三种操作:双线性插值、卷积和归一化。最后一层为卷积层,与开始卷积层不同的是,这里把relu改为sigmoid。

2.3 条件归一化

条件归一化可用来学习多种图像样式,该操作将单风格的图像转换网络转变为多风格的转换网络。多风格图像转换网络的卷积权重在不同风格间是共享的,归一化后对其对应的风格参数进行调节^[5]。具体做法就是先将输入图像映射到正态分布,再对每个固定的风格在归一化之后进行特定的缩放

和平移,所以这个网络中增加了W和B这两个参数,一个负责缩放,另一个负责平移。针对多风格的模型,只需要训练一个网络,网络整体也只是增加了W和B这两个参数,实际增加量为: $2 \times \text{风格数目} \times \text{特征映射数目}$ 。在典型的网络设置中,只有0.2%的参数代表风格的独立特征。又因为条件归一化只对缩放和移动参数起作用,所以在多个风格上训练图像转换网络所需的参数比训练多个单独网络的方法需要的参数少得多,这大大减少了模型训练需要的时间。

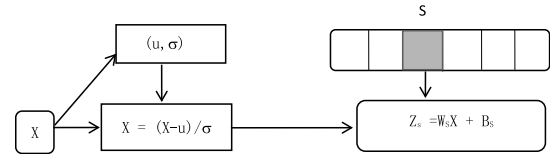


图2 条件归一化流程图

Fig.2 Conditional normalization flow chart

条件归一化把每种风格压缩为嵌入空间的一个点,多个不同点构成了最终的模型。该模型可以允许一些全新的方式组合风格,能更快地捕获新的艺术风格。

2.4 改进图像生成质量

棋盘现象:在快速迁移算法上采样过程中使用了转置卷积,这会导致生成图像出现一些类似棋盘状的伪影,尤其在颜色较深的区域,严重影响了图像的生成质量。本文采用双线性插值算法,避免了棋盘现象的产生。

伪影现象:在快速风格迁移算法中残差块的卷积层没有使用填充操作,但一般卷积层会采用零填充。零填充虽然可以控制特征图的大小,但会造成边界伪影现象,严重影响了生成图像的质量。据了解,镜像填充可以缓解这种现象,但设置较为复杂,同时会增加训练时间,不利于多风格迁移模型的训练。那么这里仅是修改卷积时的边缘填充方式,使用VALID方式进行卷积,避免了生成图像边缘发黑,因此这与镜像填充方式的作用一致,较大程度地增强了风格迁移生成图像的质量。

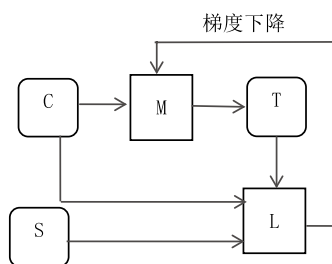
降低内容损失层:大部分的风格迁移算法都采用conv_3作为内容损失层。Gatys提出层数越低越能很好地重建内容特征,但conv_1层的内容损失很小,变化幅度相对更小不利于损失函数收敛。实验中发现采用conv_2作为内容损失层能加快收敛速度,增强生成图像质量,更适合多风格迁移。

sigmoid函数:实验中发现当最后一层卷积层采用relu时图像有黑块现象,改为sigmoid这种现象几乎消失,同时有利于将结果控制在(0,255)内。

2.5 多风格融合及快速迁移

条件归一化网络(图1)生成目标图像,再将生成图像及内容和风格图像一并传入损失函数网络计算得到总损失,损失函数值按公式(1)计算。最后通过Adam优化器最小化损失值,逐步优化模型参数,如图3所示。

$$L_u(s, t, c) = \alpha L_{con}(c, t) + \beta L_{sty}(s, t) \quad (1)$$



S: 风格图片; C: 内容图片;

T: 生成图片; L: 损失函数网络; M: 条件归一化网络

图3 条件归一化算法流程图

Fig.3 Conditional normalization algorithm flowchart

3 实验(Experiment)

3.1 实验数据和环境

本次风格迁移实验需要的数据集分为内容图片和风格图片。内容图片为MSCOCO官网的train 2014中的82783张训练图片,数据集大小为12.5G。另外需要20张风格鲜明的风格图片用于训练。

环境搭建: 由于数据集很大, CPU训练非常耗时且容易断电或网络中断影响模型训练, 因此这里采用GPU服务器进行模型训练, GPU型号为NVIDIA GeForce RTX 2080 Ti。编写语言为python, 采用TensorFlow框架, 在jupyter notebook和pycharm上运行代码。

制作tfrecord训练集: 对数据进行统一的管理是很有必要的。tfrecord文件^[6]就是对输入数据做统一管理的格式, 加上一些多线程的处理方式, 使得在训练期间对于数据管理把控的效率和灵活度都比普通暴力的方法好得多。再加之实验需求的数据量很大, 使用统一格式管理显得尤为重要。

3.2 模型训练

读取制作好的tfrecord训练集, 依照图3所示训练得到最终模型。模型训练时的风格矩阵用于提取当前风格的参数W和B。训练中提供了断点续训功能, 可及时查看当前训练结果, 保存最优模型。实验开启了多线程训练, 进一步减少了模型训练时间, 为节省磁盘空间只保存最后一代模型。

3.3 实验结果

3.3.1 图像多风格迁移

选择多张不同类别的图像为风格图像, 一张儿童图片为内容图像, 采用训练好的模型进行多风格迁移。实验结果证明, 该算法确实可以很好地每种风格特征迁移到如图4所示的内容图, 生成图如图5和图6所示。



图4 内容图

Fig.4 Content image



图5 生成图1

Fig.5 The first generated image



图6 生成图2

Fig.6 The second generated image

3.3.2 图像多风格融合

条件归一化网络算法能提取多张图像的风格特征, 并将它们按照一定比例融合。实验采用了一张灰度图像作为内容图像, 如图7所示。四张不同风格的图像令原本单调的灰度图像充满生机, 验证了条件归一化网络算法具有很好的多风格融合能力, 如图8所示。



图7 灰度图像

Fig.7 Grayscale image



图8 风格融合图像

Fig.8 Style fusion image

4 算法对比(Algorithm comparison)

4.1 普通图像迁移效果

这里采用相同的内容和风格图像, 分别利用Gatys提出的传统迁移算法、Johnson提出的快速迁移算法和本文的条件归一化网络算法生成三张迁移后的效果图像, 直观地对比三种算法对普通图像的风格迁移效果, 如图9至图11所示。

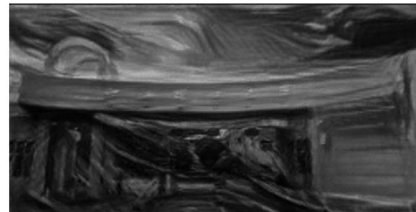


图9 传统迁移算法生成图

Fig.9 Traditional migration algorithm generation diagram



图10 快速迁移算法生成图

Fig.10 Quick migration algorithm generation diagram



图11 条件归一化算法生成图

Fig.11 Conditional normalization algorithm generation diagram

采用条件归一化网络算法生成的迁移图与其他算法的生成图在整体感官上很接近,但经过网络模型得到的图像更易捕获内容信息,如图10和图11能够保留原内容图中的文字信息,图11的文字更清晰,而图9的文字几乎被图像色彩覆盖,很难辨认。

4.2 灰度图像迁移效果

4.1节表明了条件归一化网络算法对普通图像的迁移效果要优于传统迁移算法和快速迁移算法,图12为灰度图,图13为风格图,图14和图15对比了灰度图像迁移效果。



图12 灰度图

Fig.12 Grayscale image



图13 风格图

Fig.13 Style image

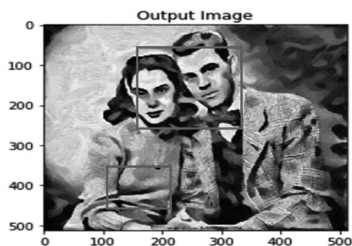


图14 传统迁移算法效果图

Fig.14 Effect diagram of traditional migration algorithm



图15 条件归一化算法效果图

Fig.15 Effect diagram of conditional normalization algorithm

实验采用了一张年代较远的灰度图像和一张色彩很浓的风格图像,对传统迁移算法和条件归一化网络算法进行了迁移效果对比。传统风格迁移算法虽然能对普通图像进行较好的风格迁移,但这种方式对灰度图像不是很理想,存在图像失真、风格渲染不均匀等现象。对比图14和图15很容易发现:图14中人物鼻部和眼部特征被光斑遮盖,甚至衣服的褶皱度也被变小。而图15几乎融合了图12的内容信息,例如,女性右侧头发的卷曲度和脸部距离与图12相比,在肉眼观察的范围内是没有差距的,这主要因为条件网络算法是通过卷积神经网络生成图像的,而在卷积池化的过程中,网络就很好地学习了图像中的鲁棒特征^[7]。当然,可采用语义分割的方式来改善图像失真问题。更为重要的是,人的脸部特征并没有因为融合风格图像而大幅度地失去,也就是在保持风格化的前提下,尽量降低内容损失。

总之,条件归一化网络下的风格迁移比传统迁移算法更大程度地减少了损失函数值,同时生成图15的速度比生成图14快了约300倍。

4.3 图像裁剪

实际应用中往往提供的风格和内容图像大小并不一致,但传统迁移算法和快速迁移算法都没有给出一种合理的裁剪方案,故迁移效果不理想。本文的条件归一化网络算法提出了一种中心裁剪的方式,即以图像中心为起点裁剪出一个正方形,该方法可调整输入图像使之大小一致,并较好地保留图像内容信息。图16对比了两种不同算法下的裁剪效果。



图16 算法裁剪效果对比图

Fig.16 Comparison of algorithm cropping effect

图16右上角为条件归一化网络算法下的生成图像,左上角为传统迁移算法生成图像。值得肯定的是,中心裁剪保留了更多的内容图特征,如条件归一化网络算法将舞者的形体美更好地展示出来。

4.4 算法耗时和空间

快速迁移算法在时间上至少快了几百倍,这一点是很有价值的。但条件归一化网络不仅加快了速度,还有效缓解了模型存储的空间压力,表1对比了三种算法在时间和模型空间

储存上的差异。

表1 算法时间、空间对比表

Tab.1 Comparison of algorithm time and space

算法类别	时间1	空间1	时间2	空间2
传统迁移算法	513 s	无	14 560 s	无
快速迁移算法	1.04 s	18 M	26.34 s	447 M
条件归一化网络算法	1.64 s	20 M	23.56 s	20 M

为了有效对比不同算法耗时，表1的时间都是在CPU服务器上的运行时间。时间1、时间2分别指生成单张图像和25张不同图像所需时间。空间1、空间2分别指生成单张图像和25张图像所需模型占用空间的大小。由表1可知：时间层次上，条件归一化网络算法所需时间与快速迁移算法差不多，都优于传统迁移算法三个数量级，且迁移的图像数目越多，这种优势越明显。空间层次上，传统迁移算法是直接梯度下降生成目标图像，故不需要模型；而快速迁移算法和本文的条件归一化网络算法都需要模型，生成一张图像时两者几乎没有差距。但条件归一化网络是支持多风格同步迁移的，只要没超过训练时的风格图像数目，模型所占空间不变，极有利于后期开发；而快速迁移算法模型占用空间会随着模型数量而线性增加，这显然对模型的移动端部署是很不利的。总之，本文提出的条件归一化网络一方面能够像快速迁移算法一样，大大缩短图像迁移所需的时间；另一方面由于模型共享，很大程度地缓解了模型存储的空间压力。

5 结论(Conclusion)

本文设计的条件归一化网络算法不仅加快了图像生成

(上接第41页)

范的正则表达式，就可以提取出环境目标的变量名。找到变量名之后，可以建立一个二维列表样式的数据结构，每一行是[采样时刻，目标1-x坐标，目标1-y坐标，目标2-x坐标，目标2-y坐标，...]，这样就可以获得所需要的数据，解决了难点(1)的问题。针对难点(2)，需要设计一个简单的播放器状态机，为了实现播放进程和单帧进行的功能，可以设计一个滑动条和按钮进行控制。利用上述方法导入某MF4文件后的三维动态显示结果如图9所示，当目标动态运动时，使用前文中方法显示目标的一些动态属性，如瞬时速度、加速度等，并在右侧使用表格和树状结构显示了一些目标的属性值。

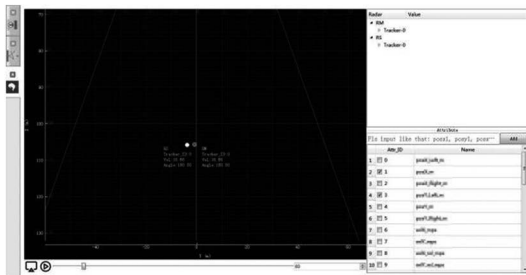


图9 三维动态显示

Fig.9 Dynamic display in three-dim

5 结论(Conclusion)

本文介绍了一种使用Qt可视化高级驾驶辅助系统测试数据的方法。本方法简单、快速，使用开源工具包，而且可以针对ADAS测试过程中的需求进行定制。经某公司验证，该软

件，更大大缓解了模型存储的空间压力，节省了计算机资源，为后期的部署应用节约了成本。更重要的是，该算法的迁移效果在很多方面要优于其他算法，且能实现风格融合。

参考文献(References)

- [1] 陈淑环,韦玉科,徐乐.基于深度学习的图像风格迁移研究综述[J].计算机应用研究,2019,36(08):2250-2255.
- [2] Leon A Gatys, Alexander SEcker, Matthias Bethge. A neural algorithm of artistic style[J]. arXiv preprint arXiv, 2015, 1508(06276):4-12.
- [3] Justin Johnson, Alexandre Alahi, Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution[J]. arXiv preprint ar Xiv, 2016, 1603(08155):1-7.
- [4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, et al. Deep residual learning for image recognition[J]. arXiv preprint arXiv, 2015, 1512(03385):2-8.
- [5] Vincent Dumoulin,Jonathon Shlens,Manjunath.Kudlur. A learned representation for artistic style[J]. In ICLR, 2017, 1610(07629):4-8.
- [6] 郑泽宇,梁博文,顾思宇.TensorFlow:实战Google深度学习框架[M].北京:电子工业出版社,2017:170-183.
- [7] Andrew Ilyas, Shibni Santurkar, Dimitris Tsipras, et al. A adversarial examples Are Not Bugs, They are Features[J]. arXiv preprint arXiv, 2019, 1905(02175):6-7.

作者简介:

陈 良(1995-),男,本科生.研究领域:深度学习,模型压缩.

件简单、实用，运行稳定，为ADAS控制算法的开发、验证工作提供了保障。

参考文献(References)

- [1] 任银行,张建龙,殷承良.基于XCP协议支持多总线的ECU标定系统的实现[J].电子技术应用,2018,44(5):72-76.
- [2] 宋维群,杨世春,李明,等.基于ASAM标准的标定数据库逻辑分析与编程实现[J].汽车技术,2012(1):10-13.
- [3] 董新宇,张春,姜承赋,等.MDF存储格式在柴油机匹配标定软件中的应用[J].现代车用动力,2011(2):29-31.
- [4] PyQtGraph官方文档及API[EB/OL]. <https://pyqtgraph.readthedocs.io/en/latest>,2020-6-10.
- [5] ASAM. MDF标准官方维基百科[EB/OL]. <https://www.asam.net/standards/detail/mdf/wiki>,2020-6-10.
- [6] Mdfreader-基于Python的MDF文件解析工具[EB/OL]. <https://github.com/ratal/mdfreader.git>.
- [7] 王硕,孙洋洋.PyQt5快速开发与实战[M].北京:电子工业出版社,2017.
- [8] 陆文周.Qt5开发及实例(第3版)[M].北京:电子工业出版社,2017.

作者简介:

杨军典(1990-),男,硕士生.研究领域:ADAS感知系统及控制策略算法.

陈凌珊(1966-),女,博士,教授.研究领域:新能源汽车技术,智能驾驶控制技术.