浙江理工大学

硕士学位论文开题报告

**题目： 基于深度学习的艺术图像分类算法研究**

学 院： 信息学院

学科专业： 软件工程

研究方向： 软件服务工程

姓 名： 杨秀芹

指导教师： 张华熊

学 号： 201720603005

入学年月： 2017年9月

选题来源： 自选课题

**报告日期 2018 年 11 月 14 日**

**目录**

[1学位论文开题的目的和意义 1](#_Toc529862098)

[2目前国内外的研究状况 1](#_Toc529862099)

[2.1国内外艺术图像分类研究情况 1](#_Toc529862100)

[2.2 卷积神经网络发展 3](#_Toc529862101)

[3学位论文研究的主要内容 4](#_Toc529862102)

[4学位论文研究的方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明） 5](#_Toc529862103)

[4.1 艺术图像数据采集 5](#_Toc529862104)

[4.2 图像预处理 5](#_Toc529862105)

[4.3 构建适用卷积神经网络 6](#_Toc529862106)

[5 本文的创新点和关键问题 6](#_Toc529862107)

[5.1本文的创新点 6](#_Toc529862108)

[5.2本文的关键问题 6](#_Toc529862109)

[6学位论文研究的工作条件 7](#_Toc529862110)

[7学位论文研究的进度计划安排 7](#_Toc529862111)

[参考文献 8](#_Toc529862112)

# 

# 1学位论文开题的目的和意义

随着数字媒体技术和互联网技术的发展，基于油画、国画等艺术作品的数字化及基于网络的分享方式正在成为主要的趋势。随着人们对艺术欣赏水平的提升，越来越多的人投入艺术绘画创作中。但是由于现有的图像资源库没有对艺术图像资源进行有效的管理和分类，使得用户在寻找资源时出现费时费力却又没有找到所需资源的情况。面对海量的数字化艺术图像，如何实现对其有效的分类与检索成为了亟待解决的问题。

随着计算机技术以及人工智能的发展，学者们提出了很多算法来对普通图像内容进行分类检索。现有的对图像进行分类的方法大都是对拍摄类图像进行分类，例如将拍摄类中属于动物类、人类、植物类、建筑类、物品类、场景类的图像进行分类。而针对艺术图像风格类型进行检索分类的实验并不是很多，使得用户想根据艺术图像性质分类检索自己所需的艺术图像，而很难在海量的图像资源库中找到自己所需的图像。

对具有绘画性质的艺术图像进行分类，以满足人们对特定艺术图像风格的检索与查询。对艺术图像资源进行内容层次的分类，不仅将人力从资源建设中解放出来，而且避免了资源的重复建设，保证了资源属性标注的客观性，有利于资源的跨平台、跨地区共享。通过对艺术图像的分类，帮助人们快速检索自己所需的艺术图像，有助于节省检索者的时间，快速而高效。艺术类图像的自动分类是现代数字图书馆、博物馆和美术馆的一项关键技术，对艺术品展览、学者进行历史研究以及用于教育和艺术图像爱好者学习欣赏具有很重要的作用。

# 2目前国内外的研究状况

## 2.1**国内外艺术图像分类研究情况**

图像分类不仅可以应用在图像分类管理和信息提取方面，还应用于手写字识别、人脸识别[1]、图像检索[2-4]、图像超分辨率应用[5]等方面。早期的对图像特征进行提取的方法主要是关注于图像的颜色、纹理和形状等特征。近年来，大量的工作用于寻找语义丰富的图像代表，其中效果很好的方法有局部聚合向量VLAD[6]、Bag-of-Words[7]、FV[8]等。但是基于这些方法的图像分类主要采用人工手动进行文本标注的方法，使用人们生活中关键词进行类别描述。但是这种文本分类的方法有很多不足：(1)图像分类的标注需要大量人工处理，而且很大程度上依赖于人的主观判断，这使得用户对图像分类的结果评判不一；(2)随着现在多媒体技术和互联网技术的快速发展，和人们对艺术欣赏水平的提升，大批的拍摄图像和艺术类绘制图像也运用而生。用关键字对海量图像进行描述，不仅异常复杂，而且不能充分描述图像视觉内容信息，依靠人工标注，效率低，而且跟不上海量图像产生的速度。

但是随着机器学习的发展，越来越多的人发现卷积神经网络对图像特征的提取比传统对图像特征进行提取的方法具有很高的准确率和效率，大量卷积神经网络已经用于对大规模图像的分类和检索中。

对大规模图像进行分类的研究中，大量的实验研究证明其方法的有效性、效率和内存使用率[9-12]。在对油画、国画的分类上，研究人员也提出了大量的方法对艺术画所属种类或所属画家进行分类。对于某种类型的绘画，根据图像中提取出的笔画信息，以区分不同风格的画家，2004年，Li等人[13]提出了用二维隐马尔可夫模型对图像中的大部分区域分析，并捕获关键区域笔画的特征，为艺术家建立的二维MHMMs的混合模型可进一步用于对图像进行分类和比较。2006年，Jiang[14]等利用图像低层的纹理、形状、颜色、边缘特征，将中国传统画分为工笔画和写意画，Shen[15]则将图像分成4×4的子块，对各个子块分别提取颜色、纹理、形状、颜色布局图像底层特征信息，并将提取出的特征输入RBF神经网络进行训练，然后对RBF神经网络输出的向量进行汉明距离计算以实现对输入的图像所属的艺术家进行分类。2009年，Yao等人[16]提出根据Morrisseau艺术图像中曲线的流畅性，判别真画与假画之间的区别。传统中国画由画的主体、落款印章和题记三部分组成。2010年，Bao等[17]根据传统中国画中的题记部分语义信息，对题记部分的颜色和结构进行特征提取。2013王征等[18]根据国画中纹理、颜色、形状等特征采用传统方法进行异构稀疏特征提取，采用监督的方式对特征进行学习，但是特征只有96维，不足以整体描述国画的特征。不同画家运笔的力道及要表现的情感的差异，使得画中笔画的粗细可以展现画家个人的风格，不仅可以从绘画的笔画、颜色等进行特征提取，还可以根据灰度直方图[19, 20]提取出的信息，用卷积神经网络对不同绘画风格的绘画者分类。从2012年Krizhevsky等[21]提出的AlexNet网络对ImageNet LSVRC-2010数据集进行分类训练，训练的分类结果以显著的优势赢得了ILSVRC 2012比赛，越来越多的人研究卷积神经网络对各类图像的分类效果问题。2015年，Sun等[22]根据中国水墨画中笔画的风格特点，提取出线条明显的区域并用CNN进行笔画特征提取，对水墨画艺术家进行分类。2016年，Sun等[23]采用蒙特卡洛凸壳特征选择模型整合基础特征描述子并使用支持向量机方法来对中国传统国画进行艺术家分类。中国画中工笔画与写意画在表现手法上存在区别，2017年高峰等[24]采用SIFT特征检测子和边缘检测得到图像关键区域，通过关键区域视觉特征和领域内部差异性的描述，采用级联分类器分析得出工笔画和写意画在表现手法上的不同。2018年，曹建收等[25]结合AlexNet网络中Dropout方法和GoogleNet网络[26, 27]中增加网络深度与宽度的Inception的思想构建了一个适用于对油画进行分类的网络模型。

## 2.2 卷积神经网络发展

随着网络性能的提升和迁移学习方法的使用，卷积神经网络的相关应用也逐渐向复杂化和多元化发展。在分类精度方面，在AlexNet[21]将ImagNet的图像分类准确度大幅度提升到84.7%之后，不断有改进的卷积神经网络模型被提出并刷新了AlexNet的记录，具有代表性的网络有VGG[29]、GoogLeNet[30-32]、ResNet[33]和BN-inception[26]等。最近，由微软提出的ResNet已经将ImageNet的图像分类准确度提高到了96.4%。Thomas等[34]对迄今为止的网络结构搜索空间、网络结构应用策略及其应用性能评估做了评估，在大规模数据量上等方面进行特征提取的效率在不断提高。而且随着艺术风格迁移[35-37]的研究，艺术类图像的多层风格特征可以被迁移到内容图像中。传统损失函数softmax多用于以上卷积神经网络中，其主要是对样本整体损失做调整，而没有分析样本类内损失以及类与类之间的损失。为增强类间的判别力，黄旭等[38]根据LDA（linear discriminant analysis）思想构建一种新的损失函数LDloss来参与卷积神经网络的训练，来最小化类内特征距离最大化类间特征距离，但是这种线性特征映射方法对于非线性分布结构问题没有很好的处理效果。Deng等人[39]选择3个样本即Anchor样本、正样本和负样本，通过惩罚Anchor样本与正负样本之间的距离来训练网络，但是样本对选择不当，会影响最终实验效果。文献[40]考虑各类别到类中心的距离，引入Center Loss用于人脸识别，但是没有考虑各类中心与总体样本的距离度量。在Center Loss基础上，Calefati等[41]提出一种约束条件，使学习到的特征比Center Loss具有更大的类间距离和更小的类内距离。

# 3学位论文研究的主要内容

现实中的数字图像在数字化和传输过程中常受到成像设备与外部环境噪声干扰等影响，因此选择合适的滤波方法对艺术图像去噪，以尽可能保留原始图像的真实性。用HSV颜色空间[28]将各类艺术图像的颜色、亮度和饱和度信息提取出。随着卷积神经网络的发展，研究人员发现扩展卷积神经网络的宽度与深度，对图像特征提取与分类或检索等的效果很好，但是不断增加网络的深度会出现网络退化的缺点，因此本文需要调整一个合适的卷积神经网络对艺术图像进行特征提取并用引入一个判别性的损失函数，以最大化类间的距离和最小化类内的距离。

本学位论文的重要内容包括：

第一章 绪论。主要介绍对艺术类图像进行分类的需求以及我们可能要面临的问题、研究背景和意义，和国内外对各类艺术图像进行分类的研究现状及研究成果。

第二章 基础知识介绍。介绍各类艺术图像各自具有的特点及各类艺术图像之间的主要差别。介绍主流神经网络对图像分类的应用及其优缺点，及介绍本论文中用到的卷积神经网络和损失函数。

第三章 数据进行预处理的阶段。我们直接从网上得到的数据可能是含有很大噪声，我们首先对得到的数据进行预处理，包括用中值滤波去噪及HSV颜色空间特征处理。

第四章 预测模型的提出和建立。将原始数据转换为有监督学习的样本形式，然后用卷积神经网络对预处理过的各类艺术图像进行特征提取。使用Keras框架实现，得出预测结果。

第五章 对实验的结果进行分析。使用其他的预测模型对原始数据进行预测作为对比模型。

第六章 对文章进行总结，并对下一步的工作进行展望。

# 4学位论文研究的方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明）

## 4.1 艺术图像数据采集

本论文的实验数据主要通过使用现在常用的python网络爬虫技术获取。数据获取的来源是一些对艺术图像开放性展览的主流网站。从网络上获取的图像大小、图像风格可能不是实验所需要的或是不属于这一类的艺术图像，还要将实验图像数据进行筛查。现在拟爬取油画、国画、水彩画、水粉画和版画各3000张。

## 4.2 图像预处理

从网络上获取的艺术类图像在数字化和传输过程中容易受到电子设备和外部环境的噪声干扰，噪声的存在可能会影响卷积神经网络对其特征的提取。因此，这就要求使用一系列滤波和图像增强算法对图像进行处理，尽可能保留原始图像的信息的同时，去除信号中无用的信息，并增强识别各类艺术图像的特征。

中国画，油画、版画、水粉画、水彩画各自具有自身特色的色调、纹理、饱和度、线条、亮度等特征。中国画，盛用线条、不重对画面背景的渲染，画中多留有空白，且以墨代色，使墨色产生丰富而细微的色度变化，也就是“墨分五色”。油画不同于中国画，其颜色丰富，颜料覆盖力强，具有不透明性，欧洲传统油画以明暗色彩为造型手段，立体感强、光感强、质感强烈。版画，主要是艺术家构思艺术图案，通过选择不同的材料进行制版和印刷而产生的艺术作品。它的颜色数目有限，颜色亮度与纯度较低，颜色间的过渡明显，刻画的形象不够生动。水彩画是通过水调和透明颜料作画的一种绘画形式，其特点是色彩透明，一层颜色覆盖另一层可以产生不同的视觉效果。水粉画是通过水调和粉质颜料作画的绘画形式，颜色透明性介于水彩画和油画之间。

中国画与部分水彩画、油画及部分水粉画在人眼视觉上看颜色特征区分度不大甚至有交叉的情况，在深入到像素级就会显现出它们的色块性。人眼是通过感知颜色的亮度、色调和饱和度来区别物体，RGB颜色空间不适合进行图像处理和分析。根据油画、国画、水彩画、水粉画和版画之间在色度、亮度及饱和度方面的特征，用HSV空间颜色对艺术图像的RGB颜色空间进行转换，可以进一步突出各类艺术图像的色调、亮度及饱和度的特征差别。

## 4.3 构建适用卷积神经网络

仅仅简单地增加网络复杂程度，会遇到一系列的瓶颈，如：过拟合问题，网络退化问题等。针对具体问题对网络结构及参数进行调整，并通过实验选出最优的网络。卷积神经网络性能的提升需要依靠更加合理的网络结构设计，根据各类艺术图像的特点和各类神经网络对特征提取的优缺点，微调一个卷积神经网络对各类艺术图像进行特征提取，使其克服过拟合的问题的同时，还要在深度上进行优化，避免部分网络层没有得到完善训练。图像的特征提取出，还要对提取出的图像进行分类，传统用softmax损失函数来反向传播不断迭代来优化网络参数，降低整体样本分类损失。本文根据各艺术类图像的特点，结合softmax损失函数并引入center loss损失函数和约束条件，进一步最大化各类间的距离，最小化各类内的距离，增强神经网络分类的判别力。

# 5 本文的创新点和关键问题

## 5.1本文的创新点

1. 在已有的网络模型基础上，通过实验对其宽度和深度进行调整，以适用艺术图像的特征提取和分类，将5类艺术图像作为训练集输入神经网络，实现各类艺术图像的定性分析。
2. 一组便于卷积神经网络学习，能够很好区分不同艺术类的特征描述。
3. 本文引入Git Loss损失函数到实验的卷积神经网络中，使学习到的特征具有更大的类间距离和更小的类内距离。

## 5.2本文的关键问题

其一，如何从网上获取大量的各类艺术图像，直接从网上得到的艺术图像是有噪声的，这需要对各类艺术图像进行去躁，要分析各个元素之间的关系，对原始数据的处理就会很重要关键；

其二，中国画和水彩画、油画和水粉画在普通人视觉上难以区分，非艺术专业的人将训练集分类，可能会混淆各类艺术图像，因此要请相关艺术的人对训练集进行评估区分；

其三，模型在训练阶段考虑过拟合问题，训练时间如果过长，如何提高模型训练效率。

# 6学位论文研究的工作条件

通过学校购买的数据库，能够找到并阅读国内外的参考文献等资料，参考其中与研究问题相关的有价值的技术方法。

其次，对机器学习、神经网络及深度学习算法等具有深入的了解。要会用python等编程语言以及Keras深度学习框架实现实验所需的程序代码。通过阅读文献和专著来学习和研究。

实验用到的硬件：Nvidia Tesla P100-12G GPU一块，Intel Xeon E5-2620 V4 CPU 两颗，DDR4 2400MHz 16 G 内存四条。

深度学习框架采用 Keras 2.1.5，其后端使用 TensorFlow 1.7.0，Python 版本为 3.5.5。

# 7学位论文研究的进度计划安排

* 2018年11月~2019年2月：收集查阅中、外相关课题资料，了解课题背景，熟悉卷积神经网络及图像处理的基本知识。收集实验所需的艺术类图像数据，完成论文的准备工作。
* 2019年3月~2019年7月：据搜集的资料，实现系统总体框架搭建工作，深入研究图像处理和卷积神经网络优化。
* 2019年8月~2019年10月：完善程序，测试代码，并对检测数据结果进行分析比较，对课题工作进行总结并撰写出论文初稿。
* 2019年11月~2019年12月：修改论文，完善论文并准备毕业答辩

# 参考文献

[1] Shen C, Jin Z, Zhao Y, et al. Deep Siamese Network with Multi-level Similarity Perception for Person Re-identification[J]. Proceedings of the 2017 ACM on Multimedia Conference. 2017: 1942-1950.

[2] Corbiere C, Ben-Younes H, Rame A, et al. Leveraging Weakly Annotated Data for Fashion Image Retrieval and Label Prediction[J]. arXiv:1709.09426.2017

[3] Deng D, Wang R, Wu H, et al. Learning deep similarity models with focus ranking for fabric image retrieval[J]. Image and Vision Computing. 2018, 70: 11-20.

[4] Alzu'Bi A, Amira A, Ramzan N. Content-based image retrieval with compact deep convolutional features[J]. Neurocomputing. 2017, 249: 95-105.

[5] Dong C, Loy C C, He K, et al. Image Super-Resolution Using Deep Convolutional Networks[J]. arXiv:1501.00092. 2015.

[6] Arandjelovic R, Zisserman A. All about VLAD[J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2013: 1578-1585.

[7] Jiang F, Hu H M, Zheng J, et al. A hierarchal BoW for image retrieval by enhancing feature salience[J]. Neurocomputing. 2016, 175(PartA): 146-154.

[8] Perronnin F, Liu Y, Anchez J S, et al. Large-Scale Image Retrieval with Compressed Fisher Vectors[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017, 44: S29-S30.

[9] Yang P, Sun X, Li W, et al. SGM: Sequence Generation Model for Multi-label Classification[J]. Proceedings of the 27th International Conference on Computational Linguistics. 2018: 3915-3926.

[10] Penatti O, B. Silva F, Valle E, et al. Visual word spatial arrangement for image retrieval and classification[J]. Pattern Recognition. 2014, 47(2): 705-720.

[11] Wang H, Wang J. An Effective Image Representation Method Using Kernel Classification[J]. 2014 IEEE 26th International Conference on Tools with Artificial Intelligence. 2014: 853-858.

[12] 汪珊娜，张华熊，康锋. 基于卷积神经网络的领带花型情感分类[J]. 纺织学报.2018, 39(8): 121-127.

[13] Li J, Wang J Z. Studying digital imagery of ancient paintings by mixtures of stochastic models[J]. IEEE Transactions on Image Processing. 2004, 13(3): 340-353.

[14] Jiang S, Huang Q, Ye Q, et al. An effective method to detect and categorize digitized traditional Chinese paintings[J]. Pattern Recognition Letters. 2006, 27(7): 734-746.

[15] Shen J. Stochastic modeling western paintings for effective classification[J]. Pattern Recognition. 2009, 42(2): 293-301.

[16] Yao L, Li J, Wang J Z. Characterizing elegance of curves computationally for distinguishing Morrisseau paintings and the imitations[J]. Proceedings International Conference on Image Processing, ICIP. 2009: 73-76.

[17] Bao H, Liang Y E, Liu H Z, et al. A novel algorithm for extraction of the scripts part in traditional Chinese painting images[Z]. 2010, V2-26 - V2-30.

[18] 王征，孙美君，韩亚洪，等. 监督式异构稀疏特征选择的国画分类和预测[J]. 计算机辅助设计与图形学学报. 2013, 25(12): 1848-1855.

[19] 盛家川. 基于直方图的水墨画艺术风格研究[J]. 计算机工程与应用. 2014, 50(21): 1-3.

[20] Sheng A J, Jiang J. Style-based classification of Chinese ink and wash paintings[J]. Optical Engineering. 2013, 52(9): 93101.

[21] Krizhevsky Alex, Ilya Sutskever, Geoffrey E.Hinton. ImageNet classification with deep convolutional neural networks[Z]. 2012，1097-1105.

[22] Sun M, Zhang D, Ren J, et al. Brushstroke based sparse hybrid convolutional neural networks for author classification of Chinese ink-wash paintings[Z]. 2015，626-630.

[23] Sun M, Zhang D, Wang Z, et al. Monte Carlo Convex Hull Model for classification of traditional Chinese paintings[J]. Neurocomputing. 2016, 171: 788-797.

[24] 高峰，聂婕，黄磊，等. 基于表现手法的国画分类方法研究[J]. 计算机学报. 2017, 40(12): 2872-2882.

[25] 曹建收，陈光喜，任夏荔，等. 基于深度学习的油画分类网络模型[J]. 桂林电子科技大学学报. 2018, 38(1): 65-68.

[26] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[J]. arXiv:1502.03167. 2015.

[27] Iino N, Takamatsu W, Iino A, et al. Going Deeper with Convolutions Christian[J].In CVPR.2015.

[28] 马玲，张晓辉. HSV颜色空间的饱和度与明度关系模型[J]. 计算机辅助设计与图形学学报. 2014, 26(8): 1272-1278.

[29] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recongnition[J]. arXiv:1409.1556. 2015.

[30] Chung C, Patel S, Lee R, et al. Implementation of an integrated computerized prescriber order-entry system for chemotherapy in a multisite safety-net health system[J]. American Journal of Health-System Pharmacy. 2018, 75(6): 398-406.

[31] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[J]. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17). 2017: 4278-4284.

[32] Goodman P, Namdeo A, Galatioto F, et al. Rethinking the Inception Architecture for Computer Vision[J]. 2014-16th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Proceedings. 2014.

[33] He K, Sun J. Deep Residual Learning for Image Recognition[J]. In CVPR.2015: 770-778.

[34] Elsken T, Metzen J H, Hutter F. Neural Architecture Search: A Survey[J].arXiv:1808.05377. 2018.

[35] Jing Y. Neural Style Transfer: A Review[J]. arXiv:1705.04058. 2018.

[36] Huang H, Zang Y, Li C F. Example-based painting guided by color features[J]. Visual Computer. 2010, 26(6-8): 933-942.

[37] Gatys L A, Ecker A S, Bethge M, et al. A Neural Algorithm of Artistic Style[J]. arXiv:1508.06576. 2015.

[38] 黄旭，凌志刚，李绣心. 融合判别式深度特征学习的图像识别算法[J]. 中国图像图像学报. 2018, 23(4): 0510-0518.

[39] Deng D, Wang R, Wu H, et al. Learning deep similarity models with focus ranking for fabric image retrieval[J]. Image and Vision Computing. 2018, 70: 11-20.

[40] Shen L, Lin Z, Huang Q. A Discriminative Feature Learning Approach for Deep Face Recognition Yandong[J]. Proceedings of the 14th European Conference on computer Vision. 2015, 1: 499-515.

[41] Calefati A, Janjua M K, Nawaz S, et al. Git Loss for Deep Face Recognition[J]. arXiv:1807.08512. 2018.