

基于卷积神经网络的唐卡尊像自动分类研究

陈玉红, 刘晓静*

(青海大学 计算机技术与应用系, 青海 西宁 811601)

摘要:唐卡作为中国的非物质文化遗产之一,受到越来越多人的关注。如何快速、准确地对唐卡中的尊像进行分类,对于唐卡的研究以及数字化保护传承极其重要。因此,该文提出了一种改进的基于卷积神经网络基本结构的唐卡尊像自动分类方法。通过手动采集及爬虫技术等收集唐卡图像,构建唐卡尊像数据集,并作为网络模型的输入数据。在保留原有卷积神经网络基本结构的前提下,在传统卷积神经网络结构每组的隐藏层中加入批量归一化层,改善模型的训练效率,并以 Relu 作为卷积池化层的激活函数,在最后一层全连接输出层前面加入 Dropout 层,减少过拟合,全连接输出层使用 Softmax 作为激活函数,而损失函数则使用交叉熵,使分类效果更好,同时采用 Adam 优化方法来进行模型的优化,并应用在唐卡尊像分类方面。最终在自己建立的唐卡数据集上进行实验,分类准确率高达 94.7%,比其他典型方法高出约 3%,分类效果更佳,更有利于唐卡这种非物质文化遗产的数字化保护。

关键词:非物质文化遗产;数字化保护;卷积神经网络;图像分类;唐卡尊像数据集

中图分类号:TP181

文献标识码:A

文章编号:1673-629X(2021)12-0167-08

doi:10.3969/j.issn.1673-629X.2021.12.028

Research on Automatic Classification of Thangka Images Based on Convolutional Neural Network

CHEN Yu-hong, LIU Xiao-jing*

(Department of Computer Technology and Application, Qinghai University, Xining 811601, China)

Abstract: As one of China's intangible cultural heritage, Thangka has attracted more and more attention. How to quickly and accurately classify Thangka images is extremely important for the research and digital protection of Thangka. Therefore, we propose an improved automatic classification method for Thangka images based on the basic structure of the convolutional neural network. Thangka images are collected by manual collection and crawler technology, and the Thangka image data set is constructed and used as the input data of network model. Under the premise of retaining the basic structure of the original convolutional neural network, each group of the traditional convolutional neural network structure adds a batch normalization layer to the hidden layer to improve the training efficiency of the model and uses Relu as the activation function. A Dropout layer is added before the last fully connected output layer to reduce overfitting and fully connected output layer uses Softmax as the activation function, and the loss function uses cross-entropy to make the classification effect better. At the same time, the Adam optimization method is used to optimize the model and applied to the classification of Thangka images. Finally, the experiments on the Thangka data set established by ourselves show that the classification accuracy rate is as high as 94.7%, which is about 3% higher than other typical methods. The classification effect is better, which is more conducive to the digital protection of Thangka, an intangible cultural heritage.

Key word: intangible cultural heritage; digital protection; convolutional neural network; image classification; dataset of Thangka images

0 引言

唐卡是用颜料或其他材料涂在彩色丝绸上的宗教卷轴画。它是一种具有地方特色的绘画形式^[1],也是中国重要的非物质文化遗产之一。唐卡内容丰富,包括建筑物、医药学、天文历算、传奇故事等^[2]。由于唐

卡生产材料、涂料、地理环境、自然环境和宗教文化的特殊性,仅存在于少数地区且现存数量较少。此外,唐卡内容广泛,色泽鲜艳,整个画面层次分明,前景和背景区分不明显,构图复杂,结构严谨,一幅唐卡画像首先按照造像量度标准起稿,面部五官、头、胸、腰等各个

收稿日期:2021-01-14

修回日期:2021-05-18

基金项目:国家自然科学基金(61862053);青海省科技厅应用基础研究项目(2021-ZJ-717)

作者简介:陈玉红(1995-),女,硕士,研究方向为信息可视化与媒体计算;通讯作者:刘晓静(1978-),女,硕士,副教授,CCF会员(33794M),研究方向为图像处理、虚拟现实及三维可视化。

部位都有严格的要求。唐卡画家通常需要多年的实践才能绘制出真正的唐卡^[3],并且唐卡画像保存有一定的难度,难以处理受潮、水渍、霉斑、脏化等问题,因此数量较少。以上两个主要原因加上唐卡是中国的非物质文化遗产,要加大力度研究唐卡的数字化保护,为保护唐卡贡献自己的一份力量。

分类时较其他类型图像存在的挑战:

(1)唐卡图像和一般图像相比,内容丰富,构图复杂。如图 1(b)所示,同一幅唐卡图像中,各个尊像的手势、坐姿、服装都有不同之处,人物之外的地方,细节精细,构图元素非常复杂且背景干扰大。在唐卡图像当中,背景复杂,构图元素极多,对唐卡尊像分类干扰较大,这使得分类难度加大。

(2)类内差大,即同一类的物体外观特征差别比较大。如图 1(c)、(d)所示,从唐卡分类规则得知,手拿宝剑的都是文殊菩萨,但是外观却差别很大。

(3)类间模糊性大,即不是同一类物体的实例具有一定的相似性。如图 1(e)、(f)所示,分别为黄财神、财宝天王,但相似性较大。

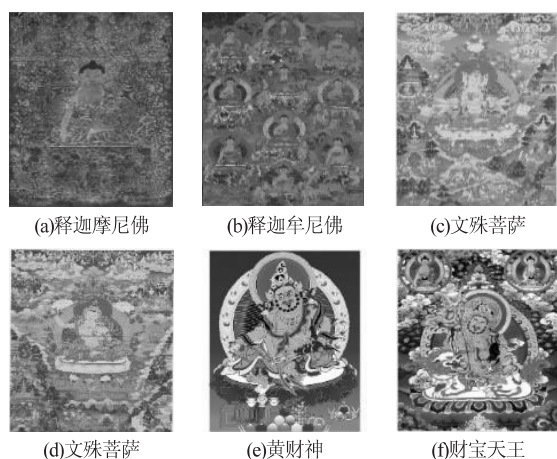


图 1 不同的唐卡尊像图像

该文首次构建了一个有关唐卡尊像的数据集,首次应用卷积神经网络对唐卡尊像图像进行分类。在唐卡专业分类知识的前提下,通过设计的 CNN 网络模型,根据内容对唐卡尊像进行分类,为研究人员提供了便利,让越来越多的人理解和保护唐卡。

该文的主要贡献在于:构建了尊像唐卡分类数据集,并提出了一个浅层的 CNN 模型进行唐卡尊像分类,该 CNN 模型结合了 AlexNet^[4]、VGG16^[5]、ResNet^[6]模型的优点。最后,在同一个唐卡数据集上将优化的模型和其他典型分类模型进行了比较。

1 相关工作

图像分类分为基于特征学习和深度学习的图像分类方法,但对唐卡图像的研究较少。下面将分别介绍

传统常见的图像分类方法、基于深度学习的图像分类方法和对唐卡图像的研究。

周依莲提出一种特征选择策略,在得到最优特征集的同时也可以约束不相关特征,从而实现了 CT 图像的不同类型肺癌分类,准确率达到 86% 左右^[7];李奇峰提出在遥感影像分类中采用不同的方式进行纹理特征提取,然后和光谱特征相结合,并利用支持向量机 (support vector machine, SVM) 分类器进行遥感影像分类的方法,分类精度提升 10%^[8];黄万继将尺度不变特征变换 (scale-invariant feature transform, SIFT) 和 SVM 相结合应用到场景识别方面,与其他传统方法相比,分类准确率达 85.12%^[9];晋远等人用贝叶斯分类器进行家电设备识别,对每个设备类别的条件概率估计,最终识别正确率达 91%^[10]。

除此以外,图像分类还包括电器标签分类^[11]、交通标志分类^[12-13]、花卉图像分类^[14]、果蔬图像分类^[15-16]等方面,但近几年,应用深度学习进行图像分类的研究也愈发增多,尤其是在 CNN 基础上进行分类研究的^[17-18]。

在零售情况未知的情况下,用架构相同,但是矩阵权重不同的两个九层卷积神经网络对水果进行分类,两个网络分别分类水果类别和区分水果品种,最终对于同一品种的水果分类准确度达到 99.78%^[19];童浩然将 ImageNet 数据集和 MNIST 数据集用于改进和优化 CNN,在使用 Inception 结构时分别使用了 Dropout 和 Adam 算法,最终识别率得到提高,达 96.53%^[20];谢潇雨提出卷积神经网络模型的批量归一化处理,之后经过全连接层的传递,并通过输出层获得最终的分类结果,从而提高了入侵检测系统的准确率和检出率^[21];任飞凯通过局部二值模式 (local binary patterns, LBP) 特征脸作为 CNN 的输入,同时采用 Adam 算法优化模型,融入 Nesterov 动量思想并通过改进二阶估计偏差修正来使学习率呈现单调递减的状态,实验结果显示优化后的模型收敛性更好,人脸识别效果更佳,准确率高达 91.05%^[22];江白华在 CNN 模型结构中引入 DenseNet 稠密连接结构,并使用批归一化与 Dropout 层,选择 Adam 优化网络参数,Relu 作为激活函数,实现了人脸在线识别,并通过与 VGG16、ResNet14 模型对比,验证该模型的性能优异,在线识别准确率达 92.5%^[23];黄雪丽等人用 SMOTE 技术扩增数据,将数据直接输入到 CNN 模型中,国画图像的分类方法是基于 Relu 和 Sigmoid 激活函数的组合,准确率达 91.55%^[24];朱等人改进 CNN,将全连接的层数减少一层,并用具有七个缺陷标签的 Softmax 分类层替换了 Softmax 分类器。卷积层和池化层的权重参数在预训练模型中共享,此模型对于水泥混凝土桥梁

表面缺陷的分类准确率为 98.96%^[25]。

卷积神经网络有较多改进版本,主要方法就是采用 Adam 算法优化模型,通过批归一化与 Dropout 改善网络,缓解过拟合,不同作者有不同的网络模型,但是在唐卡方面的工作并不是很多。国内主要是对破损唐卡进行修复^[26],提出的修复方法主要适用于对称性唐卡;再者就是唐卡图像的检索以及局部图像的识别^[27],李晓杰将局部累加直方图和传统局部累加图进行对比,研究发现基于 HSV 空间的局部累加直方图方法更适用于唐卡图像的检索^[28];王铁君等人提出一种融合 Hu 矩和局部二进制模式的图像特征提取方法,在此基础上,提出了一种基于距离密度的改进的 K 近邻分类算法,待识别唐卡法器所属的种类与周围邻居点的距离越近,相关度越高^[29];王铁君等人基于改进的 SVM 分类方法,根据主题对唐卡图像进行分类,然后从两个层次上标记唐卡中数量最多的圣像类唐卡,根据监督学习算法实现自动识别与标注唐卡中的法器、头饰和手势,得到标注关键词,与已建立的推理规则进行查询推理,实现唐卡图像的全局二次标注^[30-31]。谢辉和其他人仅用了简单的分类规则,他们将唐卡分为四类,一种是基于不同的背景:红色、金色、黑色和彩色;另一个是根据内容来判断,将唐卡分为佛、菩萨、宗师等,没有更详细的分类^[32]。刘华明等人将唐卡头饰通过自我编码进行监督学习,然后在卷积池化操作过程之后输入标记的训练样本以训练 Softmax 分类器,此论文方法在唐卡头饰分类方面具有较好的分类性能^[33]。

解辉等人仅对尊像类唐卡根据内容进行了简单的分类,并没有更详细的分类,因此,该文在此基础上对尊像类唐卡进行了更详尽的分类,并建立了分类数据集。从唐卡的复原,目标识别与检索的研究角度来看,尚无关于尊像唐卡图像分类与识别的研究。因此,该文提出了一种改进的 CNN 模型,用于对唐卡人像进行分类。实验证明,该模型在唐卡尊像图像分类中具有比典型模型更好的效果。该实验为唐卡研究人员提供了便利,使越来越多的人了解唐卡,并加入保护唐卡的行列。越来越多的人可以理解、探索和研究唐卡文化和艺术的真正价值,继承和发扬唐卡文化的优良传统,更有利于对这一非物质文化遗产进行数字保护。

2 数据集

2.1 唐卡分类规则

根据唐卡相关书籍《世界最美唐卡—三界故事》、《中国唐卡》学习得到的圣像类唐卡相关知识,将唐卡根据主尊的类别进行分类,根据有关书籍分类知识的总结归纳一共将唐卡分为六大类,其中每一类别又可

以具体细分。

在实验当中,一共选取了 13 个小类进行实验,其中,如图 2(a)所示,释迦牟尼佛为黄色身,发髻高耸,身披袈裟,结跏趺坐于法台上;图 2(b)为四臂观音,共有四只手,一般中间两手合掌于胸前,右上手持水晶念珠,左手持莲花;图 2(c)为白度母,主要形象特征是身体为白色,面有三目,上身袒露,左右手手心中各有一眼,左手于胸前执莲花;图 2(d)为药师佛,虽然和释迦牟尼佛很相似,但是也有不同之处,药师佛身体一般为蓝色,右手膝前执一果树枝,左手脐前捧佛钵。根据前述圣像类唐卡人物的分类知识以及其他尊像分类的知识,将图 2 所示的唐卡人物等进行分类。



图 2 不同的唐卡尊像图像

2.2 获取原始唐卡图像

唐卡图像通过多种渠道获取,途径一:通过爬虫技术从网上下载唐卡图像,然后手动筛选,将不是唐卡图像、图像不完整,残缺较多的图像剔除;途径二:将唐卡相关书籍《故宫唐卡图典》、《西藏唐卡(精)》、《热贡年都乎唐卡艺术》上的唐卡图像用扫描仪进行扫描得到;途径三:通过走访唐卡绘制大师,了解唐卡相关知识,并征得同意后进行唐卡图像的拍摄;整个数据集的收集花费时间较长。

2.3 构建唐卡数据集

由于唐卡图像数量有限且不均匀,难以满足神经网络模型需要大量数据的要求,因此对数量比较少的图像进行数据增强,通过镜像翻转、亮度增强、颜色增强、对比度增强、添加高斯噪声、椒盐噪声等操作扩充数据集,使测试集在数量扩大的同时大致均匀,其中,镜像翻转操作包括上下翻转和左右翻转;亮度增强包括使图片变亮和变暗,分别取原来图片明暗度的 1.5

4 实验及分析

4.1 实验步骤

唐卡人物分类的流程主要包括图像预处理、模型构建以及模型预测。

步骤一:唐卡尊像图像预处理。

(1)将样本集中的所有图像进行镜像翻转、亮度增强、亮度减弱、颜色增强、添加高斯噪声、椒盐噪声等图像增强的预处理操作;

(2)生成对应的数据列表,对图片进行缩放,裁剪为 100×100 的尺寸,即图片维度为 $(3, 100, 100)$ 。然后将图片数组进行归一化处理,得到 0 与 1 之间的数值;

(3)将所有的圣像类唐卡图像样本输入,并随机抽取每个小类里的十分之一的样本作为测试集图像,其他部分作为训练集样本。

步骤二:构建 CNN 基本模型。

(1)所构建的 CNN 基本网络结构主要包括三组卷积层、池化层和完全连接层;

(2)将学习率初始化为 0.01, Batch size 大小设定为 128 进行实验,若准确率不再提高,则将学习率每次降低 10 倍, Batch size 大小根据 2^n 进行调整,得到一个基本的网络模型,最终学习率设为 0.000 1, batch size 设为 64, 经过一段时间的训练,最终损失函数收敛较快,准确率不断处于上升阶段;

(3)加入 BatchNorm 层和 Dropout 层进行进一步的实验,手动调整学习率以及 Batch size 的数值,并用验证集对训练进行评估,不断优化,保存好训练的模型,然后将卷积池化层的可调参数进行实验。

步骤三:模型预测。

- (1)配置好预测的模型,准备好要预测的图片;
- (2)加载之前训练好的预测模型;
- (3)把图片放入预测模型,输出最终预测结果。

4.2 实验结果及分析

该实验主要在配备 8 线程 CPU, NVIDIA V100 GPU, 16 GB 显存, 100 GB 磁盘和 32 GB RAM 的服务器上运行,并且训练测试过程是使用 python 实现的。

首先在训练集上进行实验以最大程度地提高训练集的准确性,然后在测试集上进行测试以获得最终的准确率。要在选定的训练集上进行训练,首先在基础上调整学习率和 Batch size,然后调整池化层的相关参数,以下是实验数据。

经过不断测试,最终学习率选择 0.000 1 比较合适,一共训练 30 轮。首先暂定滤波器大小为 3, 池化层大小为 2, 池化层步长为 2, Batch size 选择 64, 激活类型选择 Relu 函数。

为了测试滤波器大小、池化层大小、池化层步长以

及 Batch 大小对于圣像类唐卡人物分类的准确率的影响,进行了以下几个实验,实验主要采用单一变量法进行,即进行其中一项对最终分类结果的影响,设定其他项保持不变:

首先测试滤波器大小对实验结果的影响,保持池化层大小、池化层步长为 2, Batch 大小为 64, 改变滤波器大小,对比分类准确率。

表 1 滤波器大小对实验结果的影响

filter_size	pool_size	pool_stride	Batch_size	Accuracy/%
2	2	2	64	94.16
3	2	2	64	93.86
4	2	2	64	93.18
5	2	2	64	91.69

由表 1 可以看出,在池化层大小、池化层步长以及 Batch 大小保持不变的情况下,滤波器大小设置为 2 准确率更高,所以当 filter_size=2 时,保持池化层步长不变,测试池化层大小对实验结果的影响。

池化层大小对分类准确率的影响

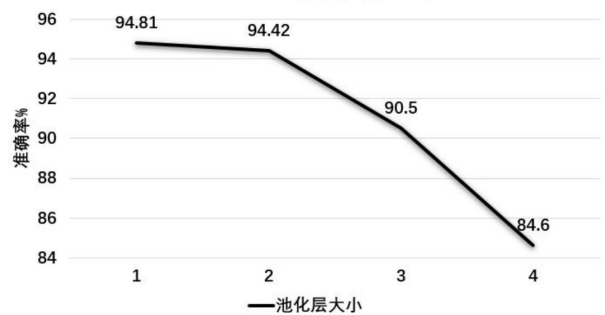


图 4 池化层大小对分类准确率的影响

如图 4 所示,当滤波器大小,池化层步长和批次大小保持不变时,池化层大小将用作变量。实验结果表明,当池化层大小为 1 时,分类准确率最高,达到 94.81%。

池化层步长对分类结果的影响

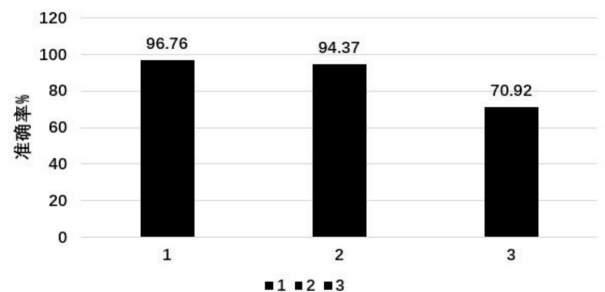


图 5 池化步长对分类准确率的影响

从图 5 中可以看出,当滤波器大小为 2 时,池化层大小为 1, 将池化层步长设置为 1, 则精度最高,达到 96.76%。

此外,批次值的大小也会影响最终分类结果的准确性和效率。在实际训练中,将所有训练集划分为相

同的训练集,然后在第二次训练期间,将所有分割的批次训练集依次放入网络中,直到完成所有批次训练集的训练为止,然后执行下一轮迭代。通过实验数据总

结批次的大小和准确性之间的关系,以进行更好的训练。实验数据如表2所示。

表2 Batch大小对实验结果的影响

Batch_size	pool_size	pool_stride	filter_size	Accuracy/%
10	1	1	2	81.21
50	1	1	2	90.50
100	1	1	2	96.53
200	1	1	2	96.19
300	1	1	2	96.28

由表2可知:当Batch的值小于50时,准确率只有81.21%,并且在训练过程中花费的时间明显比其他Batch值的时候更长,当Batch在50及以上时,基本保持不变,Batch越大,训练所花费时间越多,所以最终实验中采用Batch为64比较合理。

在训练集上调整各类参数,使最终的训练精度达到最高以后,在测试集上进行测试。如上所述,文章中采用每一类当中的十分之一作为最终测试集的数据。

具体的分类结果如图6所示。



图6 部分分类结果展示

根据最终的预测结果,获得模型的准确性、召回率、精确度,对于采集的样本,可以直接知道实际情况下哪些数据为正,哪些数据为负。通过使用样本数据运行分类模型结果,还可以知道模型认为哪些数据是正数据,哪些数据是负数据,下面是四个基本指标:

(1) 真实值为正,并且模型认为该数字为正 (True Positive = TP);

(2) 真实值为正,但模型认为该数字为负 (False Negative = FN);

(3) 真实值为负,但模型认为该数字为正 (False Positive = FP);

(4) 真实值为负,并且模型认为该数字为负 (True Negative = TN);

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \quad (3)$$

准确率的计算公式见公式(3),精确度和召回率的计算公式见公式(4)和公式(5):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

在多分类问题中,在计算召回率和准确率时,将要

为每个类别计算的类别视为正类别,而将其余类别视为负类别。各个类别的召回率、准确率和精确度如表 3 所示。

表 3 各类别的召回率、精确度和准确率

类别	召回率/%	精确度/%	准确率/%
十一面千手千眼观音	99.0	88.6	93.5
药师佛	97.7	100	98.8
尊胜佛母	96.3	92.8	94.5
文殊菩萨	99.2	86.2	92.2
大威德	99.1	100.0	99.5
释迦牟尼佛	85.9	100.0	92.4
莲花生大师	100.0	86.0	92.5
准提菩萨	100.0	97.2	98.6
白度母	94.2	98.0	96.1
宗喀巴	100.0	96.6	98.3
大白伞盖佛母	88.9	92.3	90.6
普贤菩萨	83.6	97.2	89.9
四臂观音	97.5	86.3	88.6
均值	95.5	93.9	94.7

由性能评价指标 Accuracy 可得:大威德金刚的得分高达 99.5%,而四臂观音的得分最低,为 88.6%,主要原因是大威德金刚的图像在色彩以及构造上是最复杂的,与其他类别的差别较大,比较容易区分;但是四臂观音的造像和普贤菩萨比较相似,面部表情、头饰以及手势比较相似,所以分类错误的图像较多。另外,由于整个训练集当中每个类别的数量不均匀,释迦牟尼佛的数量最多,因此,其他类别当中也有较多的图像被分类为释迦牟尼佛。

4.3 与其他分类算法的比较

为了验证文中提出方法的性能,在最佳参数条件下,将该方法与其他典型的图像分类方法进行了比较,如图 7 所示。实验结果表明:在相同数据集下,该方法的准确率比其他几种方法高约 3%,分类效果更好。

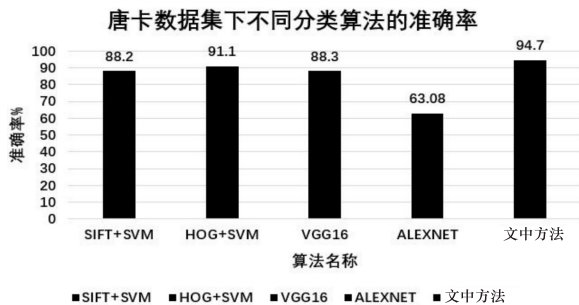


图 7 文中方法与其他典型方法准确率的比较

5 结束语

该文主要研究基于卷积神经网络模型的唐卡人像

自动分类方法,首先构建一个唐卡尊像分类数据集,然后在此基础上进行基于卷积神经网络模型的唐卡尊像图像分类实验。实验证明,在唐卡尊像图像分类中,提出模型的准确率提高了 3% 以上,更适用于唐卡尊像图像分类。减少分类所花费的人力和物力,对唐卡非物质文化遗产进行数字保护,使更多人参与唐卡的保护。

该方法的优点是使用了相对较新的卷积神经网络来重组网络结构层,实现了较浅的网络结构分类模型的构建,并且在自己构建的数据集上进行测试。然而该方法也存在一定的缺陷,即训练模型花费时间太久,系统界面功能单一。在以后的工作当中,将以提高训练效率为重点,并在此次实验的基础上继续进行唐卡人物的检测,将其他神经网络模型(如 yolov3、SSD 等)进行改进并应用到唐卡人物的分类及人物、法器、头饰等的检测当中,为唐卡这种非物质文化遗产的数字保护贡献一份力量。

参考文献:

- [1] MAKKUNI R. The electronic sketch book of Tibetan Thangka painting [J]. The Visual Computer, 1989, 5 (4): 227-242.
- [2] QIAN J, WANG W. Religious portrait Thangka image retrieval based on gesture feature [C]//IEEE 2009 Chinese conference on pattern recognition (CCPR). Nanjing, China: IEEE, 2009: 1-5.
- [3] 次仁央啦. 略谈馆藏唐卡的保护 [J]. 西藏科技, 2020 (1): 31-33.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60 (6): 84-90.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Computer Science, 2014 (2): 1-14.
- [6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, Nevada, USA: IEEE, 2016: 770-778.
- [7] 周依莲. 基于仿滤波子的 CT 图像特征选择对非小细胞肺癌的鉴别 [J]. 临床检验杂志 (电子版), 2020, 9 (1): 213-214.
- [8] 李奇峰. 结合多特征描述和 SVM 的遥感影像分类研究 [D]. 郑州: 郑州大学, 2015.
- [9] 黄万继. 基于 SIFT 特征和 SVM 的场景分类 [D]. 大连: 大连理工大学, 2011.
- [10] 晋远, 孙红三, 叶紫, 等. 基于大数据 Bayes 分类的家电设备识别算法 [J]. 建筑科学, 2017, 33 (4): 31-38.
- [11] 张治国, 李德平, 柳宁. 电器标签分类的 SVM 方法研究 [J]. 机电工程技术, 2019, 48 (12): 1-4.

- [12] 王 博,李亚文.一种改进的 CNN 交通标志识别算法[J].计算机与数字工程,2019,47(12):3082-3088.
- [13] CIRESAN D,MEIER U,MASCI J,et al. Multi-column deep neural network for traffic sign classification[J]. Neural Networks,2012,32:333-338.
- [14] 尹 红,符 祥,曾接贤,等.选择性卷积特征融合的花卉图像分类[J].中国图象图形学报,2019,24(5):762-772.
- [15] 冯语姗.果蔬图像的混合粒度分类研究[D].合肥:中国科学技术大学,2016.
- [16] ZENG G. Fruit and vegetables classification system using image saliency and convolutional neural network [C]//2017 IEEE 3rd information technology and mechatronics engineering conference (ITOEC). Chongqing, China: IEEE, 2017: 613-617.
- [17] WANG J, YANG Y, MAO J, et al. CNN-RNN: a unified framework for multi-label image classification [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, Nevada, USA: IEEE, 2016: 2285-2294.
- [18] WEI Y, XIA W, LIN M, et al. HCP: a flexible CNN framework for multi-label image classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(9):1901-1907.
- [19] RUDNIK K, MICHALSKI P. A vision-based method utilizing deep convolutional neural networks for fruit variety classification in uncertainty conditions of retail sales[J]. Applied Sciences, 2019, 9(19):3971.
- [20] 童浩然.基于深度学习的图像分类的研究[D].淮南:安徽理工大学,2018.
- [21] 谢潇雨.基于卷积神经网络的入侵检测模型研究[D].南京:南京邮电大学,2019.
- [22] 任飞凯.基于卷积神经网络人脸识别研究与实现[D].南京:南京邮电大学,2019.
- [23] 江白华.基于深度学习的人脸识别研究[D].淮南:安徽理工大学,2019.
- [24] 黄雪丽,王小华,杨 冰,等.基于 CNN 的中国绘画图像分类[J].杭州电子科技大学学报,2017,37(2):45-50.
- [25] 陈济楠,李少波,高 宗,等.基于改进 CNN 的年龄和性别识别[J].计算机工程与应用,2018,54(16):135-139.
- [26] ZHU Song. An intelligent classification model for surface defects on cement concrete bridges [J]. Applied Sciences, 2020, 10(3):972.
- [27] WANG W, JIA Y. Damaged region filling and evaluation by symmetrical exemplar-based image inpainting for Thangka [J]. EURASIP Journal on Image & Video Processing, 2017, 2017(1):1-13.
- [28] WANG W. Study of thangka image retrieval and multimedia presentation management system [C]//2009 fifth international conference on intelligent information hiding and multimedia signal processing. Kyoto, Japan: IEEE, 2009:981-984.
- [29] 李晓杰.基于内容的唐卡图像检索技术研究[D].兰州:西北民族大学,2011.
- [30] 王铁君,王维兰.基于多特征融合的唐卡图像法器识别方法[J].计算机工程,2016,42(3):198-203.
- [31] 王铁君,王维兰.基于本体的唐卡图像标注方法[J].吉林大学学报:工学版,2020,50(1):289-296.
- [32] 解 辉,王维兰.唐卡图像分类知识库的初步设计与构建[J].微计算机信息,2009,25(19):277-279.
- [33] LIU Huaming, BI Xuehui, WANG Xiuyou, et al. Automatic classification of Thangka headdresses based on convolutional depth neural networks [C]//Proceedings of the 2017 international conference on machine learning and soft computing. Ho Chi Minh City, Vietnam: ACM, 2017:105-110.
- [34] 程 冰.基于卷积神经网络的自动标注技术的研究[J].电子世界,2019(16):124-126.
- [35] YANAI K, KAWANO Y. Food image recognition using deep convolutional network with pre-training and fine-tuning [C]//2015 IEEE international conference on multimedia & expo workshops (ICMEW). Turin, Italy: IEEE, 2015:1-6.
- [36] LEAL-TAIXÉ L, CANTON-FERRER C, SCHINDLER K. Learning by tracking: siamese CNN for robust target association [C]//2016 IEEE conference on computer vision and pattern recognition workshops (CVPRW). Honolulu, HI, USA: IEEE, 2016, :418-425.
- [37] 金 昊,康宇哲,齐希阳,等.基于 Faster R-CNN 的高压电线缺陷检测方法[J].计算机应用,2019,39(z2):97-102.
- [38] 余晓龙,林国平.基于标签相关性的卷积神经网络多标签分类[J].闽南师范大学学报:自然科学版,2019,32(2):19-25.
- [39] 肖 欢,阳习党,朱 威.一种基于 BP 神经网络的输电线路故障检测系统[J].信息通信,2019(4):100-102.
- [40] 张璐璐.基于卷积神经网络的人脸面部表情识别方法研究[D].石家庄:河北科技大学,2019.
- [41] LI Y, YUAN Y. Convergence analysis of two-layer neural networks with ReLU activation [C]//Neural information processing system (NIPS). Long Beach, California, USA: Curran Associates Inc., 2017:30-35.