

Retrieval of Optical Remote Sensing Image Content Based on Convolution Neural Network

Ji Tengfei^{1,2}, Zhao Yameng², Yu Tao², Zhou Bing¹, Huang Xiangzhi²

1. School of Computer and Information Engineering, Henan University, Kaifeng 457000, China

2. Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100020, China

Abstract

The insufficiency of image detection method based on bottom feature extraction is analyzed and an optical remote sensing image content retrieval method based on convolution neural network is proposed. Firstly, convolutional neural network is used to train the rssc7-master remote sensing image data set of the training sample. Image features are extracted from the sample to build the image feature library. Then, Softmax classifier is used to classify the feature map to achieve precise classification and improve the accuracy of the model. Then, the model is tested with the help of the sample map or the actual optical remote sensing image map, and good experimental results are obtained. Experimental results show that this method is effective in the content retrieval of optical remote sensing images and has high retrieval accuracy.

Keywords: Remote Sensing Image Content Retrieval; Image Classification; Convolution Neural Network; Dropout; Softmax Classifier; RSSCN7 - master Data Set

基于卷积神经网络的光学遥感图像内容检索*

姬腾飞^{1,2}, 赵亚萌², 余涛², 周兵¹, 黄祥志²

1. 河南大学计算机与信息工程学院, 河南省 开封市 475000

2. 中科院遥感与数字地球研究所, 北京市 朝阳区 100020

摘要:分析了基于底层特征提取的图像检测方法的不足,提出了一种基于卷积神经网络的光学遥感图像内容检索方法。首先使用卷积神经网络对训练样本 RSSCN7-master 遥感图像数据集进行训练,从样本中提取图片特征,构建图片特征库,然后借助 Softmax 分类器进行特征图的分类,以达到精确分类的目的,提高模型的准确度。然后借助样本图或者实际光学遥感影像图对生成的模型进行检测,取得了不错的实验结果,测试准确率较高。实验表明该方法在光学遥感图像内容检索上的有效性,具有较高的检索准确度。

关键词: 遥感图像内容检索; 图像分类; 卷积神经网络; Dropout; Softmax 分类器; RSSCN7-master 数据集

引言

随着我国遥感卫星行业的迅猛发展,和不断进步的遥感技术水平,遥感行业所取得的成果也越来越丰硕,光学遥感图像就是遥感行业的主要成果之一,其具有海量数据,覆盖面积大、范围广、时间连续性高的特点。当前,遥感数据的用户群体其类别和数量不断增加,为了更好更充分地利用这些光学遥感图像,满足应用需求的日益多样化,对于遥感图像内容检索办法提出了更高的要求。

人们对于遥感图像的检索方法进行了各种各样的探索,例如张鑫[1]利用等面积的矩形环来提取颜色特征进而进行对图像的检索;郭庆文[2]根据等价关系 R 提取图像区域的特征向量,进而对图像进行检索分类;孙君顶

*十三五民用航天预研项目“天地一体化卫星应用评价技术”(课题号 Y7K00100KJ);国家重点研发计划项目“弱光照海洋水色遥感探测机理与模型研究”(课题号: Y6AC260010);所创新项目“全球地球综合观测遥感产品与验证”(课题 Y7SG0600CX)

[3]提取图像的颜色特征并结合其空间特征,进行相似性彩色图像检索;宋琳琳[4]分别利用 HSI 分量和带灰度查分基元矩阵提取颜色特征和纹理特征,进而实现图像检索等,可见早期的研究主要集中在色彩、形状和纹理等底层特征的提取上。但是这些基于底层特征提取的图像检测方法难以准确、完善的表达人们对于图像特征的真实诉求,即所谓“语义鸿沟”存在于底层特征与高层语义之间。基于这一问题,温超[5]提出用多层次的反馈方法处理图像低层特征与高层语义间的联系,反馈机制的提出与引入,如支持向量机(Support Vector Machine, SVM),在解决小样本、非线性及高维模式识别中表现出许多特有的优势,常用来处理用户反馈信息,但是核函数是 SVM 的关键所在,核函数的不同就决定了其 SVM 算法的不同,限制了 SVM 的应用。随着遥感技术的迅速发展,遥感图像不仅数量持续增高,而且质量也得到了大幅提高,具有尺度多样性、视角特殊性,背景复杂度高等特点,使得现有的基于内容的图像检索方法面临新的挑战。

1. 相关技术介绍

近年来,在机器学习领域,深度学习成为研究的热点,而卷积神经网络(Convolutional Neural Network, CNN)是基于深度学习理论的一种人工神经网络,其特殊的网络结构具有对复杂背景、光照、角度变化等不敏感的特性,近来广泛地应用于计算机视觉中,包括检测、分割、分类等任务,为图像特征的获取提供了新思路,获取到的特征能够更好地表达图像所包含的信息:晁安娜[6]等提出一种改进的基于卷积神经网络的遥感图像目标识别算法,克服了视角变化、噪声、遮挡等因素的影响,实现了遥感图像中飞机目标识别;李亚飞[7]等提出了一种基于卷积神经网络的多源多特征融合的方法,将遥感图像的空间结构、光谱、纹理等特征进行有效融合,实现了遥感图像分类;黄洁[8]等利用卷积神经网络,避免了复杂的特征选择和繁琐的提取过程,这是利用卷积神经网络可自主提取图像特征进行学习的优点,实现了在复杂海况背景图像下的舰船目标检测。相比传统的分类方法,卷积神经网络直接以图片为输入,对其进行特征提取和模型训练,得到遥感图像的分类信息,快速高效。

本文学习前人对卷积神经网络的研究思路,提出了一种基于卷积神经网络的光学遥感图像内容检索方法。该方法利用卷积神经网络的卷积核,对输入的图像进行自动的特征提取,然后其权值共享和池化的特性,以及 Dropout 方法的使用,保证了网络的有限参数,精简网络,然后利用 Softmax 分类器进一步明确并缩小检索范围,从而提高了检索的速度和准确性。

2. 模型方法

本文遥感图像内容检索算法的工作流程如图 1 所示,模型基于 Tensorflow 框架,使用 Python 语言进行实现。TensorFlow 是谷歌基于 DistBelief 进行研发的第二代人工智能学习系统,相比于其他深度学习框架,如 Keras、Caffe、Torch、Theano 等,Tensorflow 框架使用简单,具有可自行设计神经网络、支持自动求导等优点。

这个方法中,首先利用搭建好的卷积神经网络结构去训练预先准备好的数据集,称为模型的训练过程,在这个过程中,通过不断的训练,以实现对数据集的参数学习,即对数据集进行特征收集,构建其数据集特有的数据集特征库,直到网络模型收敛或是达到了预先设定的训练步数为止。然后,输入检测图像,根据检测图像的特征和类别,对训练得到的模型进行准确率检测,以验证取得的特征库的准确,进而证明所构建网络的准确性,实现基于卷积神经网络的光学遥感图像内容检索,达到实验的目的。

2.1 卷积神经网络

卷积神经网络是当下流行的一种人工神经网络,因其具有图像与卷积网络的结构可以完美融合、特征采集与分类同时进行、以权值共享来简化参数数量,结构简单适应性强的特点,所以卷积神经网络正广泛应用于自算计图像识别,经过扩展的卷积神经网络也应用于视频分析和自然语言的处理,用卷积神经网络处理图片并不针对每个像素进行处理,而是以图片上每一小块的区域为单位,这样做加强了获取的图片信息的连续性,使神经网络能“看”到一片区域,而不是一个点,加深了网络对图片的理解,当然卷积神经网络看到的也不是色彩缤纷的图片,而是一排排的数字该网络,比较典型的应用便是阿尔法围棋(AlphaGo)的诞生。

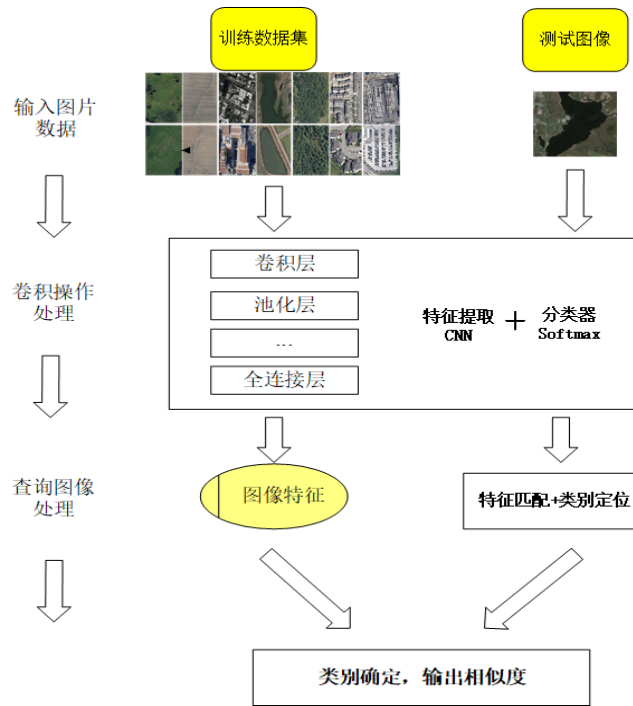


图 1 算法流程图

2.1.1 网络结构

Lenet-5 是一种典型的卷积神经网络，它包含了深度学习的基本模块：卷积层，池化层，全连接层，是深度学习模型的基础，本文就是在学习 Lenet-5 神经网络的基础上，经过自己的学习与模拟，得到的一种训练模型，通过实验数据进行验证，取得了相对不错的结果。LeNet-5 共有 7 层，不包含输入层，每层都包含可训练参数；每个层有多个 Feature Map，每个 Feature Map 通过一种卷积滤波器提取输入的一种特征，然后每个 Feature Map 有多个神经元。其网络模型如图二所示。

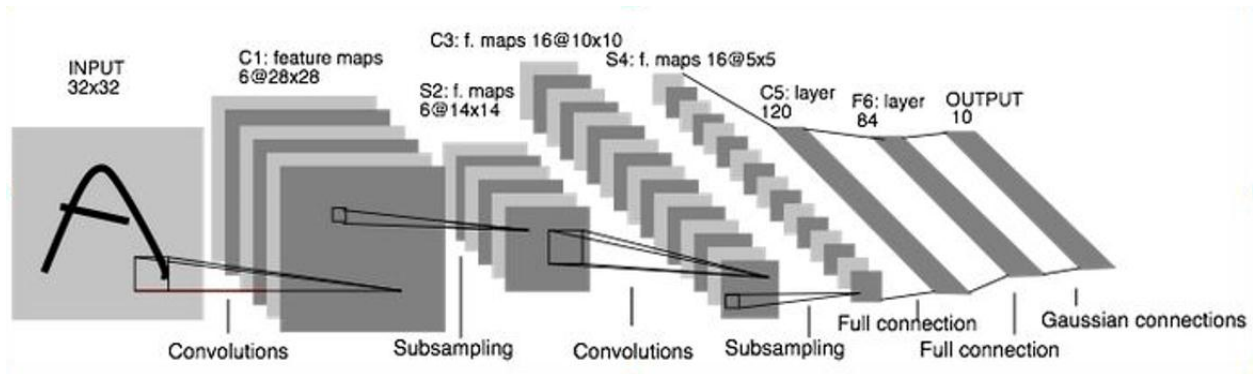


图 2 Lenet-5 网络模型结构图

由图二可以看到 Lenet-5 的 7 层网络模型，本文实验采用的卷积神经网络结构大体和 Lenet-5 网络差不多，分为了卷积层，池化层，卷积层，池化层，全连接层，全连接层和 Softmax 层。卷积层由若干卷积核组成，用来获取图片的不同特征，原则上每种卷积核都能提取到一种不同的特征，卷积操作的过程可以用公式(1.1)进行表示：

$$X_j^l = f(\sum_{i \in M_j} x_i^{l-1} \times K_{ij}^l + b_j^l) \quad (1)$$

其中, x_j^l 表示第 l 层的第 j 个特征图, f 表示激励函数, M 是输入特征图的集合, $\times$ 表示卷积运算, K 表示卷积核, b 表示偏置项。经过卷积后要进行池化操作, 池化的作用是降低模型的参数数量和模型的复杂度, 池化层通常也叫做降采样层, 池化操作分为最大池化法, 平均池化法和随机池化法, 这里采用的是最大池化。

2.1.2 激活函数

在神经网络中, 我们采用卷积的方式来处理图像, 就是对每个卷积的图像赋予一个权值, 这是一个线性的操作。但对于我们的样本来说, 却不一定是线性可分的, 所以我们需要将样本进行线性变化, 或者引入非线性因素来贴合样本, 弥补线性模型所不能解决的非线性问题, 激活函数便由此而生, 增加模型的非线性因素。

激活函数有 Sigmoid 函数、Tanh 函数和 ReLU 函数等, 目前最常用的是 ReLU 函数。Sigmoid 函数, 如公式 (1.2) 所示, 由于其饱和时梯度值非常小和输出值不是以 0 为中心值的缺点, 目前已经被淘汰了:

$$f(x) = \frac{1}{1+e^{-x}} \quad (2)$$

Tanh 函数中, 如公式(1.3)所示, $\sigma(x)$ 为 Sigmoid 函数, 仍然具有参数饱和的问题:

$$\tanh(x) = 2\sigma(2x - 1) \quad (3)$$

ReLU 函数, 如公式(1.4)所示, 该函数很大程度的解决了优化深层神经网络时的梯度耗散问题, 其函数表示图像如图 3 所示:

$$f(x) = \max(0, x) \quad (4)$$

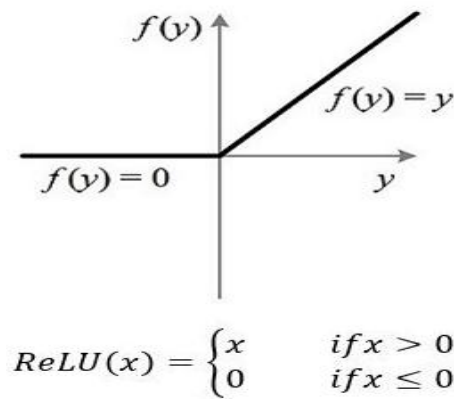


图 3 ReLU 函数图像

2.1.3 Dropout 层

本文的网络结构上设置有 Dropout 层。由于之前卷积核的卷积操作, 网络模型从图片中学习了大量的特征, 虽然有池化操作来简化网络的参数数量, 但为了保证自己的网络具有良好的泛化能力, 保证网络的鲁棒性, dropout 的设置具有必要性, 其工作原理就是在训练的时候当需要进行全权更新时, 通过设定 dropout 概率让一部分神经节点随机失活。

2.2 Softmax 回归

本文神经网络的最后是 softmax 层, 也就是分类专用的层, 使用一个概率来表示待分类对象有多大概率属于某个类, 在使用的时候需要给出输入样本特征 x 和相应的样本类别 y , 假设有 K 个类别, 则有 $y \in \{1, 2, 3, \dots\}$,

K}。对于输入的 x ，我们需要计算它属于每一个类别 j 的概率值 $p(y = j|x)$ ，所以在 Softmax 回归中，假设函数 $h_{\theta}(x)$ 将会输出一个 k 维的向量来表示输入属于每一个类别的概率。定义假设函数 $h_{\theta}(x)$ 如公式 (1.5) 所示：

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k \exp \theta_j^T x^{(i)}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (5)$$

其中， $\frac{1}{\sum_{j=1}^k \exp \theta_j^T x^{(i)}}$ 是归一化系数，保证右面概率值和为 1， θ 是需要学习的模型参数向量。定义如下的示性函数，如公式 (1.6) 所示：

$$\begin{aligned} 1\{\text{表达式的值为真}\} &= 1 \\ 1\{\text{表达式的值为假}\} &= 0 \end{aligned} \quad (6)$$

Softmax 回归的损失函数如下公式 (1.7) 所示：

$$J(\theta) = -\frac{1}{M} \left[\sum_{i=1}^m \sum_{j=0}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] \quad (7)$$

其中， m 表示样本数量， $1\{\cdot\}$ 是示性函数，即公式 (1.6) 所示。

本文在进行测试图像测试模型时候，采用训练好的 Softmax 分类器对测试图片进行分类处理，显示其所属类别。

3. 实验与结果

3.1 准备工作

前面说过，本文采用 Tensorflow 框架，使用 Python 进行编程实现，硬件环境为 i7-4790k CPU，4.00HZ，两块 NVIDIA Tesla K20c GPU，32G 内存；软件环境为 Windows 7 旗舰版，Tensorflow 框架安装在 Anaconda Navigator 集成环境上，程序使用该集成环境上的 Spyder 编译器进行处理，安装了 Numpy, OpenCV, Matplotlib 等相关支持组件。实验所用的数据来自于 RSSCN7-master 遥感图片数据集，该数据集由 2800 张 400*400 的 7 类遥感图像组成，分别是 aGrass、bField、cIndustry、dRiverLake、eForest、fResident、gParking，每一类有 400 张，格式为 jpg。

3.2 实验阶段

实验的处理流程如图 4 所示：

实验设置的最大迭代次数为 10000，初始学习率设置为 0.0001，每 2000 此调整一次学习率，每 50 迭代显示一次训练情况。在预处理的时候，由于内存的限制，需要设置批次，并且设置每批读取的图片个数，这里设置的是 16，然后生成队列，并且将图片队列与标签队列相对应，然后将这些 list 转化成 Tensorflow 能够识别的格式进行输入。这些在 inpu_data.py 中实现。然后进行卷积神经网络的结构设计，其中卷积层使用的卷积核的大小为 5*5，包括池化层全连接层等，这些都在 model.py 中完成。然后模型的训练和测试则分别由 training.py 和 test.py 中完成。

3.3 模型参数设计分析

影响模型准确率的关键在于卷积神经网络模型结构设计以及其网络参数的设置，所以在进行网络参数确定的时候特进行了对比试验。由前人的实验可以得到，池化可以简化网络模型的参数，Dropout 层可以优化网

络的性能，防止其过拟合，所以这里的测试主要用于确定网络结构及卷积核的尺寸。这里采用最大池化法，其池化层大小为 2*2，Dropout 的阈值设为 0.25, 其具体如表 2 所示：

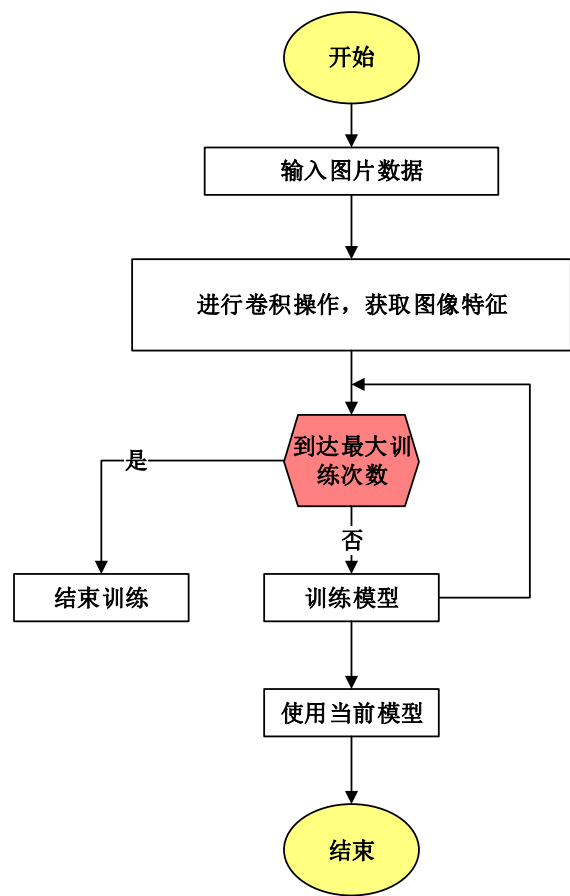


图 4 完整实验流程

表 2 模型结构与参数对比分析

	模型一	模型二	模型三	模型四
网络结构	Conv(16*3*3)	Conv(32*3*3)	Conv(32*5*5)	Conv(64*5*5)
	Pooling(2*2)	Pooling(2*2)	Pooling(2*2)	Pooling(2*2)
	Dropout(0.25)	Dropout(0.25)	Dropout(0.25)	Dropout(0.25)
	Conv(32*3*3)	Conv(64*3*3)	Conv(64*5*5)	Conv(128*5*5)
	Pooling(2*2)	Pooling(2*2)	Pooling(2*2)	Pooling(2*2)
	Dropout(0.25)	Dropout(0.25)	Dropout(0.25)	Dropout(0.25)
训练时间 (h)	12.82	11.41	8.41	8.86
分类准确度	0.964436	0.966103	0.964723	0.975146

表 2 中，可以看到其四个模型的准确率相差不大，但是由于本次实验所用的遥感影像数据集较大，其影像模型选择的主要因素就是训练时间因素。从模型二和模型三的对比可以看到，卷积核尺寸为 5*5 的模型，有相对更短的训练时间，从模型三和模型四看到，卷积后得到 64 个特征图的时间要长于 32 的时间，但其结果准确率却要高出 0.1，所以实验中卷积神经网络结构最终采用模型四。

3.4 实验结果及分析

结合训练模型结构进行测试，得到了训练模型的准确率趋势折线图和误差折线图，分别如图 5，6 所示，可以看到，模型在 700 步的时候就已经趋于收敛了，但是由于训练次数少，模型处于不稳定状态，直到 Step5000 的时候，模型趋于稳定的状态，让其继续按照模型设置的次数进行训练，得到最终的训练模型。

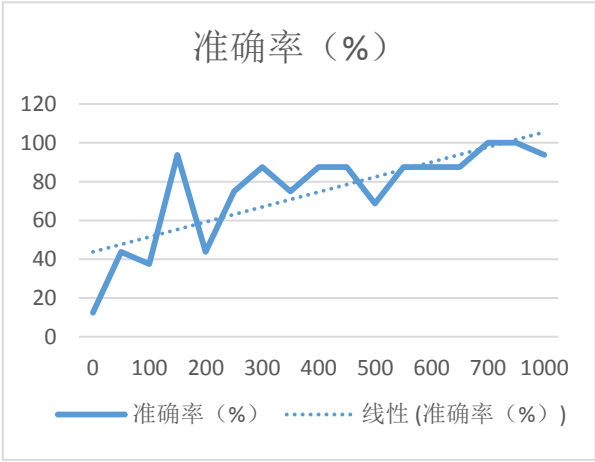


图 5 准确率趋势示意图

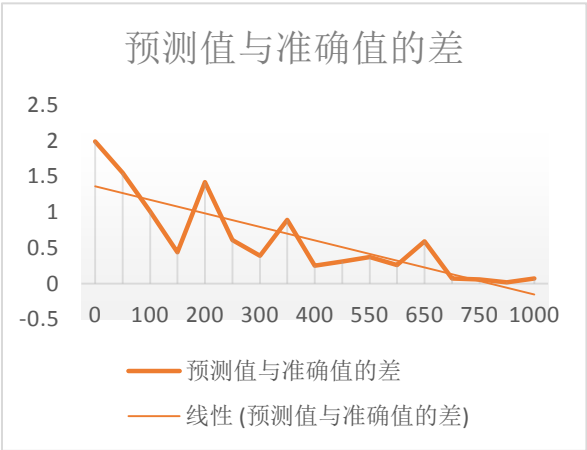


图 6 误差变化趋势示意图

模型训练结束，得到其训练模型，然后进行模型的测试工作:利用测试图片对模型进行测试，测试图片如图 7 (a)、(b) 所示，可以从数据集中提取图片做测试，也可以自己准备。

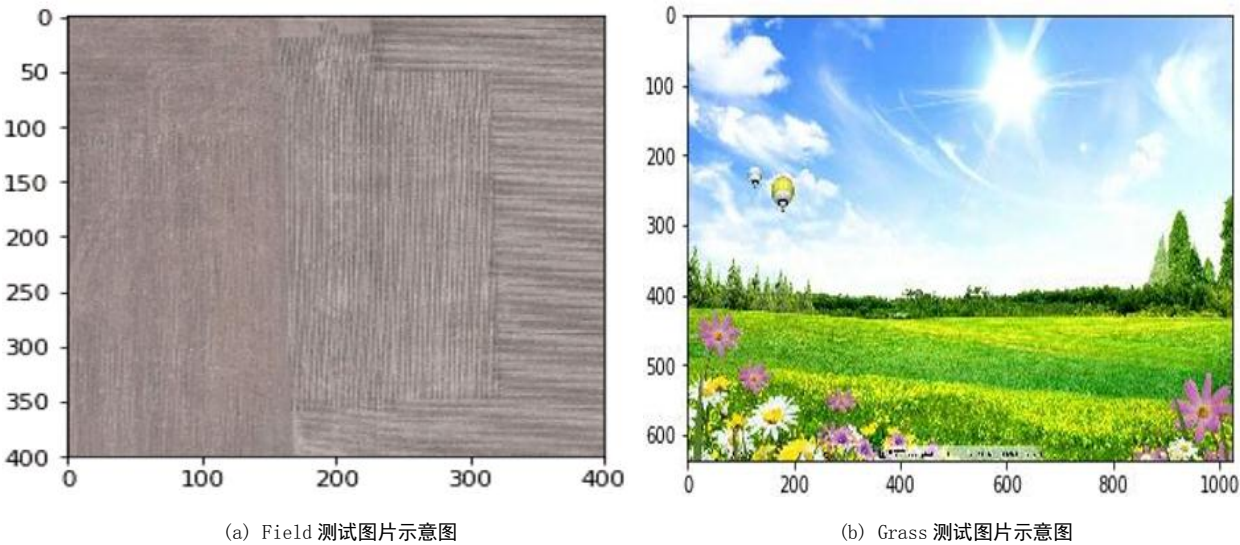


图 7 测试图片展示

其中，图 (a) 通过训练模型的加载，测试结果为：Field 的概率是 0.997538；图 (b) 为：Grass 的概率是 0.984161，可以看到测试结果准确率达到预期。

为了更加直观的定量分析本文的实验结果，特统计实验数据如下表 1 所示，从中可以看出，随着测试数量的增多，其测试准确度的下降在可接受的误差内，当测试数据达到 40 次时，其平均准确率达到了 96.1831%，通过分析数据集发现，其 fResident 因类似于 cIndustry，导致其准确率不是十分理想，但其他分类物识别精度都很高，除开 fResident 类后其他类平均准确率达到了 97.4967%。

表 1 数据集检索准确度结果

测试图像数	aGrass	bField	cIndustry	dRiverLake	eForest	fResident	gParking
10	1	1	0.980010	0.943326	0.983326	0.913398	0.983151
20	0.984230	1	0.979873	0.941139	0.978812	0.925533	0.985001
30	0.985531	1	0.978622	0.938971	0.970129	0.896298	0.971166
40	0.984161	0.997538	0.976839	0.937482	0.969909	0.883010	0.977875

在与其他基于底层特征检索方法对比的过程中，如与基于颜色和纹理特征的检索方法相比，本文方法也取得了一定的优势，其中基于颜色特征检索采用直方图的形式，纹理特征则采用灰度共生矩阵，其效果如图 8 所示，可以直观看到对 7 类地形图像的检索，本文方法均具备明显优势。

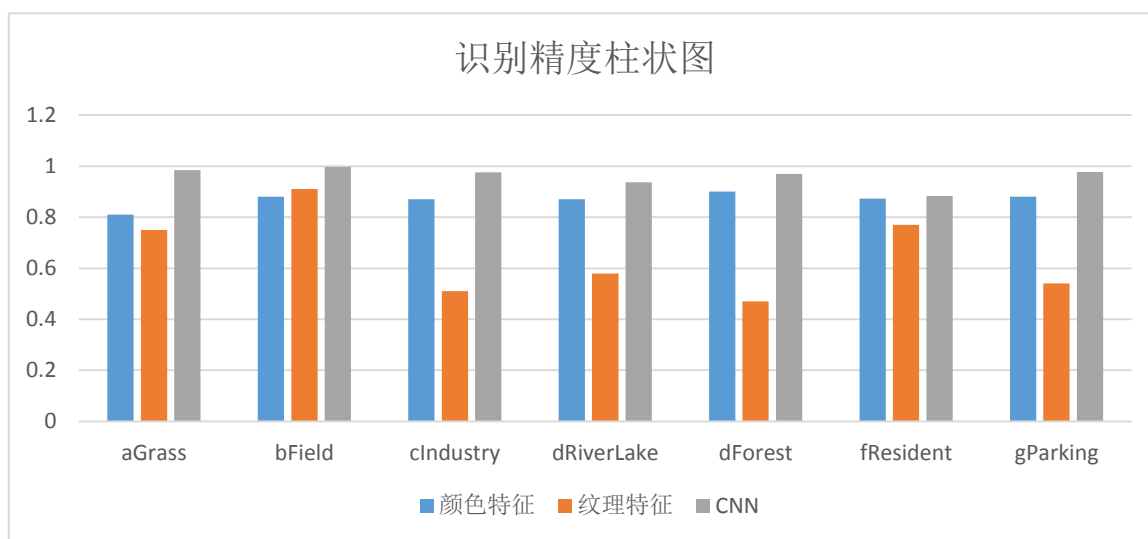


图 8 遥感地物分类结果对比图

4. 结论

本文方法是作用于使用卷积神经网络处理 RSSCN7-master 遥感图片数据集，通过实验证明，使用卷积神经网络可以有效处理遥感光学图像的分类问题，给我们在之后进行光学遥感图像内容检索时候提供了一种思路。本文方法在取得不错准确率的前提下，通过实验可以看到，其训练时间相对较长，所以如何进一步优化模型，使其训练速度进一步提高，以及增强本文方法的实用性，将成为下一步自己的研究方向。

REFERENCES

- [1] 张鑫, 温显斌, 等. 基于颜色特征的图像检索方法研究 [J]. 计算机科学, 2012, 39 (11): 243-245
- [2] 郭庆文, 王国胤, 等. 多粒度的图像检索方法研究 [J]. 计算机科学, 2013, 40 (1): 298-301
- [3] 孙君顶, 崔江涛, 等. 基于颜色和形状特征的彩色图像检索方法 [J]. 中国图象图形学报, 2004, 9 (7): 820-827
- [4] 宋琳琳, 王相海. 小波域颜色和纹理特征提取及图像检索方法研究 [J]. 小型微型计算机系统, 2009, 30 (10): 2078-2084
- [5] 温超, 耿国华. 基于内容图像检索中的“语义鸿沟”问题 [J]. 西北大学学报(自然科学版), 2005, 35 (5): 536-540
- [6] 晁安娜, 刘坤. 基于卷积神经网络的遥感图像飞机目标识别 [J]. 微型机与应用, 2017, 36 (22): 66-69, 73.
- [7] 李亚飞, 董红斌. 基于卷积神经网络的遥感图像分类研究 [J]. 智能系统学报, DOI: 10.11992/tis.201706078.
- [8] 黄洁, 姜志国, 等. 基于卷积神经网络的遥感图像舰船目标检测 [J]. 北京航空航天大学学报, 2017, 43 (9): 1841-1848

【作者简介】



¹姬腾飞（1992-），男，汉族，硕士研究生，
研究方向：计算机应用技术
Email: Jtf6666666@163.com