

文章编号: 2095-2163(2020)10-0023-04

中图分类号: TP394.1

文献标志码: A

# 基于卷积神经网络的植物叶片树种识别研究与实现

边缘, 孔小莹, 张莉, 边世正, 李瑞改

(东北林业大学 信息与计算机工程学院, 哈尔滨 150040)

**摘要:** 随着人工智能的迅速发展, 深度学习方向的算法性能逐渐提高, 推动了深度学习在各个领域的应用。本文使用卷积神经网络算法建立树种识别模型, 以叶片作为模型输入数据。本文所建立的模型在公开的 Flavia 数据集中的识别准确率在 90% 以上, 达到了应用要求, 本模型的设计对林业有一定的实际应用价值。

**关键词:** 深度学习; 卷积神经网络; 树种识别

## Research and Realization of Wood Species Recognition Based on Convolutional Neural Network

BIAN Yuan, KONG Xiaoying, ZHANG Li, BIAN Shizheng, LI Ruigai

(College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China)

**[Abstract]** As the field of artificial intelligence develop rapidly, the performance of deep learning algorithm is constantly improved, the application of deep learning in various fields is promoted greatly. In this paper, convolution neural network algorithm is used to establish tree species identification model, and leaves are used as model input data. The model established in this paper Flavia The recognition accuracy of data set is 90% Above, the application requirements are met. The design of this model has certain practical application value to forestry.

**[Key words]** Deep learning; Convolutional neural network; Wood species recognition

### 0 引言

树种识别一直是林学的热门研究内容<sup>[1]</sup>。由于叶片图像具有叶片形状特征多、辨识度高、区分度大、数据采集容易等优点, 因此是当前树种识别领域常用的数据源。但人工树种识别存在着工作量大、工作效率低等问题, 因此, 本文采用模式识别与人工智能等领域的技术进行基于叶片图像的自动分类, 具有相当大的应用前景。

由于深度学习领域大量研究者不断对已有算法进行优化, 卷积神经网络算法的性能、准确率以及对不同问题的适应性均有提高, 因而被大量应用在各个领域的图像分类问题上。在以往树种识别的研究中, 有基于传统特征提取的分类器算法、基于人工神经网络的深度学习分类器算法以及传统特征提取与人工神经网络相结合的复合分类算法。唐钦设计的基于 K-means 和 SVM 传统分类器的复合算法在自采集的 22 种 8 800 张树叶样本数据集上取得了 89.64% 的识别率<sup>[2]</sup>。但是传统分类器算法需要复杂的特征提取, 要求模型的使用者对识别任务的数

据有深刻的理解, 需要大量的先验知识和前期反复的数据分析、特征工程工作。且传统分类器算法可容纳的特征维度小、数量少, 识别难度远远大于深度学习, 在准确率方面也更容易遇到瓶颈。杨蒙蒙设计的 Hu 矩和灰度共生矩阵与 BP 神经网络结合的分类算法, 在 UCI 数据库 15 种树叶样本数据集上取得了 80.97 的识别率<sup>[3]</sup>。然而人为提取特征输入神经网络的操作, 所提取的特征不管从数量还是质量都远低于卷积神经网络自动提取的特征, 没有有效发挥神经网络模型的优势, 在准确率方面仍有较大的提升空间。赵鹏超等设计了一种浅层卷积神经网络, 在自采集的 4 种树叶数据集上的识别率在 95% 以上<sup>[4]</sup>。由于其处理的分类任务种类少、样本数量不足、难度低, 因此缺乏足够的说服力; 也因为种类和样本数量的不足使得模型难以被有效应用在现实情境中。上述模型均为浅层网络模型, 有一定的参考意义, 也各自有其不足之处。在 ILSVRC 赛事的影响下, 深度学习蓬勃发展, 浅层的卷积神经网络已逐渐被深层卷积神经网络所取代<sup>[5]</sup>。深层卷积神

**基金项目:** 2019 年度东北林业大学省级创新项目 (201910225226, SJGY20170145)。

**作者简介:** 边缘 (1998-), 男, 本科生, 主要研究方向: 机器学习; 李瑞改 (1979-), 女, 硕士, 讲师, 主要研究方向: 3S 技术及其林业应用。

**通讯作者:** 李瑞改 Email: lirg751@163.com

**收稿日期:** 2020-05-12

神经网络在模型可容纳的原始输入数据规模、所需训练的参数量、模型的训练难度以及实际的模型准确率等各方面的表现都优于浅层网络。

## 1 卷积神经网络

### 1.1 概念

卷积神经网络 (Convolutional Neural Networks, CNN) 是一种特殊的多层前馈神经网络<sup>[6]</sup>。卷积神经网络被广泛应用在以二维数据为输入的任务中,有着相对固定的网络结构。其起源于上世纪 60 年代 Hubel 和 Wiesel 等受猫的大脑视觉皮层系统启发<sup>[7]</sup>。视觉皮层的研究成果,为深入研究神经网络等相关内容提供了理论研究方向,为卷积神经网络此后的不断发展奠定了基础,CNN 在 1990 年首次实现。LeCun 等建立了反向传播算法训练的卷积神经网络模型 LeNet<sup>[8]</sup>。该网络模型最初是为了研究识别手写的邮编数字,被作为经典的卷积神经网络被后人不断研究,并在此基础上优化。与传统分类算法相比,CNN 可以不经过人工的预处理或预先提取特征,直接从原始数据中自动提取特征识别图像种类,对不同种数据都有很强的适应,使得特征提取工作不再完全依赖于相应行业专家丰富的专业知识与经验。同时 CNN 具有局部连接与权值分享的特点<sup>[9]</sup>。这二种特点对于降低网络模型的复杂度起到重要作用,很大程度上减小了模型所需的参数的规模和训练过程中的运算量,使得模型拥有更快的训练速度和较容易的训练难度。

### 1.2 网络结构

卷积神经网络由卷积层、池化层以及全连接层组成,网络结构示意图如图 1 所示。不同于输入一维数据的传统神经网络,CNN 模型以二维数据(即图像)作为输入,通过卷积层提取特征,减少参数数量;通过池化层降低卷积层输出特征图像的维数,即降维处理;最后提取到的特征图像输入到全连接层中进行分类。CNN 具有强大的特征提取能力,参数较少,卷积层与池化层的网络结构可以被设计得很深。

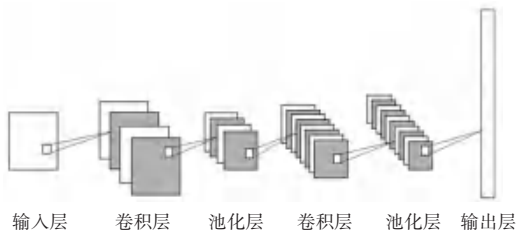


图 1 卷积神经网络结构示意图

Fig. 1 Structure diagram of convolutional neural network

### 1.3 卷积

卷积运算的原理是用一个卷积核,遍历原图像,得到特征图像。其中卷积核是一个预先设定好大小的权值矩阵  $w$ ,卷积运算即卷积核与原图像局部的对应值乘积的和加上偏置  $b$ ,经过激活函数计算得到输出的特征值,在卷积核遍历原图像后得到特征图像。

定义输入矩阵  $X(n \times n)$ ,卷积核为  $W(k \times k)$ ,输出矩阵为  $Y$ ,则卷积运算公式(1)为:

$$y_{ij} = \sum_{b=0}^k \sum_{a=0}^k x_{(i+1)(j+b)} \cdot w_{ab} \quad (1)$$

卷积核的权值通过反向传播算法训练,因此 CNN 具有学习图像特征的能力。

### 1.4 池化

卷积操作所提取出的原图像对应的特征图像实际上不需要经过池化操作即可用于全连接层分类。但碍于计算机能力的限制,直接输入到全连接层的特征图像数据量过大,需要训练大量参数,导致训练时间过长,训练难度陡升等问题。

为了解决数据量过大的问题,在卷积神经网络结构中添加了池化层。

池化层通过池化操作来减小数据量。池化操作利用了图像数据相邻像素之间有关联性的原理,通过预先设定好大小为  $n \times n$  尺寸的正方形(以像素为单位)遍历图像,在原图像的每个正方形内分别计算其结果,计算方式有求平均值和最大值二种,最后输出一个由计算结果所组成的特征图像,图像的长宽均为原图像长宽除以  $n$ 。2×2 最大池化原理示意,如图 2 所示。

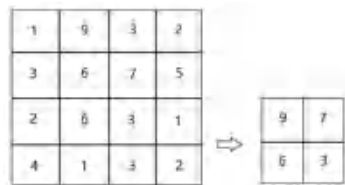


图 2 2×2 最大池化原理示意图

Fig. 2 Schematic diagram of 2×2 maximum pooling principle

在实际应用中,常见的池化操作分为最大池化与平均池化二种。池化操作可以有效减小网络所需的参数量,同时最大限度保持特征的质量。在过往的研究中最大池化在很多实验中都取得了较好的效果。本论文采用的是 2×2 最大池化,一次池化操作可以将数据的规模缩小至原来的四分之一。

## 2 数据集与预处理

实验选用 Flavia 植物叶片数据集,如图 3 所示。该数据集共有 32 种植物类别,1 907 个尺寸均为 1600×1200 样本图像,图像背景纯白无阴影。

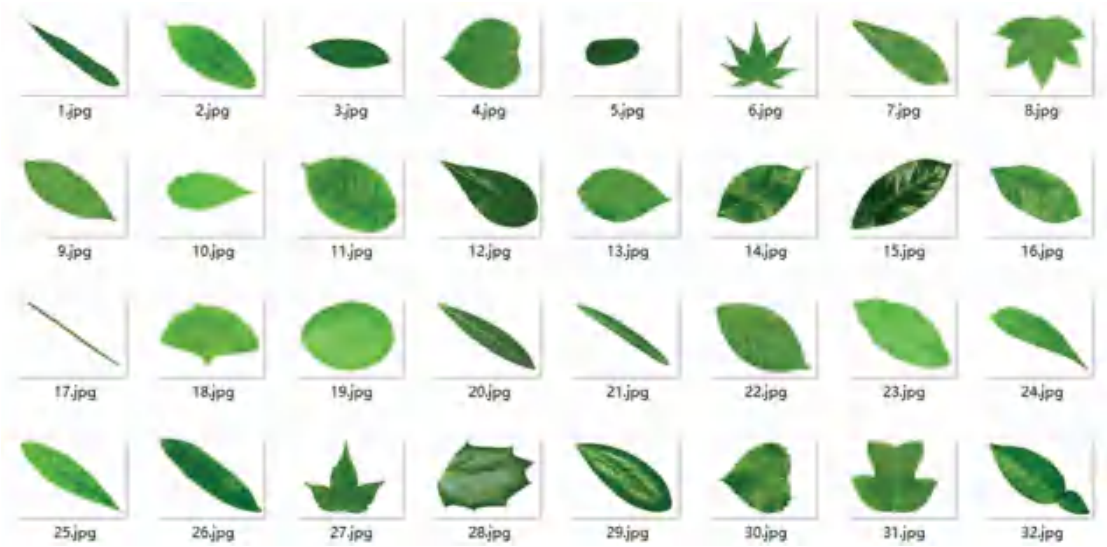


图 3 Flavia 植物叶片数据集

Fig. 3 Flavia leaf data set

树叶主要通过形态特征区分种类, 颜色影响较小, RGB 三通道图像极大地增多了数据量, 增加了神经网络的负担, 因此实验将 Flavia 的 RGB 叶片图像转换为灰度图, 如图 4 所示, 保留树叶形态特征、减少数据量, 使数据集可以使用更为复杂的网络模型。



图 4 RGB 叶片图像转换灰度图例

Fig. 4 RGB leaf image conversion grayscale legend

RGB 图像转换灰度图为公式 (2) :

$$L = R \times 0.299 + G \times 0.587 + B \times 0.114. \quad (2)$$

其中, R、G、B 分别代表 RGB 图像的红、绿、蓝 3 个通道的图像, L 代表灰度图。

对叶片数据标准化处理, 使得神经网络在训练时加速收敛, 如公式 (3) 所示:

$$x_{scale} = \frac{x - \mu}{\sigma}. \quad (3)$$

其中,  $\mu$  为样本数据的均值;  $\sigma$  为样本数据的标准差;  $x$  为图像矩阵;  $x_{scale}$  为标准化处理后得到的图像矩阵。

### 3 网络模型与训练

本实验软件环境为 Windows 10 64 位系统, 采用 Keras 深度学习框架, 使用 Python 作为编程语言。计算机内存为 8GB, 搭载 Inter (R) Core (TM) i5-7400 CPU @ 3.00GHz x4 处理器, 同时使用 Nvidia

GTX1050 显卡加速网络训练。

本文设计的网络模型共有 7 层卷积层, 4 层池化层以及 2 层全连接层, 如图 5 所示。其中卷积层选用了  $11 \times 11$ ,  $5 \times 5$  和  $3 \times 3$  种尺寸的卷积核, 池化层均选用  $2 \times 2$  最大池化层, 第一层全连接层设置了 4 096 个神经元, 第二层全连接层输出 32 种植物叶片分类结果。

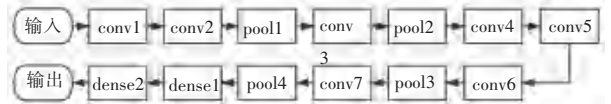


图 5 本文设计的网络模型

Fig. 5 The network model designed in this paper

本实验使用交叉验证对模型的性能进行验证, 训练集与验证集比例设置为 8 : 2, 共 1 525 张训练样本, 以及 382 张测试样本, 为了保障实验的准确性, 训练样本与测试样本均随机选取。训练过程中采用了衰减的学习率, 设置初值为 0.001, 每两轮若验证集 loss 值未下降, 则学习率衰减为原来的一半。

### 4 结果与分析

经过多次交叉验证, 模型识别的平均准确率为 91.16%, 最低准确率为 90.31%, 如图 6 所示; 最高准确率为 92.15%, 如图 7 所示, 模型的平均训练时间约为 36~37 min。准确率受到随机选取的训练集样本因素的影响, 同时 Flavia 数据集中不同示例图片中叶片方向的不一致也是可能影响模型识别准确率的因素之一。但多次实验中准确率均高于 90%, 能够达到林业相关应用领域中对树种识别技术的要求。

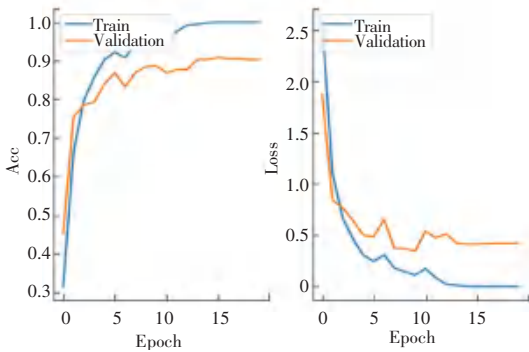


图6 准确率最低模型训练情况

Fig. 6 Model training with the lowest accuracy

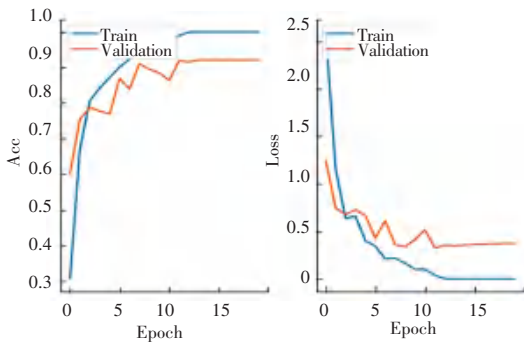


图7 准确率最高模型训练情况

Fig. 7 Model training with the highest accuracy

## 5 结束语

本文基于卷积神经网络设计了一种植物叶片树

种分类算法,该算法利用不同尺寸的卷积核随网络深度由大尺寸到小尺寸进行卷积以识别局部特征,然后利用 $2\times 2$ 最大池化核池化操作以降低数据规模并提高算法的鲁棒性,并在 Flavia 数据集上进行了模拟实验,模型准确率高于 90%。表明本文算法是一种基于植物叶片进行树种分类的有效方法。

## 参考文献

- [1] 陈明健,陈志泊,杨猛,等. 叶片传统特征和距离矩阵与角点矩阵相结合的树种识别算法[J]. 北京林业大学学报, 2017, 39(2):108-116.
- [2] 唐钦. 基于纹理和颜色特征的植物叶片识别方法研究[D]. 杭州:浙江大学,2015.
- [3] 杨蒙蒙. 基于混合生物地理学优化算法和 BP 神经网络的树叶分类系统[D]. 南京:南京师范大学,2018.
- [4] 赵鹏超,戚大伟. 基于卷积神经网络和树叶纹理的树种识别研究[J]. 森林工程,2018,34(1):56-59.
- [5] 周飞燕,金林鹏,董军. 卷积神经网络研究综述[J]. 计算机学报,2017,40(6):1229-1251.
- [6] 卢宏涛,张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理,2016(1):1-17.
- [7] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex [J]. The Journal of Physiology,1962,160(1):106-154.
- [8] LECUN Y, DENKER J S, HENDERSON D, et al. Handwritten digit recognition with a back-propagation network [C]// Advances in Neural information Processing Systems. Colorado, USA:[s.n.],1990:396-404.
- [9] 吕恩辉. 基于卷积神经网络的图像分类研究[D]. 徐州:中国矿业大学,2019.

(上接第 22 页)

## 4 结束语

针对说话人确认系统中现有得分规整方法的不足,本文提出基于对数似然值归一化(LLN)的得分规整方法。对每个测试对得分,充分利用其测试语音与集中所有说话人模型的得分做出规整,使同一测试语音对目标说话人模型和非目标说话人模型的得分差距拉大;该方法不需要预留额外开发集来估计规整参数,在后端得分域即可进行,因此也不受限于系统所使用的特征参数和模型;与已有的 Znorm、Tnorm 和 ZTnorm 得分规整方法能够很好互补,使不同测试语音或不同说话人模型的得分分布一致的同时,扩大“True”和“False”两类得分距离,在系统的统一门限下,获得更好的确认准确率,使说话人确认系统的性能进一步提高。

## 参考文献

- [1] AISIKAER R, WANG D, LI L, et al. Score domain speaking rate normalization for speaker recognition[J]. Qinghua Daxue Xuebao/journal of Tsinghua University, 2018, 58(4):337-341.
- [2] SINGH S. Support Vector Machine Based Approaches For Real

- Time Automatic Speaker Recognition System [J]. International Journal of Electrical & Computer Engineering, 2018, 8(5):2804.
- [3] LEU F Y, LIN G L. An MFCC-Based Speaker Identification System [C]// IEEE International Conference on Advanced Information Networking & Applications. Taipei: IEEE, 2017: 1055-1062.
- [4] SELMA Ozaydin. Design of a text independent speaker recognition system[C]// International Conference on Electrical & Computing Technologies & Applications. Ras Al Khaimah: IEEE, 2017:1-5.
- [5] LEI Z, LUO J, WAN Y, et al. A Mahalanobis Distance Scoring with KISS Metric Learning Algorithm for Speaker Recognition [C]// Lecture Notes in Computer Science. Switzerland: Springer, Cham, 2015.
- [6] 古斌,郭武. 说话人确认中基于无监督聚类的得分规整[J]. 数据采集与处理, 2019, 34(5):837-843.
- [7] PISANI P H, POH N, ANDRE C.P.L.F. DE CARVALHO, et al. Score normalization applied to adaptive biometric systems [J]. Computers & Security, 2017, 70(9): 565-580.
- [8] ARONOWITZ H. Speaker recognition using matched filters[C]// 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Shanghai: IEEE, 2016:5555-5559.
- [9] MAURYA A, AGGARWAL R K. Speaker recognition for noisy speech in telephonic channel [C]// International Conference on Applied & Theoretical Computing & Communication Technology. Bangalore: IEEE, 2016: 451-456.