

DOI: 10.13332/j.1000-1522.20160035

基于分层卷积深度学习系统的植物叶片识别研究

张 帅 淮永建

(北京林业大学信息学院)

摘要:深度学习已成为图像识别领域的研究热点。本文以植物叶片图像识别为研究对象,对单一背景和复杂背景图像分别给出了优化预处理方案;设计了一个8层卷积神经网络深度学习系统分别对PI@antNet叶片库和自扩展的叶片图库中33 293张简单背景和复杂背景叶片图像进行训练和识别,并与传统基于植物叶片多特征的识别方法进行了比较分析。实验证明:本文提供的CNN+SVM和CNN+Softmax分类器识别方法对单一背景叶片图像识别率高达91.11%和90.90%,识别复杂背景叶片图像的识别率也能高达34.38%,取得了较好的识别效果。利用本文实现的分层卷积深度学习识别系统在数据量大而无法做出更多优化的情况下,叶片图像的识别率更高,尤其是针对复杂背景下的叶片图像,取得了极佳的识别效果。

关键词:植物识别;叶片图像;特征提取;支持向量机;深度学习

中图分类号:S718.4; TP391.4 **文献标志码:**A **文章编号:**1000-1522(2016)09-0108-08

ZHANG Shuai; HUAI Yong-jian. **Leaf image recognition based on layered convolutions neural network deep learning.** *Journal of Beijing Forestry University* (2016) 38(9) 108-115 [Ch, 22 ref.] School of Information, Beijing Forestry University, Beijing, 100083, P. R. China.

Deep learning has been recently becoming research hotspot in the field of image recognition. In this study, plant leaf images were used as recognition objects. Plant leaf images were divided into single and complex background and were treated by using image segmentation method. We design a deep learning system which includes eight layers of Convolution Neural Network (CNNs) to identify leaf images. And then the deep learning system was tested with 33 293 leaf sample images which come from PI@antNet libraries and our extending leaf libraries for image training and recognition. Compared with the traditional identification methods, the classifier provided in this paper has achieved better recognition effect. For CNN+SVM and CNN+Softmax recognition system, the simple background leaf recognition rate reaches 91.11% and 90.90%, and the recognition rate of complex background reaches 34.38%. The system has higher recognition rate for the large amount of leaf images with no more optimization, and has a higher recognition rate especially for the recognition of complex background images.

Key words plant recognition; leaf image; feature extraction; SVM; deep learning

近年来,基于图像分析的植物识别分类技术已成为植物信息学的研究热点。植物的分类与识别一般以植物的形态、纹理、颜色等特征为主要依据,即根据花、果实、枝干、茎、叶等器官的形态纹理特征进行分类。植物叶片纹理、颜色和形态结构各异,是区别植物物种的主要依据。植物叶片存活时间较长、四季变化分明、图像采集方便。通过叶片对植物进

行分类研究已成为目前许多学者的研究方向。

在国内,王丽君等^[1]开发了基于图像多特征融合的观叶种类识别系统,该系统提取叶片图像的颜色、形状和纹理特征,基于支持向量机(SVM)算法实现识别,识别率达到了91.41%。杨天天等^[2]针对7种柳属(*Salix*)植物叶片图像分析有效叶片特征数字化指标,通过单纯的叶片相关特征进行多特

收稿日期:2016-01-21 修回日期:2016-05-06

基金项目:森林景观及林业生产过程仿真关键技术研究(2015ZCQ-XX)。

第一作者:张帅。主要研究方向:图像处理。Email: piaofeitest@qq.com 地址:100083北京市海淀区清华东路35号北京林业大学信息学院。

责任作者:淮永建,教授,博士生导师。主要研究方向:虚拟植物。Email: huaiyj@bjfu.edu.cn 地址:同上。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

征融合比较,综合判别正确率达到了 90.8%。张宁等^[3]基于叶片的几何、纹理特征,提出了克隆选择算法融合 KNN 分类器的分类方法,有效地提高了识别率。邓立苗等^[4]针对玉米(*Zea mays*)叶片,单独进行了 48 个叶片特征的提取,通过支持向量机算法完成系统,识别率高达 96%。翟传敏等^[5]提出基于叶缘与叶脉分数维特征的植物叶片识别方法,将分形理论引入叶片识别领域。邹秋霞等^[6]完成了基于 Android 手机的植物叶片分类系统的应用。研究与应用并行,是未来发展的一个新方向。在国外, CLEF 论坛(Cross Language Evaluation forum)自从 2011 年起组织了多次植物图像训练分类识别比赛,并且在 2014 年的比赛中增加了植物不同器官以及复杂拍照情况识别条件,2015 年,识别植物物种库已经高达 1 000。其中在 ImageCLEF 2011 中, Villena-Román 等^[7]利用尺度不变特征转换(SIFT)进行识别研究,证明了 SIFT 在植物图像识别领域的有效性。Backes 等^[8]利用复杂网络算法(Complex Network method)分析叶片特征并进行识别,在对不同分辨率以及一定采样噪声(noisy)的叶片图像识别上有很好的效果。在 LifeCLEF 2015 中, Ghazi 等^[9]利用 PCANet 算法进行识别研究,其中在扫描叶片图像上取得了 90.49% 的识别率,但是针对普通叶片图像,依然只有 33.80% 的识别率。

尽管植物分类识别方法的研究已经取得了许多进步,但是依然存在许多问题。基于图像分析的叶片识别过程中,提取特征一般由人工分析确定,不同植物种属差异性没有得到解决,用同样特征识别不同植物数据集会产生差异性,同时在复杂背景拍摄情况下,传统识别方法识别率明显下降。

基于卷积神经网络的深度学习算法,可以自主学习叶片特征减少人工干预,对于复杂背景叶片图像能排除噪声干扰等,提高图像识别效率。本文基于卷积神经网络算法构建了一个 8 层的深度学习叶片识别系统,并且利用 Pl@ antNet 叶片库及自主扩展的植物叶片数据来训练样本数据,完成识别率测试。为了提高识别率,对单一背景和复杂背景的叶片图像分别给出了图像预处理方案,并将基于深度学习的识别系统与 SIFT 算子和叶片图像特征的多分类器识别系统进行对比分析,验证了算法的有效性。实验证明,本文提供的 CNN + SVM 和 CNN + Softmax 分类器识别方法对单一背景叶片图像识别率能够达到 91.11% 和 90.90%,识别复杂背景叶片图像的识别率也能达到 34.38%,取得了较好的识别效果。

1 叶片图像预处理

基于叶片图像进行识别,首先需要对叶片图像进行分割处理,以排除采样噪声,减少识别误差。分割是将叶片部分从背景中分离。虽然图像整体可以提供更多识别信息,但在大多数识别过程中,对主体识别物的干扰将会显著降低识别率^[10]。本文的叶片实验样本取自 Pl@ antNet 叶片库^[11],该库涵盖了法国以及周边国家 1 000 种植物物种,包含 28 883 张叶片图像。本文作者自主扩展了该库样本数据集,将中国一些常见植物的叶片图像加入了该库以验证算法的准确性和普适性,最终训练集及测试样本共计 33 293 张叶片图像,共计 231 种。本叶片库叶片采集图像分为单一背景叶片图像(single background leaf picture)和复杂背景叶片图像(complicate background leaf picture)两大类,并且分别对其进行图像预处理,保留有效的叶片信息图像,作为实验和测试的数据集,同时将数据集增加标签信息。

单一背景叶片图像,均为纯色背景上有唯一的叶片图像,将叶片背景颜色统一为白色可以提高分类特征的精确性。将图像灰度化之后应用大津算法^[12]计算自适应阈值,该方法可以快速准确地分割目标区域。计算完成后将灰度图中小于阈值的部分标记为背景部分,并且将对应的原彩色图像所有背景部分设置为白色,以此来提取叶片的无背景轮廓图(见图 1)。

对于复杂背景下的叶片图像,其叶片图像周围充斥着复杂的背景图像以及干扰噪声,难以用上述方法对叶片做出准确的分割处理。部分叶片图像可以用分水岭算法等分割出叶片,但是这样不具有普适性,在去除噪声后,可以将重点关注部分标注出来用于后续处理,而不再考虑将叶片整体分割出来,采用感兴趣区域(ROI)检测方法^[13]分割可以取得不错的效果。首先将图像灰度化,之后应用高斯滤波器采用公式(1)进行图像滤波处理。

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

式中: σ 为方差,本文采用 $\sigma = 3$ 处理图像。

提取包含叶片的图像 ROI 矩阵作为初始 ROI 矩阵。ROI 矩阵将图像从左上角到右下角分割。将图像像素的行数和列数设为定位坐标(r, l),则左上角坐标为 $(\frac{r}{6}, \frac{l}{6})$,右下角坐标为 $(\frac{5r}{6}, \frac{5l}{6})$ 。

应用 ROI 进行分割标记后得到处理结果(图 2)。



图1 单一背景叶片图像(左)和分割后图像(右)

Fig. 1 Single background leaf picture (left) and segmentation image (right)

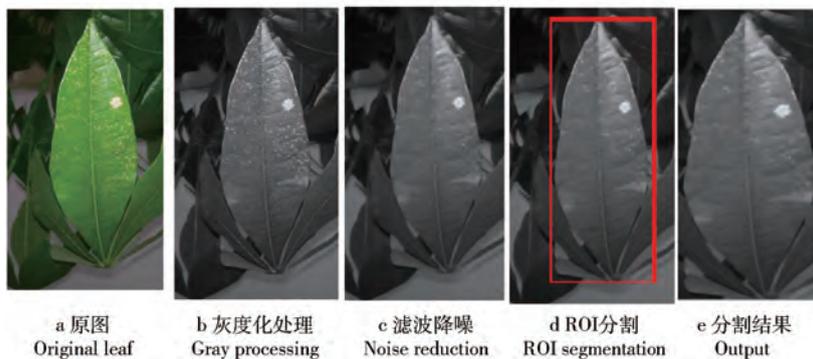


图2 感兴趣区域(ROI)分割处理过程

Fig. 2 Processing of region of interest (ROI) segmentation

2 卷积神经网络

卷积神经网络^[14]是人工神经网络的一种,具有权值共享网络,使其更类似生物神经网络,降低网络模型复杂度,减少了权值数量。利用图像直接作为网络的输入避免了传统识别算法中复杂的特征提取和数据重构过程^[15]。卷积神经网络是多层神经网络,每层由多个二维平面组成,每个平面由多个独立神经元组成。卷积神经网络由输入、卷积层、子采样层、连接层、分类层和输出等模块组成,通常采用卷积层和下采样层交替构成深度学习结构。

卷积网络本质上是一种输入到输出的映射,它能够学习大量的输入与输出之间的映射关系,而不需要任何输入和输出之间精确的数学表达式,只要用已知的模式对卷积网络加以训练,最终得到各层对于图像不同的表达形式,实现图像的深层次理解和表示^[16]。

卷积神经网络的训练过程可以分为两个阶段,第一个是向前传播阶段(feedforward pass),上一层的输出即为当前层的输入,并通过激活函数逐层传递,从样本集中取一个样本 (X, Y_p) ,将样本输入网络,计算得到相应实际输出 O_p 。整个网络的实际计算输出^[16]可表示为:

$$O_p = F_n(\dots(F_2(F_1(XW_1)W_2)\dots)W_n) \quad (2)$$

当前层的输出可表示为:

$$x^l = f(W^l x^{l-1} + b^l) \quad (3)$$

式中: l 代表网络层数, W_1, W_2, \dots, W_n 代表已经训练好的映射权值矩阵, W^l 代表已经训练好的、当前网络层的映射权值矩阵, b^l 为当前网络的加性偏置, $f()$ 是激活函数,需要在完成卷积映射之后通过激活函数对输出进行调整压缩,常见的激活函数包括 sigmoid 函数和双正切函数。

$$\text{sigmoid}(a) = 1/(1 + e^{-a}) \quad (4)$$

$$\text{tanh}(a) = (e^a - e^{-a})/(e^a + e^{-a}) \quad (5)$$

式中: a 为倾斜参数。

第二阶段是反向传播阶段 (backpropagation pass), 与传统的 BP 算法类似, 通过误差函数反向传播, 误差可以看作是神经元的基的灵敏度 (sensitivities) 利用随机梯度下降法对卷积参数和偏置进行优化调整。定义灵敏度^[16]如下:

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u} \frac{\partial u}{\partial b} = \delta \quad (6)$$

式中: b 为神经元的基, E 为误差, u 为一个节点的全部输入; 由于 $\partial u/\partial b = 1$, 所以 $\partial E/\partial b = \partial E/\partial u = \delta$, 即神经元基的灵敏度 $\partial E/\partial b = \delta$ 和误差 E 对一个节点全部输入 u 的导数 $\partial E/\partial u$ 是相等的。可以推出第 l 层的灵敏度公式如下:

$$\delta^l = (\mathbf{W}^{l+1})^T \delta^{l+1} g'(\mathbf{u}^l) \quad (7)$$

式中: \circ 表示每个元素相乘。

输出层的神经元灵敏度公式为:

$$\delta^l = f'(\mathbf{u}^l) \circ (y^n - t^n) \quad (8)$$

式中: t^n 为第 n 个样本的期望输出, y^n 为第 n 个样本的实际输出。

最后, 对每个神经元运用 δ 函数进行权值更新, 即将神经元输出运用 δ 进行缩放, 对于第 l 层, 误差对于该层每一个权值 (组合为矩阵) 的导数是该层的输入 (即上一层的输出) 与该层的灵敏度 (每个神经元的 δ 组合成一个向量的形式) 的叉乘。然后, 得到的偏导数乘以一个负学习率就是该层神经元的权值更新, 如公式 (9) 和公式 (10) 所示。

$$\frac{\partial E}{\partial \mathbf{W}^l} = \mathbf{x}^{(l-1)} (\delta^l)^T \quad (9)$$

$$\Delta \mathbf{W}^l = -\eta \frac{\partial E}{\partial \mathbf{W}^l} \quad (10)$$

实际上, 对于每一个权值 (W) _{ij} 都有一个特定的学习率 η_{ij} 。卷积神经网络与一般 BP 神经网络的最大区别除其自身的深度结构之外, 还采用稀疏连接

(局部感受) 以及权值共享的方法来进行进一步降低训练参数, 减少计算复杂度。稀疏连接, 指的是每种卷积核只与图像中的某个特定区相连接, 即每种卷积核只卷积图像的一部分, 接下来再在其他层中将这局部卷积特征联系在一起, 既符合图像像素在空间上的关联性, 又降低了卷积参数的数量。而权值共享则是使得每种卷积核的权值均相同, 通过增加卷积核的种类来提取图像多方面特征。卷积过程可以表述为用一个可训练的滤波器 f_x 卷积计算一个输入的图像, 然后加一个偏置 b_x , 得到卷积层 C_x 。子采样过程为每邻域 4 个像素求和变为一个像素, 然后通过标量 W_{x+1} 加权, 再增加偏置 b_{x+1} , 然后通过一个激活函数, 产生一个缩小 4 倍的特征映射图 S_{x+1} 。所以, 从一个平面到下一个平面的映射可以看作是卷积运算, S 层可看作是模糊滤波器, 起到二次特征提取的作用。隐层与隐层之间空间分辨率递减, 而每层所含的平面数递增, 这样可用于检测更多的特征信息。图 3 为典型的 LeNet-5 卷积神经网络结构。

3 分层卷积深度学习系统设计

卷积神经网络适用于复杂的图像特征, 并且具有很好的鲁棒性, 对图像平移、比例缩放、倾斜或其他形式的变形具有高度不变性, 基于卷积神经网络 (CNNs) 架构叶片识别系统可以有效规避人工特征选取等问题。目前常用的卷积神经网络类库及框架有 Caffe, Theano 和 Torch, 3 个框架各有所长。本文依据卷积神经网络原理以及 Theano 类库设计实现了分层深度卷积神经网络系统。对于叶片图像, 叶片颜色也是很重要的识别指标, 因此本文采用多通道 (Channel) 方法, 将图像颜色分层处理为 RGB 3 个通道作为输入图像。

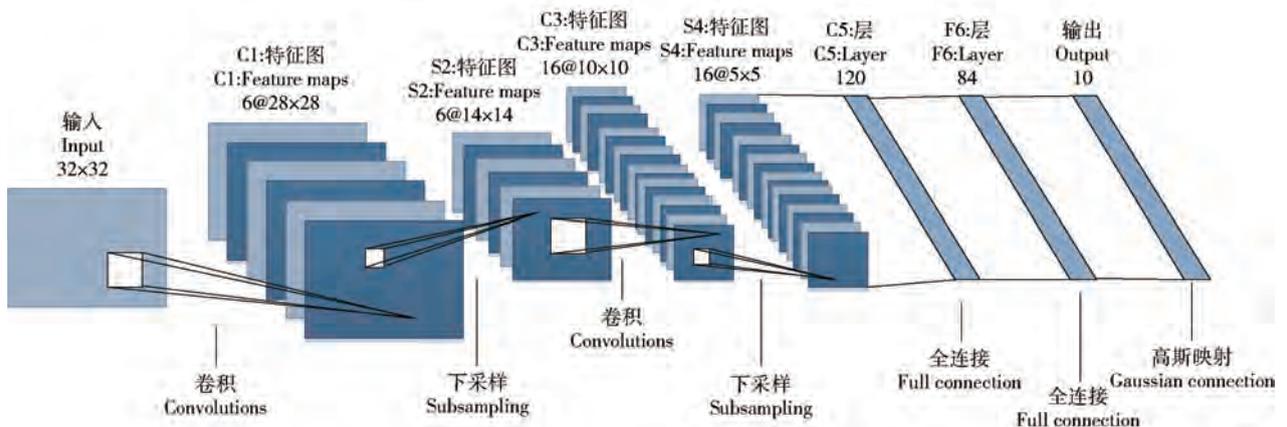


图3 LeNet-5 卷积神经网络结构

Fig. 3 LeNet-5 Convolutions neural network structure

本文设计深度学习叶片识别系统的卷积神经网络包括2个卷积层、2个子采样层、2个局部对比度

归一化层、1个全连接层、1个分类层。本网络的深度学习框架如图4所示。

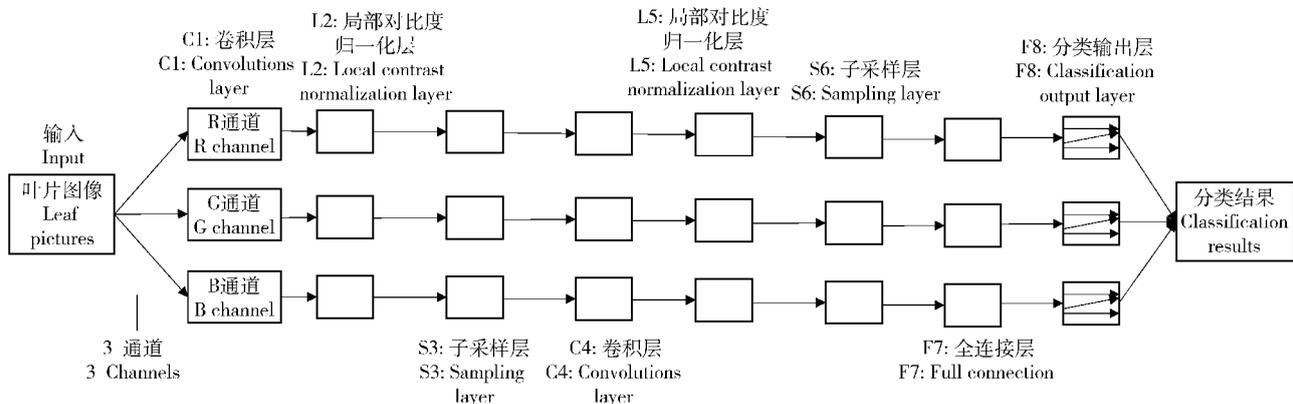


图4 分层卷积深度学习框架

Fig. 4 Layered convolutions deep learning framework

图4框架中的网络层一共有8层(C1、L2、S3、C4、L5、S6、F7、F8)。每个层有多个特征图(feature map),每个特征图通过一种卷积滤波器提取输入的一种特征,然后每个特征图有多个神经元。其中,Cn层是卷积层^[17],完成对输入的特征提取,假设给定分辨率为 $r \times c$ 的大尺寸图像,将其定义为 x_{large} 。首先通过从大尺寸图像中抽取的 $a \times b$ 的小尺寸图像样本 x_{small} 训练稀疏自编码,计算 $f = \sigma(W_1 x_{small} + b_1)$ 得到 K 个特征值,其中 σ 是sigmoid型函数, W_1 和 b_1 是可视单元及隐含单元之间的权重和偏置。对于每一个 $a \times b$ 大小的图像,计算出对应的值,然后对这些值进行卷积,就可以得到 $K \times (r - a + 1) \times (c - b + 1)$ 卷积后的特征矩阵。卷积层中同一个特征图使用相同的卷积核。对于C1层,采用32个卷积核,设定卷积核为 5×5 大小,原图像分辨率全部缩放为 32×32 ,用每个卷积核与原始输入图像进行卷积,这样特征图大小为 $(32 - 5 + 1) \times (32 - 5 + 1) = 28 \times 28$,所需要的参数个数为 $(5 \times 5 + 1) \times 32 = 832$,式中1为偏置参数,连接数为 $832 \times 28 \times 28 = 652\ 288$ 。

由于特征过多容易引起过拟合(over-fitting),所以Sn层为子采样层。本系统采用最大值池化算法(max-pooling)将不同位置特征进行聚合统计,计算图像选定区域上的某个特征的最大值作为特征值。对于S2层,使用缩放因子参数为2,获得100个 30×30 大小的特征图。Ln是局部对比度归一化层(local contrast normalization layer)^[18],该层主要进行的是局部减、除运算归一化,将特征图中相邻特征进行局部竞争,迫使不同特征图的同一空间位置特征竞争。减法归一实质上就是对应位置的值减去邻域各像素的加权后的值,权值是为了区分与该位置距

离不同而影响不同,权值可以由一个高斯加权窗来确定。除法归一化实际上先计算每一个特征图在同一个空间位置的邻域的加权求和的值,计算所有值的均值,然后每个特征图在该位置的值被重新计算为该点的值除以均值。对于不同的卷积层、采样层以及局部对比度归一化层,工作原理一样,只是随着层数加深,提取的特征更加抽象也更具表达力。F7层是全连接层,计算输入向量和权重向量之间的点积,好处是可以增加特征值的准确性,通过多个连接的神经元来进一步特化特征。

F8是分类输出层,传统卷积神经网络在特征提取完成之后多采用全连通的Softmax回归模型^[19]进行分类。Softmax回归模型是logistic回归模型在多元分类问题上的推广。在多元分类问题中,类标签 y 可以取 k 个不同的值,对于训练集 $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$,有 $y_i \in \{1, 2, \dots, k\}$ 。对于给定的测试输入 x ,用假设函数针对每一个类别 j 估算出概率值 $p(y = j | x)$,则对于 k 个不同的样本,假设函数 $h_\theta(x)$ 形式如下:

$$h_\theta(x_i) = \begin{bmatrix} p(y_i = 1 | x_i; \theta) \\ p(y_i = 2 | x_i; \theta) \\ \vdots \\ p(y_i = k | x_i; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x_i}} \begin{bmatrix} e^{\theta_1^T x_i} \\ e^{\theta_2^T x_i} \\ \vdots \\ e^{\theta_k^T x_i} \end{bmatrix} \quad (11)$$

式中 $\theta_1, \theta_2, \dots, \theta_k$ 是模型的参数,为方便起见,使用符号 θ 来表示全部的模型参数,实现Softmax回归时,将 θ 用一个 $k(n+1)$ 的矩阵来表示,该矩阵由 $\theta_1, \theta_2, \dots, \theta_k$ 按行列列起来得到; $e^{\theta_j^T x_i}$ 对概率分布进行归一化,使得所有概率之和为1。因此,Softmax假设函数输出的是一个 k 维列向量,每一个维度的值

代表那个类别出现的概率。

支持向量机(SVM)^[20]是一种依赖核函数的机器学习分类算法,引入惩罚变量 C 解决过拟合问题。通过对训练样本几倍交叉验证来调节 C 及核函数的相关参数优化拟合问题,产生的支持向量有的在超平(曲)面上,有的在两个超平(曲)面之间。为了解决过拟合而摒弃,以此来规避数据偏斜问题。本文选用的径向基(RBF)形式核函数如下:

$$K(x, x_i) = \exp\left\{-\frac{\|x - x_i\|^2}{\sigma^2}\right\} \quad (12)$$

本系统尝试实现了 SVM 及 Softmax 分类器两种分类方法,其中 SVM 选取默认参数 $C = 10$, $\sigma = 0.03846154$ 。最终发现两种分类器针对本文实现的系统以及训练集识别率差距不大,SVM 略有提高。为了提高系统的准确性,本系统在卷积层采用了 Dropout 策略^[21],也就是在模型训练时随机让网络某些隐含层节点的权重不变化。不变化的那些节点可以暂时认为不是网络结构的一部分,但是它的权重值保留,下次样本输入时随机恢复权重变化。

单个卷积神经网络采用分段学习率训练方法^[22],初始学习率参数设置为 0.0001,通过多次迭代训练,观察网络的准确率,当准确率不再显著提高时,继续调整学习率的值,以此类推,待整个卷积神经网络的网络准确率趋于稳定后停止训练。所有层的动量系数均设置为 0.9,权重衰减率为 0.004。迭代训练到整体系统识别率不再显著提高时即认为系统可以进行测试实验。此时开始使用单一背景叶片和复杂背景叶片预处理后的测试数据集批量测试结

果,将命中率记录为最终结果即可。

4 实验结果与讨论

图 5 是本文设计的叶片图像训练和识别系统的框图。实验素材采用了 Pl@ antNet 叶片库,并且通过数码相机等数字图像采集设备增加了国内常见的叶片图像数据,去掉了原叶片库中叶形过于相似的几种植物图像,最终训练集及测试样本共计 33 293 张叶片图像,其中单一背景叶片图像 13 743 张,复杂背景下的叶片图像 19 550 张,针对两类不同的图像进行相应的预处理,最终生成分辨率为 32×32 的带标签图像。选取其中 12 369 张叶片扫描图以及 16 630 张复杂背景叶片图像作为训练集训练数据,用 1 374 张叶片扫描图以及 2 920 张复杂背景叶片图像作为测试数据测试结果。单一背景叶片图像统一叫做 single leaf,复杂背景叶片图像统一叫做 complicate leaf。

为了提高样本训练质量,优化样本集,叶片图像训练之前先进行了数据增强(data augmentation),通过对叶片图像进行平移、旋转和缩放等操作生成更多图像以扩充数据集。同时,为了比对本框架跟其他系统的识别效果,本文将设计实现的 8 层分层深度学习 CNN 卷积神经网络算法与基于 SIFT 算子的叶片识别系统^[7]、基于叶片图像特征的多分类器识别系统^[1](包括 BP 神经网络、KNN 和 SVM 3 种分类器)的叶片图像识别结果进行了对比分析,对于背景单一和复杂的叶片图像都获得了较好的识别效果。

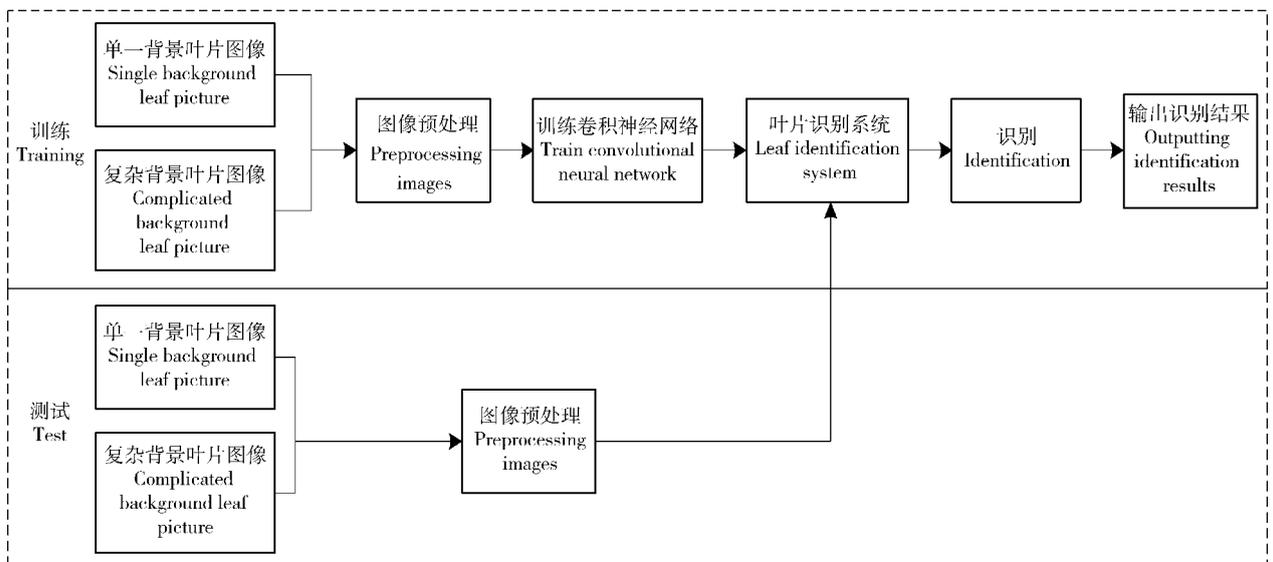


图 5 叶片识别系统结构

Fig. 5 Leaf identification system architecture

首先,利用分层卷积神经网络学习算法分别对叶片图像分割前和分割后的识别效果进行分析。对

于未分割的叶片图像,预处理时直接缩放为 32×32 像素作为输入图像,分割前和分割后的图像识别分

类器均采用 Softmax 回归模型,得到的识别结果见表 1。从表 1 可以看出,图像分割后的叶片识别率明显提高,并且针对复杂背景的叶片图像提升最明显。由于复杂背景中卷积神经网络很容易将噪声等作为特征保留,产生相应误差,所以需要对其进行分割。

表 1 CNN 识别分割与未分割叶片图像识别率比较

Tab. 1 Comparison of recognition rate of segmentation image and unprocessed image in CNN %

分类 Classification	分割 Segmentation image	未分割 Unprocessed image
单一背景叶片图像 Single leaf	90.90	72.13
复杂背景叶片图像 Complicated leaf	34.38	17.88

将经过分割的叶片库数据集应用 8 层 CNN 算法与 SIFT 算子、BP 神经网络、KNN、SVM 等叶片识别系统进行对比分析,得到的识别率见表 2。

表 2 不同叶片识别系统识别率比较

Tab. 2 Recognition rate of different identification systems %

识别系统 Identification system	单一背景叶片图像 Single leaf	复杂背景叶片图像 Complicate leaf
SIFT 特征识别 SIFT features recognition	31.15	8.49
BP 神经网络识别 BP recognition	83.26	12.26
KNN 特征识别 KNN features recognition	75.47	9.90
SVM 特征识别 SVM features recognition	88.64	15.38
卷积神经网络识别 CNN recognition	90.90	34.38

从表 2 可以分析得出:传统的基于叶片图像特征多分类器识别系统在识别单一背景叶片图像时与卷积神经网络算法识别率比较接近;但对于复杂背景叶片图像,分层卷积神经网络深度学习算法识别率明显提高,主要原因是卷积神经网络通过训练能自动生成特征。本文实现的有监督识别过程中,针对复杂背景叶片图像进行的优化处理起到了一定的收敛效果,而传统的识别系统缺乏对复杂背景噪声干扰的分辨能力,容易受到干扰。未来基于卷积神经网络的分层深度学习算法在复杂背景下的图像识别将会更加优化。

本文实现的分层卷积神经网络深度学习识别系

统分别采用了 SVM 和 Softmax 分类器,对分割前后的叶片图像进行了分类实验,得到的结果见表 3。

表 3 SVM 和 Softmax 分类器识别率比较

Tab. 3 Recognition rate of two different classifiers %

分类器 Different classifiers	单一背景叶片图像 Single leaf	复杂背景叶片图像 Complicate leaf
SVM 分类器 CNN + SVM	91.11	31.78
Softmax 分类器 CNN + Softmax	90.90	34.38

卷积神经网络输出层的分类器可以进行相应的优化,在本文实验中,并没有将 SVM 作为主要分类器,因为 SVM 惩罚系数以及核函数参数等设置没有通用的确认方法,需要网格法进行一定范围内的穷举,但是对于不同样本情况,收敛情况亦会发生变化。如表 3 结果所示,针对单一背景叶片图像的识别过程,SVM 分类器分类特征最终结果略优于 Softmax 分类器,但是对于复杂背景叶片图像则不如 Softmax 分类器表现稳定。

5 结论与展望

本文基于卷积神经网络构建了一个 8 层的深度学习叶片识别系统,并且利用 PI@ antNet 叶片库及自主扩展数据集训练叶片样本数据,完成识别率测试。首先,探讨了叶片图像预处理的算法及思路,并且通过实验对比证明了深度学习过程中预处理的必要性;其次,将构建的识别系统与传统的基于叶片图像特征多分类器识别系统进行了对比分析,证明本文实现分层卷积深度学习识别系统在数据量大而无法做出更多优化的情况下,叶片图像的识别率更高,尤其是针对复杂背景下的叶片图像,取得了极佳的识别效果;最后,尝试对深度学习网络的分类器进行了改进,通过实验对比发现,SVM 作为防过拟合的分类器之一,可以在卷积神经网络分类层中有所改进。

本文实现的 8 层深度学习叶片识别系统依然有改进空间,各层参数等均采取的默认值,权值参数调优过程依然值得改进。同时,图像分割处理部分依然可以作为未来研究的重点之一。复杂叶片图像背景下识别率不到 40%,改进空间依然很大,同时对于叶形过于相似的植物分类识别将是植物分类中面临的挑战。

参 考 文 献

- [1] 王丽君,淮永建,彭月橙. 基于叶片图像多特征融合的观叶植物种类识别[J]. 北京林业大学学报,2015,37(1):55-61.

- WANG L J, HUAI Y J, PENG Y C. Method of identification of foliage from plants based on extraction of multiple features of leaf images[J]. *Journal of Beijing Forestry University*, 2015, 37(1): 55-61.
- [2] 杨天天, 潘晓星, 穆立蕾. 基于叶片图像特征数字化信息识别7种柳属植物[J]. *东北林业大学学报*, 2014, 42(12): 75-79.
YANG T T, PAN X X, MU L Q. Identification of seven *Salix* species using digital information analysis of leaf image characteristics[J]. *Journal of Northeast Forestry University*, 2014, 42(12): 75-79.
- [3] 张宁, 刘文萍. 基于克隆选择算法和 K 近邻的植物叶片识别方法[J]. *计算机应用*, 2013, 33(7): 2009-2013.
ZHANG N, LIU W P. Plant leaf recognition method based on clonal selection algorithm and K nearest neighbor[J]. *Journal of Computer Applications*, 2013, 33(7): 2009-2013.
- [4] 邓立苗, 唐俊, 马文杰. 基于图像处理的玉米叶片特征提取与识别系统[J]. *中国农机化学报*, 2014, 35(6): 72-75, 79.
DENG L M, TANG J, MA W J. Feature extraction and recognition system of maize leaf based on image processing[J]. *Journal of Chinese Agricultural Mechanization*, 2014, 35(6): 72-75, 79.
- [5] 翟传敏, 汪青萍, 杜吉祥. 基于叶缘与叶脉分数维特征的植物叶识别方法研究[J]. *计算机科学*, 2014, 41(2): 170-173.
ZHAI C M, WANG Q P, DU J X. Plant leaf recognition method based on fractal dimension feature of outline and venation[J]. *Computer Science*, 2014, 41(2): 170-173.
- [6] 邹秋霞, 郜鲁涛, 盛立冲. 基于 Android 手机和图像特征识别技术的植物叶片分类系统的研究[J]. *安徽农业科学*, 2015, 43(11): 367-369.
ZOU Q X, GAO L T, SHENG L C. Study on plant leaves classification system based on android mobile phone and image feature recognition technology[J]. *Journal of Anhui Agricultural Sciences*, 2015, 43(11): 367-369.
- [7] VILLENA-ROMÁN J, LANA-SERRANO S, CRISTÓBAL J C G. Daedalus at ImageCLEF 2011 plant identification task; using SIFT keypoints for object detection[C]//Conference and Labs of the Evaluation forum 2011. Amsterdam: Clef Labs & Workshop, 2011.
- [8] BACKES A R, CASANOVA D, BRUNO O M. A complex network-based approach for boundary shape analysis[J]. *Pattern Recognition*, 2009, 42(1): 54-67.
- [9] GHAZI M M, YANIKOGLU B, APTOULA E, et al. Sabanci-okan system in LifeCLEF 2015 plant identification competition[C]//Conference and Labs of the Evaluation forum 2015. [2015-10-20]. <http://ceur-ws.org/Vol-1391/43-CR.pdf>.
- [10] YANIKOGLU B, APTOULA E, TIRKAZ C. Sabanci-okan system at ImageClef 2012; combining features and classifiers for plant identification[C]//Conference and Labs of the Evaluation forum 2012. Rome: Clef, 2012.
- [11] LifeCLEF 2015 plant task[EB/OL]. [2015-04-15]. <http://www.imageclef.org/lifeclef/2015>.
- [12] OTSU N. An automatic threshold selection method based on discriminate and least squares criteria[J]. *Denshi Tsushin Gakkai Ronbunshi*, 1979, 63: 349-356.
- [13] BOSCH A, ZISSERMAN A, MUOZ X. Image classification using random forests and Ferns[C]//IEEE International Conference on Computer Vision. Rio de Janeiro: IEEE, 2007: 1-8.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1106-1114.
- [15] LÉCUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [16] BOUVRIE J. Notes on convolutional neural networks[J/OL]. (2006-11-22)[2015-10-11]. <http://cogprints.org/5869>.
- [17] LECUN Y, BENGIO Y, HINTON G. Deep learning. [J]. *Nature*, 2015, 521: 436-444.
- [18] LYU S, SIMONCELLI E P. Nonlinear image representation using divisive normalization[C]//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage: IEEE, 2008: 1-8.
- [19] GOLD S, RANGARAJAN A. Softmax to softassign; neural network algorithms for combinatorial optimization[J]. *Journal of Artificial Neural Networks*, 1995, 2(4): 381-399.
- [20] 张学工. 关于统计学习理论与支持向量机[J]. *自动化学报*, 2000, 26(1): 32-42.
ZHANG X G. Introduction to statistical learning theory and support vector machines[J]. *Acta Automatica Sinica*, 2000, 26(1): 32-42.
- [21] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J/OL]. (2012-07-03)[2015-10-11]. <http://arxiv.org/abs/1207.0580>.
- [22] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.

(责任编辑 冯秀兰
责任编辑 赵秀海)