

DOI: 10.13332/j.1000-1522.20150267

基于多特征融合和深度信念网络的植物叶片识别

刘念 阚江明

(北京林业大学工学院)

摘要:基于叶片数字图像的植物识别是自动植物分类研究的热点。但是随着植物种类的增加,传统的分类方法由于提取的特征比较单一或者分类器结构过于简单,导致叶片识别率较低。为此,本文提出使用纹理特征结合形状特征进行识别,并且使用深度信念网络构架作为分类器。纹理特征通过局部二值模式、Gabor滤波和灰度共生矩阵方法得到。而形状特征向量由Hu氏不变量和傅里叶描述子组成。为了避免过拟合现象,使用“dropout”方法训练深度信念网络。这种基于多特征融合的深度信念网络的植物识别方法,在Flavia数据库中,对32种叶片的识别率为99.37%;在ICL数据库中,对220种叶片的识别率为93.939%。这表明相比一般的叶片识别方法,此方法鲁棒性更强,并且识别率更高。

关键词:叶片识别;傅里叶描述子;局部二值模式;深度信念网络

中图分类号:TP391.4 **文献标志码:**A **文章编号:**1000-1522(2016)03-0110-10

LIU Nian; KAN Jiang-ming. **Plant leaf identification based on the multi-feature fusion and deep belief networks method.** *Journal of Beijing Forestry University* (2016) **38**(3) 110-119 [Ch, 32 ref.] School of Technology, Beijing Forestry University, Beijing, 100083, P. R. China.

Plant identification based on digital leaf image is a hot topic in research of automatic classification of plants. However, with the increase in the number of plant species, the leaf recognition rate is low due to the single trait extraction and simple structure classifier in the traditional classification methods. This study applied the combination of texture features and shape features for identification, using the deep belief networks (DBNs) as the classifier. Texture features are derived from local binary patterns, Gabor filters and gray level co-occurrence matrix while shape feature vector is modeled using Hu Moment invariants and Fourier descriptors. In order to avoid overfitting, we trained the DBNs with “dropout” method. The proposed algorithm was tested on the Flavia dataset, and the recognition rate was 99.37% for 32 species, while on the ICL dataset the recognition rate was 93.939% for 220 kinds of leaves. The experimental results illustrated that the proposed method has stronger robustness and higher recognition rate compared to the traditional methods.

Key words leaf identification; Fourier descriptor; local binary patterns; deep belief networks

传统的植物识别方法有形态学、解剖学、植物化学分类法、细胞分类法等^[1]。这些传统的植物分类方法需要具有专业植物学知识的专家才能运用。现在植物图像越来越多,专家的数量难以满足大量图像处理情况的需求,从而自动识别植物受到越来越多的关注。随着相机和计算机技术的发展,清晰的植物图像很容易被采集到,使用计算机进行植物图像识别是目前植物分类的研究热点^[2]。

植物叶片具有多样性,并且比花和果实有更长的生存期,因此在计算机辅助的植物识别中经常使用叶片图像作为研究对象。2003年祁亨年等^[3]提出了计算机辅助的植物识别技术。2006年Neto等^[4]使用叶片的傅里叶特征和判别分析方法对大豆(*Glycine max*)、向日葵(*Helianthus annuus*)等4种植物进行了分类,总体的植物识别率为88.4%。早期的研究者们多用长宽比、矩形度与圆形度、偏心

收稿日期:2015-07-21 修回日期:2015-09-24

基金项目:国家自然科学基金项目(30901164)。

第一作者:刘念。主要研究方向:图像处理、模式识别。Email:bjfuln@163.com 地址:100083北京市海淀区清华东路35号北京林业大学工学院。

责任作者:阚江明,博士,教授。主要研究方向:机器视觉、智能信息处理。Email:kanjm@bjfu.edu.cn 地址:同上。

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

率、面积凹凸比、周径比等基本的几何特征和其他形态学特征,如中心轮廓距离、傅里叶特征、Hu氏不变量等对叶片进行识别^[5-7]。后来,研究者使用小波滤波器、灰度共生矩阵、分形维度等方法得到叶片纹理特征进行识别^[8]。如Cope等^[9]使用Gabor滤波器组提取纹理特征对32种叶片进行识别,识别率为85.16%。有些研究发现将纹理特征与轮廓特征结合,叶片识别率更高^[10-11]。如Wu等^[12]使用纹理与轮廓特征结合识别Flavia中32种叶片,识别率为93.82%。虽然目前的叶片识别方法很多,但是一般要识别的种类越多,识别率越低。Novotny等^[13]对ICL数据库中的220种叶片进行试验,识别率仅79.68%。

地球上的植物种类繁多,现在的识别准确率不能满足应用。提高识别准确率需要解决以下2个问题:1)特征提取问题。某一种特征对某一叶片数据库有很好的效果,却不一定适用于其他数据库。所以,提取的特征需要能充分地表达叶片的纹理和形状特征。2)分类器设计问题。需要识别的叶片种类越来越多,而目前常用的植物叶片分类器有最近邻分类器、BP神经网络和支持向量机,这些分类器

在处理较高维数的特征向量或者数量较多的训练样本时,收敛的速度较慢,所以难以达到理想效果。提高分类器的结构复杂度才能满足应用的需要。深度信念网络可以用来建立复杂的分类器模型,是目前深度学习领域的研究热点^[14],尚未有研究者使用深度信念网络识别多种类的植物叶片。

本文提出使用局部二值模式(local binary patterns, LBP)、灰度共生矩阵(gray level co-occurrence matrix, GLCM)、Gabor滤波后的纹理特征与Hu氏不变量、傅里叶算子轮廓特征结合进行叶片识别,并且使用深度信念网络(deep belief networks, DBNs)作为分类器构架,最终提高叶片识别的准确率。并将此方法在常用的2个叶片数据库——Flavia数据库^[12]、ICL数据库^[15]中进行验证。

1 叶片特征

1.1 图像预处理

由于图像的纹理特征对图像的大小和方向敏感,所以在提取纹理特征前,需要对图像进行预处理。预处理的步骤如图1所示。

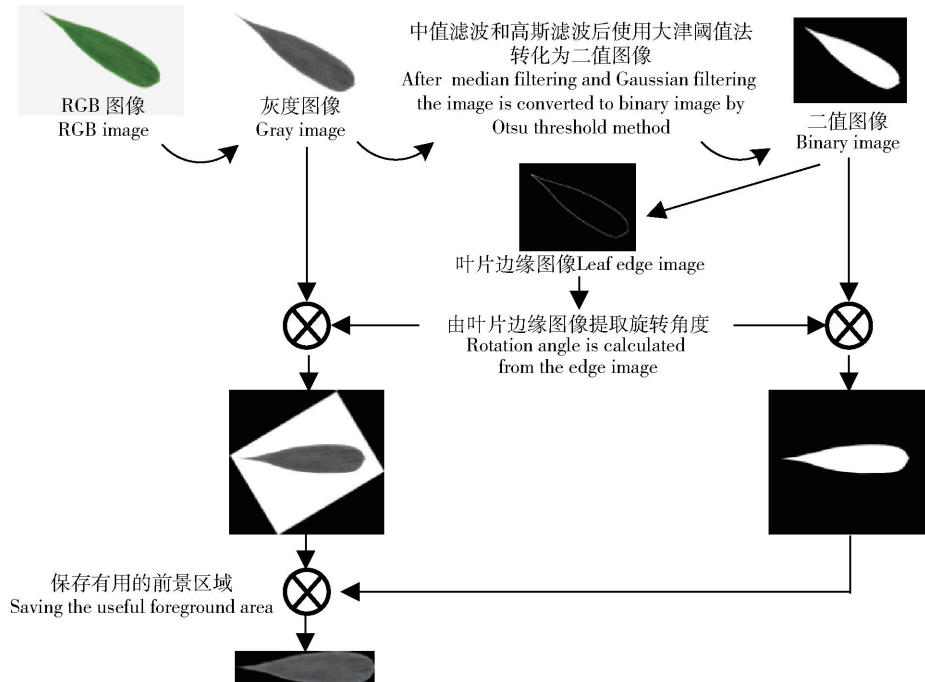


图1 预处理的步骤

Fig. 1 Preprocessing steps

由于本文提取的特征与颜色无关,所以首先将彩色图像(RGB图像)转换为灰度图像。经中值滤波和高斯滤波去掉噪声点后,用大津阈值法将灰度图像转化为二值图像。使用二值图像是为方便提取边缘,并且二值图像的前景部分可作为提取叶片区域的标识。然后使用形态学方法对二值图像进行腐

蚀处理,腐蚀后的图像与原二值图像进行与运算可以得到叶片边缘图像。找到边缘上相距最远的2点。计算最远2点所在的直线与图像横轴的夹角,以这个夹角作为旋转的角度,分别对二值图像和灰度图像进行旋转。二值图像的前景部分为叶片所在的区域,从而在灰度图的旋转图像中只保留这个区

域,得到去掉多余部分的预处理图像结果。

1.2 LBP 特征

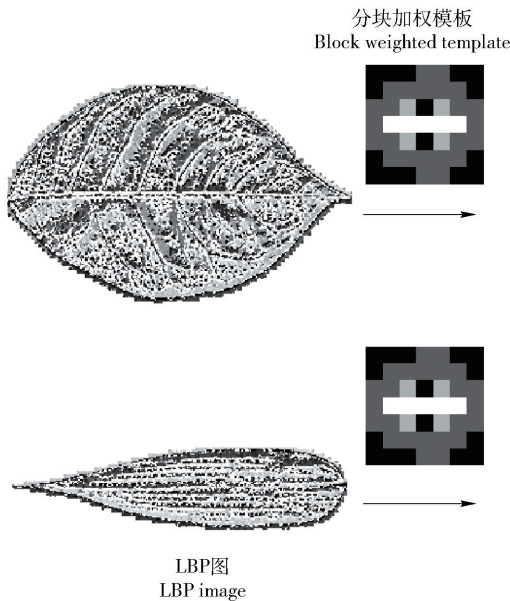
LBP 是一种有效的纹理描述方式,在人脸识别^[16]中经常被使用。基本的 LBP 算子定义在 3×3 的窗口,将相邻的 8 个像素的灰度值与中心点进行比较,若灰度值大于中心点的值,则标记为 1^[17]。用圆形邻域代替正方形邻域,改进后的 LBP 允许在半径为 R 的圆形邻域内有任意多个像素点。我们用的是半径为 2 的圆上的 8 个点。

LBP 相关的基本符号定义如下: g_c 表示局部区域中心点的灰度值, $g_p (p=0,1,\dots,7)$ 对应于中心点周围等距分布的点, (x_c, y_c) 表示中心点的坐标。以 (x_c, y_c) 为中心的 LBP 局部区域纹理计算方法如式

$$(1) \text{ 所示,其中函数 } s(g_p - g_c) = \begin{cases} 1, & g_p \geq g_c \\ 0, & g_p < g_c \end{cases}.$$

$$\text{LBP}(x_c, y_c) = \sum_{p=0}^7 s(g_p - g_c) 2^p \quad (1)$$

对于邻域有 8 个点的 LBP 算子将会产生 2^8 种



LBP 值。邻域内采样点数的增加会使二进制模式的种类数呈指数增加。过多的模式种类对于纹理的表达是不利的。用 Uniform 模式对 LBP 的模式种类进行降维。Uniform 模式定义为局部二进制模式所对应的二进制数 0 与 1 之间跳变小于等于 2 次时,所对应的二进制就称为 1 个 Uniform 模式。除 Uniform 模式外的都归为另一大类。对于 8 个采样点,二进制模式由原始的 256 种减少为 59 种,降低特征维数但保留了大多数信息。

将每张图分成 7×7 个子块,并且赋予不同的权重,分别计算 LBP 直方图,然后依次合成 1 个直方图,所以特征一共是 $49 \times 59 = 2891$ 维(如图 2 所示,分块加权模板的示意图中白色的权值高,黑色的权值低;左侧是叶片图像计算出 LBP 后的图像,右侧是与左侧叶片对应的归一化后的 2891 维 LBP 特征)。特征向量维数过高时,不利于分类器训练。试验中我们使用主成分分析方法将 LBP 特征向量由 2891 维降到 100 维,此时可保留 99% 的信息。

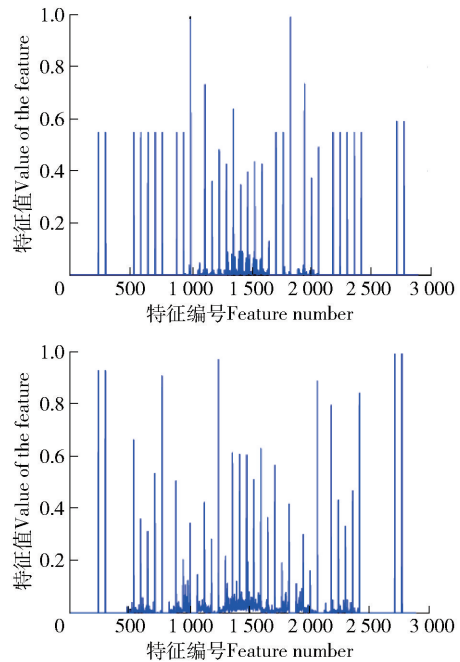


图2 叶片的 LBP 特征

Fig. 2 LBP features of leaves

1.3 Gabor 特征

Gabor 滤波器在图像识别中有广泛的应用,不同参数的 Gabor 滤波器组能够捕捉图像中不同的空间频率、空间位置以及方向等的局部结构信息^[18]。空间域的 Gabor 滤波器可以看作一个被 Gaussian 函数调制的正弦平面波,二维 Gabor 滤波器 $\varphi(x, y)$ 的定义如式(2)所示:

$$\varphi(x, y) = \frac{\alpha\beta}{\pi} \exp(-(\alpha^2 x_r^2 + \beta^2 y_r^2)) \exp(j2\pi f x_r) \quad (2)$$

式中: (x, y) 是图像中点的坐标; $\alpha = f/\gamma, \beta = f/\eta, f$ 是调制正弦平面波的频率, γ 和 η 分别是在 x 轴和 y 轴方向的差额; $x_r = x \cos \theta + y \sin \theta, y_r = -x \sin \theta + y \cos \theta, \theta$ 为滤波器的角度; j 为复数算子。图像 $I(x, y)$ 用 Gabor 滤波器滤波相当于用 Gabor 滤波器做卷积。卷积图像 $G(x, y) = I(x, y) \varphi(x, y), G(x, y)$ 为复数输出。

先将图像变换到统一大小,然后用 3 个方向的 Gabor 滤波器进行滤波,每张滤波后的图像都分成 4×4 的子块,计算每 1 块的均值和方差作为特征。

分别经过 45° 、 90° 、 135° 滤波器滤波后的图像见图 3, 提取的特征维数是 $3 \times 4 \times 4 \times 2 = 96$ 。

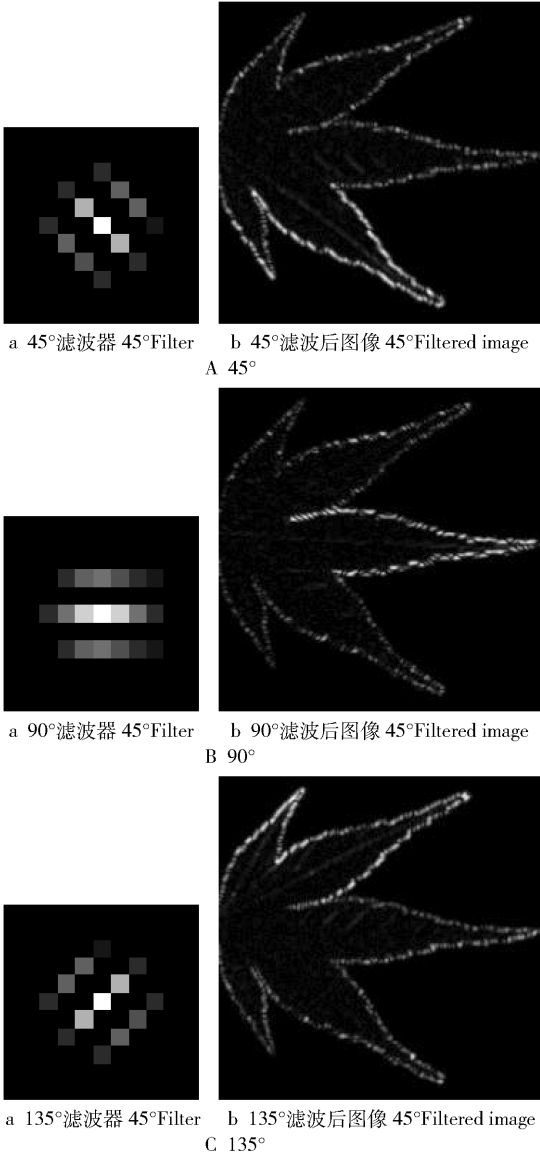


图3 Gabor 滤波图像

Fig. 3 The Gabor filtered images

1.4 Hu 氏不变量

Hu 氏不变量依据统计矩原理计算得到^[5]。对函数 $f(x, y)$ 的二维 Hu 氏 $p + q$ 阶统计量定义如式 (3), 式中 p, q 为非负整数。

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (3)$$

考虑质心 (x_c, y_c) 时, 则式 (3) 变成式 (4):

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - x_c)^p (y - y_c)^q f(x, y) dx dy \quad (4)$$

式中: $x_c = \frac{M_{10}}{M_{00}}, y_c = \frac{M_{01}}{M_{00}}$ 。此时的统计量具有平移不变性, 依照式 (5) 计算后可以保证尺度不变性。

$$\eta_{pq} = \mu_{pq} / \mu_{00}^\gamma \quad (5)$$

式中: 标准化因子 $\gamma = \frac{p+q}{2} + 1$ 。

Hu 等根据标准化的中心统计量推导出 7 个具有旋转、平移、尺度不变性的量, 称作 Hu 氏不变量, 见式 (6) ~ (12)。这些特征量可以用来描述叶片图像特征。

$$\varphi_1 = \eta_{20} + \eta_{02} \quad (6)$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (7)$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \quad (8)$$

$$\varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \quad (9)$$

$$\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{21} + \eta_{30})^2 - (\eta_{21} + \eta_{03})^2] \quad (10)$$

$$\varphi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (11)$$

$$\varphi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{12})^2] + (3\eta_{21} - \eta_{30})(\eta_{03} + \eta_{12}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] \quad (12)$$

在我们的试验中, 计算预处理后剪裁图像的 7 个 Hu 氏不变量, 并取其绝对值的对数值为特征。为了满足深度信念网络的输入, 还需使特征值归一化到 $[0, 1]$ 。

1.5 灰度共生矩阵

灰度共生矩阵计算相邻像素不同灰度值组合出现的次数, 反映了图像相邻像素关于方向和灰度梯度的信息。图像的灰度级数越多, 灰度共生矩阵越大。在灰度共生矩阵的基础上可以提取对比度、逆差矩、能量、熵、相关性^[19]。在计算之前要对灰度共生矩阵进行归一化^[20]:

$$p_{GLCM}(x, y) = \frac{G_{GLCM}(x, y)}{\sum_{x=0}^{L-1} \sum_{y=0}^{L-1} G_{GLCM}(x, y)} \quad (13)$$

式中: $p_{GLCM}(x, y)$ 为归一化后的灰度共生矩阵, $G_{GLCM}(x, y)$ 为原灰度共生矩阵值, L 为灰度共生矩阵级数。

对比度又称惯性矩, 可以理解为图像的清晰度。逆差矩反映了局部纹理特点^[21]。能量也称角二阶矩, 是灰度共生矩阵各个元素的平方和, 它能反应灰度变化和分布的均一性程度。熵能描述图像中纹理的复杂程度。根据相关性可得到图像中的灰度线性依赖关系。这 5 个量的计算方法依次如式 (14) ~ (18) 所示。

$$f_1 = \sum_{x=y=0}^{L-1} (x-y)^2 \left\{ \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p_{GLCM}(x, y) \right\} \quad (14)$$

$$f_2 = \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} \frac{p_{GLCM}(x, y)}{1 + (x-y)^2} \quad (15)$$

$$f_3 = \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p_{\text{GLCM}}(x, y)^2 \quad (16)$$

$$f_4 = - \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p_{\text{GLCM}}(x, y) \log p_{\text{GLCM}}(x, y) \quad (17)$$

$$f_5 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L-1} xy p_{\text{GLCM}}(x, y) - u_1 u_2}{\delta_1^2 \delta_2^2} \quad (18)$$

式中: $u_1 = \sum_{x=0}^{L-1} x \sum_{y=0}^{L-1} p_{\text{GLCM}}(x, y)$, $u_2 = \sum_{y=0}^{L-1} y \sum_{x=0}^{L-1} p_{\text{GLCM}}(x, y)$, $\delta_1^2 = \sum_{x=0}^{L-1} (x - u_1)^2 \cdot \sum_{y=0}^{L-1} p_{\text{GLCM}}(x, y)$, $\delta_2^2 = \sum_{y=0}^{L-1} (y - u_2)^2 \cdot \sum_{x=0}^{L-1} p_{\text{GLCM}}(x, y)$ 。

在试验中我们首先将预处理后的图像转换为16个灰度级,然后计算 0° 、 45° 、 90° 、 135° 4个方向的灰度共生矩阵。计算4个灰度共生矩阵的 $f_1 \sim f_5$ 5个量的均值和方差,得到10个特征值。

1.6 傅里叶特征

使用傅里叶特征表示叶片轮廓是比较常用的方法,傅里叶特征具有旋转不变性^[13]。将预处理后的叶片边缘图像沿顺时针或者逆时针提取叶片边缘点坐标,排序为序列 $(x_0, y_0), (x_1, y_1), \dots, (x_{K-1}, y_{K-1})$ 。其中 (x_0, y_0) 是起始点坐标, K 是边界点数。每个坐标可以表示成1个复数 $s(k) = x(k) + jy(k)$ $(k = 1, 2, \dots, K)$ 。对其进行傅里叶变换得到系数 $a(u)$ $(u = 0, 1, 2, \dots, K-1)$:

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K} \quad (19)$$

使用 $a(u)$ 进行傅里叶反变换可以得到还原边缘。如果系数只取一部分,则得到边缘估计,系数越多得到的轮廓估计越精确^[2]。 $a(u)$ 是复数,我们取复数的模为特征值。为了消除不同尺度图像对结果的影响,令 $a(u) = a(u)/n^2$ $(n = K/2)$ 。由于系数会随着 K 升高下降得很快,所以令 $a(u) = 10(|w| + 1)a(u)/n^2$ $(w = -n, -n+1, \dots, n-1)$ 。这样处理后的傅里叶特征值具有尺度和旋转不变性。为了得到平移不变性,令 $s(k) = (x(k) - x_c) + j(y(k) - y_c)$,其中 (x_c, y_c) 为前景图形的中心坐标。为了适应DBNs网络的输入值在0到1之间要将傅里叶系数归一化,得到的傅里叶特征见图4。

2 深度信念网络

深度学习与机器学习相关,使用多层的非线性信号处理技术来实现特征提取和模式识别等功能^[22]。深层网络可以构建复杂的分类器结构^[23],但是由于层数多导致训练困难。2006年Hinton等^[24]提出了DBNs,并且使用逐层贪婪学习算法解

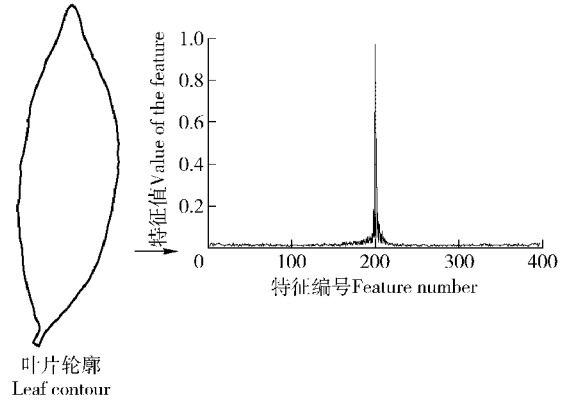


图4 叶片的傅里叶特征

Fig.4 Fourier features

决了深层网络训练困难的问题,从此深度学习成为人工智能领域的热点^[14]。目前深度学习已被应用在人脸识别、语音识别、图像识别等领域^[25]。深度学习构架有DBNs、卷积神经网络等,本文使用DBNs解决叶片识别问题。

2.1 DBNs训练方法

DBNs是由多层的限制玻尔兹曼机(Restricted Boltzmann Machine, RBM)组成。训练DBNs网络首先要逐层训练RBM,称为DBNs的预训练。第1层RBM的可见层输入是样本的特征向量,第2层RBM的可见层输入是第1层隐含层的输出,依次往下。根据限制玻尔兹曼机的能量函数模型和贝叶斯公式,可以得到隐含层的输出概率模型^[26-27],见式(20)、(21)。

$$P(\mathbf{h}_i = 0 | \mathbf{x}) = \frac{1}{1 + \exp(\mathbf{c}'_i + \mathbf{W}_i \mathbf{x})} \quad (20)$$

$$P(\mathbf{h}_i = 1 | \mathbf{x}) = \frac{\exp(\mathbf{c}'_i + \mathbf{W}_i \mathbf{x})}{1 + \exp(\mathbf{c}'_i + \mathbf{W}_i \mathbf{x})} = \text{sigm}(\mathbf{c}'_i + \mathbf{W}_i \mathbf{x}) \quad (21)$$

式中: \mathbf{h} 为隐含层单元矩阵, \mathbf{x} 为可见层单元矩阵, \mathbf{c} 是隐含层偏置矩阵, \mathbf{W} 是隐含层与可见层的连接权重。sigm表示sigmoid函数,上角标的撇号表示矩阵的转置, i 表示第 i 个单元。一般隐含层有 $\mathbf{h}_i \in \{0, 1\}$ 。同理可得,可见层的概率模型:

$$P(\mathbf{x} | \mathbf{h}) = \prod_i P(x_i | \mathbf{h}) \quad (22)$$

当可见层是二值单元时,输出模型:

$$P(x_j = 1 | \mathbf{h}) = \text{sigm}(\mathbf{b}'_j + \mathbf{W}_j \mathbf{x}) \quad (23)$$

式中: \mathbf{W}_j 是 \mathbf{W} 的第 j 列, \mathbf{b} 是可见层偏置。

训练RBM时,使用最大似然估计求解参数不易收敛,并且计算困难,所以使用对比散度(Contrastive Divergence, CD)方法来调整参数:

$$\begin{cases} \mathbf{W} \leftarrow \mathbf{W} + \varepsilon(\mathbf{h}_1 \mathbf{x}'_1 - Q(\mathbf{h}_2 = 1 | \mathbf{x}_2) \mathbf{x}'_2) \\ \mathbf{b} \leftarrow \mathbf{b} + \varepsilon(\mathbf{x}_1 - \mathbf{x}_2) \\ \mathbf{c} \leftarrow \mathbf{c} + \varepsilon(\mathbf{h}_1 - Q(\mathbf{h}_2 = 1 | \mathbf{x}_2)) \end{cases} \quad (24)$$

式中: ε 为训练RBM时的学习率, $Q(\mathbf{h}|\mathbf{x})$ 为近似的后验模型用来推导和训练。

对比散度主要进行交替的Gibbs采样,如图5所示, $P(\mathbf{h}^{k-1}|\mathbf{h}^k)$ 表示生成模型中的概率分布。通常经过1次对比散度调整即可^[28]。经过预训练后,调整好初始DBNs中各层RBM的参数。然后使用反向传播算法对整体参数进行微调即可。

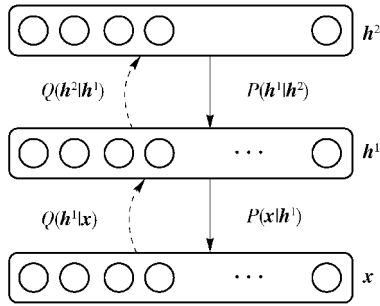


图5 DBNs的预训练过程

Fig.5 Pre-training process of the deep belief networks

2.2 使用“dropout”方法的DBNs

训练深度信念网络时,如果测试样本较少,容易产生过拟合现象。过拟合易造成训练的网络对测试样本产生错误的分类结果。“dropout”方法^[29]可以有效减少过拟合的产生。直观来说,“dropout”方法就是在前向传播时,隐含层单元以一定的概率被清零,相应的,该隐含层单元在反向传播时其误差项也是0。这样就避免了隐含层之间有依赖性,可以一定程度上防止某些特征要在另一些特征成立的情况下才有效。“dropout”方法可以看作是一种特殊的模型平均,类似Bagging和Boosting。但是其他方法需要计算多个分类器模型,在训练和测试时计算复杂,而“dropout”方法最后得到的是一个分类器模型。在试验中我们首先逐层训练DBNs得到网络的初始权重,然后使用“dropout”方法对网络进行循环微调。

3 试验

为了验证方法的有效性,我们在典型的叶片图像数据库Flavia数据库^[12]和ICL数据库^[13]中进行试验。试验使用windows 7系统,Matlab(2012a版本)环境,处理器为intel® core™ i5-4200 CPU@ 1.6 Ghz 2.3 GHz。

3.1 Flavia数据库

Flavia数据库现有32个物种的1907个样本。图像大小均为1600×1200,彩色图像,分辨率为300-DPI。我们使用使用数据库中每种叶片的前50张图像,每种叶片随机选择45张作为训练样本,5张作为测试样本。

预处理后图像的大小不一,在提取LBP特征前要将图像归一化为300×300。方法是使原图中较长边变为300进行缩放,缩放后的空余部分背景值为0。同理提取灰度共生矩阵前将图像归一化。在提取Gabor特征时,因为Gabor滤波器模板设置为14×14,滤波后图像边会减少14像素,并且要将滤波后的图像分成4×4区域进行处理(即需要图像边长被4整除),所以Gabor滤波前将预处理图像归一化到294×294。Hu氏不变量、傅里叶特征不受尺度影响,所以计算这2个特征时,直接对预处理后的图像提取特征。

400维傅里叶特征,100维LBP特征,7个Hu氏不变量,96个Gabor特征值,10个灰度共生矩阵特征量,总特征维数为613维。试验中,训练RBM时学习率为1,分组训练大小设置为32,一般对比散度循环1次即可^[28],反向传播微调时学习率为1,动量为0.5。经过多次试验得出识别率较好的DBN网络隐含层单元数为700,微调循环次数为50,dropout概率设置为0.2。试验结果见表1。其中概率神经网络(probabilistic neural network, PNN)出自文献^[12],径向基概率网络(radial basis probabilistic neural network, RBPNN)出自文献^[31],支持向量机(support vector machine, SVM)分类器出自文献^[30],最近邻分类器(nearest neighbor classifier, 1-NN)是Novotny等^[13]的实验。通过对比可以看出本文的方法识别率最高,可以达到99.37%。

从图6可以看出:使用不同的DBNs结构进行识别,隐含层单元为700时,错误率最低;使用dropout训练的网络比没有使用该方法的网络识别率要高。

3.2 ICL数据库

ICL(Intelligent Computing Laboratory)数据库是中国科学院智能计算实验室的植物叶片图像数据库^[14]。目前包括220个物种的16851个样本。数据库中每种叶片的数目不一,最少的是五叶地锦(*Parthenocissus quinquefolia*),只有26个样本;最多的是紫穗槐(*Amorpha fruticosa*)有1078个样本。为使样本数量差别不过大,我们使用的每种叶片样本不超过100个,每种中随机选取3个作为测试样本,其他的作为训练样本。

数据库中叶片图像均为彩色图像,图像大小不一,在计算LBP和Gabor特征前对预处理之后的图像归一化到400×400,方法同Flavia试验。Hu氏不变量、傅里叶特征不受尺度影响,所以直接对预处理后的图像进行特征提取。ICL数据库试验中DBNs的学习率同Flavia试验,对比散度训练1次。

表1 Flavia 数据库试验结果

Tab.1 Flavia dataset results

使用的特征 Features	类别数 Number of classes	分类器 Classifier	识别率 Recognition rate
几何特征 + 形态学特征 Geometric features + Morphological features	32	概率神经网络 PNN	90.30%
形状特征 + 纹理特征 + 颜色特征 + Zernike 统计矩 Shape features + Vein features + Color features + Zernike moments	32	径向基神经网络 RBPNN	93.82%
形状特征 + Fourier 特征 + 多尺度距离矩阵 Shape features + Fourier descriptors + Multiscale distance matrix	32	支持向量机 SVM	96.00%
Fourier 特征 + Hu 不变量 Fourier + Hu	32	最近邻分类器 1-NN	93.66%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	32	深度信念网络 DBNs	98.75%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	32	深度信念网络(使用“dropout”方法) DBNs(with“dropout”)	99.37%

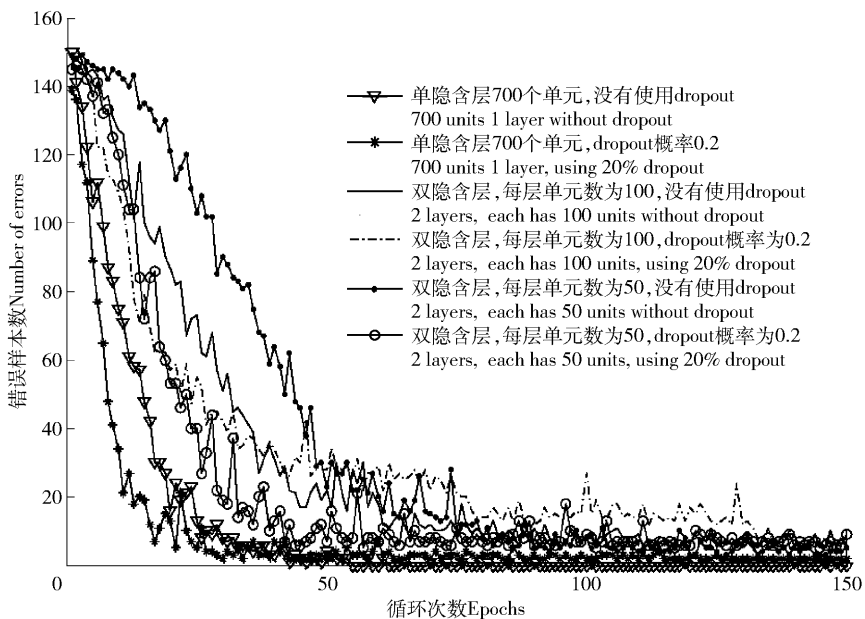


图6 DBNs 的测试结果

Fig.6 Test errors of different DBNs

类别数为 16、50、220 种时,分组训练大小依据选择的样本数量不同分别设置。16 种和 50 种的叶片种类是随机选择的,识别率结果为 3 次试验的平均值,此时的 DBNs 网络并未使用 dropout,识别率已经很高。类别数为 220 时,隐含层单元个数设置为 260,循环微调次数设置为 600 可以达到收敛,识别率为 87.424%。识别 220 种叶片时,使用“dropout”方法,隐含层权重更新的概率设置为 0.2,双隐含层,每层的单元数为 100,微调循环 1 300 次基本达到收敛,识别率为 93.939%。与其他使用 ICL 数据库的研究进行对比,结果如表 2 所示。其中最近邻方法

识别 16 种叶片出自文献[32],匹配方法出自文献[15],使用最近邻分类器识别 220 种叶片出自文献[13]。

我们提出的基于多特征融合和深度信念网络的植物叶片识别方法由于充分提取了叶片图像的特征,所以能够尽量多地利用叶片信息;并且我们使用的深度信念网络相比传统分类器能够建立更加复杂的分类模型。经过适当的调整参数后,在所需识别种类数较多时分类效果优于其他方法。增加了“dropout”方法后,对 220 叶片的识别率明显提高。

表2 ICL 数据库试验结果

Tab.2 ICL dataset results

使用的特征 Features	类别数 Number of classes	分类器 Classifier	识别率 Recognition rate
正交局部判别样条曲线嵌入技术 Orthogonal locally discriminant spline embedding techniques	16	最近邻分类器 1NN	94.29%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	16	深度信念网络 DBNs	97.94%
多尺度距离矩阵 Multiscale distance matrix	50	匹配方法 Matching	80.88%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	50	深度信念网络 DBNs	96.00%
Fourier 特征 + Hu 不变量 Fourier + Hu	220	最近邻分类器 1NN	79.68%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	220	深度信念网络 DBNs	87.42%
Fourier 特征 + LBP 特征 + Hu 不变量 + Gabor 特征 + 灰度共生矩阵特征 Fourier + LBP + Hu + Gabor + GLCM	220	深度信念网络(使用“dropout”方法) DBNs (with “dropout”)	93.94%

3.3 对比试验

为比较不同方法在不同数据库中的时间成本,我们进行了对比试验。分别使用最近邻方法、支持向量机、深度信念网络和有“dropout”的深度信念网络对 Flavia 中 32 种叶片和 ICL 数据库中 220 种叶片进行识别。结果见表 3,其中识别时间是识别 1

个样本所需的平均时间。对 Flavia 数据库, DBN 的结构为单隐含层,隐含层单元数为 700,循环 50 次, dropout 概率为 0.1。对 ICL 数据库, DBN 的结构为双隐含层,隐含层单元数为 100,循环 1 000 次, dropout 概率为 0.2。

表3 时间成本对比表

Tab.3 Comparison table of the time cost

时间 Time	Flavia				ICL			
	最近邻分 类器	支持向 量机	深度信念 网络	深度信念网络 (使用“dropout”方法)	最近邻分 类器	支持向 量机	深度信念 网络	深度信念网络 (使用“dropout”方法)
	1NN	SVM	DBNs	DBNs with “dropout”	1NN	SVM	DBNs	DBNs with “dropout”
训练时间 Training time	0.268 8	0.480 8	98.7	94.6	27.331 1	40.794 3	5 045	5 012
识别时间 Recognition time	1.87×10^{-6}	3.37×10^{-4}	8.87×10^{-5}	6.18×10^{-5}	1.06×10^{-4}	1.38×10^{-2}	8.21×10^{-5}	7.75×10^{-5}

训练 Flavia 数据库使用的样本数为 1 440,而训练 ICL 数据库使用的样本数为 12 061,所以训练 ICL 数据库所用的时间比 Flavia 要长。尽管相对最近邻分类器和支持向量机,使用深度信念网络需要的训练时间比较长,但是使用 DBNs 所需要的识别时间是很短的。对图像识别而言,训练时间影响不大,实际应用中主要考虑识别时间。深度信念网络的识别时间在亚微秒级,能够满足一般应用的需要。此外,可以使用并行处理或者 GPU 加速的方法减少深度信念网络的训练时间。

4 结论与讨论

为了提高叶片识别的识别率,本文提出将 LBP

特征、Gabor 特征、灰度共生矩阵得到的纹理特征与 Hu 氏不变量、傅里叶轮廓特征结合作为特征向量,并且使用深度信念网络作为分类器进行叶片识别。在典型的叶片图像数据库——Flavia 和 ICL 数据库中试验,结果表明本文的方法相比传统方法识别准确率更高。随着叶片种类的增多,使用“dropout”方法的深度信念网络相比传统的模型平均方法计算量更少,并且也有效防止了深度信念网络过拟合现象的产生,从而提高了识别率。深度信念网络隐含层数、隐含层单元数等的设置一般依据经验,几乎没有理论证明哪种参数得到的结果是最好的,随着对深度信念网络研究的深入,将来的叶片识别准确率会进一步提高。

参 考 文 献

- [1] 肖雪洋. 植物叶片图像识别特征的研究和在线识别系统实现 [D]. 合肥:中国科学技术大学, 2011.
XIAO X Y. Study on features of plant leaf image recognition and realization of online recognition system [D]. Hefei: University of Science and Technology of China, 2011.
- [2] COPE J S, CORNEY D, CLARK J Y, et al. Plant species identification using digital morphometrics: a review [J]. *Expert Systems with Applications*, 2012, 39(8): 7562-7573.
- [3] 祁亨年, 寿韬, 金水虎. 基于叶片特征的计算机辅助植物识别模型 [J]. *浙江林学院学报*, 2003, 20(3): 281-284.
QI H N, SHOU T, JIN S H. Leaf characteristics-based computer-aided plant identification model [J]. *Journal of Zhejiang Forestry College*, 2003, 20(3): 281-284.
- [4] NETO J C, MEYER G E, JONES D D, et al. Plant species identification using elliptic fourier leaf shape analysis [J]. *Computers and Electronics in Agriculture*, 2006, 50(2): 121-134.
- [5] WANG X F, HUANG D S, DU J X, et al. Classification of plant leaf images with complicated background [J]. *Applied Mathematics and Computation*, 2008, 205(2): 916-926.
- [6] 侯铜, 姚立红, 阚江明. 基于叶片外形特征的植物识别研究 [J]. *湖南农业科学*, 2009(4): 123-125, 129.
HOU T, YAO L H, KAN J M. Plant recognition research based on shape features of leaf [J]. *Hunan Agricultural Sciences*, 2009(4): 123-125, 129.
- [7] ANDRADE I M, MAYO S J, KIRKUP D, et al. Comparative morphology of populations of *Monstera adans* (Araceae) from natural forest fragments in Northeast Brazil using elliptic fourier analysis of leaf outlines [J]. *Kew Bulletin*, 2008, 63(2): 193-211.
- [8] DU J X, ZHAI C M, WANG Q P. Recognition of plant leaf image based on fractal dimension features [J]. *Neurocomputing*, 2013, 116: 150-156.
- [9] COPE J S, REMAGNINO P, BARMAN S, et al. Plant texture classification using gabor co-occurrences [M]. Berlin Heidelberg: Springer, 2010: 669-677.
- [10] CHAKI J, PAREKH R, BHATTACHARYA S. Plant leaf recognition using texture and shape features with neural classifiers [J]. *Pattern Recognition Letters*, 2015, 58: 61-68.
- [11] 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): 96-104.
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): 96-104.
- [12] WU S G, BAO F S, XU E Y, et al. A leaf recognition algorithm for plant classification using probabilistic neural network [C] // 2007 IEEE international symposium on signal processing and information technology. Cairo: IEEE, 2007: 11-16.
- [13] NOVOTNY P, SUK T. Leaf recognition of woody species in central Europe [J]. *Biosystems Engineering*, 2013, 115(4): 444-452.
- [14] 余凯, 贾磊, 陈雨强, 等. 深度学习的昨天, 今天和明天 [J]. *计算机研究与发展*, 2013, 50(9): 1799-1804.
YU K, JIA L, CHEN Y Q, et al. Deep learning: yesterday, today, and tomorrow [J]. *Journal of Computer Research and Development*, 2013, 50(9): 1799-1804.
- [15] HU R, JIA W, LING H, et al. Multiscale distance matrix for fast plant leaf recognition [J]. *IEEE Transactions on Image Processing*, 2012, 21(11): 4667-4672.
- [16] 袁宝华, 王欢, 任明武. 基于完整 LBP 特征的人脸识别 [J]. *计算机应用研究*, 2012, 29(4): 1557-1559.
YUAN B H, WANG H, REN M W. Face recognition based on completed local binary pattern [J]. *Application Research of Computer*, 2012, 29(4): 1557-1559.
- [17] 黄非非, 李见为, 王玮, 等. 结合小波分析和 LBP 算子的人脸描述与识别 [J]. *重庆工学院学报 (自然科学版)*, 2009, 23(1): 102-108.
HUANG F F, LI J W, WANG W, et al. Human face description and recognition combining wavelet analysis and LBP operator [J]. *Journal of Chongqing Institute of Technology (Natural Science)*, 2009, 23(1): 102-108.
- [18] 沈琳琳, 纪震. 采用精选 Gabor 小波和 SVM 分类的物体识别 [J]. *自动化学报*, 2009, 35(4): 350-355.
SHEN L L, JI Z. Gabor wavelet selection and SVM classification for object recognition [J]. *ACTC Automatic Sinici*, 2009, 35(4): 350-355.
- [19] 刘丽, 匡纲要. 图像纹理特征提取方法综述 [J]. *中国图象图形学报*, 2009, 14(4): 622-635.
LIU L, KUANG G Y. Overview of image textural feature extraction methods [J]. *Journal of Image and Graphics*, 2009, 14(4): 622-635.
- [20] 阚江明, 王怡萱, 杨晓微, 等. 基于叶片图像的植物识别方法 [J]. *科技导报*, 2010, 28(23): 81-85.
KAN J M, WANG Y X, YANG X W, et al. Plant recognition method based on leaf images [J]. *Science and Technology*, 2010, 28(23): 81-85.
- [21] 任国贞, 江涛. 基于灰度共生矩阵的纹理提取方法研究 [J]. *计算机应用与软件*, 2014, 31(11): 190-192, 325.
REN G Z, JIANG T. Study on GLCM-based texture extraction methods [J]. *Computer Applications and Software*, 2014, 31(11): 190-192, 325.
- [22] 戴礼荣, 张仕良. 深度语音信号与信息处理: 研究进展与展望 [J]. *数据采集与处理*, 2014, 29(2): 171-179.
DAI L R, ZHANG S L. Deep speech signal and information processing: research process and prospect [J]. *Journal of Data Acquisition and Processing*, 2014, 29(2): 171-179.
- [23] 柴瑞敏, 曹振基. 基于 Gabor 小波与深度信念网络的人脸识别方法 [J]. *计算机应用*, 2014, 34(9): 2590-2594.
CHAI R M, CAO Z J. Face recognition algorithm based on Gabor wavelet and deep belief networks [J]. *Journal of Computer Applications*, 2014, 34(9): 2590-2594.
- [24] HINTON G, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [25] 郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展 [J]. *中国图象图形学报*, 2014, 19(2): 175-184.

- ZHENG Y, CHENG Q Q, ZHANG Y J. Deep learning and its new progress in object and behavior recognition [J]. *Journal of Image and Graphics*, 2014, 19(2): 175-184.
- [26] PALM R B. Prediction as a candidate for learning deep hierarchical models of data [D]. Kongens Lyngby: Technical University of Denmark, 2012.
- [27] BENGIO Y. Learning deep architectures for AI[J]. *Foundations and Trends in Machine Learning*, 2009, 2(1): 1-127.
- [28] HINTON G. A practical guide to training restricted Boltzmann machines[J]. *Momentum*, 2010, 9(1): 926.
- [29] SRIVASTAVA N, HINTON G E, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. *The Journal of Machine Learning Research*, 2014, 15(1): 1929-1958.
- [30] SARI C, AKGUL C B, SANKUR B. Combination of gross shape features, fourier descriptors and multiscale distance matrix for leaf recognition [C] // 55th International symposium on ELMAR. Zadar: IEEE, 2013: 22-26.
- [31] KULKARNI A H, RAI H M, JAHAGIRDAR K A, et al. A leaf recognition technique for plant classification using RBPNN and Zernike moments[J]. *International Journal of Advanced Research in Computer and Communication Engineering*, 2013, 2(1): 984-988.
- [32] LEI Y K, ZOU J W, DONG T, et al. Orthogonal locally discriminant spline embedding for plant leaf recognition [J]. *Computer Vision and Image Understanding*, 2014, 119: 116-126.

(责任编辑 李文军
责任编辑 李文彬)