

文章编号: 1003-8701(2015)01-0051-03

基于SVM和形态学的作物病斑图像分割方法

王献锋, 王震, 王旭启, 张善文

(西京学院应用统计与理学系, 西安 710123)

摘要:作物叶片病斑图像分割是作物病害自动识别的一个重要步骤, 为了提高传统的基于阈值或聚类的叶片病斑分割方法的分割效果, 提出了一种基于支持向量机(SVM)和形态学的病斑分割方法。首先利用SVM进行病斑图像分割, 再利用开运算和闭运算来消除病斑图像中边缘的不连续性、病斑内部的小噪声和小洞。最后, 通过对黄瓜细菌性角斑病图像进行试验, 结果表明, 所提出分割方法具有较好的分割效果。

关键词:病斑图像分割; 病害叶片图像; 支持向量机(SVM); 开运算和闭运算

中图分类号: TP391.41

文献标志码: A

DOI:10.16423/j.cnki.1003-8701.2015.01.012

Crop Spot Image Segmentation based on SVM and Morphology

WANG Xian-feng, WANG Zhen, WANG Xu-qi, ZHANG Shan-wen

(Department of Applied Science, Xijing University, Xi'an 710123, China)

Abstract: Crop leaf spot image segmentation is the important steps in crop disease automatic recognition. To improve the leaf spot segmentation performance of the traditional threshold or clustering methods, a spot segmentation based on SVM and morphology was proposed in the paper. The spot image segmentation was formulated by SVM. The discontinuity edge, small noise, small hole and the small hole inside the lesion image were eliminated by the opening and closing algorithms of morphology. The experimental results showed that this approach outperformed other methods and was effective for cucumber leaf disease segmentation.

Keywords: Spot image segmentation; Disease leaf image; Support vector machine; Opening and closing algorithms.

对于基于病害叶片的作物病害检查研究而言, 病害叶片中病斑分割是病害检查的一个重要步骤, 分割效果的好坏将会直接影响到病害叶片信息的提取和最终的检查结果。只有对病害叶片图像中的病斑进行有效的分割, 病害识别特征提取和参数测量的精确性才能得以保证, 从而使得更高层次的图像分析和病害识别成为可能^[1-2]。赖军臣等^[3]引入了高斯模型的Markov随机场分割模型, 对玉米叶部病斑图像进行分割试验。该算法对分辨率高、像素太多的图像的处理效果较差。邵乔林等^[4]对田间不同光照下的叶片图像进行了研究, 得到适应不同光照条件下的分类决策

函数, 该函数可以减少阴影和植物残渣对图像分割带来的影响。陈占良等^[5]针对目前植物病害染病程度判别中以目测为主的主观随意的缺陷, 提出了一种基于计算机图像处理技术的病害分级方法。李宗儒等^[6]以手机拍摄图像, 进行苹果病虫害识别, 提取叶片病斑的H方差, 以H-S直方图特征作为病斑的颜色特征, 用计盒维数法提取病斑纹理特征, 进行病虫害识别, 但该方法比较复杂。支持向量机(Support Vector Machine, SVM)是20世纪90年代Vapnik等人在统计学理论的基础上开发出来的一种有效的机器学习方法^[7]。目前SVM已经在分类、时间预测、函数估计等领域得到了广泛应用^[8]。本文将SVM应用于作物病害叶片的病斑分割中, 取得了较好的分割效果。

1 支持向量机(SVM)

SVM最初是为寻找线性可分情况下两类分类问题的最优超平面而提出的。其基本思想是, 在

收稿日期: 2014-08-20

基金项目: 国家自然科学基金(61473237); 陕西省教育厅科研计划项目(2013JK1145、12JK1077); 西京学院科研基金项目(XJ130244、XJ130245)

作者简介: 王献锋(1965-), 男, 硕士, 副教授, 主要研究方向为模式识别与智能算法。

正确划分训练集的超平面中,选出分类间隔最大的一个平面作为最优决策超平面^[9]。在d维欧氏空间中,分类超平面方程表示为:

$$w \cdot x - b = 0, w \in R^d, b \in R \dots\dots\dots (1)$$

式中, $w \cdot x - b$ 为待求的平面方程。

归一化式(1),使得线性可分的样本集 (x_i, y_i) , $i=1, \dots, n$, $x \in R^d$, $y \in \{+1, -1\}$ 满足 $y_i[(x_i, w) - b] - 1 > 0$, $i=1, \dots, n$, 此时分类间隔为 $2/\|w\|$ 。则,线性SVM模型描述为:

$$\begin{aligned} \min_{w,b} & \frac{1}{2} \|w\|^2 \\ \text{s.t.} & y_i[(x_i, w) - b] - 1 > 0, \dots\dots\dots (2) \\ & i=1, \dots, n \end{aligned}$$

当分类样本不是线性可分情况时,SVM采用了一种核函数来完成分类。根据Lagrange乘子法和Wolfe对偶原理,得到式(2)在非线性情况下的优化问题的对偶问题如下:

$$\begin{aligned} \max_a & \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j K(x_i, x_j) \\ \text{s.t.} & \sum_{i=1}^n a_i y_i = 0 \dots\dots\dots (3) \end{aligned}$$

$$0 \leq a_i \leq C, i=1, 2, \dots, n$$

式中, a_i 为Lagrange乘子, $K(x_i, x_j)$ 核函数。

由 a_i 可得原始问题的解:

$$w^* = \sum_{i=1}^n a_i^* y_i x_i \dots\dots\dots (4)$$

$$b^* = y_j - \sum_{i=1}^n a_i^* y_i (x_i \cdot x_j) \dots\dots\dots (5)$$

对应于 $a_i^* > 0$ 的训练样本输入 x_i 称为支持向量,由式(4)和(5)得,在SVM中最终起主要决定作用的只是少量的支持向量样本^[10]。常用的核函数有线性核函数 $K(x, x') = x \cdot x'$ 、多项式核函数 $K(x, x') = (x \cdot x' + 1)^d$ 、高斯核函数 $K(x, x') = \exp(-0.5 \times (\|x - x'\|/s)^2)$ 和 Sigmoid 核函数 $K(x, x') = \tanh(v(x \cdot x') + c)$ 等。核函数的不同会导致分割效果的不同,所以核函数的选择在分割效果中也很重要。高斯核函数因泛化能力好而应用广泛,但随着核函数的研究工作不断深入,核函数的选择也越来越多。在实际应用中,一般选取线性核函数来使用SVM。

2 基于SVM和形态学的叶片病斑分割方法

SVM已经广泛应用于图像分割中^[11-15]。本文利用SVM和形态学中的开闭运算进行叶片中病斑图像分割,分割流程图见图1。

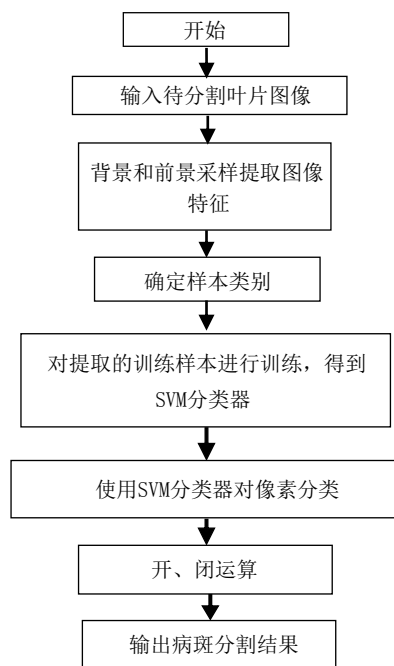


图1 基于SVM的图像分割流程图

分别对待分割叶片图像的正常叶片部分和病斑部分的像素进行标记,选取图像的正常样本和病斑样本的颜色特征,生成训练集;选取合适核函数和惩罚系数,进行SVM训练,生成SVM的训练模型,得到支持向量,再逐一提取每个像素点的特征,产生样本待分类样本集,计算每一个样本到对应超平面的距离值,然后将像素归为正常和病斑,完成对病斑图像的分割。

3 实验与分析

在RGB颜色空间选取作物病害叶片的绿色正常部位和病斑部位各20个像素点作为训练样本,确定训练集,同时将病害叶片的正常部分和病斑部分样本点的像素点的各颜色分类的颜色值、均值和方差作为训练样本集的分割特征,进行叶片正常部位和病斑部位的边缘区分。图2为黄瓜叶片病斑的分割结果。由图2(c)可以看出,分割出来的病斑图像中存在大量噪声,出现了边缘的不连续和叶片中间的小孔等现象。针对这一现象,可以利用数学形态学对病斑图像进一步平滑处理。其基本思想是用一个结构元素作为工具来观察和提取图像特征,判断这个元素是否能够适当的放入图像内部。其中,最基本的操作是膨胀和腐蚀。但在图像分割的实际应用中,更多的以各种组合的形式来表示膨胀和腐蚀。本文为了获得更好的病斑分割效果,进一步对分割的病斑图像进行形态学中的开运算、闭运算:

$$A \circ B = (A \square B) \oplus B \quad \cdots \cdots \cdots (6)$$

$$A \square B = (A \oplus B) \square B \quad \cdots \cdots \cdots (7)$$

式中, $A \circ B$ 和 $A \square B$ 分别为开运算和闭运算, $\square$ 和 $\oplus$ 分别表示形态学腐蚀和膨胀运算。

利用式(6)的开运算可以消除病斑图像中的小孔和毛刺及病斑边缘的不连续;利用式(7)的闭运算进行病斑图像中的小洞删除。由此可以有效消除病斑图像中的边缘的不连续性、病斑内部的小噪声和病斑内部的小洞。在本图像中,经过多次尝试,选取圆盘型结构元素。图2(d)为经过开、闭运算后的病斑分割结果。由图2(c)与图2(d)比较可以看出,图2(d)的病斑分割效果较好,病斑图像明显更趋于人眼的视觉习惯。

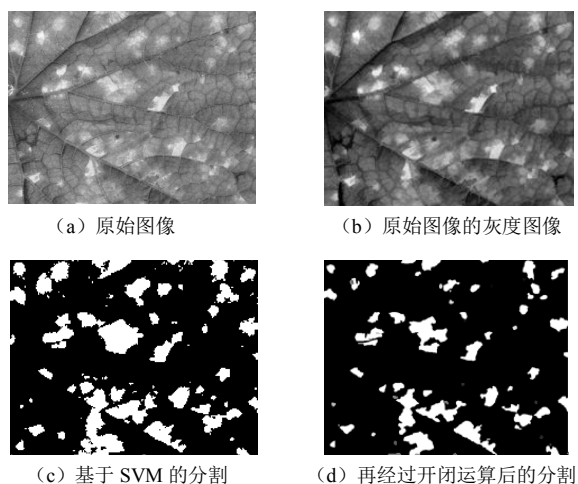


图2 基于SVM和形态学的病斑分割结果

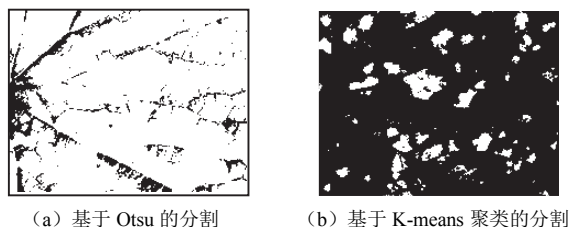


图3 基于Otsu和K-means聚类的病斑分割结果

为了验证所设计方法的有效性,将所提出的基于SVM的病斑分割方法与传统的阈值法Otsu和K-means聚类^[16-18]的分割方法进行比较,其中Otsu为采用最大津法自动选定阈值,K-means方法中引入了空间信息的特征值。图3为采用不同方法进行病斑分割的结果。表1为不同方法进行病斑分割时各图像的误分割率,其中黄瓜细菌性角斑病为在图库中的32幅黄瓜细菌性角斑病图像进行分割时的误分割率均值。由图2、图3和表1可以看出,所采用的方法在黄瓜叶片病斑分割

上具有较好的鲁棒性,分割性能受光照,噪声影响相对较小。

表1 三种方法分割病斑图像的结果比较

分割结果	本文方法	Otsu 阈值法	K-means 聚类法
分割区域数	78	652	236
误分割率	5.6%	37.9%	15.7%

4 结 论

由于作物致病病原物种类不同而使其病害叶中的病斑部分的纹理、形状、颜色有所不同,由此可以利用计算机视觉技术进行作物病害叶片分割、自动识别及病害诊断。病斑图像分割是病斑识别和诊断的基础。本文提出了一种基于SVM和形态学的叶片病斑图像分割方法。首先利用SVM进行病斑图像分割,然后对分割出的病斑图像进行开、闭运算,消除病斑图像中的边缘的不连续性、病斑内部的小噪声和小洞。通过对黄瓜细菌性角斑病图像进行试验,结果表明所提出分割方法具有较好的鲁棒性,与传统的采用阈值和聚类进行分割的方法相比,本文的分割效果得到了明显提高。

参考文献:

- [1] Camargo A, Smith J S. An image-processing based algorithm to automatically identify plant disease visual symptoms[J]. Biosystems Engineering, 2009, 102(1): 9-21.
- [2] Tushar H Jaware, Ravindra D Badgujar, Prashant G Patil. Crop disease detection using image segmentation[J]. World Journal of Science and Technology, 2012, 2(4): 190-194.
- [3] 赖军臣,汤秀娟,谢瑞芝,等.基于G-MRF模型的黄瓜叶斑病害图像的分割[J].中国农业科学,2010,43(7):1363-1369.
- [4] 邵乔林,安秋.基于邻域直方图的黄瓜田绿色植物图像分割方法[J].江西农业学报,2011,23(5):126-128.
- [5] 陈占良,张长利,沈维政,等.基于图像处理的叶斑病分级方法的研究[J].农机化研究,2008(11):73-75.
- [6] 李宗儒,何东健.基于手机拍摄图像分析的苹果病害识别技术研究[J].计算机工程与设计,2010,31(13):3051-3053.
- [7] 薛志东,王燕,李利.SVM图像分割方法的研究[J].微计算机信息,2007,23(8):306-308.
- [8] Vapnik V. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1995.
- [9] Tong S, Koller D. Support vector machine active learning with applications to text classification[J]. The Journal of Machine Learning Research, 2002(2): 45-66.
- [10] 邓乃杨,田英杰.数据挖掘中的新方法-支持向量机[M].北京:科学出版社,2005.

(下转第60页)

理成分。在世界动物地理区划上蒙古高原处于古北界范围之内,蒙古高原合垫盲蝽以古北界成分占主体。在中国动物地理区划上,蒙古高原的中国部分(主要为内蒙古)跨蒙新、华北、东北3个区,其分布的合垫盲蝽也应以这3个区成分为主。而本文的分析结果与蒙古高原的自然地理及环境条件是相吻合的。但由于蒙古高原跨越3个国家,地形地貌及生态环境复杂多样。目前,在许多条件较差的地区还未进行系统的调查与采集,甚至完全是空白。因此,已知种类还有很大的增加分布记录的空间,未知种类也有待于发现,特别是对蒙古国的种类还需更大范围的调查和采集。本文对蒙古高原合垫盲蝽亚科昆虫种类组成及区系分析仅为初步结果,有待以后进一步完善。

参考文献:

- [1] Zhang X Y. NDVI spatial pattern and its differentiation on the Mongolian Plateau[J]. *Journal of Geographical Sciences*, 2009, 19 (4): 403-415 .
- [2] Cassis G, Schuh R T. Systematics, biodiversity, biogeography, and host associations of the Miridae (Insecta: Hemiptera: Heteroptera: Cimicomorpha)[J]. *Annual Review of Entomology*, 2012 (57): 377-404 .
- [3] Schuh R T. Systematic Catalog of Plant Bugs (Insecta: Heteroptera: Miridae). 2013. Available on-line: <http://research.amnh.org/pbi/catalog/> .
- [4] 石 凯,安瑞军,张 宁. 内蒙古合垫盲蝽亚科昆虫生物地理初探[J]. *环境昆虫学报*, 2013, 35(3): 289-293 .
- [5] Kerzhner I M, Josifov M. Catalogue of the Heteroptera of the Palearctic Region. In: Aukema B., Christian R. (Eds). *Cimicomorpha II* (Vol. 3)[M]. The Netherlands Entomological Society, Amsterdam, 1999: 204-277 .
- [6] Kerzhner I M. Key to the Insects of the Far East of the USSR. In: Lehr P. A. (Ed). *Family Miridae* (Vol. 2)[M]. Nauka, Leningrad, 1988: 778-857 .
- [7] 张荣祖. 中国动物地理[M]. 北京: 科学出版社, 1999: 1-347 .
- [8] 康 乐,马 耀,谢炳仁,等. 内蒙古自治区螽亚目(Tettigoniodea, Orthoptera)昆虫区系及地带性分布的研究[J]. *昆虫分类学报*, 1990, 7(3-4): 157-170 .
- [9] 李鸿昌,马 耀,张卓然,等. 内蒙古蝗总科 Acridoidea 区系组成及其区域分布的研究[J]. *昆虫分类学报*, 1990, 7(3-4): 171-193 .
- [10] 能乃扎布,齐宝瑛. 内蒙古蝽科昆虫的区系分析[J]. *内蒙古师大学报(自然科学汉文版)*, 1987(3): 22-31 .
- [11] 任国栋,杨晋宇,于有志. 内蒙古拟步甲与地理分布[J]. *河北大学学报(自然科学版)*, 2000, 20 (增刊): 18-25 .
- [11] Cortes C, Vapnik V. Support vector machine[J]. *Machine learning*, 1995, 20(3): 273-297 .
- [12] Scholkopf B, Burges B, Smola A. *Advances in Kernel Methods-Support Vector Learning*[M]. Cambridge: MIT Press, 1999 .
- [13] Platt J C. 12 Fast Training of Support Vector Machines using Sequential Minimal Optimization[M]. 1999: 185-208 .
- [14] Shin H, Cho S. Fast pattern selection for support vector classifiers[M]. *Advances in Knowledge Discovery and Data Mining*. Springer Berlin Heidelberg, 2003: 376-387 .
- [15] Cervantes J, Li X, Yu W, et al. Support vector machine classification for large data sets via minimum enclosing ball clustering[J]. *Neurocomputing*, 2008, 71(4): 611-619 .
- [16] Chen Y, Zhou X S, Huang T S. One-class SVM for learning in image retrieval[A]. In: *Processing of Image Processing*, 2001. *Proceedings*[C]. 2001 International Conference on. IEEE, 2001: 34-37 .
- [17] 吴高巍,陶 卿,王 珏. 基于后验概率的支持向量机[J]. *计算机研究与发展*, 2005, 42(2): 196-202 .
- [18] Lin C F, Wang S D. Fuzzy support vector machines[J]. *Neural Networks, IEEE Transactions on*, 2002, 13(2): 464-471 .

(上接第 53 页)