

## 基于支持向量机的小麦条锈病和叶锈病图像识别

李冠林 马占鸿 王海光\*

(中国农业大学 农学与生物技术学院,北京 100193)

**摘要** 为了解决生产中小麦条锈病和叶锈病症状难以区分的问题,提高识别率和精度,提出了一种基于支持向量机和多特征参数的小麦条锈病和叶锈病图像分类识别方法。利用图像裁剪方法获取典型症状的子图像,采用中值滤波算法对图像进行去噪,利用 *K*-means 硬聚类算法实现病斑分割,提取病斑区域的形状、颜色和纹理特征空间的 50 个特征参数,设计支持向量机分类器进行分类识别。根据优选的 26 个特征参数,利用以径向基函数作为核函数的支持向量机对这 2 种小麦锈病图像进行识别。结果表明:训练样本识别率均为 96.67%,测试样本识别率均为 100%;与其他核函数相比,径向基核函数最适合于这 2 种小麦锈病的识别。所提出的基于支持向量机的方法可有效地进行小麦条锈病和叶锈病的图像识别。

**关键词** 小麦条锈病; 小麦叶锈病; 支持向量机; 图像识别; 特征提取; 核函数

中图分类号 TP 391.41; S 126

文章编号 1007-4333(2012)02-0072-08

文献标志码 A

## Image recognition of wheat stripe rust and wheat leaf rust based on support vector machine

LI Guan-lin, MA Zhan-hong, WANG Hai-guang\*

(College of Agronomy and Biotechnology, China Agricultural University, Beijing 100193, China)

**Abstract** It is very important to discriminate wheat stripe rust and wheat leaf rust quickly and accurately for forecast and integrated management of the diseases. In this study, a new method based on supporting vector machine (SVM) and multiple feature parameters of their images was proposed for recognition of two kinds of wheat rusts. Sub-images of visual symptoms were acquired using image cutting. The image de-noising was performed with median filtering. The diseased region was then segmented by *K*-means clustering algorithm. Fifty feature parameters from shape-related, color-related and texture-related features were extracted as inputs of the SVMs to identify the best classification model. The results showed that, using the SVMs with radial basis function (RBF) kernel based on the selected twenty-six features, the recognition rates of wheat stripe rust and wheat leaf rust were both 96.67% for the training sets, and 100% for the tested sets. It was thus evident that RBF kernel function was the most suitable method for image recognition of these two kinds of wheat rusts. The image recognition method based on SVM and multiple features could successfully discriminate wheat stripe rust from wheat leaf rust.

**Key words** wheat stripe rust; wheat leaf rust; support vector machine; image recognition; feature extraction; kernel function

小麦条锈病(由 *Puccinia striiformis* f. sp. *tritici* 引起)和叶锈病(由 *P. recondita* f. sp. *tritici* 引起)是常见的 2 种小麦锈病,发生面积广,危害严

重,这 2 种病害在小麦苗期的发生情况对于病害预测预报和综合治理具有重要意义。但是,小麦苗期发病时,这 2 种病害症状极其相似,容易混淆,在实

收稿日期: 2011-08-20

基金项目: 公益性行业(农业)科研专项经费项目(200903035); 国家自然科学基金项目(31071642)

第一作者: 李冠林,硕士研究生, E-mail: cauli@yahoo. cn

通讯作者: 王海光,讲师,主要从事植物病害流行病学和宏观植物病理学研究, E-mail: wanghaiguang@cau. edu. cn

际生产中难以区分<sup>[1]</sup>。一般地,小麦锈病的诊断主要依靠植保技术人员进行田间实时诊断或通过实验室病原鉴定进行诊断,往往会出现偏差而导致误判。因此,有必要对小麦苗期条锈病和叶锈病的诊断识别进行研究,以获得针对这 2 种病害的快速、准确的诊断识别方法。不同的小麦锈病由于其致病的病原体不同而使病斑的形状、颜色和纹理特征可能存在差异,可以利用图像处理技术得到病斑的分割图像并提取相关特征,采用支持向量机等模式识别方法对病害种类进行识别。

支持向量机(support vector machine, SVM)是一种基于统计学习理论的模式识别方法,在解决小样本、非线性和高维模式识别问题中表现出许多特有的优势<sup>[2]</sup>,其在植物病害图像识别领域已有一些应用报道。田有文等<sup>[3-5]</sup>利用支持向量机,应用形状、颜色和纹理不同特征组合对葡萄病害进行了图像识别研究,应用色度矩特征对黄瓜病害进行了图像识别研究。宋凯等<sup>[6]</sup>利用支持向量机对玉米叶部病害进行了分类识别。任东等<sup>[7]</sup>利用线性组合核函数的支持向量机对黄瓜叶部病害获得了良好的图像识别效果。Camargo 等<sup>[8]</sup>利用支持向量机研究了基于 52 个特征参数的植物病害图像识别问题,得到了最优识别率。但是,基于支持向量机的小麦锈病的图像识别研究尚未见报道。

本研究旨在利用图像处理技术对苗期小麦条锈病和叶锈病图像进行图像预处理、病斑分割以及相关特征提取,以特征参数向量作为支持向量机的输入项对这 2 种小麦锈病进行分类识别研究,利用逐步判别法对每一个特征参数进行贡献度检验,以便优选出有效特征组合,最终得到最优识别率,为苗期小麦条锈病和叶锈病的快速、准确地诊断识别提供可靠依据和途径。

## 1 小麦条锈病和叶锈病图像获取与处理

### 1.1 病害图像获取与预处理

小麦条锈病菌和叶锈病菌均为专性寄生菌,需要采用人工接种的方法诱导小麦发病<sup>[9]</sup>。小麦条锈病发病叶片在中国农业大学植物病害流行病学实验室的人工气候室内通过人工育苗、接种、培养和病叶采

集获得。小麦叶锈病典型症状发病叶片由中国农业科学院植物保护研究所提供。选取表面洁净的小麦病害叶片,采用 CanonA710 IS 数码相机垂直于叶片进行 1~55 cm 广角近距离自动拍摄,以获得清晰的叶部病害正投影图像。图像大小为 2 592 × 1 944,格式为 JPEG,24 位图。

由于小麦叶片为长椭圆形,小麦锈病原始图像中叶片位于图片中间,为了便于图像处理,需要对图像进行裁剪以得到原始图像的子图像。裁剪得到小麦条锈病和叶锈病典型症状子图像各 50 幅,大小为 400 × 300,格式为 JPEG,24 位图。图像采用中值滤波算法进行去噪。

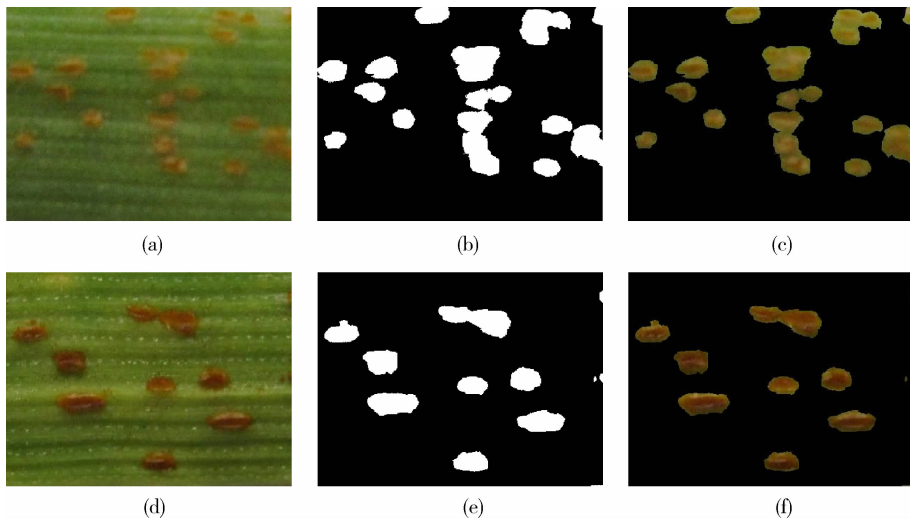
### 1.2 病害彩色图像分割

小麦锈病图像由叶片病斑区域和正常区域组成,图像分割就是要将病斑区域从图像中分割出来,以便进行有效特征的提取。本研究采用 K-means 硬聚类算法(hard C means, HCM)作为小麦锈病图像的分割处理算法,该算法是 Mac Queen 提出的一种非监督实时聚类算法<sup>[10]</sup>,其基本思想是根据图像中色彩空间不同的颜色确定不同色彩所在的区域,从而达到图像分割的目的。具体的病害图像分割实现方法已作报道<sup>[11]</sup>。利用 K-means 硬聚类算法对小麦条锈病和叶锈病图像进行病斑的二值化分割(用于形状特征的提取)和彩色分割(用于颜色和纹理特征的提取),分割结果如图 1 所示。

从病害图像分割结果可以看出,K-means 硬聚类算法能够较为准确地将病斑区域从图像中分割出来,分割效果比较理想,分割精度较高,得到的小麦锈病二值分割图像和彩色分割图像可用于后续病害图像特征提取。

## 2 小麦条锈病和叶锈病图像特征提取

本研究从小麦条锈病和叶锈病图像上提取了形状、颜色和纹理 3 个特征空间的共 50 个特征参数,作为后续分类识别的待选特征参数,包括 4 个形状特征参数、21 个颜色特征参数和 25 个纹理特征参数,其中,颜色灰度特征和纹理特征是基于 RGB 和 HSV 颜色模型的,如表 1 所示。



(a) 小麦条锈病;(b) 小麦条锈病二值分割;(c) 小麦条锈病彩色分割;  
(d) 小麦叶锈病;(e) 小麦叶锈病二值分割;(f) 小麦叶锈病彩色分割。

图1 小麦条锈病和叶锈病图像分割结果

Fig. 1 Results of image segmentation of wheat stripe rust and wheat leaf rust

表1 小麦条锈病和叶锈病图像识别待选的特征参数

Table 1 Candidate features for image recognition of wheat stripe rust and wheat leaf rust

种类	参数	数量
形状	面积( $S$ )	1
	周长( $L$ )	1
	圆度( $C$ )	1
	复杂度( $E$ )	1
颜色灰度	RGB模型的低阶灰度值( $R/G/B$ 、 $R''/G''/B''$ 、 $R'''/G'''/B'''$ )	9
	RGB模型的分量比值( $r$ 、 $g$ 、 $b$ )	3
	HSV模型的低阶灰度值( $H/S/V$ 、 $H''/S''/V''$ 、 $H'''/S'''/V'''$ )	9
纹理	RGB模型的灰度共生矩( $f_1$ 、 $f_2$ 、 $f_3$ 、 $f_4$ ( $d=1, \theta=0^\circ, 45^\circ, 90^\circ$ 和 $135^\circ$ ))	12
	HSV模型的灰度共生矩( $f_1$ 、 $f_2$ 、 $f_3$ 、 $f_4$ ( $d=1, \theta=0^\circ, 45^\circ, 90^\circ$ 和 $135^\circ$ ))	12
	分形盒子维数量( $D$ (盒子大小= $4 \times 4$ ))	1

## 2.1 病斑形状特征提取

鉴于植物病理学知识,本研究采用面积、周长、圆度和复杂度4个形状特征进行病斑形状特征提取。其中,面积和周长是绝对特征参数,圆度和复杂度是相对特征参数。

1)面积( $S$ )。  $S$  为分割图像中所有病斑区域所包含的像素总数,其计算公式如下:

$$S = \sum_{x=0, y=0}^N f(x, y) \quad (1)$$

其中,  $N$  表示病斑区域,  $f(x, y)$  表示病斑区域的单

位坐标。

2)周长( $L$ )。  $L$  为分割图像中所有病斑区域边缘轮廓线上像素总数,其计算公式如下:

$$L = \sum_{i=1}^N \Delta l_i \quad (2)$$

其中,  $N$  表示病斑区域,  $\Delta l_i$  表示病斑区域边缘轮廓线上的单位坐标。

3)圆度( $C$ )。  $C$  描述病斑区域接近圆的程度,介于0和1之间。  $C$  值越大,表示病斑区域越接近圆形。其计算公式如下:

$$C = \frac{4\pi S}{L^2} \quad (3)$$

4)复杂度( $E$ )。 $E$ 描述病斑区域复杂性和离散程度,其值大于1。 $E$ 值越大,表示病斑区域越复杂、离散程度越大。其计算公式如下:

$$E = \frac{L^2}{S} \quad (4)$$

## 2.2 病斑颜色特征提取

图像的颜色信息可用它的矩来表示,图像有用信息一般集中在低阶矩中<sup>[12]</sup>。本研究利用基于RGB颜色模型的 $R$ 、 $G$ 和 $B$ 分量灰度值的一、二阶和三阶矩和各分量的颜色比值 $r$ 、 $g$ 和 $b$ 以及基于HSV颜色模型的 $H$ 、 $S$ 和 $V$ 分量颜色值的一、二阶和三阶矩作为颜色特征参数,其中, $H$ 采用角度量比值表示。 $R$ 、 $G$ 和 $B$ 灰度值是绝对颜色特征参数,其他是相对值参数。

$R$ 、 $G$ 、 $B$ 和 $H$ 、 $S$ 、 $V$ 的一阶矩是指彩色分割图像中病斑区域的颜色平均值,其计算公式如下:

$$\mu = \frac{1}{P} \sum_{j=1}^P f_j p(f_j) \quad (5)$$

$R$ 、 $G$ 、 $B$ 和 $H$ 、 $S$ 、 $V$ 的二阶矩是指彩色分割图像中病斑区域的颜色平方差值,其计算公式如下:

$$\sigma^2 = \frac{1}{P} \sum_{j=1}^P (f_j - \mu)^2 p(f_j) \quad (6)$$

$R$ 、 $G$ 、 $B$ 和 $H$ 、 $S$ 、 $V$ 的三阶矩是指彩色分割图像中病斑区域的颜色立方差值,其计算公式如下:

$$\epsilon^3 = \frac{1}{P} \sum_{j=1}^P (f_j - \mu)^3 p(f_j) \quad (7)$$

以上(5)、(6)和(7)各式中, $P$ 表示病斑区域总像素, $f_j$ 表示病斑区域内的任意坐标, $p(f_j)$ 表示 $f_j$ 点的颜色值。

RGB分量颜色比值 $r$ 、 $g$ 和 $b$ 是指RGB颜色空间各分量的颜色比值,计算公式如下:

$$\begin{aligned} r &= R/(R+G+B) \\ g &= G/(R+G+B) \\ b &= B/(R+G+B) \end{aligned} \quad (8)$$

## 2.3 病斑纹理特征提取

本研究利用基于RGB颜色模型的 $R$ 、 $G$ 和 $B$ 分量灰度共生矩的对比度、相关性、能量和同质性,基于HSV颜色模型的 $H$ 、 $S$ 和 $V$ 分量灰度共生矩的对比度、相关性、能量、同质性以及二值图像 $4 \times 4$ 盒子分形维数作为图像病斑区域的纹理特征参数。灰度共生矩各个特征参数计算值为 $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 和 $135^\circ$

共4个方向计算值的平均值。

1)对比度。图像对比度用于衡量纹理的清晰程度,是图像中某相对位置下两像素灰度差大的情况出现多与少的量度。纹理越深,对比度越小。其计算公式如下:

$$f_1 = \sum_{n=0}^{P-1} n^2 \left\{ \sum_{i=0}^{P-1} \sum_{j=0}^{P-1} p(i,j) \right\} \quad (9)$$

其中

$$n = |i - j|$$

2)相关性。该参数用于衡量灰度共生矩阵的元素在行的方向或列的方向的相似程度,是一种纹理灰度相关统计量,其计算公式如下:

$$f_2 = \sum_{i=0}^{P-1} \sum_{j=0}^{P-1} (i,j)^2 p(i,j) \quad (10)$$

3)能量。能量亦称二阶矩,是图像灰度分布均匀性的度量。当灰度共生矩阵中的元素分布较集中于主对角线时,说明从局部区域观察图像的灰度分布是均匀的。图像纹理越粗,其能力值越大。其计算公式如下:

$$f_3 = \sum_{i=0}^{P-1} \sum_{j=0}^{P-1} p^2(i,j) \quad (11)$$

4)同质性。该参数描述区域内各像素之间纹理的一致性,其计算公式如下:

$$f_4 = \frac{\sum_{i=0}^{P-1} \sum_{j=0}^{P-1} p(i,j)}{1 + (i-j)^2} \quad (12)$$

以上(9)、(10)、(11)和(12)各式中, $P$ 表示病斑区域总像素, $(i,j)$ 表示病斑区域内的任意坐标, $p(i,j)$ 表示 $(i,j)$ 点的颜色值。

5)分形维数。病害图像病斑区域二值分形维数用于衡量病斑区域分形特征,盒子大小为 $4 \times 4$ 。其计算公式如下:

$$D = \lim_{r=0} \frac{\ln N_r(P)}{\ln r} \quad (13)$$

式中, $P$ 表示病斑区域总像素, $r$ 表示病斑区域内的任意像素, $N_r$ 表示盒子大小。

## 3 利用支持向量机进行图像识别

### 3.1 支持向量机基本原理

支持向量机是由Vanpik<sup>[13]</sup>提出的一种新的非常有应用潜力的分类技术,其被认为是目前针对小样本统计估计和预测学习的最佳理论。

在线性可分的情况下,就会存在一个超平面使

得训练样本  $(x_i, y_i)$  完全分开, 最优超平面是使得每一类数据与超平面距离最近的向量与超平面之间的距离最大的平面。在特征数目特别大的情况, 可以利用二次规划的方法求出对应的最优分类函数为<sup>[14-18]</sup>:

$$f(x) = \text{sgn}\{(\omega^* \cdot x) + b^*\} = \text{sgn}\left\{\sum_{i=1}^n a_i^* y_i (x_1 \cdot x) + b^*\right\} \quad (14)$$

其中,  $\text{sgn}$  为符号函数,  $a^* = (a_1^*, \dots, a_n^*)$  是 Lagrange 乘子,  $b^*$  为分类的阈值。

在线性不可分的情况下, 需要引入错分惩罚常数  $c$  和非负的松弛项  $g$ , 通过核函数把样本映射到一个高维特征空间, 并在此空间中运用原空间的函数实现内积运算求出最优超平面。对应的最优分类函数变为:

$$f(x) = \sum_{i=1}^n a_i^* y_i K(x_i \cdot x) + b^* \quad (15)$$

其中,  $K$  为核函数。

常用的 SVM 核函数有:

1) 线性核函数

$$K(x, y) = x \cdot y \quad (16)$$

2) 多项式核函数

$$K(x, y) = [(x \cdot y) + 1]^d \quad (17)$$

3) 径向基核函数

$$K(x, y) = \exp\left\{-\frac{|x - y|^2}{\sigma^2}\right\} \quad (18)$$

4) sigmoid 核函数

$$K(x, y) = \tanh[K(x \cdot y) + m] \quad (19)$$

### 3.2 SVM 分类器设计

1) 特征参数向量数据的归一化处理。为了加快数据训练时的收敛性, 统一样本的统计分布性, 需要对数据进行归一化处理, 一般地是在  $0 \sim 1$  之间进行概率分布的统计。

2) 特征参数向量的降维处理。由于本研究采用的特征参数较多, 属于高维空间的分类问题, 需要对特征参数向量进行降维处理, 以实现高维向低维的转化, 将原特征组合转化为新的特征组合。

3) SVM 参数的寻优选择。SVM 的识别效果受到核函数、 $c$  和  $g$  等参数的影响。为了得到最优识别率, 需要对 SVM 各个参数进行寻优选择。

4) 分类识别。通过上述特征数据处理和参数选择之后, 对样本进行训练识别分类和测试识别分类。

5) 特征参数优选。在 SVM 分类识别过程中, 不同的特征参数对识别效果的贡献不一样, 需要在识别过程中利用逐步判别法对待选的特征参数贡献度进行检验, 以选出最佳特征参数组合, 从而得出最优识别率。

在本研究中, 任意选取小麦条锈病和叶锈病  $400 \times 300$  像素的子图像 60 幅作为训练样本 (2 种病害各 30 幅), 其余 40 幅 (2 种病害各 20 幅) 作为测试样本, 对每幅图像进行预处理、病斑分割和特征提取, 利用单一特征参数、不同特征空间的参数组合和 50 个特征参数总组合分别作为 SVM 的输入项, 通过上述设计的分类方法对病害图像进行分类识别, 并利用逐步判别法对每一个特征参数进行贡献度检验, 以选出最优特征参数组合, 实现小麦条锈病和叶锈病的图像识别, 得出最优识别率。并利用不同的 SVM 核函数对样本进行分类测试, 以确定哪种核函数最适合于小麦条锈病和叶锈病的分类识别。

## 4 结果与分析

### 4.1 特征优选

本研究在 MATLAB 7.6 中借助 SVM 工具箱, 以默认的径向基函数作为核函数, 利用上述方法对小麦条锈病和叶锈病图像样本的分类识别结果如表 2 所示。其中, 以单一特征参数作为输入项得到的识别率记为单一识别率, 以特征参数组合作为输入项得到的识别率记为组合识别率, 50 个特征参数的识别率记为总识别率, 优选的特征参数识别率记为最优识别率。表 2 可见, 每个单一特征的识别率不同, 有些特征参数的单一识别率较高, 有些单一特征参数的单一识别率较低。而且不同的特征参数组合的识别率也不相同。颜色特征组合中二阶矩的识别效果最好; 纹理特征组合中, 基于 HSV 颜色模型的灰度共生矩的识别效果比基于 RGB 颜色模型的识别效果好。

识别结果表明: 没有贡献度的特征参数有 18 个, 负贡献度的特征参数有 6 个, 正贡献度的特征参数有 26 个, 利用这 26 个正贡献度的特征参数作为 SVM 的输入项, 可以使识别效果达到最理想的效果。利用 SVM 对训练样本的学习可以得出最优参数为:  $c=64, g=0.0039$  (表 2)。对 60 幅训练样本的最优识别率为 96.67%, 对 40 幅测试样本的最优识别率达到了 100%。

表 2 小麦条锈病和叶锈病图像识别特征优选结果

Table 2 Results of selecting features for image recognition of wheat stripe rust and wheat leaf rust

序号	特征组合	特征参数	单一识别率/%	组合识别率/%	总识别率/%	交叉检验			
1	形状	面积	93.33/90.00	93.33/92.50	95.00/100	○			
2		周长	91.67/90.00			○			
3		圆度	85.00/90.00			○			
4		复杂度	83.33/92.50			○			
5	颜色一阶矩	$R$ 一阶矩	60.00/45.00	100/67.50		○			
6		$G$ 一阶矩	58.33/75.00			○			
7		$B$ 一阶矩	73.33/72.50			+			
8		比值 $r$	76.67/60.00			○			
9		比值 $g$	66.67/47.50			+			
10		比值 $b$	81.67/57.50			+			
11		$H$ 一阶矩	75.00/55.00			+			
12		$S$ 一阶矩	73.33/42.50			+			
13		$V$ 一阶矩	50.00/22.50			+			
14		颜色二阶矩	$R$ 二阶矩			63.33/35.00	95.00/90.00		+
15			$G$ 二阶矩			68.33/60.00			○
16			$B$ 二阶矩			66.67/50.00			+
17			$H$ 二阶矩			86.67/85.00			+
18	$S$ 二阶矩		68.33/52.50	-					
19	$V$ 二阶矩		56.67/40.00	○					
20	颜色三阶矩	$R$ 三阶矩	80.00/62.50	81.67/75.00		+			
21		$G$ 三阶矩	76.67/57.50			○			
22		$B$ 三阶矩	68.33/60.00			+			
23		$H$ 三阶矩	81.67/77.50			-			
24		$S$ 三阶矩	85.00/55.00			-			
25		$V$ 三阶矩	93.33/65.00			○			
26	RGB 纹理	$R$ 对比度	88.33/92.50	98.33/87.50		+			
27		$R$ 相关性	70.00/60.00			-			
28		$R$ 能量	91.67/92.50			○			
29		$R$ 同质性	90.00/95.00			+			
30		$G$ 对比度	88.33/92.50			○			
31		$G$ 相关性	73.33/62.50			+			
32		$G$ 能量	93.33/90.00			○			
33		$G$ 同质性	85.00/95.00			+			
34		$B$ 对比度	91.67/97.50			-			
35		$B$ 相关性	71.67/65.00			+			
36		$B$ 能量	98.33/80.00			+			
37		$B$ 同质性	93.33/97.50			+			

续表

序号	特征组合	特征参数	单一识别率/%	组合识别率/%	总识别率/%	交叉检验	
38	HSV 纹理	$H$ 对比度	88.33/95.00	100/95.00		○	
39		$H$ 相关性	86.67/100			+	
40		$H$ 能量	91.67/92.50			+	
41		$H$ 同质性	90.00/95.00			○	
42		$S$ 对比度	88.33/95.00			+	
43		$S$ 相关性	88.33/87.50			+	
44		$S$ 能量	91.67/92.50			○	
45		$S$ 同质性	91.67/95.00			+	
46		$V$ 对比度	88.33/92.50			+	
47		$V$ 相关性	71.67/65.00			-	
48		$V$ 能量	91.67/92.50			+	
49		$V$ 同质性	90.99/97.50			○	
50		分形维数	分形维数	93.33/90.00	93.33/90.00		+

注:  $R, G$  和  $B$  分别为 RGB 颜色模型的  $R, G$  和  $B$  分量;  $r, g$  和  $b$  分别为 RGB 颜色模型的分量的颜色比值;  $H, S$  和  $V$  分量分别为 HSV 颜色模型的  $H, S$  和  $V$  分量。对识别结果有正贡献记为“+”;对识别结果无贡献记为“○”;对识别结果有负贡献记为“-”。

## 4.2 不同核函数的识别

利用不同核函数的 SVM 对小麦条锈病和叶锈病图像样本进行分类识别,对基于不同核函数的 SVM 分类性能进行比较,识别结果如表 3 所示。结果表明,对于小麦条锈病和叶锈病,径向基核函数的分类性能最好且稳定,对这 2 种病害样本训练、测试的平均识别率达到 98.34%;Sigmoid 核函数和线性核函数的分类性能次之,平均识别率分别为

96.25%和 93.75%;多项式核函数的分类性能最差,平均识别率仅为 65.42%。所以,基于径向基核函数的 SVM 图像识别方法最适合于基于多特征的小麦条锈病和叶锈病的分类识别。

表 3 可示,选择基于径向基核函数的 SVM 识别方法,利用优选的 26 个病斑区域特征参数作为 SVM 的输入项可以有效实现小麦条锈病和叶锈病的分类识别,训练样本识别率均为 96.67%,测试样

表 3 基于不同核函数的 SVM 对小麦条锈病和叶锈病的图像识别结果

Table 3 Results of image recognition of wheat stripe rust and wheat leaf rust using the SVMs with different kernel functions

核函数	病害类型	样本数量		正确识别数量		识别率/%	
		训练样本	测试样本	训练样本	测试样本	训练样本	测试样本
线性函数	小麦条锈病	30	20	30	15	100	75
	小麦叶锈病	30	20	30	20	100	100
多项式函数	小麦条锈病	30	20	30	20	100	100
	小麦叶锈病	30	20	8	7	26.67	35
径向基函数	小麦条锈病	30	20	29	20	96.67	100
	小麦叶锈病	30	20	29	20	96.67	100
Sigmoid 函数	小麦条锈病	30	20	29	20	96.67	100
	小麦叶锈病	30	20	28	19	93.33	95

本识别率均为 100%。

## 5 结论与讨论

1) 利用基于  $K\_means$  硬聚类算法的图像分割方法实现了小麦条锈病和叶锈病的病斑区域分割, 得到了病斑的二值分割图像和彩色分割图像, 较为准确地将病斑区域从图像中分割出来, 分割效果比较理想, 分割精度高。

2) 提取了小麦条锈病和叶锈病病斑区域形状、颜色和纹理特征空间的 50 个特征参数, 并设计了 SVM 分类器, 利用提取的特征参数和参数组合作为 SVM 的输入项对小麦条锈病和叶锈病图像样本进行了分类识别, 并利用逐步判别法对每个特征参数的贡献度进行了检验, 优选出 26 个有效特征参数, 利用这 26 个特征参数对小麦条锈病和叶锈病图像样本进行识别得出最优识别率, 对 60 幅训练样本的最优识别率为 96.67%, 对 40 幅测试样本的最优识别率达到了 100%。

3) 利用基于不同核函数的 SVM 对样本进行了分类识别, 结果表明, 对于小麦条锈病和叶锈病, 径向基核函数的分类性能最好且稳定。综合考虑, 选择基于径向基核函数的 SVM 识别方法, 利用优选的 26 个病斑区域特征参数可有效实现小麦条锈病和叶锈病的快速、准确诊断识别提供了一种方法和途径。

4) 利用支持向量机对小麦条锈病和叶锈病的图像识别是基于小样本的识别分类, 特征量越多对分类器的要求就越高, 所以, 以尽量少的特征参数作为支持向量机的输入项。本研究提取的 50 个特征可能存在重复现象, 得到的具有正贡献度的有效特征参数中也可能存在表义相似的特征, 对这些特征参数的精细筛选以确定更少的有效特征参数量还需进一步研究。

5) 所用小麦条锈病和叶锈病图像均为通过人工接种获得的小麦苗期单一病害发病的图像, 不存在小麦条锈病和叶锈病混合发生的情况, 也不存在与其他病害混合发生的情况, 并且所用的病害图像均为症状明显时的小麦叶片图像, 所以, 本研究所采用的小麦条锈病和叶锈病图像识别方法对于病害混合发生或病害处于潜育状态、没有明显症状表现时获

得的病害图像是否可以识别尚需进一步研究和验证。

致谢 感谢中国农业科学院植物保护研究所刘太国副研究员为本研究提供了小麦叶锈病典型症状发病叶片。

## 参 考 文 献

- [1] 侯明生, 黄俊斌. 农业植物病理学[M]. 北京: 科学出版社, 2007
- [2] 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
- [3] 田有文, 李天来, 李成华, 等. 基于支持向量机的葡萄病害图像识别方法[J]. 农业工程学报, 2007, 23(6): 175-179
- [4] 田有文, 张长水, 李成华. 基于支持向量机和色度矩的植物病害识别研究[J]. 农业机械学报, 2004, 35(3): 95-98
- [5] 田有文, 牛妍. 支持向量机在黄瓜病害识别中的应用研究[J]. 农机化研究, 2009, 31(3): 36-39
- [6] 宋凯, 孙晓艳, 纪建伟. 基于支持向量机的玉米叶部病害识别[J]. 农业工程学报, 2007, 23(1): 155-157
- [7] 任东, 于海业, 王纪华. 基于线性组合核函数支持向量机的病害图像识别研究[J]. 农机化研究, 2007, 29(9): 41-43
- [8] Camargo A, Smith J S. Image pattern classification for the identification of disease causing agents in plants[J]. Computers and Electronics in Agriculture, 2009, 66(2): 121-125
- [9] 李振岐, 曾士迈. 中国小麦锈病[M]. 北京: 中国农业出版社, 2002
- [10] Selim S Z, Ismail M A. K-means-type algorithm: a generalized convergence theorem and characterization of local optimality [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, 6(1): 81-87
- [11] 李冠林, 马占鸿, 黄冲, 等. 基于  $K\_means$  硬聚类算法的葡萄病害彩色图像分割方法[J]. 农业工程学报, 2010, 26(增刊 2): 32-37
- [12] 岑喆鑫, 李宝聚, 石延霞, 等. 基于彩色图像颜色统计特征的黄瓜炭疽病和褐斑病的识别研究[J]. 园艺学报, 2007, 34(6): 1425-1430
- [13] Vapnik V N. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995
- [14] 李国正. 支持向量机导论[M]. 北京: 电子工业出版社, 2004
- [15] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167
- [16] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2002
- [17] 范立南, 韩晓微, 张广渊, 等. 图像处理与模式识别[M]. 北京: 科学出版社, 2007
- [18] Chapelle O, Haffner P, Vapnik V N. Support vector machines for histogram-based image classification[J]. IEEE Transactions on Neural Networks, 1999, 10(5): 1055-1064