

# 基于可见光谱和支持向量机的黄瓜叶部病害识别方法研究

李鑫星<sup>1</sup>, 朱晨光<sup>1</sup>, 白雪冰<sup>1</sup>, 毛富焕<sup>1</sup>, 傅泽田<sup>1,2</sup>, 张领先<sup>1\*</sup>

1. 中国农业大学信息与电气工程学院, 北京 100083  
2. 农业部农业信息获取技术重点实验室, 北京 100083

**摘要** 以黄瓜叶部病害作为研究对象, 基于可见光谱反射率差异识别黄瓜叶部病害, 研究基于 SVM 的黄瓜叶部病害识别预测模型。采用小波变换进行数据预处理; 选取 Otsu、边缘分割法和 K 均值聚类三类分割方法进行病斑分割, 比较错分率和运行时间, K 均值聚类方法更适合黄瓜叶部病斑分割; 提取纹理、颜色和形状特征参数, 共 15 个特征参数; 通过交叉验证选择最优参数  $c$  和  $g$ , 对核函数参数进行优化处理, 并通过比较线性核、多项式核、RBF 核等不同核函数情况下 SVM 的正确识别率, 确定 RBF 核 SVM 模式识别方法能够更精准地识别黄瓜叶部病害。并将基于 SVM 与另外两种常见的黄瓜叶部病害识别方法, BP 神经网络和模糊聚类进行比较, 结果表明, 基于 SVM 的识别模型对霜霉病的正确识别率为 95%, 白粉病和褐斑病的正确识别率均为 90%, 平均诊断正确率为 92%; 该模式识别方法识别效果最佳, 运行时间最短, 为基于可见光谱的黄瓜病害识别模型提供参考。

**关键词** 可见光谱; 黄瓜叶部病害; 病害识别; 支持向量机; BP 神经网络

**中图分类号:** O433.4   **文献标识码:** A   **DOI:** 10.3964/j.issn.1000-0593(2019)07-2250-07

## 引言

黄瓜作为人们日常必需的蔬菜之一, 由于病害治理不及时引起的品质下降和产量减少, 严重影响了黄瓜种植的经济效益。常见的黄瓜病害有: 霜霉病、角斑病、白粉病等, 这些病害早期发生时, 就可能出现叶部症状。然而, 在农业实际生产过程中, 农业专家往往不能给农民提供及时的帮助<sup>[1]</sup>。因此, 基于可见光谱技术对黄瓜叶部病斑进行识别, 排除诊断主观性和对专家的依赖性, 进而对病害进行诊断, 成为现代农业病害诊断的一个研究热点。

近年来, 许多专家学者对基于光谱技术的作物叶片病害识别进行了研究。杨兴川<sup>[2]</sup>等探究植物叶片病害与光谱信息的相关性, 并建立植物叶片病害的预测模型。Xie<sup>[3]</sup>采用高光谱技术研究了茄子叶部病害的光谱和纹理特征, 结果表明, 光谱和纹理特征在茄子叶片早期疫病检测中是有效的。Yuan 等<sup>[4]</sup>利用光谱反射率差异区分健康植株与患病植株, 建立了高识别精度的判别模型。Zhang<sup>[5]</sup>等利用光谱技术分别对植

物叶片中某种元素含量进行测定, 为植物叶片病害监测提供了技术手段。基于 SVM 的识别方法在病害识别时训练样本少, 具有很好的分类性能和泛化能力, 是一种常用的模式识别方法。宰松梅<sup>[6]</sup>、Yang<sup>[7]</sup>等建立了基于光谱成像技术的叶片病害 SVM 识别模型。张建华<sup>[8]</sup>提出了一种基于最优二叉树支持向量机的蜜柚叶部病害识别方法, 使训练时间和识别精度都得到优化。

为了能够快速、准确地获取黄瓜叶部病害信息, 本文利用作物染病叶片上病斑与正常叶片在可见光范围的光谱差异, 以黄瓜病害叶片可见光图像为研究对象, 所用图片为天津农科院等合作单位的植保专家协助拍摄, 结合计算机视觉技术来进行病害信息的快速获取, 研究基于可见光图像的黄瓜病害识别方法。建立 SVM 病害识别方法, 其中采用的核函数影响着分类的精准度, 通过交叉验证选择最优参数  $c$  和  $g$ , 对核函数参数进行优化处理, 并比较 3 种核函数情况下 SVM 的正确识别率; 并与 BP 神经网络和模糊聚类识别模型进行对比, 选择识别效果最优的方法, 为基于可见光谱的黄瓜病害识别模型提供参考。

收稿日期: 2017-12-02, 修订日期: 2018-05-15

基金项目: 国家自然科学基金项目(31271618)资助

作者简介: 李鑫星, 1983 年生, 中国农业大学信息与电气工程学院副教授   e-mail: lxxcau@cau.edu.cn

\* 通讯联系人   e-mail: zlx131@163.com

# 1 实验部分

## 1.1 试验数据采集

试验结果是从天津市农科院植保所农业创新基地日光温室 6 号棚中采集到黄瓜褐斑病、白粉病、霜霉病叶片，并由植保专家协助拍摄，每种病害照片各 70 幅。图像采集于晴天，采集时间为：2016 年 5 月 15 日早上 4:00—6:00，图像采集设备为 Cannon SX710HS 数码相机，采用自动模式。在获取图后，将图的尺寸统一处理为 1 200 × 900 并保存为 JPEG 格式。图像处理算法采用 Matlab2017a 编程实现。

## 1.2 样本数据的选择

通过总结分析国内外相关研究现状<sup>[9]</sup>，并结合所采集的黄瓜可见光图像数量，选用训练集与测试集 50:20 的比例。从采集的黄瓜叶部病害图像中，挑选其中病斑特征较明显、正常光照条件下、正常角度下的病害图像，挑选霜霉病、白粉病、褐斑病各 50 幅图像为训练样本，每种病害各 20 幅为测试样本。在作物叶片特征参数研究方面，贾建楠<sup>[10]</sup>提取了面积、周长、矩形度等 10 个黄瓜叶部病斑形状特征，并将其运用于黄瓜病害的识别；丁永军<sup>[11]</sup>提取了面积、矩形度、圆形度等 4 个番茄冠层叶片形状特征参数，为基于多光谱图像分析的番茄叶片营养素含量检测提供了一定的技术支持；Bashish<sup>[12]</sup>提取了植物病斑的熵、惯性矩、逆差矩等 10 个纹理特征，作为后续病害识别的依据，结果正确率达到 93%。Muthukannan<sup>[13]</sup>提取了惯性矩、能量、相关度、逆差矩、面积等 5 个特征参数，进行植物叶部病害识别；Verma<sup>[14]</sup>提取了惯性矩、相关性、能量、熵、R/G/B 一阶矩、R/G/B 二阶矩等水稻叶片特征参数。Singh<sup>[15]</sup>提取了惯性矩、能量、逆差矩、熵等 4 个植物叶片纹理参数，达到了较高的病害正确识别率。邓立苗<sup>[16]</sup>圆形度，RGB 和 HSV 颜色空间的 R, G, B, H, S 和 V 六个分量的一阶矩、二阶矩，熵等玉米叶片参数，实现了玉米品种的识别在总结分析国内外相关文献的基础上，在植物叶片特征参数提取领域内，很多学者提取面积、周长、矩形度、惯性矩、逆差矩，5 个形状特征参数；能量、熵、圆形度、纵横轴比，4 个纹理特征参数；R/G/V 一阶矩、R/G/V 二阶矩，6 个纹理特征参数，因此，本研究提取该 15 个特征值。另外，对 RGB 颜色空间下的霜霉病、褐斑病和白粉病的各分量灰度图进行处理，发现通过 R 和 G 分量灰度图比较清楚看到黄瓜病斑，受光照影响较小。

## 1.3 数据预处理

小波分析具有低熵性、多分辨率、去相关性和选基灵活

性的特点，广泛应用于去除光谱背景噪声、仪器干扰方面。由于经过软阈值函数的作用，小波系数在小波域就比较光滑了，得到的图像很平滑。

因此利用软阈值估计小波系数，软门限去噪滤波器阈值表达式为式(1)，小波系数表达式为式(2)，完成表达式为式(3)。完成软门限去噪算法处理后，各高频子带进行小波反变换。

$$\lambda = \sigma \sqrt{2\log(N)} \tag{1}$$

$$W_{j,k} = \begin{cases} \text{sign}(W_{j,k}) (|W_{j,k}| - \lambda), & |W_{j,k}| \geq \lambda\sqrt{2} \\ 0, & |W_{j,k}| \leq \lambda \end{cases} \tag{2}$$

其中，N 是图像像素个数，σ 为方差。

$$f(i,j) = \text{sign}(y(i,j)) \times (|y(i,j)| - \lambda) = \begin{cases} y(i,j) - \lambda, & y(i,j) > \lambda \\ 0, & y(i,j) = \lambda \\ y(i,j) + \lambda, & y(i,j) < \lambda \end{cases} \tag{3}$$

## 1.4 支持向量机

支持向量机(SVM)是二十世纪九十年代兴起的一种机器学习方法，它遵循结构风险最小化原则，能解决传统机器学习中在小样本、非线性等情形下常见的陷入局部最优以及过学习等问题，对于非线性建模、解决样本量偏少和数据挖掘领域具有很强的能力。SVM 使用的核函数通过将数据映射到高维空间，来解决在原始空间中线性不可分的问题。在 SVM 分类中，不同的核函数，分类精度也不同，本研究通过交叉验证选择最优参数 c 和 g，对核函数参数进行优化处理，并对比 3 种核函数，建立识别效果最好的核函数 SVM 识别模型。

# 2 结果与讨论

## 2.1 基于可见光图像的病斑分割

采用 K 均值聚类算法，使用最大类间方差确定的阈值作为 K 均值聚类算法初始聚类中心的数据筛选参考，K 均值聚类方法的公式如式(4)。采用大津法、Sobel 算子分割和 Canny 算子分割与 K 均值聚类算法进行对比，如图 1 和图 2 所示，其中图中标题“最终结果”是 K 均值聚类算法的效果图。

$$\text{peakDis} = \frac{\sum_{i=0}^M \text{dis}(i)}{M} \tag{4}$$

随机选取 30 幅黄瓜叶部病害样本进行病斑分割，对三种分割结果进行对比分析，如表 1 所示。

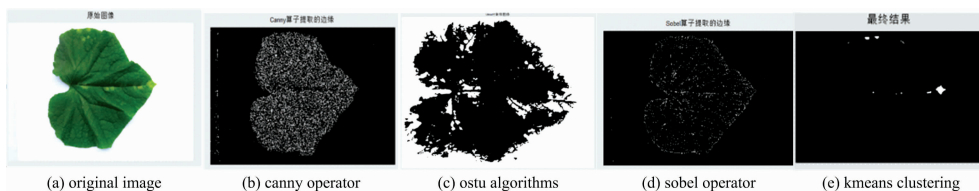


图 1 霜霉病分割方法结果对比

Fig. 1 Comparison of the downy mildew segmentation results

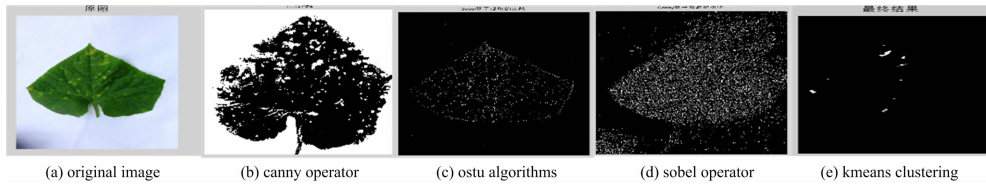


图 2 褐斑病分割方法结果对比

Fig. 2 Comparison of the target leaf spot segmentation results

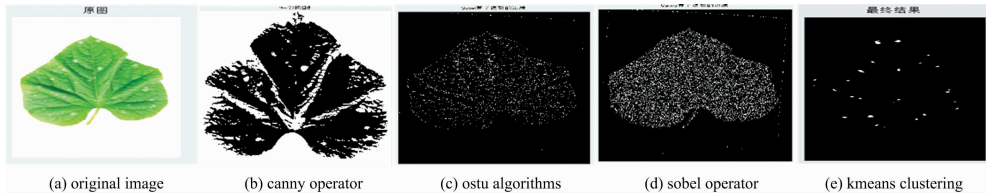


图 3 白粉病分割方法结果对比

Fig. 3 Comparison of the powdery mildew segmentation results

表 1 4 种分割方法效果对比

Table 1 Comparison of segmentation results

分割方法	平均错分率/%	运行时间/s
最大类间方差法	43.71	1.41
Canny 算子边缘分割	59.85	1.37
Sobel 算子边缘分割	52.15	1.23
K 均值聚类	7.27	1.12

对提取出来的病斑数据进行复原,得到彩色病斑图,如图 4 所示。

从图 1—图 4 可看出,基于 K 均值聚类算法能够完整的保留病斑信息,同时也会把部分健康叶片保留下来,虽有可能形成一定的干扰,但并不会影响最终的结果。

## 2.2 基于 SVM 的核函数的选择

### 2.2.1 核函数参数优化处理

使用不同的核函数表现为不同的 SVM 算法。常用的核函数主要有:线性核函数、多项式核函数、径向基核函数等,如下所示:

#### ① 线性核函数

$$K(X, Y) = \langle X, Y \rangle \quad (5)$$

#### ② 多项式核函数

$$K(X, Y) = (\langle X, Y \rangle + c)^q \quad (6)$$

式(5)和式(6)中:  $c$  为常数;  $q$  为多项式指数。当  $c=0$ ,  $q=1$  时,该核函数就是线性核函数。

#### ③ 径向基核函数

$$K(X, Y) = \exp\left\{-\frac{\|X - Y\|^2}{2\sigma^2}\right\} \quad (7)$$

式(7)中:  $\|X - Y\|$  为两个向量之间的距离;  $g = 1/2\sigma^2$  是特定范围内的常数。

### 2.2.2 核函数参数优化处理

在 MATLAB 中可利用程序段随机生成在  $[0, 100]$  范围内的参数  $c$  和在  $[-100, 100]$  范围内的参数  $g$ 。对于取定的  $c$  和  $g$  把训练样本作为原始数据样本利用常见的  $K$  折交叉验证法(K-CV)得到在此组  $c$  和  $g$  下训练集验证分类准确率,

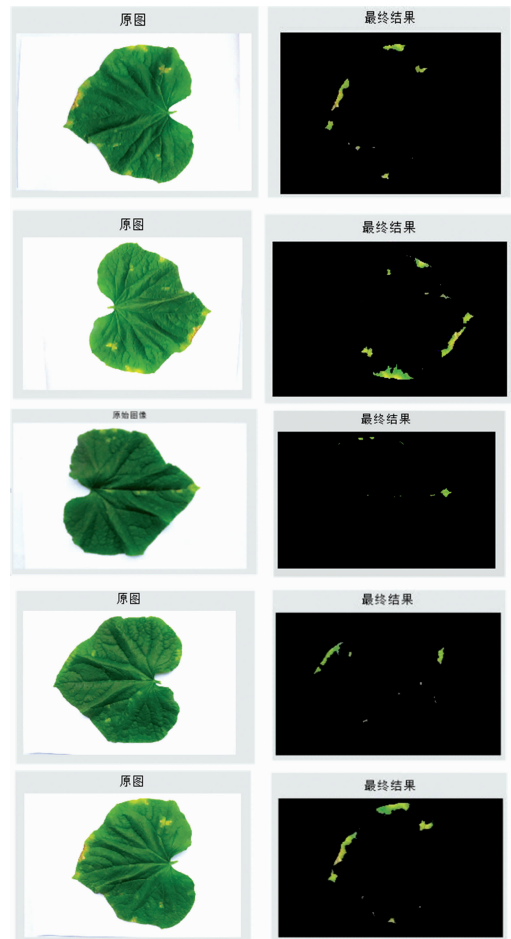


图 4 病斑特征提取结果部分图像

Fig. 4 The detection results of the lesion image

最终取最高准确率的那组  $c$  和  $g$  作为最佳参数,若存在最高准确率相同的情况,选择  $c$  最小时搜索到的第一组  $c$  和  $g$ ,因为过高的  $c$  会导致过学习状态发生。将  $c$  和  $g$  划分网格进行搜索,采用 K-CV 方法将训练样本大致分为  $k$  组,以此时的

训练样本作为测试样本，其他作为训练样本，记录下此时的准确率。由于必须将  $c$  和  $g$  的值进行离散化查找，在这里将  $c$  和  $g$  在 2 的指数范围网格内进行查找。其中训练样本和训练样本标签是必须输入的参数，其他的采用默认值即可。可以先在大的范围内寻找最佳参数  $c$  和  $g$ ，参数粗略选择结果如图 5 所示。

图 5 中  $x$  轴表示  $c$  取以 2 为底的对数后的值， $y$  轴表示  $g$  取以 2 为底的对数后的值，等高线表示的取相应的  $c$  和  $g$  后对应的 K-CV 方法的准确率，从图 5 看出，可将  $c$  的范围缩小至  $2^{-2} \sim 2^1$ ，同时  $g$  的范围可以缩小至  $2^{-4} \sim 2^1$ ，即可进行精细参数选择，参她精细选择结果图如图 6 所示，由图 6 可知， $c$  取值为 0.13， $g$  取值为 0.13 为最优参数。

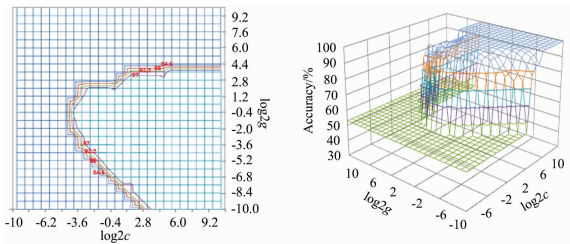


图 5 参数  $c$  和  $g$  粗略选择结果

Fig. 5 Parameter  $c$  and  $g$  rougi selection results chart

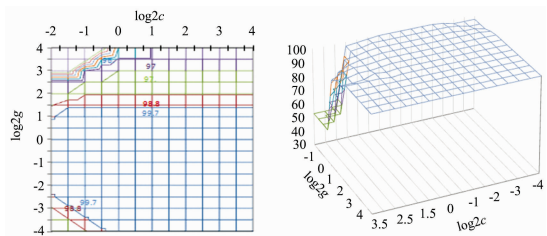


图 6 参数  $c$  和  $g$  精细选择结果图

Fig. 6 Parameter  $c$  and  $g$  fine selection results chart

2.2.3 多种核函数分析对比

此时在不同核函数情况下的正确识别率如图 7 所示。

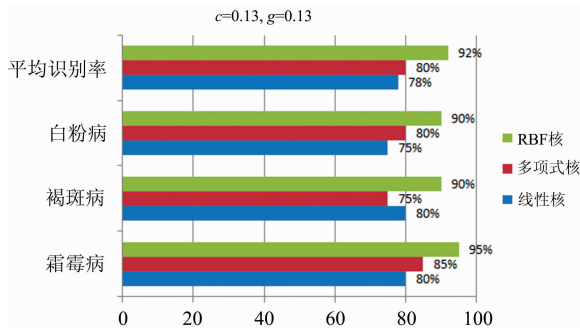


图 7 不同核函数的病害正确识别率 (%)

Fig. 7 The correct identification rate of diseases with different kernel functions(%)

由图 7 实验结果可知，当选用上文中确定的  $c$  为 0.13， $g$  为 0.13 的参数值，并采用 RBF 核函数时，病害平均正确识别率最高 91.667%；因此本研究采用 RBF 核函数作为 SVM 的核函数，惩罚参数  $c=0.13$ ，核函数参数  $g=0.13$ 。

2.3 三种识别方法的模式比较分析

模式识别是对时间和空间不同方面的信息分布，通过某种模型方法，将具体的模式分类到不同的类别中去。随着近些年来学者们对模式识别技术研究不断深入，常用的模式识别方法除了 SVM 识别方法外，还有 BP 神经网络识别方法和模糊聚类识别方法。

利用这 3 种常见模式识别方法分别构建黄瓜叶部病害识别模型，并通过分析比较 3 种不同模式识别方法对黄瓜霜霉病、褐斑病、白粉病的正确识别率和对应的运行时间，验证出最佳的模式识别方法。本研究构建 BP 神经网络模型：首先结合本研究，确定一个 3 层 BP 神经网络，输入层含 15 个神经元(15 个特征参数)，输出层含 3 个神经元(3 种病害)，通过最小二乘法拟合值确定隐含层节点个数的初始值 8，选用试凑法确定最佳的节点数，当隐含层节点数为 16 时，收敛速度适中、误差小，收敛适中且分类误差率小。在 MATLAB 软件的 BP 网络工具箱中包含的多种工具函数，进行 BP 网络分析和设计。本研究构建模糊聚类模型：根据黄瓜病害识别的特征向量及其他约束条件，要将待识未知病害图像分配到各自的模式类中去。试验中分类器的设计分为样本训练和模糊聚类。样本训练就是为了得到标准的模糊模式，使其与未知病害图像数据进行比较，从而对输入的图像数据分类。根据农业植保专家提出的黄瓜这三种病害(霜霉病、褐斑病、白粉病)诊断标准，选择 150 个样本(三种病害各 50 个样本)作为训练样本集，得到训练样本的特征向量矩阵，再把 150 个样本划分成  $k$  个区域(设  $k=3$ )。样本训练结束后，模式库中就保存了已知样本的模式特征，这些特征将作为输入黄瓜未知病害图像的分类依据。利用模糊 C-均值算法对黄瓜病害进行聚类分析，该算法基于最小方差准则的模糊聚类。算法流程如下：(1)输入病害的 150 个样本；(2)提取 15 个特征值，组成特征向量，并将其归一化，组成模糊向量；(3)计算 3 个模式中的聚类中心；(4)输入待识别样本；(5)提取样本的 15 个特征值，将特征向量归一化，转换为模糊向量；(6)计算样本与各聚类中心的贴进度，将其指派到最大贴进度的聚类中心的类中。

2.3.1 正确识别率的分析

将 60 个测试样本(霜霉病、褐斑病、白粉病各随机 20 个)所提取的特征值作为 SVM 和 BP 神经网络和模糊聚类三种分类器的输入，进行各模型验证测试。图 8—图 10 中横坐标表示测试样本数，纵坐标表示病害种类，“1”表示霜霉病实际值，“2”表示褐斑病实际值，“3”表示白粉病实际值；

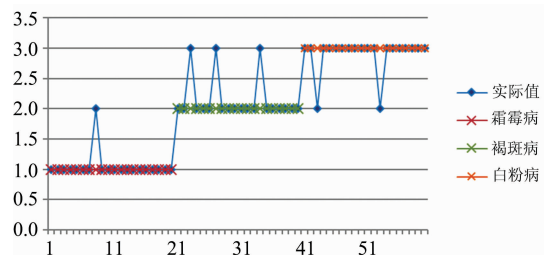


图 8 SVM 病害识别正确识别率对比

Fig. 8 Classification results of SVM identification model

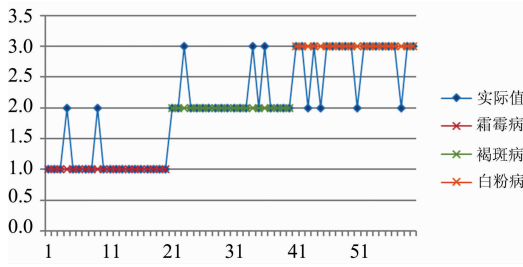


图 9 BP神经网络病害识别正确识别率对比

Fig. 9 Classification results of BP neural network identification model

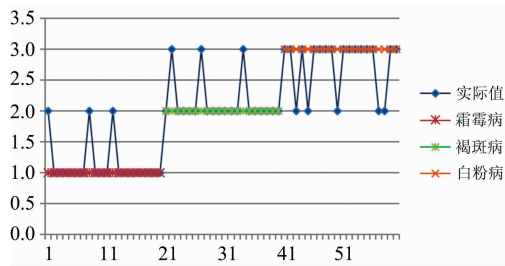


图 10 模糊聚类病害识别正确识别率对比

Fig. 10 Classification results of fuzzy clustering identification model

“蓝色线”表示预测值。以下是基于正确识别率对识别结果进行的比较与分析。

从上述病害分类识别结果图可看出, 基于 SVM 的霜霉病分类识别中 20 个测试数据有 1 个出现分类识别错误, 故正确识别率为 95%; 同理可知, 基于 SVM 的褐斑病分类识别中 20 个测试数据有 2 个出现分类识别错误, 正确识别率为 90%; 白粉病与褐斑病相同, 也有 2 个错误识别, 正确识别率为 90%。由于白粉病与褐斑病在形状、纹理特征方面有较高的相似性, 出现了白粉病和褐斑病的识别错误, 在以后的研究中, 会加强特征参数提取方面的研究。基于 BP 神经网络的霜霉病分类识别中 20 个测试数据中出现了 2 个分类识别错误, 故正确识别率为 90%; 同理可知, 褐斑病分类识别中 20 个测试数据有 3 个出现分类识别错误, 正确识别率为 85%; 白粉病分类识别中 20 个测试数据有 4 个出现分类识别错误, 正确识别率为 80%。基于模糊聚类的霜霉病和褐斑病分类识别中均出现了 3 个分类识别错误, 故正确识别率均为 85%; 白粉病分类识别中 20 个测试数据有 5 个出现分类识别错误, 正确识别率为 75%。

将上述结论进行总结, 如表 2 所示。

### 2.3.2 运行时间的分析

运行时间多少直接影响识别效率的高低, 因此在考虑正确识别率的同时, 运行时间的长短也要作为评判方法优劣的必要因素。

从表 3 可以看出, 识别同种病害时, 无论是白粉病、褐斑病还是霜霉病都是 SVM 运行时间明显加快; 若不考虑病害种类, 仍然是采用支持向量机模式识别方法运行时间最短。

表 2 基于不同分类器 3 种黄瓜病害的正确识别率

Table 2 Correct identification rate of three kinds of identification model

	正确识别数/幅			识别率/%		
	SVM	BP	模糊聚类	SVM	BP	模糊聚类
霜霉病	19	18	17	95	90	85
褐斑病	18	17	17	90	85	85
白粉病	18	16	15	90	80	75
平均	18	17	16	92	85	82

表 3 基于不同分类器的 3 种黄瓜病害识别的运行时间(s)

Table 3 Disease identification operation time three kinds of identification model

	SVM	BP	模糊聚类
霜霉病	3.98	7.31	7.30
褐斑病	4.03	7.47	7.62
白粉病	4.22	8.06	7.45

### 2.3.3 综合比较与分析

综合以上两点分析指标, 通过柱状图的表达方式, 可以更加直观地对不同病害不同识别方法进行比较, 如图 11、图 12 所示。

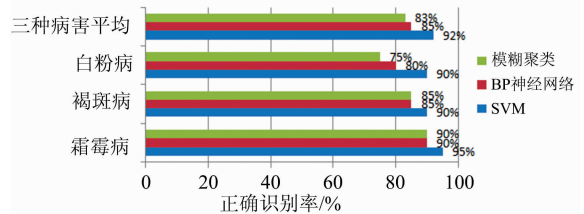


图 11 基于不同分类器 3 种黄瓜病害分类识别的正确识别率 (%)

Fig. 11 The correct recognition rate of three cucumber diseases classification based on different classifiers

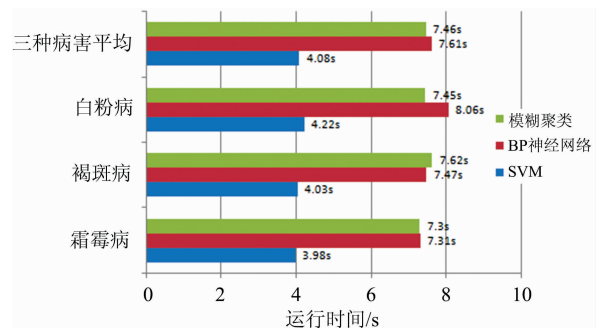


图 12 基于不同分类器的 3 种黄瓜病害识别的运行时间(s)

Fig. 12 The operation times of three cucumber disease identification based on different classifiers

试验结果表明, 对于白粉病, 模糊聚类分类方法识别准确率为 75%, 识别效果较差, 而基于 SVM 分类方法有较好

的识别效果,识别准确率高达 90%;对于褐斑病来说,模糊聚类法和 BP 神经网络的识别准确率均为 85%,基于 SVM 分类方法的识别准确率为 90%;对于霜霉病来说,模糊聚类和 BP 神经网络均有较好的识别效果,识别准确率均为 90%,而基于 SVM 分类方法识别准确率高达 95%,识别效果最佳。对于黄瓜以上三种病害的识别基于 SVM 识别方法准确率均高于另两种识别方法,可见基于支持向量机的模式识别方法是一种可行的识别方法。从运行时间上看,在同样条件下基于 SVM 的三种病害识别速度比 BP 神经网络和模糊聚类识别速度都快,且有一定差距。综合考虑正确识别率和识别速度,分析可得出基于 SVM 的识别方法更适合黄瓜叶部病害识别。

### 3 结 论

研究了基于可见光光谱和支持向量机的黄瓜病害识别方

法。采用小波变换进行去噪处理;结合可见光光谱的特征,对基于  $K$  均值的模糊聚类算法处理后叶部病斑图像进行了特征提取;为了提高黄瓜病害识别的准确率,采用交叉验证法实现了对核函数参数的优化,并分析了多种核函数,根据比较结果,采用正确识别率最高的 RBF 核函数;为了验证基于 SVM 的黄瓜识别模式效果,对 3 种常见黄瓜叶部病害识别模型的模式识别方法(SVM、BP 神经网络、模糊聚类)进行了比较,得出基于 SVM 的识别方法更适合黄瓜叶部病害识别,为基于可见光光谱的设施蔬菜病害诊断的方法研究提供参考。

### References

- [ 1 ] MA Jun-cheng, WEN Han-jie, LI Xin-xing, et al(马浚诚, 温皓杰, 李鑫星, 等). Transactions of the Chinese Society of Agricultural Machinery(农业机械学报), 2017, 48(2): 195.
- [ 2 ] YANG Xing-chuan, LUO Hong-xia, ZHAO Wen-ji, et al(杨兴川, 罗红霞, 赵文吉, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(9): 2873.
- [ 3 ] Xie C, He Y. Sensors, 2016, 16(5): 676.
- [ 4 ] Yuan L, Huang Y, Loraamm R W, et al. Field Crops Research, 2014, 156: 199.
- [ 5 ] Zhang X, Liu F, He Y, et al. Biosystems Engineering, 2013, 115(1): 56.
- [ 6 ] ZAI Song-mei, WEN Ji, GUO Dong-dong, et al(宰松梅, 温季, 郭冬冬, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2011, 27(3): 237.
- [ 7 ] Yang C, Odvody G N, Fernandez C J, et al. Precision Agriculture, 2014, 16(2): 201.
- [ 8 ] ZHANG Jian-hua, KONG Fan-tao, LI Zhe-min, et al(张建华, 孔繁涛, 李哲敏, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2014, 30(19): 222.
- [ 9 ] Mohanty S P, Hughes D P, Salathé M. Frontiers in Plant Science, 2016, 7: 1419.
- [ 10 ] JIA Jian-nan, JI Hai-yan(贾建楠, 吉海彦). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2013, 29(supl): 115.
- [ 11 ] DING Yong-jun, ZHANG Jing-jing, LEE WonSuk, et al(丁永军, 张晶晶, LEE WonSuk, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2017, (9): 32.
- [ 12 ] Bashish D, Braik M, Bani-Ahmad S. Information Technology Journal, 2011, 10(2): 267.
- [ 13 ] Muthukannan K, Latha P, Selvi R P, et al. Journal of Engineering and Applied Sciences, 2015, 10(4): 1913.
- [ 14 ] Verma T, Dubey S. Advances in Computational Sciences and Technology, 2017, 10(5): 721.
- [ 15 ] Singh V, Misra A K. Information Processing in Agriculture, 2017, 4(1): 41.
- [ 16 ] DENG Li-miao, TANG Jun, MA Wen-jie(邓立苗, 唐俊, 马文杰). Journal of Chinese Agricultural Mechanization(中国农机化学报), 2014, 35(6): 72.

# Recognition Method of Cucumber Leaves Diseases Based on Visual Spectrum and Support Vector Machine

LI Xin-xing<sup>1</sup>, ZHU Chen-guang<sup>1</sup>, BAI Xue-bing<sup>1</sup>, MAO Fu-huan<sup>1</sup>, FU Ze-tian<sup>1,2</sup>, ZHANG Ling-xian<sup>1\*</sup>

1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, China

**Abstract** In this paper, we used cucumber leaves disease as the research object, and identified cucumber leaves disease based on the difference of visible spectral reflectance. The support vector machine recognition is an efficient recognition method, which is always used as identification model. For cucumber leaves diseases, if we constructed the support vector machine based on digital image process, we can get accurate and efficient recognition. Consequently, this paper studied the cucumber leaves disease recognition method based on support vector machine. Firstly, the method of wavelet domain denoising was applied to image denoising. The segmentation results were compared with *K* mean clustering, OTSU and edge segmentation. The results showed that *K*-means clustering method was more accurate. We extracted texture, color and shape feature parameters, 15 feature parameters. Then, the optimal parameters of *c* and *g* were selected by cross-validation, and the parameters of the kernel function were optimized and using RBF kernel to construct SVM classifier. By comparing the linearity kernel, polynomial kernel and RBF kernel of the SVM recognition's correct rate, we got that the RBF kernel is most accurate for the recognition of the cucumber leaf disease. Therefore, we used RBF kernel to construct SVM classifier. Finally, there was an identification model of cucumber leaf disease which was based on SVM classifier, and two other efficient identification models, back Propagation neural network, fuzzy clustering identification model. We constructed three kinds of identification models through comparing the correct recognition rate and running time. The results of the test showed that the cucumber downy mildew's correct recognition rate based on SVM classifier was 95%. The correct recognition rate of cucumber powdery mildew and brown spot was 90%, and the average diagnosis accuracy was 92%. In addition, the method running time was the shortest. In summary, the results show that, among the three recognition methods, cucumber leaves disease recognition based on the SVM classifier is the most suitable, demonstrating that the method can be used to rapidly identify cucumber leaves diagnosis based on visual spectrum.

**Keywords** Visible spectrum; Cucumber leaves diseases; Disease recognition; Support vector machine; Back propagation neural network

(Received Dec. 2, 2017; accepted May 15, 2018)

\* Corresponding author