

农艺与调制

李增盛, 孟令峰, 王松峰, 等. 基于图像处理的烟叶烘烤阶段判别模型优选[J]. 中国烟草学报, 2022, 28 (2). LI Zengsheng, MENG Lingfeng, WANG Songfeng, et al. Selection of optimum discriminant model in tobacco curing stage based on image processing[J]. Acta Tabacaria Sinica, 2022, 28(2). doi: 10.16472/j.chinatobacco. 2021.178

基于图像处理的烟叶烘烤阶段判别模型优选

李增盛^{1,2}, 孟令峰¹, 王松峰^{1*}, 高峻³, 徐小洪⁴, 朱先洲³, 杨超⁴, 汪伯军⁴,
王爱华¹, 孟霖¹, 刘自畅¹, 杜海娜^{1,2}, 刘浩^{1,2}, 孙福山¹

1 中国农业科学院烟草研究所, 农业部烟草生物学与加工重点实验室, 青岛 266101;

2 中国农业科学院研究生院, 北京 100081;

3 四川省烟草公司凉山州公司, 四川 西昌 615000;

4 重庆烟草科学研究所, 重庆 400715

摘要:【背景和目的】烟叶烘烤阶段的自动判别是建立智能化烟叶烘烤系统的重要环节。为实现烘烤阶段的精确识别和操控, 提升烟叶烘烤的精准度。【方法】提取烘烤过程中整夹烟叶图像的 11 种颜色特征和 8 种纹理特征, 分别对颜色特征和纹理特征进行变量聚类分析, 以 10 为距离, 将提取的颜色特征和纹理特征各分为 2 类。利用相关性分析筛选出每类特征中与烘烤阶段相关性最强的 1 个特征组成特征子集 (R/G、I*、灰度平均和惯性), 作为模型输入, 分别利用基于遗传算法的支持向量机 (GA-SVM)、基于粒子群算法的反向传播 (PSO-BP) 神经网络和极限学习机 (ELM) 进行烟叶烘烤阶段的分类识别研究。【结果】以优选后 4 个图像特征作为模型输入时, 所建立的 GA-SVM 模型的测试集判别准确率为 93.27%, PSO-BP 神经网络模型的测试集判别准确率为 89.35%, ELM 模型的测试集判别准确率为 85.05%。【结论】基于遗传算法的 SVM 模型烘烤阶段识别效果优于基于粒子群算法的 BP 神经网络模型, 基于粒子群算法的 BP 神经网络模型识别效果优于 ELM 模型。

关键词: 烤烟; 烘烤阶段; 图像处理; 特征模型; 智能烘烤

目前密集烘烤过程中烟叶的变黄与干燥状态主要依靠肉眼观察和手触摸为基础进行人为判断, 存在较强的主观性, 使得烘烤工艺执行中存在超前或滞后现象^[1-2], 因此实现烟叶烘烤阶段自动精确判别和操控, 提升烟叶烘烤的精准度是目前烟叶烘烤研究的重点。

近几年机器视觉在农业应用上取得了良好成绩^[3], 在烟叶生产中的应用大多集中在烟田病虫害诊断^[4-7]、烟叶成熟度判别^[8-10]、烤后烟的分级识别^[11-13]以及烟叶复烤^[14-16]等方面。前人对烘烤过程中烟叶的形态特征变化进行了初步探索, 段史江等^[17]建立了烘烤过程中烟叶含水率的 BP 神经网络预测模型和基于遗传算法

的 LS-SVM 预测模型, 预测结果与实测值的相关系数分别为 0.894 和 0.998; 张丽英等^[18]进行了烘烤过程中烟叶叶绿素和类胡萝卜素与烟叶颜色特征的逐步回归分析, 所建立的回归方程决定系数 R^2 分别为 0.9766 和 0.6325; 郭朵朵^[19]以烤烟叶片含水量、叶绿素和类胡萝卜素 3 个指标为因变量, 以叶片颜色参数 Lab、长度收缩率和宽度收缩率为自变量, 建立了 3 个线性回归预测模型, 决定系数 R^2 分别为 0.899、0.733 和 0.713。但机器视觉技术在烟叶烘烤阶段判别的应用还鲜见报道。

目前常用的机器学习方法有支持向量机 (SVM)、

基金项目: 中国农业科学院科技创新工程 (ASTIP-TRIC03); 中国烟草总公司重点项目 (110202102007); 中国烟草总公司四川省公司科技重点项目 (SCYC202012); 中国烟草总公司重庆市公司科技项目 (B20202NY1335)

作者简介: 李增盛 (1996—), 研究生, 主要研究方向为烟草智能烘烤技术, Tel: 0532-88702076, Email: 1349466902@qq.com

通讯作者: 王松峰 (1979—), 副研究员, 主要研究方向为烟草智能烘烤技术和设备, Tel: 0532-88701009, Email: wangsongfeng@caas.cn

收稿日期: 2021-09-14; **网络出版日期:** 2022-02-11

反向传播神经网络 (BP 神经网络) 和极限学习机 (ELM) 等。其中 SVM 是一种基于统计学理论和结构最小风险原理基础上的机器学习方法, 能够较好地解决小样本学习、非线性以及高维等实际问题, 在模式识别中得到广泛应用^[20]; BP 神经网络是一种多层前向神经网络, 运用误差方向传播算法^[21], 具有分布式处理、自学习和自组织能力强等优点^[22]; ELM 是一种用于单隐层前馈神经网络的学习算法, 其隐含层节点的权重为随机或人为给定, 且不需要更新, 学习过程仅计算输出权重, 具有泛化能力强、运算时间短的特点^[23]。本研究采用 GA-SVM、PSO-BP 神经网络和 ELM 三种模型对烟叶烘烤阶段进行判别比较, 得出优选模型, 以期图像处理技术在烟叶智能烘烤中的应用提供理论依据和技术支持。

1 材料与方法

1.1 试验材料

试验于 2020 年在四川省凉山州西昌市中国农业科学院西南试验基地进行, 选取肥力中等的试验田, 供试烤烟品种为云烟 87, 行距 1.2 m, 株距 0.5 m, 按照优质烟生产技术规范进行栽培。选取大田长势整齐、落黄一致的成熟烟叶。烘烤烤房为热泵密集烤房, 装烟室规格长宽高为 8.0 m×2.7 m×3.5 m, 装烟 3 层 2 路, 装烟量为 4000~4300 kg。图像采集设备采用专用广角耐高温高清摄像头和标准拍摄用光源。

1.2 图像采集

如图 1 所示, 在烤房中层设置温湿度传感器, 并在正对烟叶 30 cm, 上挡风板高出烟夹 10 cm 的位置安装配备标准光源的图像采集设备 (图 2), 调整摄像头的角度和高度, 每隔 5 min 在线实时采集烘烤过程中的整夹烟叶图像, 上传至服务器。

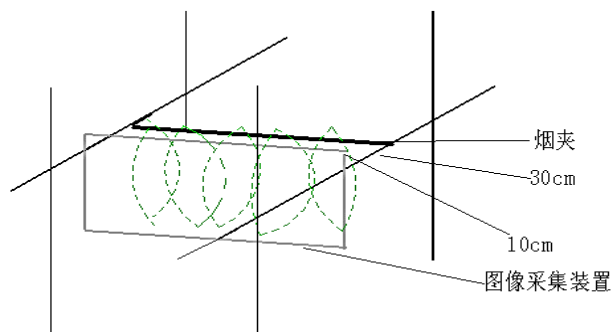


图 1 图像采集装置安装示意图

Fig.1 Schematic diagram of image acquisition device installation

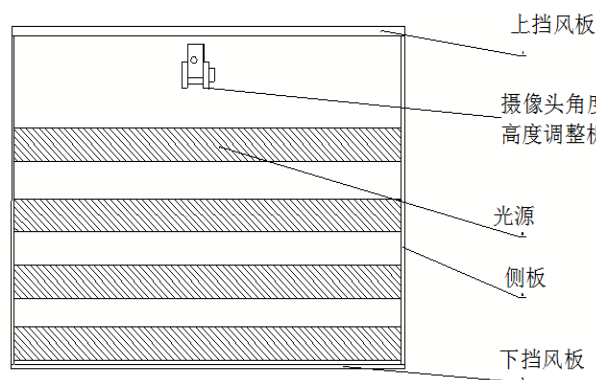


图 2 图像采集装置结构图

Fig.2 Structure Diagram of image acquisition device

1.3 图像预处理

烟叶图像受背景、光照等因素影响产生噪声, 且烘烤过程中烟叶的颜色和形状差异导致图像质量发生变化, 需对采集图像进行改善处理, 增强烟叶图像的有效信息, 提高图像的辨识度。本研究分别采用均值滤波和中值滤波对采集图像进行去噪, 结果表明使用中值滤波器能够改善烟叶图像的质量。图像分割是进行图像分析和理解的关键, 由于本研究中图像采集方式是通过在烤房中装置摄像头直接拍摄, 拍摄环境背景相对复杂, 干扰因素较多, 所以使用 MATLAB 自带的阈值分割程序进行基于 YCbCr 颜色空间的图像分割^[24], 最后运用合成运算获得与背景分割后的完整烟叶彩色图像, 同时对无法通过阈值分割去掉的部分背景进行掩膜操作, 只保留烟叶区域^[25]。

1.4 图像特征提取

烟叶失水皱缩程度和颜色变化情况是判别烟叶烘烤进程的主要依据。烟叶的颜色变化可以用颜色特征来反应, 本研究在 RGB 颜色空间的基础上提取了 RGB 的均值和基于 R、G、B 颜色矩阵的几种组合特征 $2G-R-B$ 、 $G/(R+G+B)$ 、 R/G 和 $G-R$ ^[18]以及基于 $l^*a^*b^*$ 颜色空间的 $l^*a^*b^*$ 的均值和其特征组合 a^*/b^* ^[26]共 11 个颜色特征; 烟叶的皱缩程度可以用图像的纹理特征展现, 本研究提取了烘烤过程中整夹烟叶图像基于灰度梯度共生矩阵的能量、灰度平均、梯度平均、相关度、灰度熵、梯度熵、惯性矩和逆差矩^[27]共 8 个纹理特征。

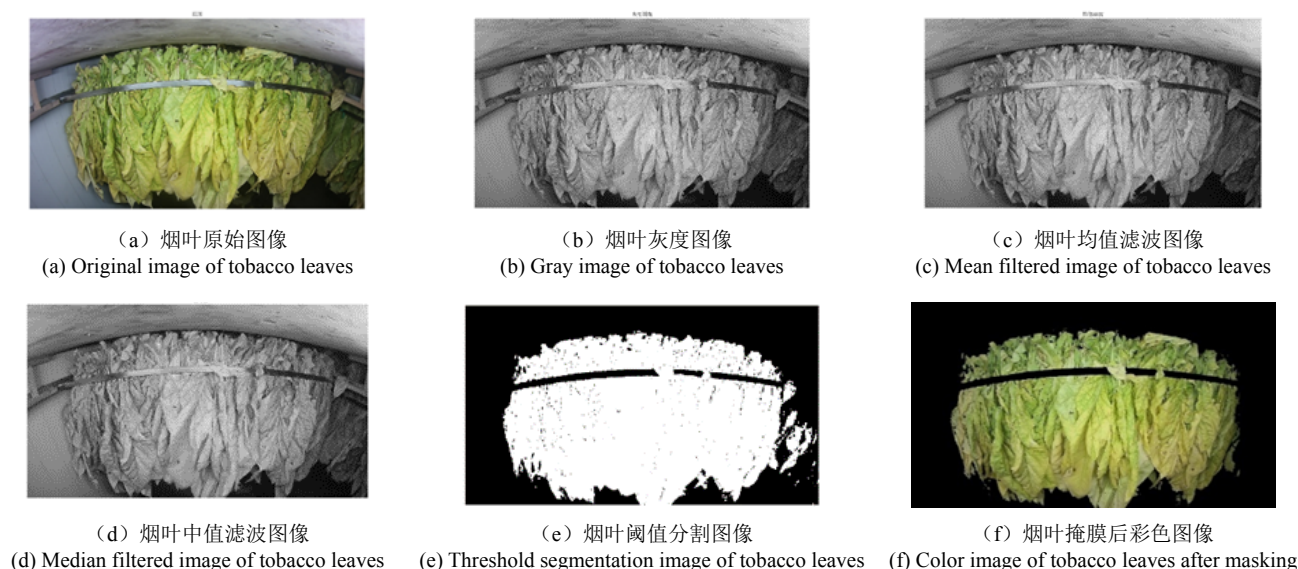


图3 烟叶预处理图像

Fig.3 Preprocessed image of tobacco leaf

1.5 数据处理与判别模型的建立

参考三段式烘烤工艺、五段五对应烘烤工艺^[28]和8点式烘烤工艺^[29]的基础上,从烟叶的变黄到干筋,将整个烘烤过程采集的图像分成10个阶段,即:常温~36℃末(第1阶段)、36℃末~38℃末(第2阶段)、38℃末~40℃末(第3阶段)、40℃末~42℃末(第4阶段)、42℃末~45℃末(第5阶段)、45℃末~47℃末(第6阶段)、47℃末~50℃末(第7阶段)、50℃末~54℃末(第8阶段)、54℃末~60℃末(第9阶段)、60℃末~65℃末(第10阶段)。数据集共计2139个样本,采用隔三选一法,将样本集划分为训练集和测试集,分别包含样品数据组1604个和535个。

为保证模型训练的快速性、准确性以及模型的泛化性,将上述提取的11个颜色特征和8个纹理特征分别进行变量聚类,然后将烟叶图像特征与烟叶烘烤阶段进行相关性分析,筛选出每类特征中与烘烤阶段相关性最强的1个图像特征作为训练样本。选取不同图像特征组合作为模型输入,建立极限学习机(ELM)、基于遗传算法的支持向量机(GA-SVM)和基于粒子群算法的反向传播神经网络(PSO-BP)分类模型对烟叶烘烤阶段进行分类比较。数据处理软件为SAS 9.4和MATLAB 2020a软件。

SVM是一种监督方式的机器学习算法,本文选用SVM核函数中不随参数变化而变化的RBF核函数^[30],同时惩罚因子 c 和核函数中的参数 g 是影响支持向量机性能的两个主要参数,其中 c 表示对误差的宽容度,

其参数的大小对最优分类面的位置有很大影响,参数 g 隐含地决定了数据映射到新的特征空间后的分布^[31]。利用模拟自然界中适者生存的进化现象的遗传算法对支持向量机中 c 和 g 两个参数进行筛选寻优^[32-34],以最高的交叉验证集判别准确率为寻优目标,在遗传算法中将种群数量设置为30,最大迭代次数为50,参数 c 和 g 从0.1~200进行寻优取值^[35]。

BP神经网络是基于误差反向传播的多层前向神经网络。本文以隐含层节点数目5为步长,从0到60对BP隐含层节点数进行寻优。选取premnmx函数,将样本数据进行归一化处理。设定训练误差目标为0.0001,终止训练步数为1000,另外,BP神经网络的初始权值和阈值运用粒子群算法(PSO)进行优化,PSO中的种群大小设置为30,迭代次数设置为50,确保分类的精准度。

ELM是一类用于单隐层前馈神经网络的学习算法。本研究选用Sigmoid函数,以隐层神经元数目20为步长,交叉验证集判别准确率最高为寻优目标进行隐层神经元数目寻优,并用10折交叉法进行验证。

2 结果

2.1 烟叶图像特征优选

2.1.1 烟叶图像特征的聚类分析

利用SAS 9.4将提取的颜色特征和纹理特征分别进行变量聚类分析,结果如图4和图5所示:

以10为距离,可以将提取的11个颜色特征值分

为两类，其中 $G/(R+G+B)$ 、 a^*/b^* 、 $G-R$ 、 a^* 、 $2G-R-B$ 、 b^* 和 R/G 为一类； R 、 G 、 B 和 I^* 为一类。同样以 10 为距离将提取的 8 个纹理特征分为梯度熵、逆差矩、能量、灰度熵、梯度平均、灰度平均、相关和惯性 2 类。

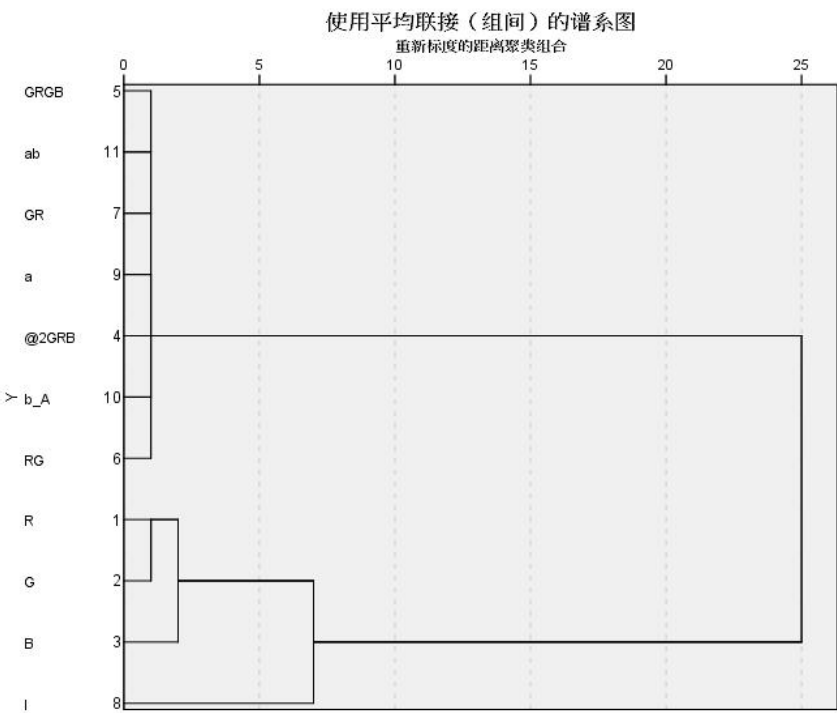


图 4 颜色特征值的聚类分析
Fig.4 Cluster analysis of color feature values

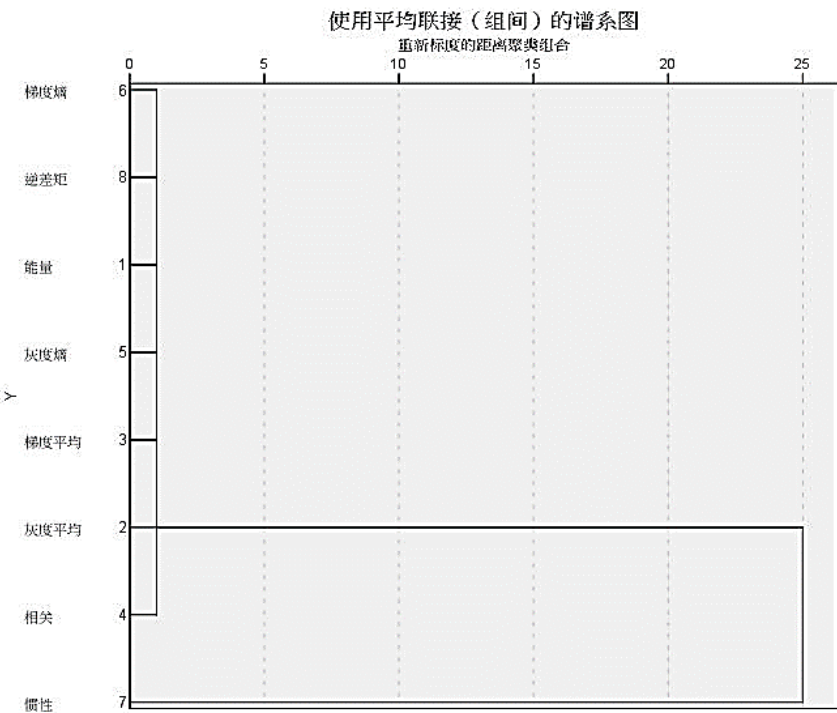


图 5 纹理特征值的聚类分析
Fig.5 Cluster analysis of texture feature values

2.1.2 烟叶图像特征与烘烤阶段的相关性分析

利用 SAS 9.4 将提取的各烟叶特征值作为待选变量，分别与主导变量烘烤阶段进行相关性分析，结果如表 1 所示，烟叶的颜色特征值与烘烤阶段的相关性均达到显著关系，类别 1 中 7 个颜色特征中 R/G 与烘烤阶段的相关性最强，相关性系数达到 0.903；类别 2 中 I*与烘烤阶段相关性系数的绝对值最大，为-0.857。

烟叶的纹理特征值与烘烤阶段的相关性均呈极显著关系，类别 1 中相关性最强的纹理特征为灰度平均，与烘烤阶段呈负相关；类别 2 中惯性的相关性系数为 -0.859。因此，选用 2 个颜色特征 R/G 与 I*和 2 个纹理特征灰度平均与惯性共 4 个烟叶图像特征组成特征子集作为建立烟叶烘烤阶段判别模型的输入变量。

表 1 烟叶图像特征值与温度阶段的皮尔逊相关性
Tab.1 The Pearson Correlation of Tobacco Leaf Image Feature Value and Temperature Stage

颜色特征类别 Color feature category	颜色特征 Color feature	皮尔逊相关性 Pirsson correlation	纹理特征类别 Texture feature category	纹理特征 Texture feature	皮尔逊相关系数 Pirsson correlation
类别 1	G/(R+G+B)	-0.715**	类别 1	梯度熵	-0.739**
	a*/b*	0.859**		逆差矩	0.484**
	G-R	-0.820**		能量	0.440**
	a*	0.757**		灰度熵	-0.570**
	2G-R-B	-0.534**		梯度平均	0.669**
	b*	-0.372**		灰度平均	-0.827**
	R/G	0.903**		相关	0.679**
类别 2	R	-0.681**	类别 2	惯性	-0.859**
	G	-0.842**			
	B	-0.053*			
	I*	-0.857**			

注：**表示在 0.01 水平（双侧）上极显著相关。
Note: ** means extremely significant correlation at the 0.01 level (bilateral).

2.2 烟叶烘烤阶段 GA-SVM 分类模型的建立结果

分别利用提取的全部 19 个图像特征、与烘烤阶段相关性系数的绝对值大于 0.7 的 10 个图像特征和经过聚类分析和相关性分析优选出的 4 个颜色纹理特征子集作为 GA-SVM 模型的输入变量进行烘烤阶段分类判别。其中，以优选后的 4 个图像特征作为模型输入，其参数寻优结果如图 6 所示，寻出最优的 g 和 c 分别为 99.97 和 89.77。

上述 3 种输入情况下建立的 GA-SVM 的分类结果如表 2 所示，当选用全部 19 个图像特征作为 GA-SVM 判别模型的输入时，模型分类效果最佳，训练集、交叉验证集和测试集的判别准确率分别为 99.75%、96.63%和 98.50%；当选用相关性分析后 10 个图像特征作模型的输入时，模型效果次之，训练集的判别准确率为 99.38%，交叉验证集的判别准确率为 96.13%，测试集判别准确率为 97.57%，总体准确率达到 97.69%；当选用优选后的 2 个颜色特征和 2 个纹理特

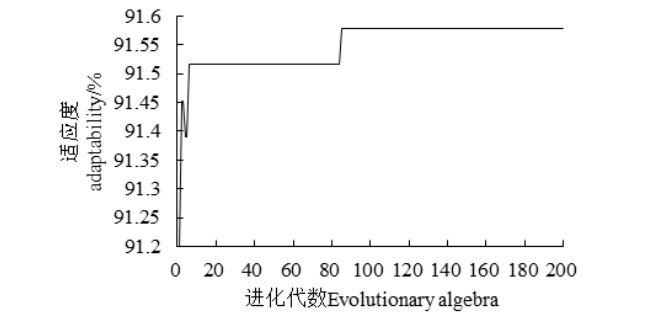


图 6 GA 参数寻优图
Fig.6 Optimization diagram of parameter based on genetic algorithm

征作为模型输入时，模型的训练集、交叉验证集和测试集的判别准确率为 95.57%、91.58%和 93.27%。

以优选后 4 个图像特征作为模型输入，GA-SVM 测试集的烘烤阶段分类情况如图 7 所示，前 4 阶段预测结果最优，第 5 阶段和第 6 阶段次之，第 7 阶段、第 8 阶段、第 9 阶段和第 10 阶段表现略差。

表 2 GA-SVM 分类结果
Tab.2 GA-SVM classification results

分类模型 Classification model	特征数 Feature quantity	模型参数 Model parameters		训练集判别准确率 Discrimination accuracy of training set /%	交叉验证集判别准确率 Accuracy of cross-validation set discrimination /%	测试集判别准确率 Test set discrimination accuracy /%
		误差惩罚因子 C	核函数参数 gamma			
GA-SVM	19	53.51	89.60	99.75	96.63	98.50
GA-SVM	10	91.06	99.54	99.38	96.13	97.57
GA-SVM	4	89.77	99.97	95.57	91.58	93.27

注：(1) 特征数目中，19 为提取的全部图像特征；10 为与烘烤阶段相关性系数的绝对值大于 0.7 的 10 个图像特征；4 为经聚类分析和相关性分析后优选出的 4 颜色纹理特征。(2) 子集判别准确率=子集机器识别正确的样品数目/子集总样品数量*100%。下同。

Note: (1)Regarding the number of features, 19 represents the number of all extracted image features; 10 represent the number of image features with the absolute value of correlation coefficient with the curing stage greater than 0.7; 4 are 4 represent the number of colors selected after cluster analysis and correlation analysis texture features. (2) Accuracy of subset discrimination = the number of samples correctly identified by the subset machine/the total number of samples in the subset*100% .The same below.

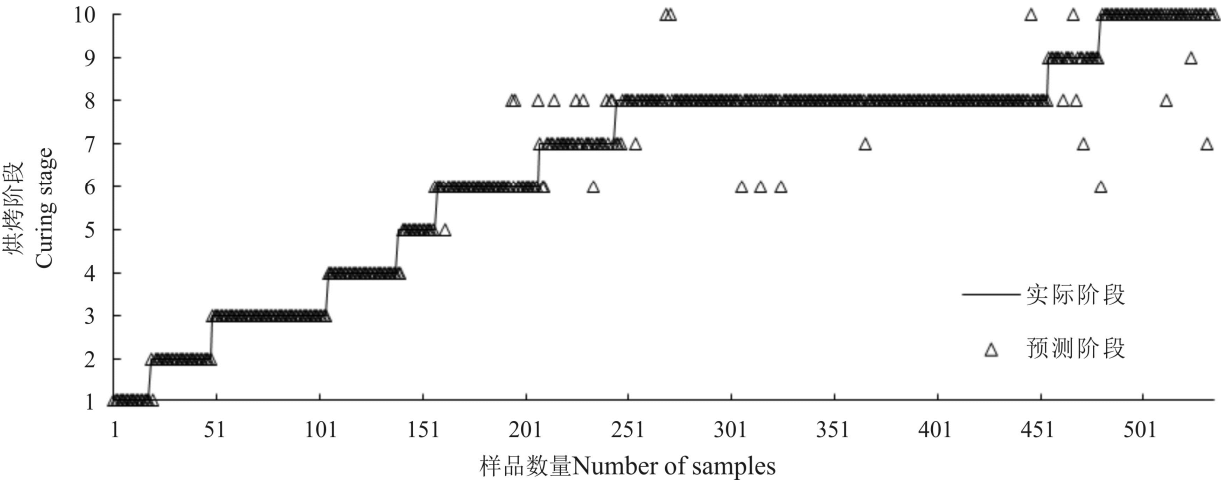


图 7 GA-SVM 测试集的实际分类和预测图
Fig.7 Actual classification and prediction map of GA-SVM test set

2.3 PSO-BP 神经网络模型建立结果分析

分别利用提取的全部 19 个图像特征、与烘烤阶段相关性系数的绝对值大于 0.7 的 10 个图像特征和经过聚类分析和相关性分析优选出的 4 个颜色纹理特征作为 PSO-BP 模型的输入变量进行分类判别时，首先以隐含层节点数目 5 为步长，从 0 到 60 对 BP 隐含层节点数进行寻优，在上述 3 种输入情况下，建立寻优后 BP 模型的隐含层节点数分别为 35、40 和 50，其中以优选后 4 个图像特征作为输入构建模型时，其寻优结果如图 8 所示。

在上述 3 种输入情况下，所建立的 PSO-BP 模型结果如表 3 所示。当选用相关分析后 10 个图像特征作为 PSO-BP 神经网络判别模型的输入时，模型分类效

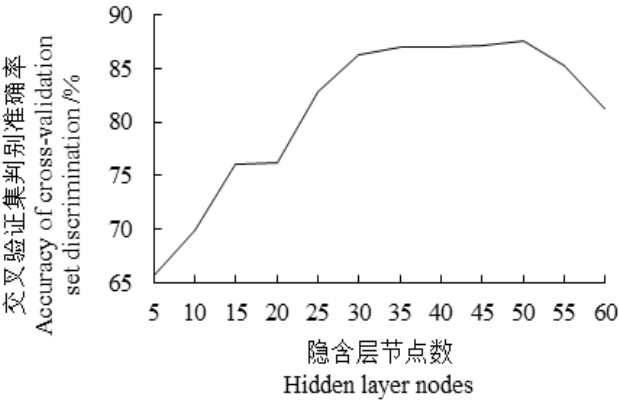


图 8 BP 隐含层节点数寻优图
Fig.8 Selection of the optimum number of nodes in BP hidden layer

果最佳，模型的训练集判别准确率、交叉验证集的判别准确率和测试集的判别准确率分别为 97.57%、95.63%和 95.70%；当选用全部 19 个图像特征作为模型输入时，模型效果次之，训练集、交叉验证集和测试

集的判别准确率分别为 96.38%、97.19%和 95.14%。当选用优选后的 2 个颜色特征和 2 个纹理特征作为模型输入时，模型的训练集、交叉验证集和测试集的判别准确率分别为 91.14%、87.54%和 89.35%。

表 3 PSO- BP 模型分类结果
Tab.3 Classification results of PSO-BP model

分类模型 Classification model	特征数 Feature quantity	隐含层节点数 Hidden layer node number	训练集判别准确率 Discrimination accuracy of training set /%	交叉验证集判别准确率 Accuracy of cross-validation set discrimination /%	测试集判别准确率 Test set discrimination accuracy /%
PSO-BP	19	35	96.38	97.19	95.14
PSO-BP	10	40	97.57	95.63	95.70
PSO-BP	4	50	91.14	87.54	89.35

以优选后 4 个图像特征作为模型输入，PSO-BP 神经网络测试集的烘烤阶段分类情况如图 9 所示，前

4 阶段表现最优，第 5 阶段和第 6 阶段次之，第 7 阶段、第 8 阶段、第 9 阶段和第 10 阶段表现略差。

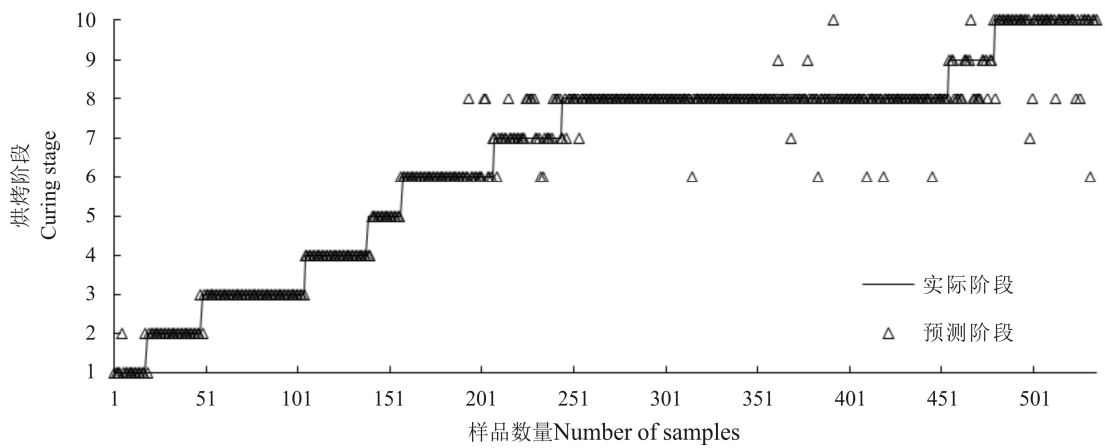


图 9 PSO-BP 测试集的实际分类和预测图
Fig.9 Actual classification and prediction map of PSO-BP test set

2.4 ELM 模型建立结果分析

分别利用提取的全部 19 个图像特征、与烘烤阶段相关性系数的绝对值大于 0.7 的 10 个图像特征和经过聚类分析和相关性分析优选出的 4 个颜色纹理特征子集作为 ELM 模型的输入变量进行分类判别，均以隐层神经元数目 20 为步长，以最高的交叉验证集判别准确率为寻优目标进行隐层神经元数目寻优。在上述 3 种输入情况下，建立寻优后的 ELM 模型的隐层神经元数目分别为 200、220 和 220，其中以优选后 4 个图像特征作为输入构建模型，寻优结果如图 10 所示。

在上述 3 种输入情况下，所建立的寻优后 ELM 模型结果如表 4 所示，在 ELM 模型中当选用全部 19 个图像特征作为判别模型的输入时，模型分类效果最佳，训练集、交叉验证集和测试集的判别准确率分别为

94.26%、92.56%和 94.21%；当选用与烘烤阶段相关性较高 10 个图像特征作模型的输入时，模型效果次之，

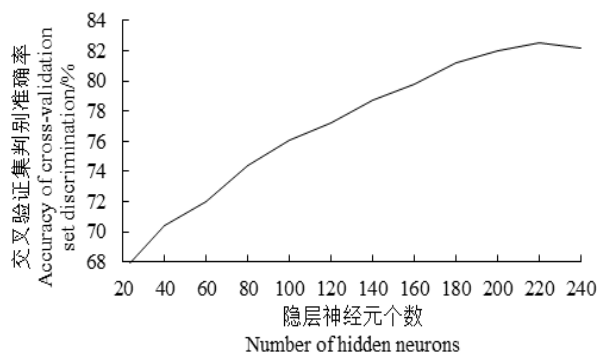


图 10 ELM 隐层神经元数目寻优图
Fig. 10 Selection of the best number of neurons in the hidden layer of ELM

训练集的判别准确率为 92.76%，交叉验证集的判别准确率为 91.06%，测试集判别准确率为 91.96%。当选用优选后的 2 个颜色特征以及 2 个纹理特征作为模型输入时，模型的训练集、交叉验证集和测试集的判别准确率分别为 85.40%、82.53%和 85.05%。以优选后 4

个图像特征作为模型输入，ELM 测试集的烘烤阶段分类情况如图 11 所示，前 4 阶段预测效果最优，第 5 阶段和第 6 阶段次之，第 7 阶段、第 8 阶段、第 9 阶段和第 10 阶段表现略差。

表 4 ELM 模型分类结果
Tab.4 ELM model classification results

分类模型 Classification model	特征数 Feature quantity	隐层神经元数目 Number of hidden neurons	训练集判别准确率 Discrimination accuracy of training set/%	交叉验证集判别准确率 Accuracy of cross-validation set discrimination/%	测试集判别准确率 Test set discrimination accuracy /%
ELM	19	200	94.26	92.56	94.21
ELM	10	220	92.76	91.06	91.96
ELM	4	220	85.40	82.53	85.05

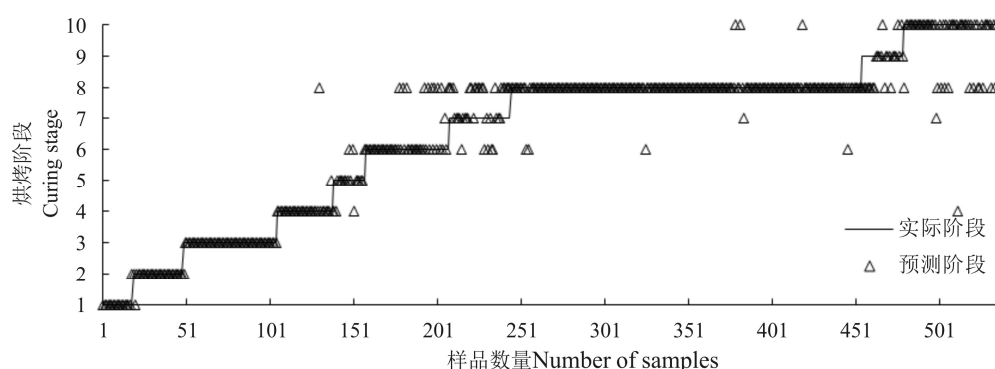


图 11 ELM 测试集的实际分类和预测图

Fig.11 Actual classification and prediction map of ELM test set

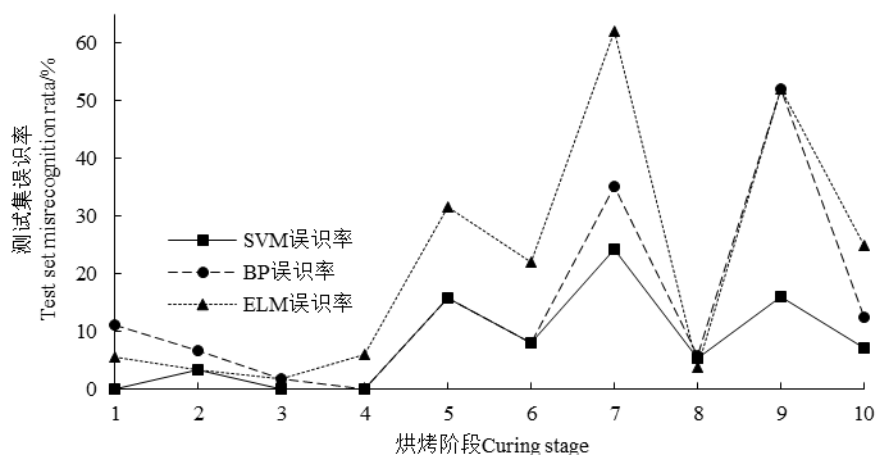
2.5 以优选后特征作为模型输入时不同算法的烘烤阶段识别结果对比

以选用优选后 4 个图像特征作模型输入，不同算法的烘烤阶段识别情况对比如表 5 和图 12 所示，3 种分类模型前 4 阶段的识别错误率较低，均在 10% 以下；第 5 阶段和第 6 阶段次之，第 7 阶段、第 8 阶段、第 9 阶段和第 10 阶段误识率较高。这主要是由于本研究

设计的前 4 个阶段为烟叶烘烤的变黄期，第 5~6 阶段为烟叶烘烤的主要定色时期，第 7~10 阶段为烟叶烘烤的定色后期和干筋期。在烘烤过程中，烟叶外观特征的显著变化都集中在变黄期和定色前期^[36-38]，识别出的图像的颜色纹理特征参数的大幅变化也主要反映在变黄期和定色前期，导致定色后期和干筋期的模型误识率相对偏高。

表 5 3 种模型各个烘烤阶段识别结果
Tab.5 Recognition results of the three models at curing stage

判别模型 Discriminant model	误识率 misrecognition rate	烘烤阶段 curing stage										总计 total
		1	2	3	4	5	6	7	8	9	10	
GA-SVM	n/N	0/18	1/30	0/56	0/34	3/19	4/50	9/37	11/210	4/25	4/56	36/535
	P/%	0	3.33	0	0	15.79	8.00	24.32	5.24	16.00	7.14	6.73
PSO-BP	n/N	2/18	2/30	1/56	0/34	3/19	4/50	13/37	12/210	13/25	7/56	57/535
	P/%	11.11	6.67	1.79	0	15.79	8.00	35.14	5.71	52.00	12.5	10.65
ELM	n/N	1/18	1/30	1/56	2/34	6/19	11/50	23/37	8/210	13/25	14/56	80/535
	P/%	5.56	3.33	1.79	5.88	31.58	22.00	62.16	3.81	52.00	25.00	14.95



注: (1) n: 对应阶段识别错误的样本个数 (个); N: 对应阶段样本总数 (个); P: 误识率。(2) 误识率=阶段识别错误的样本个数/对应阶段样本总数*100%。

Note: (1) n: the number of samples with identification errors in the corresponding stage (pieces); N: the total number of samples at the corresponding stage (pieces); P: the misrecognition rate. (2) Misrecognition rate = number of samples with stage recognition errors/total number of samples in the corresponding stage*100%.

图 12 测试集各模型烘烤阶段误识率对比图

Fig.12 Comparison of the misrecognition rate of each model in the test set during the curing stage

3 讨论

目前密集烘烤过程烟叶的变黄与干燥状态依旧需要依靠人眼和手等感官为基础进行人为判断, 存在较强的主观性和盲目性^[39]。烘烤过程中烟叶颜色、形态特征的变化是指导烘烤操作的重要依据, 图像处理技术可以将烟叶图像的颜色形态等外观特征进行量化^[40-42]。本研究利用广角耐高温高清摄像头实时采集烘烤过程中烟叶变化图像, 通过图像识别技术提取整夹烟叶图像的 11 个颜色特征和 8 个纹理特征, 分别选用全部 19 个图像特征、与烘烤阶段相关性系数的绝对值大于 0.7 的 10 个图像特征和经过聚类分析和相关性分析优选出的 4 个颜色纹理特征作为 GA-SVM、PSO-BP 神经网络和 ELM 分类模型的输入, 开展烟叶烘烤阶段的模型预测, 结果表明, GA-SVM 的烘烤阶段识别准确率均达到 90% 以上, PSO-BP 神经网络的判别准确率在 90% 左右, ELM 模型判别准确率在 85% 以上。为保证模型训练的快速性、准确性以及泛化性, 选用优选后 2 个颜色特征和 2 个纹理特征作为预测模型输入, 所建立的 3 种判别模型测试集预测精确度分别为 93.27%、89.35% 和 85.05%。3 种分类模型均表现为前 4 阶段的识别错误率较低, 第 5 阶段和第 6 阶段次之, 第 7 阶段到第 10 阶段误识率较高, 其中前 4 阶段为烟叶烘烤的主要变黄期, 关系到烟叶颜色的变化情况,

直接影响烤烟质量, 是智能烘烤自动控制的关键阶段; 第 5~6 阶段为烟叶烘烤的主要定色时期, 起到固定烟叶颜色的作用, 对烤烟质量影响较大; 第 7~10 阶段为烟叶烘烤过程中的定色后期和干筋期, 此时烟叶烘烤的主要目的是使烟筋失水干燥, 属于智能烘烤自动控制的次要阶段, 因此虽然第 4 阶段判别效果略差, 但对烤烟质量的影响较小。本研究结果表明, 基于遗传算法的 SVM 模型烘烤阶段识别效果优于基于粒子群算法的 BP 神经网络模型, 基于粒子群算法的 BP 神经网络模型识别效果优于 ELM 模型。原因在于 BP 神经网络和 ELM 这一类神经网络识别方法, 是采用非线性函数形成的超平面在输入空间内进行分割, 而各烘烤阶段颜色纹理特征的空间分布是较为接近的, 尤其是相邻的两个烘烤阶段之间, 因此难以达到十分精确的识别, 这也是误识别大概率出现在相邻烘烤阶段的原因。而 SVM 利用核函数代替高维空间的映射, 将原始数据映射到一个更高维的数据空间, 使得原本线性不可分的问题转化为高维空间中线性可分的样本, 降低了构造最优分类超平面的难度, 可以显著提高识别精度^[43-44]。

4 结论

本研究表明图像处理技术能有效地量化烘烤过程中烟叶颜色及表面皱缩、卷曲、光滑程度等物理形态

特征的变化,对于判断烘烤过程、指导烘烤操作具有实际意义。所建立的3类烘烤阶段预测模型表现为基于遗传算法的SVM模型烘烤阶段识别效果优于基于粒子群算法的BP神经网络模型,基于粒子群算法的BP神经网络模型识别效果优于ELM模型。各个模型的建立情况均表现为前6个阶段识别精度较好,第7到第10阶段的识别精度相对较差,这为后续研究提供了方向,可以单独寻找干筋期特征,建立独立的分类判别模型。

参考文献

- [1] 段史江, 朱红根, 傅宗仁, 等. 烤烟密集烘烤智能控制技术进展[J]. 江西农业学报, 2013, 25(2): 107-109.
DUAN Shijiang, ZHU Honggen, FU Zongren, et al. Research Progress in Intelligent Control Technology of Tobacco Bulk Curing [J]. Acta Agriculturae Jiangxi, 2013, 25(2): 107-109.
- [2] 贺帆, 王涛, 武圣江, 等. 密集烘烤烤烟不同品种烟叶质地和颜色变化[J]. 核农学报, 2014, 28(9): 1647-1655.
HE Fan, WANG Tao, WU Shengjiang, et al. Study on the Change of Texture and Color of Different Flue-cured Tobacco Varieties during the Bulk-curing [J]. Journal of Nuclear Agricultural Sciences, 2014, 28(9): 1647-1655.
- [3] 李东洋, 谢琳. 基于计算机视觉的农业自动化技术研究[J]. 农村实用技术, 2020, (12): 19-20.
LI Dongyang, XIE Lin. Research on Agricultural Automation Technology Based on Computer Vision [J]. Rural Practical Technology, 2020, (12): 19-20.
- [4] 张红涛, 朱洋, 谭联, 等. 基于FA-SVM技术的烟草早期病害识别[J]. 河南农业科学, 2020, 49(8): 156-161.
ZHANG Hongtao, ZHU Yang, TAN Lian, et al. The Recognition of Early Tobacco Disease Based on FA-SVM Technology [J]. Henan Agricultural Sciences, 2020, 49(8): 156-161.
- [5] 朱洋, 许帅涛, 李艺嘉, 等. 基于机器学习的烟草蛙眼病赤星病识别[J]. 粮食科技与经济, 2019, 44(5): 114-116.
ZHU Yang, XU Shuitao, LI Yijia, et al. Recognition of Tobacco Frog Eye Alternaria Based on Machine Learning [J]. Technology and Equipment, 2019, 44(5): 114-116.
- [6] 马俊红, 刘冬梅, 李永亮, 等. 烟草病虫药害智能识别基准数据集构建及三维注意力模型设计[J]. 中国烟草学报, 2021, 27(05): 52-60.
MA Junhong, LIU Dongmei, LI Yongliang, et al. Construction of benchmark data set for intelligent identification of tobacco pests, diseases, phytotoxicity and design of three dimensional attention model [J]. Acta Tabacaria Sinica, 2021, 27(05): 52-60.
- [7] 张文静, 孙秀朋, 乔永亮, 等. 基于InceptionV3的烟草病害识别[J]. 中国烟草学报, 2021, 27(05): 61-70.
ZHANG Wenjing, SUN Xiupeng, QIAO Yongliang, et al. Tobacco disease identification based on InceptionV3 [J]. Acta Tabacaria Sinica, 2021, 27(05): 61-70.
- [8] 谢滨瑶, 祝诗平, 黄华. 基于BPNN和SVM的烟叶成熟度鉴别模型[J]. 中国烟草学报, 2019, 25(1): 45-50.
XIE Binyao, ZHU Shiping, HUANG Hua. Model for identification of tobacco leaf maturity based on BPNN and SVM [J]. Acta Tabacaria Sinica, 2019, 25(1): 45-50.
- [9] 谢滨瑶. 基于图像处理的烟叶成熟度鉴别方法[D]. 重庆: 西南大学, 2020.
XIE Binyao. Identification methods of tobacco maturity based on image processing [D]. Chongqing: Southwest University, 2020.
- [10] Qiang W, Si H P, Lei X, et al. Design of a tobacco leaf maturity degree detection device based on image processing technology[C]// International Workshop on Materials. 2016.
- [11] 刘东来. 基于机器视觉的烟叶分级关键技术分析[J]. 农业技术与装备, 2020, (11): 18-19, 22.
LIU Donglai. Analysis of Key Technologies of Tobacco Leaf Classification Based on Machine Vision [J]. Agricultural Technology & Equipment, 2020, (11): 18-19, 22.
- [12] 邓晨曦. 基于机器视觉的烟叶自动分级方法研究[J]. 科技创新导报, 2020, 17(12): 39-40.
DENG Chenxi. Research on automatic classification of tobacco leaves based on machine vision [J]. Science and Technology Innovation Herald, 2020, 17(12): 39-40.
- [13] Zhang F, Zhang X. Classification and Quality Evaluation of Tobacco Leaves Based on Image Processing and Fuzzy Comprehensive Evaluation [J]. Sensors, 2011, 11(3): 2369-2384.
- [14] 崔云月, 管一弘, 孙娜, 等. BP神经网络在烟梗长短梗率检测中的应用[J]. 软件导刊, 2021, 20(2): 63-67.
CUI Yunyue, GUAN Yihong, SUN Na, et al. The Application of BP Neural Network in the Determination of the Stalk Length and Stem Rate [J]. Software Guide, 2021, 20(2): 63-67.
- [15] 李阳萱. 烟草异物图像检测技术研究及仿真[D]. 绵阳: 西南科技大学, 2018.
LI Yangxuan. Research and Simulation on Image Recognition of Tobacco Impurities [D]. Mianyang: Southwest University of Science and Technology, 2018.
- [16] 高铁功, 蒋太耕, 甘建军. 烟草异物智能剔除系统中相机优化系统的研究[J]. 科技创新导报, 2019, 16(27): 83, 85.
GAO Tiegong, JIANG Taigeng, GAN Jianjun. Research on the camera optimization system in the intelligent removal system of tobacco foreign matter [J]. Science and Technology Innovation Herald, 2019, 16(27): 83, 85.
- [17] 段史江, 宋朝鹏, 马力, 等. 基于图像处理的烘烤过程中烟叶含水量检测[J]. 西北农林科技大学学报(自然科学版), 2012, 40(5): 74-80.
DUAN Shijiang, SONG Zhaopeng, MA Li, et al. Prediction of tobacco leaf's water contents during bulk curing process based on image processing technique [J]. Journal of Northwest A & F University (Natural Science Edition), 2012, 40(5): 74-80.
- [18] 张丽英, 鲜兴明, 杨杰, 等. 烘烤过程中烟叶颜色特征参数与色素含量的关系[J]. 烟草科技, 2013, (08): 85-90.
ZHANG Liying, XIAN Xingming, YANG Jie, et al. Relationship Between Pigment Concentration and Color Attributive Parameters of Tobacco Leaves During Flue-curing [J]. Tobacco Science & Technology, 2013, (08): 85-90.
- [19] 郭朵朵. 基于机器视觉的烟草烘烤品质检测及控制系统的设计[D]. 杨凌: 西北农林科技大学, 2017.
GUO Duoduo. Detection of tobacco baking quality and design of control system based on machine vision [D]. Yangling: Northwest A & F University, 2017.
- [20] 张淑雅, 赵一鸣, 李均利. 基于SVM的图像分类算法与实现[J]. 计算机工程与应用, 2007, 43(25): 40-42, 48.
ZHANG Shuya, ZHAO Yiming, LI Junli. Algorithm and implementation of image classification based on SVM [J]. Computer Engineering and Applications, 2007, 43(25): 40-42, 48.

- [21] 顾金梅, 吴雪梅, 龙曾宇, 等. 基于 BP 神经网络的烟叶颜色自动分级研究[J]. 中国农机化学报, 2016, 37(4): 110-114.
GU Jinmei, WU Xuemei, LONG Zengyu, et al. Automatic classification research on tobacco leaf color based on BP neural network[J]. Journal of Chinese Agricultural Mechanization, 2016, 37(4): 110-114.
- [22] 韩永贵, 高阳, 韩磊, 等. 基于模糊规则的人工神经网络模拟新疆杨蒸腾耗水[J]. 应用生态学报, 2020, 31(05): 1525-1534.
HAN Yonggui, GAO Yang, HAN Lei, et al. Modeling water consumption of Populus bolleana by artificial neural network based on fuzzy rules [J]. Chinese Journal of Applied Ecology, 2020, 31(05): 1525-1534.
- [23] 王锦亚, 李振业, 倪超. 基于机器视觉的实木地板在线分色识别算法[J/OL]. 林业工程学报, 2021, 6(5): 135-139.
WANG Jinya, LI Zhenye, NI Chao. Machine vision algorithms of online color recognition and classification for solid wood flooring products [J]. Journal of Forestry Engineering, 2021, 6(5): 135-139.
- [24] 付云凤. 基于阈值的图像分割研究[D]. 重庆: 重庆大学, 2013.
FU Yunfeng. The Study of Image Segmentation Based on Thresholds [D]. Chongqing: Chongqing University, 2013.
- [25] 谢旭. 基于机器视觉的水稻品种识别系统设计与实现[D]. 哈尔滨: 黑龙江大学, 2019.
XIE Xu. Design and implementation of rice variety recognition system based on machine vision[D]. Harbin: Heilongjiang University, 2019.
- [26] 李波, 张仲文, 章程, 等. 浅谈不同颜色模型在烟叶颜色数字化中的运用[J]. 天津农业科学, 2021, 27(07): 48-51.
LI Bo, ZHANG Zhongwen, ZHANG Cheng, et al. Application of Different Color Models in Tobacco Color Digitization[J]. Tianjin Agricultural Sciences, 2021, 27(07): 48-51.
- [27] 陈琪. 温室植物病害监测和识别系统研究[D]. 天津: 天津科技大学, 2019.
CHEN Qi. Research on Diseases Monitoring And Recognition System of Greenhouse Plant[D]. Tianjin: Tianjin University of Science & Technology, 2019.
- [28] 孙福山, 王松峰, 王爱华, 等. 五段五对应烟叶烤香密集烘烤精准工艺: CN102599624A[P]. 2012-07-25.
SUN Fushan, WANG Songfeng, WANG Aihua, et al. Five segments of five correspond to the precise process of intensive roasting of tobacco leaves: CN102599624A[P]. 2012-07-25.
- [29] 徐秀红, 王传义, 刘昌宝, 等. “8 点式精准密集烘烤工艺”的创新集成与应用[J]. 中国烟草科学, 2012, 33(5): 68-73.
XU Xiuhong, WANG Chuanyi, LIU Changbao, et al. Foundation and Application of Eight-point Curing Technology [J]. Chinese Tobacco Science, 2012, 33(5): 68-73.
- [30] 杨倩, 高晓阳, 武季玲, 等. 基于颜色和纹理特征的大麦主要病害识别研究[J]. 中国农业大学学报, 2013, 18(5): 129-135.
YANG Qian, GAO Xiaoyang, WU Jiling, et al. Research on the identification of major barley diseases based on color and texture features[J]. Journal of China Agricultural University, 2013, 18(5): 129-135.
- [31] 汪世义, 陶亮, 王华彬. 支持向量机和遗传算法的人脸识别方法[J]. 计算机工程与应用, 2009, 45(12): 164-166, 232.
WANG Shiyi, TAO Liang, WANG Huabin. Face recognition based on Support Vector Machine and Genetic Algorithm. Computer Engineering and Applications[J], 2009, 45(12): 164-166, 232.
- [32] 朱超岩, 姚晓东. 基于遗传算法优化的支持向量机在变压器故障诊断中应用[J]. 仪表技术, 2019, (3): 21-23, 32.
ZHU Chaoyan, YAO Xiaodong. Application of the Support Vector Machine in the Fault Diagnosis on the Transformer Based on the Optimal Genetic Algorithm [J]. Instrumentation Technology, 2019, (3): 21-23, 32.
- [33] 马创, 王尧, 李林峰. 基于遗传算法与支持向量机的水质预测模型[J]. 重庆大学学报, 2021, 44(07): 108-114.
MA Chuang, WANG Yao, LI Linfeng. A water quality prediction model based on genetic algorithm and SVM [J]. Journal of Chongqing University, 2021, 44(07): 108-114.
- [34] 杨旭, 纪玉波, 田雪. 基于遗传算法的 SVM 参数选取[J]. 辽宁石油化工大学学报, 2004, 24(1): 54-58.
YANG Xu, JI Yubo, TIAN Xue. Parameters Selection of SVM Based on Genetic Algorithm [J]. Journal of Liaoning Petrochemical University, 2004, 24(1): 54-58.
- [35] 孟令峰, 朱荣光, 白宗秀, 等. 基于手机图像的不同贮藏时间下冷却羊肉的部位判别[J]. 食品科学, 2020, 41(23): 21-26.
MENG Lingfeng, ZHU Rongguang, BAI Zongxiu, et al. Discrimination of Chilled Lamb from Different Carcass Parts at Different Storage Times Based on Mobile Phone Images [J]. Food Science, 2020, 41(23): 21-26.
- [36] 王爱华, 王松峰, 孙福山, 等. 变黄期阶梯升温烘烤工艺对多酚类及相关物质的影响[J]. 中国烟草科学, 2016, 37(2): 59-64.
WANG Aihua, WANG Songfeng, SUN Fushan, et al. Effects of the Step-temperature Yellowing Curing Technology on Polyphenol and Related Substance in Tobacco Leaves [J]. Chinese Tobacco Science, 2016, 37(2): 59-64.
- [37] 赵应虎, 王涛, 何艳辉, 等. 烘烤过程中烤烟外观与内在质量变化研究进展[J]. 作物研究, 2013, 27(6): 700-704.
ZHAO Yinghu, WANG Tao, HE Yanhui, et al. Research Advance in Changes of Appearance Characteristics and Internal Quality of Flue-cured Tobacco during Curing [J]. Crop Research, 2013, 27(6): 700-704.
- [38] 詹军, 周芳芳, 张晓龙, 等. 密集烤房与普通烤房烘烤过程中环境和烟叶颜色的变化[J]. 河南农业大学学报, 2015, 49(1): 17-21.
ZHAN Jun, ZHOU Fangfang, ZHANG Xiaolong, et al. Changes of curing environment and color of tobacco leaves in bulk curing barn and traditional flue-curing barn during curing process [J]. Journal of Henan Agricultural University, 2015, 49(1): 17-21.
- [39] 段史江. 图像处理技术在烤烟烘烤过程中的应用研究[D]. 郑州: 河南农业大学, 2012.
DUAN Shijiang. Research on Application of Image Processing Technology during Curing Process of Flue-Cured Tobacco [D]. Zhengzhou: Henan Agricultural University, 2012.
- [40] 庄珍珍. 基于机器视觉的烟叶自动分级方法研究[D]. 重庆: 西南大学, 2016.
ZHUANG Zhenzhen. Method for Automatic Grading of Tobacco Based on Machine Vision [D]. Chongqing: Southwest University, 2016.
- [41] 庄珍珍, 祝诗平, 孙雪剑, 等. 基于机器视觉的烟叶自动分组方法[J]. 西南师范大学学报(自然科学版), 2016, 41(4): 122-129.
ZHUANG Zhenzhen, ZHU Shiping, SUN Xuejian, et al. On Method of Tobacco Automatic Grouping Based on Machine Vision [J]. Journal of Southwest Normal University (Natural Science Edition), 2016, 41(4): 122-129.
- [42] Yin Y, Xiao Y, Yu H. An image selection method for tobacco leave grading based on image information[J]. Engineering in Agriculture Environment & Food, 2015, 8(3): 148-154.
- [43] 宋彦, 谢汉垒, 宁井铭, 等. 基于机器视觉形状特征参数的祁

门红茶等级识别[J]. 农业工程学报, 2018, 34(23): 279-286.
SONG Yan, XIE Hanlei, NING Jingming, et al. Grading Keemun black tea based on shape feature parameters of machine vision [J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(23): 279-286.

[44] 胡勇. 基于图像处理的试纸识别技术研究[D]. 厦门: 厦门理工学院, 2018.
HU Yong. Research on Paper Testing Technology Based on Image Processing [D]. Xiamen: Xiamen University of Technology, 2018.

Selection of optimum discriminant model in tobacco curing stage based on image processing

LI Zengsheng^{1,2}, MENG Lingfeng¹, WANG Songfeng^{1*}, GAO Jun³, XU Xiaohong⁴, ZHU Xianzhou³, YANG Chao⁴,
WANG Bojun⁴, WANG Aihua¹, MENG Lin¹, LIU Zichang¹, DU Haina^{1,2}, LIU Hao^{1,2}, SUN Fushan¹

1 Institute of Tobacco Research of CAAS, Key Laboratory of Tobacco Biology and Processing, Ministry of Agriculture, Qingdao 266101, China;

2 Graduate School of Chinese Academy of Agricultural Sciences, Beijing 100081, China;

3 Liangshan Tobacco Company of Sichuan Province, Xichang, Sichuan 615000, China;

4 Chongqing Tobacco Science Research Institute, Chongqing 400715, China

Abstract: [Background] The automatic discrimination in the tobacco curing stage is an important link for the establishment of an intelligent tobacco leaf curing system. This study aims to realize the accurate identification and control in the curing stage and improve the accuracy of tobacco curing. [Methods] In this study, 11 color features and 8 texture features of the complete tobacco leaf images during the curing process were extracted, then variable cluster analysis of the color features and texture features was carried out, and finally the extracted color feature values and texture features were divided into two categories with a distance of 10. Correlation analysis was performed to filter out one feature the strongest correlation with the curing stage from each categorize of feature to form a feature subset (R/G, I*, gray average and inertia), which was used as model input. Then support vector machine based on genetic algorithm (GA-SVM), the particle swarm algorithm back propagation (PSO-BP) neural network and the extreme learning machine (ELM) were used for classification and recognition in the tobacco curing stage. [Results] By using the selected four image features as model inputs, the test set discrimination accuracy rates of the established GA-SVM model, the PSO-BP neural network model and the ELM model were 93.27%, 89.35%, and 85.05%, respectively. [Conclusion] The SVM model based on the genetic algorithm in the curing stage has the best recognition, followed by the BP neural network model based on the particle swarm algorithm, and the ELM model ranking last.

Keywords: flue-cured tobacco; curing stage; image processing; feature model; smart curing

*Corresponding author. Email: wangsongfeng@caas.cn

《中国烟草科学》2022年第1期目次

· 遗传育种

烤烟新品种贵烟5号的选育及特征特性
..... 王仁刚, 王自力, 林世锋, 等

· 生物技术

双向等位基因特异性PCR技术在烟草SNP分型中的应用
..... 林世锋, 王仁刚, 潘飞, 等
烟草OSCA基因家族鉴定及非生物胁迫诱导表达模式分析
..... 邓智超, 田冬冬, 宋青松, 等

· 栽培营养

不同火土灰用量对烟苗生长发育的影响
..... 肖艳松, 杨如意, 廖雅桦, 等
深耕配施有机肥对云南保山植烟土壤肥力的快速提升效应
..... 何京, 王津军, 尹兴盛, 等

· 生理生态

镉胁迫对不同植烟土壤硝化速率及N₂O排放的影响
..... 许琳, 刘跃东, 高加明, 等
烟草废弃物生物炭元素组成及其重金属安全性研究
..... 杨继鑫, 张久权, 郜军艺, 等

· 植物保护

河南烟草主要病毒发生种类及侵染类型分析
..... 白静科, 何雷, 吴彦辉, 等
杀菌剂与无致病力烟草青枯菌对噬菌体存活的影响
..... 崔晓东, 王伊玲, 李红霞, 等
抗烟草疫霉活性木霉与芽孢杆菌共培养体系的构建与优化
..... 张希芬, 林伟, 李清钰, 等
多菌灵降解菌的构建及对烟叶残留降解效果评价
..... 陈芊如, 余佳敏, 胡希好, 等

· 品质化学

沂蒙丘陵生态区蜜甜焦香烤烟外观特征及与烟叶品质的关系
..... 左伟标, 蔡宪杰, 王皓, 等
烤烟柔韧性与内在品质的相关分析
..... 张书伟, 过伟民, 陈祥, 等
酶解预处理工艺在烟叶茄尼醇提取中的应用
..... 黄振瑞, 柳颖, 黎志德

· 现代烟草农业

基于Sentinel-2遥感影像的烟草种植信息精准提取
..... 薛宇飞, 张军, 张萍, 等