

引用格式:

陈婷, 赵晓琳, 张冀武, 盖小雷, 张晓伟, 刘宇晨, 王燕, 龙杰. 基于GA-RELM多特征优选的烟叶多部位正反面识别方法[J]. 湖南农业大学学报(自然科学版), 2025, 51(1): 113–122.

CHEN T, ZHAO X L, ZHANG J W, GAI X L, ZHANG X W, LIU Y C, WANG Y, LONG J. Multi-site front and back recognition method of tobacco leaf based on multi-feature optimization using GA-RELM[J]. Journal of Hunan Agricultural University(Natural Sciences), 2025, 51(1): 113–122.

投稿网址: <http://xb.hunau.edu.cn>



## 基于 GA-RELM 多特征优选的烟叶多部位正反面识别方法

陈婷<sup>1</sup>, 赵晓琳<sup>1</sup>, 张冀武<sup>2</sup>, 盖小雷<sup>3</sup>, 张晓伟<sup>2</sup>, 刘宇晨<sup>2\*</sup>, 王燕<sup>3</sup>, 龙杰<sup>2</sup>

(1.昆明理工大学机电工程学院, 云南 昆明 650500; 2.云南省烟草质量监督检测站, 云南 昆明 650106; 3.云南省烟草烟叶公司, 云南 昆明 650217)

**摘要:** 针对现有烟叶分级模型多基于平整烟叶的正面特征构建, 分级模型准确率和实用性较低的问题, 提出一种基于遗传算法-正则化极限学习机(GA-RELM)多特征优选的烟叶多部位正反面识别方法。首先, 对自然状态下的烟叶进行多尺度正反面特征提取, 构建正反两面数据集, 根据特征重要性和特征间的潜在关系, 实现特征降维并构建新特征组合。其次, 对正则化极限学习机(RELM)进行隐藏层偏置寻优, 以提高模型实际应用性和分类精度。结果表明: 与原极限学习机(ELM)相比, GA-RELM对自然状态下的烟叶正反面和多部位正反面的分类精度分别提高了0.84%和7.88%, 运算时间分别减少2.56 s和5.72 s; 与其他烟叶分级算法相比, GA-RELM在准确率、精确率、召回率、F1评分等多个指标上表现出明显优势。

**关键词:** 烤烟; 烟叶分级; 多特征优选; 遗传算法; 正则化极限学习机

中图分类号: TS42

文献标志码: A

文章编号: 1007-1032(2025)01-0113-10

## Multi-site front and back recognition method of tobacco leaf based on multi-feature optimization using GA-RELM

CHEN Ting<sup>1</sup>, ZHAO Xiaolin<sup>1</sup>, ZHANG Jiwu<sup>2</sup>, GAI Xiaolei<sup>3</sup>,  
ZHANG Xiaowei<sup>2</sup>, LIU Yuchen<sup>2\*</sup>, WANG Yan<sup>3</sup>, LONG Jie<sup>2</sup>

(1.Factory of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 2.Yunnan Tobacco Quality Supervision and Testing Station, Kunming, Yunnan 650106, China; 3.Yunnan Tobacco Company, Kunming, Yunnan 650217, China)

**Abstract:** To address the issues of low accuracy and practicality in existing tobacco leaf grading models, which are mostly built upon the front-side features of flattened tobacco leaves, a multi-site front and back recognition method was proposed using a genetic algorithm-regularized extreme learning machine(GA-RELM). Firstly, multi-scale features from both side of tobacco leaves in their natural state were extracted to construct dataset. Feature importance and relationships were analyzed to reduce dimensionality and construct optimized feature combinations. Secondly, the hidden layer biases of the regularized extreme learning machine(RELM) were optimized to enhance model accuracy and applicability. The results showed that compared with the original extreme learning machine(ELM), GA-RELM improved classification accuracy by 0.84% and 7.88% for front/back leaves and other parts, respectively, while reducing the computation time by 2.56 s and 5.72 s, respectively. GA-RELM outperformed other grading algorithms in accuracy, precision, recall and F1-score.

收稿日期: 2023-11-26

修回日期: 2023-12-11

基金项目: 云南省重大科技专项计划(202002AC080001); 中国烟草总公司云南省烟草公司科技计划重点项目(2020530000241003、2021530000241012)

作者简介: 陈婷(1971—), 女, 云南昆明人, 硕士, 副教授, 主要从事企业集成及信息化工程研究, [kmct@163.com](mailto:kmct@163.com); \*通信作者, 刘宇晨, 博士, 工程师, 主要从事烟叶智能分级研究, [957670272@qq.com](mailto:957670272@qq.com)

**Keywords:** flue-cured tobacco; tobacco leaf grading; multi-feature optimization; genetic algorithm(GA); regularized extreme learning machine(RELM)

烤烟是烟叶制品的原材料,在烟叶加工成成品之前,准确划分烤烟等级是提升烟叶制品质量的关键一步。烤烟分类主要参照国标 GB 2635—1992<sup>[1]</sup>,烟叶的颜色、纹理等外在因素是影响烟草等级的关键。YIN 等<sup>[2]</sup>对被人工展开和压平的烟叶正面进行特征提取,后采用 Fisher 判别分析对 6 个等级的烟叶进行分级,正确率达到 86.5%。刘建新等<sup>[3]</sup>根据烟叶图像对烟叶 3 种常见病斑进行分类,准确率均在 90%以上,但同一病斑在烟叶正反面上的颜色、大小等都存在差异。张永安等<sup>[4]</sup>根据不同产地的烤烟烟叶正反面颜色的色差,分析其对感官质量的影响,发现烟叶等级 C2F 与 C3L 之间的色差显著,而 C2F、C3F、C4F 等级之间色差不显著。李雨芮等<sup>[5]</sup>通过分析红花大金元烟叶正反面色值的分布特征和差异规律,对烤烟进行产地划分,发现昆明产区的烟叶正反面色差最小。这些烤烟等级评定、部位分类、病斑识别、成熟度分级等识别分类任务都需根据叶片正反面特征差异构建正反面模型,因此,对烤烟进行正反面识别分类能够进一步满足烟叶分级、分类任务的需求。

将烟叶按类型分组是烟叶分级的前提,而烟叶部位则是分组的首要考虑因素。要确保烟叶等级评定的准确性,首先必须准确评定烟叶部位<sup>[6]</sup>。徐森等<sup>[7]</sup>提出一种基于改进的 Mask R-CNN 的识别方法,对多片散放烟叶同步进行部位识别,精确率为 97.32%;赵晨等<sup>[8]</sup>分别提取叶片 5 个绝对和相对形态特征后对部位进行分类,准确度为 93.75%。目前,研究者对烤烟部位进行智能化识别模型构建<sup>[9]</sup>时,大多采用经过预处理的平整烟叶正面数据,但实际收购的为烘烤过的叶片,在烘烤和收集中会发生卷曲、翻折等情况,由于模型缺乏自然状态下烟叶正反面识别过程,在实际应用中其准确率会有所降低。因此,在自然状态下提取烟叶正反面的叶面特征,在此基础上建立分级模型,有助于提升模型的实际应用效能。

随着叶片分类中多特征提取技术的发展,如何在众多特征中筛选出对分类有贡献的特征并对其进行融合是目前研究者们所面临的问题。为了提高特征融合能力,HE 等<sup>[10]</sup>提出一种基于偏移窗口自

注意力金字塔特征融合的烟草分类模型 FSWPNet,烟叶平均分类准确率为 75.8%。姚学练等<sup>[11-12]</sup>基于主成分分析对特征进行筛选后,特征维度大幅下降。DASARI 等<sup>[13]</sup>定义了一种基于卷积神经网络(CNN)的烤烟烟叶分级方法,将图像分辨率由 1 450 像素×1 680 像素降维到 256 像素×256 像素后,将烟叶分为 3 类,全局效率达到 85.10%。目前研究多采用传统特征降维方法和传统分类器或神经网络构建烤烟分级模型,其效率和准确度存在一定提升空间。传统降维方法仅根据特征权重分析特征重要性,无法分析特征与分类任务之间的潜在关系,也无法根据多特征之间的联系构建新的特征组合。

近年来,深度学习和机器学习在植物叶片分类上的应用取得了较多进展。LI 等<sup>[14]</sup>采用粒子群优化-极限学习机(PSO-ELM)算法对烟叶进行无损快速分级,发现 PSO-ELM 的准确率比 ELM 的高 3.6%以上。赵晨等<sup>[8]</sup>构建 KNN(K 最近邻)、LR(逻辑回归)、SVM(支持向量机)、BP(神经网络)等烟叶部位分类模型并对分类结果进行对比,发现基于 BP 神经网络模型的识别效果最好,识别准确率为 93.75%。鉴于机器学习算法具有较好的可解释性,能够清晰地解释分类决策的原因,对于数据样本量和计算机配置要求较低,本文作者提出一种基于 GA-RELM 多特征优选的烟叶多部位正反面识别方法,并将其与原 ELM 和其他烟叶分级算法进行对比,以期提高烟叶识别模型的准确率与实用性。

## 1 数据采集

### 1.1 图像采集

采用佳叶烤烟智能分级设备进行烤烟烟叶正反面数据集的采集和制作。以 2022 年在云南省收购的烟叶(主要产自云南昆明、保山、大理等地区,品种为云烟 87)为样本。共采集了包含正面、反面共计 3 786 张烟叶样本图像,由烤烟分级专家根据 GB 2635—1992 标准对烟叶真实等级进行标记,各等级烟叶正面图像数量如表 1 所示。图 1 所示为自然状态下烟叶各部位正反面图像。为方便对比,将烟叶上部正反面分别以 B、Bf 表示,中部正反面以

C、Cf 表示，下部正反面以 X、Xf 表示。为了扩充数据集，对样本图像进行了镜像、旋转和翻转等操作，使得图片样本数量达到了 22 716 张。最终，扩充后的训练集和测试集按 8 : 2 的数量比例进行划分。

表 1 各等级烟叶正面图像数量

级别	数量/张	级别	数量/张	级别	数量/张
B1F	126	B2F	280	B3F	163
C2F	205	C3F	418	C4F	224
X2F	236	X3F	241		

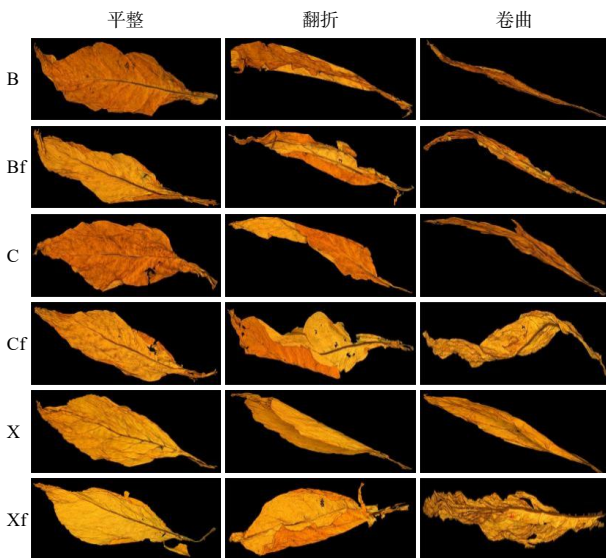


图 1 自然状态下烟叶各部位正反面图像

Fig.1 Positive and negative images of different parts of tobacco leaves in natural state

### 1.2 特征提取

1) 灰度共生矩阵。采用对比度、差异性、同质性、能量、相关性、角二阶矩 6 个特征参数<sup>[15]</sup>，分

别提取角度为 0°、45°、90°、135°，步长为 2、4、8、16 的烟叶纹理特征向量，共 96 维。

2) Gabor 小波。分别以角度 45°、90°，频率 0.6、0.1 以及内核尺寸 3、5 构成 8 种参数组合，并对其数据进行归一化处理<sup>[16]</sup>。

3) 颜色特征。在 HSV 和 RGB 这 2 种颜色空间中分别计算烟叶图像在 3 个低阶颜色矩上的色彩分布情况，所提取的烟叶颜色特征维度共 9 维。

4) 形状特征提取。使用 Hu 不变矩来提取烟叶的形状特征，对烟叶多部位正反面 7 个不变矩求取均值。

## 2 烟叶多部位正反面识别模型构建

### 2.1 特征工程

特征工程是将原始数据转化为更能反映问题本质的特征的过程，通过对特征进行分析和选择来提高模型对不可见数据的预测精度。对于一个选定的学习算法来说，哪一个特征是有效或无效是无法得知的，过多无用的数据不但会造成数据冗余，还会对模型精度造成一定影响。

#### 2.1.1 多特征融合

为了验证多特征提取较于单一特征提取在识别上的差距，展现特征融合和降维的必要性，运用支持向量机分类器(SVM)分别对单一特征和多特征正反面识别以及多部位正反面识别的效果进行对比，结果如表 2 所示。由表 2 可见，单一特征提取的运行时间较短，但精度较低；若采用全部特征提取，则在提高精度的同时也会大幅增加时间成本。

表 2 单一特征和多特征的识别效果对比

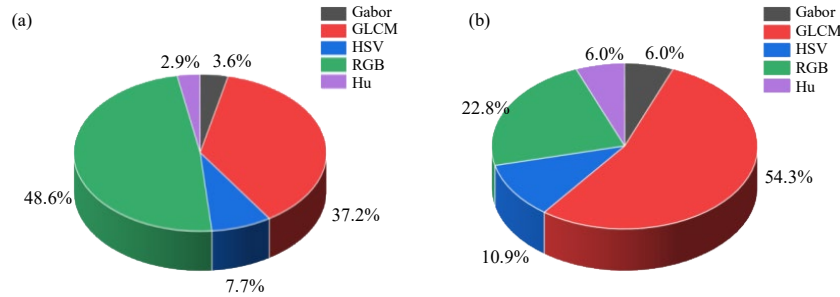
特征	正反面识别		多部位正反面识别		特征维度/维
	准确率/%	时间/s	准确率/%	时间/s	
Gabor	53.82	0.27	21.50	0.34	8
GLCM	75.59	11.20	58.04	22.31	96
Gabor+GLCM	72.55	12.44	57.65	22.07	104
RGB	91.42	0.14	66.75	0.33	9
HSV	89.84	0.22	70.84	0.40	9
RGB+HSV	93.66	0.16	71.76	0.36	18
Hu 矩	58.57	0.26	34.16	0.35	7
Gabor+GLCM+RGB+HSV+Hu 矩	93.93	6.17	76.25	14.21	129

### 2.1.2 特征重要性

如何对大量特征进行选择 and 降维, 找出对模型最有贡献度的特征, 一个重要的评价指标便是特征重要性。递归特征消除法(RFE)通过递归的方式为每

个特征分配一个外部权重估计, 逐步缩小特征集合。

本研究将所提取的 129 个特征作为初始特征集, 运用 RFE 进行建模, 采用随机森林模型计算每个特征的重要性。不同特征的重要性分布如图 2 所示。



(a) 正反面特征的重要性; (b) 多部位正反面特征的重要性。

图 2 不同特征的重要性

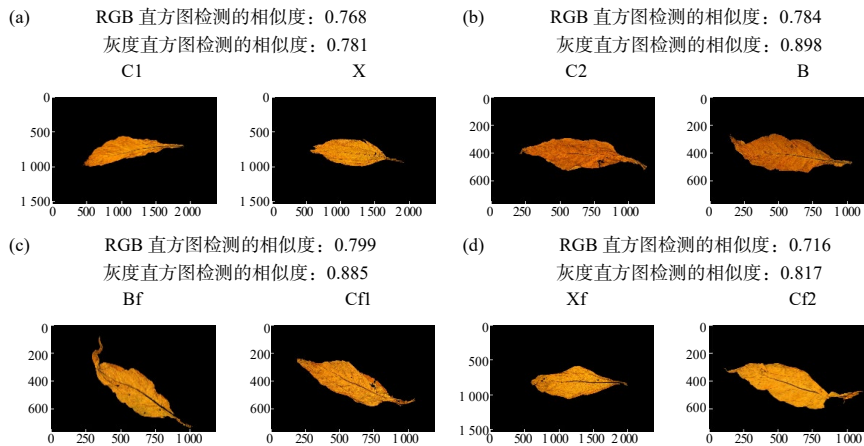
Fig. 2 Importance of different features

由图 2 可见:

1) 在 2 类特征中, 颜色(RGB)特征重要性占比均较大, 特别是在正反面特征中更是如此。这表明颜色在正反面特征和多部位正反面特征中具有较强的鲁棒性, 对于识别烟叶正反面具有更高的贡

献度。

2) 在进行多部位正反面分类时, 纹理特征重要性大于颜色重要性。这是由于不同部位正反面的颜色具有一定相似性, 仅依靠颜色难以对其进行部位划分。不同部位正反面的颜色相似度如图 3 所示。



(a) 正面中部 C1 和下部 X 颜色相似度; (b) 正面中部 C2 和上部 B 颜色相似度; (c) 反面上部 Bf 和中部 Cf1 颜色相似度; (d) 反面下部 Xf 和中部 Cf2 颜色相似度。

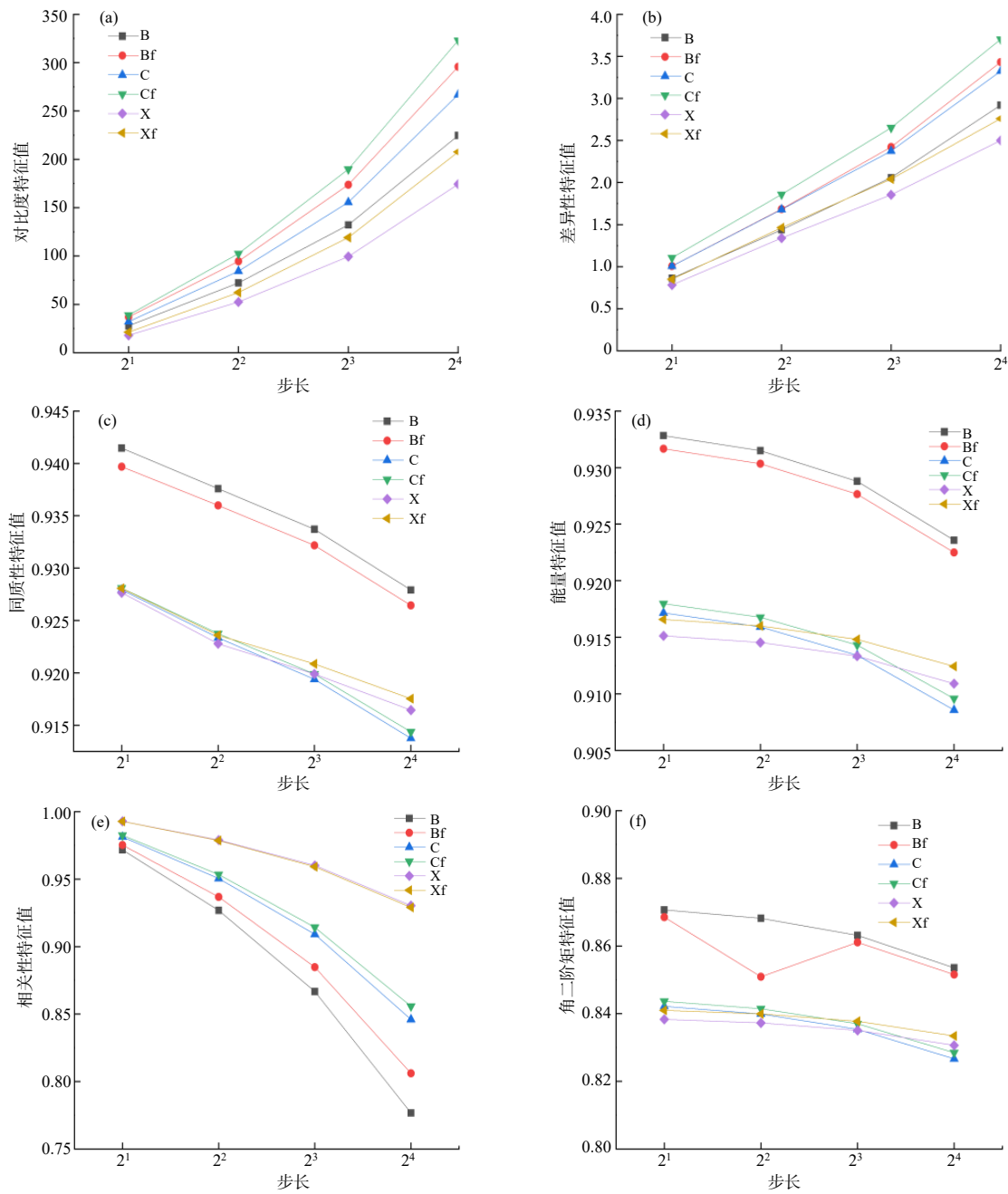
图 3 颜色相似度

Fig.3 Color similarity

### 2.1.3 特征参数优选

灰度共生矩阵和 Gabor 特征维度较高, 部分特征对模型贡献率较低, 若将全部特征都放入模型中, 则会导致数据冗余。因此, 针对不同分类任务, 选择合适的步长、角度及特征点尤为重要。

1) 灰度共生矩阵参数选取。分别对上、中、下部的烟叶正面和反面图像进行特征参数提取, 对所提取到的 3 个部位正反面的特征值求取均值, 结果如图 4 所示。



(a) 对比度特征值; (b) 差异性特征值; (c) 同质性特征值; (d) 能量特征值; (e) 相关性特征值; (f) 角二阶矩特征值。

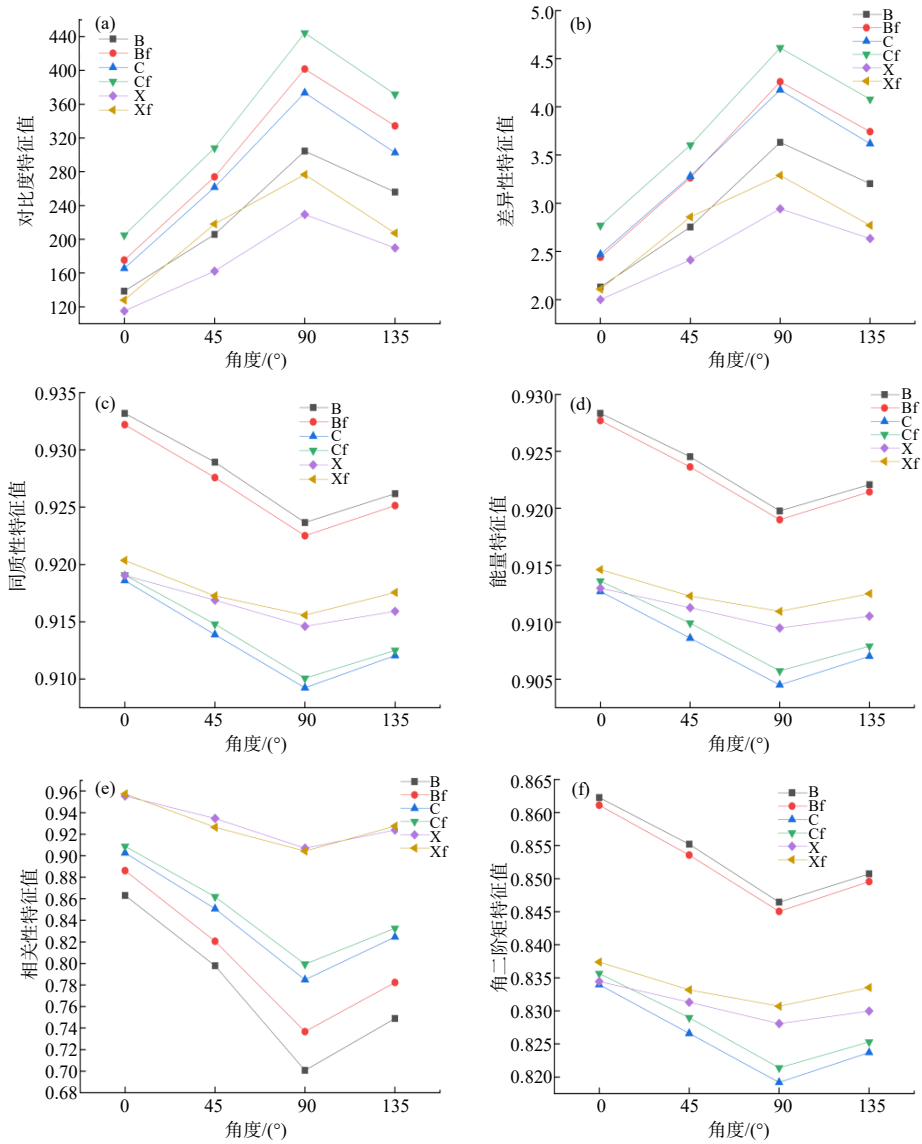
图 4 各特征值随步长的变化

Fig. 4 Changes of characteristic values with step size

由图 4 可知：当步长为 16 时，各特征对于不同类别烟叶的区分效果最明显，步长为 2 时的分类效果较差，步长为 4 和 8 时的特征值在不同类别中均出现不同程度的重叠。由图 4(a)和图 4(b)可见：对比度和差异性这 2 个特征对于烟叶纹理分类识别

的效果较好，说明烟叶各部位纹理图像较清晰。

选定步长为 16，分别对不同角度下烟叶上、中、下部的正面和反面烟叶图像进行特征提取并取均值，结果如图 5 所示。



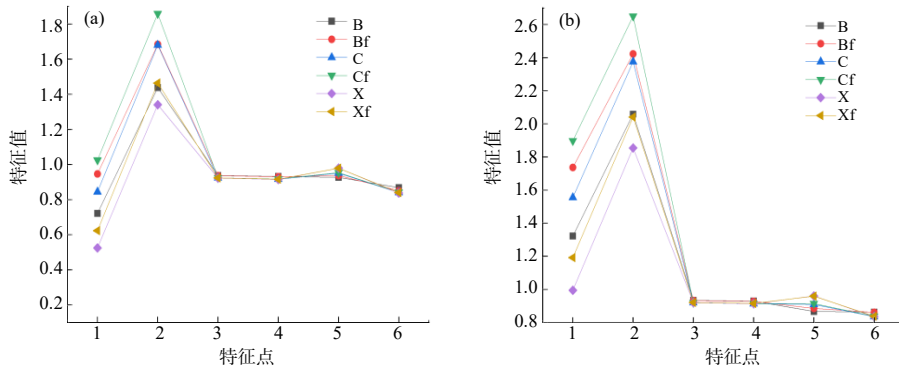
(a) 对比度特征值; (b) 差异性特征值; (c) 同质性特征值; (d) 能量特征值; (e) 相关性特征值; (f) 角二阶矩特征值。

图 5 各特征值随角度的变化

Fig. 5 Changes of characteristic values with angle

由图 5 可知, 当角度为 90°、135°时, 各特征对于烟叶 3 个部位的正反面的区分效果均较为明显。

步长分别为 4、8 时的多部位特征值如图 6 所示, 其中特征点编号 1~6 分别表示对比度、差异性、



(a) 步长为 4; (b) 步长为 8。

图 6 步长为 4 和 8 时的多部位特征值

Fig. 6 Multi-part eigenvalues with steps of 4 and 8

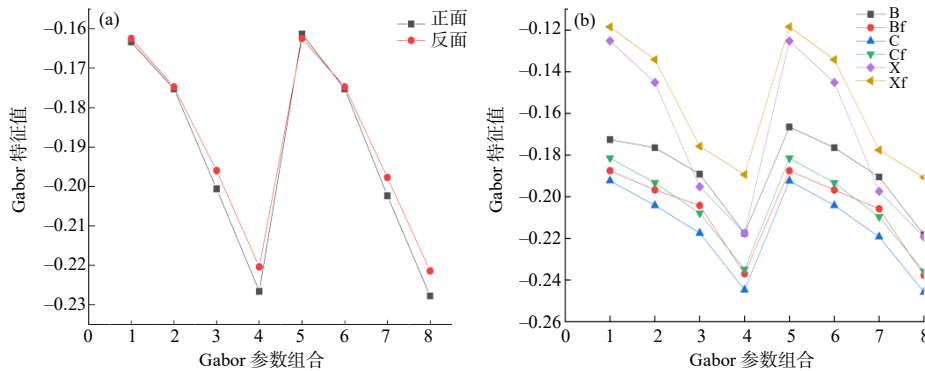


同质性、能量、相关性、角二阶矩。由图 6 可见特征点 1(对比度)的识别分类性能较好,所以对其保留特征点 1 的特征参数。综合上述试验分析结果,将特征维度由 96 维降到 20 维。

2) Gabor 参数特征选取。Gabor 参数组合如表 3 所示。进一步分析正反面和多部位正反面识别所采用的 8 种参数组合的特征值,结果如图 7 所示。

表 3 特征参数组合

Table 3 Characteristic parameter combination			
Gabor 参数组合	频率	角度/(°)	内核大小
1	0.6	45	5
2	0.6	90	5
3	0.1	45	3
4	0.1	90	3
5	0.6	45	3
6	0.6	90	3
7	0.1	45	5
8	0.1	90	5



(a) 正反面特征值随参数组合的变化; (b) 多部位正反面特征值随参数组合的变化。

图 7 正反面特征值随参数组合的变化

Fig.7 Changes of positive and negative characteristic values with parameter combinations

由图 7(a)可以看出,仅区分烟叶正反面时各参数组合均可实现区分效果。由图 7(b)可见:当特征参数组合为 4 和 8 时,折线有较多重叠,难以对 6 类烟叶进行区分。以上结果表明,在对于多部位烟叶进行纹理提取时,角度为 90°、频率为 0.1 的参数组合对多部位烟叶正反面的分类贡献度较差,据此,将特征维度由 8 维降到 6 维。

2.1.4 递归特征消除交叉验证法 (RFE) 特征优选

完成特征参数优选后,特征维度由 129 维降到 51 维,为保证特征重要性的合理性和准确性,加入 CV(交叉验证)函数对 RFE 的特征重要性进行交叉验证,从而找到最优的特征数量。由于本试验所提取的原始特征数量较多,因此,将步长设置为 0.1,每次排除 10%的特征;折数设置为 10 折,每次测试折数选择为 9 折。通过试验可以得出正反面最优的特征组合为 9 个,多部位正反面最优的特征组合为 25 个。正反面和多部位正反面优选特征的重要性分别如图 8 和图 9 所示。

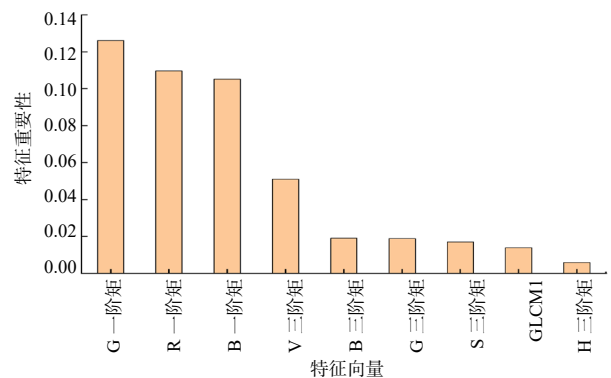


图 8 正反面优选特征的重要性

Fig.8 Importance of positive and negative optimization features

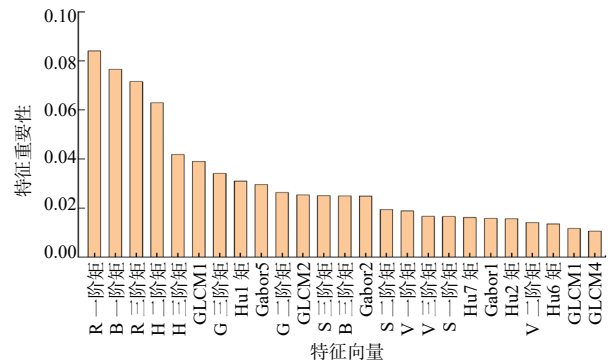


图 9 多部位正反面优选特征的重要性

Fig.9 Importance of multi-site positive and negative optimization features

对比图 8 图 9 可知, 烟叶正反面之间的色差较大, 颜色作为最直观的特征, 在正反面识别上有较好的效果, 但进一步区分部位时, 需加入其他特征。在对于颜色特征的选取上, 图 8 中的 RGB 特征重要性排名靠前, 图 9 中 HSV 的颜色通道被选取的概率较大, 说明在多部位正反面间 RGB 相似度较高的情况下, HSV 更适合对多部位正反面进行识别。

### 2.1.5 特征选择合理性验证

为验证优选特征的正确性与合理性, 将正反面最优特征组合命名为“正反面 9”, 多部位正反面最优特征组合命名为“多部位 25”, 随机抽取的 19 个特征的组合命名为“随机 19”, 使用 SVM 分类器对特征进行识别分类, 结果如表 4 所示。由表 4 可见: 合适的特征组合可以提升分类准确率, 组合中过多的特征反而会对分类准确率造成一定影响。

表 4 特征组合识别效果

特征组合	正反面识别		多部位正反面识别		特征 维度/维
	准确率/%	时间/s	准确率/%	时间/s	
正反面 9	93.79	0.55	68.07	1.17	9
多部位 25	93.21	0.27	74.80	0.50	25
随机 19	67.01	0.43	45.91	0.46	19

## 2.2 GA-RELM模型构建

RELM是极限学习机的扩展, 通过添加正则化项可以有效地控制模型的复杂度, 防止过拟合。与传统的BP神经网络算法相比, ELM具有学习速度快、泛化性能好的优点, 但同样容易陷入局部最优的问题。本研究通过引入遗传算法(GA)来搜索RELM模型的最佳权重和偏置值, 以最小化输出误差。GA是一种模拟自然选择和遗传机制的优化算法, 它通过不断进化产生更好的解决方案, 该算法对经过RFECV特征优选后的数据进行交叉、变异和可迭代收敛等操作, 找寻数据集与模型之间最匹配的权重和偏置值。GA-RELM的集成算法流程如表5所示。

表 5 GA-RELM 集成算法

Table 5 GA-RELM integration algorithm

算法 1 GA-RELM
输入: 数据集
输出: best_hidden_weight, best_b
步骤 1 构建神经网络
定义激活函数 $f_c(x, p=5)$ , 其中 $x$ 为输入, 用于计算激活函数值
定义损失函数 $f_l(W, C=0.1)$ , 用于计算目标函数值以及对应的隐藏层偏执向量 $b$
定义预测函数 $f_p(W, b, C=0.1)$ , 根据输入的权重矩阵 $W$ 和偏执向量 $b$ , 计算预测值
步骤 2 构建遗传算法
定义初始化, 将传入的参数赋值给类的属性, 以便在后续的遗传算法中使用
定义将种群转换为神经网络的矩阵的函数
population_2D_decimal=np.zeros((self.hidden,self.weight)) // 输出隐藏层*权重分布的矩阵
计算种群的适应能力
返回一维数组
在每个种群中迭代:
rat_fitness[i]=abs(fitness[i])/sum_fitness // 计算种群适应度权重
population=new_population // 完成自然选择和淘汰
if (random.random(<self.pc): then
继续 // 迭代每个隐藏层和权重层
结束 // 完成染色体交换
if (random.random(<self.pm): // 随机生成一个数和变异率比较
population[i][mpoint]==1:
population[i][mpoint]=0 // 进行 1 到 0 的变换或 0 到 1 的变换
结束 // 实现基因变异
定义 best_fitness 进行对比
if fitness[i]<best_fitness: // 如果适应度大于 0 且比当前最好的适应度还要好
return (best_individual, best_fitness) // 则更新最优秀的基因
调用主函数
输出 best_hidden_weight, best_b

“//”后的文字为算法注释。



算法设定保留6个最优隐藏层偏置，为保留 RELM的泛化能力，根据正反面和多部位正反面进行隐藏层偏置范围选取，选取范围分别为-0.3~0.6，-0.4~0.4。

### 3 模型验证与分析

通过对比试验验证 RFECV-GA-RELM 算法集成对于叶片正反面识别的效果，所用机器学习方法均采用正反面 9 个特征、多部位正反面 25 个特征的数据集，训练集和测试集的数据划分比例为 8 : 2。

#### 3.1 模型改进试验

通过试验对比原 ELM、RELM 以及改进的 GA-RELM 对烟叶正反面和多部位正反面分类识别的效果，其中，原 ELM 和 RELM 都未进行隐藏层偏置寻优，本研究提出的 GA-RELM 加入寻优算法，采用相同数据集划分，结果如表 6 所示。

表 6 模型改进前后对烤烟正反面识别的效果对比

Table 6 Comparison of recognition effects of front and back of flue-cured tobacco before and after model improvement

模型	正反面识别		多部位正反面识别	
	准确率/%	时间/s	准确率/%	时间/s
ELM	97.729 3	6.64	84.782 1	156.61
RELM	97.490 3	4.79	86.163 0	155.85
GA-RELM	98.547 8	3.03	91.459 9	131.76

由表 6 可见：RELM 相较于原 ELM 在准确率和运行效率上都具有一定提升，说明 RELM 相较于 ELM 在保证分类正确率的同时能更好地控制模型的复杂度。改进的 GA-RELM 相较于 RELM 在多部位正反面识别中的准确率提高了约 5.3 个百分点，运行时间减少 24.09 s，这说明采用 GA 算法对隐藏层进行偏置寻优，既保留了 ELM 的泛化能力又提高了分类准确率，同时还节约了 RELM 对于隐藏层偏置的搜索时间。

#### 3.2 本研究算法与其他算法对比

目前叶片分类多采用 SVM、KNN、BP 神经网络等分类方法与主成分分析(PCA)相结合的方式，通过正反面识别来进行烟叶部位分类的研究较少。本研究针对 GA-RELM 算法与目前烟叶分类常用的深度学习算法(Resnet50)<sup>[17]</sup>、支持向量机 SVM<sup>[11]</sup>、KNN(K 最近邻)<sup>[18]</sup>和 BP 神经网络<sup>[12]</sup>展开对比试验研究，分析各种算法的准确率、精确率、召回率、F1 评分和时间这 5 个指标。其中，Resnet50 采用相同数据集进行深度学习，批处理大小为 4，进行 30 次迭代训练，分类结果如表 7 所示。

表 7 不同分类算法对烤烟正反面的识别效果

Table 7 Recognition effects of different classification algorithms on front and back of flue-cured tobacco

分类算法	正反面识别					多部位正反面识别				
	准确率/%	精确率/%	召回率/%	F1 评分/%	时间/s	准确率/%	精确率/%	召回率/%	F1 评分/%	时间/s
Resnet50	85.330 0	85.560 0	85.560 0	85.550 0		55.890 0	60.050 0	57.520 0	58.360 0	
SVM	97.050 4	96.830 2	97.000 0	97.000 0	9.780 0	77.611 6	77.259 7	77.000 0	77.000 0	15.400 0
KNN	88.817 9	88.816 6	88.819 9	88.817 4	0.270 0	79.612 9	80.246 1	81.452 4	80.595 9	0.490 0
BP	97.476 6	95.865 1	95.872 0	95.864 9	468.000 0	81.178 7	82.240 3	82.380 9	82.257 6	449.000 0
GA-RELM	98.547 8	98.283 6	98.283 4	98.283 4	3.030 0	91.459 9	91.299 8	90.552 7	90.876 0	131.760 0

由表 7 可见：KNN 在分类效率上有较大优势，但 KNN 的分类器过于简单，导致分类精度较低；SVM 和 BP 神经网络在烤烟烟叶正反面的分类上具有较好性能，但是二者在进一步针对部位分类时的效果较差。SVM 在处理小样本数据时通常表现更好。BP 神经网络需要通过反向传播算法不断调整权重和偏置，直到达到收敛条件，需要迭代多次才能收敛，运行效率较低；同时 BP 神经网络泛化能力较差，容易陷入局部最优。Resnet50 对于正反面

的分类效果与文献[17]中的分类效果存在一定差距，这可能是因为：一方面，由于模型所学习的特征对于烟叶中部正反面(C、Cf)的区分效果较差，造成整体正确率降低；另一方面，深度学习算法对样本数据的依赖度较大，文献[17]中所运用的数据规模较大，其模型也取得了较好的识别效果，但在实际应用中，烟叶分级的样本数据量大小不一，采用此方法会导致模型识别效果不稳定。相较于其他方法，本文方法在正反面和多部位正反面的识别中效

果更优,且对样本量的依赖较小,可为后续烟叶分级分类研究提供参考依据。

#### 4 结论与讨论

烤烟分级方法要运用在实际中,需考虑叶片所拥有的多种状态,因此,在数据采集过程中,采集自然状态下的3种常见叶态(平整、翻折和卷曲),且不进行任何预处理操作,可以最大程度地贴近实际应用环境。

针对自然状态的烤烟多部位正反面识别方法,GA-RELM模型对自然状态的烤烟正反面和多部位正反面的分类精度分别为98.5478%和91.4599%,其运算时间相较于原ELM分别减少2.56s和5.72s。所采用的特征工程方法,在实现特征降维的同时深入挖掘了多特征组合之间的关系以及各类叶片与所提取特征间特有的联系,正反面识别和多部位正反面识别特征维度由129维分别降至9维和25维,在保证准确率的同时,极大地提高了运行效率。特征参数优选分析结果可进一步为烤烟的特征提取提供参考,提高特征提取效率。本研究所提算法能够实现自然状态下烟叶多部位正反面的识别分类,且可在未区分正反面的数据集上进行烤烟部位分类和分级。

#### 参考文献:

- [1] GB 2635—1992. 烤烟[S].
- [2] YIN Y, XIAO Y J, YU H C. An image selection method for tobacco leave grading based on image information[J]. *Engineering in Agriculture, Environment and Food*, 2015, 8(3): 148–154.
- [3] 刘建新,曾婧,徐可,等. 基于形态学和小波变换的烟叶病斑分割[J]. *图学学报*, 2018, 39(5): 933–938.
- [4] 张永安,马彩娟,刘立博,等. 不同产地烤烟叶片颜色正反面色差及其对感官评吸质量的影响[J]. *安徽农业科学*, 2020, 48(13): 198–201.
- [5] 李雨芮,李佛琳,杨文奎,等. 烤烟红花大金元品种烟叶色度值特征分析[J]. *山西农业科学*, 2023, 51(8): 888–895.
- [6] 蔡杰,朱国蕾,孙晓燕,等. 基于不同部位烟叶外观特征与内在质量分析研究[J]. *农业与技术*, 2019, 39(13): 27–28.
- [7] 徐淼,朱波,刘宇晨,等. 基于改进Mask R-CNN的多片烟叶部位的同步识别[J]. *湖南农业大学学报(自然科学版)*, 2023, 49(2): 170–175.
- [8] 赵晨,王涛,郭伟雄,等. 基于机器视觉的烤烟烟叶部位的智能识别[J]. *湖南农业大学学报(自然科学版)*, 2023, 49(4): 405–411.
- [9] 王戈,丁冉,徐玮杰,等. 计算机视觉和智能识别技术在烤烟烟叶分级中的应用[J]. *计算机与应用化学*, 2019, 36(5): 548–553.
- [10] HE Z F, CHEN G C, ZHANG Y H, et al. Pyramid feature fusion through shifted window self-attention for tobacco leaf classification[J]. *Expert Systems with Applications*, 2023, 230: 120601.
- [11] 姚学练,贺福强,平安,等. 基于PCA-GA-SVM的烟叶分级方法[J]. *烟草科技*, 2018, 51(12): 98–105.
- [12] 杨述斌,董春林,王锋,等. 基于PCA-SSA-BP神经网络的烟叶分级方法[J]. *自动化与仪表*, 2023, 38(2): 70–75.
- [13] DASARI S K, PRASAD V. A novel and proposed comprehensive methodology using deep convolutional neural networks for flue cured tobacco leaves classification[J]. *International Journal of Information Technology*, 2019, 11(1): 107–117.
- [14] LI R D, ZHANG X B, LI K Q, et al. Nondestructive and rapid grading of tobacco leaves by use of a hand-held near-infrared spectrometer, based on a particle swarm optimization-extreme learning machine algorithm[J]. *Spectroscopy Letters*, 2020, 53(9): 685–691.
- [15] 祖丽皮亚·艾尼,麦合甫热提,努尔毕亚·亚地卡尔,等. 灰度共生矩阵特征加权融合的维文签名鉴别[J]. *计算机工程与设计*, 2018, 39(4): 1195–1200, 封3.
- [16] ARIVAZHAGAN S, GANESAN L, PRIYAL S P. Texture classification using Gabor wavelets based rotation invariant features[J]. *Pattern Recognition Letters*, 2006, 27(16): 1976–1982.
- [17] 鲁梦瑶,周强,姜舒文,等. 基于深度学习与多尺度特征融合的烤烟烟叶分级方法[J]. *中国农机化学报*, 2022, 43(1): 158–166.
- [18] 张慧,张文伟,张永毅,等. 基于高光谱与纹理融合的烤烟分类方法研究[J]. *中国烟草学报*, 2022, 28(3): 72–80.

责任编辑:伍锦花  
英文编辑:柳正