

DOI: 10.13652/j.spjx.1003.5788.2024.60122

基于三维荧光光谱和 ISSA-SVM 的食用植物油鉴别

张 静¹ 齐国红¹ 陈景召¹ 曹晓丽¹ 李莉莉²

(1. 郑州西亚斯学院, 河南 郑州 451100; 2. 河南农业大学, 河南 郑州 450046)

摘要: [目的] 提高食用植物油的分类精度, 建立基于三维荧光光谱和 ISSA-SVM 的食用植物油鉴别模型。[方法] 结合三维荧光光谱特征信息, 运用改进的麻雀搜索算法优化 SVM 模型参数, 构建一个融合三维荧光光谱信息特征和 ISSA-SVM 模型的食用植物油鉴别方法。[结果] 与 SVM 模型、GA-SVM 模型、PSO-GA 模型和 SSA-SVM 模型相比, ISSA-SVM 模型的食用植物油分类精度最高, 为 100%。[结论] ISSA-SVM 模型具有更高的收敛效率、系统稳定性以及避免局部最优解的能力, 可以有效应对复杂多变的样本分类任务。

关键词: 支持向量机; 麻雀搜索算法; 三维荧光光谱; 食用植物油

Classification of edible vegetable oils based on three-dimensional fluorescence spectroscopy and ISSA-SVM

ZHANG Jing¹ QI Guohong¹ CHEN Jingzhao¹ CAO Xiaoli¹ LI Lili²

(1. Zhengzhou Sias College, Zhengzhou, Henan 451100, China; 2. Henan Agricultural University, Zhengzhou, Henan 450046, China)

Abstract: [Objective] To improve the classification accuracy of edible vegetable oils, an identification model based on three-dimensional fluorescence spectroscopy and ISSA-SVM was established. [Methods] Combining the feature information of three-dimensional fluorescence spectroscopy, an improved sparrow search algorithm was used to optimize the parameters of the SVM model, constructing an edible vegetable oil identification method that integrates the characteristics of three-dimensional fluorescence spectroscopy information and the ISSA-SVM model. [Results] Compared with the SVM model, GA-SVM model, PSO-SVM model, and SSA-SVM model, the classification accuracy of the ISSA-SVM model for edible vegetable oils reached 100%. [Conclusion] The ISSA-SVM model has higher convergence efficiency, system stability, and the ability to avoid local optimal solutions, which can effectively cope with complex and variable sample classification tasks.

Keywords: support vector machine; sparrow search algorithm; three-dimensional fluorescence spectroscopy; edible vegetable oils

市场上食用植物品种繁多, 掺杂造假现象频发, 严重损害消费者利益。传统的检测方法如色谱法、质谱法等, 但存在操作复杂、耗时长、成本高等问题, 难以满足快速、大规模检测的需求。三维荧光光谱技术是由激发波长、发射波长和荧光强度构成的三维坐标所表征的矩阵光谱, 其原理是物质分子在受到短波长激发光照射后, 吸收光子能量跃迁至激发态, 再以辐射跃迁方式回到基态并发出长波长荧光。该技术具有灵敏度高、分辨率高及

非破坏性等优点, 被广泛应用于生物医学、环境监测、材料科学等领域^[1-2]。支持向量机(support vector machine, SVM)作为一种强大的分类工具, 在非线性、高维及小样本问题分析中具有较强的优势, 但其性能往往受到参数选择的影响。SVM 在结合荧光光谱技术应用于油类污染物分类识别、食用植物油种类区分、橄榄油掺假检测以及农药残留检测等方面均展现出了广泛的应用前景和巨大的潜力。徐瑞煜等^[3]结合三维荧光光谱运用遗传算法优

基金项目: 河南省科技厅科技攻关项目(编号: 232102221029); 河南省教育厅高校品牌专业建设项目(编号: 教政法[2016]896号)

通信作者: 张静(1983—), 女, 郑州西亚斯学院讲师, 硕士。E-mail: zhaqq1124@163.com

收稿日期: 2024-06-18 改回日期: 2024-10-08

化 SVM 实现了 3 种不同组合多环芳烃的分类识别。Sikorska 等^[4]利用同步荧光光谱结合 K 最近邻法和线性判别分析,成功区分不同食用植物油种类。Poulli 等^[5]通过同步荧光光谱与偏最小二乘回归模型,实现了橄榄油中葵花籽油掺假的检测。Guzmán 等^[6]利用荧光光谱结合 PCA 和三向偏最小二乘回归,快速检测出植物油中乳脂掺假。Elena 等^[7]通过荧光光谱评估橄榄油质量,关联其荧光发射量与多项质量参数,实现了橄榄油品质全面评估及来源追溯。Stokes 等^[8]评估了同步荧光技术在摩洛哥坚果油掺假分析中的可行性,通过优化分类器参数和多元校正处理,实现了掺假油的定性定量检测。王书涛等^[9]将三维荧光光谱与改进蚱蜢算法优化 SVM 结合起来,提出了一种油类污染物分类识别的新方法。王书涛等^[10]利用荧光光谱技术与粒子群优化支持向量机(PSO-SVM)算法相结合,该方法能有效避开干扰色素的影响,准确测定混合溶液中胭脂红含量,与误差逆向传播算法相比,PSO-SVM 的预测结果更为准确,相关系数更高,平均回收率更高,均方根误差更低。陈珏等^[11]提出了一种利用荧光光谱技术和机器学习算法检测大白菜中吡虫啉含量的新方法。通过采集特定激发波长下的农药残留光谱数据,经预处理和特征筛选,构建了基于 SVM 的预测模型,并采用麻雀搜索算法(sparrow search algorithm, SSA)优化 SVM 模型参数。结果表明,联合使用卷积平滑与标准正态变量校正的预处理、连续投影算法对遗传算法提取的特征波长进行二次降维,以及 SSA 优化的 SVM 模型,可以实现高精度的吡虫啉含量检测。吴梅凤等^[12]利用三维荧光光谱法结合机器学习技术快速检测橄榄油中掺假廉价油的新方法。通过采集橄榄油及其掺入不同浓度大豆油、玉米油、棕榈油后的荧光光谱数据,采用多种光谱预处理方法进行数据处理。基于多种机器学习算法(K 近邻、随机森林、支持向量机、偏最小二乘法和卷积神经网络)构建了橄榄油定量掺假模型。结果表明,基于偏最小二乘法的定性模型在判别掺假橄榄油方面表现最佳,而标准差标准化预处理结合随机森林算法的定量模型在预测掺假量方面表现最优。该方法具有快速、实时、低成本的特点。为提高食用植物油种类鉴别的精度,研究拟提出一种基于三维荧光光谱和改进麻雀算法优化 SVM 的食用植物油种类鉴别方法,以期食用植物油的快速识别提供依据。

1 相关理论

1.1 支持向量机

针对训练样本 (x_i, y_i) ,引入松弛变量 $\xi_i, \hat{\xi}_i$ 后的 SVM 模型为^[12]:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) \tag{1}$$

$$\text{s.t.} \begin{cases} f(x_i) - y_i \leq \epsilon + \xi_i \\ y_i - f(x_i) \leq \epsilon + \hat{\xi}_i \\ \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, m \end{cases}$$

式中:

- ϵ ——不敏感损失因子;
- C ——惩罚参数;
- w ——权值向量。

在优化支持向量机(SVM)模型过程中,运用一种数学工具——拉格朗日乘子法,将原始优化问题转化为一个更易于处理的对偶问题形式:

$$\max_{\alpha, \hat{\alpha}} \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) y_i - \sum_{j=1}^m (\hat{\alpha}_j + \alpha_j) \epsilon - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) K(x_i, x_j), \tag{2}$$

式中:

- $\alpha_i, \hat{\alpha}_i, \alpha_j, \hat{\alpha}_j$ ——拉格朗日乘子;
- $K(x_i, x_j)$ ——核函数。

RBF 核函数数学式为:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2g^2}\right), \tag{3}$$

式中:

- g ——核函数宽度。

求解式(2),SVM 分类模型可表示为^[13]:

$$f(x, \hat{\alpha}, \alpha) = \text{sgn}\left\{\sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) K(x_i, x_j) + b\right\}, \tag{4}$$

式中:

- b ——超平面偏差。

1.2 麻雀搜索算法

假设 N 只麻雀构成的种群矩阵 X 为:

$$X = [x_1, x_2, \dots, x_N]^T, x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,d}], i = 1, 2, \dots, N, \tag{5}$$

式中:

- d ——待求解优化问题的维数;
- T ——矩阵转置符号;
- x_i ——第 i 个麻雀的位置。

N 只麻雀个体构成的适应度值矩阵为:

$$F_x = [f(x_1), f(x_2), \dots, f(x_N)]^T, f(x_i) = [f(x_{i,1}), f(x_{i,2}), \dots, f(x_{i,d})], \tag{6}$$

式中:

- $f(x_i)$ ——第 i 只麻雀的适应度值。

在 SSA 算法中,发现者的位置更新式为^[14]:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \times \exp\left(-\frac{i}{\alpha \times T_{\max}}\right), R_2 < S_T \\ X_{i,j}^t + Q \times L, R_2 \geq S_T \end{cases}, \quad (7)$$

式中:

t, T_{\max} ——当前迭代次数和最大迭代次数;

Q ——正态分布的随机数且 $Q \in [0, 1]$;

$X_{i,j}^t$ ——第 i 只麻雀在第 t 次迭代时在第 j 维位置上的信息;

L —— $1 \times d$ 的矩阵(矩阵中的元素均为 1);

α ——随机数且 $\alpha \in (0, 1)$;

R_2 ——预警值且 $R_2 \in [0, 1]$;

S_T ——安全值且 $S_T \in [0.5, 1]$ 。

当觅食环境的安全阈值 R_2 低于预设的警戒标准 S_T 时,表明该区域天敌稀少,此时发现者将采取更为广泛的搜索策略以寻找食物。若 R_2 达到或超过 S_T 水平,则预示着觅食区域周边存在天敌活动的迹象,部分敏锐的麻雀会立即发出警报信号,引领整个麻雀群体到其他区域进行觅食活动。

跟随者的位置更新式为^[15-16]:

$$X_{i,j}^{t+1} = \begin{cases} Q \times \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right), i > \frac{N}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \times A^+ \times L, i \leq \frac{N}{2} \end{cases}, \quad (8)$$

式中:

X_p^{t+1} ——第 $t+1$ 次迭代时发现者的最佳位置;

X_{worst}^t ——第 t 次迭代时的全局最差位置;

A^+ —— $1 \times d$ 的矩阵(矩阵中的元素为 1 或 -1,且 $A^+ = A^T(AA^T)^{-1}$)。

为了提升种群的适应性和多样性,随机挑选种群中 10%~20% 的个体作为预警者,预警者的位置更新式为:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \times |X_{i,j}^t - X_{\text{best}}^t|, f_i > f_g \\ X_{i,j}^t + h \times \left[\frac{X_{i,j}^t - X_{\text{worst}}^t}{(f_i - f_w) + \epsilon} \right], f_i = f_g \end{cases}, \quad (9)$$

式中:

β ——步长控制因子且 $\beta \in (0, 1)$;

X_{best}^t ——第 t 次迭代时的全局最佳位置;

h ——运动方位控制因子且 $h \in [-1, 1]$;

ϵ ——最小的常数(防止出现分母为 0 的情形);

f_i, f_g, f_w ——第 i 次迭代的适应度值、最优和最差的适应度值。

若 $f_i > f_g$,麻雀个体将处于种群边缘;若 $f_i = f_g$,位于种群核心区域的麻雀,能够敏锐地发现危险,并倾向于向种群的其他部分聚集,以寻求集体保护和安全感。

1.3 改进 SSA

在标准 SSA 算法中,其初始种群个体是随机产生的。由于这种随机性,初始种群在解空间内的分布往往缺乏均匀性,不仅无法全面覆盖整个解空间,还可能因此导致

算法在搜索过程中过早地收敛于局部最优解,从而降低了算法的全局搜索能力和收敛速度。为了克服这一局限性,采用佳点集法优化 SSA 算法的初始种群生成^[17]。假定所求优化问题的可行域为:

$$x \in [l, u], [l, u] = \{x \in R^d | l_k \leq x_k \leq u_k, k = 1, 2, \dots, d\}. \quad (10)$$

若存在一个 d 维空间的单位立方体 G ,在 G 空间中取一点 $r = (r_1, r_2, \dots, r_d)$,其中 $r_k = \{2\cos(2\pi k/p), 1 \leq k \leq d\}$, p 是满足 $(p-3)/2 \geq d$ 的最小素数,此时 r 为单位立方体 G_d 空间中的佳点,那么 r 在 G 空间中的佳点集 $p_M(i) = \{\{r_{1i}\}, \{r_{2i}\}, \dots, \{r_{di}\}\}$, $i = 1, 2, \dots, M$,其中 M 为佳点的数量。佳点集在可行域中的映射式为:

$$x_k^i = l_k + p_M(i)_k(u_k - l_k), \quad (11)$$

式中:

x_k^i ——第 i 个麻雀个体在第 k 维上映射值。

针对初期阶段发现者快速收敛至全局最优解可能引发的局限性问题,通过引入基于历史全局最优解的动态调整策略与非线性递减的惯性权重系数改进发现者的位置更新公式。此设计旨在平衡算法的收敛效率与探索能力,既保留了算法快速逼近最优解的优势,又显著降低了早熟收敛至局部最优的风险,进而增强了搜索过程的全局覆盖性与解的精确度。改进的发现者更新式为^[18-19]:

$$w = w_{\text{start}} + (w_{\text{start}} - w_{\text{end}}) \times \left[\frac{2t}{T_{\max}} - \left(\frac{t}{T_{\max}} \right)^2 \right], \quad (12)$$

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t + w \times (f_{j,g}^t - X_{i,j}^t) \times \text{rand}, R_2 < S_T \\ X_{i,j}^t + Q \times L, R_2 \geq S_T \end{cases}, \quad (13)$$

式中:

$w_{\text{start}}, w_{\text{end}}$ ——惯性权重的初始值和终止值。

1.4 ISSA-SVM 模型

在构建基于 SVM 模型的食用植物油鉴别系统中,识别结果的精确性深受其关键参数——惩罚参数 C 与核函数参数 g 的影响。当 C 值设置过高时,模型易陷入过拟合状态,即过分拟合训练数据而丧失对新样本的泛化能力; C 值过低则可能导致模型欠拟合,无法充分捕捉数据特征。同样地,核函数参数 g 的调节也至关重要:较小的 g 值往往能提升模型的泛化性能,使其在面对未见数据时表现更佳;而 g 值过大则可能削弱这种能力。为优化 SVM 模型在食用植物油鉴别任务中的性能,引入 ISSA 算法来自动调整这两个关键参数。通过 ISSA 算法的全局搜索能力,旨在找到最优的 C 和 g 组合,以实现模型在识别精度与泛化性能之间的最佳平衡。

2 食用植物油鉴别模型

2.1 目标函数

为提高食用植物油鉴别的准确率,提出一种基于三维荧光光谱技术与 ISSA-SVM 模型的食用植物油鉴别方

法,优化目标为:

$$\min f(C, g) = \sqrt{\frac{1}{n} \sum_{i=1}^n [x(i) - x_p(i)]^2}$$

$$\text{s.t. } C \in [C_{\min}, C_{\max}]$$

$$g \in [g_{\min}, g_{\max}]$$

式中:

n ——训练集样本的数量;

$x(i), x_p(i)$ ——第 i 个训练样本的食用植物油实际类别和预测类别;

$[C_{\min}, C_{\max}], [g_{\min}, g_{\max}]$ ——惩罚参数 C 和核参数 g 的寻优区间。

2.2 算法步骤

基于三维荧光光谱与 ISSA-SVM 的食用植物油鉴别算法步骤可具体地描述为:

步骤一:数据预处理与划分

加载食用植物油的三维荧光光谱数据集,并进行归一化处理以消除量纲差异对模型训练的影响。随后,采用合理的比例将数据集划分为训练集和测试集,确保模型既能充分学习特征,又能有效评估其泛化能力。

步骤二:模型与算法参数初始化

准备阶段,设定 SVM 模型的关键参数(惩罚参数 C 和核参数 g)为待优化目标,并引入改进的麻雀搜索算法(ISSA)进行参数寻优。初始化 ISSA 算法参数:最大迭代次数 T_{\max} 、发现者比例 P_D 、侦察者比例 S_D 、预警值 S_T 、种群规模 N 。由于优化目标为两个参数(惩罚参数 C 和核参数 g),因此设定初始种群维数为 $\text{dim}=2$ 。

步骤三:适应度评估与最优解追踪

对于 SSA 算法中的每个麻雀个体,将其代表的参数组合(C 和 g 的值)应用于 SVM 模型,并使用训练集数据评估模型的性能。按式(14)计算每个个体的适应度值,识别并记录当前迭代中的最优适应度值及其对应的最优参数组合。

步骤四:位置更新策略

根据 SSA 算法原理,将麻雀种群分为发现者、跟随者和预警者 3 类。对于发现者,按式(3)的策略更新其位置;跟随者则按式(4)调整自身位置;预警者则依据式(5)的策略调整其位置。

步骤五:迭代优化与终止条件判断

重复执行步骤三和步骤四,直至达到预设的最大迭代次数 T_{\max} 。每次迭代后,检查是否满足终止条件,若满足,则输出当前找到的最优适应度值及其对应的 SVM 模型参数(C 和 g);反之,返回步骤二。

步骤六:模型训练与测试

将 ISSA 算法寻优得到的最佳 SVM 参数应用于模型训练,并使用测试集数据评估训练后模型的性能。通过对比实际分类结果与预测结果,验证该算法在食用植物

油鉴别中的有效性和准确性。基于三维荧光光谱与 ISSA-SVM 的食用植物油鉴别算法流程如图 1 所示。

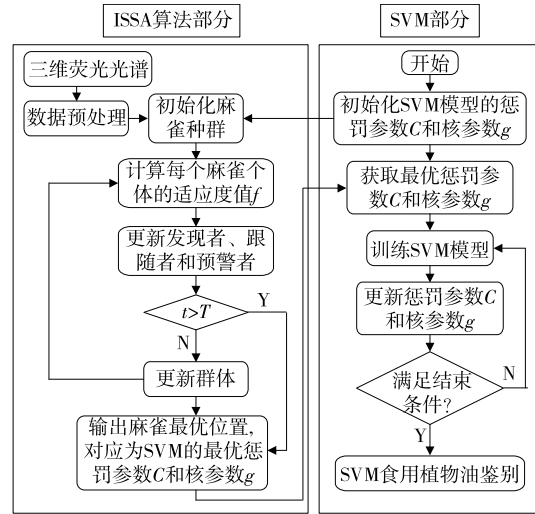


图 1 优化 SVM 模型的食用植物油鉴别流程图
Figure 1 Flowchart of edible vegetable oil identification based on optimized SVM model

2.3 光谱数据预处理

选择市售某品牌菜籽油、大豆油和玉米油为研究对象,共采集 58 组数据,29 组作为训练样本,其余作为测试样本。在荧光光谱分析中,由于光谱仪内部组件如激发光源和激发单色仪的固有系统误差,直接测量的荧光光谱往往难以精准反映真实的光谱特性,尤其是在峰值定位和荧光强度评估上易产生偏差。因此,实施精确的荧光检测前,对仪器进行细致的激发与发射校正至关重要,这是确保数据准确性的基础步骤。此外,待测样品溶剂的特性以及光散射现象(包括瑞利散射和拉曼散射)也是不可忽视的干扰因素。这些干扰会显著影响光谱数据的纯净度,甚至掩盖样品本身的关键光谱特征,以菜籽油溶液为例(图 2),散射光谱的存在严重扭曲了光谱的真实面貌。在数据分析前,必须采取去散射处理。空白扣除法结合缺损数据重构(MDR)技术是一种高效的解决方案,能显著减少散射对光谱数据的干扰。经过这一校正和去散射流程后,原本被掩盖的荧光峰和光谱特征得以清晰展现,不仅增强了数据的逻辑性,也显著降低了因仪器误差和外界干扰导致的数据偏差。

3 试验方法与结果

3.1 仪器与参数

采用日本 Hitachi 公司的 F-7000 型荧光光谱仪作为分析工具,为了全面捕获食用植物油的荧光特征,激发波长扫描区间为 250~400 nm,发射波长扫描范围为 270~450 nm,可确保激发波长相较于发射波长超前 20 nm,有

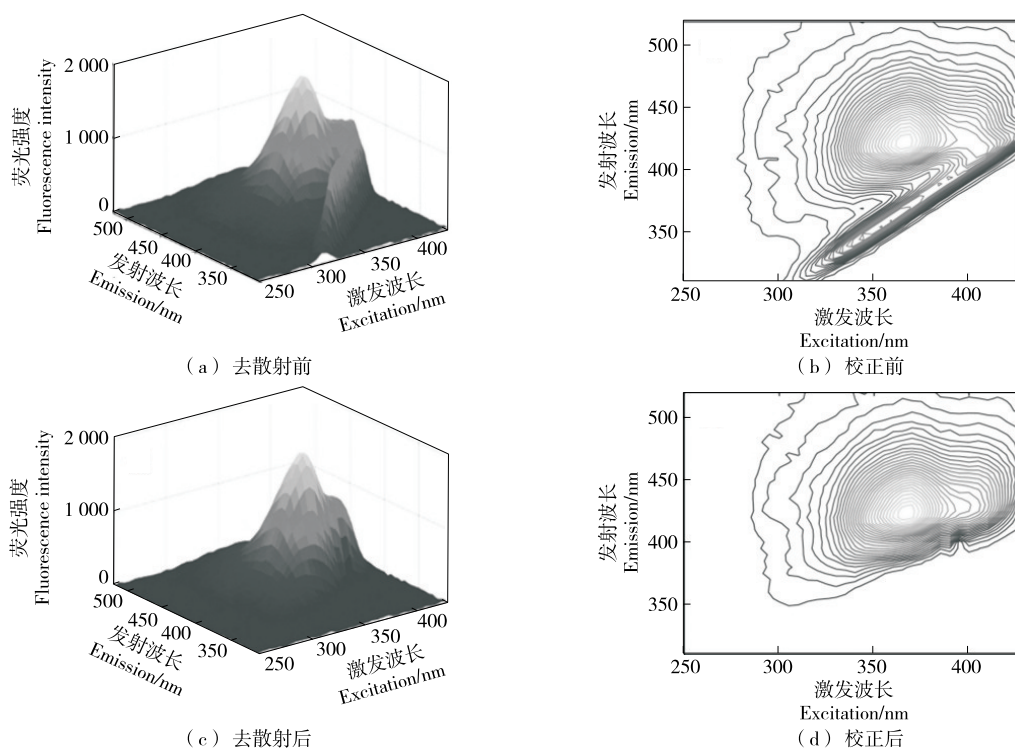


图2 菜籽油三维荧光光谱图

Figure 2 Three-dimensional fluorescence spectra of rapeseed oil

效规避瑞利散射对试验结果的潜在干扰。同时,为了提升数据解析的精细度,设置5 nm的激发与发射扫描步长,并统一将狭缝宽度调至2.5 nm。为了进一步增强数据的可靠性与准确性,数据重复扫描3次。

3.2 结果与分析

菜籽油、大豆油和玉米油的荧光光谱图和等高线图如图3所示。将菜籽油、大豆油和玉米油的类别标签编码分别设为1、2、3,菜籽油、大豆油和玉米油的荧光光谱图和等高线图作为SSA-SVM的输入,食用植物油的类别标签作为SSA-SVM的输出,建立三维荧光光谱与SSA-SVM模型的食用植物油鉴别模型。

传统的SVM模型参数设定为:惩罚参数 C 为100,核参数 g 为0.1,分类结果如图4所示。图4(a)中,菜籽油表现出色,所有样本均被正确分类,准确率达到100%。大豆油的分类效果相对较弱,准确率为77.8%,显示出较高的误分类率。玉米油的分类表现较为稳定,准确率为90%,但仍有提升空间。图4(b)中,两个大豆油样本被错误地识别为玉米油,可能源于两类油在某些特征上的相似性。一个玉米油样本则被误判为菜籽油,可能是由于特征选择或模型参数设置上的局限性。此次分类的最佳适应度为89.67%,反映了模型在训练数据上的整体拟合效果,共有3个样本被错误分类,最终的分类识别准确率为91.18%。

3.3 不同算法优化SVM对比

不同算法通用参数设定为:种群规模 $N=20$ 、进化代数 $T_{\max}=100$;ISSA和SSA算法^[20-21]中,发现者比例 $P_D=0.2$ 、预警值 $S_r=0.8$ 、侦察者比例 $S_D=0.2$ 、 $w_{\text{start}}=0.9$ 、 $w_{\text{end}}=0.2$;粒子群算法(PSO)^[22-23]中,惯性权重 $w=0.2$ 、学习因子 $c_1=c_2=2$;遗传算法(GA)^[24-25]中,交叉概率 $p_c=0.3$,变异概率 $p_m=0.7$ 。

由图5可知,ISSA展现出卓越的收敛性能,在第2代迭代即达到最佳适应度值,并保持稳定;其平均适应度值自第3代起紧密跟随最佳适应度曲线,显示出高度的稳定性和效率。SSA在第8代迭代后达到最佳适应度,其平均适应度曲线在前期(37代前)呈连续振荡,之后趋于稳定并与最佳适应度曲线重合。PSO的收敛速度介于ISSA与SSA之间,于第4代迭代后达最佳适应度并趋于稳定,但平均适应度与最佳适应度之间存在一定的差距。GA展现出较慢的收敛特性,其最佳适应度曲线经历多个阶段后在第12代左右趋于稳定,而平均适应度则持续处于震荡状态。经ISSA优化后的SVM参数使得分类识别率达到100%,表现出最佳的优化效果。SSA的最佳适应度略低,为97.06%,显示出其优化能力的次优性。PSO与GA均能达到94.12%的最佳适应度,但在收敛速度和稳定性上,PSO表现更佳。

不同算法优化SVM模型得到的参数代入SVM模型进行分类,分类结果如图6和表1所示。由图6可知,

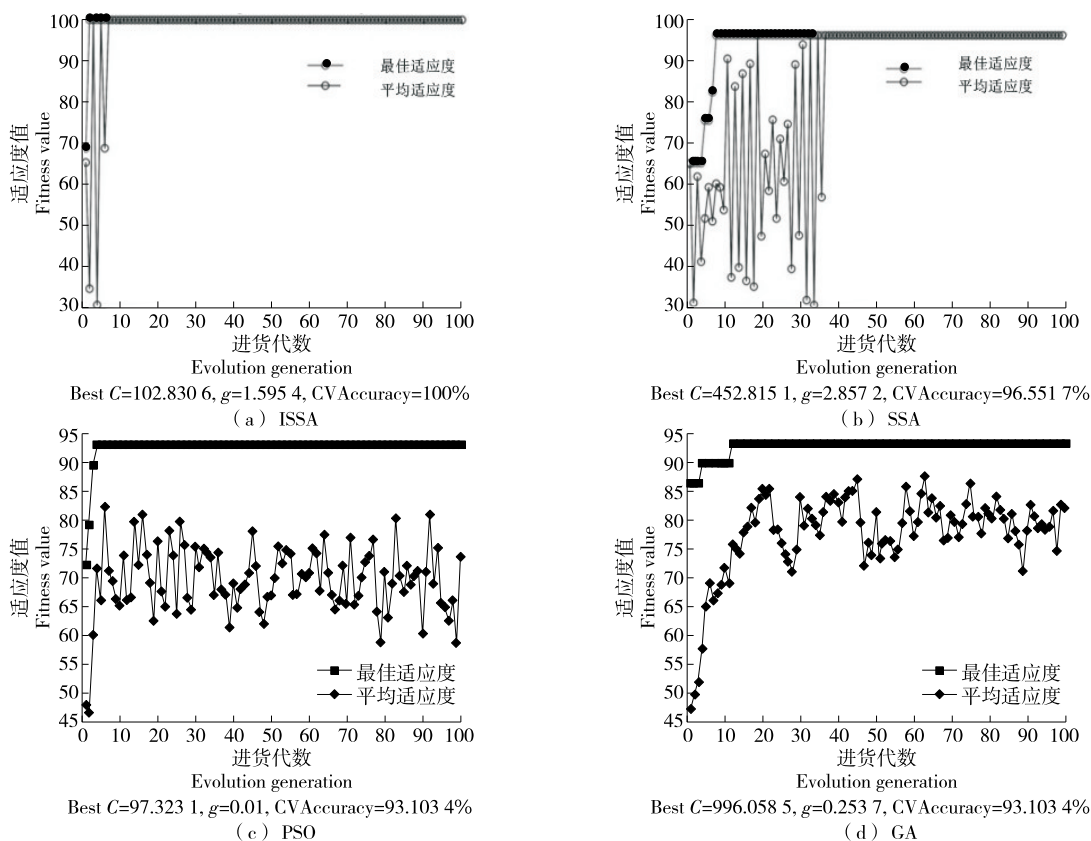


图5 ISSA、SSA、PSO和GA的适应度曲线

Figure 5 Fitness curves for ISSA, SSA, PSO and GA

一个相对较高的水平,均为93.10%,显示了这两种优化算法在提升SVM分类能力上的积极作用。

由表1可知,在食用植物油鉴别任务中,引入改进型麻雀搜索算法(ISSA)与支持向量机(SVM)相结合的模型(ISSA-SVM)展现出了卓越的性能。具体表现为:

(1) 准确率显著提升。ISSA-SVM的准确率高达95.96%,与SSA-SVM、PSO-SVM、GA-SVM和未经优化的SVM相比,分别提升了4.35%,5.92%,5.23%,16.23%,表明ISSA在优化SVM模型参数方面相较于其他算法具有显著优势,大幅提高了鉴别任务的准确性。

(2) 检测能力优越。在检测率方面,ISSA-SVM表现出色,检测率达95.38%。与SSA-SVM、PSO-SVM、GA-SVM及SVM相比,其检测率分别提升了4.28%,5.26%,

5.66%,15.52%,进一步验证了ISSA-SVM在提升模型检测能力方面的有效性。

(3) 错误率显著降低。ISSA-SVM的错误率仅为4.15%,与DE-SVM、PSO-SVM、GA-SVM和SVM相比,其错误率分别降低了2.19%,2.87%,2.96%,5.01%,表明ISSA-SVM在减少误判、提高鉴别精度方面具有显著优势。

综上,ISSA-SVM在食用植物油鉴别任务中相较于其他优化算法结合的SVM模型及未经优化的SVM模型具有更高的准确率、更强的检测能力和更低的错误率。

3.4 不同噪声水平对比分析

为了应对复杂环境下的食用植物油鉴别,对原始数据添加噪声,对比不同噪声水平下的鉴别准确率,结果见表2。

表1 不同算法对比结果

Table 1 Comparison results of different algorithms

算法	准确率/%	检测率/%	错误率/%	计算时间/s	SVM参数
ISSA-SVM	95.96	95.38	4.15	52.47	$C=102.8306, g=1.5954$
SSA-SVM	91.61	91.10	6.34	60.04	$C=452.8151, g=2.8572$
PSO-SVM	90.04	90.12	7.02	75.03	$C=97.3231, g=0.01$
GA-SVM	90.73	90.62	7.11	71.45	$C=996.0585, g=0.2537$
SVM	79.73	79.86	9.16	91.51	$C=100, g=5$

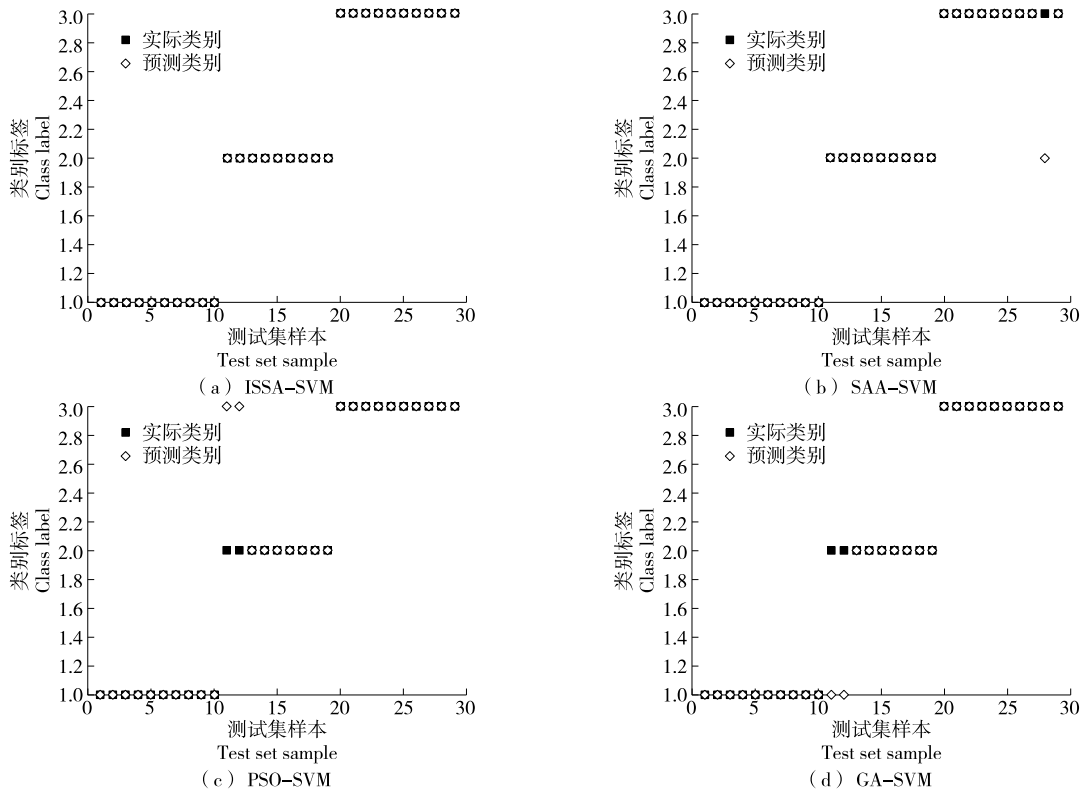


图 6 不同算法优化 SVM 模型 的分类结果

Figure 6 Classification results of SVM models optimized by different algorithms

表 2 不同噪声水平对比结果

Table 2 Comparison results of different noise levels

噪声水平(SNR)	准确率/%	噪声水平(SNR)	准确率/%
5	96.65	20	91.62
10	92.98	30	90.17

由表 2 可知,随着噪声水平的增加,食用植物油鉴别准确率呈下降趋势,但仍保持在较高水平,说明该模型具有较高的稳定性和抗噪声能力。

4 结论

为了提高食用植物油的分类精度,结合三维荧光光谱特征信息,运用改进的麻雀搜索算法优化支持向量机(SVM)模型参数,构建了一个融合三维荧光光谱信息特征和 ISSA-SVM 模型的食用植物油鉴别方法。与 SVM 模型、GA-SVM 模型、PSO-GA 模型和 SSA-SVM 模型相比,ISSA-SVM 模型的食用植物油分类精度高达 100%。ISSA-SVM 模型具有更高的收敛效率、系统稳定性以及避免局部最优解的能力,可以有效应对复杂多变的样本分类任务。后续可对比其他三维荧光光谱的预处理与特征提取技术。

参考文献

[1] 程朋飞,朱燕萍,潘金燕,等. 三维荧光光谱结合 IGOA-SVM 分类鉴别油类污染物[J]. 光谱学与光谱分析, 2024, 44(4):

1 031-1 038.

CHENG P F, ZHU Y P, PAN J Y, et al. Classification and identification of oil pollutants based on three-dimensional fluorescence spectroscopy combined with IGOA-SVM[J]. Spectroscopy and Spectral Analysis, 2024, 44(4): 1 031-1 038.

[2] 王书涛,刘娜,程琪,等. 三维荧光光谱结合 GA-SVM 对多环芳烃的分类鉴别[J]. 光谱学与光谱分析, 2020, 40(4): 1 149-1 155.

WANG S T, LIU N, CHENG Q, et al. Classification and identification of polycyclic aromatic hydrocarbons based on three-dimensional fluorescence spectroscopy combined with GA-SVM[J]. Spectroscopy and Spectral Analysis, 2020, 40(4): 1 149-1 155.

[3] 徐瑞端,朱焯炜,胡扬俊,等. 三维荧光光谱结合 PCA-SVM 对几种浓香型白酒的鉴别[J]. 光谱学与光谱分析, 2016, 36(4): 1 021-1 026.

XU R Y, ZHU Z W, HU Y J, et al. Identification of several strong aroma wine types based on three-dimensional fluorescence spectroscopy combined with PCA-SVM [J]. Spectroscopy and Spectral Analysis, 2016, 36(4): 1 021-1 026.

[4] SIKORSKA E, GÓRECKI T, KHMELINSKII I V, et al. Classification of edible oils using synchronous scanning fluorescence spectroscopy[J]. Food Chemistry, 2005, 89 (2): 217-225.

[5] POULLI K I, MOUSDIS G A, GEORGIU C A. Synchronous

- fluorescence spectroscopy for quantitative determination of virgin olive oil adulteration with sunflower oil[J]. *Analytical & Bioanalytical Chemistry*, 2006, 386(5): 1 571-1 575.
- [6] NTAKATSANE M P, LIU X M, ZHOU P. Short communication: rapid detection of milk fat adulteration with vegetable oil by fluorescence spectroscopy[J]. *Journal of Dairy Science*, 2013, 96(4): 2 130-2 136.
- [7] GUZMÁN E, BAETEN V, PIERNA J A F, et al. Evaluation of the overall quality of olive oil using fluorescence spectroscopy [J]. *Food Chemistry*, 2015, 173: 927-934.
- [8] STOKES T D, FOTEINI M, BROWNFIELD B, et al. Feasibility assessment of synchronous fluorescence spectral fusion by application to argan oil for adulteration analysis[J]. *Applied Spectroscopy*, 2018, 72(3): 432-441.
- [9] 王书涛, 刘诗瑜, 王志芳, 等. 基于ICSO-SVM和三维荧光光谱的山梨酸钾浓度检测[J]. *光谱学与光谱分析*, 2020, 40(5): 1 614-1 619.
- WANG S T, LIU S Y, WANG Z F, et al. Detection of potassium sorbate concentration based on ICSO-SVM and three-dimensional fluorescence spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(5): 1 614-1 619.
- [10] 王书涛, 彭涛, 李明珊, 等. 基于三维荧光光谱和PSO-SVM对胭脂红含量的测定[J]. *光谱学与光谱分析*, 2019, 39(1): 150-155.
- WANG S T, PENG T, LI M S, et al. Determination of carmine content based on three-dimensional fluorescence spectroscopy and PSO-SVM[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(1): 150-155.
- [11] 陈珏, 李佳琮, 刘翠玲, 等. 荧光光谱技术结合机器学习算法检测大白菜中吡虫啉含量[J]. *食品安全质量检测学报*, 2023, 14(13): 134-140.
- CHEN J, LI J C, LIU C L, et al. Determination of imidacloprid in cabbage by fluorescence spectroscopy combined with machine learning algorithms[J]. *Journal of Food Safety and Quality*, 2023, 14(13): 134-140.
- [12] 吴梅凤, 李杨, 李敏敏, 等. 三维荧光光谱法快速检测橄榄油中掺假廉价油[J]. *食品安全质量检测学报*, 2024, 15(5): 289-297.
- WU M F, LI Y, LI M M, et al. Rapid detection of adulterated cheap oil in olive oil by three-dimensional fluorescence spectrometry[J]. *Journal of Food Safety and Quality*, 2024, 15(5): 289-297.
- [13] XUE J, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. *Systems Science & Control Engineering: An Open Access Journal*, 2020, 8(1): 22-34.
- [14] 黄崇富, 常宇, 刘力超. 基于改进鲸鱼算法和神经网络的食品机器人能耗轨迹规划[J]. *食品与机械*, 2022, 38(9): 108-113.
- HUANG C F, CHANG Y, LIU L C. Energy consumption path planning of food robot based on improved whale algorithm and neural network[J]. *Food & Machinery*, 2022, 38(9): 108-113.
- [15] TUERXUN W, CHANG X, GUO H Y, et al. Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm[J]. *IEEE Access*, 2021, 9: 69 307-69 315.
- [16] 王少英. 基于改进型循环神经网络算法的食品包装智能实时识别系统研究[J]. *食品与机械*, 2023, 39(9): 110-116.
- WANG S Y. Research on intelligent real-time identification system of traffic signs based on improved recurrent neural network algorithm[J]. *Food & Machinery*, 2023, 39(9): 110-116.
- [17] DORNAIKA F, BOSAGHZADEH A, SALMANE H, et al. Graph-based semi-supervised learning with local binary patterns for holistic object categorization[J]. *Expert Systems with Applications*, 2014, 41(17): 7 744-7 753.
- [18] LIU W, DUAN C W, YU B, et al. Multi-pose pedestrian detection based on posterior HOG feature[J]. *Acta Electronica Sinica*, 2015, 43(2): 217-224.
- [19] 赖嘉雯, 汪宇玲, 蔡晓宇, 等. 基于CNN-SVM的多维度信息融合半月板撕裂分类方法[J]. *波谱学杂志*, 2023, 40(4): 423-434.
- LAI J W, WANG Y L, CAI X Y, et al. Multidimensional information fusion method for meniscal tear classification based on CNN-SVM[J]. *Chinese Journal of Magnetic Resonance*, 2023, 40(4): 423-434.
- [20] 李旭东, 李艳军, 曹愈远, 等. 基于CNN-SVM的飞机EHA故障诊断算法研究[J]. *西北工业大学学报*, 2023, 41(1): 230-240.
- LI X D, LI Y J, CAO Y Y, et al. Study on fault diagnosis algorithms of EHA based on CNN-SVM[J]. *Journal of Northwestern Polytechnical University*, 2023, 41(1): 230-240.
- [21] 方明明, 刘静. 基于回归卷积神经网络的近红外光谱苹果脆片品质评价方法研究[J]. *食品科技*, 2020, 45(7): 303-308.
- FANG M M, LIU J. Evaluation method of apple chips quality by near infrared spectroscopy based on regressive convolutional neural network[J]. *Food Science and Technology*, 2020, 45(7): 303-308.
- [22] 王浩云, 李晓凡, 李亦白, 等. 基于高光谱图像和3D-CNN的苹果多品质参数无损检测[J]. *南京农业大学学报*, 2020, 43(1): 178-185.
- WANG H Y, LI X F, LI Y B, et al. Research on non-destructive detection of apple multi-quality parameters based on hyperspectral imaging technology and 3D-CNN[J]. *Journal of Nanjing Agricultural University*, 2020, 43(1): 178-185.
- [23] LIU Y, ZENG L, HUANG Y. An efficient HOG-ALBP feature for pedestrian detection[J]. *Signal Image and Video Processing*, 2014, 8(1): 125-134.
- [24] XUE D X, ZHANG R, FENG H, et al. CNN-SVM for microvascular morphological type recognition with data augmentation[J]. *Journal of Medical and Biological Engineering*, 2016, 36(6): 755-764.
- [25] VALAPPIL N K, MEMON Q A. CNN-SVM based vehicle detection for UAV platform[J]. *International Journal of Hybrid Intelligent Systems*, 2021, 17(1): 1-12.