

基于近红外高光谱成像技术的塑料分类(特邀)

胡锡敦, 尹禄*, 杨钦晨, 王乐**

中国计量大学光学与电子科技学院, 浙江 杭州 310018

摘要 塑料因其可塑性与低成本在日常生活与工业中被广泛使用,然而这也带来环境污染与资源浪费等问题,因此塑料分类成为重要研究课题。为验证高光谱成像技术在塑料分类中的可行性,采用近红外高光谱成像技术(NIR-HSI),比较了1100~1650 nm波段数据在9种常见塑料分类中的效果。涵盖K邻近法(K-NN)、支持向量机(SVM)、粒子群算法训练的SVM(PSO-SVM)、遗传算法优化的SVM(GA-SVM)等机器学习方法。通过验证数据筛选模型准确率后,将其应用于高光谱图像,通过可视化分类对比原始图像评估模型效果。结果显示,基于欧氏距离、余弦相似度的K-NN和GA-SVM分类效果最佳,验证数据的精度分别达到96.14%、96.21%和98.67%,在可视化分类上也呈现出良好效果。高光谱成像技术在塑料分选中具有很高的应用价值,只需获取特定塑料的光谱数据并进行适当处理,即可对不同颜色、形状、工艺的同类塑料制品进行有效区分。

关键词 近红外高光谱成像; 塑料分类; 机器学习; 可视化分类

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP232402

Classification of Plastics Based on Near-Infrared Hyperspectral Imaging Technology (Invited)

Hu Xidun, Yin Lu*, Yang Qinchen, Wang Le**

College of Optics and Electronic Science and Technology, China Jiliang University,
Hangzhou 310018, Zhejiang, China

Abstract Plastics are widely used in daily life and industry because of their plasticity and low costs. However, they cause problems, such as environmental pollution and resource waste, and plastic classification has become an important research topic. Near-infrared hyperspectral imaging (NIR-HSI) is used to compare the effect of 1100–1650 nm band data in classifying nine common plastics to verify the feasibility of hyperspectral imaging in plastic sorting. Machine learning methods such as the *K*-neighborhood method (*K*-NN), support vector machine (SVM), SVM trained by particle swarm algorithm (PSO-SVM), and SVM optimized by genetic algorithm (GA-SVM) are used. After verifying the accuracy of the data screening model, it is applied to hyperspectral images, and the model effect is evaluated by comparing the original images through visual classification. The results show that the *K*-NN and GA-SVM based on the Euclidean distance and cosine similarity are the most effective in classification, and the accuracy of the validation data reaches 96.14%, 96.21%, and 98.67%, respectively. Good results are also presented in the visualization classification. The experiment demonstrates that hyperspectral imaging technology has high application value in plastic sorting. This can effectively differentiate similar plastic products based on color, shape, and process by acquiring the spectral data of specific plastics and processing them appropriately.

Key words near-infrared hyperspectral imaging; plastic classification; machine learning; visual classification

1 引言

塑料凭借着可塑性和低成本的优势,成为了社会

发展的重要角色,生活中随处可见,但伴随而来的是大量的废弃塑料。OECD《全球塑料展望报告》^[1]指出,2019年全球的塑料垃圾达3.53亿吨,仅9%得到了回

收稿日期: 2023-10-31; 修回日期: 2023-11-19; 录用日期: 2023-11-27; 网络首发日期: 2023-12-12

基金项目: 国家自然科学基金(51832005,62075203,62305320,1210042018)、国家重点研发计划项目(2021YFC3340400)、浙江省科技计划项目(2022C01127,2021C05005)、浙江省自然科学基金(LQ23A040007)

通信作者: *calla@cju.edu.cn; **yinlu890622@163.com

收,焚烧占比 19%,弃入海洋、土壤等高达 70%。提高回收利用率的先决条件是对塑料垃圾进行有序分类。

我国塑料分类的方式仍旧以人工分选法为主,这种方式不仅效率低下,还影响劳动者身体健康。为节省人力物力,提高分类效率,学者们提出了多种分类方法^[2]:力学分选;光学分选;电磁分选;萃取分离。其中,光学分选具有以下优势:自动化程度高;分选塑料快;分选效率高;错误分选率低;人工成本低;绿色环保。

塑料分选的光学方法包括拉曼光谱法(RS)^[3]、激光诱导荧光技术(LIF)^[4]、激光诱导击穿光谱法(LIBS)^[5]。这些方法有着各自的优缺点:RS对多种类型的样品,包括固体、液体和气体都有很好的适应性,但在一些情况下,样品本身的荧光会对光谱的解析产生影响,对于较厚的样品检测准确度会下降;LIF只能检测具有荧光特性的物质,对环境光的干扰敏感,需要在受控环境中进行实验;LIBS有高的准确度,但相对于其他技术,设备和维护成本较高,需要样品表面干净平整才能获得准确的分析结果。

本文基于高光谱成像技术(HSI)^[6]建立塑料分选模型,HSI是一种具有非常多窄波段的影像数据技术,成像时会得到很多人眼看不到的光谱信息,采集到的是具备物体信息和光谱信息的三维数据立方。凭借着不破坏物体本身和扫描效率高的特点,其在农业^[7-9]、环境监测^[10-11]、医学^[12-13]等领域发展迅速,展现了可观的应用前景和潜在研究价值。2017年,天津大学化学工程与技术学院 Zheng等^[14]利用波长范围为 1000~2500 nm 的高光谱相机对 6 种塑料展开分类研究,利用主成分分析选择特征波段并构建 Fisher 判别模型,最终得到了分类精度为 100% 的模型。2022年,奥胡斯大学生物与化学工程系 Henriksen等^[15]利用波长范围为 955~1700 nm 的高光谱相机对 12 种塑料展开研究,利用无监督机器学习建立分类模型,结合工业塑料添加剂成分,对同种塑料存在的光谱偏差最终导致的分类错误作出了解释。

但上述研究存在一定局限性,训练样本和验证样本为同一批样本,建模过程包含了所有样本的信息。为更全面验证高光谱成像技术在塑料分选中的价值,本实验利用 9 种类别的塑料在近红外波段中光谱曲线的差异性,结合多种机器学习方法,构建了分类识别模型。训练样本的形貌单一,验证样本的形貌多样。

2 实验部分

2.1 材料

塑料样品选自生活常见的 9 种材质,分别为聚对苯二甲酸乙二醇酯(PET)、聚乙烯(PE)、聚氯乙烯(PVC)、聚丙烯(PP)、聚苯乙烯(PS)、聚碳酸酯(PC)、聚甲醛(POM)、丙烯腈-丁二烯-苯乙烯共聚物(ABS)、尼龙(PA)。由于黑色样本的光吸收较强,故未选择黑色样本进行测试。

实验所需样本包含训练样本和验证样本。训练样

本和验证样本的材料种类均为 9 种,各种类的颜色统计如表 1 所示。训练样本为规整的 4 cm×4 cm×3 mm 的塑料片。验证样本为市场内可以购买到的塑料制品,由于高光谱采集系统的视场限制,对部分塑料制品进行了裁剪处理。

表 1 实验样本的种类及颜色统计
Table 1 Species and color statistics of experimental samples

Species	Color (training sample)	Color (testing sample)
PET	Transparent	Transparent, transparent-green
PE	White	White, blue, red, yellow, brown
PVC	Transparent	White, transparent
PP	White, dark yellow	White, blue, brown, transparent
PS	White	Transparent-orange, white
PC	White, transparent	Transparent, transparent-pink
POM	White	White
ABS	White, beige	Dark yellow
PA	White	White

本实验选择的样本包含了许多透明的塑料,通过多次测试后发现,在黑色金属载物台上,透明材料在近红外波段的采集效果较差。为解决采集问题,由于不同材料在不同波段对 PVC 底板的反射率有不同的吸收情况,在实验采集前将待测样本放在一块白色 PVC 底板上,再放到载物台上进行采集,通过测定底板透过不同透明材料的反射率达到检测透明材料光谱相对反射率的目的。塑料样本过薄也会对数据采集造成困难,材料对底板反射率的吸收率过小,无法体现在光谱维度上,可以通过堆叠薄片来提高对底板反射率的吸收强度达到检测目的。模型训练的样本具体如图 1 所示。本实验选择规整的塑料片作为训练集,其工艺简单,材质密度均匀,更能体现近红外波段内不同塑料光谱的差异性,进而延展至塑料制品。

近红外高光谱相机提取光谱信息时可以有效规避颜色带来的光谱差异,验证样本的多样性,初步验证高光谱成像技术在塑料分选上的可行性、实用性。

模型验证样本的具体情况如图 2 所示,为了增加样本的多样性,本实验选择的验证样本有以下特点:每种塑料的颜色多样,详情可见表 1;样本的表面工艺存在差异,包括光滑面和磨砂面;样本形态差异大,形状多样,厚度不一,表面不平整。

2.2 仪器

实验数据的采集工作在杭州彩谱科技有限公司完成,使用的仪器型号为 FS-15 短波近红外高光谱相机,如图 3 所示。该相机的光谱范围为 900~1700 nm,光谱波段数为 254,是推扫式成像高光谱相机。

本实验的高光谱采集系统如图 4 所示,主要结构为高光谱相机、光源、载物台和计算机。采集图像时,将待测物平稳地放在移动载物台上,通过计算机的控制程序完成扫描,获取所需的高光谱图像。

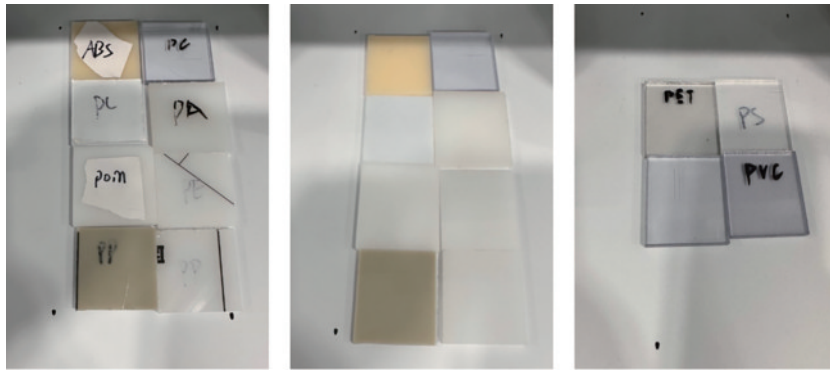


图1 训练样本
Fig. 1 Training samples

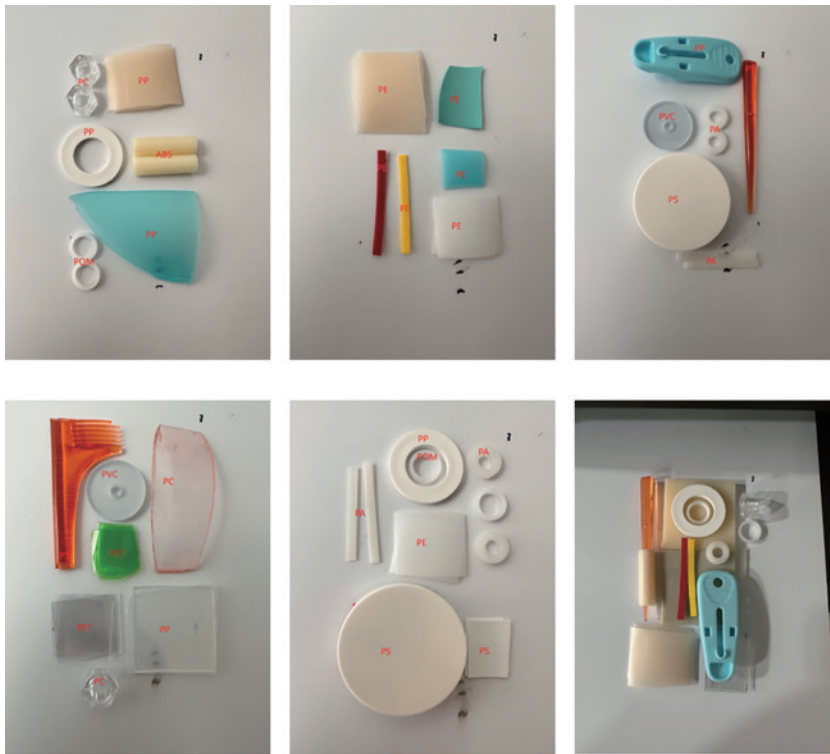


图2 验证样本
Fig. 2 Testing samples



图3 彩谱 FS-15 短波近红外高光谱相机
Fig. 3 Color spectrum FS-15 short-wave near-infrared hyperspectral camera

2.3 模型建立及验证

本实验的主要目的是创建针对实验选用的9种塑料的分类模型,图4为高光谱采集系统,图5为塑料分类识别模型建立流程。

2.3.1 图像采集

图像采集在黑暗环境下进行,可避免杂散光的影响。在正式采集高光谱图像前需要对相机进行足够时间(30 min)的预热,可以有效避免基线漂移对数据采集造成影响。由于高光谱相机工作时存在暗电流,在对塑料样品进行扫描前,还需要完成相机的黑白校准以获得准确的测试数据。黑白标定公式为

$$R = \frac{I_1 - I_2}{I_3 - I_2}, \quad (1)$$

式中: R 表示校正后的图像数据; I_1 表示原始图像数据; I_2 表示99%反射率图像数据; I_3 表示0%反射率图像数据。

2.3.2 光谱提取

本实验在图像采集时存在高光谱相机视场较小的情况,需对样本进行多次采集,并将采集到的图像进行拼接。图6为需要拼接的训练样本图像,图7为6幅图

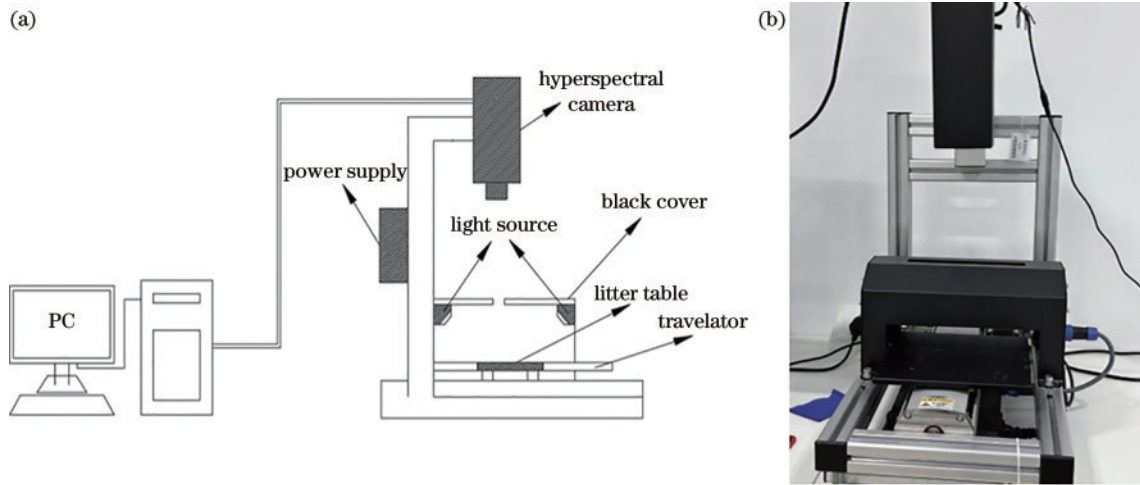


图 4 高光谱采集系统。(a)示意图;(b)实物图

Fig. 4 Hyperspectral acquisition system. (a) Schematic diagram; (b) physical diagram

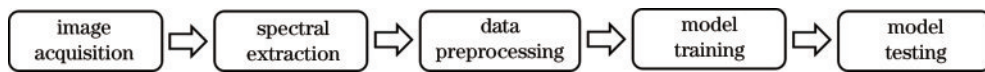


图 5 塑料分类识别模型建立流程

Fig. 5 Process of establishing plastic classification and identification model

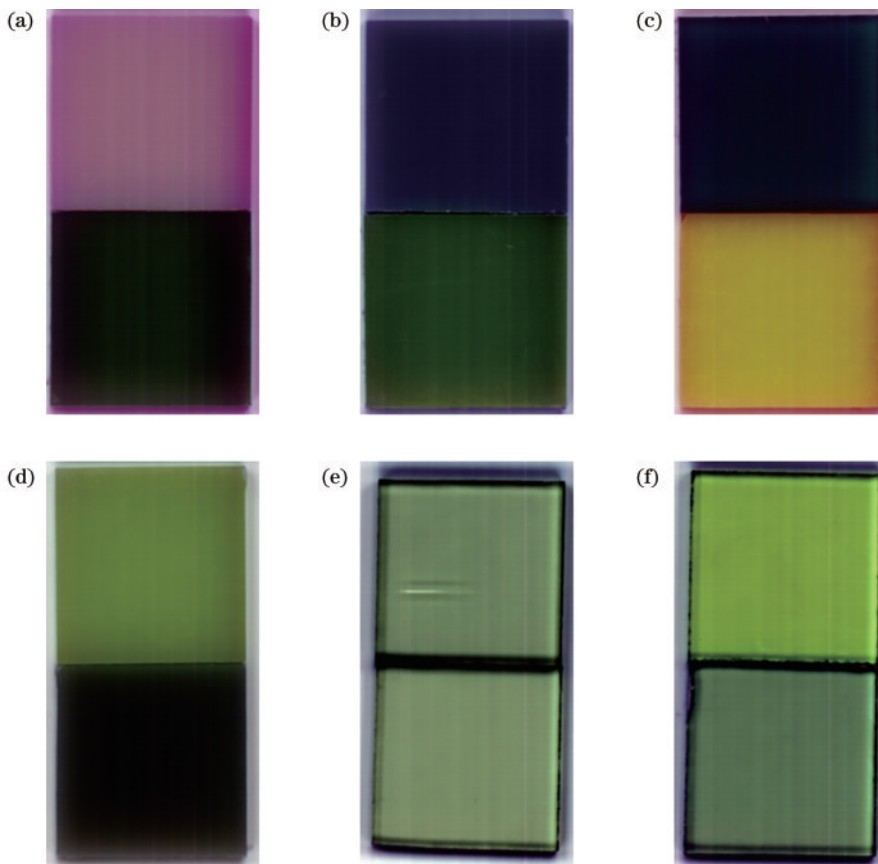


图 6 需拼接的训练集高光谱图像。(a) PC (white), PE; (b) ABS (white), PP (white); (c) PA, POM; (d) ABS (beige), PP (dark yellow); (e) PC (transparent), PS; (f) PVC, PET

Fig. 6 Hyperspectral images of the training set to be spliced. (a) PC (white), PE; (b) ABS (white), PP (white); (c) PA, POM; (d) ABS (beige), PP (dark yellow); (e) PC (transparent), PS; (f) PVC, PET

像拼接的最终效果。同理,对验证图像进行图像拼接处理,结果如图 8 所示。

高光谱相机采集到的图像为三维数据立方,包含了图像范围内所有的信息,为方便后续模型训练的展

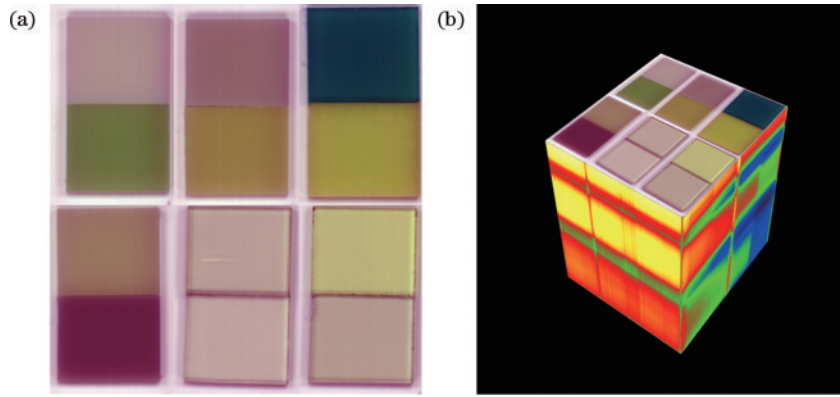


图 7 训练集高光谱图像拼接结果。(a)训练图像;(b)训练图像立方

Fig. 7 Results of hyperspectral image stitching in training set. (a) Training images; (b) training image cube

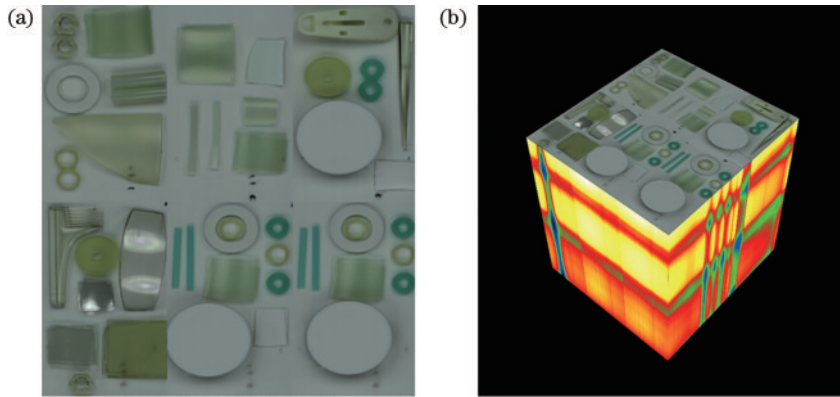


图 8 验证集高光谱图像拼接结果。(a)验证图像;(b)验证图像立方

Fig. 8 Results of hyperspectral image stitching in testing set. (a) Testing images; (b) testing image cube

开,需要先提取各类塑料样本的光谱,以便于接下来的数据处理。首先提取高光谱图像感兴趣区域(ROI),

如图 9 所示,以规则的图形选取各类塑料样本的 ROI 后,对 ROI 的数据输出进行光谱数据的预处理。

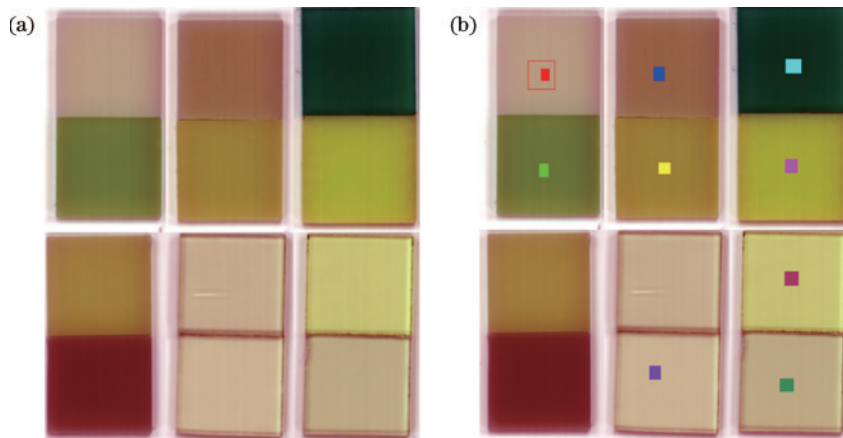


图 9 塑料的 ROI 提取示意图。(a)待提取图像;(b)ROI 提取

Fig. 9 ROI extraction of plastics. (a) Images to be extracted; (b) ROI extraction

2.3.3 数据预处理

数据预处理的目的是减小环境、噪声等因素的影响,提高模型训练的精度,处理方法多种多样,本实验选择的预处理方式如下:

- 1) 对光谱数据进行裁剪;
- 2) 对导入的数据进行中值滤波;

- 3) 对滤波后的数据进行多项式平滑处理(SG 平滑);
- 4) 对平滑后的数据进行标准正态变化(SNV);
- 5) 对数据进行归一化。

通过观察光谱数据后发现该实验测得的数据,在 900~1100 nm 波段的光谱数据对于颜色依旧存在敏

感性, 1650~1700 nm 波段的光谱一定程度上受到噪声影响, 故裁剪波段为 1100~1650 nm, 保留波段数 171 个。设置中值滤波器配合 SG 平滑, 以在保留光谱原始特征的情况下最大程度消除噪声。SNV 处理的

目的为消除样本表面散射以及光程变换对近红外波段反射光谱的影响。最后对数据集进行归一化, 让反射率数据以区间 [0, 1] 来表达各光谱曲线。图 10 为裁剪完成的光谱数据预处理前后的对比。

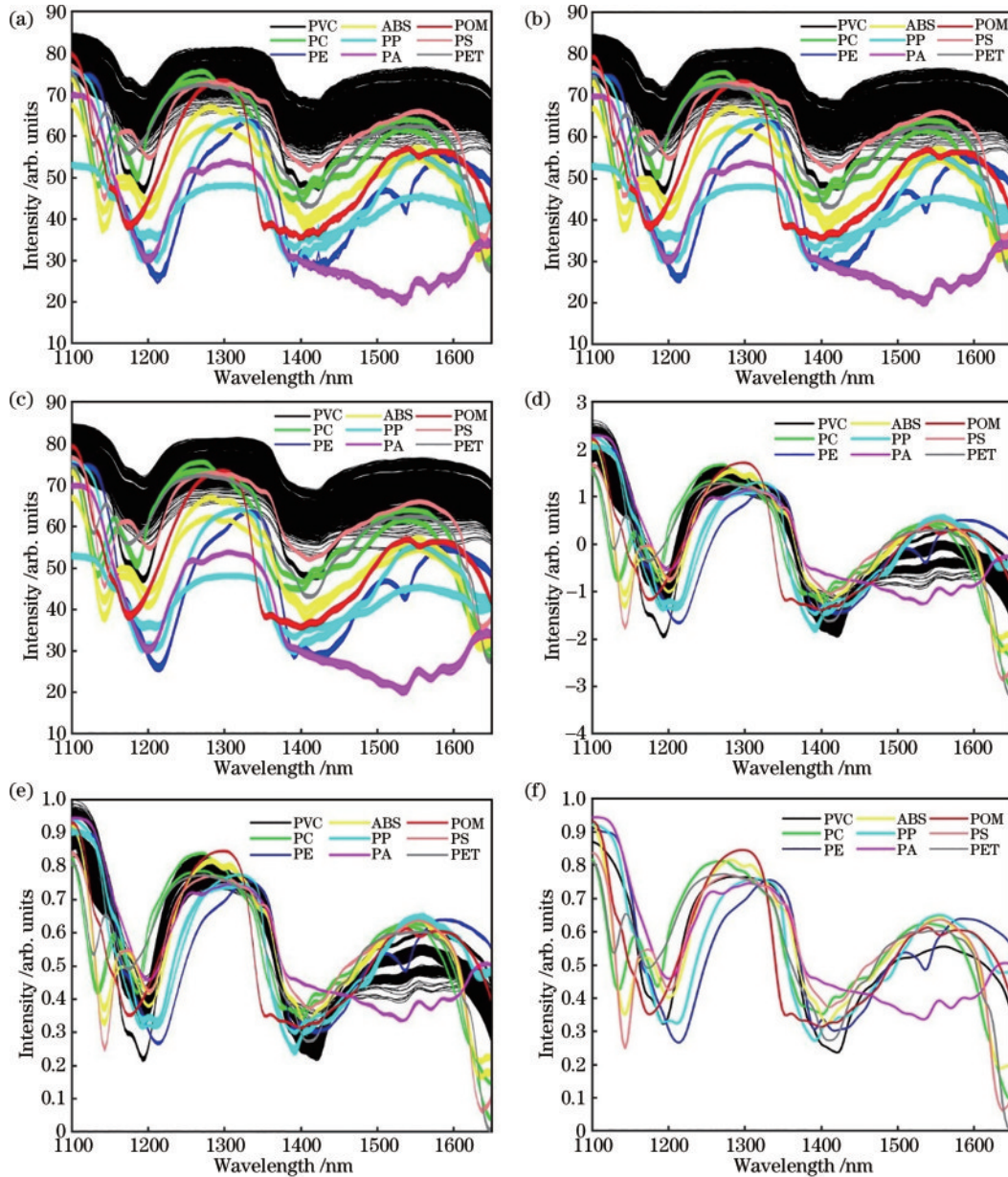


图 10 预处理前后的光谱曲线。(a)原始光谱;(b)中值滤波;(c)中值滤波-SG平滑;(d)中值滤波-SG平滑-SNV;(e)中值滤波-SG平滑-SNV-归一化;(f)平均光谱

Fig. 10 Spectral curves before and after preprocessing. (a) Raw spectra; (b) median filtering; (c) median filtering-SG; (d) median filtering-SG-SNV; (e) median filtering-SG-SNV-normalized; (f) average spectra

2.3.4 机器学习

本实验比较了多种机器学习算法在塑料分类模型上的效果。

2.3.4.1 K-NN算法

K-近邻(K-NN)算法^[16]是一种基于实例的、非参数的监督学习算法,其基本原理如下:

1) 邻居选择:对于一个给定的未知样本,K-NN算法首先找到训练集中与该样本最接近的K个邻居,其

中,K是一个预先设定的正整数。

2) 距离度量:K-NN可以使用不同的距离度量来衡量样本之间的相似性。

3) 分类问题:对于分类问题,K-NN采取多数表决的策略,即将K个最近邻样本中出现次数最多的类别标签赋给未知样本。

2.3.4.2 SVM算法

支持向量机(SVM)^[17-18]是一种强大的监督学习算

法。SVM 通过核函数来处理非线性问题, 以下是 SVM 算法的基本原理:

1) 线性可分情况

假设我们有一个二分类问题, 希望将两类样本分开。如果两类样本可以被一个超平面完全分开, 那么我们称这种情况为线性可分。在二维空间中, 超平面就是一条直线; 在三维空间中, 它是一个平面; 在更高维度的空间中, 它是一个超平面。SVM 的目标是找到一个最优的超平面, 使得训练样本点到超平面的距离(间隔)最大化。支持向量是离超平面最近的训练样本点, 它们决定了超平面的位置。

2) 软间隔分类

在实际问题中, 很多时候数据并不是完全线性可分的, 即有一些噪声或异常值。为了处理这种情况, 引入了“软间隔分类”。

软间隔分类允许一些样本分类错误, 但会尽量使得间隔最大化。可以使用一个正则化参数(通常记为 C)来平衡间隔的大小和错误分类的惩罚。

3) 核函数

SVM 通过核函数来处理非线性问题。核函数允许将样本从原始特征空间映射到一个更高维的特征空间, 使得原始样本在新的特征空间中可以被线性分隔。常用的核函数包括线性核、多项式核、高斯核(径向基函数核)等。

2.3.4.3 PSO-SVM 算法

粒子群优化(PSO)是一种基于群体智能的优化算法, 灵感来自鸟群或鱼群等群体的行为。每个个体被称为粒子, 它们在解空间中移动, 并根据自身经验和群体的经验来搜索最优解。

PSO-SVM 将粒子群优化与支持向量机相结合, 通过优化 SVM 的超参数(如惩罚参数 C 、核函数参数等)来寻找最优的超平面。主要步骤包括: 初始化粒子群; 计算适应度; 更新个体最优; 更新全局最优; 更新粒子速度和位置; 迭代并获取最优解

2.3.4.4 GA-SVM 算法

遗传算法(GA)是一种基于自然选择和遗传机制的优化算法, 通过模拟自然界中的进化过程来搜索最优解。它涉及种群、基因编码、交叉、变异等概念。

GA-SVM 将遗传算法与支持向量机相结合, 通过优化 SVM 的超参数(如惩罚参数 C 、核函数参数等)来寻找最优的超平面。主要步骤包括: 初始化种群; 计算适应度; 选择; 交叉; 变异; 更新种群; 迭代并获取最优解。

2.3.5 模型验证

本实验在模型验证上分成两个步骤进行: 1) 导入验证数据集和训练模型进行分类预测, 计算分类精度, 根据分类精度对模型进行优化和筛选; 2) 导入达到精度要求的模型和验证图像, 对整幅图像进行分类预测, 将分类得到的结果用颜色编码显示, 并与原始图像进行对比, 得到最终的分类效果。具体流程如图 11 所示。

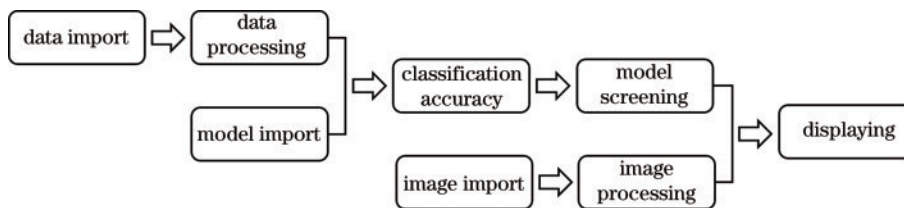


图 11 分类模型验证流程示意图

Fig. 11 Schematic diagram of classification model testing process

3 数据分析

3.1 基于多种距离度量的 K-NN 算法的分类效果

本实验采用 K-NN 算法进行模型训练, 选择了 6 种具有较好相似度区分的距离度量方法, 分别为欧式距离(Euclidean)、城市街区距离(City block)、切比雪夫距离(Chebyshev)、余弦相似度(Cosine)、光谱角相似度(SAM)、光谱信息散度(SID)。通过多次训练, 得到 6 种不同度量方法的分类识别模型。为验证模型的初步准确度, 导入验证数据集和相应的类别标签, 并计算模型的准确度。这种综合采用多种距离度量方法的模型设计, 有助于在不同情境下更全面地评估相似度, 提高分类模型的鲁棒性和适用性。

分类精度如表 2 所示, 不同距离度量在训练模型时都有 100% 的区分度, 当样本出现多样性后, City

表 2 多种距离度量的 K-NN 算法的分类精度统计

Table 2 Classification accuracy statistics of K-NN algorithm for various distance measures

Model type	Accuracy rate of training set / %	Accuracy rate of testing set / %
K-NN (Euclidean)	100	96.14
K-NN (City block)	100	23.46
K-NN (Chebyshev)	100	18.80
K-NN (Cosine)	100	96.21
K-NN (SAM)	100	93.69
K-NN (SID)	100	74.32

block、Chebyshev 两种距离算法明显不适用, SID 的准确率也偏低, 只有 74.32%。其余模型都达到了 90% 以上的准确度。其中, Euclidean 和 Cosine 距离算法有着最好的效果, 达到了 96.14% 和 96.21% 的验证准确率。

通过生成预测分类对照图和混淆矩阵查看具体分类情况,直观反映分类错误。从图 12 可以发现,错误集中在 PC 和 PET 两个材料上。其中,SAM 在 PC 材料的分类准确率只有 25.8%。其余材料都具有较好的分类效果。为此将验证图像完整导入后再查看分类

情况,观察分类问题出现的区域。

将验证图像导入,通过颜色编码显示分类效果,如图 13 所示。分类错误集中在弧形的透明 PC 片上。整体上 Euclidean、Cosine 的分类效果较好,SAM 在测试样本的轮廓区域出现的分类错误较多。分析可知,反

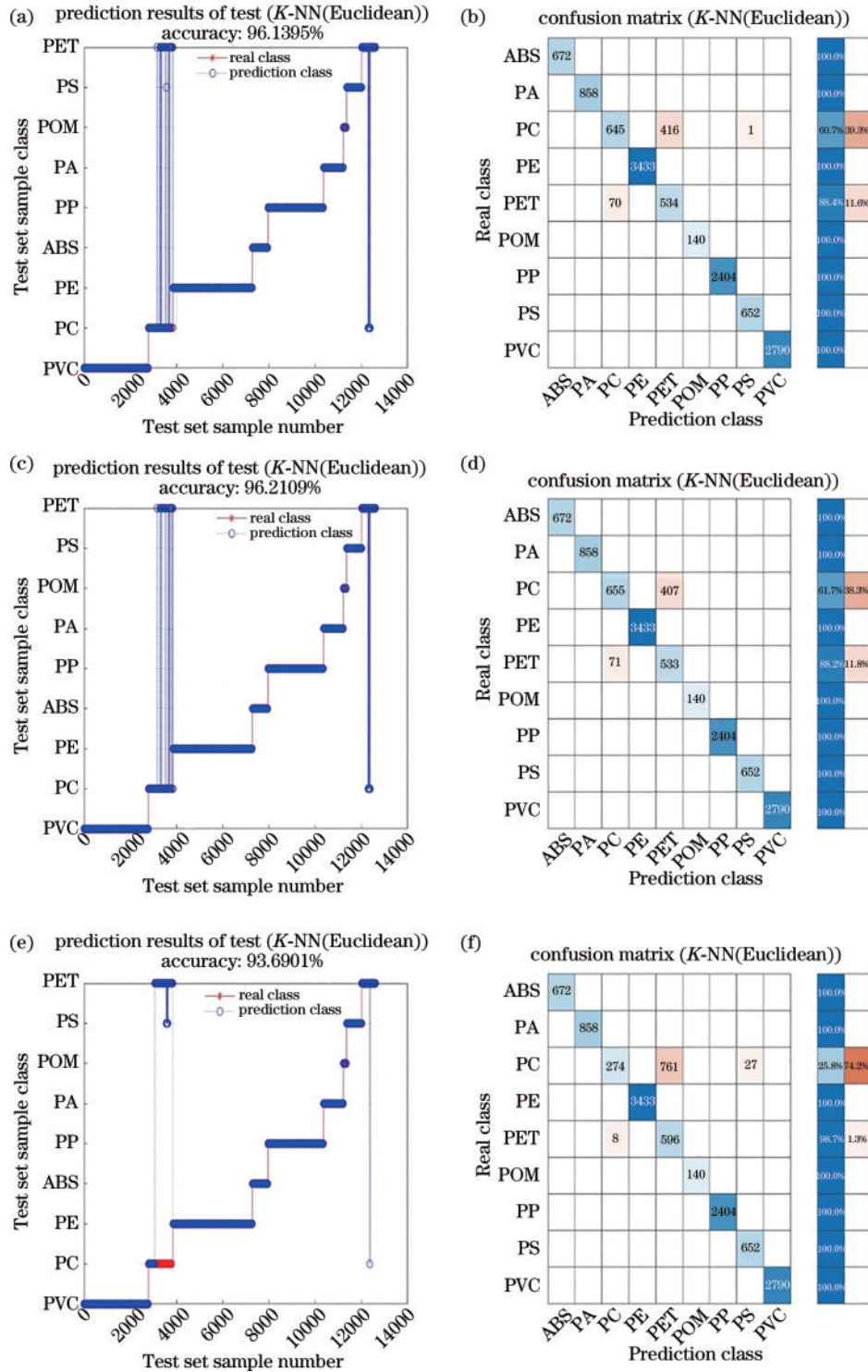


图 12 基于 Euclidean、Cosine、SAM 的 K-NN 的分类精度。(a) K-NN (Euclidean) 预测结果;(b) K-NN (Euclidean) 混淆矩阵;(c) K-NN (Cosine) 预测结果;(d) K-NN (Cosine) 混淆矩阵;(e) K-NN (SAM) 预测结果;(f) K-NN (SAM) 混淆矩阵
 Fig. 12 Classification accuracy of K-NN based on Euclidean, Cosine, and SAM. (a) K-NN (Euclidean) prediction results; (b) K-NN (Euclidean) confusion matrix; (c) K-NN (Cosine) prediction results; (d) K-NN (Cosine) confusion matrix; (e) K-NN (SAM) prediction results; (f) K-NN (SAM) confusion matrix

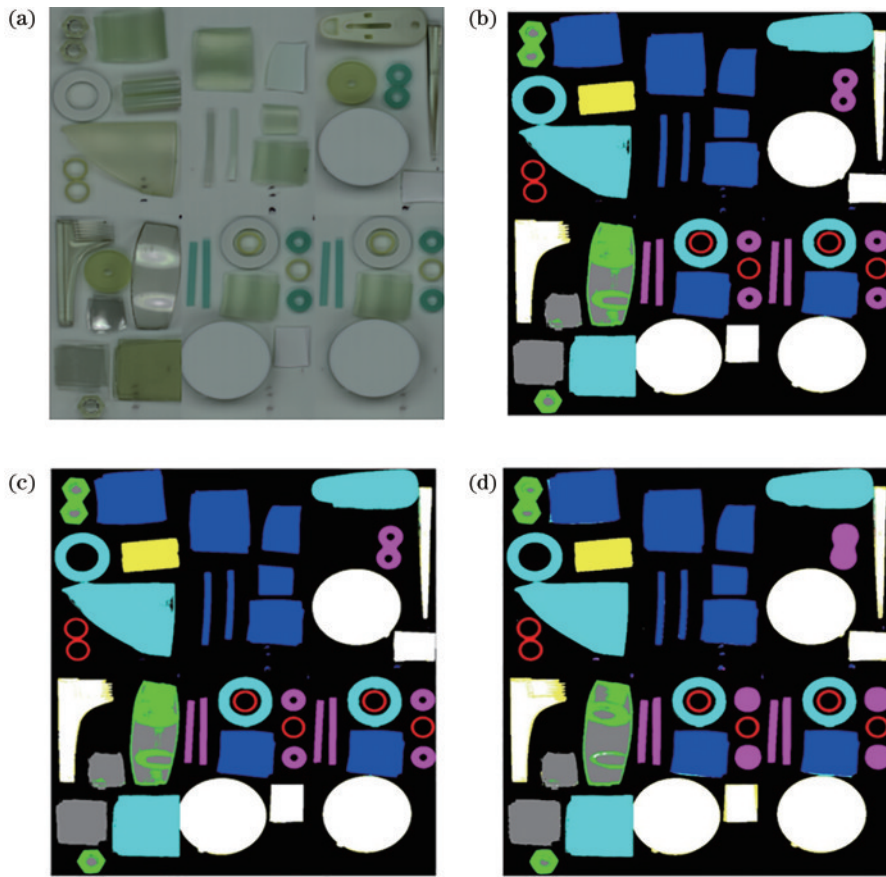


图 13 基于 Euclidean、Cosine、SAM 的 K-NN 的分类效果对比。(a) 验证图像；(b) K-NN (Euclidean)；(c) K-NN (Cosine)；(d) K-NN (SAM)

Fig. 13 Comparison of classification effects of K-NN based on Euclidean, Cosine and SAM. (a) Testing images; (b) K-NN (Euclidean); (c) K-NN (Cosine); (d) K-NN (SAM)

光现象对分类结果会造成严重影响,阴影和散射光等可通过预处理算法配合合适的机器学习方式消除。

3.2 基于 SVM 算法、PSO-SVM 算法、GA-SVM 算法的分类效果

本实验在使用 SVM 训练模型时,通过粒子群优化算法和遗传优化算法结合 SVM,优化了 SVM 算法参数,从而获取最优解。分类精度如表 3 所示。

表 3 SVM 算法、PSO-SVM 算法、GA-SVM 算法分类精度统计
Table 3 Classification accuracy statistics of SVM algorithm, PSO-SVM algorithm, and GA-SVM algorithm

Model type	Accuracy rate of training set / %	Accuracy rate of testing set / %
SVM-Liner	100	96.01
PSO-SVM-Liner	100	96.42
GA-SVM-Liner	100	98.67

通过绘制预测结果比对图和混淆矩阵可以直观反映分类情况,从图 14 可以发现,这 3 种 SVM 算法都具有较高的准确度,且 PSO-SVM 算法和 GA-SVM 算法通过优化参数提高了分类准确度。

同 K-NN 模型验证方法一样,为了验证模型分类效果,将验证图像完全导入,以颜色编码显示最终分

类情况,如图 15 所示。比较发现,SVM 和 PSO-SVM 的分类错误多,PSO-SVM、GA-SVM 对部分环状样本识别错误。可能原因是样本本身体积小,中间的环口小,容易受自身散射光、阴影等因素的影响。对比 3 种模型的样本轮廓可知,GA-SVM 的识别效果最佳,显示轮廓清晰。

3.3 复杂图像模型分类效果

通过上述数据分析可以得到,K-NN (Euclidean)、K-NN (Cosine) 及 GA-SVM-Liner 的分类效果最佳。为进一步查看模型分类效果,导入相对复杂的高光谱图像。分类效果如图 16 和图 17 所示。

观察可视化效果可以得到结论:图 16 的样本摆放无序杂乱,但样本数量较少,在这种环境下 K-NN (Euclidean)、K-NN (Cosine)、GA-SVM 的分类效果依旧出色,其中,GA-SVM 在样本轮廓的识别上效果更好;图 17 材料以堆叠形式出现,所有分类模型都出现了不同程度的分类错误,主要错误集中在塑料片上的两条竖直摆放的 PE 矩形条,GA-SVM-Liner 在分类上也出现了较多错误,集中在 PET 材料和样本阴影上。

通过分析这两幅图像的分类效果,可以得到以下结论:单个视场下,少量样本可以有效避免散射光和

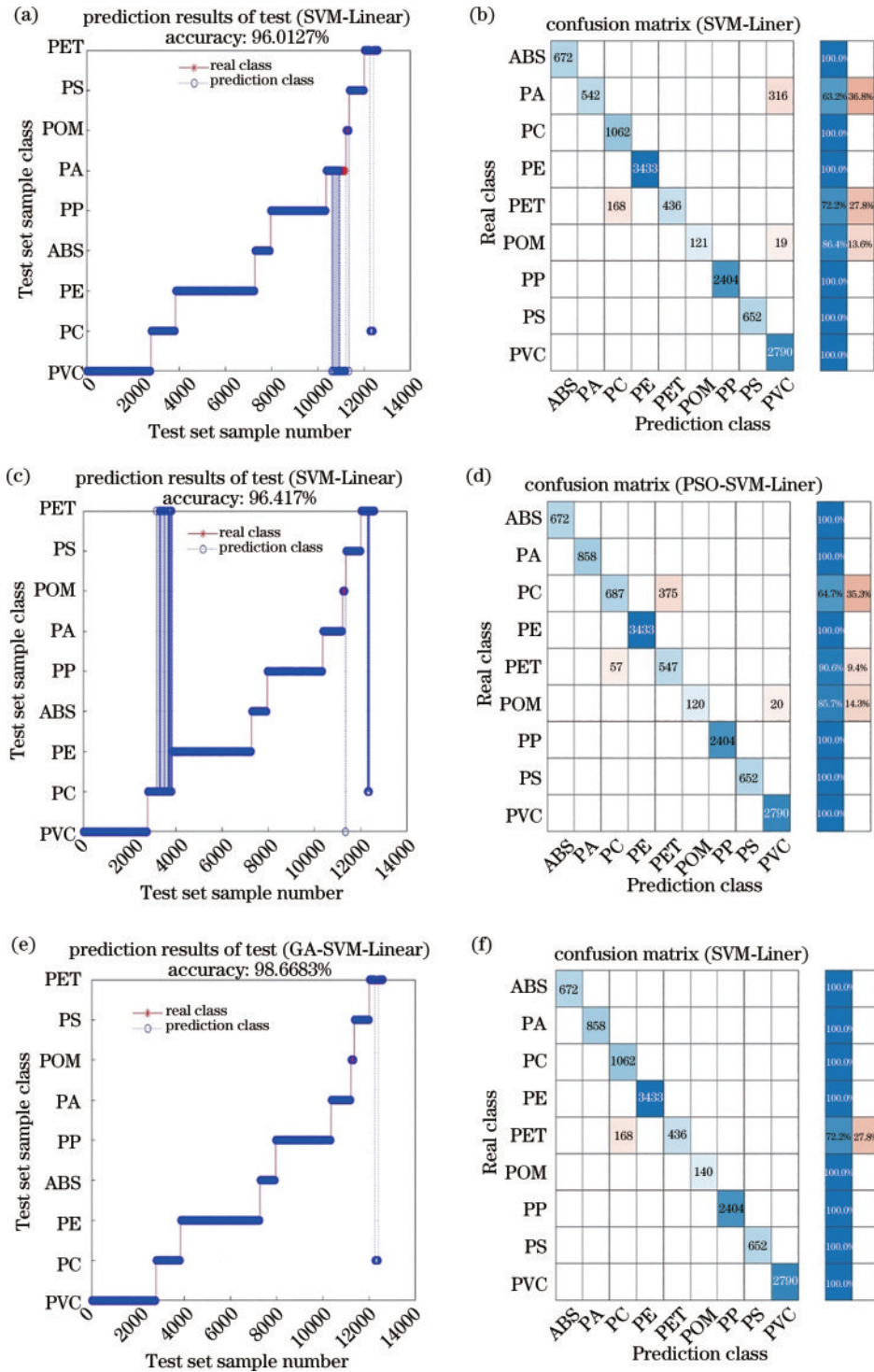


图 14 SVM、PSO-SVM、GA-SVM 分类精度。(a) SVM-Liner 预测结果；(b) SVM-Liner 混淆矩阵；(c) PSO-SVM-Liner 预测结果；(d) PSO-SVM-Liner 混淆矩阵；(e) GA-SVM-Liner 预测结果；(f) GA-SVM-Liner 混淆矩阵

Fig. 14 SVM, PSO-SVM, and GA-SVM classification accuracy. (a) SVM-Liner prediction results; (b) SVM-Liner confusion matrix; (c) PSO-SVM-Liner prediction results; (d) PSO-SVM-Liner confusion matrix; (e) GA-SVM-Liner prediction results; (f) GA-SVM-Liner confusion matrix

阴影影响,处于视场中心区域的样本也可以有效避免视场边缘光照不均匀带来的反光现象。比较 6 类机器学习方法可知,GA-SVM 在样本边缘识别上有绝对的优势,K-NN (Euclidean)、K-NN (Cosine)可以有效减小光照不均带来的影响。

3.4 总结与分析

本实验得到了 6 种高验证准确率模型,均达到了 90% 以上的准确率。对比 6 种模型,将数据汇总分析,汇总内容为模型的验证集准确率和验证集中出现的错误种类,以及该类错误率,具体见表 4。

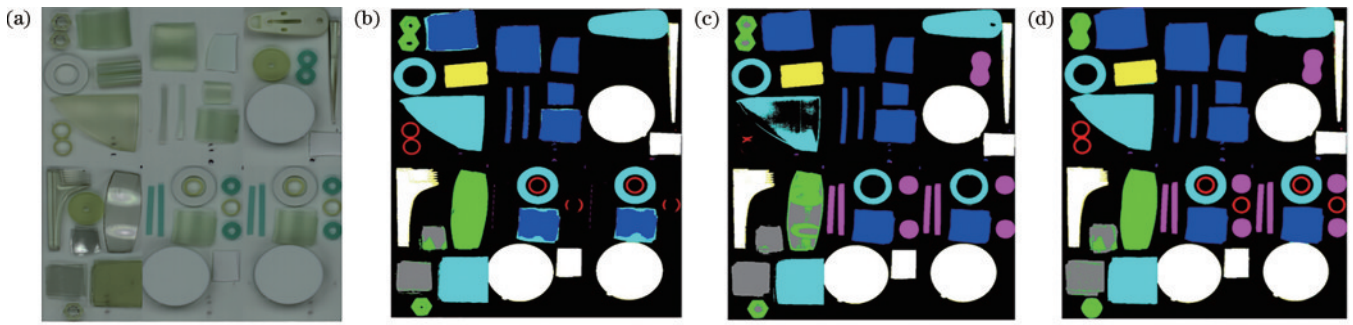


图 15 SVM、PSO-SVM、GA-SVM 模型的分分类效果对比。(a)验证图像;(b) SVM-Liner;(c) PSO-SVM-Liner;(d) GA-SVM-Liner
Fig. 15 Comparison of classification effects of SVM, PSO-SVM, and GA-SVM models. (a) Testing images; (b) SVM-Liner; (c) PSO-SVM-Liner; (d) GA-SVM-Liner

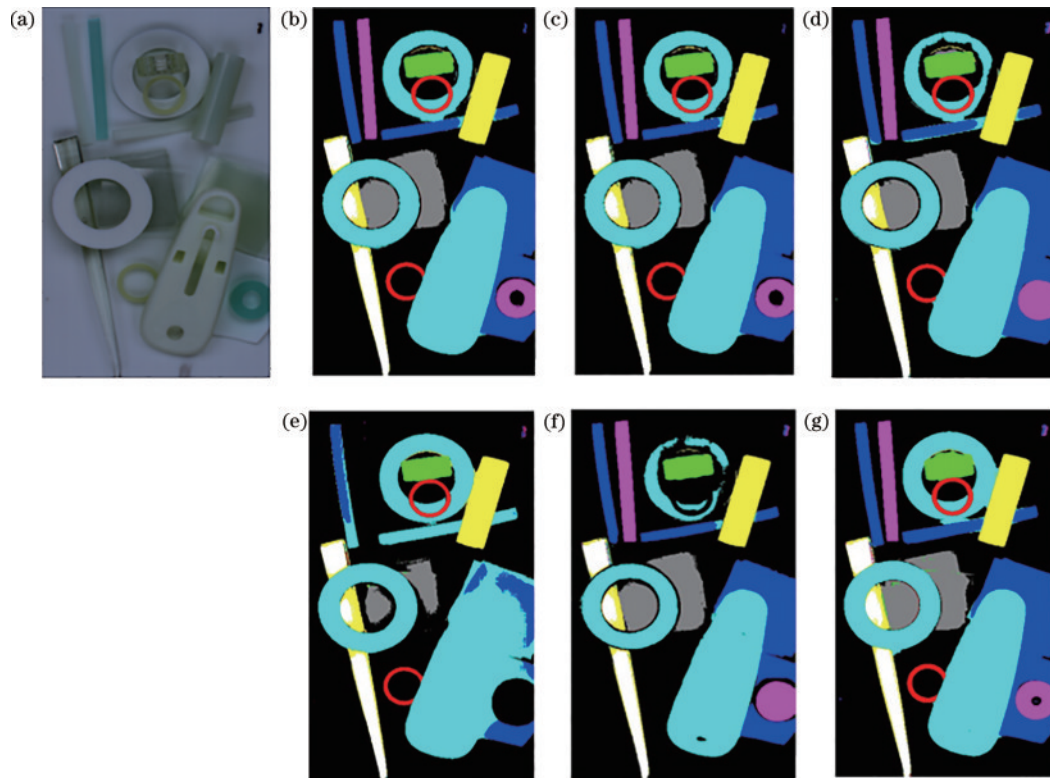


图 16 3 种模型的分分类效果对比。(a)验证图像;(b)K-NN (Euclidean); (c)K-NN (Cosine); (d) K-NN (SAM); (e) SVM-Liner;
(f) PSO-SVM-Liner; (g) GA-SVM-Liner
Fig. 16 Comparison of classification effects of three models. (a) Testing images; (b) K-NN (Euclidean); (c) K-NN (Cosine); (d) K-NN (SAM); (e) SVM-Liner; (f) PSO-SVM-Liner; (g) GA-SVM-Liner

表 4 塑料分选模型精度汇总
Table 4 Summary of plastic sorting models

Model type	Accuracy rate / %	Error type & error rate
K-NN(Euclidean)	96.14	PC(39.3%),PET(11.6%)
K-NN(Cosine)	96.21	PC(38.3%),PET(11.8%)
K-NN(SAM)	93.69	PC(74.2%),PET(1.3%)
SVM-Liner	96.01	PA(36.8%),PET(27.8%),POM(13.6%)
PSO-SVM-Liner	96.42	PC(35.3%),PET(9.4%),POM(14.3%)
GA-SVM-Liner	98.67	PET(27.8%)

本实验对 3 幅图像进行了分类可视化操作,一张规整摆放的拼接图像,两张堆叠摆放的复杂图像,并根据分类可视化的准确度和精细度,给出效果好、效

果一般、效果差等 3 种评价,6 种模型的可视化效果见表 5。可以看出:GA-SVM-Liner 的分类精度最高,可视化效果最好;K-NN(Euclidean)、K-NN(Cosine)

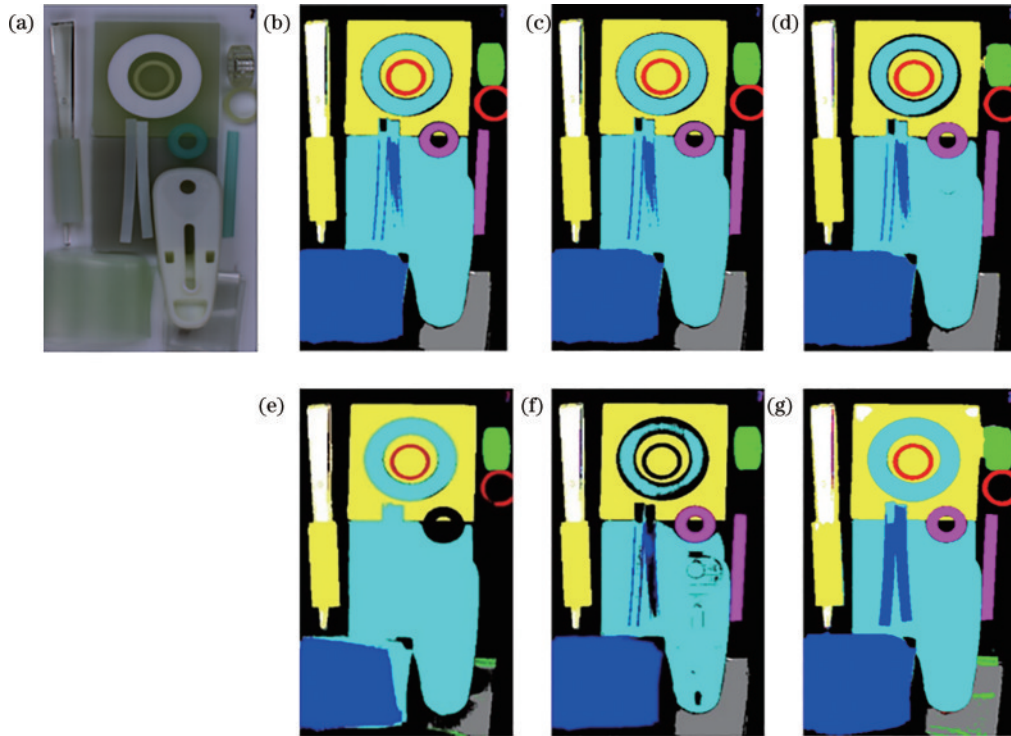


图 17 3 种模型分类效果对比。(a)验证图像;(b)K-NN (Euclidean);(c)K-NN (Cosine);(d)K-NN (SAM);(e)SVM-Liner;(f)PSO-SVM-Liner;(g)GA-SVM-Liner

Fig. 17 Comparison of classification effects of the three models. (a) Testing images; (b) K-NN (Euclidean); (c) K-NN (Cosine); (d) K-NN (SAM); (e) SVM-Liner; (f) PSO-SVM-Liner; (g) GA-SVM-Liner

表 5 塑料分选模型可视化效果汇总

Table 5 Summary of plastic sorting model visualization effects

Model type	Visualization (regular)	Visualization (complex 1)	Visualization (complex 2)
K-NN(Euclidean)	Effective	Effective	Generally
K-NN(Cosine)	Effective	Effective	Generally
K-NN(SAM)	Effective	Ineffective	Ineffective
SVM-Liner	Ineffective	Ineffective	Ineffective
PSO-SVM-Liner	Ineffective	Ineffective	Ineffective
GA-SVM-Liner	Effective	Effective	Effective

的分类精度较高,整体可视化效果较好;K-NN (SAM)、SVM-Liner、PSO-SVM-Liner的分类精度高但可视化效果欠佳。

4 结 论

本实验采用近红外高光谱成像技术,利用塑料方片获取光谱信息,并成功建立了多个模型。在同一实验环境下对各种塑料制品进行测试,其中,GA-SVM 的分类准确率达到了 98.67%。这表明近红外高光谱成像技术在塑料分选领域具有显著的实用价值和广阔的发展前景。这对于实现环保和可持续发展目标具有重要意义,因为它有助于提高塑料回收的效率,减少对自然资源的需求,并减小塑料垃圾对环境造成的负面影响。

然而,实验中也存在一些未完全解决的问题,本实

验中光谱特性相近的两种材料 PET 和 PC,由于反光、不平整、高度差等问题,分类错误率较高。此外,空心环状样本的空心区域也容易受到阴影的影响,导致样本形状识别错误。

为了提高实用性和鲁棒性,将高光谱成像技术真正引入工业分选,这些问题需要在后续的研究中得到解决,后续可以通过改进高光谱系统的光源光照结构,优化样本散射光、阴影处理、噪声分离等算法,也可以通过优化采集系统,让样本在采集过程中集中在视场中间区域,从而提高图像质量。

参 考 文 献

- [1] OECD. Global plastics outlook: economic drivers, environmental impacts and policy options[R]. Paris: OCED Publishing, 2022.
- [2] 陈飞扬, 陈万锁. 废塑料的分类分选、预处理及回收现状[J]. 中国资源综合利用, 2021, 39(10): 118-134.
Chen F Y, Chen W S. Classification, pretreatment and recycling status of waste plastics[J]. China Resources Comprehensive Utilization, 2021, 39(10): 118-134.
- [3] 苑宁之, 陈少华, 牟涛涛. 拉曼光谱结合深度学习算法的塑料分类的研究[J]. 光学仪器, 2023, 45(5): 35-43.
Yuan N Z, Chen S H, Mu T T. Research on classification of plastics by Raman spectroscopy combined with deep learning algorithm[J]. Optical Instruments, 2023, 45(5): 35-43.
- [4] Wang X, Zhao N J, Yin G F, et al. Classification and

- identification of plastic with laser-induced fluorescence spectroscopy based on back propagation neural network model[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(10): 3136-3141.
- 王翔, 赵南京, 殷高方, 等. 基于反向传播神经网络的激光诱导荧光光谱塑料分类识别方法研究[J]. *光谱学与光谱分析*, 2019, 39(10): 3136-3141.
- [5] Osoro B O, Ndegwa R, Ombati W N, et al. Identification of microplastics in plastic bottled drinking water using laser-induced breakdown spectroscopy (LIBS) and Raman spectroscopy[J]. *Asian Journal of Research and Reviews in Physics*, 2023, 7(2): 15-24.
- [6] 张彬, 孙岩森, 李晓杰, 等. 高光谱目标分类技术与展望[J]. *红外*, 2023, 44(8): 1-12.
- Zhang B, Sun Y S, Li X J, et al. Development and prospect of hyperspectral target classification technology [J]. *Infrared*, 2023, 44(8): 1-12.
- [7] Zhang L, Wang Y Q, Wei Y G, et al. Near-infrared hyperspectral imaging technology combined with deep convolutional generative adversarial network to predict oil content of single maize kernel[J]. *Food Chemistry*, 2022, 370: 131047.
- [8] Fabiyi S D, Vu H, Tachtatzis C, et al. Varietal classification of rice seeds using RGB and hyperspectral images[J]. *IEEE Access*, 2020, 8: 22493-22505.
- [9] 杨杰锴, 郭志强, 黄远. 高光谱成像在蔬菜性状无损检测中的研究进展[J]. *激光与光电子学进展*, 2022, 59(10): 1000002.
- Yang J K, Guo Z Q, Huang Y. Research progress of hyperspectral imaging in nondestructive testing of vegetable traits[J]. *Laser & Optoelectronics Progress*, 2022, 59(10): 1000002.
- [10] Wei L F, Wang Z, Huang C, et al. Transparency estimation of narrow rivers by UAV-borne hyperspectral remote sensing imagery[J]. *IEEE Access*, 2020, 8: 168137-168153.
- [11] 李武耀, 买买提·沙吾提, 买合木提·巴拉提. 基于分数阶微分的土壤有机质含量高光谱反演研究[J]. *激光与光电子学进展*, 2023, 60(7): 0730005.
- Li W Y, Mamat S, Maihemuti B. Fractional differential-based hyperspectral inversion of soil organic matter content[J]. *Laser & Optoelectronics Progress*, 2023, 60(7): 0730005.
- [12] Liew S C, Choo C K, Lau J W M, et al. Monitoring water quality in Singapore Reservoirs with hyperspectral remote sensing technology[J]. *Water Practice and Technology*, 2019, 14(1): 118-125.
- [13] Nageswaran K, Nagarajan K, Bandiya R. A novel algorithm for hyperspectral image denoising in medical application[J]. *Journal of Medical Systems*, 2019, 43(9): 291.
- [14] Zheng Y, Bai J R, Xu J N, et al. A discrimination model in waste plastics sorting using NIR hyperspectral imaging system[J]. *Waste Management*, 2018, 72: 87-98.
- [15] Henriksen M L, Karlsen C B, Klarskov P, et al. Plastic classification via in-line hyperspectral camera analysis and unsupervised machine learning[J]. *Vibrational Spectroscopy*, 2022, 118: 103329.
- [16] Wei X L, Li W, Zhang M M, et al. Medical hyperspectral image classification based on end-to-end fusion deep neural network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, 68(11): 4481-4492.
- [17] Blanzieri E, Melgani F. Nearest neighbor classification of remote sensing images with the maximal margin principle [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(6): 1804-1811.
- [18] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778-1790.