

引用格式: XU Cheng, LI Fang, CHEN Feng, et al. Rapid Classification of Laser Induced Breakdown Spectroscopy of Titanium Alloys[J]. Acta Photonica Sinica, 2022, 51(4):0430001

许铖,李芳,陈锋,等. 钛合金的激光诱导击穿光谱快速分类[J]. 光子学报,2022,51(4):0430001

钛合金的激光诱导击穿光谱快速分类

许铖¹,李芳¹,陈锋²,张登²,邓凡³,郭连波³

(1 武汉工程大学 机电工程学院 光学信息与模式识别湖北省重点实验室,武汉 430205)

(2 华中科技大学 武汉光电国家研究中心,武汉 430074)

(3 华中科技大学 光学与电子信息学院,武汉 430074)

摘要:利用激光诱导击穿光谱技术快捷、实时、原位、微损、多元素同步分析的优势,对相同牌号下不同国标编号的钛合金进行分析,以实现钛合金快速精确的分类识别。用激光诱导击穿光谱技术采集钛合金于各激光强度与延时下的光谱,通过对比 6 个特征处的峰值强度与信噪比获得最优条件。结合 K 最邻近算法,对最优条件下采集的 TC4 钛合金数据进行处理,并优化其模型参数,从而实现同牌号钛合金的分类识别研究。结果表明,采用基于数据处理与模型优化相结合的方法,可以将相同牌号钛合金的分类准确率从 84.15% 提升至 99.14%,同时训练时间从 1 232.41 s 降低至 83.91 s,分类性能获得显著提升。研究成果有望实现对相同牌号钛合金的快速精准分类识别,具有广泛的应用前景。

关键词:光谱学;钛合金分类;激光诱导击穿光谱;K 最邻近算法;TC4 钛合金

中图分类号:O433.4

文献标识码:A

doi:10.3788/gzxb20225104.0430001

0 引言

钛及钛合金是 20 世纪 50 年代发展起来的一种密度小、强度高、耐蚀耐热性好的金属材料,因其独特的优点,近年来被广泛用于航空航天、医疗卫生、交通运输和高端制造等重要领域^[1-4]。不同种类的钛合金由于性能存在差异适用于不同领域,即使是同牌号的钛合金在混淆使用时也易造成严重后果。例如 TC4(Ti-4Al-4V,钛-6 铝-4 钒)钛合金,自 20 世纪中期开始作为医疗植入物广泛应用于用医疗领域,而当其中 Al 元素含量过高时,会对人体造成不可逆转的损害。在国家一级标准物质的分类中,TC4 钛合金按照元素含量被分为 GBW02503~GBW02508 六种编号不同的种类,故用于医疗领域的钛合金通常是 Al 含量较低的 TC4-1(GBW02503,Al 元素含量 3.90%)或 TC4-2(GBW02504,Al 元素含量 4.67%),而非 Al 含量较高的 TC4-5(GBW02507,Al 元素含量 6.78%)或 TC4-6(GBW02508,Al 元素含量 7.03%)。因此,对外表十分相似而性能存在差异的钛合金进行快速精准分类识别具有十分重要的研究和应用价值。

目前,钛合金的分类检测方法种类繁多,但存在着各自的局限。现有检测方法主要分为消解法和非消解法两类。消解法包括滴定法、原子吸收光谱法等,在测试过程中需要对样品进行破坏性处理,无法做到微损检测。而非消解法如 X 射线荧光光谱分析、火花光谱直读法等,则对环境要求十分苛刻(如需要真空环境等),且部分检测方法的激发源存在辐射,对人体存在损害。因此,探寻一种新型的钛合金快速分类识别检测方法成为当前的研究热点。

激光诱导击穿光谱^[5-7](Laser-induced Breakdown Spectroscopy, LIBS)技术,又称激光探针技术,是一种新型的基于原子发射光谱的元素分析技术。其原理是利用激光烧蚀样品产生等离子体,然后采集等离子体的光谱用以定性定量分析。LIBS 技术具有快捷、实时、原位、微损、多元素同时分析等优点,可以检测任何

基金项目:国家自然科学基金(No.62075069)

第一作者:许铖(1996-),男,硕士研究生,主要研究方向为基于激光诱导击穿光谱的有色金属快速分类。Email: xvcheng0612@qq.com

导师(通讯作者):李芳(1982-),女,教授,博士,主要研究方向为超快纳米光子学和激光先进制造技术。Email: lifang@wit.edu.cn

郭连波(1980-),男,教授,博士,主要研究方向为多模态智能激光感知技术。Email: lbguo@hust.edu.cn

收稿日期:2021-11-09;录用日期:2022-01-25

<http://www.photon.ac.cn>

物态(气态^[8]、固态、液态^[9])的样品,近年来广泛应用于冶金工业成分分析^[10]、环境监测与保护^[11]、农业与食品安全检测^[12-13]等领域。2011年美国发射的好奇号火星探测器和2020年中国发射的天问一号火星探测器均装备了LIBS系统对火星执行勘探任务。

LIBS技术的独特优势使其获得了国内外研究人员的关注^[14-16]。SUNGHO S等^[17]利用LIBS系统收集铝合金、铜合金、不锈钢和铸钢的光谱,之后对全光谱数据进行主成分分析(Principal Component Analysis, PCA),结果中除对铜合金的最高分类结果为94.7%外,其余均达到100%的分类准确度;ABERKANE S M等^[18]将LIBS与K最邻近(K-Nearest Neighbor, KNN)、支持向量机(Support Vector Machines, SVM)以及人工神经网络(Artificial Neural Network, ANN)结合,对锌基合金进行快速精确分类研究,均获得优秀的结果;OLLILA A M等^[19]在对各种地质样本进行分类时,使用LIBS系统结合偏最小二乘判别分析(Partial Least Squares Discrimination Analysis, PLS-DA),最终结果拥有接近100%的分类准确度;刘凯等^[20]对4种塑料样品识别时,利用PCA对通过LIBS获取的数据进行处理,精确地识别出聚对苯二甲酸乙二醇酯(Polyethylene Terephthalate, PET),但是对另外三种材料稍有混淆;郭连波等^[21]利用LIBS技术测量钢铁中钒、钛元素的含量,运用基本定标法和内定标法建立了钢铁样品中V和Ti含量的光谱分析定标曲线,采用效果更好的内定标法时,V和Ti的拟合相关系数 R^2 分别达到0.9952和0.9921,对检测样品中V和Ti元素的测定相对误差均可降低到4.0%以下;周中寒等^[22]利用基于光纤激光器的LIBS系统(Fiber-LIBS)对6种不同牌号的铝合金进行分类研究,结果显示经过SVM训练后的分类模型可以使得分类准确度达到99.83%。

通过前期文献调研可以发现,研究者们对于合金分类的研究主要集中于不同种类、不同牌号的合金,对同牌号钛合金的LIBS分类研究鲜见报道,因此本文采用KNN算法结合LIBS技术,实现对同牌号钛合金的快速精准分类。

1 实验

1.1 实验仪器

使用的LIBS装置的示意图如图1所示。样品被放置在Y110TA150电控位移平台上,每次采集光谱后,样品可以随位移平台移动,避免激光束在同一位置重复烧蚀。位移平台的每次移动,都要确保每个烧蚀坑之间间隔0.5 mm,并烧蚀出一个 20×20 的点矩阵。激光光源采用北京镭宝(Beamtech)光电技术有限公司生产的调Q开关Nd:YAG激光器(Nimma-400,波长:532 nm,脉冲持续时间:8 ns,最大重复频率:10 Hz)。激光输出后被 45° 反射镜反射,并透过石英透镜(焦距:150 mm)后聚焦到样品表面。激发产生的等离子发射光用采集头(深圳青锐科技有限公司, SPECC100)收集,采集头内置一个焦距为50 mm、直径为2.54 cm的紫外熔融石英平凸透镜,用于将等离子体光辐射耦合进光纤。并通过光纤传输至六通道光谱仪(Avantes, AvaSpec-ULS4096CL-EVO,光谱范围:196~874 nm,最小门宽:9 μ s,分辨率:0.07~0.12 nm)。激光器和

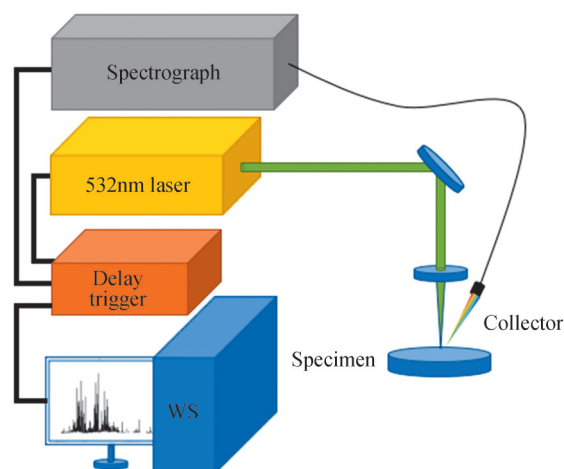


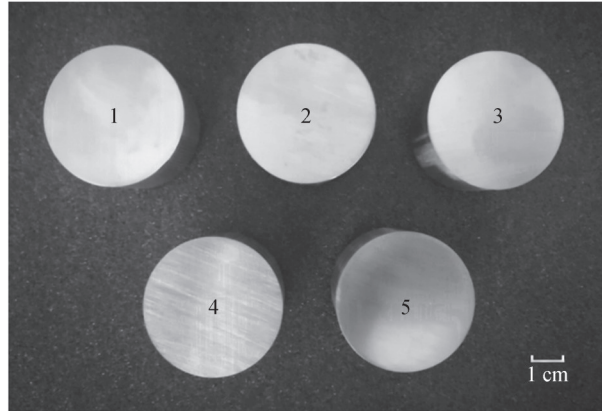
图1 实验装置示意图

Fig. 1 Diagram of experimental setup

光谱仪由数字延迟发生器(Stanford Research Systems, DG645)进行触发控制,整个实验在开放式环境中进行。后期进行数据处理与模型训练的计算机为华硕笔记本电脑,处理器型号: Intel(R) Core(TM) i7-8750H CPU @ 2.20 GHz,内存 16 GB。

1.2 实验样品

使用的实验样品牌号为 TC4(Ti-6Al-4V) 钛合金,具体为标样 TC4-1(GBW02503)、TC4-2(GBW02504)、TC4-3(GBW02505)、TC4-5(GBW02507)与控样 TC4KY 五种样品。在采集光谱前,每份样品都进行打磨处理并用酒精清洗,以去除表面杂质和氧化层。样品实物如图 2 所示,每种样品的元素含量如表 1 所示。每种样品采集 400 幅,共获得 2 000 幅钛合金的 LIBS 光谱。



TC4-1(1)、TC4-2(2)、TC4-3(3)、TC4-5(4)、TC4KY(5)

图 2 TC4(Ti-4Al-4V) 钛合金实物图
Fig. 2 Picture of TC4 titanium alloy

表 1 钛合金编号及元素含量

Table 1 Titanium alloy number and element content

Material	Number	Concentration of element/%					
		Al	V	Fe	Si	C	Ti
TC4-1	GBW02503	3.90	5.56	0.39	0.28	0.16	89.62
TC4-2	GBW02504	4.67	5.01	0.31	0.20	0.12	89.69
TC4-3	GBW02505	5.38	3.41	0.24	0.12	0.10	90.75
TC4-4	GBW02507	6.78	3.85	0.13	0.09	0.02	89.13
TC4KY	—	6.24	4.08	0.05	0.02	0.01	89.60

1.3 实验参数优化

1.3.1 能量优化

激光烧蚀样品表面时,激光功率密度与 LIBS 光谱质量紧密相关,因此需要对激光能量、离焦量等实验参数进行优化。离焦量优选为 -1 mm,能量优化时所采用的延时根据前期实验的经验选取 2 μ s。依据前期文献调研与经验^[23],在进行激光能量优化时,选取表 2 中的 6 条谱线。以五点寻峰法获取峰值强度并将其平均值作为信号,选取谱线附近无谱线干扰的背景计算信噪比,通过峰值强度与信噪比的大小与变化趋势来决定最佳能量强度。信噪比计算公式为

$$\text{SNR} = \frac{S}{N} \quad (1)$$

式中, S 为选取谱线的峰值强度均值, N 为选取的背景范围内所有数据的标准差。

表 2 为选取谱线的波长、跃迁几率 A_k 、下能级能量 E_1 与上能级能量 E_2 。

每次调整单位脉冲能量强度后,在样品 TC4-1 上进行 5×5 的矩阵扫描激发,获得 25 幅 LIBS 光谱。每幅光谱在特征处使用五点寻峰法获得其峰值,再计算信噪比,最终获得 25 幅光谱的峰值和信噪比的均值与误差,绘制出如图 3 所示的峰值强度与信噪比随单位脉冲能量变化的折线图。

表2 谱线波长、跃迁几率及能级能量

Table 2 Spectral wavelength, transition probability and energy level

Element	Wavelength/nm	A_k/S^{-1}	E_1/eV	E_2/eV
Al I	394.401	4.99×10^7	0	3.14
Al I	396.152	9.85×10^7	0	3.14
V II	311.071	1.58×10^8	0.35	4.33
V II	355.680	6.40×10^7	1.13	4.61
Ti I	395.820	4.88×10^7	0.05	3.18
Ti I	399.864	4.81×10^7	0.05	3.15

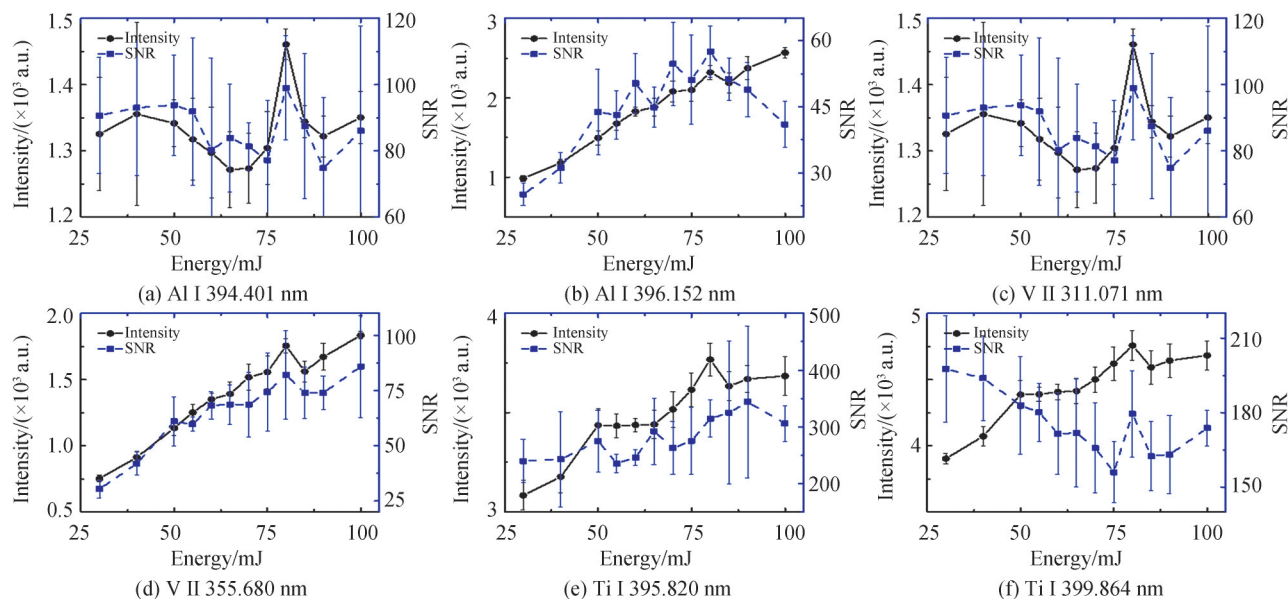


图3 峰值强度与信噪比随单位脉冲能量的变化

Fig. 3 The change of peak intensity and signal to noise ratio with unit pulse energy

如图3所示,横轴为单发激光脉冲能量强度,纵轴(左)为特征处的峰值强度,纵轴(右)为信噪比。黑色实线为峰值强度与单发激光脉冲能量之间的关系,蓝色虚线为信噪比与单发激光脉冲能量之间的关系。可以看出:随着单发激光脉冲能量的增加,峰值强度和信噪比也随之提高,当单发激光脉冲能量大小为80 mJ时,大多特征处的峰值强度最高且信噪比最大,故选取80 mJ作为单脉冲输出能量,在此条件下,功率密度为 $3.18 \times 10^{10} \text{ W/cm}^2$ 。

1.3.2 延时优化

光谱仪的触发延时对光谱质量亦有着很大的影响。延时过短时,韧致辐射导致的光谱背景较大,光谱质量较差;延时过长时,等离子体已经逸散,此时采集的光谱难以与背景区分,所以对延时进行优化显得十分重要。仍然以样品TC4-1钛合金为例进行延时优化。

在单脉冲能量为80 mJ的激光作用下,每次调整采集延时后,对TC4-1样品进行 5×5 的矩阵扫描,并进行光谱累加。特征谱线选择与能量优化过程时所选相同。计算峰值和信噪比的均值与误差,通过两种参数的大小与变化趋势来决定最佳触发延时。

实验结果如图4所示,横轴为触发延时,纵轴(左)为峰值强度,纵轴(右)为信噪比。黑色实线为峰值强度与触发延时之间的关系,蓝色虚线为信噪比与触发延时之间的关系。可以看出:随着触发延时的增加,峰值强度逐渐降低,而多数特征处的信噪比在2 μs 时会达到峰值,之后开始逐渐下降。所以最终采用的光谱仪触发延时为2 μs 。

优化后的全谱图如图5所示,通过全谱可以观测到样品中各元素的发射光谱。

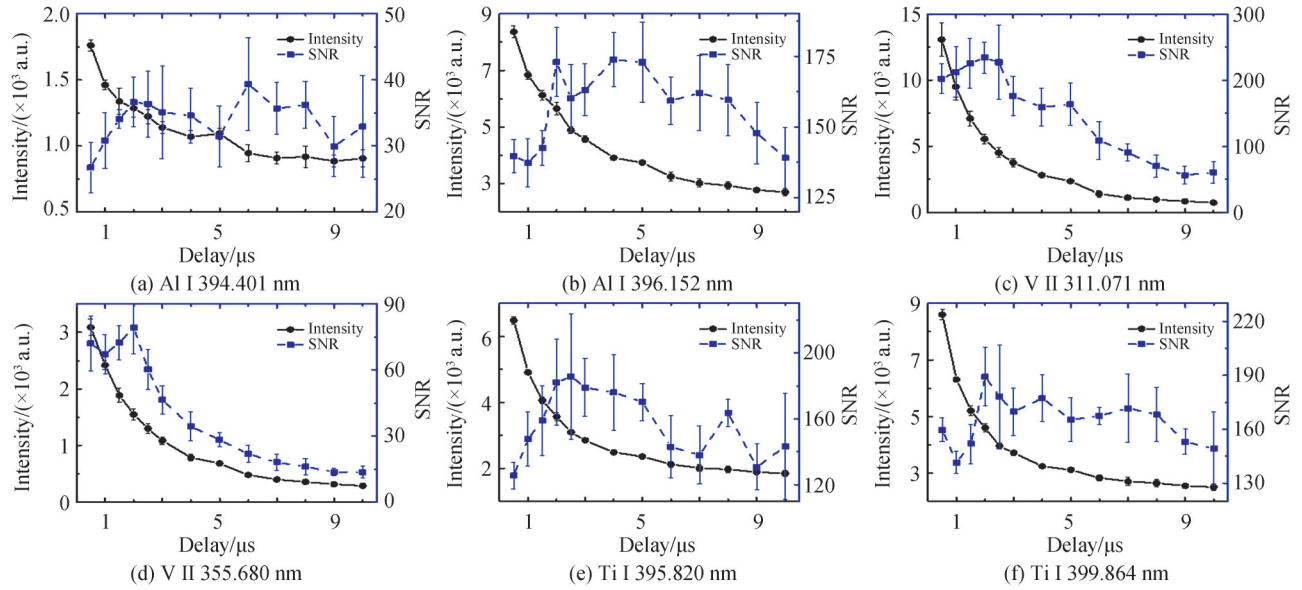


图4 峰值强度与信噪比随触发延时的变化
Fig. 4 The change of peak intensity and signal to noise ratio with trigger delay

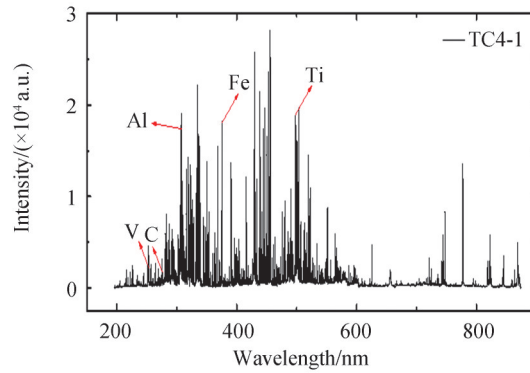


图5 TC4-1全谱图
Fig. 5 TC4-1 full spectrum

2 结果与讨论

在前期研究中,使用不同的分类算法对不同牌号的金属材料进行了分类研究。其中无监督算法的K-means效果极差,分类准确度仅56.05%;有监督算法中,决策树的分类准确度为88.4%,朴素贝叶斯分类(Naive Bayes Classifier, NBC)的结果为98.6%,KNN和SVM的结果为100%,但KNN的训练速度略快于SVM。本文中的研究样品为同牌号的钛合金,与前期研究有一定相似度,故选取KNN作为本次研究中训练模型的算法。KNN是一种经典的分类算法,非常适合LIBS光谱的快速分类研究及应用^[24]。KNN的原理是通过识别未知点X附近K个已知点的类别,来决定X属于哪一个种类。本文中用的是MATLAB中的机器学习与深度学习APP Classification Learner来实现分类。

2.1 原始光谱分类结果

首先对原始LIBS光谱进行分类处理,对钛合金样品采集的2000幅LIBS光谱按照3:1的比例划分训练集与测试集,共获得训练集1500幅LIBS光谱与测试集500幅LIBS光谱。基于快捷、简便的思想,结合前期研究,选取Classification Learner已有模板中的加权KNN对训练集进行分类梳理。加权KNN参数:K值过小会导致整体模型变得复杂,容易发生过拟合,K值过大则相反,会使模型变简单,同时对现有训练集的训练误差增大,故优先选取K值为中间值10的加权KNN;加权KNN的距离度量是欧氏距离,是最常用的距离度量;加权KNN与其他模板最大的不同是其距离权重为两点之间反距离的平方,即两点之间欧式距离越大,

其权重越小且变小的幅度远大于距离权重为反距离时的情况。训练集分类时,进行十折交叉验证,即将训练集分为十组,使用其中一组数据验证其余九组数据训练出的模型,最终通过十次结果的平均正确率来对算法精度进行估算,从而避免过拟合现象出现。最终获得如图6所示的混淆矩阵。绿色框中的数字代表分类正确的光谱,红色框中的数字代表分类错误的光谱,白色框为未分类错误的部分。

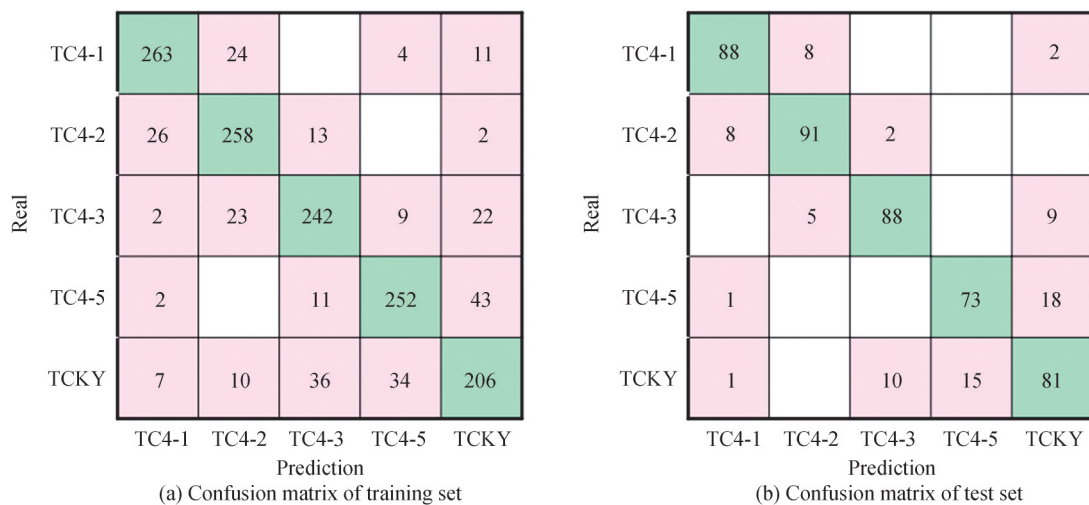


图6 原光谱的训练集进行加权KNN分类后的混淆矩阵与测试集分类结果的混淆矩阵

Fig. 6 Confusion matrix of training set of original spectrums after weighted KNN classification and confusion matrix of test set classification result

从图6中的混淆矩阵可以计算出,训练集交叉验证准确率为81.40%,1500幅训练集光谱中有1221幅分类正确。其中TC4KY因其元素含量与4种标样之间并无线性关系,故是容易产生分类错误的一种样品;而TC4-1、TC4-2与TC4-3三种样品之间,由于其元素浓度接近,也是容易分类错误的几组样品。将测试集的数据代入分类模型,获得的最终结果为:测试集分类准确度84.20%,500幅测试集光谱中有421幅分类正确,模型训练所用时间为1232.41s。

通过分析可知,采用原始数据与加权KNN直接进行分类存在两方面的不足,一是分类准确度不高,测试集的分类准确度只有84.20%;二是训练时间过长,总时长达二十多分钟。因此,从数据与模型两个方面进行优化,进而提高分类准确度,减少分类模型的训练时间。

2.2 数据优化

数据的优化主要从三个方面进行:1)通过数据筛选,剔除劣质光谱,进而避免其对结果带来的影响;2)通过数据归一化,减少实验环境和实验参数波动的影响;3)通过主成分分析,进行数据降维,提高分类准确率并减少模型训练时间。

数据的筛选是通过在元素浓度适中的Al、V、Fe三种元素的LIBS谱线中,选取如表3中的10处峰值高、信噪比大的谱线,由于谱线的波动近似趋近于正态分布,所以在这些波长处使用 3σ 法进行异常值剔除,最终剩余1866幅光谱。为方便实施模型训练时的十折交叉验证,随机选取1400幅光谱为训练集,剩余466幅光谱为测试集。

表3 光谱筛选用特征谱线

Table 3 Characteristic lines for spectral screening

Element	Al I	Al II	V I	V II	Fe II
Wavelength/nm	226.90	281.62	411.15	268.87	238.24
	396.15		439.55	310.23	
				319.07	259.99

最大值和最小值对于数据归一化的结果有很大影响,而剔除异常值时已将数据过大和过小的光谱筛去,确保了归一化的效果,所以剩余的每张光谱进行归一化后,结果预计会显著提升。归一化公式为

$$X'_i = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (2)$$

式中, $\min(X)$ 为光谱中峰值强度最小值, $\max(X)$ 为其中峰值强度最大值。

在此基础上,进一步研究发现原光谱共有 24 564个维度,冗余数据很多,严重降低了分类模型的训练速度,对于分类结果的准确度也有很大影响,故需要对数据进行降维处理。PCA通过正交变换将一组可能存在相关性的 n 维变量转换为一组线性不相关的 k 维变量,重新构造出来的 k 维特征称为主成分,在包含了原 n 维数据的主要信息的同时,数据将获得有效的降维处理,精简了数据。在保证总解释方差高于 98%的情况下,选取解释方差前 20 维的特征,输入 Classification Learner 重新进行分类模型训练。图 7 为数据处理之后分类结果的混淆矩阵。

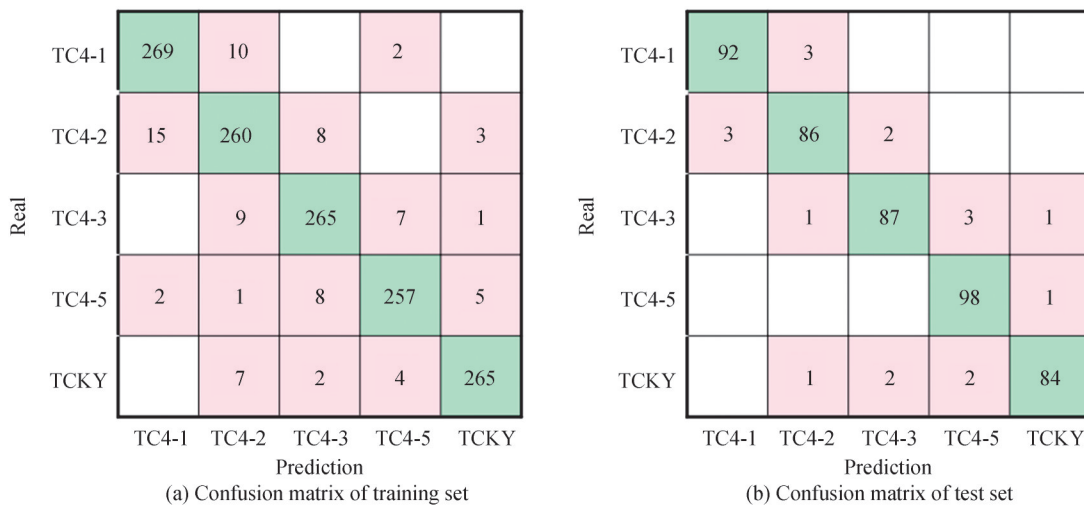


图 7 数据处理后的光谱训练集进行加权 KNN 分类后的混淆矩阵与测试集分类结果的混淆矩阵

Fig. 7 Confusion matrix of weighted KNN classification for spectral training set after data processing and confusion matrix of test set classification result

从图 7 中的混淆矩阵可以算出,训练集交叉验证准确率为 94.00%,1 400 幅训练集光谱中有 1 316 幅分类正确。原本最易分类错误的 TC4KY,其分类准确度大幅度提高。而因 TC4-1 与 TC4-2 之间的元素浓度最为接近,所以最后的主要分类错误集中在 TC4-1 与 TC4-2 之间。将测试集的数据代入分类模型,获得的最终结果为:测试集分类准确度 95.92%,466 幅测试集光谱中有 447 幅分类正确,模型训练时间为 84.76 s。

分类结果大幅度优化,测试集准确度提升至 95% 以上,模型训练时间缩小至原数据结果的 6.88%。对于 TC4-1 和 TC4-2 之间因元素浓度接近而容易混淆的问题,将通过模型参数的更改进行进一步优化。

2.3 模型参数优化

KNN 模型可以优化的参数主要有临近点数 K 、距离度量、距离权重三种。临近点数 K 是 KNN 的核心参数。在样品的元素浓度相近时,其数据点在图像上的表现往往是纠缠在一起的,不易区分。而通过选择合适的 K 值可解决此问题,进而大幅度提升分类结果的准确性。对于本文出现的 TC4-1 与 TC4-2 之间的混淆问题,可以通过增加 K 值,扩大邻点选择范围,进而提高分类准确度。由于 K 值过高时,结果将不会出现大幅波动,所以最优 K 值将在 1 到 100 范围内选择。距离度量为两点之间的距离计算方式,不同的距离度量下,两点之间的数据不同,因此,将在最常用的欧氏距离与最适合高维数据的余弦相似度中选取最优者。距离权重则从等距离、反距离及反距离平方中选出。图 8 为三种参数的优化结果。

图 8 中,横坐标为临近点数 K ,六条线分别为两种距离度量与三种距离权重的组合,纵坐标为最终结果测试集的准确率。可以看出,随着临近点数的增加,测试集准确度也随之增加,并在 $K > 20$ 后增幅趋于平缓。而在距离度量与距离权重的六种组合中,当 K 值足够大时,选取余弦相似度和反距离平方相结合时结果最好。最终选择 $K = 100$ 、距离度量为余弦相似度、距离加权为反距离平方,结果由 95.92% 提高至 99.14%。其 KNN 分类结果的混淆矩阵具体如图 9 所示。

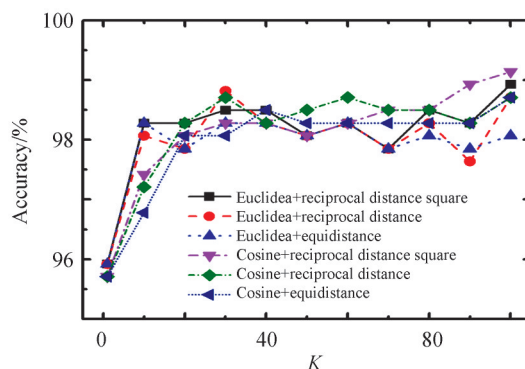


图8 三种参数的优化结果

Fig. 8 Optimization results of three parameters

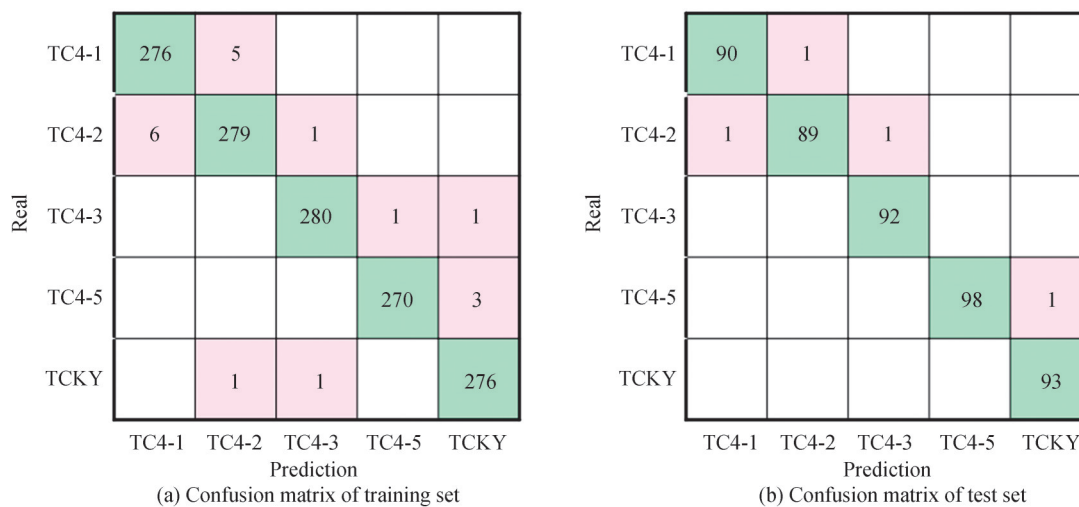


图9 最终分类结果的训练集混淆矩阵与测试集分类结果的混淆矩阵

Fig. 9 Training set confusion matrix of final classification result and confusion matrix of test set classification result

将测试集的数据代入分类模型,最终结果为:训练集交叉验证准确率98.64%,1400幅训练集光谱中仅有19幅分类错误;测试集分类准确度99.14%,466幅测试集光谱中仅有4幅分类错误;模型训练时间为83.97 s。

2.4 结果对比

表4为运用不同算法获得的最优结果。可以看出KNN依旧是进行同牌号钛合金分类的最优算法。KNN优化前后的效果对比如表5所示。实验结果表明,在经过数据与分类模型的优化之后,训练集交叉验证准确率提高了17.24%,测试集分类准确度达到了99.14%,提高了14.94%,分类模型训练时间减少了93.19%,意味着采用KNN算法结合LIBS技术对同牌号钛合金快速分类结果非常理想。

表4 不同算法间的结果对比

Table 4 Comparison of results between different algorithms

	K-means	Decision tree	NBC	SVM	KNN
Training time/s	43.86	79.28	78.95	91.18	83.91
Training set cross validation accuracy/%	53.34	81.6	97.9	98.4	98.64
Test set classification accuracy/%	58.31	82.91	96.79	98.5	99.14

表5 优化结果对比
Table 5 Comparison of optimization results

	Raw data	Data processing	KNN model optimization
Training time/s	1 232.41	84.79	83.91
Training set cross validation accuracy/%	81.40	94.00	98.64
Test set classification accuracy/%	84.20	95.92	99.14

3 结论

使用LIBS技术采集TC4钛合金样品的光谱,通过对数据进行筛选、归一化、PCA降维的处理,在大幅缩短训练时间的基础上,提高了分类的准确性;通过对分类模型参数的优化(优化后参数为: $K=100$,余弦相似度,反距离平方),最终稳定将分类准确度提升至99%左右。分类研究的最终结果表明:在开放式环境中,基于LIBS技术原位、微损、快捷、实时、多元素同时检测等优点,通过采用数据处理与模型优化相结合的方法,将分类模型的训练时间降低至83.91 s,分类的准确度提升至99.14%,实现了对于钛合金的快速原位微损的分类分析。本研究证明了LIBS技术在同牌号钛合金快速精准分类领域具有良好的应用前景。

参考文献

- [1] JIA Hong, LU Fusheng, HAO Bin. China titanium industry development report 2020[J]. Titanium Industry Progress, 2021, 38(2):34-41.
贾翊, 逯福生, 郝斌. 2020年中国钛工业发展报告[J]. 钛工业进展, 2021, 38(2):34-41.
- [2] JIN Hexi, WEI Kexiang, LI Jianming, et al. Research progress of titanium alloy for aviation [J]. Transactions of Nonferrous Metals Society of China, 2015, 25(2): 280-292.
金和喜, 魏克湘, 李建明, 等. 航空用钛合金研究进展[J]. 中国有色金属学报, 2015, 25(2): 280-292.
- [3] MA Zhongxian, FENG Juning, HU Zhijie. Research progress of titanium and titanium alloy profiles [J]. World Nonferrous Metal, 2016(24):52-53.
马忠贤, 冯军宁, 胡志杰. 钛及钛合金型材研究进展[J]. 世界有色金属, 2016(24):52-53.
- [4] GUO Jiaming, LIANG Jinglong, SHEN Haitao, et al. Preparation methods and application progress of biomedical titanium alloy materials[J]. Hot Working Technology, 2021(20):30-34.
郭佳明, 梁精龙, 沈海涛, 等. 生物医用钛合金材料制备方法及应用进展[J]. 热加工工艺, 2021(20):30-34.
- [5] GUO Lianbo, ZHANG Deng, SUN Lanxiang, et al. Development in the application of laser-induced breakdown spectroscopy in recent years: a review[J]. Frontiers of Physics, 2021, 16(2): 1-25.
- [6] YU Yang, HAO Zhongqi, LI Changmao, et al. Application of support vector machine algorithm in plastic identification by laser-induced breakdown spectroscopy[J]. Chinese Journal of Physics, 2013, 62(21): 215201.
于洋, 郝中骥, 李常茂, 等. 支持向量机算法在激光诱导击穿光谱技术塑料识别中的应用研究[J]. 物理学报, 2013, 62(21): 215201.
- [7] HU Zhenlin, CHEN Feng, ZHANG Deng, et al. A method for improving the accuracy of calibration-free laser-induced breakdown spectroscopy by exploiting self-absorption[J]. Analytica Chimica Acta, 2021, 1183: 339008.
- [8] HE Yaxiong, ZHOU Wenqi, KE Chuan, et al. Review of laser induced breakdown spectroscopy in gas detection [J]. Spectroscopy and Spectral Analysis, 2021, 41(9):2681-2687.
何亚雄, 周文琦, 柯川, 等. 激光诱导击穿光谱技术在气体检测中的研究综述[J]. 光谱学与光谱分析, 2021, 41(9): 2681-2687.
- [9] MA Shixiang, TANG Yun, MA Yuyang, et al. The pH effect on the detection of heavy metals in wastewater by laser-induced breakdown spectroscopy coupled with a phase transformation method [J]. Journal of Analytical Atomic Spectrometry, 2020, 35(1): 198-203.
- [10] LIN Jingjun, LIN Xiaomei, GUO Lianbo, et al. Identification accuracy improvement for steel species using a least squares support vector machine and laser-induced breakdown spectroscopy[J]. Journal of Analytical Atomic Spectrometry, 2018, 33(9): 1545-1551.
- [11] WU Jiasen, GAO Qing, WEI Gongxiang, et al. Optical properties and laser-induced breakdown spectroscopy analysis of Al-or co-doped CuO thin films prepared on glass by radio-frequency magnetron sputtering[J]. Thin Solid Films, 2021, 722: 138572.
- [12] CHU Yanwu, TANG Shisong, MA Shixiang, et al. Accuracy and stability improvement for meat species identification using multiplicative scatter correction and laser-induced breakdown spectroscopy [J]. Optics Express, 2018, 26(8): 10119-10127.
- [13] ZHAO Zhifang, YU Hanyue, ZHANG Siyu, et al. Visualization accuracy improvement of spectral quantitative analysis

- for meat adulteration using Gaussian distribution of regression coefficients in hyperspectral imaging[J]. *Optik*, 2020, 212: 164737.
- [14] ZHANG Dacheng, FENG Zhongqi, WEI Kuan, et al. Remote laser-induced breakdown spectroscopy and its application (invited)[J]. *Acta Photonica Sinica*, 2021, 50(10):1030001.
张大成, 冯中琦, 魏宽, 等. 远程激光诱导击穿光谱技术与应用(特邀)[J]. *光子学报*, 2021, 50(10):1030001.
- [15] COSTA V C, SANTOS FERREIRA SDOS, SANTOS L N, et al. Qualitative and quantitative analysis of soils using laser-induced breakdown spectroscopy and chemometrics tools[J]. *Journal of Applied Spectroscopy*, 2020, 87: 378-386.
- [16] GUO Lianbo, ZHENG Weinan, CHEN Feng, et al. Meat species identification accuracy improvement using sample set portioning based on joint x-y distance and laser-induced breakdown spectroscopy[J]. *Applied Optics*, 2021, 60(20): 5826-5831.
- [17] SUNGHO S, YOUNGMIN M, JAEPIL L E E, et al. Signal processing for real-time identification of similar metals by laser-induced breakdown spectroscopy[J]. *Plasma Science and Technology*, 2018, 21(3): 034011.
- [18] ABERKANE S M, ABDELHAMID M, MOKDAD F, et al. Sorting zamak alloys via chemometric analysis of their LIBS spectra[J]. *Analytical Methods*, 2017, 9(24): 3696-3703.
- [19] OLLILA A M, LASUE J, NEWSOM H E, et al. Comparison of two partial least squares-discriminant analysis algorithms for identifying geological samples with the ChemCam laser-induced breakdown spectroscopy instrument[J]. *Applied Optics*, 2012, 51(7): B130-B142.
- [20] LIU Kai, WANG Qianqian, ZHAO Hua, et al. Application of laser induced breakdown spectroscopy in plastic classification[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(5):1171-1174.
刘凯, 王茜茜, 赵华, 等. 激光诱导击穿光谱在塑料分类中的应用[J]. *光谱学与光谱分析*, 2011, 31(5):1171-1174.
- [21] GUO Lianbo, ZHANG Yong, HAO Zhongqi, et al. Quantitative detection of vanadium and titanium in steel by laser-induced breakdown spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2014, 34(1): 217-220.
郭连波, 张庸, 郝中骐, 等. 钢铁中钒, 钛元素的激光诱导击穿光谱定量检测[J]. *光谱学与光谱分析*, 2014, 34(1): 217-220.
- [22] ZHOU Zhonghan, TIAN Xueyong, SUN Lanxiang, et al. Fiber-LIBS technology combined with SVM to identify aluminum alloy brand[J]. *Laser & Optoelectronics Progress*, 2018, 55(6): 063002.
周中寒, 田雪咏, 孙兰香, 等. Fiber-LIBS 技术结合 SVM 鉴定铝合金牌号[J]. *激光与光电子学进展*, 2018, 55(6): 063002.
- [23] PAN Lijian, CHEN Weifang, CUI Rongfang, et al. Quantitative analysis of aluminum alloy based on laser induced breakdown spectroscopy and radial basis function neural network[J]. *Laser & Optoelectronics Progress*, 2020, 57(19): 193002.
潘立剑, 陈蔚芳, 崔榕芳, 等. 基于激光诱导击穿光谱与径向基函数神经网络的铝合金定量分析[J]. *激光与光电子学进展*, 2020, 57(19): 193002.
- [24] YAN Xiaotao, PENG Xinying, QIN Yuzhi, et al. Classification of plastics using laser-induced breakdown spectroscopy combined with principal component analysis and K nearest neighbor algorithm[J]. *Results in Optics*, 2021, 4: 100093.

Rapid Classification of Laser Induced Breakdown Spectroscopy of Titanium Alloys

XU Cheng¹, LI Fang¹, CHEN Feng², ZHANG Deng², DENG Fan³, GUO Lianbo³

(1 Hubei Key Laboratory of Optical Information and Pattern Recognition, School of Mechanical and Electrical Engineering, Wuhan Institute of Technology, Wuhan 430205, China)

(2 Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China)

(3 College of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: Titanium alloy is a kind of metal material which with a wide range of applications, including aerospace, rail transit, medical equipment and other fields. The appearance of titanium alloys with different brands is very similar, but they are suitable for different fields. Even the properties of titanium alloys with the same brand and different numbers are different, and confusion is easy to cause serious accidents. Therefore, it is urgent to study the rapid and accurate classification and identification of titanium alloys with

the same brand. In recent years, laser-induced breakdown spectroscopy as a fast, real-time, in-situ, micro-loss, multi-element synchronous analysis of advanced detection technology, favored by researchers. Using the advantages of laser-induced breakdown spectroscopy technology, the titanium alloys with different national standard numbers under the same brand are analyzed, which can realize the rapid and accurate classification and identification of titanium alloys. The whole device used in laser-induced breakdown spectroscopy is composed of laser, spectrometer, electronic control displacement platform, workstation, acquisition head, digital delay generator and several lenses and optical fibers. According to sequential analysis, the spectra of titanium alloy under various laser intensities and different delay times were collected by the device. Combined with previous literature research and experience, six characteristic spectra with high signal strength and high signal-to-noise ratio were selected. The optimal laser intensity and trigger delay were obtained by comparing the peak intensity and signal-to-noise ratio of the six characteristics. The TC4 titanium alloy spectrum collected under the optimal conditions is divided into training set and test set according to the ratio of 3:1. The training set data is trained by a variety of algorithms, and then the test set is substituted into the trained model. Through the analysis of the two results, the advantages and disadvantages of the algorithm are determined, and the optimal algorithm is K-nearest Neighbor algorithm. The optimization of data is mainly carried out from three aspects: 1) By using 3σ method for data screening at 10 characteristic spectra, the inferior spectra with too large or too small peak values are eliminated, to avoid its impact on the results; 2) Through data normalization, reduce the impact of experimental environment and experimental parameters; 3) Through principal component analysis, the data dimension is reduced, a large number of redundant data is reduced, the classification accuracy is improved and the model training time is reduced. KNN model can optimize the parameters mainly include the number of adjacent points, distance measurement, and distance weight. The number of adjacent points is the core parameter of KNN, which determines the number of data used to determine the unknown points. Distance measure is the distance calculation method between two points. Under different distance measures, the data between two points are different. Distance weight is the relative importance of determining the distance between the known point and the unknown point. The three parameters are arranged and combined into the model to retrain, and the optimal parameter combination is determined by comparing the results of the training set and the test set. The optimal classification results are obtained through various work, and the classification and recognition of the same grade titanium alloy are realized. The results show that the classification accuracy of the same grade titanium alloy can be improved from 84.15% to 99.14% by combining data processing and model optimization, and the training time can be reduced from 1232.41 s to 83.91 s. The classification performance is significantly improved. The research results are expected to achieve rapid and accurate classification and identification of titanium alloys with the same brand, and have broad application prospects.

Key words: Spectroscopy; Titanium alloy classification; Laser-induced breakdown spectroscopy; K-nearest neighbor; TC4 titanium alloy

OCIS Codes: 300.2140; 300.6210; 300.6365