人工神经网络在激光诱导击穿光谱数据分析中的应用进展

赵文雅[1,2]，闵　红[2]，刘　曙[2*]，安雅睿[1*]，俞　进[3]

1. 上海理工大学理学院化学系，上海　200093
2. 上海海关工业品与原材料检测技术中心，上海　200135
3. 上海交通大学物理与天文学院，上海　200240

摘　要　激光诱导击穿光谱(LIBS)具有实时、远程、多元素同时分析的优点，近年来在工业在线分析领域逐渐受到关注，发挥着重要作用。但基于发射光谱本身的特性，LIBS存在光谱噪声、基线漂移、自吸收和重叠峰等现象；又由于环境变化、激光能量波动、基体效应、样品表面形貌等因素，造成光谱稳定性和重现性差。这些问题导致光谱信息与定性、定量分析之间呈非线性关系，限制了分析灵敏度和准确度。随着LIBS器件稳定性的逐渐改善，LIBS光谱数据分析方法日新月异，人工神经网络(ANN)能跟踪和识别非线性特性，自适应学习LIBS光谱特征，筛除干扰信息，在LIBS数据分析领域的应用得到飞速发展。介绍了LIBS原理、仪器结构和工作流程以及在LIBS光谱分析领域常见的神经网络模型，总结出2015年—2020年LIBS结合常见的ANN模型在地质、合金、有机聚合物、煤炭、土壤及生物等领域的具体应用，指出ANN在数据分析领域的超强能力可有效改进LIBS分析精度，提升光谱数据利用率，降低光谱采集环境要求。针对仍然有待突破的技术难点，展望了ANN在LIBS光谱深度信息挖掘、便携式专用型设备开发、技术联用等方面的发展前景。LIBS日趋成熟，但其数据分析领域仍有广阔发展空间。该综述可为机器学习在LIBS数据分析领域的应用提供参考。

关键词　激光诱导击穿光谱；人工神经网络；数据分析；应用
中图分类号：O657.3　　**文献标识码**：R　　**DOI**：10.3964/j.issn.1000-0593(2021)07-1998-07

引　言

激光诱导击穿光谱(laser-induced breakdown spectroscopy，LIBS)是一种基于等离子体的原子发射光谱，具有实时、原位、微损、远距离、多元素同时分析等优点，但噪声、基体效应、基线干扰、自吸收和激光能量波动等影响分析结果的重现性和准确性[1]。化学计量学对LIBS进行基线校正、噪声滤波、峰拟合和数据压缩[2]，提取特征信号，实现模式识别和多变量回归分析，有效推动LIBS技术的应用发展。随着LIBS应用领域进一步扩大，传统化学计量学的分析精度和适用范围逐渐受限，多种化学计量学方法往往进行多重组合后消除光谱干扰，过程精细但复杂，易引起光谱失真现象。机器学习是对传统化学计量学方法的扩充，具有强大的数据挖掘能力，加强了对低质量和重复性差的光谱信号的利用，能够避免甚至利用干扰信息进行模式识别和回归分析，使得LIBS技术的适用性进一步扩大。其中人工神经网络(artificial neural network，ANN)是一种应用较为广泛的机器学习模型，源自生物学，结合数学和物理方法从信息处理的角度对人脑神经网络进行抽象，建立自适应非线性动态系统，具有强大的输入输出非线性映射能力、自我适应能力和学习能力[3]。20世纪40年代至今，ANN理论取得快速发展，由浅层简单网络逐渐过渡到深度学习，将LIBS光谱数据分析水平提升到新的高度。

张天龙等[4]总结了化学计量学在LIBS数据预处理(基线校正、噪声滤除、重叠峰分辨和数据压缩等)的独特优势，同时指出ANN模型在LIBS数据定性分析和定量分析的优缺点。Andrade等[5]提出数据处理是连接原始数据和提取特征分析信息的桥梁，化学计量学是将LIBS数据转化为有用信息不可或缺的工具。Fu等[6]指出数学方法可以充分利用LIBS光谱数据，减少基体干扰，提高了模型的预测能力。Guo等[7]提到ANN可以一定程度上纠正LIBS光谱的非线

收稿日期：2020-07-15，**修订日期**：2020-11-21
基金项目：海关总署科研项目(2019HK074)，国家重点研发计划(2018YFF0215400)资助
作者简介：赵文雅，1994年生，上海理工大学理学院化学系硕士研究生　　e-mail：182692066@st.usst.edu.cn
*通讯作者　　e-mail：liu_shu@customs.gov.cn；anyarui@usst.edu.cn

性效应，如自吸收和基体效应。20 世纪 90 年代起，ANN 对 LIBS 进行数据分析从起步拓展到各应用领域，近年来尤其受到广泛关注，尚没有专题应用综述。本文对 2015 年—2020 年该领域公开发表的文献进行归纳总结，介绍 LIBS 和常见 ANN 模型的基本知识，从地质样品、合金材料、有机聚合物、煤炭、土壤和生物等角度分析，归纳 LIBS 结合 ANN 的应用案例及技术优势，探讨了 ANN 在 LIBS 数据分析中的聚焦点和发展方向。

1　激光诱导击穿光谱和人工神经网络的发展

LIBS 是一种以高能量激光脉冲轰击物质表面，获取物质元素成分和含量的原子发射光谱分析技术，仪器由激光光源系统、激光聚焦系统、等离子体辐射收集系统、光谱仪和控制电路及数据处理系统组成，如图 1 所示[8]。工作流程为：计算机控制激光器发射高能量激光脉冲，经光学系统折射、扩束、聚焦等，作用于样品并产生等离子体；同时，光电管感应脉冲激光，并将信号传送给延时器记录延迟时间。等离子体中跃迁原子回到基态时发射特征光信号，增强型电感耦合器件(ICCD)将其转换为电信号，传输到计算机，进行数字化处理，直观显示为横坐标为波长、纵坐标为相对强度的光谱图。

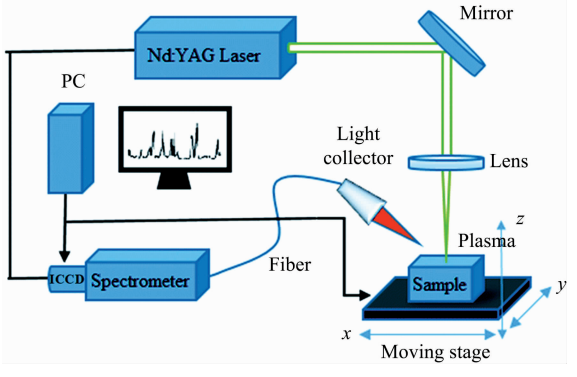

图 1　LIBS 仪器构成[8]

Fig. 1　The composition of LIBS instrument[8]

LIBS 光谱包含有用信息和干扰信息，数据解析过程复杂耗时，ANN 提升了 LIBS 数据的分析效率，拓展了 LIBS 的发展空间。LIBS 应用领域常见的 ANN 模型包括以下 5 种：

误差反向传播人工神经网络(back propagation artificial neural network，BPANN)在多层感知器的基础上增加误差反向传播。在正向和反向计算过程中，不断调整网络的权值和阈值，使网络的误差达到要求，显著提升了传统化学计量学在 LIBS 数据分析的速度和精确度，广泛应用于数据分类和预测，是 LIBS 领域应用最为广泛的模型。但 BPANN 在训练过程中，调节权值采用的梯度下降法收敛速度慢，容易陷入局部极小，使得 LIBS 数据分析时间延长，并且不能达到满意的精度。

20 世纪 80 年代末，有报道提出以函数逼近理论为基础的前向网络：径向基函数神经网络(radial basis function neural network，RBFNN)，借鉴生物局部调节和交叠接受区域的相关知识，通过局部接受域来执行函数映射[9]。主要改进：将隐藏节点的线性基函数转换为距离函数，并使用径向基函数作为激活函数。主要作用是将低维空间非线性可分的问题映射到高维空间，使非线性问题线性化，更利于在实际场景中的 LIBS 数据分析。

1992 年提出小波神经网络(wavelet neural network，WNN)，隐藏层转换函数变为由小波伸缩和平移得到的小波元函数。小波分析的独特优势在于信号去噪、信号奇异值检测及信号发展趋势检测等，有较强的时频特征提取能力；结合 ANN 后，对线性和非线性问题有良好的逼近能力[10]，更适用于在实际场景中，受噪声影响明显的 LIBS 信号分析。由于小波变换对时频特征信号的灵敏性，此模型更适用非线性时间序列预测问题，并且 LIBS 数据维数过高时，会使计算复杂，存在学习停滞风险，使得 LIBS 数据分析精度无法提升。

极限学习机(extreme learning machine，KLM)算法由 Huang 等在 2004 年提出。训练过程中无需调整，只需设置隐含层神经元的个数，便可获得唯一的最优专解，在 LIBS 数据分析的学习速率和泛化能力上更具优势[11]。此外，当隐含层特征映射被一个稳定的核函数[如径向基函数(RBF)]所替代时，此类学习机被称为核极限学习机(kernel extreme learning machine，K-ELM)。

卷积神经网络(convolutional neural network，CNN)由 Yann Lecun 等提出，与其他模型相比，增加了卷积区域：卷积层利用卷积核和训练权重对光谱数据进行卷积运算，通过非线性转换函数输出局部相关特征；池化层对特征进行下采样，实现对特征的进一步精炼，降低参数维度[12]。CNN 最突出的优点是整个训练过程中权值共享，降低参数存储量，提高了 LIBS 数据分析的运算速度；另外，较强的特征提取潜力节省人工提炼 LIBS 光谱数据特征过程，避免了特征不明显导致的精度不高问题。通常情况下，CNN 对原始数据进行学习可以达到令人满意的结果，简化了实验步骤。

目前，ANN 应用于 LIBS 数据分析的研究工作主要聚焦在两个方面：(1)选取不同的光谱数据作为 ANN 模型的输入变量，包括全谱、特定光谱波段、人工或算法提取的特征发射线等；(2)优化 ANN 配置参数，包括常规组合优化、运用遗传算法(GA)和粒子群算法(PSO)等进行全局寻优等。在最优输入和最佳配置下，ANN 处理 LIBS 数据时保持高效学习状态，得到最优预测结果，有利于 LIBS 技术在多领域实际场景的应用。

2　激光诱导击穿光谱结合人工神经网络的分析应用

2.1　地质样品

岩性识别：岩性的快速鉴别有利于质量监控，提高工业利用率。部分矿物会存在组成元素相同但含量不同的现象，LIBS 技术可以实现原始矿石的多元素快速检测，无需复杂

制样，但是仅依靠该技术进行元素分析并不能准确将各类矿石完全区分，而 ANN 技术分析 LIBS 光谱数据时，相较于其他模型，识别精度可以提升至较高水平。如图 2 所示，Alvarez等[13]用 LIBS 采集原矿石，保留了原始基质特征；对比树状图、PCA 和决策树对 LIBS 光谱进行聚类和特征提取；证明不同的输入变量会影响 LIBS 结合算法进行分类的性能，并与 K 近邻(KNN)算法，簇类独立软模式(SIMCA)算法和偏最小二乘判别分析(PLSDA)算法进行对比，利用灵敏度、精度、准确度、稳健性突出 ANN 较其他模型的巨大优势。有报道将 ANN 训练得到的权向量与测试样本进行相关系数分析，对比了改进后的 ANN 模型对全谱、主成分分析(PCA)和特征谱线的分类准确率，攻克了相似地质样本的聚类现象。Yang 等[14]利用主成分载荷图进行铁矿石 LIBS 光谱指纹区域提取，结合 ANN 对进口铁矿石进行原产国家和品牌的鉴别，准确度均在 99%以上。虽然 ANN 技术可以解决矿物 LIBS 数据的岩性快速识别问题，但通过研究人员的不断尝试，提取出的不同特征波段对不同模型预测精度仍存在较大的影响，ANN 较其他模型受影响程度相对较小。

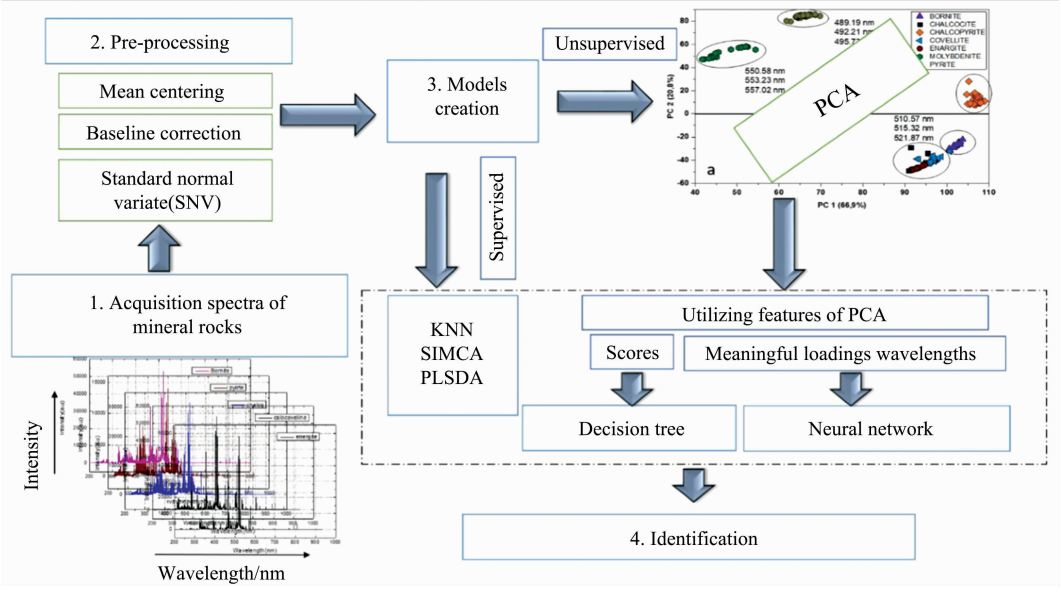

图 2 矿物 LIBS 光谱数据的典型处理流程图[13]

Fig. 2 Typical flow chart for processing mineral LIBS spectral data[13]

定量分析：地质样品含有丰富的金属元素，基质复杂，对 LIBS 技术进行定量分析带来挑战，通常情况下，定标效果并不理想，精度有限，引入 ANN 模型进行 LIBS 数据定量分析，攻克了 LIBS 技术分析地质样本的诸多缺陷。Lu 等[15]分析了不同熔沸点元素在激光烧蚀不同时间产生等离子体现象对 LIBS 特征发射线的影响，并利用 BPANN 成功预测 Ti 元素含量，证明 ANN 可有效减少元素等离子体出现时间不同而产生的非线性干扰。Ding 等[16]指出烧结铁矿石的 LIBS 光谱中存在显著的重叠峰现象和基体效应，利用多变量校准模型 K-ELM 模型能克服干扰，准确预测铁含量和碱度。胡等[17]指出 USGS 系列地质标样基体效应明显，采用 LIBS 技术结合 BPANN 对 Fe 元素定量分析的相对误差均在 6%以下，克服了自吸收、元素间干扰、激光能量波动等影响。

2.2 合金材料

废料分选：对金属废料进行回收利用有利于保护环境和发展循环经济，因此快速有效的分选技术不可或缺，LIBS 技术由于优越的在线监测优势，已经成功应用于多种废料的分选工艺。Campanella 等[18]利用“模糊”方法改进 ANN 模型，将每一类样本的 LIBS 光谱类型标签编成 4×1 列的向量(如 $*=[1\ 0\ 0\ 0]$)，并用向量代替类的数量作为输出，使 ANN 模型克服了 LIBS 常有的信号低和重复性差的问题。Kong 等[19]利用 LIBS 技术进行分区测试，Fe，Cr 和 Ni 特征线集中的波段经过 PCA 降低维数后，ANN 可以将废钢 100%区分，证明对废钢材料，即使依靠先验知识进行人工分区的 LIBS 光谱，无需复杂的特征提取，ANN 就可以获取令人满意的结果。

定量分析：合金在工业上应用广泛，由一种或多种金属和非金属熔合而成，具有复杂的元素组成。针对其元素含量，国内外均有 LIBS 结合 ANN 技术进行分析，各方案主要成果如表 1 所示。由于合金中的元素复杂多样，LIBS 光谱中常存在基线漂移、自吸收、重叠峰和谱线相互干扰等现象，导致对元素定量分析精度有限，学者们将 LIBS 与 ANN 结合，有效筛除干扰信息，实现了对合金中多种元素快速精确的分析，归因于 ANN 良好的非线性映射能力。此外，利用优化算法调整 ANN 参数，改善了收敛速度慢和陷入局部极小的问题；基于 ANN 的混合模型能有效防止过拟合，进一步提升了单一模型对合金元素定量学习的极限。

2.3 有机聚合物

硅橡胶老化程度鉴定：硅橡胶绝缘能力强，常被用于包裹电线，但老化问题会导致其性能下降、寿命缩短。Chen 等[25]利用热重分析(TGA)探究硅橡胶填料含量对老化的影响；PCA-ANN 模型分析 Si，Al，C 和 O 等元素的光谱线强

表 1　LIBS 技术结合 ANN 模型对合金元素进行分析的典型研究

Table 1　The typical study on the analysis of alloying elements by LIBS technique combined with ANN model

合金类型	元素	ANN 模型	主要成果	参考文献
碳钢	C(Ⅰ) 247.86 nm	GA-BPANN	克服相邻铁谱线干扰	[20]
碳钢；低合金钢；微合金钢	Cu；V	GA-ANN	GA 用于选取样本目标元素的特征谱线强度比，提高了 ANN 分析精度	[21]
碳钢；低合金钢；微合金钢	Cr；Ni	基于 ANN 的多谱线校正(MSLC)	用目标和基体元素的多谱线强度比训练神经网络识别等离子体脉冲的变化，克服了激发条件不稳定的问题和自吸收效应	[22]
铜镍二元合金	Ni	BPANN	训练 ANN 识别谱线强度与等离子体参数之间的基本物理关系，克服了激光波动和基体效应	[23]
青铜标样	Cu (324.75 nm；327.39 nm) Sn (303.41 nm；326.23 nm)	ELM-SVR	以 ELM 的输出为支持向量回归(SVR)输入的混合模型，解决 ELM 模型的超参数过拟合的问题，提高了定量分析的准确度	[24]

度的关系，对样品的抗侵蚀水平进行分类。由于材料有机成分含量高，LIBS 技术检测光谱仅包含元素的类型和发射峰强度，缺少物质结构特征，说明在这些领域，仅靠 LIBS 技术结合算法进行分析有一定的局限性，与其他技术互补有助于全面分析和解决问题。

塑料分选：对塑料回收再利用，突破传统手工分拣效率低的难题，已经发展成一项全球任务。Roh 等[26]利用 PCA 和独立成分分析(ICA)混合预处理算法将 LIBS 技术采集的光谱进行特征提取，提高了 RBFNN 学习性能、计算效率和泛化能力，该识别系统对黑塑料识别问题有协同效应，ANN 模型处理 LIBS 数据时，克服了黑色塑料物理状态以及噪声信息对分析的干扰，验证了识别性能与提取的特征个数之间的关系；最后指出 RBFNN 继承了传统 ANN 和模糊理论的优点，对 GA 优化其结构进行了展望。如图 3 所示，Junjuri 等[27]利用飞秒激光诱导击穿光谱(fs-LIBS)收集塑料信息，用 PCA 筛选出仅占总数据 2.5%的 10 个显著的原子和分子特征光谱，ANN 的识别率高达 100%，证明神经网络模型对于塑料 LIBS 数据的显著特征的学习效率很高，对塑料分拣做出贡献。利用化学计量学进行 LIBS 数据降维和特征提取的手段相对快速简便，传统的手动筛选特征费时费力，不合适的特征反而会降低分类器的正确率，难以应用于现场快速分拣，LIBS 技术结合 ANN 模型既能提高所提取的特征光谱的利用价值，也可以节约分析时间。

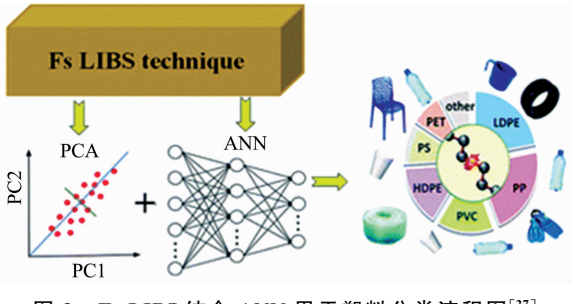

图 3　Fs-LIBS 结合 ANN 用于塑料分类流程图[27]

Fig. 3　Fs-LIBS combined with ANN flow chart for plastic classification[27]

2.4　煤炭

灰分分析：煤炭中含有丰富的矿物质，燃烧后会产生大量灰分，LIBS 技术已被证明可以有效检测灰分组成成分，快速得知煤炭的种类及特性。有报道采用独立主成分分析(ICA)-WNN 模型具有对煤灰的分类与鉴别能力，证明 WNN 可以对不同煤灰进行种类鉴别。另外，Wei 等[28]利用小波分析多尺度和局域性、ANN 自组织和自适应的优点，对煤灰 LIBS 光谱进行噪声和干扰信息滤除，分析几种主要矿物质(SiO_2，Al_2O_3，Fe_2O_3，CaO，MgO，MnO_2 和 TiO_2)的含量，证明 WNN 比传统 ANN 具有更强自适应能力、更快收敛速度和更高预测精度。

热值测量：热值是反映煤品质的又一指标。Lu 等[29]根据物理机理及数学手段挑选 LIBS 数据的分析谱线，并用 GA 优化 BPANN，提高了模型收敛速度和泛化能力，预测总热值的平均绝对误差(MAE)为 0.39 $MJ \cdot kg^{-1}$。有研究针对煤粉样本间的基体差异，采用 K-means 将样本聚类，选用离聚类中心较近的样品作为预测集，消除基体效应干扰，使其更有一般性，然后根据物理意义挑选主要元素(C，O，H，Ca，Na，K，Mg 等)发射线作为 BPANN 输入，预测结果重复性较好。

元素含量分析：煤炭作为主要能源之一，元素含量分析对煤质监测意义重大。Yan 等[30]利用 K-ELM 模型克服了传统线性模型无法解决的 LIBS 实验参数不可控波动、样品表面不均匀性和物理化学基质效应等问题，完善了传统 ANN 训练缓慢、局部最优、过拟合和结果波动大等缺陷，成功用于煤炭中 C 和 S 元素定量测试。

2.5　土壤

有机质含量预测：土壤的理化性质可以通过有机质评价，作为有机物检测，LIBS 技术往往需要结合其他技术进行信息互补分析。Xu 等[31]研究了 LIBS 技术与衰减全反射傅立叶变换中红外光谱(FTIR-ATR)的数据融合策略：分别将光谱进行处理后，直接进行拼接作为输入变量，定量分析结果优于单一光谱预测精度，说明将 LIBS 技术与其他方法进行数据融合可以进一步提高 ANN 的学习能力。

元素含量分析：土壤是植物生长的基质，对土壤中金属

元素的检测可以有效反映土壤的健康状况。国内外利用 LIBS 结合 ANN 进行的研究如表 2 所示。LIBS 原始数据量大，对其进行特征提炼可以降低维度，提高 ANN 计算速度。在文献[32-34]典型报道中，利用经验或算法选取特征谱线及相关谱线的区间，大大缩减了数据量，ANN 对土壤中的 K，Cd，Pb，Cd 和 Ag 等元素均能实现高精度预测，有效克服了基体效应和噪声干扰，表现出强大的自学习、自组织和自适应能力。

表 2 LIBS 技术结合 ANN 模型检测土壤金属含量的典型研究

Table 2 Typical studies of LIBS technology combined with ANN model to detect soil metal content

元素	ANN 模型	输入变量	主要成果	参考文献
K	CNN	主成分和延迟时间确定的时间分辨 LIBS 数据矩阵	$R_V=0.996\ 8$；$RMSE_V=0.078\ 5$	[32]
Cd	BPANN	分析谱线与谱线区间：(356. 29～346. 93 nm)	$R_C=0.999\ 9$；$R_P=0.981\ 5$	[33]
Pb，Cd	BPANN	分析谱线与谱线区间：(346. 29～346. 93 nm)；(405. 43～406. 17 nm)	$R_P=0.995\ 3$；$RMSE_P=0.145\ 2$	
Ag	BPANN	数据最后一列加入一项带有土壤类型的数据，组成广义光谱	$R_V=0.999\ 9$；ppm 量级	[34]

注：R：决定系数；RMSE：均方根误差；C：校准集；V：验证集；P：预测集

Note：R：Determination coefficient；RMSE：RMS error；C：Calibration set；V：Verification set；P：Prediction set

2.6 生物材料

植物：植物是自然界中必不可少的组成部分，品种繁多，种类间基质差异较大，LIBS 结合 ANN 技术能克服种类差异，有效分析各种植物。Peng 等[35]对比了偏最小二乘回归(PLSR)和 ELM 模型对烟叶中铜含量的标定性能，指出 ELM 是一种优于传统线性模型的多变量校准方法，能够从原始信号中提取有用的信息，在非线性情况下具有更好的预测能力，避免了烧蚀过程、基体效应和参数波动所导致的 LIBS 光谱偏差。植物医学方面，疾病快速诊断是一项具有挑战性的任务，Liu 等[36]利用 LIBS 技术诊断油菜菌核病，相比其他三种模型，ELM 预测准确率最佳，但仅 85%左右，说明对油菜这种基质，LIBS 光谱噪声、基线等干扰信息过多，仍有待开发能有效克服多种非线性干扰的化学计量学手段。Liu 等[37]指出仅依靠纤维素含量很难对生物质鉴定，利用 LIBS 结合多种化学计量学方法对四类生物质球团进行分析时，RBFNN 优势明显。此外，为进一步验证 RBFNN 模型优秀的识别性能，根据模型的预测类别和实际样本，绘制生物质球团的分类伪彩色图像，如图 4 所示。这种手段有效克服了 ANN 固有的暗箱操作缺陷，利用其他可视化技术能进一步解释算法对光谱信息的学习和特征提取能力，直观地显示出 ANN 技术的优势。

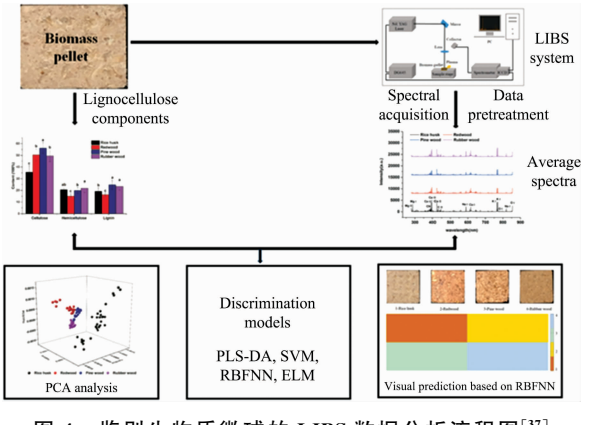

图 4 鉴别生物质微球的 LIBS 数据分析流程图[37]

Fig. 4 Flow chart of LIBS data analysis for identifying biomass pellets[37]

微生物：致病微生物包含细菌、病毒、霉菌、感染因子、变形虫和真菌等，在生活中时常出现，对人体健康存在隐患。LIBS 可以实现样本直接检测，结合 ANN 技术后，提高了微生物分析的效率和速度。Manzoor 等[38]利用 LIBS 技术获得不同元素(C，N，H，O 和 CN 等)的组合信息，经过 ANN 学习后正确鉴别了培养皿中 7 种念珠菌。对于细菌的鉴别，Prochazka 等[39]利用拉曼光谱(Raman)与 LIBS 技术化学信息互补(分子组成和元素组成)的优点，进行光谱融合，基于自组织映射算法(SOM)的 ANN 能够对融合信息实现 100%鉴别，补偿了原子发射光谱的局限，说明基于 LIBS 的技术融合可以提高细菌鉴别的精度，同时也证明 ANN 强大的多元数据分类能力。

2.7 其他领域

LIBS-ANN 技术除了解决以上领域的相关问题外，在食品质量控制(如牛奶掺假、转基因玉米鉴别)、金属冶炼质量控制(如炉渣元素定量分析)、核安全检测(如核法证学)、考古研究(如古陶瓷鉴定)、公共安全(如枪支发射残留识别、大宗物流的重金属危险品检测)、文件检查程序(如纸墨样本分类)等方面也均取得少量的应用成果。

3 总结与展望

近年来，LIBS 结合化学计量学进行数据分析的研究成为热点，ANN 模型在地质样品、合金材料、有机聚合物、煤炭、土壤、生物材料等多领域的 LIBS 光谱分析中展示出超强数据挖掘能力：有效克服和利用噪声、基体效应、基线干扰、自吸收和激光能量波动；可以从各类原始数据中提取有价值的非线性信息进行自适应学习，提高 LIBS 分析的准确度和灵敏度。然而，利用 ANN 对 LIBS 数据分析正处于探索阶段，低质量光谱经过多层非线性转换后，能达到较高的定性和定量分析的根本原因仍有待全面研究，未来该领域仍有

巨大的发展空间：

(1)ANN 在 LIBS 数据分析中的应用由浅层简单网络向深度学习发展，对于高维和数据量大的光谱数据，浅层神经网络未必能深入学习光谱特性，无法满足精度要求。网络层数增多，光谱数据经过多层映射后，特征更加凝练，有用信息更容易被充分挖掘。合适的光谱数据对分析精度具有很大影响，已报道的研究主要集中在经验挑选和算法提取(如主成分分析、独立成分分析、平均影响值等)，但这些信息提取手段与 ANN 模型彼此独立，提取特征的适用性仍有待证明，而深度学习的神经网络模型增加了特征提取层，所提取的光谱信息与其后负责模式识别和回归的隐藏层相互配合，保证了光谱特征的适用性。随着深度学习的发展，研究多种算法优化组合，进行 LIBS 信息挖掘、特征选择和回归，克服化学基体效应同时更多关注物理基体效应[40]被提上日程。此外，ANN 模型是一种暗箱操作，如何合理解释这种非线性映射关系，解析特征，是必不可少的环节。

(2)ANN 在 LIBS 数据分析中的应用将更广泛地集成于便携式专用型设备。为更好地适用于户外或工业现场测试，LIBS 仪器逐渐向便携化、集约型靠拢，现今国内外均有了实用性仪器的投用，但实现快速有效地实时监测，仍任重而道远。另外，集约型仪器需要更加智能、精确的预处理、特征提取和预测算法支撑。调研的文献中绝大部分神经网络集中在 BP 算法建立的网络模型，而其他 ANN 模型探索依旧很少，有待更全面的研发。

(3)ANN 在 LIBS 数据分析中的应用将进一步拓展到 LIBS 与其他光谱技术联用。LIBS 是一种典型的原子光谱，测量范围局限在 150～950 nm 波段，将其用于硅橡胶、微生物等领域的分析具有局限性。近期大量出现 LIBS 与拉曼(Raman)、激光烧蚀电感耦合等离子体发射光谱法(LA-ICP-OES)、X 射线荧光光谱仪(XRF)、傅里叶红外光谱仪(FT-IR)技术联用的报道，将各类技术进行信息互补，以弥补 LIBS 只能用于元素分析、光谱干扰大和对非金属不灵敏的缺点，提高检测精度。

References

[1] Fortes F J, Moros J, Lucena P, et al. Analytical Chemistry, 2013, 85(2): 640.

[2] Zhang T, Yan C, Qi J, et al. Journal of Analytical Atomic Spectrometry, 2017, 32(10): 1960.

[3] ZHU Da-qi, SHI Hui(朱大奇，史　慧). Principle and Applications of Artificial Neural Networks(人工神经网络原理及应用). Beijing: Science Press(北京：科学出版社), 2006. 1.

[4] ZHANG Tian-long, WU Shan, TANG Hong-sheng, et al(张天龙，吴　珊，汤宏胜，等). Chinese Journal of Analytical Chemistry(分析化学), 2015, 43(6): 939.

[5] Andrade D F, Pereira-Filho E R, Amarasiriwardena D. Applied Spectroscopy Reviews, 2021, 56(2): 98.

[6] Fu X, Li G, Dong D. Frontiers in Physics, 2020, 8: 68.

[7] Guo Y M, Guo L B, Li J M, et al. Frontiers of Physics, 2016, 11(5): 114212.

[8] Hu Y, Li Z, Lu T. Journal of Analytical Atomic Spectrometry, 2017, 32(11): 2263.

[9] Chen S, Cowan C N, Grant P M. IEEE Transactions on Neural Networks, 1991, 2(2): 302.

[10] HOU Yuan-bin, DU Jing-yi, WANG Mei(侯媛彬，杜京义，汪　梅). Neural Networks(神经网络). Xi'an: Xi'an Electronic Sience & Technology University Press(西安：西安电子科技大学出版社), 2007. 89.

[11] Huang G, Song S, Gupta JND, et al. IEEE Transactions on Cybernetics, 2014, 44(12): 2405.

[12] Chen J, Pisonero J, Chen S, et al. Spectrochimica Acta Part B: Atomic Spectroscopy, 2020, 166: 105801.

[13] Alvarez J, Velasquez M, Kumar Myakalwar A, et al. Journal of Analytical Atomic Spectrometry, 2019, 34(12): 2459.

[14] Yang Y, Li C, Liu S, et al. Analytical Methods, 2020, 12(10): 1316.

[15] Lu T, Hu Y, Li Z, et al. Spectrochimica Acta Part B: Atomic Spectroscopy, 2018, 143: 55.

[16] Ding Y, Yan F, Yang G, et al. Analytical Methods, 2018, 10(9): 1074.

[17] HU Yang, LI Zi-han, LÜ Tao(胡　杨，李子涵，吕　涛). Laser & Optoelectronics Progress(激光与光电子学进展), 2017, 54(5): 345.

[18] Campanella B, Grifoni E, Legnaioli S, et al. Spectrochimica Acta Part B: Atomic Spectroscopy, 2017, 134: 52.

[19] Kong H, Sun L, Hu J, et al. Plasma Science & Technology, 2015, 17(11): 964.

[20] He J, Pan C, Liu Y, et al. Applied Spectroscopy, 2019, 73(6): 678.

[21] Li K, Guo L, Li J, et al. Applied Optics, 2017, 56(4): 935.

[22] Li K, Guo L, Li C, et al. Journal of Analytical Atomic Spectrometry, 2015, 30(7): 1623.

[23] D'Andrea E, Pagnotta S, Grifoni E, et al. Applied Physics B: Lasers and Optics, 2015, 118(3): 353.

[24] Owolabi T O, Gondal M A. Journal of Intelligent & Fuzzy Systems, 2018, 35(6): 6277.

[25] Chen P, Wang X, Li X, et al. Sensors, 2019, 19(5): 1087.

[26] Roh S-B, Park S-B, Oh S-K, et al. Journal of Material Cycles and Waste Management, 2018, 20(4): 1934.

[27] Junjuri R, Gundawar M K. Journal of Analytical Atomic Spectrometry, 2019, 34(8): 1683.

[28] Wei J, Dong J, Zhang T, et al. Analytical Methods, 2016, 8(7): 1674.

[29] Lu Z, Mo J, Yao S, et al. Energy & Fuels, 2017, 31(4): 3849.

[30] Yan C, Qi J, Ma J, et al. Chemometrics and Intelligent Laboratory Systems, 2017, 167: 226.
[31] Xu X, Du C, Ma F, et al. Geoderma, 2019, 355: 113905.
[32] Lu C, Wang B, Jiang X, et al. Plasma Science & Technology, 2019, 21(3): 034014.
[33] KANG Sheng, WEI Zhao-hang, YANG Fan, et al(康 盛, 魏兆航, 杨 帆, 等). Agricultural Engineering(农业工程), 2019, 9(10): 38.
[34] Sun C, Tian Y, Gao L, et al. Scientific Reports, 2019, 9: 11363.
[35] Peng J, Ye L, Shen T, et al. Transactions of the ASABE, 2018, 61(3): 821.
[36] Liu F, Shen T, Wang J, et al. Transactions of the ASABE, 2019, 62(1): 123.
[37] Liu X, Feng X, He Y. Renewable Energy, 2019, 143A(12): 176.
[38] Manzoor S, Ugena L, Tornero-Lopez J, et al. Talanta, 2016, 155: 101.
[39] Prochazka D, Mazura M, Samek O, et al. Spectrochimica Acta Part B: Atomic Spectroscopy, 2018, 139: 6.
[40] Xu W, Chen S, Tan Y, et al. Journal of Analytical Atomic Spectrometry, 2020, 35(8): 1641.

Application Progress of Artificial Neural Network in Laser-Induced Breakdown Spectral Data Analysis

ZHAO Wen-ya[1,2], MIN Hong[2], LIU Shu[2*], AN Ya-rui[1*], YU Jin[3]
1. College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China
2. Technical Center for Industrial Product and Raw Material Inspection and Testing, Shanghai Customs, Shanghai 200135, China
3. School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract Laser-induced breakdown spectroscopy (LIBS) has the advantages of real-time, rapid, and multi-element simultaneous detection. It has attracted more and more attention in recent years and played an essential role in online industrial analysis. However, based on the emission spectrum characteristics, LIBS has spectral noise, baseline drift, self-absorption, and overlapping peaks, etc. In addition, spectral stability and reproducibility are poor due to environmental changes, laser energy fluctuations, matrix effects, and samples' surface topography. These result in the nonlinear relationship between spectral information and qualitative and quantitative analysis, limiting the analysis's sensitivity and accuracy. With the gradual improvement of LIBS devices' stability, LIBS spectral data analysis methods are also changing with each new day. Artificial neural networks (ANN) can track and identify nonlinear characteristics, adaptive learning of LIBS spectral characteristics, screening out interference information, and its application in LIBS data analysis has been rapidly developed. This paper introduces the principle, instrument structure, and working process of LIBS and common neural network model in the field of LIBS spectrum analysis, summed up the LIBS in 2015—2020 in combination with the common ANN model in geological, alloy and organic polymer, coal, soil and biological areas such as the specific application. It pointed out that ANN's super ability in the field of data analysis can effectively improve the LIBS analysis accuracy and improve the utilization rate of spectrum data, reducing the spectrum collection and environmental requirements. Given the technical difficulties that still required broken through, ANN's development prospect in LIBS spectral depth information mining, portable special equipment development, technology combination, and other aspects has prospected. LIBS is becoming more and more mature, but data analysis of this technology still has a broad space for development. This review can provide a reference for the application of machine learning in LIBS data analysis.

Keywords Laser-induced breakdown spectroscopy; Artificial neural network; Data analysis; Application

* Corresponding authors
(Received Jul. 15, 2020; accepted Nov. 21, 2020)