

基于卷积神经网络与投票机制的蒲黄炮制品近红外判别方法

陈承武¹, 王天舒^{1*}, 胡孔法¹, 包贝华², 严辉², 杨曦晨³

1. 南京中医药大学人工智能与信息技术学院, 江苏 南京 210029

2. 南京中医药大学药学院, 江苏 南京 210029

3. 南京师范大学计算机与电子信息学院/人工智能学院, 江苏 南京 210023

摘要 蒲黄炭是由香蒲花粉炮制而成, 具有止血、化瘀、通淋等多种功效, 被广泛应用于临床抗血栓, 创面和出血。然而蒲黄炭在炒炭过程中, 常常会出现炭化过轻或者炭化过重的现象, 从而出现不同炭化程度的蒲黄炭药品, 主要为轻度炭化、标准炭化与重度炭化三种不同的蒲黄炭药品。由于炭化程度不同, 蒲黄炭的凝血效果优劣不等, 其中标准炭化的蒲黄炭药品药效最优。目前, 鉴别蒲黄炭药品的方法多为人工凭借肉眼与经验进行判别。基于人工的蒲黄炭药品判别方法判别效率低, 受主观因素影响大, 判别结果不稳定, 难以区分出标准炭化的蒲黄炭。为有效地对不同炭化程度的蒲黄炭进行识别, 提出一种基于卷积神经网络与投票机制的蒲黄炮制品近红外判别方法。该方法创新性地结合深度学习与机器学习算法, 有效利用卷积神经网络强大表征提取能力的同时通过投票决策提升算法模型的泛化能力与鲁棒性。首先通过近红外光谱技术获取蒲黄炭的近红外光谱, 并通过卷积神经网络分别提取样本经过四种预处理方法所得光谱图的高阶特征, 并计算预测结果。按照样本准确率与损失值为四种预处理方法分配相应权重得到蒲黄炮制品预测模型。该模型将所得到的四种预测结果结合权重共同投票出样本的最终结果, 从而鉴别出蒲黄炭的炭化程度。实验结果表明所提方法可以有效判别蒲黄炮制品的炭化程度。当训练集所占样本比例为80%时, 预测准确率达到95.4%。所提方法与传统卷积神经网络方法、线性判别分析方法以及标准正太变量变换-线性判别分析方法相比预测准确率分别提高8.6%, 4.3%和2.6%。同时, 所提方法具有一定的稳定性, 当训练集所占样本比例大于70%时, 测试准确率高于90%; 当训练集比例仅占10%时, 预测准确性仍然能够达到约80%。

关键词 蒲黄炭; 卷积神经网络; 投票机制; 近红外光谱

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)11-3361-07

引言

蒲黄(cattail pollen)系香蒲科香蒲属, 为东方香蒲(*typha orientalis presl*)或同属植物的干燥花粉。蒲黄是一味著名的中药, 具有止血、化瘀及通淋等多种功效。制炭是将净选或切制后的药材经高温处理, 使药材外部炭化、内部保留固有性能的一种中药炮制方法。生蒲黄经过炭化制成的蒲黄炭(carbonized typhae pollen, CTP)具有明显的抗出血作用, 被广泛应用于临床抗血栓, 治疗创面和出血。然而炒炭程度对其止血作用具有重要影响。蒲黄炒炭过程极难掌握, 炭化过程中容易出现“不及”或“太过”现象, 从而产生轻度炭化、标

准炭化与重度炭化三种不同的蒲黄炭药品。这三种蒲黄炭化程度不同, 凝血效果优劣不等。当前判别蒲黄炭的方法多为凭借人工经验进行肉眼观测。然而人工判别的方法, 判别效率低, 受主观影响大, 判别结果不稳定。因此, 亟需提出一种高效、客观且准确的蒲黄炭判别方法。

近红外光谱是介于可见光和中红外之间, 波长范围为700~2500 nm的电磁辐射波。通过扫描样品的近红外光谱, 可以得到样品中有机分子含氢基团的特征信息。近红外光谱分析技术具有快速、无损与无污染等优点, 被广泛应用于食品^[1-2]、木材^[3]与药物分析^[4]等领域。Xie等^[1]使用傅里叶变换红外光谱与随机森林方法准确分类不同地区的香菇。刘亚超等^[2]运用近红外光谱技术提供了一种大米直链淀粉含量的

收稿日期: 2020-10-12, 修订日期: 2022-09-13

基金项目: 国家自然科学基金项目(82204770, 62101268), 江苏省科技计划项目青年基金项目(BK20210696), 未来网络科研基金项目(FN-SRFP-2021-ZD-24), 南京市科技计划项目(201812021)资助

作者简介: 陈承武, 2000年生, 南京中医药大学人工智能与信息技术学院本科生 e-mail: chenchengwu@njucm.edu.cn

* 通讯作者 e-mail: wangtianshu@njucm.edu.cn

快速检测方法。Unger 等^[5]利用近红外光谱技术结合分子建模数据验证了环氧树脂体系固化过程中反应基团的演化。Kartakoullis 等^[6]使用近红外光谱结合偏最小二乘回归与随机森林回归算法预测咸肉糜的脂肪、水分和蛋白质含量。Cui 等^[7]使用近红外定量分析技术鉴别区分玉米单倍体种子与二倍体种子。Mandrile 等^[8]分别利用漫反射近红外光谱、衰减全反射中红外光谱技术和电感耦合等离子体发射光谱技术鉴别了可可豆壳产地。当前近红外的判别方法大多使用传统的机器学习方法。传统的一般采用手工特征判别。而手工特征判别耗时费力,且需要研究人员熟悉特定应用领域。

深度学习是机器学习的进一步发展,通过模拟人脑进行分析学习并建立神经网络,模仿人脑的机制解释数据。深度学习是一种使计算机自动学习并提取模式特征的方法,能够降低人为设计特征产生的不完备性。卷积神经网络(convolutional neural network, CNN)是深度学习的代表算法之一。CNN 是一种深度前馈人工神经网络,具有表征学习能力,已广泛应用于图像识别^[9]、物体识别^[10]与自然语言处理^[11]等领域。Nogovitsyn 等^[12]使用深度卷积神经网络得出海马体积的自动分割算法。实验结果表明基于 CNN 的海马分割算法具有高度的稳定性与准确性。Wartini 等^[13]提出一种基于卷积神经网络的土壤特性分析算法,并获得了较高的准确率。Xu 等^[14]运用深度卷积神经网络对组织病理学图像中的上皮与基质两块区域进行了分离与分类。以上研究表明通过 CNN 进行中药分析与识别具有可行性。

因此,本文提出一种基于 CNN 与投票机制的蒲黄炮制品近红外判别方法。该方法首先通过近红外光谱技术获取蒲黄炭的光谱数据,得到原始样本空间 S_1 。分别运用标准正态变量变换(standard normal variate transform, SNV),一阶差分,Min-max 标准化三种不同的预处理方法对 S_1 进行数据预处理,得到样本空间 S_2 , S_3 与 S_4 。之后对这四个样本空间分别运用 CNN 进行样本判定,分离不同炭化程度的蒲黄炭样本,并按照判别准确率进行降序排序。依序为四种不同的样本空间赋予权重 0.5, 0.5, 0.4 和 0.3。接下来,采用 CNN 分别提取四组样本空间特征,进行特征学习。在得到四组样本的判别结果后,结合权重投票出蒲黄炭最终的炭化程度,从而实现样本的正确判定。

1 实验部分

1.1 样品制备

东方香蒲花粉采自江苏通州,凭证标本号为 20180820。花粉经南京中医药大学药学院严辉副教授鉴定,保存于中药化学实验室。

轻度炭化,标准炭化与重度炭化蒲黄炭通过炮制时间与蒲黄炭的传统饮片性状确定^[15]。所有的蒲黄样品为同一批次,共 166 份。每份取 100 g 花粉样品,在 (270 ± 10) °C 下,用简单的炒制方法对花粉进行炭化处理。将样品划分为 3 份,分别控制其炮制时间为小于 7 min,等于 7 min,大于 7 min,并将炮制后的样品结合其饮片性状确定其炭化程度。

1.2 近红外光谱数据采集

将炮制后的三份样品置于 40 °C 下干燥 3 h,减小因水分引起的光谱干扰。实验采用傅里叶近红外光谱仪(Antaris II, Thermo, US)。光谱分辨率为 8 cm^{-1} ,扫描累积数为 32,光谱范围为 $4\ 000 \sim 12\ 000 \text{ cm}^{-1}$ 。在 22~27 °C 温度范围和 60% 相对湿度的环境下,将每份样品粉末放入石英样品杯中,混合均匀,并扫描空气作为参考。每份粉末样品在相同条件下进行三次扫描,生成平均光谱。

1.3 基于 CNN 与投票机制的蒲黄炮制品判别方法

我们运用 One-Hot 编码为每个样本设置标签。若某样本 s 为轻度炭化蒲黄炭,则其标签 $l_s = [0, 0, 1]$;若某样本 s 为标准炭化蒲黄炭,则其标签 $l_s = [0, 1, 0]$;若某样本 s 为重度炭化蒲黄炭,则其标签 $l_s = [1, 0, 0]$ 。

本文提出的基于 CNN 与投票机制的蒲黄炮制品判别模型如图 1 所示。将样本空间 S 划分为训练集 S_{tr} 与测试集 S_{te} ,其中训练集 S_{tr} 用来对模型进行训练,测试集 S_{te} 用来评价模型。首先对训练集 S_{tr} 中所有样品的原始光谱进行四种不同的预处理:(1)保持不变;(2)SNV 处理;(3)一阶差分处理;(4)Min_max 处理。经四种预处理方法处理后的数据集分别记为: S_{tr1} , S_{tr2} , S_{tr3} 与 S_{tr4} 。

接下来,将 S_{tr1} , S_{tr2} , S_{tr3} 与 S_{tr4} 分别输入至 CNN 深度学习网络进行 CNN 模型训练与预测。以对不同预处理方法的性能进行评判。网络模型包括一层一维卷积池化层、一层二维卷积池化层和一层全连接层。一维卷积池化层将输入的一维向量转化成二维矩阵。一维卷积池化层包括一维卷积操作、激活操作与池化操作。其中一维卷积操作卷积核的数目为 32 个,大小为 10×1 ,卷积的步长为 6。并由修正线性单元(rectified liner unit, ReLU)完成激活操作,使神经网络中的神经元具有稀疏激活性。池化操作使用平均池化模型,每次取 10×1 的池化窗口,步长设为 2。之后,将经过一维卷积池化层得到的二维矩阵输入至二维卷积池化层,并经过二维卷积操作、激活操作与二维池化操作,将其转化成多个二维矩阵。二维卷积操作中卷积核数目为 64 个,大小为 10×32 。将二维卷积池化层得到的二维矩阵输入至全连接层,输出一维高阶向量。本文中卷积神经网络模型采用学习率衰减机制,学习率初始值设为 0.003,衰减指数为 $1/e$ 。在网络训练的过程中随时间逐步衰减学习率进行动态调整。每层权重初始值服从标准差为 0.1 的零均值高斯分布。

对训练集 S_{tr} 进行 CNN 的 train-test 过程,设 S_{tr1} , S_{tr2} , S_{tr3} 与 S_{tr4} 的预测准确率百分比分别为 a_1 , a_2 , a_3 和 a_4 。接下来,对四种不同预处理方法进行权重分配。表 1 给出四种预处理方法所对应数据集的 CNN 预测准确率与权重分配情况。其中,保持不变的数据集得到的准确率为 72%;经 SNV 预处理的数据集 S_{tr2} 得到的准确率为 88%;经一阶差分预处理的数据集 S_{tr3} 得到的准确率为 92%;经 Min_max 预处理的数据集 S_{tr4} 到的准确率为 76%。我们将准确率最高的两种预处理方法的权重值分配为 0.5;准确率第三高的预处理方法的权重值分配为 0.4;准确率最低的预处理方法的权重值分配为 0.3。

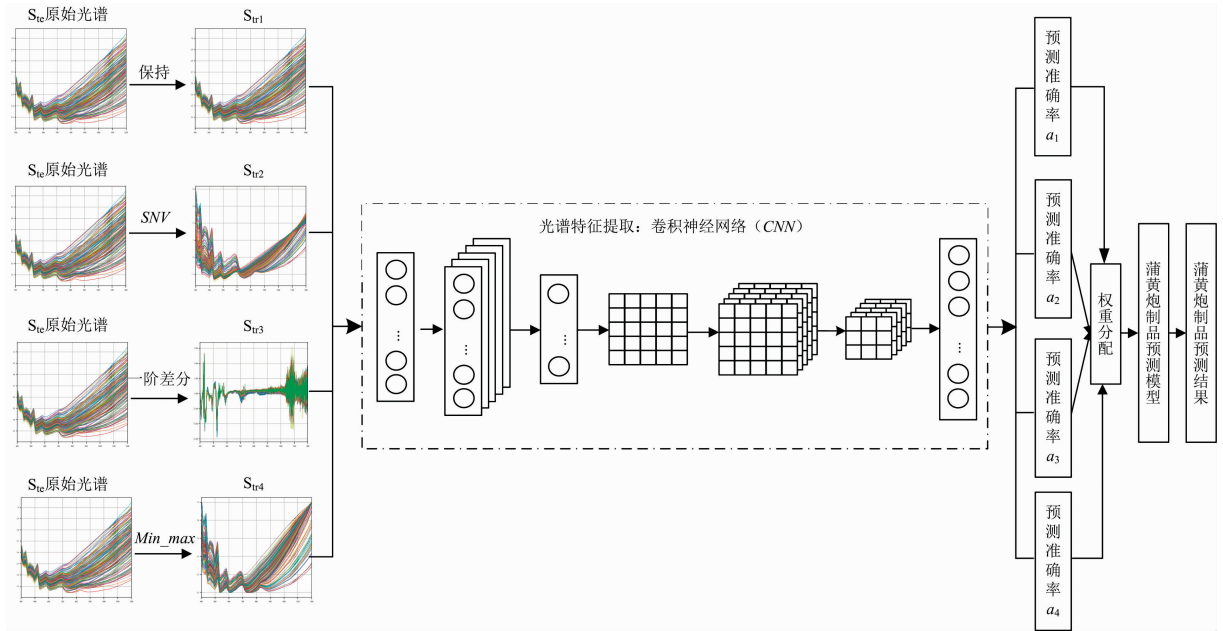


图 1 基于 CNN 与投票机制的蒲黄炮制品判别模型

Fig. 1 Identification model of processed products of Pollen Typhae based on CNN and voting mechanism

表 1 不同预处理方法的权重分配

Table 1 Weight distribution of different preprocessing methods

预处理方法	数据集	CNN 预测准确率 a	权重 ω
保持不变	S_{tr1}	$a_1 = 72\%$	$\omega_1 = 0.3$
SNV	S_{tr2}	$a_2 = 88\%$	$\omega_2 = 0.5$
一阶差分	S_{tr3}	$a_3 = 92\%$	$\omega_3 = 0.5$
Min_max	S_{tr4}	$a_4 = 76\%$	$\omega_4 = 0.4$

设测试集 S_{te} 中某个样本为 s_{te} 。首先通过四种不同预处理方法对该样本的原始光谱分别进行处理，并将计算结果分别输入至 CNN 预测模型。将 CNN 预测模型得到的预测结果分别记为： y_{te_1} 、 y_{te_2} 、 y_{te_3} 和 y_{te_4} 。其中，第 i 个预处理方法的预测结果 $y_{te_i} = [y_{te_{i1}}, y_{te_{i2}}, y_{te_{i3}}]$ 。例如，若第 i 个预处理方法的预测结果为轻度炭化蒲黄，则 $y_{te_i} = [0, 0, 1]$ ；若预测结果为标准炭化蒲黄，则 $y_{te_i} = [0, 1, 0]$ ；若预测结果为重度炭化蒲黄，则 $y_{te_i} = [1, 0, 0]$ 。

对样本 s_{te} 进行投票，计算 s_{te} 在三种不同炭化程度上所得的投票分值，记为 v_j (轻度炭化为 v_1 ，标准炭化为 v_2 ，重度炭化为 v_3)。投票分值 v_j 的计算方法如式(1)所示，其中 i 表示第 i 种预处理方法， j 表示第 j 种炭化程度。

$$v_j = \sum_{i=1}^4 \omega_i y_{te_{ij}} (4 - j) \quad (1)$$

对三种炭化程度的 v_j 值进行比较，得到 s_{te} 的最终预测结果。若 v_1 最大，则预测 s_{te} 为轻度炭化样品；若 v_2 最大，则 s_{te} 为标准炭化样品；若 v_3 最大，则 s_{te} 为重度炭化样品。

2 结果与讨论

2.1 光谱分析

三种不同炭化程度的蒲黄炭样品的近红外原始光谱数据

如图 2 所示。绿色曲线表示轻度蒲黄炭(Light degree CTP)，蓝色曲线表示标准蒲黄炭(Standard degree CTP)，红色曲线表示重度蒲黄炭(Heavy degree CTP)。由图 2 可知，近红外光谱的谱带重叠严重，整体近乎一致，吸收峰也几乎处于同一位置，没有显著的差异，难以用肉眼区分三种蒲黄炭样品的近红外光谱。因此我们对光谱进行四种不同的预处理方式，从而加大不同炭化程度的蒲黄炭近红外光谱的差异性，也提高了准确判定蒲黄炭的炭化程度可能性。

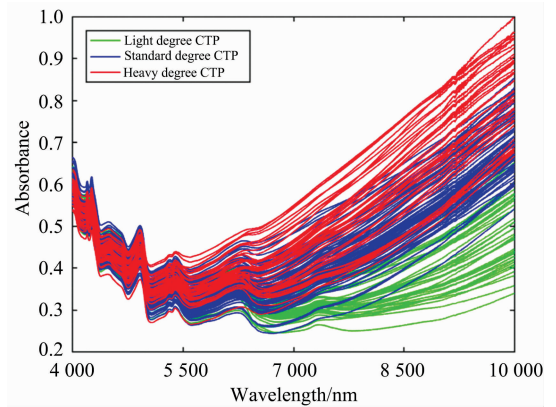


图 2 原始光谱

Fig. 2 Raw spectral

2.2 CNN 模型分析

基于 CNN 与投票机制的蒲黄炮制品判别模型中，需要将四种不同的预处理方法的计算结果输入至 CNN 卷积神经网络。预处理后的一维数据通过 CNN 的一维卷积池化层计算后，转化为 139×32 的二维矩阵。其中，一维卷积池化层的卷积核(Filter 1~16)为 32 个，其部分卷积核(Filter 1~16)如图 3 所示。

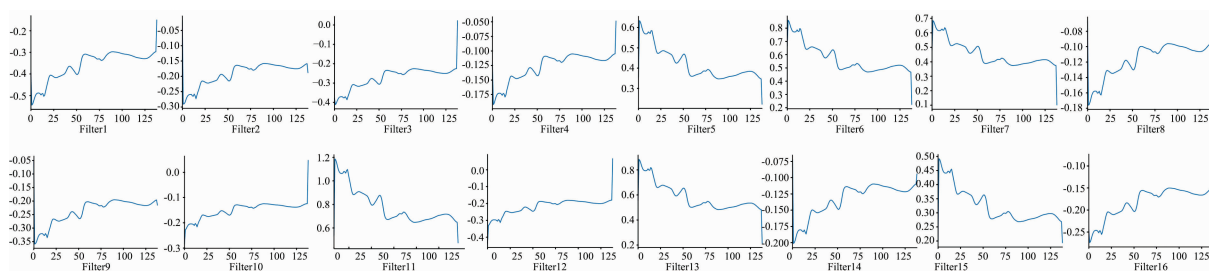


图 3 一维卷积池化层的部分卷积核

Fig. 3 Partial convolution kernel of one-dimensional convolution pool

在经过一维卷积池化操作后，二维矩阵数据经过二维卷积池化层转换为 24×64 的二维矩阵。二维卷积池化层中卷积核为 64 个，其部分卷积核(Filter 1~16)如图 4 所示。

进行模型训练所得到的特征向量。其中，红色线条对应蒲黄炭轻度炭化样本；绿色线条对应蒲黄炭标准炭化样本；蓝色线条对应蒲黄炭重度炭化样本。由图 5 可知，通过 CNN 得到的三种样品特征向量区分较为明显，重叠样本少。

图 5 展示训练集 S_{tr} 经四种预处理方法后分别输入 CNN

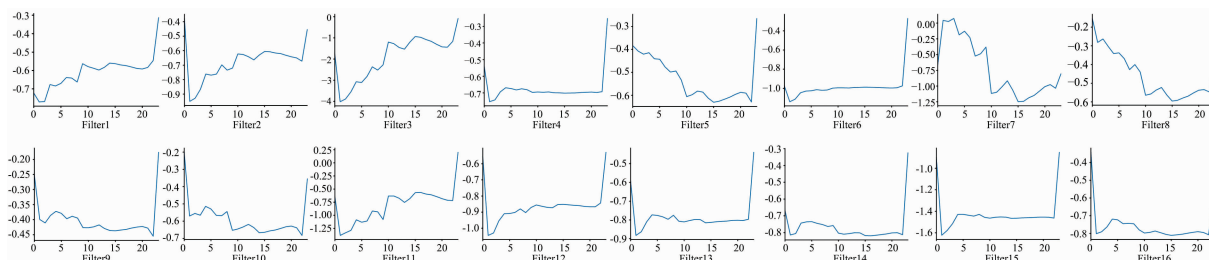


图 4 二维卷积池化层部分卷积核

Fig. 4 Partial convolution kernel of two-dimensional convolution pool

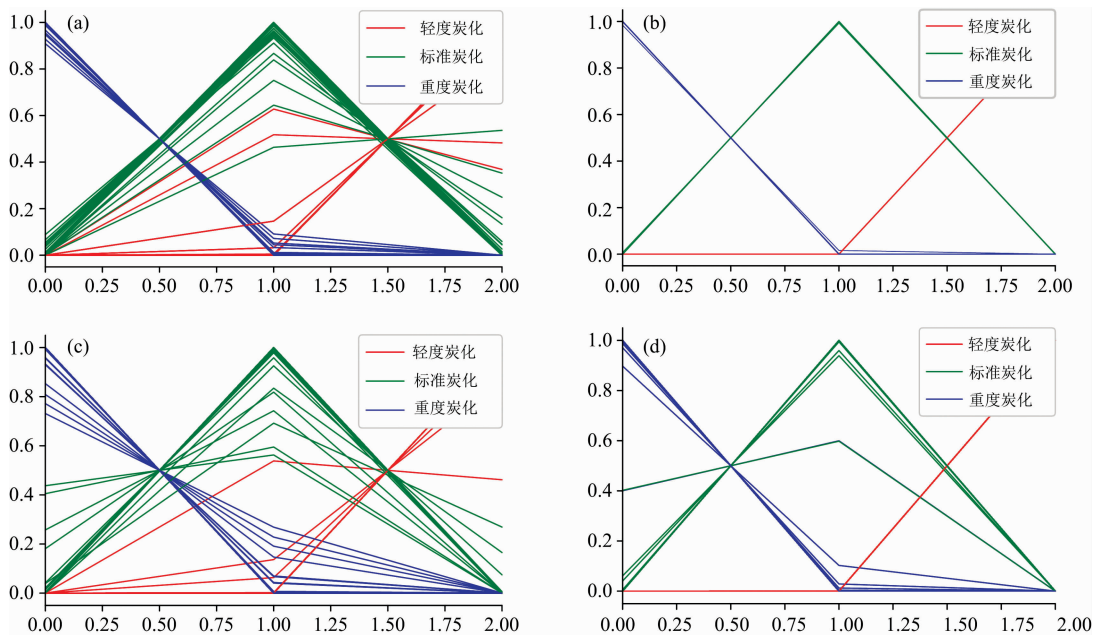


图 5 四种预处理方法的 CNN 特征向量

(a): 保持不变预处理 CNN 特征向量；(b): SNV 预处理 CNN 特征向量；
(c): 一阶差分预处理 CNN 特征向量；(d): Min_max 预处理 CNN 特征向量

Fig. 5 CNN eigenvectors of four pre-processing methods

(a): Unchanged preprocesses CNN eigenvectors; (b): SNV preprocesses CNN eigenvector;
(c): First-order difference preprocesses CNN eigenvectors; (d): Min_max preprocesses CNN eigenvectors

图 6 给出了四种预处理方式在训练 CNN 模型时测试准确率随迭代次数的变化情况, 由图 6 可看出, 测试准确率随迭代次数增加而迅速上升。为避免因随机抽样导致的实验误差, CNN 的 train-test 过程重复 50 次, 经实验发现四种预处理方式的测试准确率最终均可稳定在 80% 以上。

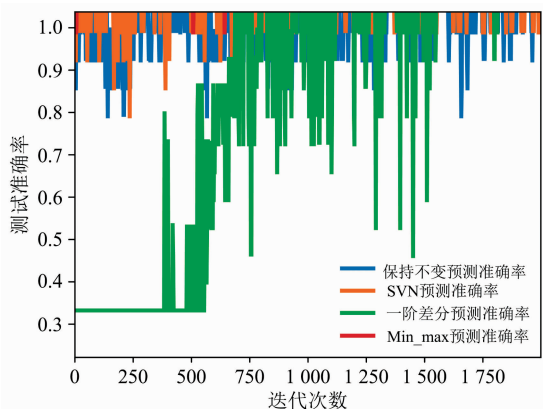


图 6 四种预处理方法的 CNN 测试准确率

Fig. 6 CNN test accuracy of four pre-processing methods

在蒲黄炮制品识别模型中, 我们采用交叉熵损失函数判别模型预测效果。从图 7 可以看出, 四种预处理方式在训练 CNN 模型时, 随着迭代次数的增加, 损失值始终以下降趋势不断递减。

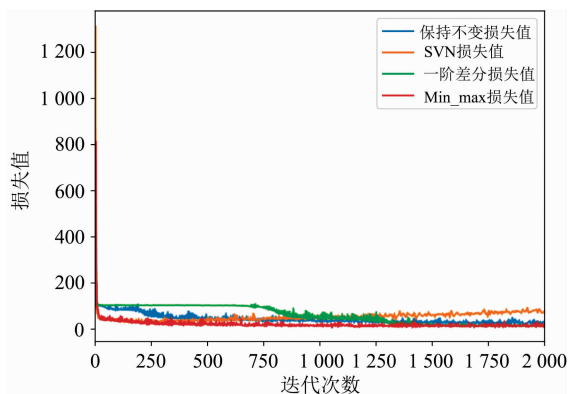


图 7 四种预处理方法的 CNN 交叉熵损失值

Fig. 7 Cross entropy loss of CNN based on four pre-processing methods

2.3 方法性能分析

为了评估本方法的性能, 我们将该方法与传统卷积神经网络 CNN, 线性判别分析方法 (linear discriminant analysis, LDA) 以及标准正交变量变换-线性判别分析方法 (standard normal variable-linear discriminant analysis, SNV-LDA) 进行对比。实验过程中, 训练集所占比例为 80%, 测试集所占比例为 20%, train-test 过程重复 50 次。四种方法的预测准确率对比如图 8 所示。由图可知, 本方法的准确率最高, 分别比 CNN, LDA 和 SNV-LDA 提高 8.6%, 4.3% 和 2.6%。

本方法需要通过训练预测模型进行训练, 而训练集的规模可以影响预测准确性。为了分析本方法的稳定性, 将训

练集所占样本比例分别设为 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% 和 90%。不同比例下的 train-test 过程重复 50 次。稳定性测试结果如图 9 所示。由图 9 可知, 当训练集所占比例高于 70% 时, 本方法的预测准确率高于 90%。同时, 当训练集仅占比 10% 时, 所提方法的预测准确率可以维持在 80% 左右。稳定性测试实验结果表明本方法具有较高的稳定性。

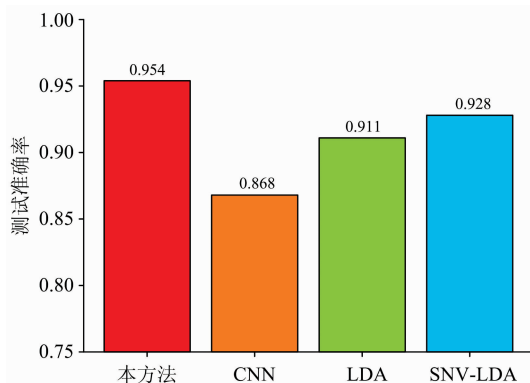


图 8 本方法与 CNN, LDA, SNV-LDA 预测准确率对比

Fig. 8 Comparison of the proposed method with CNN, LDA and SNV-LDA in terms of test accuracy

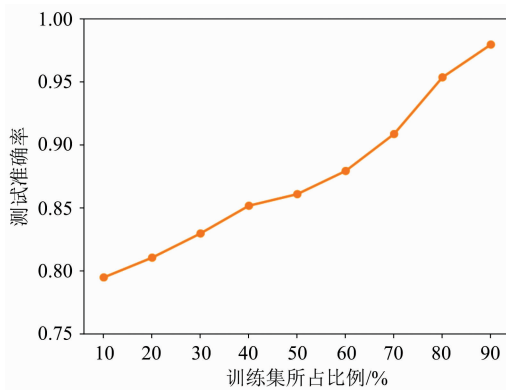


图 9 不同训练比例预测准确率

Fig. 9 Test accuracy of different training set proportions

为综合评价本方法的性能, 将 35 份外部蒲黄炮制样品输入至所建立的蒲黄炮制品预测模型进行判别。35 份蒲黄炮制样品中, 仅 2 份样品预测出错, 预测准确率达到 94.29%。由此可见, 本方法适用于蒲黄炮制品近红外判别, 能够满足标准蒲黄炭快速识别的实际应用需求。

3 结论

提出一种基于卷积神经网络与投票机制的蒲黄炮制品近红外判别方法, 高效判别重度炭化、标准炭化、轻度炭化三种不同炭化程度的蒲黄炭。蒲黄炭的原始光谱在经过四种不同的预处理方式后通过卷积神经网络模型预测出四种结果, 结合自身权重后投票出蒲黄炭的炭化程度。实验结果表明, 本工作所提出的蒲黄炮制品预测模型对测试集样本的判别准

确率达到 95.4%，识别准确性高。基于卷积神经网络与投票机制的蒲黄炮制品判别模型结合近红外光谱的实验方法可以高效识别蒲黄炭的炭化程度，为实际鉴别蒲黄炭的炭化程度

与药效和药品鉴定仪器的开发提供了技术支持。

本工作是对蒲黄炮制品快速鉴别的初步研究，后续会收集不同产地，不同采摘时间的样品，进一步研究与完善。

References

- [1] Xie Y, Zhou R R, Xie H L, et al. *International Journal of Biological Macromolecules*, 2019, 122: 1115.
- [2] LIU Ya-chao, LI Yong-yu, PENG Yan-kun, et al(刘亚超, 李永玉, 彭彦昆, 等). *Chinese Journal of Analytical Chemistry(分析化学)*, 2019, 47(5): 785.
- [3] Schimleck L R, Antony F, Mora C, et al. *Holzforschung*, 2020, 74(1): 20.
- [4] Correia R M, Domingos E, Flávia Tosato, et al. *Analytical Methods*, 2018, 10(6): 593.
- [5] Unger R, Braun U, Fankhänel, et al. *Computational Materials Science*, 2019, 161: 223.
- [6] Kartakoullis A, Comaposada J, Cruz-Carrión A, et al. *Food Chemistry*, 2019, 278: 314.
- [7] Cui Y, Ge W, Li J, et al. *Computers and Electronics in Agriculture*, 2019, 158: 358.
- [8] Mandrile L, Barbosa-Pereira L, Sorensen K M, et al. *Food Chemistry*, 2019, 292: 47.
- [9] Krizhevsky A, Sutskever I, Hinton G E, et al. *Communications of the ACM*, 2017, 60(6): 84.
- [10] Bapu J J, Florinabel D J, Robinson Y H, et al. *Earth Science Informatics*, 2019, 12(4): 525.
- [11] Bacchi S, Oakden-Rayner L, Zerner T, et al. *Stroke*, 2019, 50(3): 758.
- [12] Nogovitsyn N, Souza R, Muller M, et al. *NeuroImage*, 2019, 197: 589.
- [13] Wartini N, Budiman M, Maryam M, et al. *Geoderma*, 2019, 352: 251.
- [14] Xu J, Luo X, Wang G, et al. *Neurocomputing*, 2016, 191: 214.
- [15] Gao M, Bao B, Cao Y, et al. *Molecules*, 2019, 24: 128.

Identification Method of Pollen Typhae Processed Products Based on Convolutional Neural Network and Voting Mechanism

CHEN Cheng-wu¹, WANG Tian-shu^{1*}, HU Kong-fa¹, BAO Bei-hua², YAN Hui², YANG Xi-chen³

1. College of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing 210029, China
2. College of Pharmacy, Nanjing University of Chinese Medicine, Nanjing 210029, China
3. School of Computer and Electronic Information/School of Artificial Intelligence, Nanjing Normal University, Nanjing 210023, China

Abstract Carbonized Typhae Pollen (CTP) is processed by Pollen Typhae. It has various effects such as hemostasis, removing blood stasis and treating stranguria. It has been widely used in clinical anti-thrombosis, wounds and bleeding. However, in the process of CTP, the carbonization is often too light or too heavy, resulting in different degrees of CTP, mainly for light carbonization, standard carbonization and heavy carbonization CTP. Due to the different degrees of carbonization, the coagulation effect of CTP is different. The standard CTP has the best effect. At present, the identification of CTP mainly relies on eyes and experience. The manual method is challenging to distinguish the standard CTP because it is inefficient, volitional, and unstable. Therefore, to effectively identify CTP with different degrees of carbonization, a near-infrared identification method based on Convolutional Neural Network (CNN) and the voting mechanism is proposed. This method innovatively combines deep learning and machine learning algorithms, effectively utilizes the powerful representation extraction ability of CNN, and applies voting decisions to improve the generalization ability and robustness of the prediction model. The near-infrared spectrum of CTP is firstly obtained. Then the high-order features of the spectrum processed by four different pre-processing methods are extracted by CNN. Next, the prediction results are calculated. The weights of four pre-processing methods are allocated according to the accuracy and loss to get the prediction model. Finally, the model combines the four prediction results with the weights to identify the CTP with different degrees of carbonization. The experimental results show that the proposed method can effectively distinguish the CTP with different degrees of carbonization. When the training set occupies 80%, the test accuracy is up to 95.4%. Compared with CNN, Linear Discriminant Analysis (LDA) and Standard Normal Variable (SNV)-LDA, the proposed method improves the prediction accuracy by 8.6%, 4.3% and 2.6%, respectively. At the same time, the proposed method is

robust. When the proportion of the training set occupies more than 70%, the test accuracy is higher than 90%. When the proportion of the training set only occupies 10%, the prediction accuracy can still reach about 80%.

Keywords Carbonized typhae pollen; Convolutional neural network; Voting mechanism; Near infrared spectrum

(Received Oct. 12, 2020; accepted Sep. 13, 2022)

* Corresponding author