

基于电子舌和电子眼信息融合的贝母品种快速辨识方法

马景余, 孙涛, 王彦荣, 李鑫, 曾婉晴, 王志强

A Fast Identification Method for *Fritillaria* Varieties Based on the Fusion of Electronic Tongue and Electronic Eye Information

MA Jingyu, SUN Tao, WANG Yanrong, LI Xin, ZENG Wanqing, and WANG Zhiqiang

在线阅读 View online: <https://doi.org/10.13386/j.issn1002-0306.2024020161>

您可能感兴趣的其他文章

Articles you may be interested in

基于电子鼻与电子舌技术评价氯吡啶对甜瓜品质的影响

Effects of Forchlorfenuron on Quality of Melon Based on Electronic Nose and Electronic Tongue

食品工业科技. 2019, 40(14): 24-30 <https://doi.org/10.13386/j.issn1002-0306.2019.14.004>

基于电子舌对不同品种鲜莲子滋味成分的分析

Analysis of Taste Components in Different Varieties of Fresh Lotus Seeds Based on Electronic Tongue

食品工业科技. 2023, 44(24): 319-326 <https://doi.org/10.13386/j.issn1002-0306.2023020291>

基于电子舌评价不同品种番茄制备番茄酱的滋味品质

Evaluation of Taste Quality of Tomato Paste Prepared from Different Tomato Varieties by Electronic Tongue

食品工业科技. 2019, 40(19): 209-215 <https://doi.org/10.13386/j.issn1002-0306.2019.19.035>

基于电子舌和EEMD-WOA-LSSVM模型的红酒贮藏年限区分

Identification of Red Wine Storage Years based on Electronic Tongue and EEMD-WOA-LSSVM Model

食品工业科技. 2021, 42(19): 275-282 <https://doi.org/10.13386/j.issn1002-0306.2020120105>

基于电子舌和电子鼻结合氨基酸分析鱼香肉丝调料风味的差异

The Flavour Difference of Yu-shiang Shredded Pork Seasoning Revealed by Electronic Nose, Electronic Tongue and Automatic Amino Acid Analyzer

食品工业科技. 2022, 43(9): 48-55 <https://doi.org/10.13386/j.issn1002-0306.2021070282>

基于灰色关联度法构建西式香肠钠盐含量与电子舌的相关性

The Correlation Between Sodium Content of Western Sausage and Electronic Tongue Established Based on Grey Correlation Method

食品工业科技. 2021, 42(12): 262-268 <https://doi.org/10.13386/j.issn1002-0306.2020080134>



关注微信公众号, 获得更多资讯信息

马景余, 孙涛, 王彦荣, 等. 基于电子舌和电子眼信息融合的贝母品种快速辨识方法 [J]. 食品工业科技, 2024, 45(18): 9–18. doi: 10.13386/j.issn1002-0306.2024020161

MA Jingyu, SUN Tao, WANG Yanrong, et al. A Fast Identification Method for *Fritillaria* Varieties Based on the Fusion of Electronic Tongue and Electronic Eye Information[J]. Science and Technology of Food Industry, 2024, 45(18): 9–18. (in Chinese with English abstract). doi: 10.13386/j.issn1002-0306.2024020161

·特邀主编专栏—人工智能在食品工业中的应用 (客座主编: 闵巍庆、郭志明、朱金林)·

基于电子舌和电子眼信息融合的贝母品种快速辨识方法

马景余, 孙涛, 王彦荣, 李鑫, 曾婉晴, 王志强*
(山东理工大学计算机科学与技术学院, 山东淄博 255049)

摘要: 贝母是一种应用广泛的中药材, 其来源复杂, 品种繁多, 不同品种之间具有相似的外形特征, 传统方法难以分辨。为实现贝母品种的快速、客观鉴别, 本文提出了一种基于电子舌和电子眼结合深度学习模型的贝母品种快速辨识方法。分别使用电子舌、电子眼采集不同品种贝母的味觉指纹信息以及视觉图像信息。针对电子舌采集信号, 采用基于因果注意力机制改进的 Transformer 编码器提取信号中的时间序列特征, 提高对局部特征的提取能力。针对电子眼采集图像, 采用基于坐标注意力机制改进的 ShuffleNetV2 网络提取图像不同区域的形态特征并抑制其背景噪声。提出一种特征加权融合分类模块, 对电子舌和电子眼提取的特征信息进行自适应加权融合, 并实现对融合特征的分类识别。结果表明, 基于信息融合方法相较于单独采用电子舌和电子眼具有更好的分类效果, 其测试准确率达到 98.4%。本研究为贝母品种的快速鉴别提供了一种新方法, 并可为其它中药材的品种分类和溯源分析提供研究思路。

关键词: 电子舌, 电子眼, ShuffleNetV2, Transformer, 注意力机制, 贝母

中图分类号: TS201.1 文献标识码: A 文章编号: 1002-0306(2024)18-0009-10

DOI: 10.13386/j.issn1002-0306.2024020161



本文网刊:

A Fast Identification Method for *Fritillaria* Varieties Based on the Fusion of Electronic Tongue and Electronic Eye Information

MA Jingyu, SUN Tao, WANG Yanrong, LI Xin, ZENG Wanqing, WANG Zhiqiang*

(School of Computer Science and Technology, Shandong University of Science and Technology, Zibo 255049, China)

Abstract: *Fritillaria* is a widely used traditional Chinese medicine, with a complex source and a wide variety of medicinal materials. Different varieties have similar external characteristics, making it difficult to distinguish using traditional methods. To achieve rapid and objective identification of *Fritillaria* species, this study proposed a method for rapid identification based on electronic tongue and electronic eye combined with a deep learning model. Electronic tongue and electronic eye were utilized to collect gustatory fingerprint and visual image information from different categories of *Fritillaria*, respectively. An enhanced Transformer encoder based on causal attention mechanism was employed to extract time-series features from the ET signals and augment the ability to extract local features. Meanwhile, an improved ShuffleNetV2 network based on coordinate attention mechanism was used to extract morphological features of EE image and suppress background noise. Subsequently, a feature weighted fusion module was presented to adaptively integrate the feature information extracted from both the electronic tongue and electronic eye, and achieve classification and recognition of the fused features. The experimental results indicated that the proposed information fusion method had better classification performance compared to separate usage of electronic tongue and electronic eye, with a testing accuracy of 98.4%. This study provides a novel approach for rapidly identifying *Fritillaria* varieties, which offers research insights into

收稿日期: 2024-02-21

基金项目: 山东省自然科学基金项目 (ZR2022MF330); 教育部科技发展中心产学研创新基金项目 (2018A02010)。

作者简介: 马景余 (1999-), 男, 硕士, 研究方向: 人工智能与智能系统, E-mail: 1317797916@qq.com。

* 通信作者: 王志强 (1977-), 男, 博士, 教授, 研究方向: 人工智能、大数据、嵌入式系统, E-mail: wzq@sdu.edu.cn。

the classification and traceability analysis of other Chinese medicinal materials.

Key words: electronic tongue; electronic eye; ShuffleNetV2; Transformer; attention mechanism; fritillaria

贝母(*Fritillaria*)是一种百合科植物,具有止咳化痰、清热润肺等药用功效,在中医药领域被广泛应用^[1]。不同品种的贝母因其生长环境不同,药效成分、品质和价格存在很大差异,且极易混淆^[2]。传统的贝母品种鉴别方法主要有性状鉴别法和理化分析法。性状鉴别法主要依靠鉴定人员通过嗅觉、味觉、视觉等感官进行分辨,存在主观性强、易受干扰和稳定性差等缺点。理化分析法综合利用物理、化学等多种技术手段对样品进行分析,其中物理分析法主要采用近红外光谱^[3]、太赫兹光谱^[4]等技术对贝母的主要组分进行辨识。化学分析法主要采用固相萃取技术^[5]、薄层色谱法^[6]鉴别法等技术对贝母的化学成分进行鉴别。但此类方法前处理过程复杂,仪器设备体积大、价格昂贵,检测周期长,不适用于现场快速检测。因此,探索一种快速、低成本贝母品种鉴别方法就显得尤为重要。

电子舌和电子眼是一类利用传感器阵列结合模式识别技术对样本的味觉和视觉信息进行分析的仿生学检测方法,具有精确度高、检测速度快、操作简单、成本低廉等优点。电子舌和电子眼已经被广泛应用于中药材识别领域,杨银爱等^[7]采用电子舌结合判别因子分析法(DFA)区分不同品种莲子滋味。张晓等^[8]采用电子眼结合主成分分析法(Principal Component Analysis, PCA)对不同规格的穿心莲药材质量进行评价。但贝母的成分复杂,其品质特征在味觉和视觉上都有所体现,单独使用电子舌或电子眼难以获得完整的样本信息,影响分类效果。近年来,结合不同传感器获取的数据,通过信息融合的方法,已成为中药材品质鉴定和检测的一种有效策略。段金芳等^[9]运用电子舌和电子眼结合主成分分析(PCA)法成功优选山茱萸最佳制备时间。陈佳瑜等^[10]采用电子舌和电子眼结合正交偏最小二乘判别分析法(Orthogonal Partial Least Squares-Discriminant Analysis, OPLS-DA)对茶汤的滋味和色泽等进行综合品质分析。然而这些方法所使用的模式识别技术主要基于传统的机器学习和有限的训练样本,容易出现泛化能力不足和过拟合问题^[11]。

深度学习能够自动对大规模数据进行特征提取,并实现“端到端”的模式识别。在电子舌模式识别领域,Lu等^[12]采用卷积神经网络(CNN)结合电子舌对稻米品质进行定量分析。王首程等^[13]采用电子舌结合密集神经网络(DenseNet)对不同贮藏年限的陈醋进行了区分。然而,电子舌信号具有高维度、大变幅、含噪声等特点,使用传统深度学习模型很难从中提取隐藏的时序特征。近年来,Transformer模型在信号处理领域得到了广泛应用^[14]。其自注意力机制能够充分考虑输入序列中的上下文依赖关系,使

模型能够有效地捕捉全局信息。然而,Transformer对长序列的局部特征不敏感,导致在捕捉长距离依赖关系和处理大规模序列数据时存在一定缺陷^[15]。在电子眼模式识别方面,Gong等^[16]采用电子眼结合多语义交叉网络(TFNet)实现了对枸杞等中药的产地鉴定。王健庆等^[17]采用GoogLeNet对不同品种的中药饮片图像进行分类识别。周丽媛等^[18]采用AlexNet对不同品种的藏药材切片图像进行分类识别。但此类用于图像识别的深度学习模型存在网络深度过大或参数量较多等问题,容易出现梯度消失和梯度爆炸现象^[19]。ShuffleNetV2是一种推理速度较快的轻量级网络^[20]。相较于传统深度模型,其采用的通道混洗机制有效减小了计算量,提高了图像处理速度。但由于ShuffleNetV2网络结构相对简单,导致在捕捉复杂图像特征和抑制噪声干扰方面存在一定不足^[21]。

基于上述分析,本文提出了一种基于电子舌和电子眼结合深度学习模型的贝母品种快速辨识方法。分别采用电子舌和电子眼采集贝母样本的味觉指纹图谱信息和视觉图像信息,提出一种基于因果注意力机制^[22]改进的Transformer编码器(Encoder)对电子舌信号进行特征提取,捕捉输入序列中远距离时序依赖关系和局部特征。采用基于坐标注意力机制^[23]改进的ShuffleNetV2网络提取贝母电子眼图像中的形状和纹理特征,并抑制其噪声干扰。最后设计一种特征加权融合分类模块对电子舌和电子眼提取的关键特征信息进行融合和分类,从而实现了对贝母品种的快速辨别。

1 材料与方法

1.1 材料与仪器

贝母(川贝母、浙贝母、平贝母、伊利贝母和藏贝母) 福曦堂药店提供,密封完好,未超过保质期。

基于虚拟仪器技术的伏安电子舌系统 自行研制;睿鸿 10A-RR500 电子眼 南京南派科技有限公司。

1.2 实验方法

1.2.1 电子舌与电子眼系统 电子舌系统由本实验室自行研发^[24],电子舌结构如图1所示,系统由传感器阵列、信号调理电路、数据采集装置和上位机软件组成。其中传感器阵列由三电极系统组成,包括8个工作电极(铂、金、钛、钯、银、钨、镍和玻碳)、1个Ag/AgCl参比电极和1个铂辅助电极。其工作原理为:在辅助电极和工作电极之间进行大幅度脉冲伏安法(LAPV)扫描,其扫描电位范围为-1~1 V,扫描频率为20 Hz,阶跃电位为5 mV。在LAPV的诱

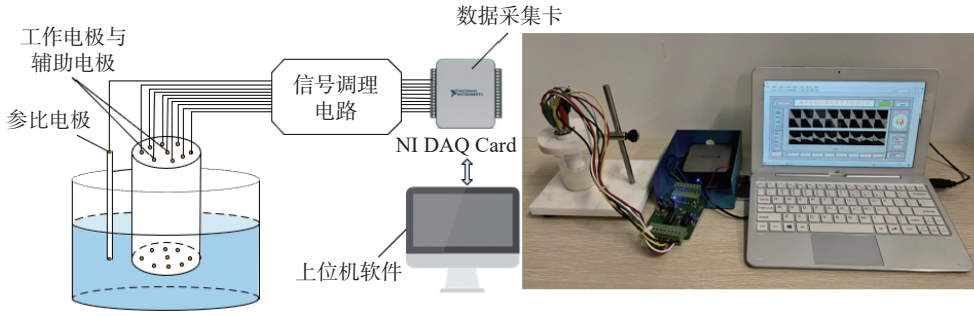


图 1 电子舌系统结构图与实物图

Fig.1 Structural diagram and physical diagram of electronic tongue system

导下,工作电极表面发生氧化还原反应并产生微弱的电流信号。该信号经信号调理电路放大后,由数据采集卡转换成二进制数据,再传输到 LabVIEW 软件进行进一步处理。

电子眼系统结构如图 2 所示,系统由上位机软件、500 万电子目镜、光电镜头、LED 灯及调节器、支架和底盘等五部分组成。其检测过程为:将实验样本放在电子眼底盘上固定,调整电子眼目镜倍数以及支架的高度使得采集的图像呈现清晰的状态为止,使用 55 mm 等效焦距进行拍摄,然后将采集到的数据传到上位机系统中进行数据处理。

1.2.2 样本数据采集 电子舌数据采集时,称取 15 g 贝母样本,经破壁机粉碎后,将样本粉末置入 50 mL 开水中,搅拌均匀,静置 50 min 后,使用滤纸

两次过滤,待样本溶液冷却至室温(25±2) °C 后取上清液转移至 50 mL 烧杯中,采用电子舌进行检测^[25],每一类贝母采集 500 个样本,共计采集 2500 个电子舌数据样本;电子眼数据采集时,将贝母样本置于设备底盘上,电子眼目镜经过矫正后对样本逐个采集图像,每个样本变换位置采集 3 次图像,每种贝母样本共采集 500 张图像,共采集 2500 张图像^[26]。

1.2.3 模式识别方法

1.2.3.1 组合模型结构 本文提出的组合模型结构如图 3 所示。采集到的电子舌一维信号与电子眼二维图像分别进行归一化和图像裁剪等数据预处理操作。然后,电子舌信号输入改进的 Transformer Encoder 模块,采用多头因果自注意力机制实现对输入

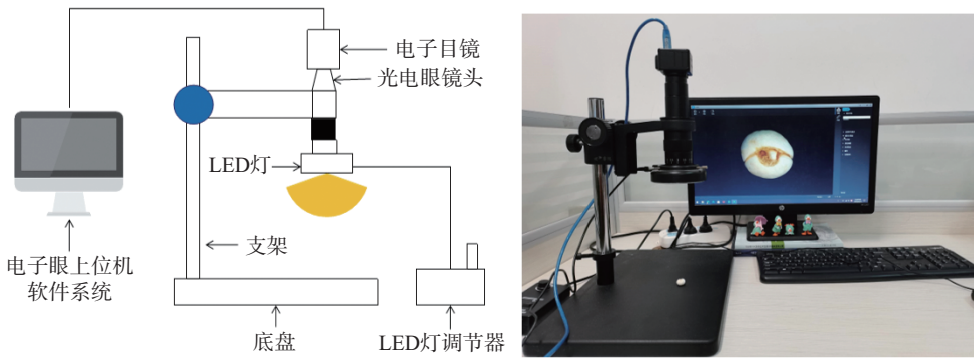


图 2 电子眼系统结构图与实物图

Fig.2 Structure diagram and physical diagram of electronic eye system

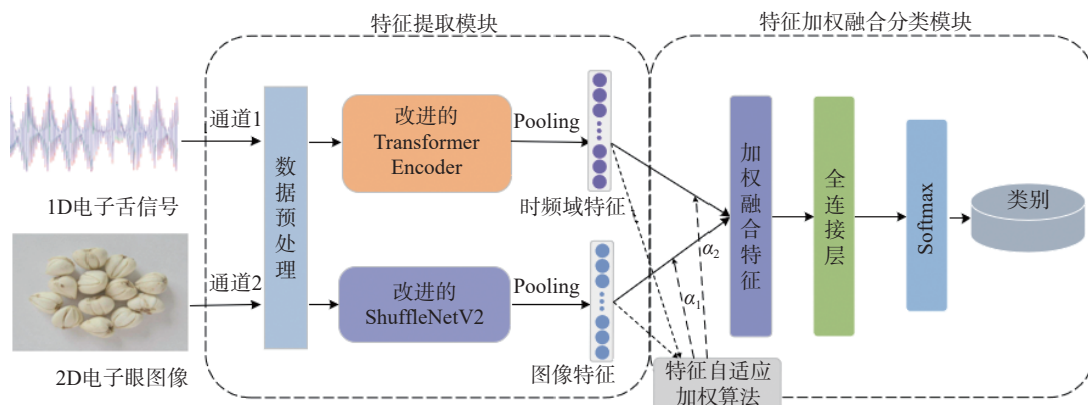


图 3 组合模型结构

Fig.3 Structure of the combination model

信号序列中长距离时序依赖关系和局部特征的有效捕捉。电子眼图像输入改进的 ShuffleNetV2 网络, 利用其坐标注意力机制模块来增强对贝母形状和纹理特征的有效提取。然后, 通过特征自适应加权算法将 Transformer Encoder 和 ShuffleNetV2 网络提取的特征进行加权融合。最后, 经过全连接层降维后, 利用 Softmax 分类器实现对贝母的分类鉴别。

1.2.3.2 改进的 Transformer 编码器 Transformer 是一种用于自然语言处理和其他序列建模任务的深度学习模型架构。传统 Transformer 模型自注意力机制适合学习序列中各时间点之间的依赖关系, 但对于序列的局部特征并不敏感^[27]。根据卷积神经网络具有局部相关性的特点, 引入因果卷积改进 Transformer 自注意力模块, 从而有效捕捉序列中的全局和局部时序依赖关系。

改进的 Transformer Encoder 主要由多头因果自注意力机制和前向传播层组成, 结构如图 4(a) 所示。电子舌信号归一化后经过位置编码和时间编码得到编码序列, 然后通过因果卷积提取信号中的局部特征, 再通过多头因果自注意力机制捕捉其中的全局时序依赖关系。最后通过两个残差、正则化操作以及一个前向传播层得到输出特征向量。

多头因果自注意力机制采用因果卷积替换传统点乘方法, 其结构如图 4(b) 所示。将经过位置编码和时间编码得到编码序列等分为 k 个模块, 每个模块中的序列分别通过三个 1×1 卷积生成 Q, K, V 矩阵。多头因果自注意力机制通过计算 Q, K, V 之间的相互关系, 为每个特征计算一个注意力分数 QK^T , 促使模型关注关键区域, 优化公式如下:

$$\text{Attention}(Q, K, V) = \text{Soft max} \left[\text{mask} \left(\frac{QK^T}{\sqrt{d_k}} \right) \right] V \tag{1}$$

式中 Q, K, V 分别为请求矩阵、主键矩阵和数值矩阵; K^T 为 K 的转置矩阵; d_k 表示请求向量的维度。 QK^T 得到的权重分数刻画的是向量之间的相关关系, 再除以 $\sqrt{d_k}$ 以缓解 Softmax 函数中的梯度消失问题。

$\text{mask}(\cdot)$ 表示遮罩操作, 遮罩将矩阵上三角元素 (不包括对角线) 设置为 $-\infty$ 。通过引入遮罩, 使 Transformer 采用的多头注意力机制能够充分挖掘并保留序列各元素之间全部的相关关系信息, 同时能够抑制噪声的干扰, 使模型聚焦于关键特征信息。

1.2.3.3 改进的 ShuffleNetV2 网络 ShuffleNetV2 是一种轻量级深度学习模型, 在保留通道混洗 (Channel Shuffle) 和分组卷积的基础上, 降低了网络模型的参数量, 提高了推理速度^[28]。但传统 ShuffleNetV2 结构较为简单, 在图像处理中易受背景及噪声的干扰。坐标注意力机制强调了位置信息的重要性, 其不仅能有效建模通道间关系, 还能利用位置信息捕捉图像中的长距离依赖关系, 能够有效地减少背景以及噪声干扰, 从而更好地获取图像的特征信息^[29]。本文将坐标注意力机制引进 ShuffleNetV2, 其结构如图 5 所示。电子眼图像首先通过一个 3×3 卷积调整特征通道数。随后, 通过一个最大池化层对特征图进行空间下采样操作。再经过 3 个改进的 ShuffleNetV2 unit 提取贝母图像视觉特征。最后, 通过一个 1×1 卷积和一个全局平均池化层进行空间维度缩减后得到输出特征图。

改进的 ShuffleNetV2 unit 由下采样单元和基本单元堆叠组成。其中, 下采样单元将特征图输入到两个分支中进行分组卷积, 每个分支均通过 1×1 卷积和 3×3 深度卷积操作, 以减小模型的参数量, 降低过拟合风险。然后, 两个分支通过坐标注意力机制 (CoordAtt) 模块增强特征图中不同通道之间的关联性, 以便更好地捕捉关键特征信息。随后经过通道拼接操作进行信息融合。最后, 通过通道混洗操作确保两分支的特征信息充分融合。基本单元输入特征通过通道分离 (Channel split) 操作将特征通道一分为二, 上侧分支经过 2 个 1×1 卷积和 1 个 3×3 深度卷积操作, 以提取特征图的空间信息, 并保持输入与输出通道相等。然后, 上侧分支经过 CoordAtt 模块后与下侧分支通过通道拼接进行信息融合。最后, 通过通道混洗操作促进通道之间的信息充分流动, 进一步提高特征信息的融合效果。

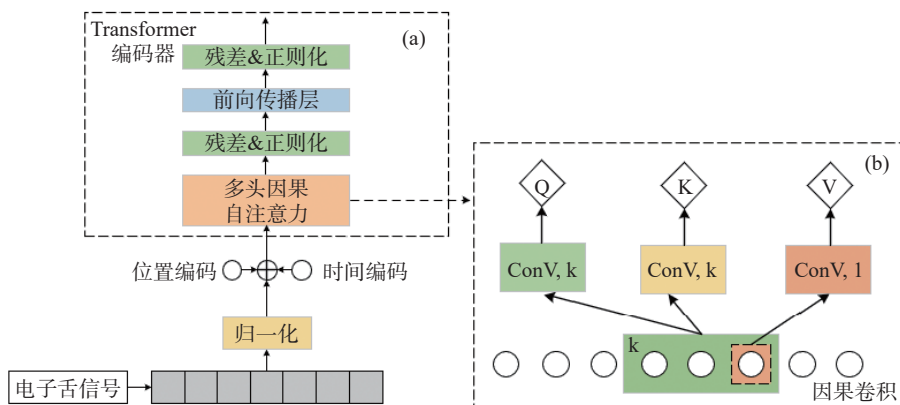


图 4 多头因果注意力机制以及因果卷积过程

Fig.4 Multi head causal attention mechanism and causal convolution process

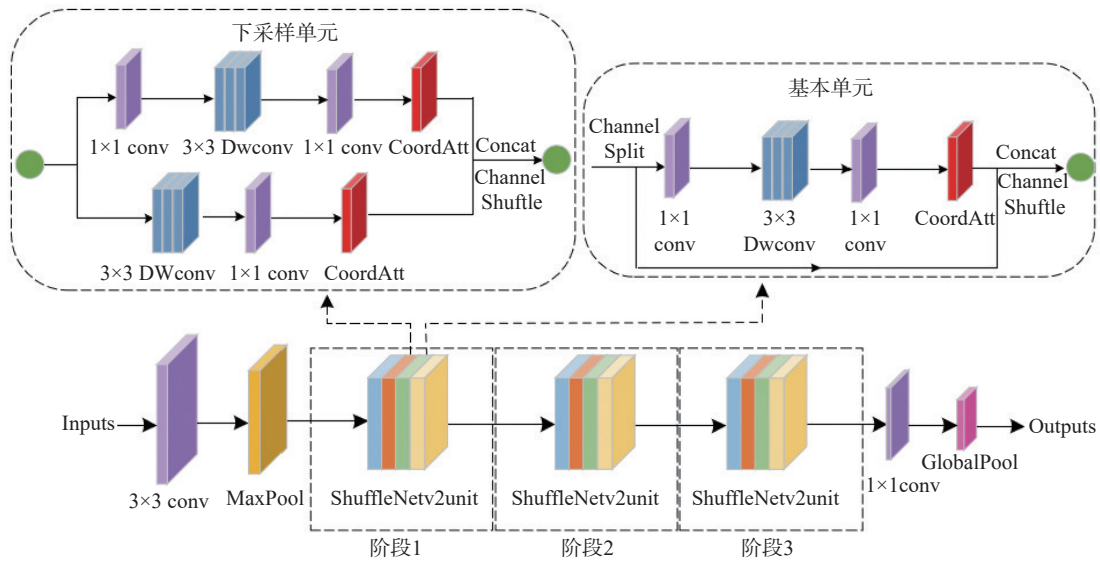


图 5 改进的 ShuffleNetV2 网络
Fig.5 Improved ShuffleNetV2 network

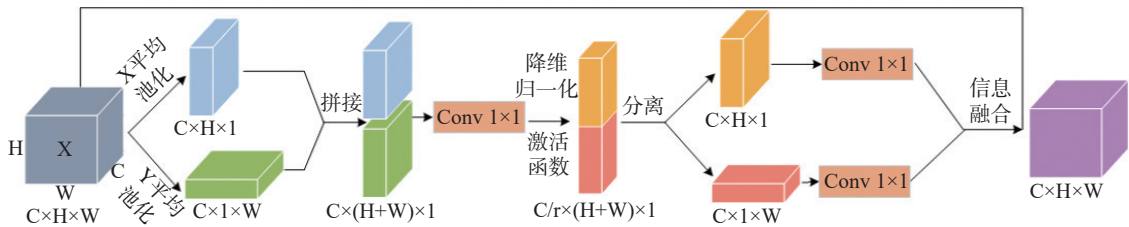


图 6 坐标注意力机制
Fig.6 Coordinate attention mechanism

坐标注意力机制结构如图 6 所示。首先对输入特征图 X 分别进行平均池化操作, 沿水平方向 X 轴和垂直方向 Y 轴分别进行编码, 得到一对大小分别为 $C \times 1 \times W$ 、 $C \times H \times 1$ 的方向感知特征图。然后, 将特征图拼接, 并采用 1×1 卷积操作、归一化和非线性激活函数对水平和垂直方向的空间信息进行编码, 得到中间特征映射。然后, 将特征映射沿空间维度拆分为 $C \times H \times 1$ 和 $C \times 1 \times W$ 的独立的特征图, 并分别通过卷积操作获得两个注意力权重, 最后与原特征图信息融合, 得到尺寸为 $C \times H \times W$ 的输出特征。

1.2.3.4 特征加权融合分类模块 传统特征融合大多采用 Concatenate 方法对特征进行简单拼接, 存在关键特征易缺失、模型分类精度低等问题^[30]。本文提出一种特征自适应加权算法, 能够根据不同特征的重要程度自适应地调整电子舌和电子眼特征融合时所占的比重, 以更好的捕捉贝母的关键特征信息, 其过程如图 7 所示。

假设电子舌特征 X_{tongue} 由 k 个特征向量组成, 表示为 $\{\omega_1, \dots, \omega_k\}$, 电子眼特征 X_{eye} 由 n 个特征向量组成, 表示为 $\{\omega_1, \dots, \omega_n\}$ 。函数 $F(\cdot)$ 分别对电子舌和电子眼特征求和, 得到相应的特征 F_{tongue} 、 F_{eye} :

$$F_{\text{tongue}} = F(\omega_i) = \sum_{i=1}^k \omega_i \quad \text{式 (2)}$$

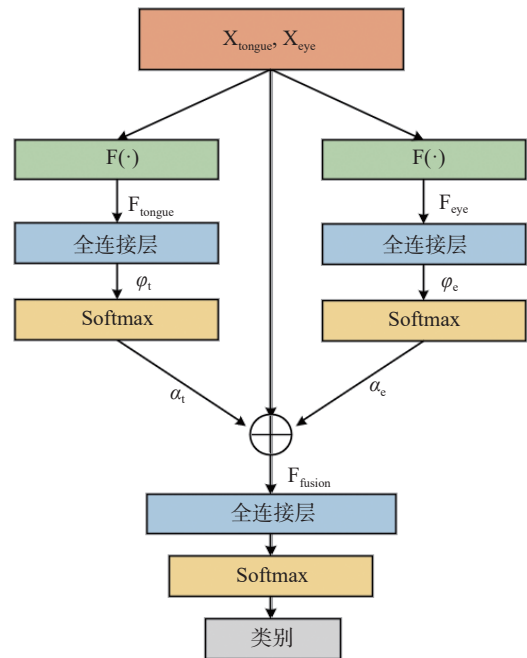


图 7 特征加权融合分类模块
Fig.7 Feature weighted fusion classification module

$$F_{\text{eye}} = F(\omega_i) = \sum_{i=1}^n \omega_i \quad \text{式 (3)}$$

将获得的特征 F_{tongue} 、 F_{eye} 通过全连接层获得输出 φ_i 、 φ_e , 然后通过 Softmax 函数获得该特征的权重 α_i 、

α_e 。Softmax 函数将电子舌和电子眼的特征权重映射到(0, 1)数值空间,各通道特征的权重和为 1, 计算为

$$\alpha_i = \text{Soft max}(\varphi_i) = \frac{e^{\varphi_i}}{\sum_{i=1}^k e^{\varphi_i}} \quad \text{式 (4)}$$

$$\alpha_e = \text{Soft max}(\varphi_e) = \frac{e^{\varphi_e}}{\sum_{i=1}^n e^{\varphi_e}} \quad \text{式 (5)}$$

加权之后的特征 ω_i^* 和融合特征 F_{fusion} 可表示为

$$F_{\text{fusion}} = \sum_{i=1}^{n+k} \omega_i^* = \sum_{i=1}^{n+k} \alpha_i \cdot \omega_i \quad \text{式 (6)}$$

信息融合后,将加权融合特征 F_{fusion} 输入到全连接层并采用 Softmax 实现目标分类。特征加权融合分类模块使得模型能够在不同网络提取的特征之间进行权衡和融合,提高模型的分类性能和鲁棒性。

1.3 数据处理

在 Pycharm 2022.2.1 开发环境下,利用 Python 3.10.9 对电子舌、电子眼数据进行分析建模,深度学习框架采用 pytorch 2.0.1。电子舌信号处理软件为 LabView 2014,绘图软件为 2018 9.5。

2 结果与分析

2.1 电子舌信号采集与预处理

电子舌信号采集结果如图 8 所示。在大幅方波脉冲的激励作用下,样本溶液在电极表面发生电化学反应,由于不同电极的独特催化作用,每个电极的信号呈现出不同的波形。总体表现出贝母样本的味觉指纹图谱信息。电子舌每个电极采集 1000 个数据点,单个贝母样本生成 8000 个数据点。将电子舌信号每类各选取 50 条样本数据作为测试集,剩余样本数据按 7:3 划分训练集和验证集,训练集样本数据量为 1575 条,验证集为 675 条。

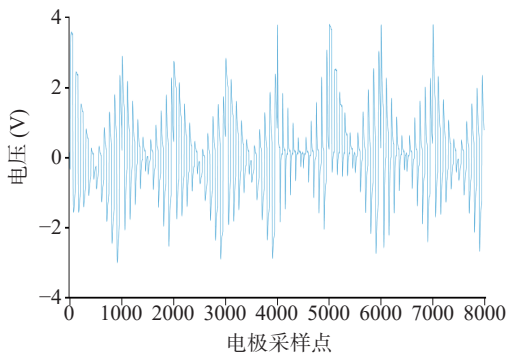


图 8 电子舌响应信号

Fig.8 Electronic tongue response signal

电子舌信号中含有噪声并且数值变化幅度较大,为了便于后续处理,采用式(7)对响应信号进行归一化处理。

$$x'_i = \frac{(x_i - x_{\min})}{(x_{\max} - x_{\min})} \quad \text{式 (7)}$$

式中: x_i 为信号中的第 i 个响应值; x_{\max} 为信号中的最大响应值; x_{\min} 为信号中的最小响应值; x'_i 为信号中第 i 个值的归一化结果。

2.2 电子眼图像采集与预处理

电子眼采集的贝母样本图像如图 9 所示,像素大小为 1920×1080 。由于贝母样本在图片中的位置及大小存在差异,会影响模型训练效果,导致模型出现过拟合问题。因此对图像进行缩放、中心裁剪和归一化等预处理,以去除冗余部分。经预处理后,贝母图像被裁剪成 224×224 像素。将电子眼图像每类各选取 50 张图像作为测试集,剩余图像按 7:3 划分训练集和验证集,其中训练集为 1575 张样本图像,验证集为 675 张样本图像。

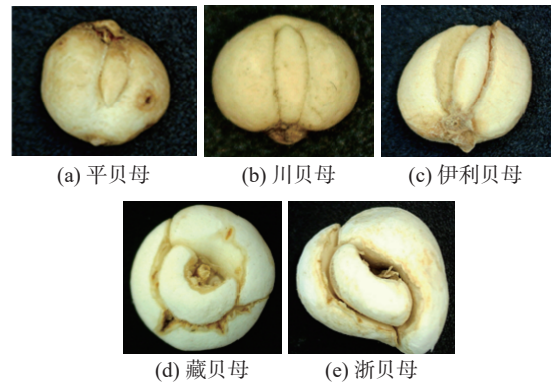


图 9 贝母电子眼图像

Fig.9 Beimu Electronic eye image

2.3 模型结构和参数优化

模型结构和参数影响系统鲁棒性和分类性能。引入贝叶斯优化算法(BOA)对组合模型进行全局优化。贝叶斯优化的目的是尽可能使用少量的目标函数评估优化函数的后验概率,得到最优参数组合^[31],贝叶斯优化的超参数组合的选择可表示为:

$$x_{i+1} = \max_{x \in X} \lambda(x, D_{i,i}) \quad \text{式 (8)}$$

$$r_i = |y^* - y_i| \quad \text{式 (9)}$$

式中: X 为决策空间; $\lambda(x, D_{i,i})$ 为采集函数; y^* 为最优解; r 为总损失; x_{i+1} 为评估点; y_i 为目标函数值。

采用 BOA 算法分别对 Transformer Encoder 模型的层数(Num_layers)、注意力机制的头数(Num_heads)、批量大小(Batch Size)以及优化方法(Optimization method)进行优化。对 ShuffleNetV2 网络的 Repeat 次数(Repeats)、批量大小(Batch Size)、学习率(Learning Rate)以及优化方法(Optimization method)进行优化。表 1 总结了两个模型超参数的搜索空间以及寻优结果。

特征加权融合分类模块对组合模型分类准确性和泛化能力具有较大影响。采用 BOA 算法对特征加权融合分类模块中的自适应初始化权重(Ones)、

表 1 Transformer Encoder、ShuffleNetV2 的 BOA 的搜索空间及寻优结果

Table 1 Search space and optimization results of BOA for Transformer Encoder and ShuffleNetV2

模型	参数	搜索空间	搜寻结果
Transformer Encoder	Num_layers	(1, 128)	2
	Num_heads	(1, 128)	4
	Batch Size	(2, 128)	32
	Optimization method	(SGD, Adama, Nadam)	Adama
ShuffleNetV2	Repeats	[(1, 128), (1, 128), (1, 128)]	[4, 8, 4]
	Batch Size	(2, 128)	16
	Learning Rate	(1×10^{-4} , 1×10^{-2})	0.00421
	Optimization method	(SGD, Adama, Nadam)	SGD

表 2 特征加权融合分类模块的搜索空间及寻优结果

Table 2 Search space and optimization results of feature weighted fusion classification module

参数	搜索空间	搜寻结果
Ones	[(0, 1), (0, 1)]	[0.31, 0.69]
Batch Size	(2, 128)	13
Learning Rate	(1×10^{-4} , 1×10^{-2})	0.00533
Optimization method	(SGD, Adama, Nadam)	Adama

批量大小(Batch Size)、学习率(Learning Rate)、以及优化方法(Optimization method)进行优化。结果如表 2 所示。

2.4 消融实验

为了验证因果自注意力机制、坐标注意力机制的效果,采用原始 Transformer、ShuffleNetV2 作为基础模型,通过消融试验验证改进模型的有效性,结果如表 3 所示。从表 3 中可以看出,在组合模型中引入因果注意力机制,模型准确率和 F1 分数分别提升了 3.2%、0.03,说明因果注意力机制能够很好地捕获信号中的时间序列特征。引入坐标注意力机制后,模型准确率和 F1 分数分别提升了 5%、0.07,证明坐标注意力机制能够有效地增强特征图中不同通道之间的关联性,提高模型对形状和纹理特征的提取能力。综合使用因果注意力机制和坐标注意力机制,模型准确率和 F1 分数分别提升了 13.2%、0.13,说明两种注意力机制能够有效配合,使模型具备更好的特征提取能力和分类效果。

表 3 消融实验结果

Table 3 Results of ablation experiment

编号	因果自注意力机制	坐标注意力机制	准确率(%)	F1-Score
0	-	-	85.2	0.85
1	√	-	88.4	0.88
2	-	√	90.2	0.92
3	√	√	98.4	0.98

2.5 模型性能对比实验

为了验证改进后的模型性能,分别对未改进和

改进 ShuffleNetV2 模型的准确率、参数量、FLOPS 和 FPS 进行对比。其中,模型参数量单位为 M,代表参数个数。浮点计算数(Floating point operations, FLOPs)用于衡量模型的计算复杂度。每秒传输帧数(Frames per second, FPS)即每秒内可以处理的图片数量,用于评估模型分析速度。结果如表 4 所示,改进的 ShuffleNetV2 较未改进的 ShuffleNetV2 模型,准确率提高了 4.2%,参数量减少了 0.21 M, FLOPS 减少了 0.03 GB, FPS 提高了 3.4 frame/s,说明改进的 ShuffleNetV2 不仅识别精度更高,同时有效的减少了模型的参数量,提高了模型的运行速度。

表 4 性能对比结果

Table 4 Results of performance comparison

模型	准确率(%)	参数量(M)	FLOPS(GB)	FPS(frame/s)
ShuffleNetV2	90.2	1.29	0.15	35.2
改进的ShuffleNetV2	94.4	1.08	0.12	38.6

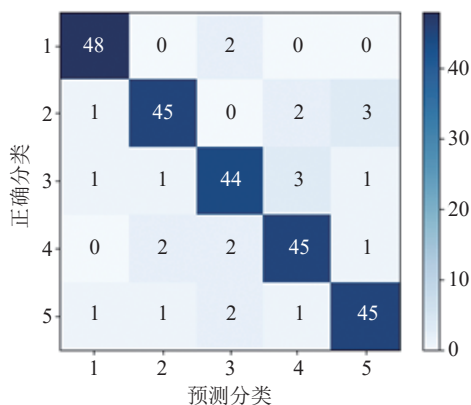
2.6 感官融合性能分析

为了验证本文提出的信息融合模型的有效性,采用电子舌、电子眼以及融合数据集分别在 Transformer、ShuffleNetV2 以及组合模型上进行训练。对五种不同贝母样本进行检测,电子舌和电子眼数据集分别有 2500 个样本,融合数据集共计 5000 个样本数据。利用测试集建立混淆矩阵,结果如图 10 所示,其中 1~5 分别代表川贝母、浙贝母、平贝母、伊利贝母和藏贝母。实验表明,单独采用电子舌检测,贝母样本总体分类准确率为 90.4%。单独采用电子眼图像检测,总体分类准确率达到 94.4%,说明与电子舌信号相比,电子眼图像中包含了更丰富的特征信息。而采用电子舌、电子眼信息融合方法总体分类准确率达到 98.4%,说明电子舌和电子眼信息具有互补性,融合数据集能够获取更多有效信息,有助于区分不同品种贝母样本。

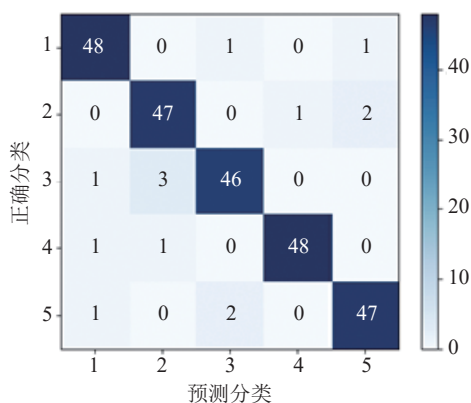
采用 t 分布随机近邻嵌入算法(Stochastic neighbor embedding, t -SNE)对 Transformer Encoder、ShuffleNetV2 提取到的特征以及融合的特征进行可视化分析,结果如图 11 所示。由图 11(a)和图 11(b)可以看出,电子舌和电子眼的每类样本基本聚集成簇,但存在部分重合点。电子眼特征可视化的效果优于电子舌,说明电子眼图像比电子舌信号含有更多差异性信息。由图 11(c)可以看出,电子舌和电子眼混合感官的每一类样本在特征空间中完全汇聚成簇,且不同类别之间分隔较大,具有明显的类间差距,说明使用融合特征分辨样本的能力优于单独使用电子舌和电子眼。

2.7 信息融合方法对比

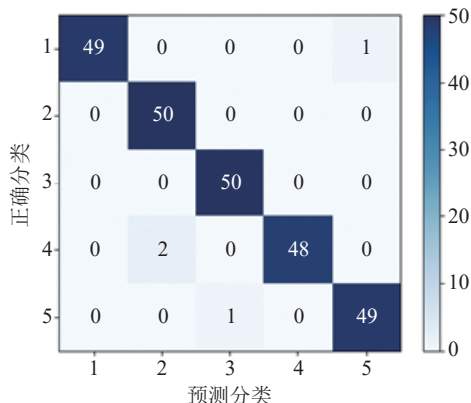
为了验证本文提出模型的性能,分别采用经过 BOA 优化参数的 VGG16、ResNet50、等深度学习模型结合电子舌(ET)和电子眼(EE)采集数据对贝母样本进行检测。采用准确率、精确率、召回



(a) 电子舌感官混淆矩阵



(b) 电子眼感官混淆矩阵



(c) 融合感官混淆矩阵

图 10 测试集混淆矩阵结果

Fig.10 Test set confusion matrix results

率、F1-Score 作为评价标准对模型分类效果进行评价,其公式为:

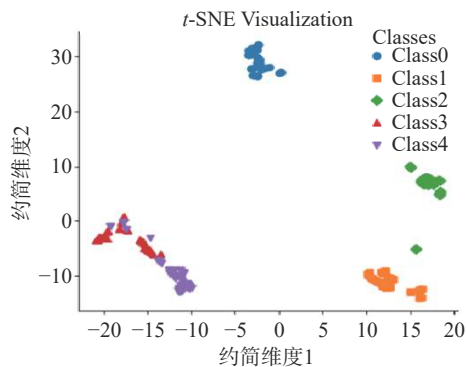
$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad \text{式 (10)}$$

$$Precision = \frac{T_p}{T_p + F_p} \quad \text{式 (11)}$$

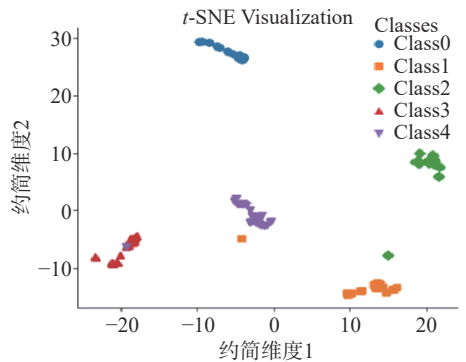
$$Recall = \frac{T_p}{T_p + F_n} \quad \text{式 (12)}$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{式 (13)}$$

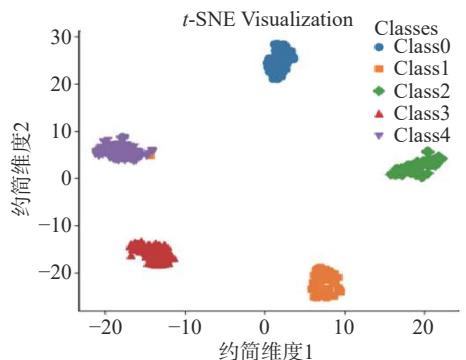
其中, T_p 为匹配正确的样本个数; F_p 为匹配不正确的样本个数; F_n 为没有找到正确匹配的样本数目;



(a) 电子舌特征可视化



(b) 电子眼特征可视化



(c) 混合感官可视化

图 11 t-SNE 特征可视化

Fig.11 Visualization of t-SNE features

T_n 为找到了但不匹配的样本数目。

不同信息融合方法的分类结果如表 5 所示,从表中可以看出,在单模态对比中,电子舌识别准确率最高的是 Transformer,较其他深度学习网络,其准确率提高了 4.8%~25.3%,精确率提高了 4.8%~25.2%,召回率提高了 4.8%~25.3%,F1 分数提高了 0.05~0.25;电子眼识别准确率最高的是 ShuffleNetV2,较其他深度学习网络,其准确率提高了 6.1%~23.6%,精确率提高了 6%~23.6%,召回率提高了 6%~23.6%,F1 分数提高了 0.06-0.23。此外,不同的特征融合策略之间存在明显差异。采用改进的 Transformer 以及 ShuffleNetV2 作为主干网络,本文模型相较于特征相加融合 (AddFusion) 以及特征拼接 (ConcatFusion) 方法,其准确率提高了 2.2%~3.4%,精确率提高了 2.1%~3.4%,召回率提高了 2.2%~3.4%,F1 分数提高了 0.02~0.03,说明本文的自适应

表 5 不同信息融合方法性能对比
Table 5 Results of ablation experiment

数据来源	模型	准确率(%)	精确率(%)	召回率(%)	F1-Score
单模态	VGG16(ET)	65.1	65.2	65.1	0.65
	VGG16(EE)	70.8	70.8	70.8	0.71
	ResNet50 (ET)	85.6	85.6	85.6	0.85
	ResNet50 (EE)	73.2	73.2	73.2	0.73
	Transformer (ET)	90.4	90.4	90.4	0.90
	Transformer (EE)	88.3	88.4	88.3	0.88
	ShuffleNetV2(ET)	78.8	78.8	78.8	0.78
	ShuffleNetV2(EE)	94.4	94.4	94.4	0.94
信息融合	Addfusion	95.0	95.0	95.0	0.95
	ConcatFusion	96.2	96.3	96.2	0.96
	本文模型	98.4	98.4	98.4	0.98

加权方法可以动态分配权重,融合更加灵活,识别效果更好,验证了本文所提出方法的有效性。

2.8 实际应用分析

为了进一步验证本文提出方法的实用性,在本地市场上分别购买川贝母、浙贝母和平贝母各 50 个样本,采用前文方法进行数据采集和处理,然后分别进行测试,实验结果如表 6 所示。结果表明,平贝母、浙贝母和川贝母的平均分类准确率为 96.2%,平均精确率为 96.3%,平均召回率为 96.2%,平均 F1 分数为 0.96。实验证明本文所提出的方法在实际应用中具有较高的检测准确性和较好的泛化性能。

表 6 测试集实验结果
Table 6 Results of test set experimental

样本	准确率(%)	精确率(%)	召回率(%)	F1-Score
平贝母	95.2	95.3	95.2	0.95
浙贝母	96.0	96.0	96.0	0.96
川贝母	97.6	97.6	97.6	0.97

3 结论

本文提出了一种基于电子舌和电子眼结合组合深度模型的贝母品种快速识别方法。利用电子舌和电子眼分别采集贝母样本的味觉和视觉信息。针对电子舌的信号特点,在 Transformer 编码器中引入多头因果自注意力机制,采用因果卷积代替点乘方法,使模型能够充分提取时间序列的局部特征,有效避免噪声干扰。针对电子眼采集信息,在 ShuffleNetV2 网络上添加坐标注意力机制以提高对图像特征的提取能力,并抑制非重要特征干扰。提出一种特征自适应加权算法,对电子舌和电子眼特征自适应调整比重,加强了可用于品种检测的特征信息,提高了贝母品种的识别准确率。本文方法在准确率、精确率、召回率、F1 分数上分别达到了 98.4%、98.4%、98.4%、0.98。同时,对市场上购买的三种实际样本的平均准确率、平均精确率、平均召回率、平均 F1 分数上分别达到了 96.2%、96.3%、96.2%、0.96。本研究提出了一种快速、低成本、高效的贝母品种识别技术,并为其它中药材的快速分类鉴别提供了新的方法。后

续,将进一步提高分类精度、降低计算复杂度,并将该方法应用于嵌入式设备中。

© The Author(s) 2024. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

参考文献

- [1] BORJIGIN G, WEI F S, JIANG S, et al. Extraction, purification, structural characterization and biological activity of polysaccharides from *Fritillaria*: A review[J]. *International Journal of Biological Macromolecules*, 2023, 242: 124817.
- [2] 张富丽,刘炜,毛建霖,等.基于核苷类物质建立川贝母药材真伪鉴别模型研究[J]. *中国中药杂志*, 2021, 46(13): 3337-3348. [ZHANG F L, LIU WEI, MAO J F, et al. Analysis of flavor components of 6 kinds of Yanshan chestnut by GC-MS combined with electronic nose/electronic tongue[J]. *China Journal of Chinese Materia Medica*, 2021, 46(13): 3337-3348.]
- [3] AN Y L, LI Y, WEI W L, et al. Species discrimination of multiple botanical origins of *Fritillaria* species based on infrared spectroscopy, thin layer chromatography-image analysis and untargeted metabolomics[J]. *Phytomedicine*, 2023, 123: 155228.
- [4] 刘燕德,徐振,胡军,等.基于太赫兹光谱技术的贝母品种鉴别方法研究[J]. *光谱学与光谱分析*, 2021, 41(11): 3357-3362. [LIU Y D, XU Zhen, HU Jun, et al. Research on variety identification of *fritillaria* based on terahertz spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(11): 3357-3362.]
- [5] 翟映红,于国强,杨永茂,等.从川贝母中检出平贝母的理化鉴别方法[J]. *华西药学杂志*, 2017, 32(6): 633-635. [ZHAI Yinghong, YU Guoqiang, YANG Yongmao, et al. Physical and chemical identification on distinguishing *Fritillariae ussuriensis* Bulbus from *Fritillariae cirrhosa* Bulbus[J]. *West China Journal of Pharmaceutical Sciences*, 2017, 32(6): 633-635.]
- [6] 王恬,谈娜娜,陆佳,等.川贝母生物碱的化学组成分析与 IL-6 结合活性成分的虚拟筛选[J]. *宝鸡文理学院学报(自然科学版)*, 2023, 43(4): 30-34. [WANG Tian, TAN Nana, LU Jia, et al. Chemical composition analysis of alkaloids form *Fritillaria cirrhosa* D. Don and virtual screening of IL-6 binding active ingredients[J]. *Journal of Baoji University of Arts and Sciences (Natural Science)*, 2023, 43(4): 30-34.]
- [7] 杨银爱,邵海燕,牛犇,等.基于电子舌对不同品种鲜莲子滋味成分的分析[J]. *食品工业科技*, 2023, 44(24): 319-326. YANG Yin'ai, GAO Haiyan, NIU Ben, et al. Analysis of taste components in different varieties of fresh lotus seeds based on electronic tongue[J]. *Science and Technology of Food Industry*, 2023, 44(24): 319-326.
- [8] 张晓,吴宏伟,于现阔,等.基于电子眼技术的穿心莲质量评价[J]. *中国实验方剂学杂志*, 2019, 25(1): 189-195. [ZHANG Xiao, WU Hongwei, YU Xiankuo, et al. Quality evaluation of *andrographis herba* based on electronic-eye technique[J]. *Chinese Journal of Experimental Traditional Medical Formulae*, 2019, 25(1): 189-195.]
- [9] 段金芳,肖洋,刘影,等.一测多评法与电子眼和电子舌技术相结合优化山茱萸蒸制时间[J]. *中草药*, 2017, 48(6): 1108-1116. [DUAN Jinfang, XIAO Yang, LIU Ying, et al. Optimization of steaming time of *Cornus officinalis* by QAMS combined with electronic-eye and electronic-tongue techniques[J]. *Chinese Traditional and Herbal Drugs*, 2017, 48(6): 1108-1116.]

- [10] 陈佳瑜, 袁海波, 沈帅, 等. 基于智能感官多源信息融合技术的滇红工夫茶汤综合感官品质评价[J]. *食品科学*, 2022, 43(16): 294–301. [CHEN Jiayu, YUAN Haibo, SHEN Shuai, et al. Comprehensive sensory quality evaluation of dianhong congou tea infusions using intelligent sensory multi-source information fusion technology[J]. *Food Science*, 2022, 43(16): 294–301.]
- [11] WANG S C, ZHANG Q, LIU C Z, et al. Synergetic application of an E-tongue, E-nose and E-eye combined with CNN models and an attention mechanism to detect the origin of black pepper[J]. *Sensors and Actuators A: Physical*, 2023, 357: 114417.
- [12] LU L, HU Z Q, HU X Q, et al. Quantitative approach of multidimensional interactive sensing for rice quality using electronic tongue sensor array based on information entropy[J]. *Sensors and Actuators B: Chemical*, 2021, 329: 129254.
- [13] 王首程, 于雪莹, 高继勇, 等. 基于电子舌和电子鼻结合 DenseNet-ELM 的陈醋年限检测[J]. *食品与机械*, 2022, 38(4): 72–80, 133. [WANG Shoucheng, YU Xueying, GAO Jiyong, et al. Age detection of mature vinegar based on electronic tongue and electronic nose combined with DenseNet-ELM[J]. *Food & Machinery*, 2022, 38(4): 72–80, 133.]
- [14] SUN Y T, MA H Z. Research progress on oil-immersed transformer mechanical condition identification based on vibration signals[J]. *Renewable and Sustainable Energy Reviews*, 2024, 196: 114327.
- [15] ZHANG S Y, LIAN C, XU B R, et al. A token selection-based multi-scale dual-branch CNN-transformer network for 12-lead ECG signal classification[J]. *Knowledge-Based Systems*, 2023, 280: 111006.
- [16] GONG A J, GUO L B, YU Y, et al. Spectrum-image dual-modality fusion empowered accurate and efficient classification system for traditional Chinese medicine[J]. *Information Fusion*, 2024, 101: 101981.
- [17] 王健庆, 戴恺, 李子柔. 基于深度学习的中药饮片图像识别研究[J]. *时珍国医国药*, 2020, 31(12): 2930–2933. WANG Jianqing, DAI Kai, LI Zirong. Research on image recognition of traditional chinese medicine decoction pieces based on deep learning[J]. *Lishizhen Medicine and Materia Medica Research*, 2020, 31(12): 2930–2933.
- [18] 周丽媛, 赵启军, 高定国. 基于注意力引导深度纹理特征学习的复杂背景藏药材切片图像识别[J]. *世界科学技术-中医药现代化*, 2022, 24(12): 4825–4832. [ZHOU Liyuan, ZHAO Qijun, GAO Dingguo. Recognition of Tibetan medicinal slices in images with complex background via attention-guided deep texture feature learning[J]. *World Science and Technology—Modernization of Traditional Chinese Medicine and Materia Medica*, 2022, 24(12): 4825–4832.]
- [19] CHEN Y, HUANG Y Q, ZHANG Z Z, et al. Plant image recognition with deep learning: A review[J]. *Computers and Electronics in Agriculture*, 2023, 212: 108072.
- [20] SHANG Y Y, XU X S, JIAO Y T, et al. Using lightweight deep learning algorithm for real-time detection of apple flowers in natural environments[J]. *Computers and Electronics in Agriculture*, 2023, 207: 107765.
- [21] 彭红星, 何慧君, 高宗梅, 等. 基于改进 ShuffleNetV2 模型的荔枝病虫害识别方法[J]. *农业机械学报*, 2022, 53(12): 290–300. [PENG Hongxing, HE Huijun, GAO Zongmei, et al. Litchi diseases and insect pests identification method based on improved ShuffleNetV2[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(12): 290–300.]
- [22] SUI Y, WANG X, WU J, et al. Causal attention for interpretable and generalizable graph classification[Z]. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington DC, USA; Association for Computing Machinery. 2022: 1696–705.10.1145/3534678.3539366.
- [23] WEN G, LI S, LIU F, et al. YOLOv5s-CA: A modified YOLOv5s network with coordinate attention for underwater target detection[J]. *Sensors*, 2023, 23(7): 3367.
- [24] 杨正伟, 张鑫, 李庆盛, 等. 基于电子舌及一维深度 CNN-ELM 模型的普洱茶贮藏年限快速检测[J]. *食品与机械*, 2020, 36(8): 45–52. [YANG Zhengwei, ZHANG Xin, LI Qingsheng, et al. A fast detection Pu-erh tea storage based on the voltammetric electronic tongue and one-dimension CNN-ELM[J]. *Food & Machinery*, 2020, 36(8): 45–52.]
- [25] YANG Z W, GAO J Y, WANG S C, et al. Synergetic application of E-tongue and E-eye based on deep learning to discrimination of Pu-erh tea storage time[J]. *Computers and Electronics in Agriculture*, 2021, 187: 106297.
- [26] 金鑫宁, 刘铭, 桑恒亮, 等. 基于电子舌和电子眼结合改进 MobileNetv3 的黄芪快速溯源检测[J]. *食品与机械*, 2023, 39(6): 37–47. [JIN Xinning, LIU Ming, SANG Hengliang, et al. Fast traceability detection of *Astragalus membranaceus* based on the combination of electronic tongue and electronic eye to improve MobileNetv3[J]. *Food & Machinery*, 2023, 39(6): 37–47.]
- [27] 李练兵, 高国强, 吴伟强, 等. 考虑特征重组与改进 Transformer 的风电功率短期日前预测方法[J]. *电网技术*, 2024, 48(4): 1466–1480. [LI Lianbing, GAO Guoqiang, WU Weiqiang, et al. A short-term day-ahead wind power prediction method considering feature recombination and improved transformer[J]. *Power System Technology*, 2024, 48(4): 1466–1480.]
- [28] JIN X, JIANG J X, LI Y, et al. Improved ShuffleNetV2 for action recognition in BPPV treatment[J]. *Biomedical Signal Processing and Control*, 2024, 88: 105601.
- [29] 游小荣, 李淑芳. 融合迁移学习和集成学习的服装风格图像分类方法[J/OL]. *现代纺织技术*: 1–8[2024-06-18]. <https://doi.org/10.19398/j.att.202401007>. [YOU Xiaorong, LI Shufang. Fashion style image classification method integrating transfer learning and ensemble learning: 1–8[2024-06-18]. <https://doi.org/10.19398/j.att.202401007>.]
- [30] 武晓春, 郜文祥. 基于 tSNE 多特征融合的 JTC 轨旁设备故障检测[J]. *铁道科学与工程学报*, 2024, 21(3): 1244–1255. [WU Xiaochun, GAO Wenxiang. Fault detection of JTC trackside equipment based on tSNE multi-feature fusion[J]. *Journal of Railway Science and Engineering*, 2024, 21(3): 1244–1255.]
- [31] 崔佳旭, 杨博. 贝叶斯优化方法和应用综述[J]. *软件学报*, 2018, 29(10): 3068–3090. [CUI J X, YANG B. Survey on Bayesian optimization methodology and applications[J]. *Journal of Software*, 2018, 29(10): 3068–3090.]