

基于视觉模型UniFormer-MS的五常大米掺假识别方法

邢晓^{1,2}, 张媛^{1,2}, 李鑫^{1,2}, 王家宁^{1,2}, 杜艳平^{1,2}, 朱磊^{1,2*}

(1. 北京印刷学院国家邮政局绿色包装技术研发中心, 北京 102600)(2. 北京印刷学院机电工程学院, 北京 102600)

摘要: 现阶段, 市场上五常稻花香大米的掺假越来越严重, 为了高效、精准地识别掺假大米, 该研究使用外形极其相似的东北长粒香作为掺假大米, 同时在手部及黑色背景上拍摄 2 100 张普通图片, 并提出了一种结合 Transformer 与 CNN 的视觉模型 UniFormer-MS。UniFormer-MS 模型融合了 Transformer 的自注意力机制与 CNN 在局部特征提取方面的优势, 构建了端到端的深度学习框架。其中, Transformer 通过全局特征建模与融合增强识别能力, 并引入改进的多头自注意力机制以提高大米真假分类的准确性。Squeeze-and-Excitation (SE) 模块则对通道特征进行加权优化, 以进一步提升模型的精确度和鲁棒性。实验结果表明, 该模型在掺假大米的识别任务中表现出色, 其准确率为 87.00%, 精确度为 87.30%, 召回率为 86.67%, F1 分数为 86.52%。与传统的 UniFormer 相比, UniFormer-MS 准确率提升了 6.92%, 精确度提升了 5.76%, 召回率提升了 5.96%, F1 分数提升了 5.89%。UniFormer-MS 模型通过提高大米表面特征的辨识精度, 高效、精准地识别掺假大米, 为食品质量监控、智能农业及生产线自动化提供了可靠技术支持, 显著提升了食品安全检测的效率与准确性。

关键词: 食品安全; 大米品种; 掺假识别; 卷积神经网络; UniFormer-MS 模型; 深度学习

文章编号: 1673-9078(2026)6-385-394

DOI: 10.13982/j.mfst.1673-9078.2026.6.0396

Adulteration Identification Method for Wuchang Rice Based on Visual Model UniFormer-MS

XING Xiao^{1,2}, ZHANG Yuan^{1,2}, LI Xin^{1,2}, WANG Jianing^{1,2}, DU Yanping^{1,2}, ZHU Lei^{1,2*}

(1.Green Packaging Technology Research and Development Center of the State Post Bureau, Beijing Institute of Graphic Communication, Beijing 102600, China)(2.School of Mechanical and Electrical Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China)

Abstract: Adulteration of Wuchang Daoxiang rice in commercial markets has emerged as a serious issue. To efficiently and accurately identify adulterated rice, this study used Northeast Changxiang rice, which exhibits a high degree of visual similarity to adulterated samples. Concurrently, 2 100 images of ordinary rice were collected against hand-held and black backgrounds. Furthermore, a multiple structural unified transformer (UniFormer-MS) visual model is proposed that combines Transformer architectures with convolutional neural networks (CNNs). The UniFormer-MS model integrates the self-attention mechanism of the Transformer with CNN-based local feature extraction to construct an end-to-end deep learning framework. Specifically, the

引文格式:

邢晓,张媛,李鑫,等.基于视觉模型UniFormer-MS的五常大米掺假识别方法[J].现代食品科技,2026,42(6):385-394.

XING Xiao, ZHANG Yuan, LI Xin, et al. Adulteration identification method for wuchang rice based on visual model UniFormer-MS [J]. Modern Food Science and Technology, 2026, 42(6): 385-394.

收稿日期: 2025-03-20; 修回日期: 2025-05-29; 接受日期: 2025-06-03

基金项目: 北京市自然基金项目 (KZ202210015020); 北京印刷学院科研平台建设项目 (KYCPT202507)

作者简介: 邢晓 (2000-), 女, 硕士研究生, 研究方向: 多传感器数据融合和信号处理、图像识别, E-mail: 126020296@qq.com

通讯作者: 朱磊 (1982-) 男, 博士, 高级工程师, 研究方向: 图像识别、智能包装设计、物流安全的智能状态识别, E-mail: zhulei@bigc.edu.cn

Transformer enhances recognition ability via global feature modeling and fusion employing an improved multi-head self-attention mechanism that improves the accuracy of rice authenticity classification. The squeeze-and-excitation (SE) module performs weighted optimization on channel features to further improve the accuracy and robustness of the model. Experimental results demonstrated the excellent performance of the model in identifying adulterated rice, achieving an accuracy of 87.00%, precision of 87.30%, recall of 86.67%, and F1 score of 86.52%. Compared with traditional UniFormer models, UniFormer-MS improved the recognition accuracy of rice surface features by 6.92%, precision by 5.76%, recall by 5.96%, and F1 score by 5.89%, thereby enabling efficient and accurate identification of adulterated rice. This study provides reliable technical support for food quality monitoring, intelligent agriculture, and production line automation, which significantly enhances the efficiency and accuracy of food safety testing.

Key words: food safety; rice cultivars; adulteration identification; convolutional neural network; multiple structural unified transformer model; deep learning

随着人民生活水平的提高, 食品安全^[1,2]问题, 尤其是食品掺假^[3,4]问题, 愈发受到广泛关注。作为全球重要的主食之一, 大米长期成为不法商贩的掺假对象^[5], 这不仅危及消费者健康^[6], 还损害了农业和食品产业的信誉。作为我国优质大米的代表, 五常大米亦面临严重的掺假问题。因此, 研发高效、准确的检测方法, 既有助于维护五常大米的品牌形象, 也对保障消费者权益和提升食品质量安全具有重要意义。

近年来, 随着人工智能技术的迅猛发展, 传统食品检测方法的局限性日益显现, 如人工检验误差较大、检测速度缓慢及成本较高等问题^[7]。尽管生物化学分析检测方法能够较为精准地揭示物质内部成分的差异并实现有效辨别, 但该类方法通常需要破坏样品, 且操作复杂、耗时费力, 难以实现大规模推广应用。Dai 等^[8]通过分子标记技术对样品进行片段化处理, 以检测食品欺诈行为; Xiao 等^[9]采用无标记蛋白质组学方法区分常规大米与有机大米; Nakamura 等^[10]利用实时荧光 PCR 及插入 / 缺失变异技术进行鉴定; Hu 等^[11]结合荧光探针反应和模式识别方法, 实现对大米掺假的有效鉴定; Chen 等^[12]则采用近红外光谱结合化学计量学手段对掺假大米进行定量分析。然而, 以上方法普遍存在样品破坏和检测周期较长等缺点, 限制了其实用性。

相比之下, 基于图像的非破坏性检测技术因其快速、准确的特点受到广泛关注。深度学习技术凭借强大的数据处理能力和自适应学习特性, 能够从大量样本中自动提取有效特征, 实现更加精准的分类与识别。特别是在计算机视觉领域^[13], 针对大米掺假问题, 深度学习模型通过对大米外观图像的分析, 能够高效识别掺假成分, 为食品检测提供智能化、精细化的解决方案。Pradana-López 等^[14]指出, 基于低要求成像条件

并结合迁移学习的方法, 可实现灵敏且稳定的大米掺假定量检测; Wang 等^[15]利用太赫兹成像结合深度学习对大米新鲜度进行识别; He 等^[16]通过 CD-IMS 与机器学习技术快速检测香米掺假; 此外, 采用深度卷积神经网络的方法对水稻品种进行分类^[17], 进一步展示了深度学习在食品检测领域的广阔应用前景, 有效提升了检测的准确性与效率。

尽管如此, 当前多数方法在数据采集阶段仍依赖复杂且高精度的专业设备, 导致检测成本高昂、流程周期冗长, 难以满足大规模、低成本应用的实际需求。因此, 迫切需要构建一种高效、便捷且具备良好可扩展性的检测体系, 实现基于普通图像采集手段对掺假大米的精准识别与分类。该方法将充分挖掘大米图像的特征特征, 致力于降低检测门槛、缩短检测时间, 推动掺假大米检测技术在实际生产监管中的广泛应用与产业化进程。

综上所述, 本研究不仅具备重要的理论价值, 同时展现出广泛的应用前景。本文基于深度学习技术^[18], 创新性地结合了自注意力机制与卷积操作, 针对图像分类任务进行了有效设计。在 UniFormer 主干网络的 stage1 和 stage2 阶段, 采用并联的双重 SE 注意力机制, 以增强模型的特征表达能力; 通过通道注意力机制动态调整各通道的重要性, 实现对关键信息的有效聚焦。相较于传统图像处理方法和现有深度学习算法^[19], 本研究旨在为大米掺假检测提供一种高效、便捷且低成本解决方案, 进而为食品安全监管工作^[20]提供坚实的技术支撑与保障。

1 材料与方法

1.1 样本准备

本研究选取了黑龙江省哈尔滨市产的五常稻花香和东北长粒香作为实验样本, 其中以五常稻花香为实

验主体，使用外形极为相似的东北长粒香作为掺假大米进行对比分析。所有实验样品均在保质期内完成检测，样品被存放于密封容器中，环境相对湿度严格控制在 30% 以下，室温维持在 20~24 ℃ 之间，以确保样品状态的稳定性。

实验硬件环境为 Windows 11 操作系统，配备 AMD Ryzen 9 7945HX 处理器和 NVIDIA GeForce RTX 4060 显卡。软件环境包括 CUDA 12.6、Python 3.10.14 以及 PyTorch 2.0 深度学习框架，所有代码开发均在 PyCharm 2023 集成开发环境中完成。模型训练时，输入图像尺寸统一设定为 128×128 像素，批处理大小（batch size）为 16，训练周期（epochs）为 20 轮。整个训练过程采用一致的数据集和训练参数设置，确保实验的可重复性和结果的可靠性。

1.2 图像获取与数据集构建

本研究中，水稻样品图像采集采用了 4 800 万像素广角镜头、1 200 万像素超广角镜头及 1 200 万像素长焦镜头，确保图像质量和细节完整性。实验共采集 2 100 张照片，均在室内自然光条件下以手动模式拍摄，光圈固定设置为 F2.4。拍摄过程中采用标准暖光，以模拟消费者在日常环境中的拍摄情境。具体拍摄环境如图 1 所示。为实现研究目标，图像采集涵盖了七种掺有不同比例东北长粒香混合大米的样本，掺假比例分别为 0%、15%、30%、45%、60%、75% 及 100%。每种比例的大米样本均在两种不同背景下拍摄 150 张照片，累计形成 2 100 张图像数据集。两种采集背景的设置过程如下：首先将样品置于黑色卡纸上，以模拟纯色背景环境；其次将样品置于手心，选取三名不同个体的手心作为背景，模拟消费者在超市购买大米时的实际场景。采集图像示例如图 2 所示。采集的原始图像格式为 HEIC，本文对其进行了标准化预处理，统一转换为 JPG 格式，图像原始分辨率为 4 928×3 264 像素。为适应模型输入要求，所有图像在输入模型前均缩放至 128×128 像素，确保数据的一致性和训练效果的稳定性。



图 1 实验拍摄环境

Fig.1 Experimental shooting environment

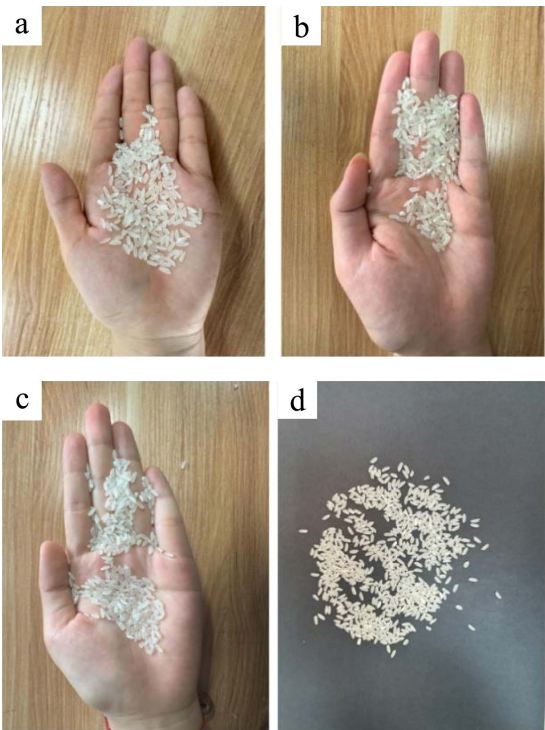


图 2 图像采集示例

Fig.2 Image acquisition example

注：(a) 以模特一手部为背景；(b) 以模特二手部为背景；
(c) 以模特三手部为背景；(d) 以黑色卡纸为背景。

1.3 模型的构建与评价指标

1.3.1 模型的构建

本模型主要实现了一个基于 Transformer 和 CNN 结合的视觉模型 UniFormer-MS，模型结构图如图 3 所示，它结合了自注意力机制^[21](Self-Attention) 和卷积操作 (Convolutional Layers)^[22]，用于处理图像分类的任务。其中 UniFormer 为模型的主结构^[23]，包含多个卷积和 Transformer 计算单元，主结构采用四个阶段进行特征提取，每个阶段有不同的 depth，分别使用 CBlock 和 SABlock 进行计算。其中前两阶段使用 CNN 计算 (CBlock)，后两阶段使用 Transformer 计算 (SABlock)，最后采用 SE 通道注意力机制^[24]，即高效多尺度注意力 (Efficient Multi-scale Attention)，一种自适应通道注意力机制。UniFormer-MS 模块由三个关键模块组成，即动态位置嵌入 (DPE)^[25]、多头关系聚合器 (MHRA) 和前馈网络 (FFN)。输入为图像的初始特征图 $x \in \mathbb{R}^{T \times H \times W}$ ，其中 T 是时间步数，H 和 W 是图像的高度和宽度。经过逐像素卷积 (DWConv)，得到特征图：

$DWConv(X)=Conv_{dw}(X)$ (1)

式中：

$Conv_{dw}$ ——表示深度可分离卷积。

对卷积后的特征图进行标准化,使用批归一化 (Batch Normalization) 或层归一化 (Layer Normalization), 公式如下:

$$\hat{X} = \frac{X - \mu}{\sigma} \tag{2}$$

式中:
 μ 和 σ ——分别为特征图的均值和标准差。

输入到 MHRA 模块后,进行自注意力计算,得到多个注意力头。每个头都计算如下:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \tag{3}$$

式中:
 Q 、 K 、和 V ——分别是查询、键和值;
 d_k ——键的维度。

最终的注意力输出为:

$$MHRA(X) = Concat(head_1, head_2, \dots, head_h) \tag{4}$$

在计算完注意力权重后,使用 SE 进行平滑处理。

其中针对模型三个核心构建阶段进行了特征可视化与解释:

① CNN 局部特征提取阶段:利用 Grad-CAM 对中间卷积层的特征图进行了可视化,展示了模型如何捕捉大米表面纹理和形态差异;

② Transformer 全局自注意力阶段:绘制了多头注意力权重热图,说明不同注意力头如何特征上分布关注重点,强化了全局特征的建模与融合;

③ SE 模块通道加权阶段:统计并可视化了每个通道的权重分布,直观反映了模型在动态调整通道响应时对关键波段和特征通道的选择机制。

综上,针对于该模型的改进部分进行总结,UniFormer-MS 模型通过将 Transformer 的全局自注意力机制与 CNN 的局部特征提取能力深度融合,构建了端到端的深度学习框架;在此基础上,针对大米真伪分类任务对多头自注意力机制进行了改进,采用多尺度特征交互与自适应权重分配策略,以更精细地捕捉微小差异;此外,引入 Squeeze-and-Excitation (SE) 模块,对各通道特征施加动态权重调整,有效突出关键信息、抑制冗余响应,从而显著提升了分类的准确率与鲁棒性;整个模型在统一的训练与推理流水线中协同优化,既简化了部署流程,又保证了各组件的最优协同效应,为大规模生产环境下的快速应用提供了强有力的技术支撑。

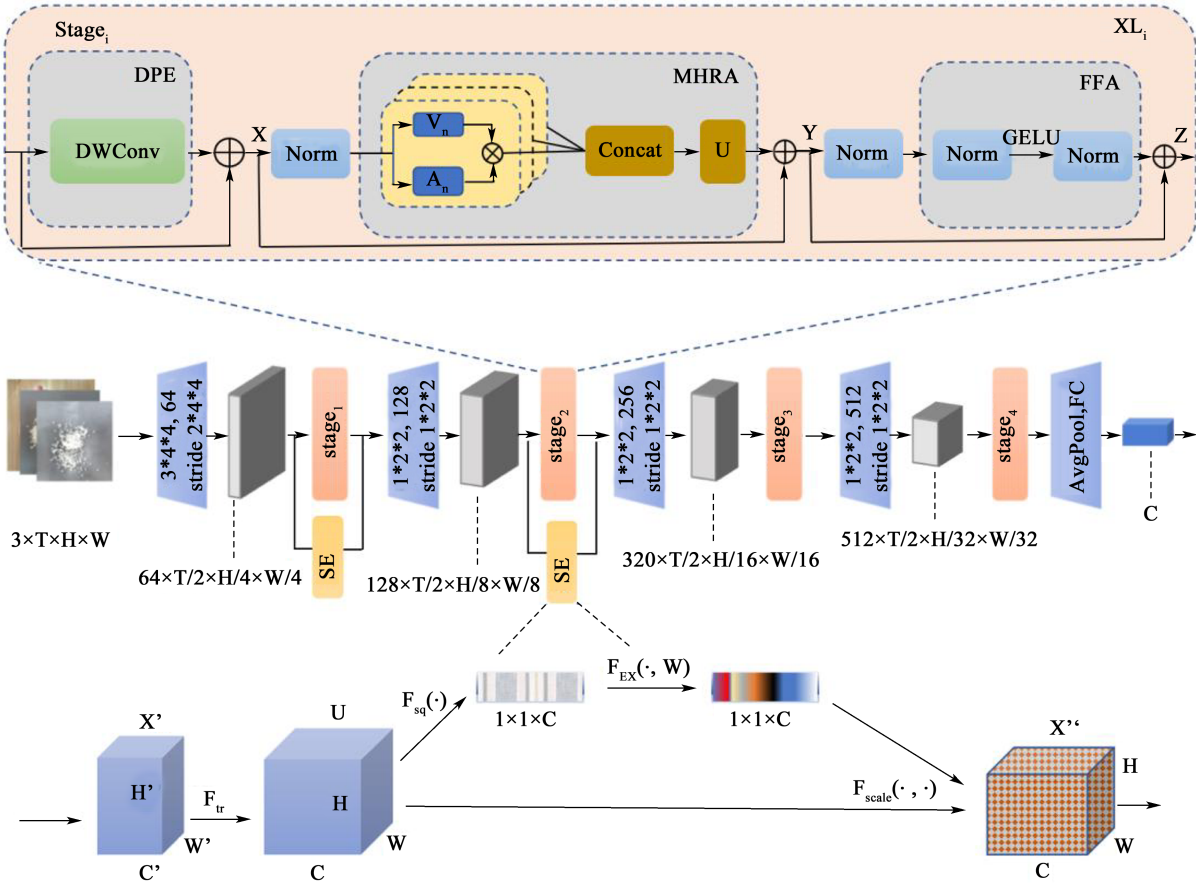


图 3 UniFormer-MS 模型结构
Fig.3 UniFormer-MS model structure

1.3.2 评价指标

在大米图像分类的过程中，分类的准确性至关重要。因此，本研究使用准确率、召回率、精确度和 F1-Score 等指标来评估模型的性能。在整个训练阶段，所有模型均基于相同的数据集和训练参数进行训练。其中 TP 为真阳性，FP 为假阳性，TN 为真阴性，FN 为假阴性。

准确度的公式为：

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$
 (5)

召回率为：

$$Recall = \frac{TP}{TP + FN}$$
 (6)

精确率为：

$$Precision = \frac{TP}{TP + FP}$$
 (7)

F1-Score 为：

$$F1-Score = 2 \frac{Precision * Recall}{Precision + Recall}$$
 (8)

1.4 数据分析

数据的收集与整理采用 excel、PowerPoint 软件进行，在作图方面，所有图表均使用 Origin 2024 软件绘制，以便直观展示实验结果与数据趋势。通过这些分析方法，确保了研究数据的严谨性与科学性，能够为后续研究提供可靠的依据。

2 结果与分析

2.1 不同品种大米原始图像特征

东北五常稻花香与东北长粒香作为典型优质籼型稻米品种，在原始图像特征维度展现出高度相似性，主要体现在色泽、纹理、完整性及杂质控制等方面。具体而言，在光学属性上，二者均具有较高的亮度均值与低色差值，整体色调洁白，油润度适中，表面反射率分布均匀，杂质与色斑数量极低，表征了成熟稳定的栽培体系与加工工艺。在表面形貌特征上，经由灰度共生矩阵（GLCM）计算的纹理一致性（Homogeneity）及能量（Energy）指标均较高，表面粗糙度参数保持在低水平区间，颗粒轮廓完整且边缘过渡自然，裂纹发生率极低，反映出两者在粒体微观结构上的一致性。形态学特征方面，五常稻花香与长粒香均呈现出高完整率（>95%）、低碎米率（<5%），颗粒在长度、宽度、长宽比等基础几何特征上虽存在微小差异，但整体分布区间高度重叠。在背景清洁度

方面，图像噪声点密度极低，异物比例可忽略，为后续基于纹理、形态与色彩特征的自动化识别奠定了良好基础。综上所述，五常稻花香与东北长粒香在宏观图像特征层面呈现出高度一致性，这一现象在一定程度上增加了基于传统特征提取方法进行品种区分的挑战性。图 4 和图 5 分别展示了东北五常稻花香和东北长粒香。



图 4 东北五常稻花香

Fig.4 Northeast wuchang rice fragrance



图 5 东北长粒香

Fig.5 Northeast long grain fragrance

表 1 消融实验结果表

Table 1 Results of ablation experiment

Model	Accuracy/%	Precision/%	Recall/%	f1-score/%
UniFormer	80.71	81.54	80.71	80.63
UniFormer+MHRA	84.50	84.63	84.52	84.50
UniFormer+SE	85.50	85.80	84.80	84.80
UniFormer+MHRA+SE	87.00	87.30	86.67	86.52

2.2 消融试验

表 1 展示了消融实验结果，系统比较了不同模型在融合多种注意力机制及模块后的性能变化。作为基准模型，UniFormer 在基本配置下已表现优异，其准确率达到 80.71%，精确度为 81.54%，召回率为 80.71%，F1 分数为 80.63%，但仍存在进一步提升的潜力。引入多头自注意力机制（MHRA）后，模型性能显著提升，准确率提升至 84.50%，精确度为 84.63%，召回率为 84.52%，F1 分数为 84.50%。该机制有效捕捉了数据中的复杂依赖关系，增强了模型的表达能力和检

测效果。进一步引入 Squeeze-and-Excitation (SE) 模块后, 模型表现再度优化, 准确率达到 85.50%, 精确度为 85.80%, 召回率为 84.80%, F1 分数亦提升至 84.80%。SE 模块通过动态调节通道权重, 强化了特征提取能力, 促进了性能提升。当 UniFormer 同时融合 MHRA 与 SE 模块时, 模型取得最佳性能表现, 准确率达到 87.00%, 精确度为 87.30%, 召回率为 86.67%, F1 分数为 86.52%。这一复合机制显著增强了模型在捕捉数据依赖性与特征表达上的能力, 实现了多项性能指标的全面提升。相关消融实验结果详见图 6。

综上所述, 消融实验充分证明了多头自注意力机制 (MHRA) 与 Squeeze-and-Excitation (SE) 模块对模型性能的显著促进作用, 尤其是两者协同应用时, 模型的准确率、精确度、召回率及 F1 分数均获得显著提升。研究结果表明, 在基础模型中引入更为复杂的注意力机制与特征处理模块, 有助于显著提升模型的识别能力和泛化性能。

2.3 不同注意机制对模型检测性能的影响

表 2 展示了不同注意力机制对模型检测性能的影响, 比较了多种注意力模块 (包括 UniFormer-MS、ECA、CBAM、CA、GGAM 及 EMA) 在准确率 (Accuracy)、精确度 (Precision)、召回率 (Recall) 和 F1 分数 (F1-score) 等指标上的表现。

从各模型表现来看, UniFormer-MS 取得了最佳效果, 准确率为 87.00%, 精确度达到 87.30%, 召回率为 86.67%, F1 分数为 86.52%, 体现了其在精度与召回之间的良好平衡, 综合性能最为突出。相较之下, ECA 的表现稍显逊色, 准确率为 83.30%, 精确度为 83.44%, 召回率为 83.33%, F1 分数为 83.20%。尽管整体未及 UniFormer-MS, ECA 在精确度与召回率的权衡方面仍表现稳定, 适用于对精度要求较高的应用场景。CBAM 的性能较弱, 准确率为 80.95%, 精确度为 81.62%, 召回率为 80.95%, F1 分数为 80.67%, 这可能与其特征提取机制相关, 限制了指标的提升空间。CA 模型的准确率为 85.23%, 精确度为 85.70%, 召回率为 85.24%, F1 分数为 85.13%, 整体优于 CBAM 和 ECA, 但仍不及 UniFormer-MS, 尤其在精确度和召回率方面存在一定差距。GGAM 表现最差, 准确率仅为 75.00%, 精确度为 75.93%, 召回率为 75.00%, F1 分数为 74.74%, 其注意力机制未能有效提升模型性能。EMA 表现较为稳定, 准确率为 85.00%, 精确度为 85.70%, 召回率为 84.50%, F1 分数为 84.40%, 与 CA 相近, 但整体仍逊于 UniFormer-MS。

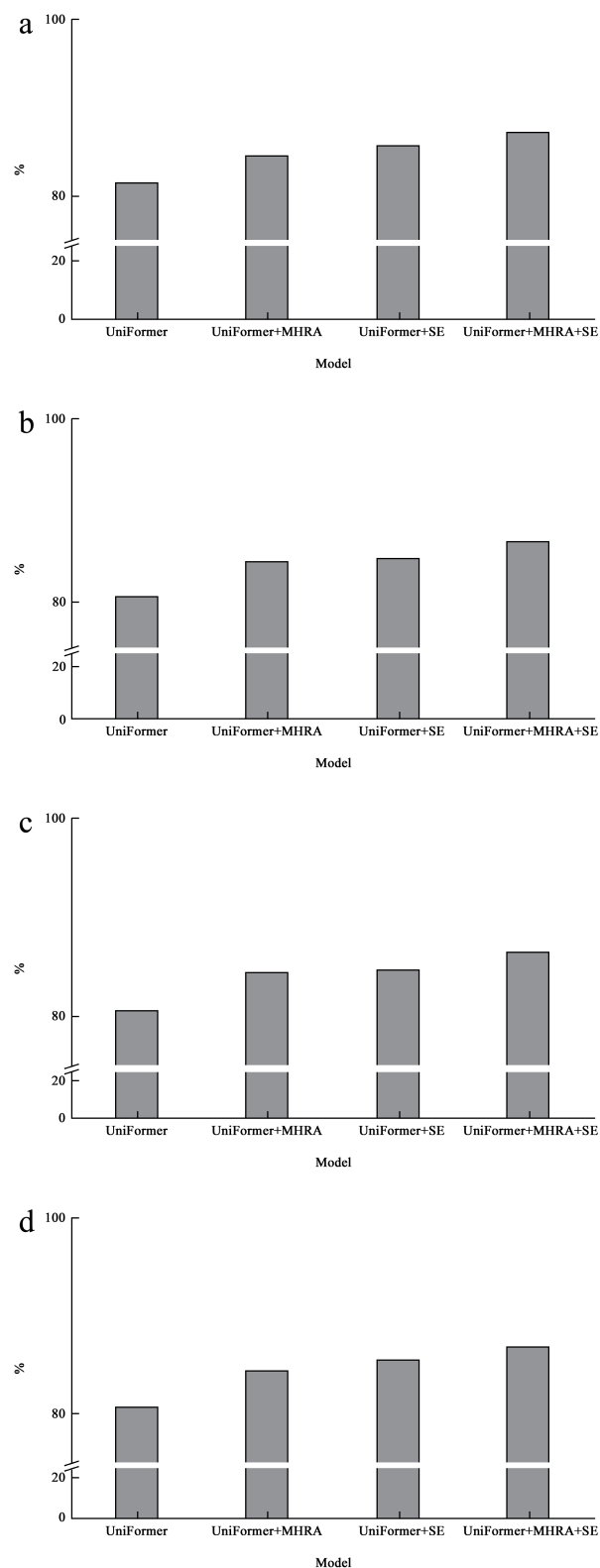


图 6 消融实验结果图

Fig.6 Results of ablation experiment

总体而言, UniFormer-MS 在各项性能指标中均表现领先, 尤其在准确率和 F1 分数方面优势明显。ECA 与 CA 展现出较强的竞争力, 而 CBAM 与 GGAM 则相对表现较弱。综合考虑, UniFormer-MS 凭借其卓越

的性能，适合应用于对分类和检测精度要求较高的任务，具备较强的实用价值和推广潜力。

表 2 不同注意机制对模型检测性能结果表
Table 2 Table of model detection performance results for different attention mechanisms

Model	Accuracy/%	Precision/%	Recall/%	f1-score/%
UniFormer-MS	87.00	87.30	86.67	86.52
ECA	83.30	83.44	83.33	83.20
CBAM	80.95	81.62	80.95	80.67
CA	85.23	85.70	85.24	85.13
GGAM	75.00	75.93	75.00	74.74
EMA	85.00	85.70	84.50	84.40

2.4 不同模型对比分析

在本段中，对照图 7 从不同模型的角度，对结果进行了详细分析。在所有模型中，UniFormer 在各个指标上表现最好，准确率为 87.00%，精确度为 87.30%，召回率为 86.67%，F1 分数为 86.52%。这表明 UniFormer 在分类任务中的综合表现是最强的，能够在准确性和召回率之间取得良好的平衡。Yolov8n 的表现也相当不错，准确率为 83.10%，精确度为 83.50%，召回率为 82.00%，F1 分数为 83.20%。尽管与 UniFormer 相比略逊一筹，但仍然保持了较高的性能，特别是在精确度和召回率上也较为接近。EfficientNetV2_S 的准确率为 81.90%，精确度为 81.93%，召回率为 81.90%，F1 分数为 81.53%。虽然稍逊于 UniFormer 和 Yolov8n，但它的表现仍然很稳定，适合于需要较高效率和较好性能的应用。MobileNetV3_s^[26] 准确率为 69.05%，精确度为 70.50%，召回率为 69.00%，F1 分数为 68.87%。相比其他模型，MobileNetV3_S 在这些指标上表现较弱，可能因为其设计旨在提高计算效率而有所牺牲。VGG19 的准确率为 81.9%，精确度为 82.90%，召回率为 81.90%，F1 分数为 82.02%。VGG19 的性能表现与 EfficientNetV2_S^[27] 相似，但略有提升，尤其是在精确度上。Yolov11 的表现相对较低，准确率为 76.60%，精确度为 76.90%，召回率为 76.50%，F1 分数为 76.20%。这表明 Yolov11 的整体性能较弱，可能不适合高要求的任务。

在本段中，又从不同的识别指标的角度对结果进行分析。在准确率（Accuracy）方面：UniFormer-MS 达到最高的 87.0%，明显优于 Yolov8n（83.1%）、VGG19/EfficientNetV2_S（均 81.9%）以及 MobileNetV3_S（69.1%）和 Yolo11（76.6%），表

明其在整体分类正确率上具有显著优势；在精确度（Precision）方面：UniFormer-MS 同样以 87.3% 领跑，其次依次为 Yolov8n（83.5%）、VGG19（82.9%）、EfficientNetV2_S（81.9%）、Yolo11（76.9%）和 MobileNetV3_S（70.5%）。这一结果说明 UniFormer-MS 识别效果更好，误报率最低；在召回率（Recall）方面：UniFormer-MS（86.7%）优势最大，Yolov8n 为 82.0%，VGG19 与 EfficientNetV2_S 均为 81.9%，Yolo11 为 76.5%，MobileNetV3_S 为 69.1%。高召回率意味着 UniFormer-MS 在捕获正样本方面最为全面；在综合指标 F1-score 上：UniFormer-MS 达到 86.5%，优于 Yolov8n 的 83.2%、VGG19 的 82.0%、EfficientNetV2_S 的 81.5%、Yolo11 的 76.2% 与 MobileNetV3_S 的 68.9%。综合来看，UniFormer-MS 在精确度与召回率之间实现了最佳平衡，故在 F1-score 上也表现最优。

总体而言，UniFormer 在准确率、精确度、召回率和 F1 分数方面均表现出色，领先于其他模型。而 Yolov8n 和 EfficientNetV2_S 则也表现稳健，适合在不同任务中使用。MobileNetV3_S 和 Yolov11 的表现较为逊色，可能更适合在对计算资源要求较低的场景中使用。

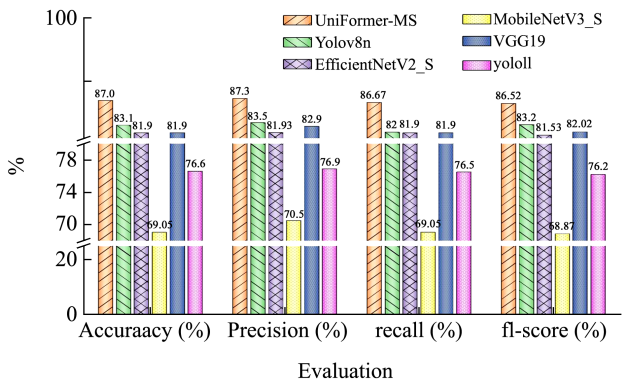


图 7 不同模型对比分析结果图

Fig.7 Comparison and analysis results of different models

2.5 模型的半定量化分析

2.5.1 半定量化分析原理

在模型预测过程中，每个类别（如 0%、15%、30% 等）的样本数量可能不同，或者模型对不同类别的识别准确率不同。如果直接对所有类别的掺假比例取简单平均值，可能会忽略某些类别的重要性（例如，某个类别的样本数量特别多，或者模型对某个类别的识别特别准确）。因此，在模型预测过程中需要计算出加权平均掺假比例，使其可以更合理地反映整体的掺假比例。加权平均掺假比例是一种统计方法，用于根

据每个类别的识别比率和对应的掺假比例，计算整体的平均掺假比例。它的核心思想是：不同类别的样本在总体中的贡献不同，因此需要根据每个类别的样本数量或识别准确率进行加权。

加权平均掺假比例的计算公式如下：

$$F=\sum_{i=0}^{n_{class}-1}(y_i\times class_to_ratio[i])$$
 (9)

softmax 函数^[28]的公式如下：

$$y_i=\frac{e^{x_i}}{\sum_{k=1}^Ne^{x_k}}$$
 (10)

式中：

y_i ——softmax 函数；

class_to_ratio[i]——类别 i 对应的掺假比例（例如，0%、15% 等）；

X_i ——输入向量的第 i 个元素；

N ——输入向量的总元素数量；

F ——加权掺假比例。

2.5.2 半定量化结果验证

在本研究中，本文通过半定量化方法对不同掺假比例的大米样本进行了验证。所使用的模型能够识别掺假成分并计算出其占比。实验中，本文们测试了五个不同掺假比例的样本，分别为 20%、40%、60% 和 80%。每个样本的分析结果如下，结果展示图如图 8 和图 9。

以黑色卡纸为背景的样本检测：

样本 A：掺假大米比例约为 20%。模型识别结果显示，掺假比例为 17.98%，其绝对误差为 -2.02%，相对误差约为 10.10%。

样本 B：掺假大米比例约为 40%。模型识别结果为 37.50%，其绝对误差为 -2.50%，相对误差约为 6.25%。

样本 C：掺假大米比例约为 60%。模型识别结果为 59.84%，其绝对误差为 -0.16%，相对误差约为 0.27%。

样本 D：掺假大米比例约为 80%。模型识别结果为 83.30%，其绝对误差为 +3.30%，相对误差约为 4.13%。

以手为背景的样本检测：

样本 A：掺假大米比例约为 20%。模型识别结果显示，掺假比例为 21.55%，其绝对误差为 +1.55%，相对误差约为 7.75%。

样本 B：掺假大米比例约为 40%。模型识别结果为 45.00%，其绝对误差为 +5.00%，相对误差约为 12.50%。

样本 C：掺假大米比例约为 60%。模型识别结果

为 60.00%，其绝对误差为 0.00%，此识别过程无误差。

样本 D：掺假大米比例约为 80%。模型识别结果为 75.00%，其绝对误差为 -5.00%。相对误差约为 6.25%。

从上述结果可以看出，模型对于不同掺假比例的大米样本均能够给出较为准确的定量化分析，尤其在掺假比例较低时，误差较小。然而，当掺假比例增大时，误差稍有增加，可能与掺假大米中不同成分的比例和颗粒形态的变化有关。总体而言，本研究中的半定量化方法在大米掺假检测中表现出了较好的精度。为使结果更加直观，将上述结果整理为表 3。

表 3 半定量化结果验证
Table 3 Semi-quantitative result validation

背景	样本	掺假比例/%	模型识别结果/%	绝对误差/%	相对误差/%
黑色卡纸	A	20	17.98	-2.02	10.10
黑色卡纸	B	40	37.50	-2.50	6.25
黑色卡纸	C	60	59.84	-0.16	0.27
黑色卡纸	D	80	83.30	3.30	4.13
手	A	20	21.55	1.55	7.75
手	B	40	45.00	5.00	12.50
手	C	60	60.00	0.00	0.00
手	D	80	75.00	-5.00	6.25

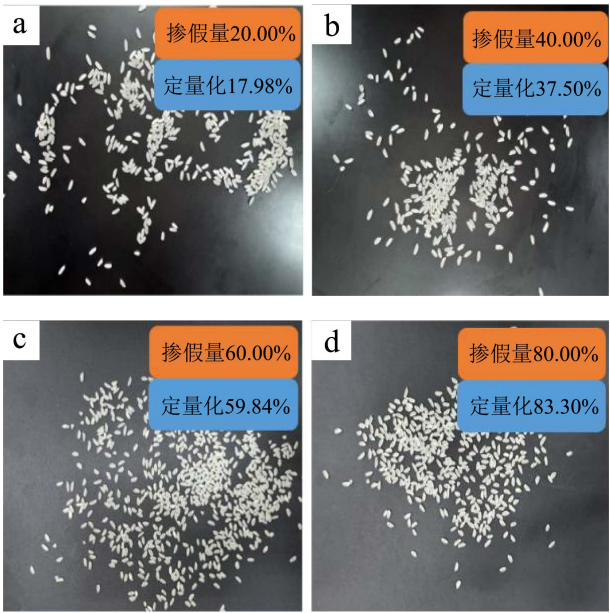


图 8 黑色背景掺假大米的半定量化结果
Fig.8 Semi quantitative results of adulterated rice with black background

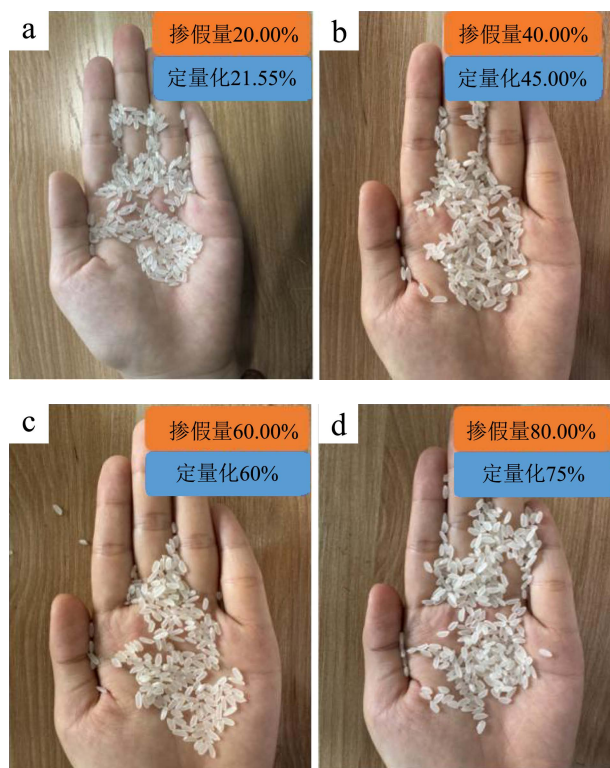


图9 手部背景掺假大米的半定量化结果

Fig.9 Semi quantitative results of adulterated rice in hand background

3 结论

本研究提出的 UniFormer-MS 模型通过动态位置嵌入 (DPE)、多头关系聚合 (MHRA) 和前馈网络 (FFN) 三大模块,并在主干网络的 stage1 和 stage2 部分并联 SE 注意力机制,将自注意力机制与卷积操作有机结合,实现了对大米表面形态特征的高效提取与表达。在包含 0%、15%、30%、45%、60%、75% 和 100% 七种掺假比例的大米图像分类实验中,UniFormer-MS 分别取得了 87.00% 的识别准确率、87.30% 的精确度、86.67% 的召回率和 86.52% 的 F1 分数,显著优于传统生物化学分析方法且无需破坏样本。该方法在标准暖光环境下使用普通摄像头即可完成快速识别,具备良好的实用性与推广价值。未来,随着算法在表面特征检测精度上的进一步提升,UniFormer-MS 可广泛应用于智能农业、食品质量检测^[29]、供应链管理和生产线自动化等领域,为食品安全^[30,31]与生产效率提供有力技术支撑。

参考文献

- [1] BARNES J B, SMITH J C, ROSS K, et al. Performing food safety inspections [J]. Food Control, 2024, 160: 110329.
- [2] WALLS H, BAKER P, CHIRWA E, et al. Food security, food safety

& healthy nutrition: are they compatible [J]. Global Food Security, 2019, 21: 69-71.

- [3] GOYAL K, KUMAR P, VERMA K. Food adulteration detection using artificial intelligence: A systematic review [J]. Archives of Computational Methods in Engineering, 2022, 29(1): 397-426.
- [4] ANAGAW Y K, AYENNEW W, LIMENH L W, et al. Food adulteration: Causes, risks, and detection techniques [J]. Sage Open Medicine, 2024, 12(8): 1250184.
- [5] ŚLIWIŃSKA-BARTEL M, BURNS D T, ELLIOTT C. Rice fraud a global problem: A review of analytical tools to detect species, country of origin and adulterations [J]. Trends in Food Science & Technology, 2021, 116: 36-46.
- [6] ISLAM S, RAHMAN M M, RAHMAN M A, et al. Inorganic arsenic in rice and rice-based diets: health risk assessment [J]. Food Control, 2017, 82: 196-202.
- [7] KIM H, KIM O W, KWAK H S, et al. Prediction model of rice eating quality using physicochemical properties and sensory quality evaluation [J]. Journal of Sensory Studies, 2017, 32(4): e12273.
- [8] DAI X, HE C, ZHOU L, et al. Identification of a specific molecular marker for the rice blast-resistant gene Pigm and molecular breeding of thermo-sensitive genic male sterile leaf-color marker lines [J]. Molecular Breeding, 2018, 38: 1-10.
- [9] XIAO R, LI L, MA Y. A label-free proteomic approach differentiates between conventional and organic rice [J]. Journal of Food Composition and Analysis, 2019, 80: 51-61.
- [10] NAKAMURA K, AKIYAMA H, KAWANO N, et al. Evaluation of real-time PCR detection methods for detecting rice products contaminated by rice genetically modified with a CPTI-KDEL-T-NOS transgenic construct [J]. Food Chemistry, 2013, 141(3): 2618-2624.
- [11] HU L, ZHANG Y, JU Y, et al. Rapid identification of rice geographical origin and adulteration by excitation-emission matrix fluorescence spectroscopy combined with chemometrics based on fluorescence probe [J]. Food Control, 2023, 146: 109547.
- [12] CHEN M, SONG J, HE H, et al. Quantitative analysis of high-price rice adulteration based on near-infrared spectroscopy combined with chemometrics [J]. Foods, 2024, 13(20): 3241.
- [13] SHAMSOLMOALI P, GARCÍA S, ZHOU H, et al. Advances in domain adaptation for computer vision. Image Vis [J]. Comput, 2021, 114: 104268.
- [14] PRADANA-LÓPEZ S, PÉREZ-CALABUIG A M, RODRIGO C, et al. Low requirement imaging enables sensitive and robust rice adulteration quantification via transfer learning [J]. Food Control, 2021, 127: 108122.
- [15] WANG Q, ZHANG Y, GE H, et al. Identification of rice freshness using terahertz imaging and deep learning [J]. In Hotonics, 2023, 10(5): 547.
- [16] HE Y, WANG M. Rapid detection of fragrant rice adulteration by CD-IMS with machine learning [J]. Microchemical Journal, 2025, 209: 112740.
- [17] DIN N M U, ASSAD A, DAR R A, et al. RiceNet: A deep convolutional neural network approach for classification of rice varieties. [J]. Expert Systems with Applications, 2024, 235: 121214.

- [18] JANIESCH C, ZSCHECH P, HEINRICH K. Machine learning and deep learning [J]. *Electronic Markets*, 2021, 31(3): 685-695.
- [19] CHANG L, WU K, CHENG H, et al. An automatic progressive chromosome segmentation approach using deep learning with traditional image processing [J]. *Medical & Biological Engineering & Computing*, 2024, 62(1): 207-223.
- [20] HOOKER N H. Food safety regulation and trade in food products [J]. *Food Policy*, 1999, 24(6): 653-668.
- [21] LIU S, WEI J, LIU G, et al. Image classification model based on large kernel attention mechanism and relative position self-attention mechanism [J]. *PeerJ Computer Science*, 2023, 9: e1344.
- [22] CHEN B, LI H, LUO W. Image processing operations identification via convolutional neural network [J]. *Science China Information Sciences*, 2020, 63(3): 240-242.
- [23] LI K, WANG Y, ZHANG J, et al. Uniformer: Unifying convolution and self-attention for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(10): 12581-12600.
- [24] SHAN C, GENG X, HAN C. Remote sensing image road network detection based on channel attention mechanism [J]. *Heliyon*, 2024, 10(18): e37470.
- [25] ZHOU X, REN Z, ZHOU S, et al. Rethinking position embedding methods in the transformer architecture [J]. *Neural Processing Letters*, 2024, 56(2): 41.
- [26] ABD ELAZIZ M, DAHOU A, ALSALEH N A, et al. Boosting COVID-19 image classification using MobileNetV3 and aquila optimizer algorithm [J]. *Entropy*, 2021, 23(11): 1383.
- [27] SUN Z, LI K, ZHENG Y, et al. Radar spectrum image classification based on deep learning [J]. *Electronics*, 2023, 12(9): 2110.
- [28] SHAO H, WANG S. Deep classification with linearity-enhanced logits to softmax function [J]. *Entropy*, 2023, 25(5): 727.
- [29] ABASI S, MINAEI S, JAMSHIDI B, et al. Dedicated non-destructive devices for food quality measurement: A review [J]. *Trends in Food Science & Technology*, 2018, 78: 197-205.
- [30] 徐国冲,吴筱薇.中国保健食品安全监管的变迁与趋势[J].*学习论坛*,2025,3:77-83.
- [31] 丁浩晗,韩瑜,宋晓东,等.机器学习在食品安全领域生物毒素预测中的应用与展望[J].*食品科学*,2025,46(15):16-26.