KITTI 对抗版聚焦自动驾驶场景，含有雷达和摄像头协同攻击的多传感器数据，这些数据集可模拟现实中的攻击场景。实验设计需包含三类关键实验：对照实验对单模态防御和多模态防御效果进行对比，确定融合技术的优势；攻击强度梯度实验安排不同大小的扰动幅度，如图像像素扰动幅度取值为 0.1 至 0.5，查看防御系统的抗攻击上限；跨场景迁移实验把医疗场景中训练的防御模型迁移到安防场景中，查看模型泛化的能力，实验的时候要控制模型结构、训练参数等变量，保证结果具有可比性。

4.3 典型防御系统的效果对比分析

效果对比需从融合逻辑以及技术路径切入，凸显差异。按融合层次划分，数据层融合防御系统在模态内攻击状况下表现更优，遭受图像噪声攻击，准确率保留率达 90% 以上，然而对跨模态攻击的防御效果较差；特征层融合系统因能捕捉跨模态关联，对跨模态攻击的防御准确率比数据层要高 15% 至 20%，不过计算耗时增加 30%；决策层融合系统响应的速度居于首位，响应时间不足 50ms，适合应对实时场景，但稳定性稍差。按技术路径划分，基于对抗训练的防御系统对已知攻击的防御效果好，白盒攻击背景下准确率下降不足 8%，但处理未知攻击时泛化能力欠佳；基于模态一致性校验的防御系统有识别新型跨模态攻击的能力，误报率不到3%,但对模态里面的细微扰动敏感。从场景应用的角度看，面对混合攻击，自动驾驶多模态防御系统的车辆控制稳定性比单模态高 40%，医疗多模态防御系统把误诊率降低到 2% 以下，为场景化的选型提供借鉴。

5 面临挑战分析与未来研究展望

当前多模态融合智能防御系统落地仍面临三大核心挑战。一是模态异构性适配难，视觉、文本、传感器等数据结构与语义的差异十分显著，现有的融合算法难以完全消除异构鸿沟，造成跨模态信息交互期间容易出现特征丢失或矛盾，损害防御决策的准确水平。二是自适应攻击防御滞后，攻击者可针对多模态融合逻辑进行破解，比如动态调换攻击模态、伪造多模态协同的干扰，而当前的防御系统多数依赖预设规则，难以实时应对新出现的攻击。

后续研究可从三方面突破。其一，探索神经符号融合架构，结合深度学习的特征提取能力与符号逻辑的推理能力，强化异构模态的融合兼容水平。其二，开展针对强化学习驱动的动态防御机制的研发，让系统借助实时沟通学习攻击模式，自适应调整融合权重，调整检测策略，应对自适应攻击。其三，优化轻量化融合算法，凭借模型蒸馏、量化等技术压缩计算成本规模，适配边缘设备需求；构建标准化多模态对抗数据集及评估指标，统一防御效果对比基准，推动技术落地。

6 结语

综上所述，多模态融合为智能防御系统应对对抗性攻击给出了核心解决方案，依靠多层级融合架构以及针对性防御措施，有效弥补了单模态系统鲁棒性方面的弱点，在关键领域的安全防护事务中展现显著价值。但目前技术仍面临着模态异构适配不容易、自适应攻击防御滞后、实时性与鲁棒性难平衡等方面的挑战，对其大规模落地形成了制约。未来要重点攻克新型融合架构与轻量化算法，完善标准化评估体系，促使多模态智能防御系统在更多高安全需求场景中实现可靠运用，为智能系统的安全运转搭建防护壁垒。

参考文献

[1] 张凯.基于AI的计算机网络安全防御系统设计与实现[J].信息与电脑,2025,37(17):103-105.

[2] 潘登科.基于对抗性攻击的AI模型安全防御体系构建与实践[J].信息记录材料,2025,26(09):179-181+184.

[3] 李俊华.基于智能技术的计算机网络安全防御系统分析[J].通讯世界,2025,32(08):68-70.

[4] 何龙.基于人工智能技术的计算机网络安全防御系统设计和实现[J].信息记录材料,2025,26(08):65-67.

 DOI: https://doi.org/10.12345/bdai.v6i6.34068

Application of Artificial Intelligence in Food Safety Testing—Image Recognition and Component Analysis

Ruiqi Wang[1] Haiwu Zheng[3*] Jijun Li[2] Liang Cheng[3]

1. Dalian University of Technology, Dalian, Liaoning, 116024,China
2. Chang'an Bank Headquarters, Xi'an, Shanxi, 710075,China
3. Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, 010018,China

Abstract

The deep integration of artificial intelligence (AI) and big data technologies provides efficient solutions for food safety testing, with image recognition and component analysis being the core application areas. This paper systematically reviews the technical framework of AI in image recognition and component analysis for food safety testing, summarises its application progress in scenarios such as food defect detection, adulteration identification, and quality grading, and addresses current challenges such as the lack of data standardisation, model generalisation capability, and poor edge computing adaptability. Combined with big data technologies, it explores development trends such as integrated regulation of AI, Internet of Things, and blockchain, lightweight model deployment, and federated learning for privacy protection, providing a reference for the integration and application of AI and big data in the field of food safety testing and academic research.

Keywords

Big data; Artificial intelligence; Food safety testing; Image recognition

人工智能在食品安全检测中的应用—图像识别与成分分析

王睿琦[1] 郑海武[3*] 李继军[2] 程亮[3]

1. 大连理工大学，中国·辽宁 大连 116024
2. 长安银行总行公司，中国·陕西 西安 710075
3. 内蒙古农业大学，中国·内蒙古 呼和浩特 010018

摘　要

人工智能（AI）与大数据技术的深度融合，为食品安全检测提供了高效解决方案，其中图像识别与成分分析是核心应用方向。本文系统梳理了人工智能在食品安全检测中图像识别与成分分析的技术框架，综述其在食品缺陷检测、掺假识别、品质分级等场景的应用进展，针对当前数据标准化缺失、模型泛化能力、边缘计算适配性差等挑战的问题，结合大数据技术展望“AI+物联网+区块链”一体化监管、轻量化模型部署、联邦学习隐私保护等发展趋势，为人工智能与大数据在食品安全检测领域的融合应用及学术研究提供一定的参考。

关键词

大数据；人工智能；食品安全检测；图像识别

1 引言

食品安全是公共卫生领域的核心问题，食品生产、加工、流通全链条中存在的掺假、污染物残留、品质不均等问题，不仅严重威胁公众健康，更制约产业可持续发展[1]。据联合国粮食及农业组织统计，全球每年因食品安全问题导致的经济损失达 1500-2000 亿美元[2]，而传统的实验室理化分析，人工视觉筛查等检测方法存在检测周期长数据处理效率低、一般主要依赖检测人员操作等固有局限，很难以满足海量食品样本快速筛查、生产线实时监控等产业实际需求，成为食品安全链条监管的技术瓶颈。

随着大数据时代的深度演进，食品产业链各环节已积累图像数据、光谱数据、传感器数据等海量多源异构数据，涵盖生产加工环节的图像数据、实验室检测的光谱 / 质谱数据、流通环节的传感器环境数据及历史监管数据等[3]，为人工智能（AI）技术的落地应用提供了核心数据支撑。本论文论述了以技术融合 + 行业应用的定位，重点阐述大数据驱动下人工智能在食品安全检测中图像识别与成分分析的技术路径、应用场景及创新方向，突出大数据对 AI 模型训练、

【作者简介】王睿琦（2004–），男，中国内蒙古呼和浩特人，在读本科，从事人工智能方向多模态融合，计算机视觉及大语言模型研究。

【通讯作者】郑海武（1991–），男，中国河北邯郸人，硕士研究生，讲师，研究方向：食品检测、微生物发酵。

优化及实际应用的支撑作用，期望能为相关领域研究者与从业者提供参考。

2 大数据与人工智能的融合路径

2.1 多源数据采集与预处理

食品安全检测数据具有多模态、海量性、高维度特征，主要包括：图像数据、成分检测数据、辅助数据。

图像数据来源于生产线上的视觉采集设备：有可见光图像、高光谱图像、红外图像，成分检测数据来源于实验室检测仪器与现场快速检测设备；有光谱数据、质谱数据、传感器数据，辅助数据来源于物联网终端与企业数据库。有食品生产环境数据、供应链溯源数据、历史检测数据。

数据预处理是技术落地的关键环节，核心步骤通过以下五个方面进行处理：

数据清洗：通过去除噪声数据、KNN 插值法填补缺失值实现数据净化，采用孤立森林、DBSCAN（异常识别率≥95%）异常检测算法，剔除异常数据；

数据标准化：采用 Z-score 标准化（适用于光谱数据）、Min-Max 归一化（适用于图像像素数据），消除不同设备、环境带来的数据差异；

数据降维：利用主成分分析（PCA）、线性判别分析（LDA）降低高维度数据计算复杂度，维度压缩比例达 60%-80%；

数据增强：图像数据采用 0° -360° 旋转、裁剪、翻转等操作，成分数据采用插值扩充比例 1:3 进行扩充，提升模型泛化能力；

数据标注：采用 LabelImg、LabelStudio 等半自动标注工具，结合人工校验（标注准确率 ≥99%），生成高质量标签数据，为模型训练提供支撑。

2.2 人工智能模型构建与优化

人工智能模型构建与优化通过计算完成模型的构建和优化，具体的方法包括：图像识别算法和成分分析算法完成。

图像识别算法：基于深度学习的端到端模型成为主流，包括卷积神经网络（CNN）及其改进模型（如 YOLO、Faster R-CNN 用于目标检测，U-Net 用于图像分割）。针对食品外观缺陷检测、品质分级等场景，通过大数据训练优化模型参数，提升对复杂形态、多变环境的适应性。利用高光谱图像与 CNN 结合的模型，可提取食品内部品质相关的特征信息，实现外观与内在品质的同步检测。

成分分析算法：融合机器学习与深度学习算法，包括传统机器学习（支持向量机 SVM、随机森林 RF、人工神经网络 ANN）与深度学习（1D-CNN、循环神经网络 RNN、Transformer）。通过挖掘大数据中成分检测数据与食品营养成分、污染物含量的映射关系，构建定性定量预测模型。利用拉曼光谱数据结合 1D-CNN 模型，可提取峰位、峰强光谱数据的时序特征，解决传统算法难以捕捉的微弱成分差异问题，在农药残留定量检测中准确率提升 10%-15%。

大数据驱动的优化策略：采用迁移学习解决小样本场景下的模型训练问题，利用 Food-101、UEC-Food-100 预训练模型建立海量公开数据集，再通过目标数据集微调；采用联邦学习实现多机构数据共享训练，在保护数据隐私的前提下，整合多方数据提升模型性能。

2.3 检测系统构建与部署应用

基于数据层与算法层，构建一体化食品安全检测系统，通过核心实时检测，数据处理、结果反馈等模块，实现样本的快速筛查、完成数据的实时分析与模型推理和向监管人员与企业提供检测结果与风险预警。针对不同应用场景，可进行实验室精准检测，以及部署云端大规模计算系统与边缘计算节点进行生产现场实时检测。

3 典型应用场景：大数据与 AI 的实战落地

3.1 图像识别

食品外观与品质的实时检测利用图像识别完成，主要通过缺陷检测与异物识别和食品种类与品质分级，食品包装与标签检测。

缺陷检测与异物识别：在食品加工环节，基于 YOLOv8 模型构建缺陷检测系统，某大型方便面企业生产线应用该系统后，单条生产线检测效率提升至 4000 个 / 小时，漏检率 0.5%（较人工检测漏检率 8%-10% 显著降低），有效解决高生产速度下的漏检问题；在便当、冷食生产中，AI 检测设备可精准识别缺料、毛发异物等问题，准确率 99.2%，部分设备支持标签检测与异物检测联动，响应时间 ≤1s。

食品种类与品质分级：利用高光谱成像结合机器学习算法，可实现果蔬成熟度、肉类新鲜度的精准分级。在食品中通过分析肉类颜色、纹理特征构建 CNN 模型，新鲜度评估准确率达 98.5%，较传统人工分级效率提升 5 倍；在农产品检测中，AI 图像识别可快速区分合格与不合格产品，分选效率提升至 2000 个 / 小时。

食品包装与标签检测：采用 Faster R-CNN 模型检测食品包装完整性与标签位置，破损识别准确率 99.0%，标签偏移检测精度 ≤0.5mm，避免因包装破损或标签错误导致的食品安全隐患，保障食品流通环节质量管控

3.2 成分分析

食品内在品质与安全的精准表征通过食品掺假检测、污染物残留检测、营养成分分析对食品内在品质与安全的精准表征。

食品掺假检测：通过光谱技术结合 AI 算法，可实现食品掺假的精准识别。拉曼光谱结合 1D-CNN 模型实现芝麻油掺假检测准确率达 100%，最低检出掺假比例 5%；在蜂蜜检测中，通过机器学习算法分析其化学成分，可有效识别蔗糖掺假行为，地理溯源准确率达 95.28%。内蒙古地区乳制品企业采用“AI+ 近红外光谱”技术，结合本地奶源成分大数据，实现蛋白质、脂肪含量的实时检测，检测周期从 2 小时缩短至 5 分钟，误差 ≤1%。

污染物残留检测：AI 技术可辅助快速检测食品中的农药残留、重金属、真菌毒素等污染物。例如，在蔬菜种植环