

# 大语言模型在植物病虫害智能诊断中应用的研究进展

王惟实 张慧源 朱诗嫣 何雄奎\*

(中国农业大学理学院, 北京 100193)

**摘要:** 植物病虫害是制约农业生产的重要因素, 依赖人工目测和专家经验的传统诊断方式存在效率低、标准不一等问题。大语言模型凭借强大的语言理解、知识整合和推理能力为构建智能化病虫害诊断系统提供了新思路。该文重点阐述大语言模型的技术基础与发展现状, 基于检索增强生成的病虫害诊断方法, 多模态大语言模型的应用, 基于智能体的诊断系统架构、提示工程与模型微调技术等研究进展, 分析现有研究的优势与不足, 并对未来发展方向进行展望, 以期对相关研究和应用提供参考。

**关键词:** 植物病虫害; 大语言模型; 检索增强生成; 多模态学习; 智能体

## Research advances in the application of large language models for intelligent diagnosis of plant diseases and insect pests

Wang Weishi Zhang Huiyuan Zhu Shiyuan He Xiongkui\*

(College of Science, China Agricultural University, Beijing 100193, China)

**Abstract:** Plant diseases and pests are major constraints on agricultural production. Traditional diagnostic methods relying on manual visual inspection and expert experience are often characterized by low efficiency and inconsistent standards. Large language models (LLMs), with their powerful capabilities in language understanding, knowledge integration, and reasoning, offer new approaches for developing intelligent diagnostic systems for plant diseases and insect pests. This paper systematically reviews the technical foundations and recent developments of large language models, retrieval-augmented generation (RAG)-based diagnostic methods, applications of multimodal large language models, agent-based diagnostic system architectures, as well as prompt engineering and model fine-tuning techniques. The advantages and limitations of existing studies are analyzed, and future research directions are discussed, with the aim of providing a reference for related research and practical applications.

**Key words:** plant disease and insect pest; large language model; retrieval-augmented generation; multimodal learning; intelligent agent

植物病虫害是制约农业生产的重要因素, 每年造成的粮食损失约占全球总产量的 20%~40% (Savary et al., 2019)。准确、及时的病虫害诊断是实施有效防治的前提。传统的病虫害诊断主要依赖专

家经验, 存在诊断效率低、专家资源分布不均、基层技术力量薄弱等问题(慕君林等, 2023)。随着人工智能技术的发展, 智能诊断系统逐渐成为解决这一难题的重要手段。

基金项目: 国家自然科学基金项目(31761133019), 国家重点研发计划项目(2022YFD2001400), 国家现代农业产业技术体系资助项目(CARS-28), 中国农业大学 2115 人才培养发展支持计划项目(2115-89052)

\* 通信作者 (Author for correspondence), E-mail: xiongkui@cau.edu.cn

收稿日期: 2025-12-16

在农业领域,大语言模型的应用尚处于起步阶段,但在农业知识问答、农技咨询、智能决策等场景应用中已显示出独特优势(赵春江,2019;Saiz-Rubio & Rovira-Más,2020;Bjerge et al.,2022)。在植物保护领域,大语言模型具有如下五大优势:可以理解农户用自然语言(包括方言、口语化表达)描述的病虫害症状;整合植物病理学、昆虫学、农药学、气象学等多学科知识并进行综合诊断;通过多轮对话逐步澄清症状细节,模拟专家咨询过程;生成详细的诊断依据和防治建议,提升可解释性;快速更新知识,适应病虫害发生的动态变化(Kuska et al.,2024;Zhu HY et al.,2024)。这些优势使大语言模型有望突破传统诊断方法的局限,构建更智能、更实用的病虫害诊断系统。

目前,国内外学者在大语言模型农业应用方面开展了一系列探索性研究,但相关成果较为分散,缺乏系统性综述,特别是针对植物病虫害诊断这一具体应用场景,大语言模型的技术适配方法、性能评估标准、实际应用效果等关键问题并未充分梳理和总结。鉴于此,本文重点阐述大语言模型的技术基础与发展现状,基于检索增强生成(retrieval-augmented generation, RAG)的病虫害诊断方法,多模态大语言模型的应用,基于智能体的诊断系统架构、提示工程与模型微调技术等研究进展,分析现有研究的优势与不足,并对未来发展方向进行展望,以期对相关研究和应用提供参考。

## 1 大语言模型概述

近年来,人工智能技术的快速发展为植物保护领域带来了新的机遇。大语言模型作为自然语言处理领域的重大突破,展现出强大的语言理解、知识整合和推理能力。自2018年双向编码器表示技术(bidirectional encoder representations from transformers, BERT)模型问世以来(Devlin et al.,2019),大语言模型经历了从早期的预训练-微调范式到当前的基础模型范式的演进;2022年ChatGPT的发布标志着大语言模型进入了新的发展阶段,其在多领域的应用潜力引起了广泛关注(OpenAI,2022)。

### 1.1 大语言模型的基本原理与架构

大语言模型本质上是基于深度神经网络的自然语言处理模型,通常采用Transformer架构作为基础框架(Vaswani et al.,2017)。Transformer架构通过自注意力机制实现了对输入序列中长距离依赖关系的有效建模,克服了传统循环神经网络在处理长序

列时的梯度消失问题。当前主流的大语言模型主要分为编码器-解码器架构(如T5)(Raffel et al.,2020)、仅编码器架构(如BERT)(Devlin et al.,2019)和仅解码器架构(如生成式预训练变换器(generative pre-trained transformer, GPT)系列)3种架构类型。仅解码器架构因其在生成任务中的优异表现而成为当前大语言模型的主流,其通过在海量文本数据上进行预训练学习语言的统计规律和知识表征,然后通过监督微调 and 基于人类反馈的强化学习(reinforcement learning from human feedback, RLHF)来对齐人类偏好(Ouyang et al.,2022)。在技术实现层面,大语言模型的训练通常包括预训练、微调和对齐3个阶段。预训练阶段在大规模无标注语料上进行自监督学习,使模型获得基础的语言理解能力;微调阶段在特定任务的标注数据上进行有监督学习,使模型适应下游任务;对齐阶段通过RLHF等技术使模型输出符合人类价值观和使用习惯(Wang HF et al.,2023);经过这3个阶段模型参数规模从最初的数亿级发展到当前的千亿甚至万亿级,参数规模的增长带来了涌现能力的出现,这使大语言模型展现出少样本学习、思维链推理等新特性(Wei et al.,2022a)。

### 1.2 大语言模型发展历程与核心能力特征

大语言模型的发展可以分为3个主要阶段。第1阶段(2018—2019年)以BERT和GPT为代表,开创了预训练-微调的新范式。BERT采用双向编码器结构,显著提升了自然语言处理(natural language processing, NLP)能力;GPT则采用单向解码器结构,通过自回归语言模型任务进行预训练,展现了强大的文本生成能力。第2阶段(2020—2021年)以GPT-3为标志,模型规模扩展至1750亿个参数(Brown et al.,2020)。这个阶段还出现了T5、BART等多种架构的大语言模型(Raffel et al.,2020;Zhou et al.,2025),这些模型在文本摘要、机器翻译等方面表现优异。第3阶段(2022年至今)以ChatGPT和GPT-4为代表,标志着大语言模型进入实用化阶段(OpenAI,2022;2023)。同期,开源社区也推出了LLaMA、Alpaca、Vicuna等多个开源大语言模型(Touvron et al.,2023),这些模型的出现降低了研究和应用门槛。

大语言模型展现出多项超越传统NLP模型的核心能力。首先是强大的知识表征能力,即通过在海量文本上预训练,模型内部隐式地存储了丰富的世界知识,包括事实性知识、常识性知识和专业领域

知识(Jiang et al., 2020)。这种能力使得大语言模型在知识问答方面表现出色。其次是上下文学习能力,即模型能根据提示词中的少量示例快速适应新任务,无需参数更新(Brown et al., 2020)。这种能力使得大语言模型可以灵活用于各种下游任务,特别适合标注数据稀缺的场景。在植物病虫害诊断领域,这意味着大语言模型可以快速学习新的病虫害类型,而不需要大量标注样本。再次是思维链推理能力,即通过生成中间推理步骤,模型能解决复杂的多步推理问题(Wei et al., 2022b)。这种能力对于病虫害诊断中需要综合多种症状特征、环境因素进行分析判断的场景特别有价值。在提示词中加入让我们一步步思考等引导语可以显著提升模型在复杂推理任务上的表现。最后是指令遵循能力,即经过指令微调的大语言模型能准确理解和执行用户的自然语言指令(Wang YZ et al., 2023)。这种能力使得模型可以作为智能助手,与用户进行多轮交互,提供个性化的诊断建议和防治方案。

### 1.3 大语言模型在农业领域的应用现状

大语言模型在农业领域的应用正在快速发展。在作物管理方面,利用大语言模型构建智能问答系统,从而为农民提供种植技术咨询服务(裴国权等, 2025);在精准农业领域,大语言模型用于解析传感器数据、生成农情报告和决策建议(Kumari et al., 2025);在农业知识管理方面,大语言模型可以从非结构化的农业文献中抽取知识、构建知识图谱(Wang & Zhao, 2024);在植物保护领域,大语言模型可用于病虫害识别与诊断。例如, Madhavi et al. (2025)利用GPT模型自动生成植物病害的文本描述;裴国权等(2025)构建了基于大语言模型的农业害虫知识问答系统。然而现有研究多局限于简单的文本处理任务,尚未充分发挥模型的推理和多模态能力。目前,大语言模型在植物病虫害诊断应用中面临几个关键挑战:一是模型的专业知识有限,通用大语言模型在植物保护专业知识方面存在不足;二是缺乏有效的图像-文本融合机制,难以充分利用病虫害的视觉特征信息;三是诊断推理过程不够透明,难以为农民提供可解释的诊断依据(Kpodo & Nejadhashemi, 2025)。

## 2 基于RAG的病虫害诊断方法

RAG是一种将外部知识库检索与大语言模型生成能力相结合的技术架构,即在生成回复前先从知识库中检索相关信息,从而有效缓解了大语言模

型的幻觉问题,进而提高了响应的准确性和可靠性(Lewis et al., 2021)。在植物病虫害诊断领域,RAG技术将海量的病虫害知识、防治经验和研究文献整合到诊断系统中,因此模型能基于最新、最准确的专业知识进行推理和诊断(杨俊等, 2025)。典型的RAG系统包括知识库构建、相似度检索和增强生成3个核心组件,其工作流程:当用户输入病虫害相关查询时,系统首先将查询转换为向量表示,然后在知识库中检索最相关的文档片段,最后将检索到的信息与用户查询一起输入大语言模型,生成准确、专业的诊断建议。

### 2.1 基于向量数据库的RAG方法

向量数据库是RAG系统中最常用的知识存储和检索方案,其核心思想是将文本信息通过嵌入模型转换为高维向量,并利用向量相似度搜索快速定位相关知识(Johnson et al., 2021)。在植物病虫害诊断中,向量数据库方法首先需要将病虫害百科知识、学术文献、防治手册等文本资料进行分块处理,然后使用BERT、Sentence-BERT等预训练语言模型将文本块编码为稠密向量(Reimers & Gurevych, 2019)。这些向量被索引存储到Faiss、Milvus、Pinecone等专用向量数据库中,这些数据库支持毫秒级的近似最近邻搜索(Wang et al., 2021)。向量数据库的检索效果很大程度上取决于嵌入模型的质量和领域适配性。DPR(dense passage retrieval)方法通过对比学习训练专用的查询编码器和文档编码器显著提升了检索的准确率(Karpukhin et al., 2020)。在农业领域,杨俊等(2025)在植物病虫害语料上对通用嵌入模型进行微调,构建了领域特化的向量表示,这使得通过叶片发黄、边缘枯萎等模糊症状描述也能准确检索到相关病害信息。此外,分块策略对检索效果也至关重要,过大的文本块会引入噪声信息,文本过小则可能丢失上下文。在实践中通常采用500~1 000字符的重叠滑窗分块方式,并保留文档的作物类型、病害类别、发生时期等元数据标签,这便于混合检索(Dai et al., 2022)。

向量数据库方法的主要优势在于检索速度快、可扩展性强,能处理百万甚至亿级规模的知识条目。然而该方法也有局限性,一是语义相似度不等同于事实相关性,可能检索到表述相似但内容无关的文档;二是缺乏对知识间逻辑关系的建模,难以处理需要多步推理的复杂诊断任务(Izcard & Grave, 2021)。

## 2.2 基于知识图谱的RAG方法

知识图谱以结构化的方式表示实体及其关系,为RAG系统提供了更精确、可解释的知识检索能力(Ji et al., 2022)。在植物病虫害诊断领域,知识图谱通过将作物、病虫害、症状、病原、防治措施等要素抽象为实体节点,并通过“引发”“表现为”“防治方法包括”等关系进行连接,形成了领域知识的语义网络(吴桓宇等, 2025)。与非结构化文本相比,知识图谱能准确表达小麦赤霉病由禾谷镰刀菌 *Fusarium graminearum* 引发、在扬花期喷施多菌灵可防治等专业知识的逻辑关系。

基于知识图谱的RAG方法通常采用实体识别-子图检索-路径推理的流程。首先,利用命名实体识别技术从用户查询中识别作物名称、症状描述等关键实体,然后在知识图谱中定位这些实体,提取其周边的子图结构(Yasunaga et al., 2021)。SubGraph-RAG方法通过图神经网络对子图进行编码,将结构化知识转换为向量表示后与大语言模型融合,实现了知识的有效注入(Li MF et al., 2025)。在植物保护领域,研究者构建了涵盖常见作物病虫害的专业知识图谱,该图谱包含数千个实体和关系三元组,结合GPT等大模型可实现智能问答和诊断推理(Zhao et al., 2024; 吴桓宇等, 2025)。

知识图谱的另一优势是支持可解释的推理路径。当系统诊断某种病害时,可以展示从症状节点到病害节点再到防治措施节点的完整推理链条,增强了结果的可信度(Yu et al., 2020)。然而知识图谱的构建和维护成本较高,需要领域专家进行知识抽取、本体建模和质量校验;同时,知识图谱的覆盖范围有限,对于新出现的病虫害或罕见症状可能缺乏相应的知识节点(Chen et al., 2020)。

## 2.3 混合RAG架构

为了综合利用向量数据库和知识图谱的优势,混合RAG架构被提出,该架构将稠密检索、稀疏检索和结构化知识检索相结合,构建多层次、互补的知识检索系统(Ma et al., 2023)。在病虫害诊断场景中,混合RAG架构通常包含以下4个关键模块:第一,向量数据库层。存储病虫害文献、案例描述等非结构化文本,支持语义相似度检索。第二,知识图谱层。存储标准化的病虫害知识,支持实体关系推理。第三,传统检索层。基于BM25等关键词匹配算法,补充字面匹配能力。第四,融合排序层。综合多路检索结果,进行相关性重排(Yang et al., 2024)。例如, Balaguer et al. (2024) 针对农业场景对比了大

模型微调和RAG方法,发现这两种方法可以相互促进,共同提高农业场景回答质量。Bai et al. (2025) 提出了 Sem-RAG 方法,该方法首先通过向量检索召回候选文档,然后利用知识图谱对候选答案进行高层次关联信息补充,最后将高质量的证据文本输入生成模型。类似思路可用于病虫害诊断。当农户描述番茄叶片出现褐色斑点,逐渐扩大并有黄色晕圈的症状时,系统首先通过向量检索找到相关文献,识别出可能的病害类型如番茄早疫病、晚疫病,然后查询知识图谱验证各病害的典型症状特征,最终综合判断,给出诊断结果。

此外,混合RAG架构还可以利用用户反馈进行动态优化。通过记录哪些检索策略在特定场景下效果更好(如症状描述模糊时,知识图谱推理更可靠;查询新型病害时,文献检索更有帮助),系统可以学习并调整不同检索模块的权重(Asai et al., 2023)。在咖啡叶病害识别中, Kumar et al. (2024) 将RAG与YOLOv8模型结合,该混合RAG架构通过从外部知识库动态检索最新的、特定领域的病害诊断信息,显著提升了诊断准确性和可靠性,特别是在农业动态环境下需要实时更新知识的复杂场景中表现出色,该混合RAG架构有效克服了大语言模型的幻觉问题。针对尼泊尔丘陵地区的柑橘产业, Lamsal et al. (2025) 将RAG架构与跨模态视觉语言模型结合,在多源农业数据环境中该混合RAG架构可显著提升病害特征关联能力,从而实现更高鲁棒性的实时诊断。然而混合RAG架构会增加系统复杂度和计算成本。如何设计高效的检索协调机制、如何处理不同来源知识的冲突、如何在性能和成本间取得平衡,仍是当前的研究重点。

综上所述,基于RAG的植物病虫害诊断方法显著提升了大语言模型在专业领域的应用效果,但仍面临诸多挑战。首先,知识库的质量直接决定诊断准确性,需要建立标准化的病虫害知识库构建规范和持续更新机制。其次,检索效率与准确性的平衡仍需优化,特别是在移动端应用场景中。再次,如何有效整合非结构化文本、结构化知识图谱和图像等多模态信息,构建统一的检索框架,是未来研究的重点(Huang et al., 2024)。从技术演进看,向量数据库RAG方法实现简单、扩展性好,适合快速原型开发;知识图谱RAG技术具有可解释性强、推理能力出色的优势,但构建成本较高;混合RAG架构能综合多种方法优势,代表了RAG技术的发展趋势。未来研究应重点关注领域自适应的嵌入模型、高效的知识

图谱构建方法以及多模态知识融合技术。此外, RAG系统的评估标准也需要进一步完善。归一化折损累计增益、平均倒数排名等传统的检索评估指标难以全面反映诊断系统的实际性能,需要结合诊断准确率、解释合理性、用户满意度等多维度指标(Es et al., 2024)。因此,建立植物病虫害诊断领域的标准测试集和评估基准对于推动该领域技术进步具有重要意义。

### 3 多模态大语言模型在病虫害诊断中的应用

随着 GPT-3 向 GPT-4 的演进,以 GPT-4 为代表的大语言模型与视觉模型的迅速发展,催生了多模态大语言模型(multimodal large language model, MLLM)这一新兴领域。MLLM以强大的大语言模型为核心,通过集成图像、文本、音频、视频等多种数据类型来解决多模态任务(Yin et al., 2024),已成为农业决策与病虫害诊断领域的重要技术途径。基于规则的传统专家系统依赖人工定义规则,难以适应农业环境的动态变化;单模态机器学习模型因数据类型单一而存在局限性,仅依赖文本数据难以识别作物的细微视觉特征,而仅依赖图像数据则无法充分融入气象和土壤等背景信息进行综合推断(Sitokonstantinou et al., 2024)。MLLM通过跨模态语义对齐与信息融合突破上述传统人工智能驱动的决策支持系统的瓶颈,借助 CLIP(Radford et al., 2021)和 BLIP-2(Li ZP et al., 2023)等预训练模型在大规模图像-文本数据上构建不同模态之间的映射关系,通过特征拼接、注意力动态加权、多模态结果投票等策略整合各模态的互补信息,显著提升决策的全面性与可靠性,在解决复杂农业种植问题的方面展现出巨大潜力。MLLM在农业病虫害诊断中应用主要依赖模态融合、领域自适应微调、知识整合三大核心技术体系,这三者协同作用实现了多源信息的有效利用与专业决策生成。

#### 3.1 模态融合技术

模态融合是 MLLM 的核心,旨在建立图像、文本、音频等不同模态数据间的语义映射,实现互补信息融合。当前主流融合策略可分为特征级融合、注意力机制融合与多模态特征融合网络、对比学习融合 3 类。特征级融合通过特征拼接、加权求和等方式直接整合不同模态的特征向量;注意力机制融合与多模态特征融合网络通过动态加权分配不同模态的贡献权重,聚焦关键信息,有效抑制冗余干扰;对

比学习融合通过跨模态对比学习构建统一特征空间,增强模态间语义一致性,提升融合精度。

模态融合技术在农业病虫害诊断中已实现多场景落地。例如,Neethirajan(2024)使用麦克风阵列采集动物声学数据,提取特征后将其输入大语言模型中,实现模式识别与健康状态分类;Mono3DVG将单目图像与语言描述结合,并将该方法用于农业领域的三维目标定位与跟踪,该研究推动了作物生长和害虫分布监测等的智能化发展(Zhan et al., 2024);Xie et al.(2021)采用融合网络对植物病害进行识别,并将图像数据与其他模态信息结合,体现了类似视觉问答(visual question answering, VQA)的多模态特性;Milioto et al.(2018)提出了一种用于作物与杂草分类的实时语义分割方法,基于此方法使用 VQA 技术通过引入基于自然语言的交互式查询进一步增强杂草识别系统的可解释性和用户友好性;Dhvale et al.(2024)开发了一种用于农作物病害识别的聊天机器人,该机器人系统将大型模型与生成对抗网络(generative adversarial network, GAN)结合以提升图像质量,采用卷积神经网络(convolutional neural network, CNN)进行精确病害检测,并利用大语言模型为农民提供专业诊断建议和解决方案;Guo et al.(2020)提出结合上下文嵌入与字形特征的实体识别方法,该方法采用三维 CNN 提取形态学上下文信息,引入加权注意力机制增强局部特征表达能力;Wang et al.(2022)将深度学习框架 MPest-RCNN 用于苹果中典型害虫的检测与计数,识别准确率达 99.11%,较 Faster RCNN 算法提升约 0.31 百分点;Catal Reis & Turk(2024)提出基于 Transformer 架构与深度可分离卷积的马铃薯叶片病害识别模型 MDSCIRNet,该模型在测试集上的识别准确率达 99.24%;Zhu et al.(2025)提出融合 MSC-ResViT、MSC-TextCNN 与 CT-CNN 的多模态 AI 模型,该模型对马铃薯早晚疫病的检测准确率达 98.43%,基于该模型开发的 PotatoGPT 在线诊断防控平台可用于番茄、茄子等茄科植物病害的检测;Zhou et al.(2021)提出了一种基于图像-文本多模态协同表示并辅以知识增强的疾病识别模型 ITK-Net,该模型可用于检测番茄、黄瓜的常见病害;吴华瑞等(2025)提出专为甘蓝作物农业知识问答设计的多模态大语言模型 Agri-QA Net,该模型整合文本、音频和图片数据,通过 BERT、声学模型、CNN 分别提取对应特征,借助 Transformer 融合层、跨模态注意力机制及领域自适应技术增强模型能力,其准确

率达89.5%,且各项指标均显著高于单模态模型;还有大量研究将GPT-4的逻辑推理与YOLO等视觉网络结合,用于诊断病虫害,并通过视觉-语言预训练实现黄瓜病害的少样本识别(Cao et al., 2023; Qing et al., 2023)。然而现有的模态融合方法在整合作物特定知识、生长阶段信息和细粒度视觉分析方面仍存在不足,易忽视农业场景中固有的复杂环境和季节动态,导致模型在复杂田间条件下的鲁棒性有待提升。

### 3.2 领域自适应微调技术

通用MLLM在农业领域的专业性能有限,需通过领域自适应微调注入专业知识,其核心目标是在不显著增加计算成本的前提下提升模型对农业垂直领域任务的适配能力,其主流策略包括3类:第1类为低秩适应(low-rank adaptation, LoRA)微调,即在不改变预训练模型主干参数的前提下,通过训练低秩矩阵实现参数高效微调,显著降低计算成本;第2类为冻结式微调与双阶段混合训练,即通过冻结部分模型层,聚焦关键参数优化,平衡模型泛化能力与专业性能;第3类为提示工程优化,即通过设计上下文感知提示词缓解模型幻觉问题,为开源模型适配农业场景提供轻量化方案。例如,针对中国苹果病虫害领域的实体类别多样、别名或缩写普遍、稀有实体识别困难等问题,Zhang et al. (2021)提出了面向该领域的基于字符的命名实体识别模型APD-CA,该模型融合词典信息与相似度增强的词表示,在APDCner语料库上的精确率、召回率和综合评价指标 $F_1$ 值均表现优异。Zhang CK et al. (2025)提出一种结合上下文感知提示词与多模式协作策略的多模态提示词微调方法,该方法利用文本与视觉提示词减轻MLLM在早期大豆种植决策中的幻觉问题,显著提高了开源模型在水培、除草与病虫害管理等栽培决策中的可靠性与实用性。Yan CQ et al. (2025)提出整合深度学习视觉模型与微调大语言模型的CDIP-ChatGLM3创新框架,该框架基于ChatGLM3-6B,采用DMT训练策略,并将LoRA与冻结式微调结合来进行优化,试验表明该方法可使模型BLEU-4得分提升33.16%,平均ROUGE F值增长27.04%,性能超越Qwen-max模型等顶尖竞品,该框架兼精准疾病识别能力与交互式精准处方生成功能,极大方便了农民。Wang YQ et al. (2025)提出一种基于VisualGLM改进的作物病虫害多模态识别大语言模型LLMI-CDP,该模型结合LoRA技术调整预训练模型权重以精准捕捉病虫害特征,借助Q-Former框

架实现语言模型与图像特征的有效模态对齐,该模型在自制数据集上的相关评价指标超过5种领先多模态大语言模型,该模型能准确识别特定作物病害并提供环境相关预防措施建议。针对传统病虫害管理低效及深度学习依赖人工标记数据、动态推理不足的问题,Zhang YQ et al. (2025)提出专门用于病虫害知识的中文大语言模型IPM-AgriGPT,该模型通过G-EA框架生成高质量问答语料库,并结合ACR-CoTD技术与LoRA技术优化基础模型,该模型在多个任务中获得了优异的评分,显示了其在农业智能和病虫害防治方面的巨大潜力。Zhao et al. (2023)提出ChatAgri框架,该框架基于ChatGPT,通过系统化的提示词构建与答案对齐策略,在无需农业领域微调的前提下实现在跨语言与低资源场景下的农业文本分类,在多语种数据与有限标注条件下该框架的表现优于基于预训练语言模型微调及提示词的学习方法,在农业视觉诊断方面表现出色。ChatAgri解决了传统预训练语言模型在农业领域训练数据稀缺、跨语言迁移能力弱及部署成本高等问题。然而,ChatAgri依赖ChatGPT等闭源应用程序接口(application programming interface, API)模型,缺乏对农业领域知识的深度融合与定向优化,难以完全满足领域专属性与可控性的双重需求。

### 3.3 知识整合

在农业智能问答系统中,具体的原始数据对于提高系统的准确性和实用性至关重要。将知识图谱的结构化数据注入大语言模型(Moisev et al., 2022)可增强生成响应的连贯性,并提供语义上下文线索(Li B et al., 2023)。采用多模态概念描述并整合多模态大语言模型可实现基于知识的可视化,从而提升VQA的性能,进而得到更精确的答案(Zha et al., 2023)。例如,Liu YF et al. (2023)构建了中文柑橘病虫害语料库,并对BERT-BiLSTM-CRF模型进行了优化,提升了模型的识别性能。Zhang CS et al. (2024)提出了一种注入实体相关视觉线索的多模态命名实体识别方法,该方法有效提高了实体识别的准确率,由于目前所构建的数据集仅涵盖有限的作物与病害类型,导致该方法领域局限性明显(Li HD et al., 2025)。

利用多模态学习和知识图谱技术整合文本、图像、传感器数据,为大语言模型提供特定领域的专业知识,展示了大语言模型与农业问答系统的集成,从而在农业场景中实现了更全面的上下文感知响应。例如,杨森等(2025)提出了Knowledge Bridger框

架,该框架通过大规模多模态联合建图与实体关系挖掘解析病理特征,该模型对新发虫害的适应性与迁移能力显著增强。

针对农业病虫害数据集匮乏的现状,学者进行了优化工作。针对农业命名实体识别中文歧义、信息源有限及跨模式融合不足等问题,Huang et al. (2025)提出采用双流实体级特征编码架构与动态跨模态门控注意(dynamic cross-modal gated attention, DCGA)机制的 AgriFuseNER 模型,同时构建了含 12 074 个注释样本的农业多模式 NER 数据集 MANER,该模型在该数据集上综合评价指标  $F_1$  得分达 90.73% 且显著优于纯文本模型、零射击 GPT 等多种模型,为农业多模态知识图谱构建等下游应用提供技术支持。针对现有数据集缺乏农业 VQA 所需问答对(QA pair)的问题,Jin et al. (2025)开发并发布了含高质量图像及作物健康、病虫害鉴定等相关详细问答对的 IP-VQA 数据集,利用多模态大语言模型为其设置基准,试验结果表明该数据集可显著提升模型在复杂农业场景中的识别与推理能力。

### 3.4 模型幻觉缓解技术

尽管 MLLM 在处理复杂数据时表现出高效性和高精度,但它们容易产生看似合理但不正确的信息,即幻觉(Bai et al., 2024),这成为制约其实际应用的核心瓶颈。此类幻觉易导致错误的推理结论与决策建议,严重削弱了 MLLM 在实际生产场景中的可落地性与实操价值。幻觉带来的负面影响是深远的,特别是在农业等领域,作物管理优化、农业资源合理分配及长期生产规划等关键农业决策的制订均高度依赖精准的病虫害诊断、环境适配等专业信息,若决策过程受模型幻觉误导,极易引发作物减产、资源浪费等实质性生产损失。因此准确的信息对于做出关于作物管理、资源分配和长期规划的明智决策至关重要。目前已有大量针对大语言模型幻觉问题的研究。Wang SH et al. (2025)将传统序列标注任务重新表述为文本生成问题,从而使基于 GPT 的模型可用于命名实体识别。Yan FR et al. (2025)提出了一种整合上下文生成的实体提示词处理框架,该框架在保持低成本、良好适应性和高鲁棒性的同时,实现了接近监督学习的识别性能。Yuan et al. (2025)提出了将物联网与大语言模型集成的 Pest-GPT 框架,该框架靠引入专家知识库来减轻幻觉,该框架已经在线农场管理平台上应用,试验结果表明基于深度学习的害虫检测模型对 17 种生物的平均检测准确率为 78.2%,害虫治理建议生成的平均

准确率为 77.33%。尽管上述研究在提高复杂农业实体识别精度方面取得了成果,但基于纯文本的实体识别方法仍受限于语言信息本身,难以融合图像、音频等多模态数据。

## 4 基于智能体的病虫害诊断系统

随着大语言模型技术的发展,基于智能体的应用范式正在成为植物病虫害智能诊断领域的新方向。智能体系统通过赋予大语言模型规划、决策、工具调用和环境交互等能力来自主完成复杂的诊断任务。与传统的单轮问答或简单的 RAG 系统相比,基于智能体的病虫害诊断系统具有多步推理、动态规划和自我反思等高级认知能力,能模拟植保专家的诊断思维过程,在面对复杂、多变的田间病虫害问题时表现出更强的适应性和准确性(Chen & Huang, 2025)。智能体架构将大语言模型作为核心大脑,配合外部工具(如图像识别模型、气象数据接口、农药数据库等)和记忆系统构建出能自主感知、推理、决策和执行的智能诊断系统(Kandamali et al., 2025)。提示工程作为智能体系统设计的关键技术,通过精心设计的提示词模板和指令引导模型进行结构化推理和准确执行任务。

### 4.1 智能体架构设计

智能体架构是病虫害诊断系统的核心框架,直接影响系统的推理能力和诊断效果。典型的病虫害诊断智能体采用推理和行动(reasoning and acting, ReAct)架构,从而将推理与行动交织进行(Mahmoud et al., 2025)。该架构包含感知模块、规划模块、执行模块和记忆模块4个核心组件。

#### 4.1.1 感知模块

感知模块负责接收和处理多源输入信息,包括农户描述的文本信息、上传的病害图片、地理位置信息和气象数据等。该模块采用多模态信息融合技术将异构数据转换为智能体可理解的统一表征(Baltrusaitis et al., 2019)。例如,Pan et al. (2025)开发了作物病害诊断智能体系统,该系统在感知模块中集成了计算机视觉模型,利用计算机视觉模型提取病斑特征,同时采用自然语言处理技术解析农户的症状描述,实现了图文信息的协同理解。

#### 4.1.2 规划模块

规划模块是智能体的决策中枢,负责将复杂的诊断任务分解为可执行的子任务序列。该模块采用思维链提示词技术引导大语言模型进行逐步推理(Wei et al., 2022a)。在病虫害诊断场景中,规划模

块通常遵循症状识别→病原鉴定→发生程度评估→防治方案制订的逻辑流程。例如, Wang et al. (2024)提出了智能体规划框架,该框架通过分层任务分解策略将病虫害诊断划分为初步筛查、精确鉴定和方案优化3个层次,每个层次又包含若干子任务,形成树状决策结构,显著提升了诊断的系统性和准确性。

#### 4.1.3 执行模块

执行模块负责调用外部工具完成具体任务。针对病虫害诊断需求,执行模块通常配备图像分类工具、相似病例检索工具、气象数据查询工具、农药数据库接口等(Schick et al., 2023)。例如, Kandamali et al. (2025)设计了棉花种植智能体,该智能体在执行模块中集成了病害图像识别API、植物病理知识库检索接口、病害发生规律预测模型等多种专业工具,该智能体根据规划模块的指令动态选择合适的工具组合,从而实现了诊断过程的自动化和智能化。

#### 4.1.4 记忆模块

记忆模块为智能体提供短期和长期记忆能力。短期记忆存储当前诊断会话的上下文信息,支持多轮交互中的信息连贯性;长期记忆则保存历史诊断案例和经验知识,用于类比推理和知识迁移。例如, Yang et al. (2025)开发的农业智能体采用双层记忆架构,该架构的短期记忆采用滑动窗口机制维护最近的对话历史,长期记忆则利用向量数据库存储典型病例,在面对相似症状时该架构能快速检索参考案例。

### 4.2 提示工程技术

提示工程技术是智能体系统设计的关键技术,其通过精心构造的提示词模板引导大语言模型生成符合预期的输出。在病虫害诊断场景中,提示词设计需要融合植物保护领域知识、诊断流程规范和推理逻辑约束(Liu PF et al., 2023; Zhu Y et al., 2024)。

角色定位提示词是基础性技术,其通过明确定义智能体的专业身份和能力边界使模型输出更符合领域规范。典型的角色定位提示词示例:你是一位具有20年经验的植物病理学专家,擅长诊断水稻、小麦、玉米等主要作物的病虫害问题。你需要根据农户提供的症状描述和图片,系统分析病害类型、发生原因和防治方法(Zamfirescu-Pereira et al., 2023)。这种明确的角色设定能激活大语言模型中与植物保护相关的知识,减少不相关信息的干扰。

少样本学习提示词通过在提示词中嵌入典型案例帮助模型理解诊断任务的输入输出模式。在病虫

害诊断中,少样本提示词通常包含2~5个标注完整的诊断案例,每个案例包括症状描述、诊断过程和结论。例如,模型在提示词中加入生成的马铃薯病虫害问题模板后,该模型在3个数据集上的综合评价指标F<sub>1</sub>得分分别为100.00%、96.70%和93.55%,优于其他模型,显著改善了马铃薯病害的诊断效果(Tang & Hu, 2025)。

思维链提示词引导模型展现中间推理步骤,增强诊断过程的可解释性(Kojima et al., 2022)。在病虫害诊断场景中,思维链提示词要求模型逐步说明“为什么排除某些疑似病害”“哪些症状支持当前诊断”等推理依据。例如, Zhang KP et al. (2024)开发了可解释小麦病害诊断系统,该系统采用思维链提示词,这使得农技人员能审查模型的推理逻辑,发现诊断错误的原因,从而提高了系统可信度。

约束条件提示词用于规范模型输出格式和内容边界。例如,在农药推荐环节,提示词中明确要求“只推荐已登记的化学农药,必须注明使用剂量、施药时期和安全间隔期,禁止推荐高毒农药”。这类约束确保了诊断建议的合规性和安全性,避免了大语言模型可能产生的不当推荐。

动态提示词优化技术根据诊断进展自适应调整提示词内容。例如, Zhang ZL et al. (2025)提出了自适应提示词引导计数器,其使用视觉提示词来改进检测和计数,在检测大规模稻飞虱数量时该策略的计数性能强,且计算开销较低。

### 4.3 系统挑战与未来

基于智能体的病虫害诊断系统代表了大语言模型应用的高级形态,但在实际部署中仍面临若干挑战。首先,复杂智能体系统的推理成本较高,多智能体协作可能涉及数十次模型调用,导致响应延迟和计算开销增加(Akata et al., 2020);针对此问题,可采用混合架构策略,将简单任务交由轻量级模型或规则系统处理,仅在复杂决策环节调用大模型智能体,从而实现效率与性能的平衡。其次,提示工程技术的有效性高度依赖于领域专家知识,优质提示词模板的构建需要植物保护专家与AI工程师深度协作,知识提取和形式化过程复杂且耗时;建立标准化的提示词库和提示工程方法论,形成可复用的技术资产,是未来研究的主要方向。再次,智能体系统的可靠性和安全性问题不容忽视。大语言模型可能产生幻觉输出,在病虫害诊断中表现为不存在的病害名称或错误的防治方法,这可能误导农户,从而造成经济损失(Ji et al., 2023);通过外部知识库验证、多

智能体交叉核验和人工审核机制可在一定程度上缓解该问题,但建立完善的输出质量保障体系仍需深入研究。此外,智能体系统的可解释性对于农业应用至关重要。农技人员和农户只有理解诊断依据才能信任和采纳建议。思维链提示词、推理过程可视化和证据溯源技术能增强系统透明度,但如何在保持推理能力的同时提供简洁易懂的解释,仍是技术难点。未来研究可借鉴可解释AI领域的方法,开发面向农业用户的解释界面和交互模式。多智能体协作机制在理论上具有优势,但实际应用中协调开销可能抵消性能收益。智能体数量、分工粒度和协作模式的优化是复杂的系统工程问题,需要在具体应用场景中进行权衡(Dorri et al., 2018)。建立智能体系统的性能评估基准和设计指南,将有助于指导实践应用。

从长远看,基于智能体的病虫害诊断系统将向着自主性更强、适应性更好的方向发展。结合机器学习技术,智能体可从诊断实践中持续学习,优化推理策略和工具使用模式(Wang JQ et al., 2023)。融合物联网技术,智能体可主动获取田间传感器数据,实现从被动诊断到主动监测预警的跨越。多智能体系统与数字农业平台的深度整合,将构建覆盖监测-诊断-决策-执行全流程的智能植保生态系统,为现代农业的可持续发展提供有力支撑。

## 5 展望

未来大语言模型在植物病虫害智能诊断领域的研究与应用将呈现多维度的发展趋势。第一,模型的专用化与轻量化将成为重要方向。在大型通用模型基础上,通过领域数据微调、参数高效化技术和边缘部署优化,可使模型更适应农业场景中带宽受限、计算资源有限等现实条件。第二,多模态感知能力将进一步增强。随着高光谱图像、无人机遥感影像、物联网传感数据的深度融合,多模态大语言模型有望实现从单一图像识别向综合环境感知的跃迁,为复杂田间场景下的病虫害诊断提供更稳健的支持。第三,知识增强将持续发挥关键作用。未来的RAG架构将进一步结合农业知识图谱、动态监测数据及专家规则,实现实时更新、可解释的诊断推理体系,从而提高预测的可溯源性与透明度。第四,基于智能体的自主式诊断系统将不断成熟。借助智能体的规划、决策与任务执行能力,病虫害诊断系统将从答复型向行动型演进,可在田间实施连续监测、智能问诊与干预策略推荐。此外,提示工程与可控生成技

术将更精细,这使模型在农业领域的行为更稳定、可靠。随着监管机制与安全评估体系的完善,大模型在实际农业生产中的部署也将更规范。

综上所述,大语言模型为植物病虫害智能诊断带来了机遇,但从研究探索到大规模应用仍需克服诸多技术、实践和社会层面的挑战。通过持续的技术创新、跨学科协作和应用实践,这一模型有望成为现代植物保护体系的重要支撑,为保障全球粮食安全和农业可持续发展做出重要贡献。未来5—10年将是该领域快速发展的关键时期,需要科研机构、高校、企业和政府部门的共同努力,推动技术从实验室走向田间地头,真正惠及广大农业生产者。

## 参考文献 (References)

- Akata Z, Balliet D, de Rijke M, Dignum F, Dignum V, Eiben G, Fokkens A, Grossi D, Hindriks K, Hoos H, et al. 2020. A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8): 18–28
- Asai A, Wu ZQ, Wang YZ, Sil A, Hajishirzi H. 2023. Self-RAG: learning to retrieve, generate, and critique through self-reflection. *arXiv preprint arXiv: 2310.11511*
- Bai B, Meng XY, Zhao CZ. 2025. Research on sem-RAG: a corn planting knowledge question-answering algorithm based on fine-grained semantic information retrieval enhancement. *Applied Sciences*, 15(19): 10850
- Bai ZC, Wang PC, Xiao TJ, He T, Han ZB, Zhang Z, Shou MZ. 2024. Hallucination of multimodal large language models: a survey. *arXiv: 2404.18930*
- Balaguer A, Benara V, de Freitas Cunha RL, de M Estevão Filho R, Hendry T, Holstein D, Marsman J, Mecklenburg N, Malvar S, Nunes LO, et al. 2024. RAG vs fine-tuning: pipelines, tradeoffs, and a case study on agriculture. *arXiv: 2401.08406*
- Baltrusaitis T, Ahuja C, Morency LP. 2019. Multimodal machine learning: a survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2): 423–443
- Bjerge K, Mann HMR, Høye TT. 2022. Real-time insect tracking and monitoring with computer vision and deep learning. *Remote Sensing in Ecology and Conservation*, 8(3): 315–327
- Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, et al. 2020. Language models are few-shot learners. *arXiv: 2005.14165*
- Cao YY, Chen L, Yuan Y, Sun GL. 2023. Cucumber disease recognition with small samples using image-text-label-based multi-modal language model. *Computers and Electronics in Agriculture*, 211: 107993
- Catal Reis H, Turk V. 2024. Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks. *Engineering Applications of Artificial*

- Intelligence, 133: 108307
- Chen D, Huang YB. 2025. Integrating reinforcement learning and large language models for crop production process management optimization and control through a new knowledge-based deep learning paradigm. *Computers and Electronics in Agriculture*, 232: 110028
- Chen XJ, Jia SB, Xiang Y. 2020. A review: knowledge reasoning over knowledge graph. *Expert Systems with Applications*, 141: 112948
- Dai ZY, Zhao VY, Ma J, Luan Y, Ni JM, Lu J, Bakalov A, Guo K, Hall KB, Chang MW. 2022. Promptagator: few-shot dense retrieval from 8 examples. arXiv: 2209.11755
- Devlin J, Chang MW, Lee K, Toutanova K. 2019. BERT: pre-training of deep bidirectional transformers for language understanding//2019 Conference of the North American Chapter of the Association for Computational Linguistics: human language technologies, pp. 4171–4186
- Dhvale C, Pawar T, Singh A, Pole S, Sabat K. 2024. Revolutionizing farming: GAN-enhanced imaging, CNN disease detection, and LLM farmer assistant//2024 2nd International Conference on Computer, Communication and Control (IC4). IEEE, pp. 1–6
- Dorri A, Kanhere SS, Jurdak R. 2018. Multi-agent systems: a survey. *IEEE Access*, 6: 28573–28593
- Es S, James J, Espinosa Anke L, Schockaert S. 2024. RAGAs: automated evaluation of retrieval augmented generation//Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: system demonstrations. ACL, pp. 150–158
- Guo XC, Zhou H, Su J, Hao X, Tang Z, Diao L, Li L. 2020. Chinese agricultural diseases and pests named entity recognition with multi-scale local context features and self-attention mechanism. *Computers and Electronics in Agriculture*, 179: 105830
- Huang J, Wang M, Cui YP, Liu J, Chen L, Wang T, Li H, Wu JM. 2024. Layered query retrieval: an adaptive framework for retrieval-augmented generation in complex question answering for large language models. *Applied Sciences*, 14(23): 11014
- Huang JZ, Hao X, Wang Y, Song RZ, Mu ZN, Chu W, Papadakis G, Niu SJ, Guo XC. 2025. Entity-level cross-modal fusion for multi-modal Chinese agricultural diseases and pests named entity recognition. *Smart Agricultural Technology*, 12: 101188
- Izacard G, Grave E. 2021. Leveraging passage retrieval with generative models for open domain question answering//Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: main volume. Online Association for Computational Linguistics, pp. 874–880
- Ji SX, Pan SR, Cambria E, Marttinen P, Yu PS. 2022. A survey on knowledge graphs: representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2): 494–514
- Ji ZW, Lee N, Frieske R, Yu TZ, Su D, Xu Y, Ishii E, Bang YJ, Madotto A, Fung P. 2023. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12): 1–38
- Jiang ZB, Xu FF, Araki J, Neubig G. 2020. How can we know what language models know. *Transactions of the Association for Computational Linguistics*, 8: 423–438
- Jin KR, Zi X, Thiagarajan K, Braytee A, Prasad M. 2025. IP-VQA dataset: empowering precision agriculture with autonomous insect pest management through visual question answering//Proceedings of the ACM on Web Conference 2025. ACM, pp. 2000–2007
- Johnson J, Douze M, Jégou H. 2021. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3): 535–547
- Kandamali DF, Porter WM, Porter E, McLemore A, Rains GC. 2025. CottonBot: an AI-driven cotton farming assistant and irrigation advisor using LLM-RAG and agentic AI tools. *Smart Agricultural Technology*, 12: 101640
- Karpukhin V, Oguz B, Min S, Lewis P, Wu L, Edunov S, Chen DQ, Yih WT. 2020. Dense passage retrieval for open-domain question answering//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). ACL, pp. 6769–6781
- Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. 2022. Large language models are zero-shot reasoners//Proceedings of the 36th International Conference on Neural Information Processing Systems. ACM, pp. 22199–22213
- Kpodo J, Nejadhashemi AP. 2025. Navigating challenges/opportunities in developing smart agricultural extension platforms: multi-media data mining techniques. *Artificial Intelligence in Agriculture*, 15(3): 426–448
- Kumar SS, Khan AKMA, Banday IA, Gada M, Shanbhag VV. 2024. Overcoming LLM challenges using RAG-driven precision in coffee leaf disease remediation//2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS). IEEE, pp. 1–6
- Kumari K, Mirzakhani Nafchi A, Mirzaee S, Abdalla A. 2025. AI-driven future farming: achieving climate-smart and sustainable agriculture. *AgriEngineering*, 7(3): 89
- Kuska MT, Wahabzada M, Paulus S. 2024. AI for crop production: where can large language models (LLMs) provide substantial value. *Computers and Electronics in Agriculture*, 221: 108924
- Lamsal RR, Acharya P, Pokhrel R, Dahal S, Sigdel V. 2025. Integrating machine learning and rag-based chatbot for mandarin orange disease detection in hilly region of Nepal. <https://doi.org/10.21203/rs.3.rs-6105402/v1>
- Lewis P, Perez E, Piktus A, Petroni F, Karpukhin V, Goyal N, Küttler H, Lewis M, Yih WT, Rocktäschel T, et al. 2021. Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv: 2005.11401
- Li B, Zhang YH, Chen LY, Wang JH, Pu FY, Yang JK, Li CY, Liu ZW. 2023. MIMIC-IT: multi-modal in-context instruction tuning. arXiv: 2306.05425
- Li HD, Wu HR, Li QX, Zhao CJ. 2025. A review on enhancing agricultural intelligence with large language models. *Artificial Intelligence in Agriculture*, 15(4): 671–685
- Li MF, Miao SQ, Li P. 2025. Simple is effective: the roles of graphs and large language models in knowledge-graph-based retrieval-augmented generation. arXiv: 2410.20724

- Li ZP, Cao Z, Li PF, Zhong Y, Li SB. 2023. Multi-hop question generation with knowledge graph-enhanced language model. *Applied Sciences*, 13(9): 5765
- Liu PF, Yuan WZ, Fu JL, Jiang ZB, Hayashi H, Neubig G. 2023. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9): 1–35
- Liu YF, Wei SQ, Huang HJ, Lai Q, Li MS, Guan LX. 2023. Naming entity recognition of citrus pests and diseases based on the BERT-BiLSTM-CRF model. *Expert Systems with Applications*, 234: 121103
- Ma XB, Gong YY, He PC, Zhao H, Duan N. 2023. Query rewriting in retrieval-augmented large language models//*Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Singapore Association for Computational Linguistics, pp. 5303–5315
- Madhavi, Singh TP, Singh SP. 2025. DeepLabV3+ with GPT: a framework for leaf disease image to text generation. *Procedia Computer Science*, 260: 353–359
- Mahmoud AAS, Shishah W, Mistry NR. 2025. ReACT\_OCRS: an AI-driven anonymous online reporting system using synergized reasoning and acting in language models. *IEEE Access*, 13: 92800–92815
- Milioto A, Lottes P, Stachniss C. 2018. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs//*2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2229–2235
- Moiseev F, Dong Z, Alfonseca E, Jaggi M. 2022. SKILL: structured knowledge infusion for large language models//*Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: human language technologies*. ACL, pp. 1581–1588
- Mu JL, Ma B, Wang YF, Ren Z, Liu SX, Wang JX. 2023. Review of crop disease and pest detection algorithms based on deep learning. *Transactions of the Chinese Society for Agricultural Machinery*, 54(Suppl. 2): 301–313 (in Chinese) [慕君林, 马博, 王云飞, 任卓, 刘双喜, 王金星. 2023. 基于深度学习的农作物病虫害检测算法综述. *农业机械学报*, 54(增刊2): 301–313]
- Neethirajan S. 2024. Decoding the language of chickens: an innovative NLP approach to enhance poultry welfare. *bioRxiv preprint*: doi. <https://doi.org/10.1101/2024.04.29.591707>
- OpenAI. 2022. ChatGPT: optimizing language models for dialogue. <https://openai.com/blog/chatgpt>
- OpenAI. 2023. GPT-4 technical report. arXiv: 2303.08774
- Ouyang L, Wu J, Xu J, Almeida D. 2022. Training language models to follow instructions with human feedback//*Proceedings of the 36th International Conference on Neural Information Processing Systems*. ACM, pp. 27730–27744
- Pan JD, Zhong RH, Xia FL, Huang JF, Zhu LC, Yang Y, Lin T. 2025. ChatLeafDisease: a chain-of-thought prompting approach for crop disease classification using large language models. *Plant Phenomics*, 7(3): 100094
- Pei GQ, Qian XY, Zhou B, Wang BJ, Liu ZG, Wu WD. 2025. Tea intelligent question-answering system using lightweight large language models. *Transactions of the Chinese Society of Agricultural Engineering*, 41(15): 174–182 (in Chinese) [裴国权, 钱雪英, 周兵, 王白娟, 刘自高, 吴文斗. 2025. 基于轻量级大语言模型的茶叶智能问答方法. *农业工程学报*, 41(15): 174–182]
- Qing JJ, Deng XL, Lan YB, Li ZK. 2023. GPT-aided diagnosis on agricultural image based on a new light YOLOPC. *Computers and Electronics in Agriculture*, 213: 108168
- Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, et al. 2021. Learning transferable visual models from natural language supervision. arXiv: 2103.00020
- Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, Zhou YQ, Li W, Liu PJ. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(1): 5485–5551
- Reimers N, Gurevych I. 2019. Sentence-BERT: sentence embeddings using Siamese BERT-networks//*Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. China Association for Computational Linguistics, pp. 3982–3992
- Saiz-Rubio V, Rovira-Más F. 2020. From smart farming towards agriculture 5.0: a review on crop data management. *Agronomy*, 10(2): 207
- Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. 2019. The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3(3): 430–439
- Schick T, Dwivedi-Yu J, Dessì R, Raileanu R, Lomeli M, Zettlemoyer L, Cancedda N, Scialom T. 2023. Toolformer: language models can teach themselves to use tools. arXiv: 2302.04761
- Sitokostantinou V, Porras EDS, Bautista JC, Piles M, Athanasiadis I, Kerner H, Martini G, Sweet LB, Tsoumas I, Zscheischler J, et al. 2024. Causal machine learning for sustainable agroecosystems. arXiv: 2408.13155
- Tang WT, Hu ZL. 2025. Potato disease and pest question classification based on prompt engineering and gated convolution. *Agriculture*, 15(5): 493
- Touvron H, Lavril T, Izacard G, Martinet X, Lachaux MA, Lacroix T, Rozière B, Goyal N, Hambro E, Azhar F, et al. 2023. LLaMA: open and efficient foundation language models. arXiv: 2302.13971
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J. 2017. Attention is all you need//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. ACM, pp. 6000–6010
- Wang HF, Li JW, Wu H, Hovy E, Sun Y. 2023. Pre-trained language models and their applications. *Engineering*, 25: 51–65
- Wang HW, Zhao RX. 2024. Knowledge graph of agricultural engineering technology based on large language model. *Displays*, 85: 102820
- Wang JG, Yi XM, Guo RT, Jin H, Xu P, Li SJ, Wang XY, Guo XZ, Li CM, Xu XH, et al. 2021. *Milvus*: a purpose-built vector data man-

- agement system//Proceedings of the 2021 International Conference on Management of Data. ACM, pp. 2614–2627
- Wang JQ, Liu ZL, Zhao L, Wu ZH, Ma C, Yu SG, Dai HX, Yang QS, Liu YH, Zhang SY, et al. 2023. Review of large vision models and visual prompt engineering. *Meta-Radiology*, 1(3): 100047
- Wang L, Ma C, Feng XY, Zhang ZY, Yang H, Zhang JS, Chen ZY, Tang JK, Chen X, Lin YK, et al. 2024. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6): 186345
- Wang SH, Sun XF, Li XY, Ouyang RB, Wu F, Zhang TW, Li JW, Wang GY, Guo C. 2025. GPT-NER: named entity recognition via large language models//Findings of the Association for Computational Linguistics: NAACL 2025. ACL, pp. 4257–4275
- Wang TW, Zhao LG, Li BH, Liu XW, Xu WK, Li J. 2022. Recognition and counting of typical apple pests based on deep learning. *Ecological Informatics*, 68: 101556
- Wang YQ, Wang FH, Chen WB, Lv BW, Liu MC, Kong XY, Zhao CJ, Pan ZC. 2025. A large language model for multimodal identification of crop diseases and pests. *Scientific Reports*, 15: 21959
- Wang YZ, Kordi Y, Mishra S, Liu A, Smith NA, Khashabi D, Hajishirzi H. 2023. Self-instruct: aligning language models with self-generated instructions//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). ACL, pp. 13484–13508
- Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, Yogatama D, Bosma M, Zhou D, Metzler D, et al. 2022a. Emergent abilities of large language models. arXiv: 2206.07682
- Wei J, Wang XZ, Schuurmans D, Bosma M. 2022b. Chain-of-thought prompting elicits reasoning in large language models//Proceedings of the 36th International Conference on Neural Information Processing Systems. ACM, pp. 24824–24837
- Wu HR, Zhao CJ, Li JC. 2025. Agri-QA net: multimodal fusion large language model architecture for crop knowledge question-answering system. *Smart Agriculture*, 7(1): 1–10 (in Chinese) [吴华瑞, 赵春江, 李静晨. 2025. 基于多模态融合大模型架构 Agri-QA Net 的作物知识问答系统. *智慧农业*, 7(1): 1–10]
- Wu HY, Chen M, Guo QH. 2025. Rice pest and disease expert system based on large language model and knowledge graph. *Transactions of the Chinese Society of Agricultural Engineering*, 41(22): 244–255 (in Chinese) [吴桓宇, 陈明, 郭钦鸿. 2025. 基于大语言模型与知识图谱的水稻病虫害专家系统. *农业工程学报*, 41(22): 244–255]
- Xie LH, Han RL, Xie SH, Chen DJ, Chen YX. 2021. Multi-view fusion network for crop disease recognition//2021 The 5th International Conference on Algorithms, Computing and Systems. ACM, pp. 121–126
- Yan CQ, Liang ZY, Cheng H, Li SY, Yang GP, Li ZW, Yin L, Qu JJ, Wang J, Wu GH, et al. 2025. CDIP-ChatGLM3: a dual-model approach integrating computer vision and language modeling for crop disease identification and prescription. *Computers and Electronics in Agriculture*, 236: 110442
- Yan FR, Yu P, Chen X. 2025. LTNER: large language model tagging for named entity recognition with contextualized entity marking. *Pattern Recognition*. 2025. Cham: Springer, pp. 399–411
- Yang J, Yang W X, Yang S, He L, Zhang D. 2025. Intelligent question-answering method for crop pest and disease based on adaptive hybrid retrieval-augmented large model. *Smart Agriculture*, <https://link.cnki.net/urlid/10.1681.S.20251009.1353.004> (in Chinese) [杨俊, 杨婉霞, 杨森, 何亮, 张娣. 2025. 自适应混合检索增强大模型的农作物病虫害智能问答方法. *智慧农业*, <https://link.cnki.net/urlid/10.1681.S.20251009.1353.004>]
- Yang S, Feng Q, Yan WB, Zhou WW, Yang WX. 2025. Recognizing few-shot crop diseases using multimodal-guided visual Transformer. *Transactions of the Chinese Society of Agricultural Engineering*, 41(6): 195–203 (in Chinese) [杨森, 冯全, 阎文博, 周文伟, 杨婉霞. 2025. 多模态引导视觉 Transformer 的小样本农作物病害识别. *农业工程学报*, 41(6): 195–203]
- Yang ST, Liu ZH, Mayer W, Ding NP, Wang Y, Huang Y, Wu PF, Li WL, Li L, Zhang HY, et al. 2025. ShizishanGPT: an agricultural large language model integrating tools and Resources//Web Information Systems Engineering: WISE 2024. 2025. Singapore: Springer, pp. 284–298
- Yang T, Mei YP, Xu L, Yu HH, Chen YY. 2024. Application of question answering systems for intelligent agriculture production and sustainable management: a review. *Resources, Conservation and Recycling*, 204: 107497
- Yasunaga M, Ren HY, Bosselut A, Liang P, Leskovec J. 2021. QA-GNN: reasoning with language models and knowledge graphs for question answering//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: human language technologies. Online Association for Computational Linguistics, pp. 535–546
- Yin SK, Fu CY, Zhao SR, Li K, Sun X, Xu T, Chen EH. 2024. A survey on multimodal large language models. *National Science Review*, 11(12): nwae403
- Yu DH, Zhu CG, Yang YM, Zeng M. 2020. JAKET: joint pre-training of knowledge graph and language understanding. arXiv: 2010.00796
- Yuan ZP, Liu K, Peng RL, Li SB, Leybourne D, Musa N, Huang H, Yang P. 2025. PestGPT: leveraging large language models and IoT for timely and customized recommendation generation in sustainable pest management. *IEEE Internet of Things Magazine*, 8(1): 26–33
- Zamfirescu-Pereira JD, Wong RY, Hartmann B, Yang Q. 2023. Why johnny can't prompt: how non-AI experts try (and fail) to design LLM prompts//Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. ACM, pp. 1–21
- Zha ZW, Wang JA, Li ZX, Zhu XR, Song W, Xiao YH. 2023. M<sup>2</sup>ConceptBase: a fine-grained aligned concept-centric multimodal knowledge base. arXiv: 2312.10417
- Zhan Y, Yuan Y, Xiong ZT. 2024. Mono3DVG: 3D visual grounding in monocular images//The Thirty-Eighth AAAI Conference on Artificial Intelligence, pp. 6988–6996
- Zhang CK, Jiang D, Wan TY, Rao Y, Jin X, Wang T, Wang XB, Li JJ,

- Zhang W, Shao X. 2025. Hallucination Alleviation-based smart decision for early soybean cultivation in greenhouse and field scenarios. *Computers and Electronics in Agriculture*, 238: 110811
- Zhang CS, Zhang LJ, Wu HR, Wang CS, Chen C, Zhu HJ, Liang FF. 2024. Chinese named entity recognition for agricultural diseases based on entity-related visual prompts injection. *Computers and Electronics in Agriculture*, 227: 109493
- Zhang JY, Guo M, Geng YJ, Li M, Zhang YL, Geng N. 2021. Chinese named entity recognition for apple diseases and pests based on character augmentation. *Computers and Electronics in Agriculture*, 190: 106464
- Zhang KP, Ma L, Cui BB, Li X, Zhang BQ, Xie N. 2024. Visual large language model for wheat disease diagnosis in the wild. *Computers and Electronics in Agriculture*, 227: 109587
- Zhang YQ, Fan QJ, Chen X, Li M, Zhao ZY, Li FZ, Guo LF. 2025. IPM-AgriGPT: a large language model for pest and disease management with a G-EA framework and agricultural contextual reasoning. *Mathematics*, 13(4): 566
- Zhang ZL, Guo HS, Zhang Y, Ke Z, Guo YH, Sun KL, Tong SS, He ZZ, Zhang L, Gui LY, et al. 2025. Towards accurate field counting of small pests with visual prompts. *Computers and Electronics in Agriculture*, 237: 110635
- Zhao B, Jin WQ, Del Ser J, Yang G. 2023. ChatAgri: Exploring potentials of ChatGPT on cross-linguistic agricultural text classification. *Neurocomputing*, 557: 126708
- Zhao CJ. 2019. State-of-the-art and recommended developmental strategic objectives of smart agriculture. *Smart Agriculture*, 1(1): 1-7 (in Chinese) [赵春江. 2019. 智慧农业发展现状及战略目标研究. *智慧农业*, 1(1): 1-7]
- Zhao XY, Chen BY, Ji MX, Wang XY, Yan YH, Zhang JM, Liu S, Ye MY, Lv CL. 2024. Implementation of large language models and agricultural knowledge graphs for efficient plant disease detection. *Agriculture*, 14(8): 1359
- Zhou C, Li Q, Li C, Yu J, Liu YX, Wang GJ, Zhang K, Ji C, Yan QB, He LF, et al. 2025. A comprehensive survey on pretrained foundation models: a history from BERT to ChatGPT. *International Journal of Machine Learning and Cybernetics*, 16(12): 9851-9915
- Zhou J, Li JX, Wang CS, Wu HR, Zhao CJ, Teng GF. 2021. Crop disease identification and interpretation method based on multimodal deep learning. *Computers and Electronics in Agriculture*, 189: 106408
- Zhu HF, Shi WM, Guo XY, Lyu ST, Yang RB, Han ZZ. 2025. Potato disease detection and prevention using multimodal AI and large language model. *Computers and Electronics in Agriculture*, 229: 109824
- Zhu HY, Lin CZ, Liu GQ, Wang DN, Qin S, Li AJ, Xu JL, He Y. 2024. Intelligent agriculture: deep learning in UAV-based remote sensing imagery for crop diseases and pests detection. *Frontiers in Plant Science*, 15: 1435016
- Zhu Y, Wang Y, Mu JY, Li Y, Qiang JP, Yuan YH, Wu XD. 2024. Short text classification with soft knowledgeable prompt-tuning. *Expert Systems with Applications*, 246: 123248

(责任编辑:张俊芳)