doi:10.6041/j.issn.1000-1298.2025.09.019

农业领域大语言模型研究进展

王耀君　徐国威　朱建军　别宇辉
（中国农业大学信息与电气工程学院，北京 100083）

摘要：大语言模型（LLMs）作为人工智能领域发展的核心驱动力，凭借其强大的语义理解、逻辑推理与多模态信息处理能力，实现了对海量非结构化数据的深度解析与智能生成。近年来，大语言模型技术凭借其在处理复杂农业信息、辅助精准决策方面的卓越表现，在提升知识获取效率、赋能病虫害智能诊断、优化生产管理等方面展现出巨大潜力，有望为推动智慧农业向更高阶的认知智能发展提供核心技术支撑。本文梳理了大语言模型从基础模型演进到应用生态繁荣的全景，分析了其在农业知识问答、多模态应用和决策支持系统中的研究现状，并指出了当前研究在数据稀缺、模型可靠性与部署成本等方面面临的挑战；进一步，归纳了以检索增强生成（RAG）、多模态融合和智能体（Agent）技术为核心的解决方案，并探讨了其在克服模型知识局限、融合多源数据等方面的关键作用；通过剖析病虫害智能诊断、精准作物管理等典型应用实例，阐明了大语言模型赋能现代农业的价值与意义。最后，对大语言模型在农业领域的未来发展进行了展望。其技术演进将聚焦于模型与知识图谱的深度结合、多模态感知能力的全面增强，以及基于多智能体的农业决策支持系统建立。这些进展将共同推动农业智能从被动的辅助工具向主动的决策伙伴角色转变，最终为实现农业生产精细化管理和可持续发展注入强大的科技动能。

关键词：农业大语言模型；检索增强生成；农业知识问答系统；智能体；多模态；农业决策支持系统
中图分类号：S24　　**文献标识码**：A　　**文章编号**：1000-1298(2025)09-0240-17　　**OSID**：

Survey of Research on Large Language Models in Agriculture

WANG Yaojun　XU Guowei　ZHU Jianjun　BIE Yuhui
(*College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: As a core driving force in artificial intelligence, large language models (LLMs) leverage their powerful capabilities in semantic understanding, logical reasoning, and multimodal information processing to achieve deep analysis and intelligent generation from vast unstructured data. In recent years, through their exceptional performance in processing complex agricultural information and aiding precision decision-making, LLM technologies have shown immense potential to enhance knowledge acquisition, empower intelligent pest and disease diagnosis, and optimize production management. They were poised to provide foundational technological support for elevating smart farming to a more advanced stage of cognitive intelligence. The landscape of LLMs was systematically reviewed from the evolution of foundational models to their flourishing application ecosystem. It provided an in-depth analysis of the current research status in agricultural question answering, multimodal applications, and decision support systems, while also identifying challenges such as data scarcity, model reliability, and deployment costs. Furthermore, it was outlined the key solutions centered on retrieval-augmented generation (RAG), multimodal fusion, and agent technology, discussing their critical roles in overcoming model knowledge limitations and integrating multi-source data. Through an analysis of typical application cases like intelligent pest and disease diagnosis and precision crop management, the value and significance of LLMs in empowering modern agriculture were clarified. Finally, it provided an outlook on the future development of LLMs in agriculture. It posited that future technological evolution would focus on the deep coupling of models with knowledge graphs, the comprehensive enhancement of multimodal perceptual capabilities, and the establishment of responsible AI governance frameworks. These advancements were

收稿日期：2025－07－01　修回日期：2025－07－21
基金项目：国家重点研发计划项目（2024YFD2000805、2023YFD1701000）、北京市乡村振兴农业科技项目（NY2502020025）和中央高校建设世界一流大学（学科）和特色发展引导专项资金项目（2025AC030、2025AC044）
作者简介：王耀君（1985—），男，副教授，博士，主要从事农业大模型技术研究，E-mail：wangyaojun@cau.edu.cn

expected to collectively drive the transformation of agricultural intelligence from a passive auxiliary tool into an active decision-making partner, ultimately providing a powerful technological impetus for achieving refined management and sustainable development in agriculture.

Key words: large language model for agriculture; retrieval-augmented generation; agricultural knowledge question-answering system; agent; multimodal; agricultural decision support system

0　引言

当前,全球农业正面临着人口持续增长、气候变化加剧、自然资源日趋紧张以及农村劳动力结构性短缺等多重严峻挑战。为保障全球粮食安全与促进农业可持续发展,世界各国正积极推动农业向数字化、智能化转型升级。在这一进程中,以物联网(IoT)、遥感技术、无人机和各类传感器为代表的现代信息技术已在农业生产中广泛部署,催生了海量的、多源异构的农业大数据。然而,如何有效驾驭这些蕴含在文本、图像、时序数据中的复杂信息,并将其转化为能够直接指导生产实践的精准知识与智能决策,已成为"智慧农业"从自动化迈向更高阶认知智能阶段所必须攻克的关键瓶颈。传统的机器学习模型虽在特定任务(如图像分类、数据预测)上取得了显著成果,但往往依赖大量标注数据,且在处理非结构化文本、融合多模态信息和进行复杂逻辑推理方面能力有限,难以满足现代农业对综合性、系统性智能解决方案的迫切需求。

正是在这一背景下,以大语言模型(Large language models, LLMs)为核心驱动的新一轮人工智能浪潮,为破解农业信息处理难题、赋能农业认知智能提供了解决方案。大语言模型通过在超大规模语料上进行预训练,掌握了前所未有的语义理解、逻辑推理、知识整合与内容生成能力。它们不仅能够深度解析海量非结构化文本(如科研文献、农技推广手册、市场报告),还能通过与视觉、传感等其他模态信息融合,构建对农业生产场景更为全面、立体的认知。与依赖特定任务训练的"狭义 AI"不同,大语言模型展现出的"通用 AI"潜力,使其能够作为一种底层技术基座,灵活适应多样化的农业应用场景,从根本上改变农业知识的获取、传播与应用方式。

近年来,将大语言模型技术引入农业领域已成为全球前沿研究热点。初步研究与应用实践表明,LLMs 在赋能农业知识服务、革新病虫害智能诊断、优化生产管理决策等方面展现出巨大潜力。各项研究旨在提升农业生产效率,更致力于将先进的科技成果惠及每一位农业从业者,为推动农业生产精细化、绿色化和智能化发展注入强大动能。

鉴于此,系统性地梳理与审视大语言模型在农业领域的应用现状、核心技术、面临挑战与未来趋势,具有重要的理论与实践意义。本文旨在全景式地展现大语言模型技术从基础模型演进到应用生态繁荣的完整脉络。本文梳理其关键基础模型的演进过程,并阐述其当前的技术生态。分析其在农业知识问答、多模态应用和决策支持系统中的具体研究与实践。在此基础上,本文剖析当前研究在数据稀缺、模型可靠性与部署成本等方面面临的核心挑战,并归纳以检索增强生成(RAG)、多模态融合和智能体(Agent)技术为核心的解决方案。本文对大语言模型在农业领域的未来发展方向进行展望,以期为相关领域的科研人员与行业实践者提供有价值的参考,共同推动智慧农业向更高水平的认知智能发展。

1　大语言模型基础模型演进

1.1　关键基础模型演进

在 Transformer 架构提出之前,序列数据处理主要依赖循环神经网络 RNN(Recurrent neural networks, RNNs)及其变体,如长短期记忆网络(LSTM)和门控循环单元(GRU)。这类模型通过顺序处理输入序列,理论上能够捕捉长距离依赖。然而,其实践应用面临两个主要限制:①顺序计算的特性阻碍了模型并行化训练,效率较低。②梯度消失或梯度爆炸问题在长序列中依然存在,影响了模型学习远距离依赖关系的能力。

2017 年,VASWANI 等[1]提出了 Transformer 架构。该架构摒弃了 RNN 的循环结构,完全基于自注意力机制(Self-attention mechanism)进行序列信息的处理。

自注意力机制允许模型在处理序列中的任一词元(token)时,直接计算该词元与序列中所有其他词元之间的关联权重,并基于此进行加权聚合,从而生成新的上下文表示。这种机制使模型能够直接建模序列中任意两个位置间的依赖关系,不受距离限制,并有效地捕捉全局信息。由于计算不依赖于前一时刻的状态,Transformer 可以进行高度并行化的训练,显著提升硬件利用率和训练速度。

原始的 Transformer 架构由一个编码器(Encoder)和一个解码器(Decoder)组成(图 1),二者均由多层相同模块堆叠而成。该架构最初在机器

翻译任务上取得了优异性能,其核心组件为后续大语言模型发展奠定了基础。

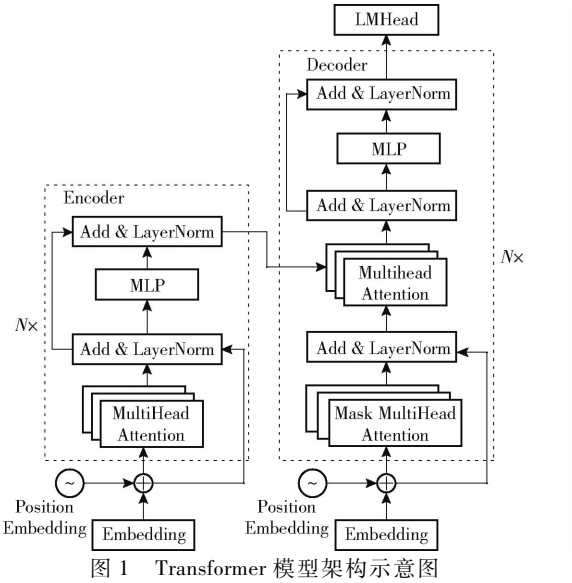


图 1 Transformer 模型架构示意图

Fig. 1 Schematic of Transformer model architecture

在 Transformer 提供了高效的架构后,如何利用大规模无标注文本数据提升模型泛化能力成为研究重点。2018 年,Google 提出的 BERT(Bidirectional encoder representations from transformers)模型[2],有效推广了 NLP 领域的“预训练-微调”(Pre-training and fine-tuning)范式。

BERT 的核心创新为预训练任务的设计。它使用 Transformer 的编码器部分(图 2),并通过以下两个任务学习深度的双向语言表示:

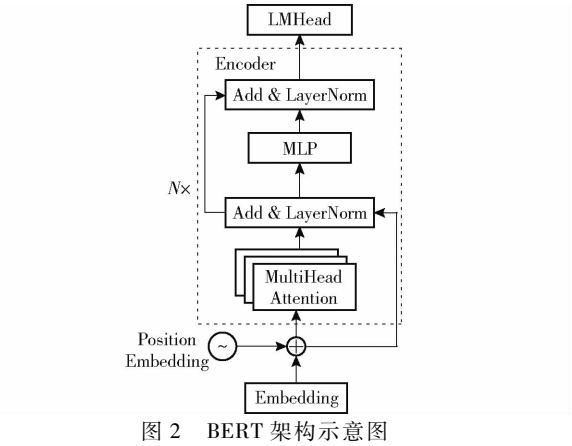


图 2 BERT 架构示意图

Fig. 2 Schematic of BERT architecture

掩码语言模型(Masked language model, MLM):在输入文本中随机屏蔽一部分词元,并训练模型利用该词元左右两侧的上下文信息来预测被屏蔽的词元。这一设计使模型必须学习词元间的深层语义与句法联系,由于同时利用了双向信息,其获得的表示是深度双向的。

下一句预测(Next sentence prediction, NSP):给定两个句子 A 和 B,判断句子 B 是否是 A 的直接后续句。此任务旨在让模型学习句子间的连贯性和逻辑关系。

通过在通用语料库上进行预训练,BERT 能够学习到丰富的通用语言知识。在处理下游具体任务(如文本分类、命名实体识别)时,仅需使用少量有标注数据对预训练模型进行微调(Fine-tuning),即可达到很高的性能。这一范式降低了为特定任务从头训练高性能模型的成本,促进了 NLP 技术的工程化应用。

BERT 的成功引发了研究界对预训练任务的进一步探索。有研究指出,BERT 的 MLM 任务以独立词元为单位进行掩码,可能忽略了文本中更高层次的语义单元。针对此问题,百度于 2019 年提出了 ERNIE(Enhanced representation through knowledge integration)模型[3]。ERNIE 的主要改进在于其知识掩码(Knowledge masking)策略,即对具有完整语义的单元(如实体、短语)进行整体掩码,促使模型学习先验知识。对于农业等专业领域,这种对知识单元进行整体建模的思路具有参考价值。

与 BERT 并行的另一条技术路线是 OpenAI 主导的 GPT(Generative pre-trained transformer)系列模型,其发展重点是文本生成(Text generation)。

GPT 模型采用 Transformer 的解码器(Decoder)架构(图 3)。其核心预训练任务是标准的自回归语言建模(Autoregressive language modeling):根据已经生成的上文内容,逐个预测下一个词元。这种从左到右的生成模式天然适用于文本生成任务。

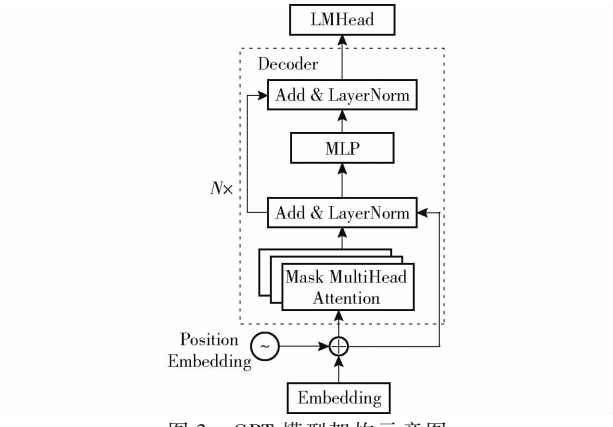


图 3 GPT 模型架构示意图

Fig. 3 Schematic of GPT model architecture

GPT 系列的发展体现了“规模定律”(Scaling law)的影响,即模型性能随着参数量、数据量和计算量的增加而持续提升:

GPT-1(2018 年)[4]:验证了生成式预训练模型在多种 NLP 任务上进行微调的有效性。

GPT-2(2019 年)[5]:显著增加了模型参数和训练数据量,模型表现出较强的零样本(Zero-shot)

能力。

GPT－3（2020 年）[6]：参数量扩展至 1 750 亿，展示出强大的情境学习（In-context learning）能力。用户仅需在输入提示（Prompt）中提供少量任务范例（Few-shot），模型即可理解任务模式并生成相应输出。

虽然 GPT－3 展示了 LLM 的强大能力，但其应用普及的一个关键节点是 ChatGPT 的发布。ChatGPT 基于 GPT－3.5/GPT－4 架构，其技术上的一个关键点在于引入了基于人类反馈的强化学习（Reinforcement learning from human feedback，RLHF）[7] 对模型进行对齐（Alignment）。

RLHF 旨在使模型输出更符合人类的指令和偏好，其流程通常包括 3 个阶段：

（1）监督微调（Supervised fine-tuning，SFT）：使用人工编写的高质量“指令－回答”数据对，对预训练 LLM 进行微调。

（2）训练奖励模型（Reward modeling，RM）：使模型针对同一指令生成多个回答，由人类对这些回答进行排序，并利用此数据训练一个奖励模型。

（3）强化学习优化（Reinforcement learning optimization）：以奖励模型作为环境，利用强化学习算法（如 PPO）进一步优化 SFT 模型。

RLHF 技术改善了人机交互的自然度和安全性，使得 LLM 技术能以更易用的形式被大众所接触，进而推动了相关技术和产业发展。

1.2　开源模型崛起与生态繁荣

在闭源商业模型快速发展的同时，开源社区也在积极推动 LLM 技术的普及，大幅降低了 LLM 的技术使用门槛，促进了技术广泛传播和二次开发。

2023 年 2 月，Meta 发布的 Llama 系列模型[8] 对开源生态产生了重要影响（图 4）。Llama 系列模型以相对较小的参数规模（如 7、13B）在多个公开基准上取得了与更大规模闭源模型相当的性能，这显示了高质量训练数据与训练策略的重要性。尽管 Llama 初版许可证限制了商业用途，但它催生了大量基于其进行指令微调的开源项目（如 Alpaca[9]、Vicuna[10]），活跃了开源社区。

随后，Meta 发布的 Llama 2[11] 提供了可免费用于研究和商业用途的版本，并包含了经过 RLHF 对齐的对话模型。Llama 2 的发布是开源 LLM 发展的一个关键节点，它为开发者提供了一个高性能、可定制的基座模型。这对农业等垂直领域具有重要意义，因为它降低了研究机构和企业利用自有领域数据进行模型微调、开发专用模型的门槛。

在 LLM 的全球发展中，中国的科技公司也推出了众多模型。阿里巴巴达摩院的 Qwen[12－16]（通义千问）系列是其中的代表性模型之一。Qwen 系列在设计上注重中英双语能力的平衡，并逐步扩展到多语言支持。其模型矩阵覆盖了从 1.8B 到 72B 等多个参数量级，并均进行了开源，为不同需求的用户提供了选择。此外，Qwen 团队也在多模态（Multimodality）方向进行了探索，推出了 Qwen－VL 等视觉语言模型。这类模型能够处理图像和文本输入，执行图文问答等任务，为 LLM 在需要视觉信息处理的农业场景（如病虫害图像识别）提供了技术路径。

除了大型科技公司，一些新兴 AI 公司也对 LLM 开源社区做出了贡献，DeepSeek AI 是其中之一。DeepSeek 最初以其代码生成模型 DeepSeek Coder[17－18] 获得关注，该模型在多个代码能力基准上表现突出。随后，DeepSeek 进一步开源了其通用对话模型[19－20]，在推理和数学等复杂任务上展现了强大的能力（图 5），为开源社区提供了又一个高性能选择，以及低成本训练方式，进一步推动了模型在复杂问题解决能力上的发展[21]。

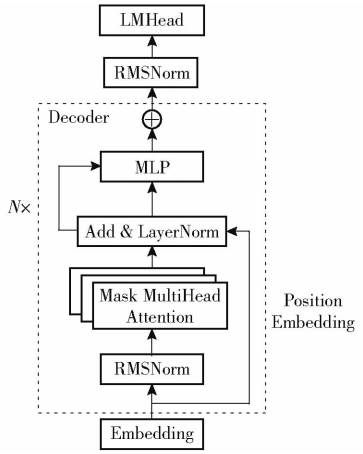


图 4　Llama 模型架构示意图

Fig. 4　Schematic of Llama model architecture

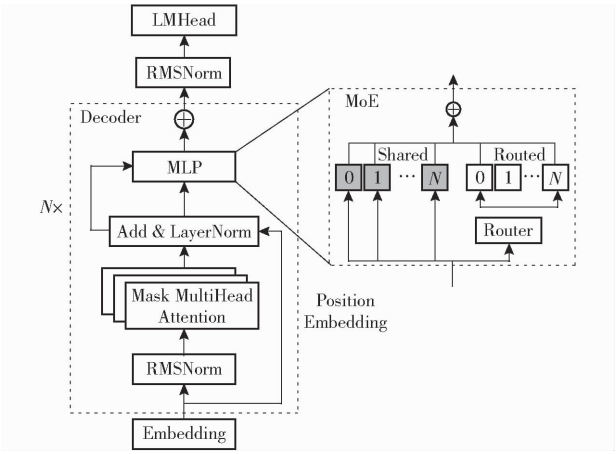


图 5　DeepSeek－R1 模型架构示意图

Fig. 5　Schematic of DeepSeek－R1 model architecture

1.3 从提示词到智能体:大语言模型生态的技术演进

随着基础模型的成熟,其应用生态也在快速演进。4 个关键技术方向相互补充,共同推动着大语言模型向更智能、更实用的方向发展。

检索增强生成(Retrieval-augmented generation, RAG)[22-24]技术通过结合外部知识库与大语言模型的生成能力,有效解决了模型知识更新滞后和幻觉问题。该技术核心思想是在生成回答前,先从相关文档库中检索出与查询最相关的信息片段,然后将这些信息作为上下文输入给语言模型进行生成。RAG 技术的优势在于能够动态获取最新信息,提高答案的准确性和可信度,同时降低重新训练大模型的成本,成为连接静态模型与动态知识的重要桥梁。

基于大语言模型的智能代理(Agent)技术代表了人工智能向自主决策和行动能力的重要进展。Agent 不仅具备理解和生成自然语言的能力,还能够制定计划、使用工具、与环境交互,并根据反馈调整策略。现代 Agent 系统通常包含感知、规划、记忆和执行等核心模块。通过多轮对话和工具调用,Agent 能够完成复杂的多步骤任务,如数据分析、代码编写、文档处理等。这种技术的发展预示着从被动语言理解转向主动任务执行的范式转变。

多模态大语言模型将文本理解能力扩展到视频、音频等多种模态,实现了跨模态的信息处理和生成。这类模型能够同时处理文本、图像、音频等不同类型的输入,并生成相应的多模态输出。多模态技术的核心挑战在于如何有效地对齐和融合不同模态的信息表示。通过预训练和微调策略,现代多模态模型在图像描述、视觉问答、文档理解等任务上展现出卓越性能。这一技术方向为 AI 系统提供了更加丰富的感知能力,使其能够更好地理解和描述现实世界。

提示词工程(Prompt engineering)[25-26]作为与大语言模型交互的重要技术,专注于设计和优化输入提示以获得最佳模型输出。有效的提示词设计能够显著提升模型在特定任务上的表现,而无需修改模型参数。提示词工程涵盖了多种技术,包括少样本学习(Few-shot learning)、思维链提示(Chain-of-thought)[27]、角色扮演等。随着模型能力增强,提示词工程正向更加结构化和自动化的方向发展,成为释放大语言模型潜力的关键技术手段。

1.4 讨论

历程始于 Transformer 架构的提出,其后以 BERT 和 GPT 系列为代表的模型分别在理解和生成任务上确立了“预训练-微调”范式和规模效应的重要性。ChatGPT 通过引入 RLHF 技术,极大地提升了模型的实用性。随后,以 Llama、Qwen、DeepSeek 为代表的开源模型的出现,显著降低了技术门槛,促进了技术的广泛应用和二次开发。

在这些基础模型之上,一个繁荣的技术生态正在形成:检索增强(RAG)解决了知识时效性问题,Agent 技术赋予了模型自主规划和行动的能力,多模态技术扩展了模型的感知边界,而提示词工程则优化了人机交互的效率。从基础模型的架构创新到开源生态的繁荣,再到应用层技术的不断涌现,大语言模型的发展脉络清晰地展现了其从理论走向实践、从通用走向专用的过程。对这些基础模型及其发展脉络的理解,是后续探讨其在农业领域具体应用的前提。

2 BERT 模型在农业领域的应用

随着信息技术的飞速发展,农业领域产生了海量的文本数据。如何从这些非结构化数据中高效地提取知识、洞察趋势,成为智慧农业发展的关键。在语言模型出现之前,以 BERT 为代表的预训练语言模型,凭借其强大的上下文理解和语义表征能力,在自然语言处理领域取得了突破性进展,并被广泛应用于解决农业领域的实际问题。研究人员利用 BERT 及其变体,在农业知识图谱构建、智能问答、文本分类与情感分析等方向取得了显著成果。例如,一项利用机器学习对过去 20 年间 120 万份关于中低收入国家小农研究分析表明,农业领域相关研究文献数量正呈指数级增长,而人工智能方法是有效梳理这些知识、发现研究空白(如营养、性别和包容性议题)的关键工具[28]。本文将围绕 BERT 模型在农业领域的具体应用加以综述。

2.1 农业知识图谱构建:实体识别与关系抽取

知识图谱(Knowledge graph)是实现智能农业的重要基础设施,而从海量文本中准确抽取出实体和关系,是构建知识图谱的核心环节。BERT 模型为这一任务提供了强有力的技术支持。

在农业命名实体识别(NER)方面,研究者致力于提升模型在专业领域的识别精度。例如,HUANG 等[29]在 2024 年提出了一种融合 BERT-wwm(Whole word masking)、双向长短期记忆网络(BiLSTM-CRF)和通道注意力机制(CA)的模型,该模型利用 BERT-wwm 对文本进行词嵌入,再通过下游网络充分捕捉上下文特征,显著提升了农业实体识别准确率、召回率和 F1 值,为构建农业知识图谱和问答系统奠定了坚实基础。

在实体关系抽取方面,BERT 同样表现出色。针对水稻表型知识图谱的构建,袁培森等[30]于 2021

年基于植物本体论对水稻的基因、环境、表型等实体间的关系进行了分类，并构建了基于 BERT 的关系抽取模型。实验结果表明，与卷积神经网络（CNN）等模型相比，BERT 模型的性能更优，精度达到 95.11%，F1 值高达 95.85%，证明了其在复杂农业实体关系抽取任务中的有效性。

此外，为了更好地适应农业领域，研究者还对 BERT 模型的微调策略和领域适应性进行了深入探讨。PANOUTSOPOULOS 等[31] 于 2024 年比较了通用 BERT、RoBERTa、Sci - BERT 以及领域专用的 Agriculture - BERT 在农业术语抽取任务上的表现，重点探索了不同微调配置（如冻结或更新特定网络层）对模型性能的影响，并发现针对不同类型的术语（已见词、同义词、新词），最佳的微调策略有所不同。该研究为农业领域语言模型优化提供了重要参考[31]。这些基础性工作最终服务于大型农业知识图谱自动化构建，如 CHEN 等[32] 于 2019 年提出的 AgriKG 项目，旨在利用自然语言处理和深度学习技术，自动从互联网的非结构化文本中识别农业实体并链接成图谱，支撑实体检索和智能问答等下游应用。

2.2　智能问答系统：提升信息检索效率

农业智能问答（Q&A）系统能够帮助农业生产者和技术人员快速获取所需信息，是知识应用的重要出口。BERT 模型的引入极大地提升了问答系统的准确性和效率。

问句分类是问答系统的关键前端模块，其准确性直接影响后续的答案检索。金宁等[33] 为改善农业问答问句短文本分类性能，提出了 BiGRU_MulCNN 分类模型。该模型利用双向门控循环单元（BiGRU）获取上下文特征，并结合多尺度卷积神经网络（CNN）进行多粒度特征提取，在包含 12 个类别的农业问答数据集上，准确率达到 95.9%。随着预训练语言模型的兴起，基于 BERT 及其变体的模型在问句意图识别上展现出强大能力。杨国峰等[34] 构建了基于 BERT 微调的常见作物病害问句分类模型。研究表明，与 Bi - LSTM 自注意力网络和 Transformer 模型相比，BERT 模型结构更简单、训练参数更少，且分类准确率高出 2 ~ 5 个百分点，在作物病害问句数据集上准确率达到 92.46%，能够高效地实现问句精准分类。然而，在病虫害等专业领域，一个问句常包含多种意图（如同时询问“症状”和“防治方法”），且存在语料不足的问题。为了克服这些挑战，韦婷婷等[35] 基于树状结构构建了农业病虫害问句的层级多标签分类体系，采用 ERNIE（BERT 的增强版）作为基础模型，并引入对抗训练方法来提高模型泛化能力，有效解决了传统单标签模型难以处理复杂问句以及数据稀疏性的难题。

在完整的问答流程中，研究者们进一步探索了如何优化 BERT 以应对农业场景的特有挑战，如语料库规模小、内容范围窄等。为此，HUANG 等[36] 提出了一种名为 Chains - BERT 的高性能模型，该模型基于半监督学习和对比学习对 BERT 网络进行改进，并结合了文本关系链来增强模型性能。该模型不仅在公开数据集上超越了多个主流模型，在农业场景的开放式问答任务中更是达到了业界领先水平，兼具高精度与高效率。

2.3　文本分类与情感分析：洞察数据价值

除知识图谱和问答系统，BERT 模型在更广泛的农业文本分类任务中也发挥着重要作用，尤其是在情感分析和特定信息检测方面。

在线农产品评论是连接生产者和消费者的重要桥梁，但其中存在的信息不对称问题常导致供需失衡。为了解决这一问题，CAO 等[37] 提出了一种基于改进 BERT 模型的农产品评论情感分析算法。该模型能够有效处理评论文本中表达不规范、特征稀疏等问题，通过提取包含完整语义信息的句子级特征向量，准确识别消费者的情感倾向。实验证明，改进后模型 F1 值较原始 BERT 模型提升 7.05 个百分点，达到 89.86%，为生产者优化产品和消费者决策提供了可靠的数据支持。

在处理特定领域且标注数据有限的文本分类任务时，研究者探索了更为先进的模型架构。例如，在植物健康灾害检测领域，由于带标签的数据稀缺，传统的微调方法效果有限。JIANG 等[38] 提出将 GAN - BERT（一种结合了生成对抗网络的半监督模型）与领域预训练模型 ChouBERT 相结合。研究发现，该组合能够有效利用未标记数据，提升文本分类器的泛化能力。尽管这种组合可能会增加训练过程的不稳定性，但研究者也针对性地提出了缓解方案，为在数据稀疏的农业子领域应用 BERT 模型提供了新的思路。

2.4　讨论

BERT 模型在农业领域的应用，其核心贡献在于处理专业的非结构化文本数据。研究表明，BERT 及其变体不仅在农业知识图谱构建（实体识别与关系抽取）、智能问答系统和文本分类与情感分析等任务中取得了优越性能，更重要的是，研究趋势已从通用模型微调转向更深层次的创新。通过领域自适应、模型融合与半监督学习等策略，研究者正有效克服农业场景下专业术语复杂、标注数据稀缺等特有挑战，为智慧农业提供了强大的技术支撑。

3 农业知识问答系统开发

随着人工智能技术,特别是大语言模型(LLM)的飞速发展,农业知识问答(QA)系统已成为智能农业领域的研究热点。这些系统旨在高效处理海量农业数据,为农业生产者和研究人员提供精准、及时的信息支持。下面阐述当前基于大语言模型开发农业知识问答系统时主要研究方向、技术路径与面临的挑战。

3.1 农业知识问答系统核心技术路径

当前农业知识问答系统的构建主要围绕 3 种核心的知识存储与利用方法:文本语料库、知识图谱和大型语言模型。这 3 种路径各有侧重,并在实践中常常相互融合[39]。

首先,基于文本语料库的方法是构建领域专用模型的基础。研究人员通过收集大量的专业文献、书籍和在线文本,构建特定领域的知识库。例如,为了增强模型在植物科学领域的专业性,YANG 等[40]构建了 PLLaMa 模型,该模型在一个包含超过 150 万篇植物科学学术文章的数据库上进行了增强训练,显著提升了模型对专业主题的理解能力。这种方法高度依赖于语料库质量与广度。

其次,知识图谱(Knowledge graph)与大语言模型的结合是提升问答准确性和可靠性的重要途径。知识图谱通过实体、关系和属性来结构化地表示知识,使信息查询更为高效和精准[39]。KUMAR 等[41]提出的 KrishiAI 架构便是这一思路的典型实现。该系统利用 LLM 将用户的自然语言问题转换为对 Neo4j 图数据库的 Cypher 查询,并将查询结果转换回自然语言答案。为了确保 LLM 生成查询的准确性,该架构还设计了一个净化器(sanitiser)模块进行验证,显著提高了系统可靠性。

最后,直接利用或微调通用大语言模型也是一条重要的技术路线。ZHAO 等[42]的研究(ChatAgri)探索了直接应用 ChatGPT 进行跨语言农业文本分类的潜力,发现即使未经特定农业数据微调,通用大模型也能有效应对数据稀缺、领域迁移等挑战,其表现可与经过微调的专业模型相媲美。这表明,通用大模型为农业信息化提供了一种低门槛、高效能的解决方案。此外,一些研究开始探索令 LLM 集成外部工具与资源,如 ShizishanGPT 模型,其设计理念便是将模型自身能力与外部工具相结合,以应对更复杂的农业问答任务[43]。

3.2 领域专用模型构建与评估

尽管通用大模型能力强大,但在处理高度专业化的农业问题时,其表现仍可能不及经过领域数据优化的专用模型。因此,构建和评估农业领域的专用模型成为一个关键研究方向。

在模型构建方面,研究者不仅关注模型的规模,也致力于解决特定场景的难题。PAN 等[44]针对农业领域少数民族语言支持不足的问题,提出了一种构建支持汉语和维吾尔语的轻量化农业大语言模型的方法。该研究通过光学字符识别(OCR)技术从农业书籍中提取中文语料,利用翻译 API 将其转换为维吾尔语料,并构建混合词表进行训练。这种方法不仅解决了维吾尔语农业语料稀缺的问题,其轻量化设计也降低了硬件要求,便于在资源受限的环境中部署。

在模型评估方面,建立领域专用的基准数据集至关重要。KPODO 等[45]针对农业推广(Agricultural extension)领域,提出了一个名为 AgXQA 的新型问答基准数据集。基于该数据集,其训练的领域专用模型 AgRoBERTa 在抽取式问答任务上的表现优于 GPT-3.5 等主流模型。更重要的是,还引入名为 AgEES 的定制化人工评估指标,通过与专家评估结果比对,进一步验证了领域专用模型的优势。这项工作凸显了在特定领域内,精细化的数据集和模型评估体系对于推动技术发展的必要性。

3.3 讨论

大语言模型在农业知识问答领域仍面临若干挑战。数据稀缺与质量问题:在特定子领域(如病虫害诊断)和少数民族语言中,高质量、结构化的数据依然匮乏[44]。如何高效、低成本地获取并处理这些数据是未来的研究重点。模型的准确性与可靠性:LLM 存在的“幻觉”问题在对信息准确性要求极高的农业领域是致命的。如 KrishiAI 中的验证模块和 PLLaMa 的国际专家小组验证等,都是确保模型回答可靠性的有益尝试[40-41]。模型部署与应用成本:大型模型训练和部署成本高昂,限制了其在广大农业场景中的应用。轻量化模型设计[44]和利用现有通用大模型 API 的策略[42]是解决这一问题的两个主要方向。领域适应与评估:如何使模型更好地适应农业领域的独特挑战,并建立科学、全面的评估体系来衡量其真实效能,是推动技术从实验室走向应用的关键[45]。

总而言之,未来的研究将趋向于构建知识图谱与大语言模型深度融合的混合智能系统,开发更多轻量化、多语言、高可靠性的领域专用模型,并建立完善的评测基准,从而真正实现人工智能技术对现代农业的赋能,助力农业可持续发展和乡村振兴。

4　农业领域大模型检索增强生成技术

随着大语言模型(LLM)在各行各业展现出巨大潜力,其在农业领域的应用也备受关注。然而,通用 LLM 通常面临知识局限性、信息过时和“幻觉”等问题,难以满足农业生产对精确性、实时性和地域特异性的高要求。为了克服这些挑战,检索增强生成(Retrieval-augmented generation, RAG)技术应运而生。RAG 通过在生成答案前从外部知识库中检索相关信息,并将其作为上下文提供给 LLM,从而显著提升模型响应的准确性、可靠性和领域相关性。下面将系统阐述 RAG 技术在农业领域的应用价值、关键技术路径与典型实践,并展望其未来发展方向。

4.1　RAG 在农业领域的价值与挑战

LLM 在农业领域的应用旨在通过智能服务提升农业生产的科学化水平,但其自身固有的局限性构成了主要障碍。首先,LLM 的知识来源于其训练数据,这些数据通常是静态且泛化的,缺乏针对特定农业领域(如特定作物的病虫害防治、特定地域的种植技术)的深度和实时知识。其次,模型在处理其知识范围外或模糊的问题时,容易产生不准确甚至完全错误的“幻觉”内容,这在指导农业生产时可能是灾难性的。

RAG 技术被认为是解决上述问题的有效途径。它将 LLM 的强大语言生成能力与外部、可控的知识库相结合,实现了知识的动态扩展。有研究明确指出,RAG 是解决 LLM 知识库有限这一核心问题的关键方法,尤其是在构建农业决策支持系统(DSS)时,它能够提升系统的可信赖度、可解释性并减少偏见[46]。另有研究认为,农业知识图谱(KGs)与 LLM 的融合具有巨大潜力,能够有效管理和结构化农业数据,而 RAG 正是实现这种互补融合的关键技术之一,为农业知识智能服务提供了新的思路[47]。

然而,将 RAG 应用于农业领域也面临挑战。农业信息浩瀚且高度碎片化,如何构建高质量、结构化的农业专用知识库是 RAG 成功应用的前提。无论是从非结构化的文本中自动构建农业知识图谱(AGKG)[48],还是整理权威农业文档,都需要大量的专业投入。

4.2　基于 RAG 的农业应用实践

近年来,研究人员已开始探索 RAG 在具体农业场景中的应用,尤其是在病虫害诊断与管理方面取得了显著成果。

在咖啡种植业中,一项研究集成了 YOLO v8 对象检测模型与 RAG 技术,开发了一套精准农业系统[49]。该系统首先通过 YOLO v8 识别咖啡叶片的疾病特征,随后利用 RAG 驱动的语言模型提供上下文感知的诊断和修复建议。这种方法不仅有效解决了 LLM 的幻觉问题,还实现了动态的疾病识别与 remediation 策略,旨在保障粮食供应和推广环保农作方式[49]。

类似地,针对尼泊尔丘陵地区的柑橘产业,研究者开发了一款基于 RAG 的聊天机器人,用于辅助柑橘病害检测[50]。该系统集成了 MobileNetV2 模型进行疾病预测,并利用一个基于 Llama－3.2 的 RAG 聊天机器人分析叶片图像,为农民提供关于疾病预防和果园管理的实时指导。该应用通过移动端界面,使先进的人工智能技术变得易于农民使用。这些实践案例表明,RAG 能够将先进的计算机视觉技术与 LLM 的推理能力有效结合,为农业生产提供精准、及时的决策支持。

4.3　RAG 技术路径与性能优化

在农业领域应用 LLM 时,开发者通常有两种主流技术路径来融入专有数据:微调(Fine-tuning)和 RAG。微调是将新知识内化到模型参数中,而 RAG 则是在推理时从外部获取知识。一项针对农业场景的深入研究对比了这两种方法,发现其并非相互排斥,而是可以相互补充、累积增益[51]。研究结果显示,对 Llama2－13B 等模型进行微调后,在特定农业知识问答任务上准确率提升超过 6 个百分点;在此基础上再结合 RAG,准确率可以进一步提升 5 个百分点。这证明了 RAG 与微调相结合是构建高性能农业 LLM 的有效策略[51]。

研究者们为了进一步优化 RAG 在农业场景中的性能,提出了更精细化的方法。例如,一项名为“知识同化”(Knowledge assimilation)的研究针对农业领域数据稀缺、SFT(监督微调)忽略农业专有词元以及 RAG 检索精度不足等问题,提出了一个知识引导的农业大语言模型(KALLM)[52]。在 RAG 层面,该研究引入了一种基于主题匹配的“自反思 RAG 机制”(self-reflective RAG),旨在提升证据检索的准确性。通过在包含 22 万问答对的高质量农业数据集上进行实验,证明该机制显著优于现有的 SFT－RAG 流程,在响应的流畅性、准确性和领域保真度方面均达到领先水平。

4.4　视觉 RAG

尽管 RAG 已在农业领域展现出巨大潜力,其未来的发展仍有广阔空间。当前主流的 RAG 系统主要依赖于文本检索,这在处理包含图表、图像和复杂布局的农业文档(如技术手册、土壤分析报告)时会造成信息损失。

针对这一问题,一项前沿研究提出了 VisRAG,

一个基于视觉的检索增强生成的应用[53]。VisRAG 不先将文档解析为纯文本,而是使用视觉语言模型(VLM)直接将包含布局和图像的文档页面作为整体进行嵌入和检索。实验证明,这种方法能够最大程度地保留和利用文档中的多模态信息,相较于传统基于文本的 RAG,在检索和生成阶段均表现出优越性,实现了 20% ~ 40% 的端到端性能提升。VisRAG 的理念为未来农业 RAG 系统的发展指明了新方向,即从单模态文本检索迈向多模态信息检索,以更全面地理解和利用丰富的农业知识载体。

4.5 讨论

作为应对大语言模型(LLM)在农业领域应用中固有局限性的核心策略,检索增强生成(RAG)技术正受到广泛关注。面对 LLM 存在的知识陈旧、领域深度不足和“幻觉”问题,RAG 通过动态链接外部权威知识库(如专业文档、知识图谱),显著提升了模型响应的准确性、实时性和可靠性。在实践层面,RAG 已成功应用于构建病虫害诊断等精准农业系统,常与计算机视觉等技术结合,提供从识别到决策的闭环支持。研究进一步揭示,RAG 与微调(Fine-tuning)并非互斥关系,而是相辅相成的技术路径,二者结合能实现性能最大化。此外,通过引入“自反思”等精细化检索机制,可以进一步优化 RAG 的性能。

展望未来,该领域的前沿方向正从纯文本检索迈向多模态的视觉 RAG,旨在更全面地理解和利用包含图表与图像的复杂农业文档,解决信息损失问题。综上,RAG 技术作为推动大语言模型(LLM)在农业领域实际应用的关键,将在未来着重与知识图谱及多模态技术深度融合,进而构建更为智能且可靠的农业决策支持系统。

5 多模态农业大模型

现代智慧农业的发展建立在对海量、多源数据的有效利用之上,即所谓的“农业大数据”[54]。这些数据不仅包含文本信息,还广泛涉及图像、遥感影像、环境传感器数据等多种形态。随着大语言模型(LLM)在自然语言处理领域取得突破,其强大的推理和生成能力开始向更多领域渗透。然而,单一文本模态显然无法满足现代智慧农业的复杂需求。因此,能够理解和处理多种数据类型的多模态大模型,已成为当前农业人工智能研究的前沿。研究表明,多模态大模型能够融合不同来源的数据,从而在作物病虫害诊断、长势监测、产量预测和自动化管理等任务中发挥关键作用[55]。

5.1 视觉-语言模型赋能农业识别与诊断

在多模态大模型的研究中,视觉-语言模型(Vision-language model, VLM)是一个核心分支,它旨在通过融合图像和文本信息,实现跨模态的理解与生成。在农业领域,传统的深度学习方法,如使用 YOLO v5 进行害虫及危害症状的精确检测[56-60]、利用卷积神经网络(CNN)进行病害分类[61-62]、构建多尺度特征融合网络(如 MSSNet)进行高分辨率遥感图像的作物分类[63],已经取得了显著成效,如图 6 所示。然而,这些模型通常输出的是类别标签或分割掩码,缺乏对病因、农事建议等深层信息的理解和生成能力。

VLM 的出现正弥补了这一不足。研究人员开始通过微调预训练的 VLM 来适应农业专业任务。例如,YU 等[64]提出了 AgriVLM 框架,该框架利用 Q-former 结构作为图像编码器和语言模型之间的桥梁,实现了农业图像与文本数据的跨模态融合。通



(a) 基于改进YOLO v5l的病虫害检测[56]　(b) 基于害虫数量评估灾害程度[61]　(c) 基于MSSNet的作物区域分类[63]

(d) 基于AgriVLM的农业问题和回答[64]　(e) 基于GPT的交互式诊断[65]　(f) 基于视觉大模型的苹果健康检测[66]

图 6 视觉语言模型农业应用示例图像

Fig. 6 Visual language model agricultural research example images

过低秩自适应(LoRA)微调技术,该模型在作物病害识别和生长阶段判断任务上准确率超过 90%。

除了端到端的 VLM,将视觉感知模型与 LLM 相结合的流水线方法也展现了巨大潜力。QING 等[65]提出了一种“GPT 辅助诊断”模式,首先用一个轻量化的 YOLOPC 网络处理农业图像,再通过创新的图文映射方法将视觉结果输入 GPT-4 进行逻辑推理,最终生成专业的农业诊断报告,推理准确率达到 90%。类似地,ROUMELIOTIS 等[66]的研究直接比较了多模态大模型 GPT-4o 与传统 CNN 模型 ResNet-50 在植物病害分类上的表现。结果显示,经过微调的 GPT-4o 准确率(98.12%)略高于 ResNet-50,且泛化能力更强,这表明通用多模态大模型在经过领域适配后,有潜力成为精准农业系统中可扩展的智能诊断核心。

5.2　多源数据融合提升分析与预测精度

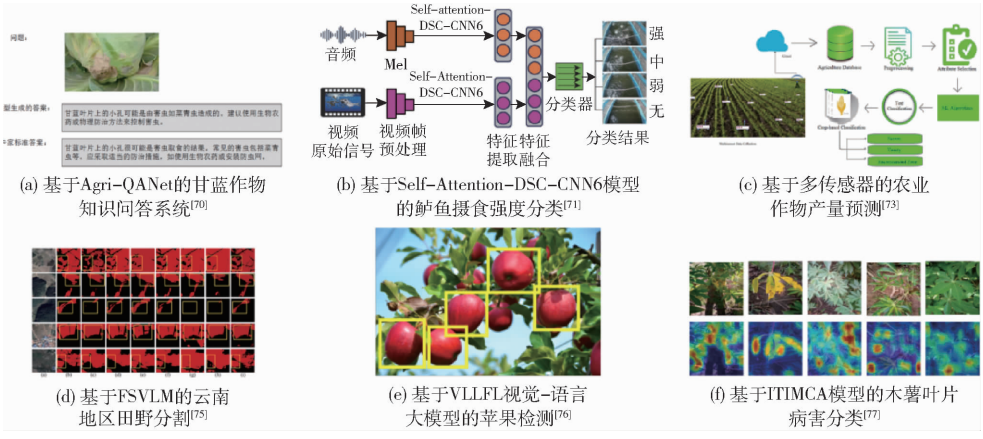
农业生产是一个复杂的系统,其状态不仅体现在作物的视觉表型上,还受到环境、气象、土壤等多维因素的共同影响。因此,融合文本、图像之外的更多模态数据,是构建全面、精准的农业决策支持系统的关键。当前的机器学习应用已经广泛涉及多传感器数据融合,例如在自主喷洒系统中,整合多光谱传感器、LiDAR 等数据的机器人和无人机,能够实现对病虫害的精准识别和农药的精确施用[67]。然而,正如李道亮等[68]所指出的,当前农业领域的多模态研究仍存在数据利用程度低、有效特征提取困难、融合方式单一等问题。因此,深入探索不同模态间的复杂交互关系,是未来通过多模态融合技术推动农业精准管理、提高生产效率的核心。

多模态大模型正在将这种融合能力提升到新的高度。LU 等[69]提出的方法将图像、文本和传感器数据进行有效整合,利用 Transformer 模型深度分析复杂的农业问题。该模型在病害检测、图像描述生成和目标检测等多项任务中均取得了优异表现,病害检测准确率高达 94%,验证了融合多源数据在提升模型性能方面的巨大潜力。吴华瑞等[70]则进一步拓展了模态的种类,设计了名为 Agri-QA Net 的作物知识问答系统,该系统能够同时处理文本、音频和图像 3 种模态的数据,通过基于 Transformer 的融合层和跨模态注意力机制,为用户提供关于甘蓝种植的专业问答服务,其多模态输入准确率达到 89.5%。李道亮等[71]提出了一种多模态融合框架,将音频(摄食音效)与视频信息结合,并引入 Self-Attention-DSC-CNN6 网络,实现对鲈鱼摄食强度分类,分类准确率高达 90.24%,表明多源数据融合在复杂环境下提升分类效果的显著优势。

除了即时传感器数据,历史和时序数据对于农业预测也至关重要。SHAMSUDDIN 等[72]开发了一个多模态深度学习模型,用于玉米产量的早期预测。该模型创新性地融合了田间的 RGB 图像、作物物候表型数据以及历史气象数据,并考虑了时间序列效应,结果表明,引入历史气象数据显著增强了模型的预测和解释能力。同样,REYANA 等[73]也展示了一种多传感器机器学习方法(MMLA),通过融合传感器数据进行文本分类,为棉花、小麦等 8 种作物提供种植建议,其采用的随机森林算法表现出良好的分类效果。虽然并非所有多模态融合研究都基于 LLM,但其融合策略具有重要的参考价值。例如,ZHOU 等[74]在研究杨树干旱胁迫评估时发现,在多种数据融合方法中,“特征层融合”比数据层融合或决策层融合表现更优,这为未来多模态农业大模型如何高效整合不同来源的特征提供了启示,具体示例如图 7 所示。

5.3　面向特定场景的创新架构与应用

随着多模态大模型技术的发展,研究人员不仅在通用任务上验证其有效性,还开始针对农业领域的特定挑战,设计创新的模型架构和应用范式,以解决数据稀疏、隐私保护和特定场景应用等实际问题。



(a) 基于Agri-QANet的甘蓝作物知识问答系统[70]　(b) 基于Self-Attention-DSC-CNN6模型的鲈鱼摄食强度分类[71]　(c) 基于多传感器的农业作物产量预测[73]

(d) 基于FSVLM的云南地区田野分割[75]　(e) 基于VLLFL视觉-语言大模型的苹果检测[76]　(f) 基于ITIMCA模型的木薯叶片病害分类[77]

图 7　多模态大模型农业研究示例图像

Fig. 7　Example images of multimodal large model agricultural research

在遥感影像分析领域，WU 等[75]首次探索了将语言引导的 VLM 应用于农田分割，构建了一个专用的农田图像-文本（FIT）数据集，并设计了农田分割模型 FSVLM。该模型利用语言描述来精确表达农田随季节和生长阶段变化的物候景观，有效弥补了传统视觉方法在理解复杂特征方面的不足，在泛化性和分割精度上均优于现有方法。

为解决农业数据分散、隐私敏感以及模型训练通信成本高的问题，LI 等[76]提出了一种基于 VLM 的轻量级联邦学习框架（VLLFL）。该框架在保护各农场数据隐私的前提下，通过联邦学习共同训练一个紧凑的提示生成器，以提升部署在不同农场上的 VLM 性能。实验证明，该框架在将 VLM 性能提升 14.53% 的同时，减少了 99.3% 的通信成本，为构建高效、可扩展且保护隐私的智慧农业应用提供了新方案。

此外，针对农业领域训练数据稀缺的普遍挑战，研究人员利用多模态预训练模型的强大能力来解决小样本学习问题。LI 等[77]提出了一种名为 ITIMCA 的方法，用于自然环境下的木薯叶部病害分类。该方法整合了 CLIP 多模态预训练和跨注意力机制，通过图像-文本对比学习，有效利用文本信息来增强模型在小样本场景下的识别能力。该模型在自建数据集上准确率达到 78.00%，证明了多模态方法在应对数据稀缺问题时的有效性。

5.4 讨论

通过融合视觉与语言信息，这些模型将农业 AI 的能力从传统的病虫害识别提升至生成诊断报告和农事建议的认知层面。同时，通过进一步整合环境传感器、气象时序等多维数据，模型在作物产量预测和长势监测等复杂任务中的分析精度与可靠性也得到了显著增强，实现了对农业生产环境更全面的理解。

在此基础上，研究正向更具针对性的创新架构演进，例如利用语言引导遥感影像分割、通过联邦学习解决数据隐私与通信成本问题，以及借助多模态预训练应对小样本难题。这些前沿探索共同表明，多模态技术正推动农业 AI 从单一感知向综合认知迈进，其发展趋势是构建能够全面分析多维信息的集成化智能系统，为智慧农业的深化发展提供核心动力。

6 基于大语言模型的农业决策支持系统

随着农业 4.0 向农业 5.0 时代的演进，数据已成为驱动现代农业发展的核心要素。高效的数据管理与分析是实现精准决策、提升生产力与可持续性的关键。近年来，以大语言模型（LLM）为代表的人工智能技术取得了突破性进展，为构建新一代农业决策支持系统提供了强大的工具。

6.1 农业智能决策系统的演进与大模型的角色

传统智能农业（Smart farming）依赖于传感器、物联网（IoT）等技术收集数据，通过数据驱动的管理模式优化资源利用，实现增产增效与环境保护[78]。然而，传统机器学习模型在处理和应用这些海量、异构的农业数据时，常面临泛化能力不足、需要大量标注数据以及对特定领域知识依赖性强等局限性[79]。

大语言模型为解决这些问题提供了新的范式。与传统模型不同，预训练的大模型（LMs）利用海量跨领域数据进行训练，能够以更少的任务特定微调来处理多样化的任务，展现出强大的语言理解、推理和生成能力[79-80]。在农业领域，大语言模型不仅限于处理文本数据，其技术理念已扩展至大型视觉模型（LVMs）、多模态大语言模型（MLLMs）和基于强化学习的大模型（RLLMs），为农业智能提供了全面的技术基础[79]。这些模型能够整合来自无人机、多光谱相机、激光雷达以及地面传感器（如土壤湿度、pH 值、温湿度传感器）等多源数据，形成对作物健康、土壤状况和环境因素的综合理解，从而为精准作物管理提供决策支持[81-83]。通过结合边缘计算，这些复杂的 AI 模型还能在田间地头实现快速响应，进一步提升农业管理的效率与自动化水平。具体示例如图 8 所示。因此，大语言模型正成为连接海量农业数据



(a) 油田地图的检测识别[78] (b) 基于地图的可变速率决策系统[81] (c) 农业平台的人机交互界面[84] (d) VineScout机器人的图形用户决策系统[85]

图 8 农业决策系统示例图像

Fig. 8 Example image of agricultural decision-making system

与精准决策之间的关键桥梁，推动农业决策系统从数据驱动向智能驱动的深刻变革。

6.2 大语言模型在农业决策中的关键应用

大语言模型在农业决策系统中的应用已从概念走向实践，并在人机交互、数据管理、知识生成和信息分析等多个层面展现出巨大潜力。

在提升人机交互友好性方面，研究人员致力于构建自然语言接口，使非专业人员也能便捷地操控复杂的农业测控平台。例如，ZHANG 等[84]提出了一种基于 BERT 的 Seq2Seq 模型，通过特殊的掩码机制，能够将用户输入的自然语言指令（如“打开一号大棚的灌溉系统”）高效地解析并转换为机器可执行的结构化 AOM（Action - Object - Member）语句，显著改善了农业系统的用户体验和操作效率。

在实现自动化数据管理与分析方面，大语言模型被用作智能“副驾驶”（Copilot），以应对现代农业中数据量巨大、异构性强且处理流程复杂的挑战。PAN 等[85]开发了一个名为 ADMA Copilot 的多智能体系统，该系统中的 LLM 控制器能够理解用户的分析意图，自主规划数据处理流程，并协调其他智能体完成数据整理、集成、加载和分析等任务，有效促进了农业数据的“可发现、可访问、可互操作和可重用”（FAIR）原则的实现。

在生成具体农事建议方面，大语言模型能够融合多源信息，提供高精度、高相关性的决策支持。PARK 等[86]设计了一种多轮提示（Multi-round prompt）框架，利用 LLMs（如 GPT - 4、Claude2）整合最新的气象预报、土壤状况和作物数据，迭代优化农业气象建议。实验证明，该方法在准确性和上下文相关性上均优于传统模型，展现了其在实际农业生产决策中的应用价值。

大语言模型在从海量文献和案例中提取、分析和综合知识方面也发挥着重要作用。ZHONG 等[87]利用 ChatGPT 构建了一个问答式信息提取框架，用于分析全球农业适应气候变化的案例数据库。该框架高效地从大量报告中提取出案例分布、气候胁迫因素、适应措施和制约因素等关键信息，从而识别出全球农业适应措施的宏观趋势，为政策制定者提供了宝贵的参考依据。

6.3 讨论

大语言模型在农业领域的广泛应用，不仅带来了技术层面的革新，也引发了关于其社会经济影响、潜在风险和未来治理的深刻思考。近年来国内外农业领域的主要大模型包括神农大模型[88]、天工开悟农业大模型[89]、范蠡大模型[90]、孺子牛大模型[91]、丰登大模型[92]、狮子山大模型（ShizishanGPT）[43]、后稷农业大模型[93]、奇稷大模型[94]以及雄小农[95]等。

从对劳动力的影响来看，LLMs 的引入正在重塑农业劳动市场的动态。研究表明，大语言模型主要在认知功能方面表现出色，能够显著增强或部分替代依赖数据分析、规划和管理的任务。然而，对于农业生产中必需的物理、心理和感官能力，LLMs 无法完全替代。因此，部分以认知任务为主的农业岗位面临被替代的风险，而另一些岗位则可能迎来增强效应，即 LLMs 自动化处理常规认知任务，使人类从业者能专注于更具创造性的非例行工作[96]。

与此同时，LLMs 的推广也伴随着一系列不容忽视的风险。DE CLERCQ 等[97]指出，农业错误信息的传播、大量农民敏感数据的收集与隐私问题，以及对农业就业岗位的潜在威胁，都是需要重点关注的议题。这些风险凸显了在技术大规模普及前，农业政策制定者有必要思考并建立相应的框架和指南，以确保 LLMs 在粮食生产系统中得到负责任的应用。国内外农业领域主要大模型如表 1 所示。

7 未来展望

随着大语言模型在农业领域的应用研究不断深化，其技术演进与产业落地正面临着从“单点突破”向“系统融合”转型的关键时期。尽管在知识问答、病虫害诊断等方面已取得显著成效，但模型在领域知识的深度、多源数据的综合感知能力以及决策的自主性与可靠性方面仍存在提升空间。这些挑战也为未来的技术创新指明了方向。围绕知识融合深度化、感知能力多模态化和决策系统自主化三大维度，展望农业大语言模型的发展路径和应用前景。

7.1 知识融合深度化

当前大模型知识局限且“幻觉”频发，通用大模型缺乏农业领域深度、实时的专业知识，在处理复杂问题时易产生不准确甚至错误的回答（即“幻觉”），严重影响其在生产决策中的可靠性。发展大语言模型与农业知识图谱的深度耦合技术，通过模型直接在知识图谱上进行结构化推理，而非仅检索文本片段，可从根本上提升回答的逻辑性与准确性，确保决策依据的可追溯与可解释。

7.2 感知能力多模态化

当前大模型对农业场景的感知能力单一，应用多局限于文本与图像处理，无法全面理解由环境、气象、土壤等多维因素构成的复杂农业系统，限制了其在精准管理中的应用深度。未来需要构建全息感知

表 1 国内外农业大模型

Tab. 1 Large language agricultural models in China and worldwide

支持模态	应用场景	模型名称	发布单位	发布时间	主要特点	文献序号
视频 + 音频 + 图像 + 文本	农业知识问答、语义理解、文本摘要生成及决策推理	神农大模型	中国农业大学	1.0 版:2023 - 12 2.0 版: 2024 - 07 - 13	广泛覆盖育种、种植、养殖、农业遥感及气象等多个农业应用场景	[88]
音频 + 图像 + 文本	农业问答大模型	天工开悟农业大模型	哈尔滨工业大学智慧农场技术与系统全国重点实验室	2024 - 11 - 02	对品种推荐、标准化种植、病虫害防治等问题进行智能回答	[89]
视频 + 音频 + 图像 + 文本	渔业大模型	范蠡大模型	中国农业大学国家数字渔业创新中心	2024 - 06 - 15	为渔业养殖工人、管理经营者和政府决策部门提供全面、精准的智能化支持	[90]
图像 + 文本	奶业全产业链智能决策	孺子牛大模型	中国农业大学	2025 - 01 - 15	基于 GLM - 4 大语言模型架构,以提升产业效率、优化资源配置、增强决策能力	[91]
图像 + 文本	种业育种知识问答与决策支持	丰登大模型	上海 AI 实验室、崖州湾国家实验室、中国农业大学	2024 - 04 - 28	首个种业大语言模型,整合分散知识、融合跨学科信息,已开放水稻知识接口	[92]
图像 + 文本	农业知识问答系统	狮子山大模型	华中农业大学、南澳大学	2024 - 09 - 20	将模型自身能力与外部工具相结合,应对复杂农业问答任务	[43]
图像 + 文本	作物栽培与农业管理	后稷农业大模型	西北农林科技大学	2025 - 06 - 18	融合 NLP 与计算机视觉能力,支持农业语义理解、图像识别、知识问答与生产决策推理	[93]
图像 + 文本	果蔬、作物全生育期智能管理	奇稷大模型	北京市农林科学院	2024 - 05 - 30	覆盖蔬菜、果树、粮经作物全生育期知识,支持图像识别与栽培问答	[94]
图像 + 文本	雄安农业知识问答与政策服务	雄小农	雄安新区管理委员会、中国科学院自动化研究所	2024 - 03	面向雄安新区涉农政策、技术推广、生产管理的本地化大模型	[95]
文本	植物科学领域知识问答	PLLaMa	加利福尼亚大学、林肯大学、中国农业科学院	2024 - 01 - 03	在超过 150 万篇植物科学学术文章数据库上进行增强训练	[40]
文本	农业知识问答系统	KrishiAI	印度理工学院、西悉尼大学	2025 - 02 - 04	利用 LLM 将自然语言问题转换为 Cypher 查询,并设计净化器模块验证生成查询的准确性	[41]
图像 + 语言	病害诊断	AgriVLM	浙江大学	2024 - 09 - 17	Q-former 桥接架构,LoRA 微调	[64]
图像 + 语言	农业诊断	GPT - 4 + YOLOPC	华南农业大学	2023 - 10	视觉结果映射到文本推理	[65]
图像 + 语言	昆虫大模型	Insect - Foundation	阿肯色大学	2025 - 02 - 14	专门的农业昆虫大模型	[98]

多模态大模型。通过融合遥感影像、物联网传感器时序数据及激光雷达点云等多源信息,建立对农田生态系统的“数字孪生”级理解,实现从宏观环境到微观作物的动态、全方位监测与预测。

7.3 决策系统自主化

现有自主决策系统应用被动且操作复杂,多为被动式问答工具,难以主动执行复杂的管理任务。同时,高级农业系统的操作对用户专业技能要求高,不利于技术普及。未来方向多为研发基于多智能体的自主决策系统。通过构建分别负责诊断、规划、市场分析等任务的专业智能体,并由一个“指挥官”智能体协同调度,用户仅需通过自然语言下达高级指令,系统即可自主完成任务规划、工具调用与执行,实现农业生产全流程的自动化与智能化。

7.4 成本与隐私问题

大型模型训练与部署成本高昂,且农业数据涉及农户隐私与生产安全,这些因素共同阻碍了技术在中小农户和资源受限地区的广泛应用。推广模型轻量化与联邦学习技术,通过模型剪枝、量化等技术降低部署门槛,使轻量化大模型可在边缘设备或普通终端上运行,更加靠近生产端实践;同时,利用联邦学习框架,在保护各方数据隐私的前提下,协同训练出性能更强的领域模型,实现技术普惠化与安全应用。

参 考 文 献

[1] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998 - 6008.

[2] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, 2019: 4171 - 4186.

[3] ZHANG Z, HAN X, LIU Z, et al. ERNIE: enhanced language representation with informative entities[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, 2019: 1441 - 1451.

[4] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training[EB/OL]. (2018 - 06 - 11)[2025 - 06 - 18]. https://openai.com/blog/language-unsupervised/.

[5] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners[EB/OL]. (2019 - 02 - 14)[2025 - 06 - 18]. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

[6] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1877 - 1901.

[7] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback[J]. Advances in Neural Information Processing Systems, 2022, 35: 27730 - 27744.

[8] TOUVRON H, LAVRIL T, IZACARD G, et al. Llama: open and efficient foundation language models[EB/OL]. (2023 - 02 - 27)[2025 - 06 - 18]. https://doi.org/10.48550/arXiv.2302.13971.

[9] TAORI R, GULRAJANI I, ZHANG T, et al. Alpaca: a strong, replicable instruction-following model[EB/OL]. (2023 - 03 - 13)[2025 - 06 - 18]. https://crfm.stanford.edu/2023/03/13/alpaca.html.

[10] CHIANG W, LI Z, LIN Z, et al. Vicuna: an open-source chatbot impressing GPT - 4 with 90% * ChatGPT quality[EB/OL]. (2023 - 03 - 30)[2025 - 06 - 18]. https://lmsys.org/blog/2023 - 03 - 30-vicuna/.

[11] TOUVRON H, MARTIN L, STONE K, et al. Llama 2: open foundation and fine-tuned chat models[EB/OL]. (2023 - 07 - 18)[2025 - 06 - 18]. https://doi.org/10.48550/arXiv.2307.09288.

[12] BAI J, BAI S, CHU Y, et al. Qwen technical report[EB/OL]. (2023 - 09 - 28). arXiv:2309.16609. https://doi.org/10.48550/arXiv.2309.16609.

[13] BAI J, BAI S, CHEN S, et al. Qwen1.5: a family of open-source language models[EB/OL]. (2024 - 03 - 06)[2025 - 06 - 18]. arXiv:2403.04255. https://doi.org/10.48550/arXiv.2403.04255.

[14] BAI J, BAI S, FAN Y, et al. Qwen2: the new generation of open-source large language models[EB/OL]. (2024 - 06 - 07)[2025 - 06 - 18]. arXiv:2406.04832. https://doi.org/10.48550/arXiv.2406.04832.

[15] CHU Y, BAI J, CHEN Q, et al. Qwen-Audio: a family of versatile audio-language models[EB/OL]. (2024 - 05 - 15)[2025 - 06 - 18]. arXiv:2405.09406. https://doi.org/10.48550/arXiv.2405.09406.

[16] YANG A, LI A, YANG B, et al. Qwen3: Qwen3 technical report[EB/OL]. (2025 - 05 - 15)[2025 - 06 - 18]. arXiv:2505.09388. https://doi.org/10.48550/arXiv.2505.09388.

[17] BI Q, LI Q, SHEN Y, et al. DeepSeek LLM: scaling open-source language models with longtermism[EB/OL]. (2024 - 01 - 05)[2025 - 06 - 18]. arXiv:2401.02954. https://doi.org/10.48550/arXiv.2401.02954.

[18] BI Q, LI Q, DAI W, et al. DeepSeek - Coder: when the large language model meets programming-the rise of code intelligence[EB/OL]. (2024 - 01 - 25)[2025 - 06 - 18]. arXiv:2401.14196. https://doi.org/10.48550/arXiv.2401.14196.

[19] LU Y, GUO Z, FAN S, et al. DeepSeek - VL: towards real-world vision-language understanding[EB/OL]. (2024 - 03 - 08)[2025 - 06 - 18]. arXiv:2403.05525. https://doi.org/10.48550/arXiv.2403.05525.

[20] BI Q, LI Q, SHEN Y, et al. DeepSeek - V2: a strong open-source MoE model[EB/OL]. (2024 - 05 - 07)[2025 - 06 - 18]. arXiv:2405.04434. https://doi.org/10.48550/arXiv.2405.04434.

[21] GUO D, YANG D, ZHANG H, et al. DeepSeek - R1: incentivizing reasoning capability in llms via reinforcement learningdang[EB/OL]. (2025 - 01 - 20)[2025 - 06 - 18]. arXiv:2501.12948. https://doi.org/10.48550/arXiv.2501.12948.

[22] LEWIS P, PEREZ E, PIKTUS A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks[J]. Advances in Neural Information Processing Systems, 2020, 33: 9459 - 9474.

[23] GAO Y, XIONG Y, GAO X, et al. Retrieval-augmented generation for large language models: a survey[EB/OL]. (2023 - 12 - 18)[2025 - 06 - 18]. arXiv:2312.10997. https://doi.org/10.48550/arXiv.2312.10997.

[24] ZHAO S, YANG Y, WANG Z, et al. Retrieval augmented generation (RAG) and beyond: a comprehensive survey on how to make your llms use external data more wisely[EB/OL]. (2024 - 09 - 23)[2025 - 06 - 18]. arXiv:2409.14924. https://doi.org/10.48550/arXiv.2409.14924.

[25] TAM D, MENON R R, BANSAL M, et al. Improving and simplifying pattern exploiting training[C]// Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2021: 4980 - 4991.

[26] SCHICK T, SCHUTZE H. Exploiting cloze questions for few shot text classification and natural language inference[EB/OL].

(2020 - 01 - 21) [2025 - 06 - 18]. arXiv:2001.07676. https://doi.org/10.48550/arXiv.2001.07676.

[27] WEI J, WANG X, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[J]. Advances in Neural Information Processing Systems, 2022, 35: 24824 - 24837.

[28] PORCIELLO J, LIPPER L, IVANINA M. Using machine learning to evaluate 1.2 million studies on small-scale farming and post-production food systems in low- and middle-income countries[J]. Frontiers in Sustainable Food Systems, 2023, 6: 1013701.

[29] HUANG Q, TAO Y, WU Z, et al. Based on BERT-wwm for agricultural named entity recognition[J]. Agronomy, 2024, 14(6): 1217.

[30] 袁培森, 李润隆, 王翀, 等. 基于 BERT 的水稻表型知识图谱实体关系抽取研究[J]. 农业机械学报, 2021, 52(5): 151 - 158.
YUAN Peisen, LI Runlong, WANG Chong, et al. Research on entity-relation extraction for rice phenotype knowledge graph based on BERT[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 151 - 158. (in Chinese)

[31] PANOUTSOPOULOS H, ESPEJO-GARCIA B, RAAIJMAKERS S, et al. Investigating the effect of different fine-tuning configuration scenarios on agricultural term extraction using BERT[J]. Computers and Electronics in Agriculture, 2024, 225: 109268.

[32] CHEN Y, KUANG J, CHENG D, et al. AgriKG: an agricultural knowledge graph and its applications[J]. Database Systems for Advanced Applications, 2019, 11448: 533 - 537.

[33] 金宁,赵春江,吴华瑞,等. 基于 BiGRU_MulCNN 的农业问答问句分类技术研究[J]. 农业机械学报, 2020, 51(5): 199 - 206.
JIN Ning, ZHAO Chunjiang, WU Huarui, et al. Research on agricultural question classification technology based on BiGRU_MulCNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 199 - 206. (in Chinese)

[34] 杨国峰, 杨勇. 基于 BERT 的常见作物病害问答系统问句分类[J]. 计算机应用, 2020, 40(6): 1580 - 1586.
YANG Guofeng, YANG Yong. Question classification for common crop disease question answering system based on BERT[J]. Computer Applications, 2020, 40(6): 1580 - 1586. (in Chinese)

[35] 韦婷婷, 葛晓月, 熊俊涛. 基于层级多标签的农业病虫害问句分类方法[J]. 农业机械学报, 2024, 55(1): 263 - 269, 435.
WEI Tingting, GE Xiaoyue, XIONG Juntao. Hierarchical multi-label classification of agricultural pest and disease interrogative question[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 263 - 269, 435. (in Chinese)

[36] HUANG Y, LIU J, LV C. Chains - BERT: a high-performance semi-supervised and contrastive learning-based automatic question-and-answering model for agricultural scenarios[J]. Applied Sciences, 2023, 13(5): 2924.

[37] CAO Y, SUN Z, LI L, et al. A study of sentiment analysis algorithms for agricultural product reviews based on improved BERT model[J]. Symmetry, 2022, 14(8): 1604.

[38] JIANG S, CORMIER S, ANGARITA R, et al. Improving text mining in plant health domain with GAN and/or pre-trained language model[J]. Frontiers in Artificial Intelligence, 2023, 6: 1072329.

[39] YANG T, MEI Y, XU L, et al. Application of question answering systems for intelligent agriculture production and sustainable management: a review[J]. Resources, Conservation and Recycling, 2024, 204: 107497.

[40] YANG X, GAO J, XUE W, et al. PLLaMa: an open-source large language model for plant science[EB/OL]. (2024 - 01 - 03) [2025 - 06 - 18]. https://doi.org/10.48550/arXiv.2401.01600.

[41] KUMAR M, WILSON S, GOEL N, et al. KrishiAI: architecture for harnessing capabilities of LLMs for delivery of accurate cultivation information to farmers[M]//SAINI M K, GOEL N, MIGUEZ M, et al. Agricultural-centric computation. Springer Nature Switzerland, 2025: 68 - 79.

[42] ZHAO B, JIN W, DEL SER J, et al. ChatAgri: exploring potentials of ChatGPT on cross-linguistic agricultural text classification[J]. Neurocomputing, 2023, 557: 126708.

[43] YANG S, LIU Z, MAYER W, et al. ShizishanGPT: an agricultural large language model integrating tools and resources[M]//BARHAMGI M, WANG H, WANG X, et al. Web information systems engineering. Springer, 2025: 284 - 298.

[44] PAN K, ZHANG X, CHEN L. Research on the training and application methods of a lightweight agricultural domain-specific large language model supporting Mandarin Chinese and Uyghur[J]. Applied Sciences, 2024, 14(13): 5764.

[45] KPODO J, KORDJAMSHIDI P, NEJADHASHEMI A P. AgXQA: a benchmark for advanced agricultural extension question answering[J]. Computers and Electronics in Agriculture, 2024, 225: 109349.

[46] VIZNIUK A, DIACHENKO G, LAKTIONOV I, et al. A comprehensive survey of retrieval-augmented large language models for decision making in agriculture: unsolved problems and research opportunities[J]. Journal of Artificial Intelligence and Soft Computing Research, 2025, 15(2): 115 - 146.

[47] GONG R, LI X. The application progress and research trends of knowledge graphs and large language models in agriculture [J]. Computers and Electronics in Agriculture, 2025, 235: 110396.

[48] WANG H, ZHAO R. Knowledge graph of agricultural engineering technology based on large language model[J]. Displays, 2024, 85: 102820.

[49] SELVA K S, KHAN A K M A, BANDAY I A, et al. Overcoming LLM challenges using RAG-driven precision in coffee leaf disease remediation[C]//2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS), 2024: 1 - 6.

[50] LAMSAL R R, ACHARYA P, POKHREL R, et al. Integrating machine learning and RAG-based chatbot for mandarin orange

disease detection in hilly region of nepal[EB/OL]. (2025-02-27)[2025-06-18]. https://doi.org/10.21203/rs.3.rs-6105402/v1.

[51] BALAGUER A, BENARA V, DE CUNHA R L F F, et al. RAG vs fine-tuning: pipelines, tradeoffs, and a case study on agriculture[EB/OL]. (2024-01-16)[2025-06-18]. http://arxiv.org/abs/2401.08406.

[52] JIANG J, YAN L, LIU H, et al. Knowledge assimilation: implementing knowledge-guided agricultural large language model [J]. Knowledge-Based Systems, 2025, 314: 113197.

[53] YU S, TANG C, XU B, et al. VisRAG: vision-based retrieval-augmented generation on multi-modality documents[EB/OL]. (2025-03-20)[2025-06-18]. http://arxiv.org/abs/2410.10594.

[54] WOLFERT S, GE L, VERDOUW C, et al. Big data in smart farming—a review[J]. Agricultural Systems, 2017, 153: 69-80.

[55] ZHU H, QIN S, SU M, et al. Harnessing large vision and language models in agriculture: a review[EB/OL]. (2024-04-10)[2025-06-18]. http://arxiv.org/abs/2404.07267.

[56] CHAKRABARTY S, SHASHANK P R, DEB C K, et al. Deep learning-based accurate detection of insects and damage in cruciferous crops using YOLOv5[J]. Smart Agricultural Technology, 2024, 9: 100663.

[57] JAFAR A, BIBI N, NAQVI R A, et al. Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations[J]. Frontiers in Plant Science, 2024, 15: 1356260.

[58] DOLATABADIAN A. Image-based crop disease detection using machine learning[J/OL]. Plant Pathology, 2025. doi:10.1111/ppa.14006.

[59] 张惠莉, 代晨龙, 任景龙, 等. 基于 YOLO v8-GSGF 模型的葡萄病害识别方法研究[J]. 农业机械学报, 2024, 55(11): 75-83.
ZHANG Huili, DAI Chenlong, REN Jinglong, et al. Research on grape disease recognition method based on YOLO v8-GSGF model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 75-83. (in Chinese)

[60] 左昊轩, 黄祺成, 杨佳昊, 等. 基于改进 YOLO v5s 的作物黄化曲叶病检测方法[J]. 农业机械学报, 2023, 54(增刊1): 230-238.
ZUO Haoxuan, HUANG Qicheng, YANG Jiahao, et al. Detection method for crop yellow leaf curl disease based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp.1): 230-238. (in Chinese)

[61] VENKATESWARA S M, PADMANABHAN J. Deep learning based agricultural pest monitoring and classification[J]. Scientific Reports, 2025, 15(1): 8684.

[62] 许景辉, 邵明烨, 王一琛, 等. 基于迁移学习的卷积神经网络玉米病害图像识别[J]. 农业机械学报, 2020, 51(2): 230-236,253.
XU Jinghui, SHAO Mingye, WANG Yichen, et al. Corn disease image recognition based on convolutional neural network with transfer learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230-236,253. (in Chinese)

[63] LU T, GAO M, WANG L. Crop classification in high-resolution remote sensing images based on multi-scale feature fusion semantic segmentation model[J]. Frontiers in Plant Science, 2023, 14: 1196634.

[64] YU P, LIN B. A framework for agricultural intelligent analysis based on a visual language large model[J]. Applied Sciences, 2024, 14(18): 8350.

[65] QING J, DENG X, LAN Y, et al. GPT-aided diagnosis on agricultural image based on a new light YOLOPC[J]. Computers and Electronics in Agriculture, 2023, 213: 108168.

[66] ROUMELIOTIS K I, SAPKOTA R, KARKEE M, et al. Plant disease detection through multimodal large language models and convolutional neural networks[EB/OL]. (2024-06-12)[2025-06-18]. https://doi.org/10.3390/preprints202406.0827.v1.

[67] TOSCANO F, FIORENTINO C, SANTANA L S, et al. Recent developments and future prospects in the integration of machine learning in mechanised systems for autonomous spraying: a brief review[J]. AgriEngineering, 2025, 7(5): 142.

[68] 李道亮, 赵晔, 杜壮壮. 农业领域多模态融合技术方法与应用研究进展[J]. 农业机械学报, 2025, 56(1): 1-15.
LI Daoliang, ZHAO Ye, DU Zhuangzhuang. Advances in multi-modal fusion techniques and applications in agricultural field [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(1): 1-15. (in Chinese)

[69] LU Y, LU X, ZHENG L, et al. Application of multimodal transformer model in intelligent agricultural disease detection and question-answering systems[J]. Plants, 2024, 13(7): 972.

[70] 吴华瑞, 赵春江, 李静晨. 基于多模态融合大模型架构 Agri-QA net 的作物知识问答系统[J]. 智慧农业(中英文), 2025, 7(1): 1-10.
WU Huarui, ZHAO Chunjiang, LI Jingchen. Crop knowledge question answering system based on multimodal fusion large model architecture Agri-QA net[J]. Smart Agriculture, 2025, 7(1): 1-10. (in Chinese)

[71] 李道亮, 李万超, 杜壮壮. 基于音视频信息融合与 Self-Attention-DSC-CNN6 网络的鲈鱼摄食强度分类方法[J]. 农业机械学报, 2025, 56(1): 16-24.
LI Daoliang, LI Wanchao, DU Zhuangzhuang. Classification method of largemouth bass feeding intensity based on audio-video information fusion and Self-Attention-DSC-CNN6 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(1): 16-24. (in Chinese)

[72] SHAMSUDDIN D, DANILEVICZ M F, AL-MAMUN H A, et al. Multimodal deep learning integration of image, weather, and phenotypic data under temporal effects for early prediction of maize yield[J]. Remote Sensing, 2024, 16(21): 4043.
[73] REYANA A, KAUTISH S, KARTHIK P M S, et al. Accelerating crop yield: multisensor data fusion and machine learning for agriculture text classification[J]. IEEE Access, 2023, 11: 20795 - 20805.
[74] ZHOU Z, ZHANG H, BIAN L, et al. Integrating sensor fusion with machine learning for comprehensive assessment of phenotypic traits and drought response in poplar species[J]. Plant Biotechnology Journal, 2025,23(7): 2464 - 2481.
[75] WU H, DU Z, ZHONG D, et al. FSVLM: a vision-language model for remote sensing farmland segmentation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 1 - 13.
[76] LI L, LI J, CHEN D, et al. VLLFL: a vision-language model based lightweight federated learning framework for smart agriculture[EB/OL]. (2025 - 01 - 16)[2025 - 06 - 18]. http://arxiv.org/abs/2501.09241.
[77] LI H, CHEN B, CHEN J, et al. ITIMCA: image-text information and cross-attention for multi-modal cassava leaf disease classification based on a novel multi-modal dataset in natural environments[J]. Crop Protection, 2025, 189: 106981.
[78] SAIZ-RUBIO V, ROVIRA-MÁS F. From smart farming towards agriculture 5.0: a review on crop data management[J]. Agronomy, 2020, 10(2): 207.
[79] SHAIKH T A, RASOOL T, VENINGSTON K, et al. The role of large language models in agriculture: harvesting the future with LLM intelligence[J]. Progress in Artificial Intelligence, 2025, 14:117 - 164.
[80] CAO Y. Cucumber disease recognition with small samples using image-text-label-based multi-modal language model[J]. Computers and Electronics in Agriculture, 2023, 211: 107993.
[81] KUMARI K, MIRZAKHANI N A, MIRZAEE S, et al. AI-driven future farming: achieving climate-smart and sustainable agriculture[J]. AgriEngineering, 2025, 7(3): 89.
[82] SOULTANE O B, SALIH-ALJ Y, ET-TAIBI B. Smart agriculture optimization: Integrating edge computing and AI for enhanced crop management[C]//2024 10th International Conference on Applied System Innovation (ICASI), 2024: 1 - 3.
[83] PERUMAL V K. Improving crop production using an agro-deep learning framework in precision agriculture[J]. BMC Bioinformatics, 2024, 25: 970.
[84] ZHANG Y, YAO S, WANG P, et al. Building natural language interfaces using natural language understanding and generation: a case study on human-machine interaction in agriculture[J]. Applied Sciences, 2022, 12(22): 11830.
[85] PAN Y, SUN J, YU H, et al. Building multi-agent copilot towards autonomous agricultural data management and analysis[C]// 2024 IEEE International Conference on Big Data (BigData), 2024: 4384 - 4393.
[86] PARK J J, CHOI S J. LLMs for enhanced agricultural meteorological recommendations[EB/OL]. (2024 - 08 - 08)[2025 - 06 - 18]. http://arxiv.org/abs/2408.04640.
[87] ZHONG J, ZHANG X, MA X. Global agricultural adaptation case database and trend analysis based on large language models [EB/OL]. (2024 - 09 - 02)[2025 - 06 - 18]. https://www.ssrn.com/abstract=4922214.
[88] 中国网财经. 中国农业大学“神农大模型 2.0”正式发布[EB/OL]. (2024 - 07 - 13)[2025 - 07 - 25]. https://finance.china.com.cn/news/20240713/6140961.shtml.
[89] 哈尔滨工业大学计算学部“天工开悟”团队. 天工开悟大模型[EB/OL]. [2025 - 07 - 25]. https://www.tgkwai.com/.
[90] 新京报客户端. 我国首个渔业大模型“范蠡大模型 1.0”发布[EB/OL]. (2024 - 06 - 16)[2025 - 07 - 25]. https://news.cau.edu.cn/mtndnew/98d7e0f834aa4f71aaa4d29a97b00290.htm.
[91] 中国农大新闻网. 首个国际奶业领域垂直大模型——孺子牛大模型- AI4DLLM 发布[EB/OL]. (2025 - 01 - 23)[2025 - 07 - 25]. https://news.cau.edu.cn/zhxwnew/c80b74a1b7fa4cae95fdff7be247895d.htm.
[92] 澎湃新闻. AI 育种大模型“丰登”发布:具备品种选育、栽培技术回答能力[EB/OL]. (2024 - 04 - 28)[2025 - 07 - 25]. https://www.thepaper.cn/newsDetail_forward_27198551.
[93] 科技日报. 西北农林科技大学发布“后稷农业大模型”[EB/OL]. (2025 - 06 - 06)[2025 - 07 - 25]. https://www.stdaily.com/web/gdxw/2025 - 06/06/content_350983.html.
[94] 世界农化网中文网. “奇稷(QeeG)大模型”——全国首个在农业领域具备千万量级知识资源的大模型发布[EB/OL]. (2025 - 06 - 03)[2025 - 07 - 25]. https://cn.agropages.com/News/NewsDetail - - - 34909.htm.
[95] 北京日报客户端. 雄安新区发布国内首个 DeepSeek 农业行业大模型[EB/OL]. (2025 - 02 - 27)[2025 - 07 - 25]. https://xinwen.bjd.com.cn/content/s67c05088e4b08edd28f58ec3.html.
[96] MARINOUDI V, BENOS L, VILLA C C, et al. Large language models impact on agricultural workforce dynamics: opportunity or risk? [J]. Smart Agricultural Technology, 2024, 9: 100677.
[97] DE CLERCQ D, NEHRING E, MAYNE H, et al. Large language models can help boost food production, but be mindful of their risks[J]. Frontiers in Artificial Intelligence, 2024, 7: 1326153.
[98] TRUONG T D. Insect-Foundation: a foundation model and large multimodal dataset for vision-language insect understanding [J/OL]. International Journal of Computer Vision. http://doi.org/10.1007/s11263-025-02521-4.