

# 基于多模态知识图谱的水稻施肥期判别方法

袁立存 周俊 戈为溪 郑彭元

(南京农业大学工学院, 南京 210031)

**摘要:** 目前水稻施肥时间的确定依赖于传统经验与人工巡田观察的综合判断, 难以满足当前农业智能化的发展。为此, 提出了一种基于多模态知识图谱的水稻施肥期判别方法, 综合利用了文本形式的经验信息和图像形式的视觉信息进行施肥期确定。首先构建单模态水稻施肥知识图谱, 利用依存句法分析提取返青肥、分蘖肥、穗肥、粒肥4个施肥期对应的跨模态特征短语, 结合 Chinese CLIP 模型得到它们与图像的匹配度以及与施肥期节点的权重后组成新的带有跨模态节点的三元组, 完成多模态水稻施肥知识图谱的构建; 然后基于多模态知识图谱计算输入信息的综合匹配度, 使用大田采集的图像进行交叉验证, 综合评估判别方法的准确性和稳定性确定各施肥期的判定阈值, 实现对该输入的施肥期判别。以实际采集的各施肥期当日及前、后5d的600幅图像测试判别方法的准确率, 结果表明, 基于多模态知识图谱的水稻施肥期判别方法总体准确率达到86.2%, 其中粒肥时期判别准确率最高, 为90.1%。该施肥期判别方法同时利用文本、图像两种模态的信息, 提高了信息利用率, 在真实场景下具有判别能力, 为水稻施肥期自动确定提供参考。

**关键词:** 水稻; 施肥期; 多模态; 知识图谱; CLIP 模型

中图分类号: TP391

文献标识码: A

文章编号: 1000-1298(2024)09-0163-11

OSID:



## Rice Fertilization Period Discrimination Method Based on Multi-modal Knowledge Graph

YUAN Licun ZHOU Jun GE Weixi ZHENG Pengyuan

(College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

**Abstract:** Currently, the determination of the optimal fertilization time for rice relies heavily on a combination of traditional experience and manual field inspection, which struggles to meet the demands of modern agricultural intelligence. In response, a method for rice fertilization period discrimination was introduced based on a multi-modal knowledge graph, integrating textual experiential information and visual cues for determining the fertilization period. Initially, a single-modal knowledge graph for rice fertilization was constructed. On this basis, cross-modal feature phrases corresponding to the four fertilization periods (re-greening, tillering, heading, and grain-filling) were extracted by using dependency syntax analysis. These phrases were then combined with the Chinese CLIP model to determine their match with images and their respective weights for the fertilization periods, forming new triplets with cross-modal nodes. This led to the creation of a multi-modal rice fertilization knowledge graph. Subsequently, the multi-modal knowledge graph was used to calculate the comprehensive matching degree of input information, and field-collected images were utilized for cross-validation. This process comprehensively evaluated the accuracy and stability of the discrimination method, thereby determining the decision thresholds for each fertilization period. The discrimination methods accuracy was tested by using 600 images captured on the day of each fertilization period and five days before and after. Results showed that the overall accuracy rate of the multi-modal knowledge graph-based rice fertilization period discrimination method was 86.2%, with the highest accuracy rate of 90.1% during the grain-filling period. By utilizing both textual and visual modalities, this method enhanced information utilization

收稿日期: 2023-12-19 修回日期: 2024-02-06

基金项目: 江苏省现代农机装备与技术示范推广项目(NJ2021-59)

作者简介: 袁立存(1993—), 男, 博士生, 主要从事知识图谱和决策系统研究, E-mail: 18260068128@163.com

通信作者: 周俊(1974—), 男, 教授, 博士生导师, 主要从事农业装备智能化技术与农业机器人研究, E-mail: zhoujun@njau.edu.cn

and demonstrated discriminative capability in real-world scenarios, offering a reference for the automated determination of rice fertilization periods.

**Key words:** rice; fertilization period; multi-modal; knowledge graph; CLIP model

0 引言

水稻作为我国第二大粮食作物,保持其稳产与增产是国家粮食安全的重要保障<sup>[1-2]</sup>。随着“智慧农业”<sup>[3]</sup>与“无人化农场”<sup>[4]</sup>等概念的提出,水稻生产的各环节也逐步向智能化发展。施肥管理是水稻生产的重要环节之一,它包括施肥方式、施肥量与施肥期等要素<sup>[5]</sup>。其中,正确选择水稻施肥期可以最大限度地提高水稻产量和品质,减少对土壤和水环境的污染,同时也可以降低成本和提高农民收益<sup>[6]</sup>。因此,开展水稻施肥期自动判别研究具有重要价值。

目前,对于水稻施肥期的判断主要采用人工巡田的方式,依赖于传统的观测和经验判断。这种方法主观性强且重复性差,难以满足缺乏种植经验的农户和水稻生产环节智能化的需求。已有研究者使用水稻多光谱数据根据植被指数的差异实现水稻生育期的辨别<sup>[7-8]</sup>。多光谱数据的采集成本高且受光照条件影响较大,使用 RGB 图像结合深度学习算法从颜色和纹理特征的角度识别水稻关键生育期也取得了较好表现<sup>[9-10]</sup>。尽管图像数据在一定程度上能够反映水稻生长阶段信息,但它在数据维度和环境适应性上存在限制。施肥期的判别还需综合考虑水稻品种、种植地点等其他因素。

知识图谱作为一种强大的信息整合工具,能很好地理解和组织大量的文本信息,在种植推荐<sup>[11]</sup>、农业知识问答<sup>[12]</sup>、作物病虫害诊断<sup>[13]</sup>等复杂的农业应用场景上具备优势。然而,仅有文本信息难以全面反映实际农业环境,同时也限制输入输出只能为单一文本模态。多模态知识图谱(Multi-modal knowledge graph, MMKG)通过整合文本、图像等多个模态的信息丰富了知识表达的维度,在农业检索问答系统<sup>[14]</sup>上展现了语义理解的深度与更好的性能。目前,将知识图谱应用在水稻施肥期判别应用场景的研究较少。现有成果通过构建单模态水稻施肥案例知识图谱进行检索与推理<sup>[15]</sup>,或是构建多模态水稻施肥知识图谱辅助决策<sup>[16]</sup>。前者可以根据案例得到施肥期的区间或相关的客观文本描述,后者根据多维度数据采集系统生成定量的施肥处方,但对于具体施肥期的确定仍然没有明确方案。

综上,为实现水稻施肥期自动且相对准确的判别,本文提出一种基于多模态知识图谱的水稻施肥

期判别方法,首先构建单模态水稻施肥知识图谱,利用 CLIP (Contrastive language-image pre-training) 模型<sup>[17]</sup>对水稻关键施肥期的图像和对应描述短语进行图文匹配,从而将图像和水稻施肥知识图谱的关键节点通过带有权重的跨模态特征描述短语连接,以实现水稻施肥多模态知识图谱的构建。利用多模态水稻施肥知识图谱,由水稻生长相关的图像或文本输入确定具体的水稻施肥期。

1 数据来源

1.1 文本数据

文本数据主要来源于国家水稻数据中心网、IRRI 水稻知识库、农业种植网中包含 360 个水稻品种的水稻施肥相关知识,中国知网中与水稻施肥相关的 196 篇文献以及标准 QX/T 468—2018《农业气象观测规范——水稻》。其中网页数据通过 Python 爬虫工具获取,文献和标准使用 FineReader 工具转化为可编辑文本数据。

1.2 图像数据

图像数据来源于 Python 软件从 1.1 节网络数据资源获取的对应施肥期的 153 幅水稻相关配图和 3 099 幅大田采集图像。

大田图像采集地点为:江苏省南京市高淳区禾田农业合作社与久保田合作试验田块(31.277 15°N, 118.849 88°E)和江苏省淮安市洪泽区三河镇祥发合作社智慧农场试验田块(33.140 06°N, 118.813 45°E),图 1 为试验田块俯视图。

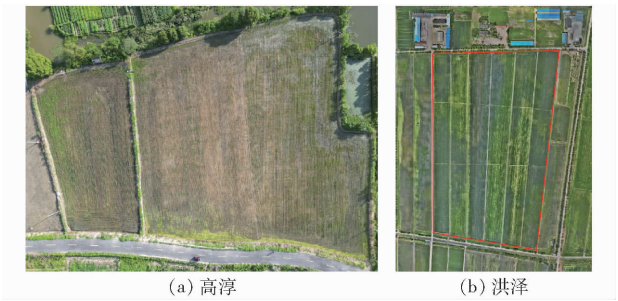


图 1 高淳试验田块和洪泽试验田块  
Fig. 1 Experimental plots in Hongze and Gaochun

试验田均按照合作社种植经验进行施肥,本文与合作社农业专家沟通后在 2023 年 6 月至 10 月期间于各施肥日前后 5 d 及当天进行图像采样。使用大疆 mini3 pro 型无人机遥感平台,在晴朗天气光强稳定的 12:00—14:00 获取各施肥时期的 RGB 正射影像和倾斜影像,飞行高度 5 m;倾斜影像拍摄调整

无人机云台为 45°。各时期正射影像与倾斜影像比例为 1:1。

高淳试验田种植品种为南粳 46,其采用侧深施肥技术在返青期末施返青肥,故分别在分蘖肥、穗肥和粒肥施用日:7 月 7 日、8 月 11 日和 8 月 29 日当天及前后 5 d,采集田间水稻 RGB 图像共 1 526 幅;洪泽试验田种植品种为南粳 5718,该合作社按照以往种植习惯未施粒肥,故分别在返青肥、分蘖肥和穗肥施用日:7 月 3 日、7 月 24 日和 8 月 10 日当天及前后 5 d,采集田间水稻 RGB 图像共 1 573 幅,图 2 为两处数据采集点的部分图像示例。

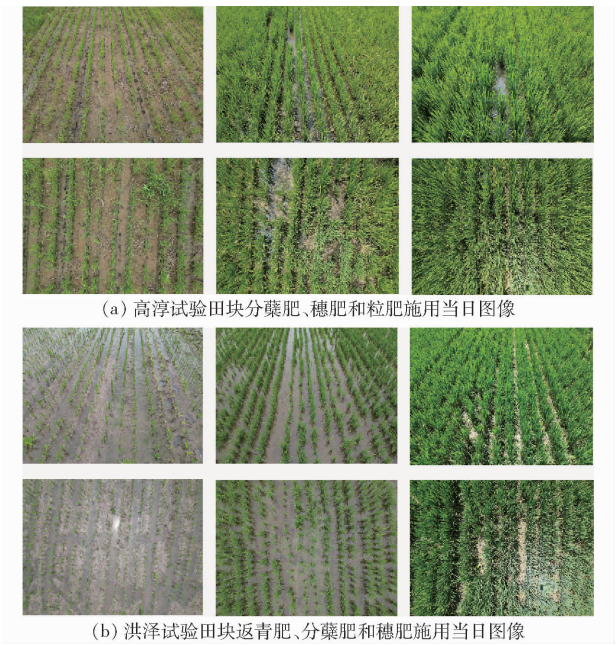


图 2 部分图像数据示例  
Fig. 2 Display of selected image data

## 2 水稻施肥期判别方法

### 2.1 方法总体框架

基于多模态知识图谱的水稻施肥期判别方法总体框架如图 3 所示。

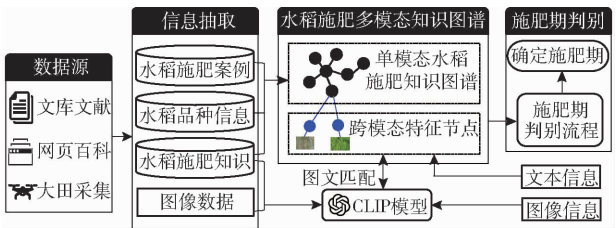


图 3 基于多模态知识图谱的水稻施肥期判别方法总体框架  
Fig. 3 Overall framework of rice fertilization period discrimination method based on multi-modal knowledge graph

本文对数据源进行信息抽取与处理后得到结构化的水稻施肥案例数据、品种信息数据、水稻施肥知识数据以及图像数据。利用前 3 种数据构建文本形式的单模态水稻施肥知识图谱,该知识图谱包含水

稻不同施肥期的相关节点。从水稻施肥知识中提取水稻不同施肥期的特征短语,利用 CLIP 模型与对应的图像数据进行图文匹配,得到相应的跨模态特征短语权重后与图像和单模态水稻施肥知识图谱中的相关节点连接,完成水稻施肥多模态知识图谱的构建。在进行施肥期判断时,可以输入文本与图像单一或者双模态信息:图像信息通过 CLIP 模型与跨模态特征短语进行匹配度计算;文本信息直接与知识图谱中的文本节点进行匹配度计算。最后通过综合匹配度的施肥期判定阈值确定该输入信息是否符合某一施肥期。

### 2.2 水稻施肥多模态知识图谱构建

多模态知识图谱旨在有效利用各模态的丰富信息来提高知识利用效率<sup>[18]</sup>,在处理多样化的农业场景时能更好地适应复杂的条件。本文综合利用文本模态的水稻品种、种植地区等信息和图像模态的水稻施肥期特征提升知识表达的完整性,以期在施肥期判别任务中具有较好的准确性。文本形式的单模态知识图谱构建技术已趋于成熟<sup>[19]</sup>,现阶段农业领域内已有构建完成的文本知识图谱<sup>[20-23]</sup>。本文提出的水稻施肥多模态知识图谱构建方法可以利用现有或构建的单模态知识图谱,结合 CLIP 模型将单模态知识图谱拓展为文本、图像两种模态形式的多模态知识图谱。

#### 2.2.1 单模态水稻施肥知识图谱构建

本文在进行单模态水稻施肥知识图谱构建时结合课题组经验<sup>[15]</sup>,分为本体构建、实体抽取、知识存储 3 个主要步骤:

(1)本体构建。本文结合国家水稻数据中心网站中的概念与施肥应用场景,采用自顶向下<sup>[24]</sup>的本体构建方式构建水稻施肥知识本体,包括水稻品种、种植地区、种植实例、施肥方案等,图 4a 为构建完成的本体。

(2)实体抽取。以中国知网、国家水稻数据中心为数据源进行实体抽取。采用与文献[15]相同的改进 RoBERTa<sub>l</sub>-wwm-CASREL 模型从获取的文本数据中得到包含水稻施肥信息的实体三元组数据。

(3)知识存储。采用 Neo4j 数据库作为存储与可视化工具,将水稻施肥三元组数据导入到数据库存储查询,图 4b 为构建完成水稻施肥知识图谱的局部示例。

#### 2.2.2 基于依存句法分析的跨模态特征短语提取

一般来说,水稻施肥时期包含基肥、返青肥、分蘖肥、穗肥和粒肥 5 个阶段<sup>[25]</sup>,基肥施用时间相对固定,仅与当地种植习惯有关,因此本文施肥期并未

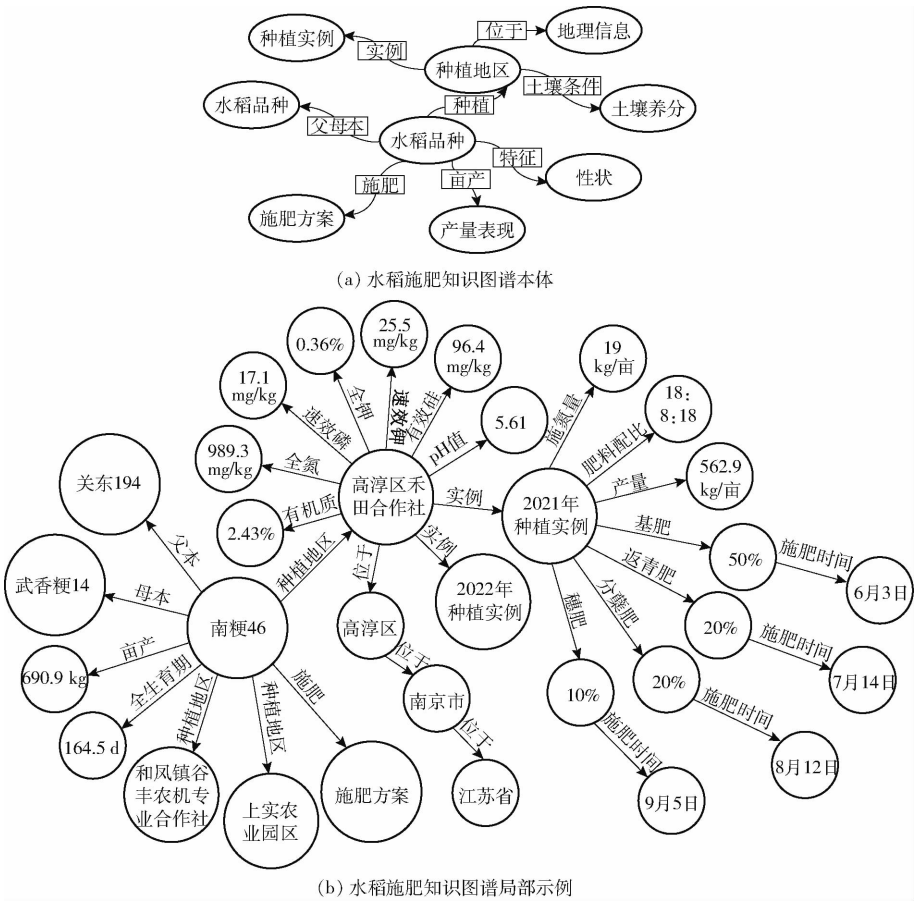


图 4 单模态水稻施肥知识图谱本体与图谱局部示例

Fig. 4 Ontology of unimodal rice fertilization knowledge graph and partial graph display

考虑基肥阶段。其他各阶段水稻生理特征与大田形态特征均有较大差别。

本文基于依存句法分析提取各施肥期的特征短语。依存句法分析是通过识别句子中词与词之间的依存关系来分析文本的句法结构<sup>[26]</sup>。该方法中每个句子只有一个根节点 (root)，其他词中处于支配

地位的为中心词,属于从属地位的为从属词。本文采用 HanLP 为依存句法分析工具,以“水稻进入返青期的标志是叶片颜色变绿。”为例,其句法结构如图 5 所示。其中谓语动词“是”为句子中的唯一独立成分作为根节点,其他词语均从属于某个词,如“标志”从属于“进入”。

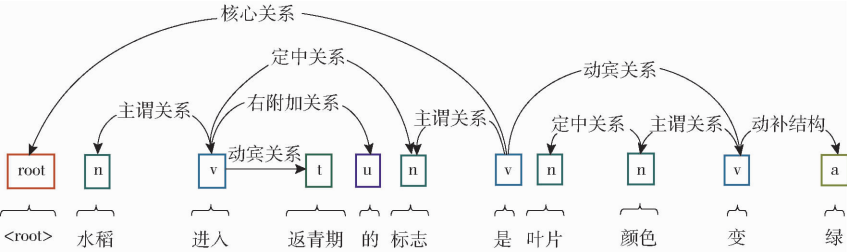


图 5 依存句法分析示意图

Fig. 5 Schematic of dependency syntax analysis

使用依存句法分析提取符合条件的词语组合。本文关注的特征短语是以修饰词为主的不同施肥期水稻特征的描述词。因此,结合数据源中文本描述的特点制订了二元依存关系及对应的词性组合规则,如表 1 所示。使用 HanLP 工具与规则表自动提取 4 个施肥期的特征短语,经人工检查后增减部分词组即得到关于水稻施肥期描述的跨模态特征短语集合。

2.2.3 基于 CLIP 模型的多模态知识图谱

多模态知识图谱构建的关键是不同模态的数据融合,于本文而言就是图像和文本的关联。CLIP 模型是 OpenAI 于 2021 年开源的用于匹配图像和文本的预训练神经网络模型,它在遥感<sup>[27]</sup>和农业领域<sup>[28]</sup>内的图像分类任务已有成功应用。它利用网络上收集到的 4 亿对图文对,分别使用一个文本编码器和一个图像编码器对文本和图像进行编码,通

表 1 二元依存关系及对应词性规则

Tab. 1 Binary dependency relations and corresponding part-of-speech rules

| 依存关系类型 | 词性组合        | 描述                  | 示例          |
|--------|-------------|---------------------|-------------|
| 形容词修饰  | n - a       | 形容词修饰名词             | 翠绿的叶片       |
| 名词主语   | n - v - n   | 名词作为动词的主语           | 籽粒呈现绿色      |
| 形容词补语  | n - v - a   | 形容词作为动词的补语          | 叶片颜色变深      |
| 副词修饰   | adv - a - n | 副词修饰形容词，<br>形容词修饰名词 | 略微卷曲的<br>叶片 |
| 数词修饰   | num - q - n | 数词与量词修饰名词           | 三片叶子        |

过对比学习的方式<sup>[29]</sup>将正相关的文本-图像对靠近,将负相关的文本-图像对推远,从而使相同语义的文本和图像映射到同一个共享的语义空间<sup>[30]</sup>,实践表明其具有强大的零样本学习能力<sup>[31]</sup>。为保留模型良好的泛化性能,本文使用未调整的中文版本的 SOTA 模型——Chinese CLIP<sup>[32]</sup>作为图文匹配工具,构建多模态知识图谱。

2.2.2 节中提取了 4 个施肥期对应的跨模态特征短语集合,其中每一个施肥期都对应大量不同的跨模态特征短语,不同的特征短语在表征某一施肥期时的精准性具有差异,因此需要确定跨模态特征短语集合中每一个元素的权重。

对于 1.2 节中的图像数据,每个施肥期使用  $m$  幅图像数据。分别用 Q、F、H、L 表示返青肥、分蘖肥、穗肥、粒肥,以返青肥为例图像编号为 imgQ\_1 到 imgQ\_  $m$ ,其特征短语共  $n$  个编号,为 phraseQ\_1 到 phraseQ\_  $n$ 。

(1) 对于每幅图像与每个描述短语,使用 Chinese CLIP 模型分别对它们编码后得到特征向量  $I_1, I_2, \dots, I_m$  和  $P_1, P_2, \dots, P_n$ 。

(2) 计算所有图文组合之间的匹配度  $M$ , 值范围为  $[0, 1]$ , 其中 1 表示完美匹配, 0 表示完全不匹配, 得到一个  $m \times n$  的匹配度矩阵<sup>[33]</sup>, 如图 6 所示。

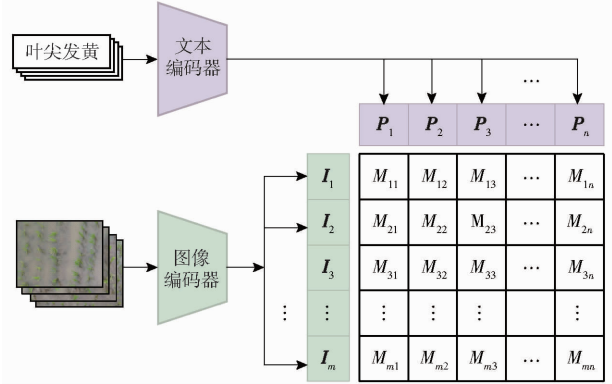


图 6 基于 CLIP 模型的匹配度矩阵示意图

Fig.6 Schematic of matching matrix based on CLIP model

(3) 对于每个特征短语, 计算它与所有图像的

平均匹配度  $A$ , 即

$$A_i = \frac{\sum_{k=1}^m M_{ki}}{m} \quad (i = 1, 2, \dots, n) \tag{1}$$

式中  $M_{ki}$ ——第  $k$  幅图像和第  $i$  个特征短语匹配度  
 $A_i$ ——第  $i$  个特征短语平均匹配度

(4) 将特征短语平均匹配度标准化到  $[0, 1]$  范围内, 得到权重  $w'$ , 即

$$w'_i = \frac{A_i - A_{\min}}{A_{\max} - A_{\min}} \tag{2}$$

式中  $A_{\min}$ ——第  $i$  个特征短语平均匹配度最小值  
 $A_{\max}$ ——第  $i$  个特征短语平均匹配度最大值  
 $w'_i$ ——第  $i$  个特征短语权重

(5) 得到 4 个施肥期对应特征短语的权重后, 考虑到某些特征短语同时存在于多个施肥期, 那么它并非某个施肥期的独特特征, 根据经验引入惩罚项  $1.2 - 0.2f$  与权重相乘, 然后对权重进行归一化, 最终确定特征短语权重  $w$ , 即

$$w_i = \frac{w'_i (1.2 - 0.2f)}{\sum_{i=1}^n w'_i} \tag{3}$$

式中  $f$ ——特征短语重复出现频次

$w_i$ ——第  $i$  个特征短语归一化权重

本文设置 4 个施肥期节点“返青肥”、“分蘖肥”、“穗肥”、“粒肥”与单模态知识图谱中的“施肥方案”节点连接。多模态知识图谱分为 A-MMKG 和 N-MMKG 两种<sup>[34]</sup>, 前者是多模态数据作为实体的属性值, 适用于属性丰富、需要细致描述的场景; 后者是将多模态数据直接作为实体, 适用于关系密集、需要揭示实体间复杂关系的场景。在本文的应用场景中, 图像之间相对独立, 并无复杂关系连接, 因此采用第 1 种构建方式更为合适。将各施肥期使用的图像作为属性值与这 4 个施肥期节点连接, 其三元组形式为 (施肥期, has\_image, 图像)。各施肥期对应的跨模态特征短语同时与图像节点及施肥期节点相连: 其与图像节点连接的关系是步骤 (2) 中得到的匹配度; 其与施肥期节点连接的关系是式 (3) 中得到的权重。它们的三元组形式分别为 (跨模态特征短语, 匹配度, 图像) 和 (施肥期, 权重值, 跨模态特征短语)。三元组中的图像采用本地 URL 形式, 将所有包含跨模态特征节点的三元组导入 Neo4j 后得到构建完成的多模态水稻施肥知识图谱。图 7 为该多模态知识图谱的局部示例。

2.3 水稻施肥期判别流程

基于构建好的水稻施肥多模态知识图谱进行水稻施肥期判别, 流程如图 8 所示。

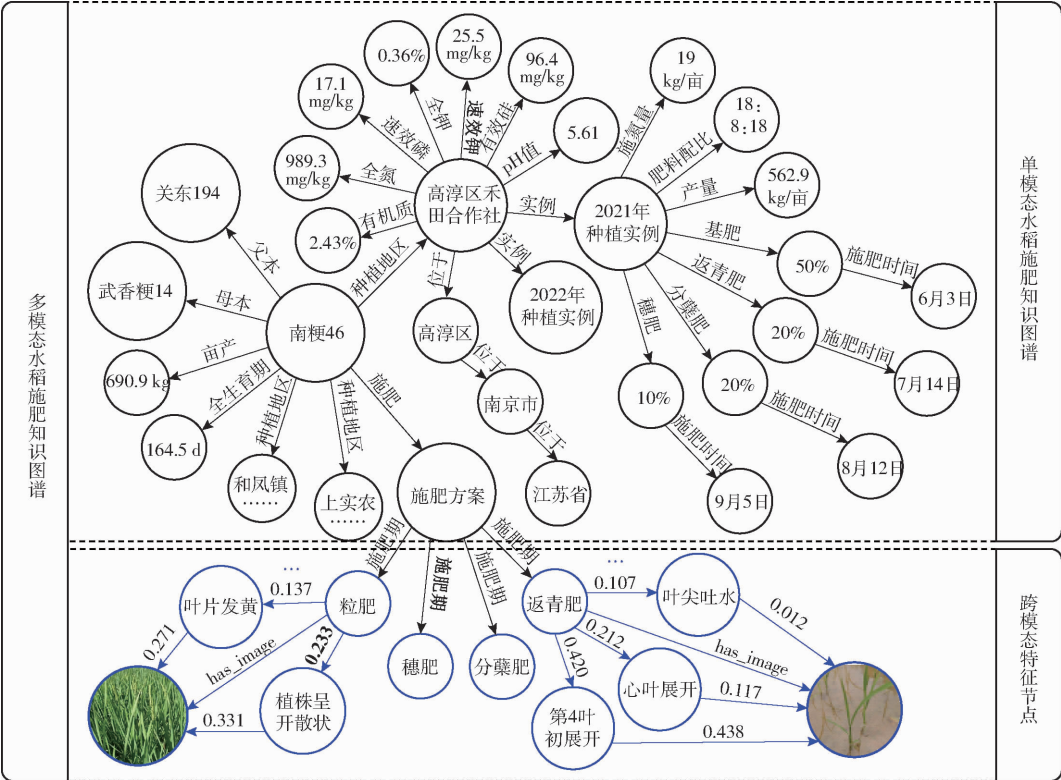


图 7 多模态水稻施肥期知识图谱局部示意图

Fig. 7 Partial schematic of multi-modal rice fertilization period knowledge graph

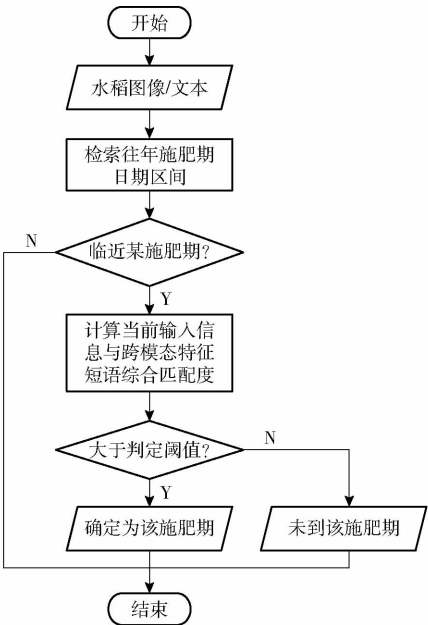


图 8 水稻施肥期判别流程图

Fig. 8 Flowchart of rice fertilization period discrimination process

(1) 当有水稻图像或文本输入需要进行施肥期判别时,根据种植地点和种植品种,首先在知识图谱中检索对应品种在对应地区的施肥案例中各施肥期的日期区间,得到各施肥期的大致施肥时间。

(2) 根据当前实际日期判断该日期是否在某一施肥期的日期区间内,若在则已经临近施肥期。

(3) 计算当前输入信息与跨模态特征短语的匹

配度。对于图像信息直接利用 CLIP 模型得到该图像与施肥期对应跨模态特征短语的匹配度,将它们分别与对应权重相乘再求和得到一个综合匹配度;对于文本信息则先基于 2.2 节进行词组提取,然后计算各词组与跨模态特征短语的余弦相似度,将与之相似度最大的跨模态特征短语的权重求和作为综合匹配度;对于混合信息则分别求出对应综合匹配度再求和。

(4) 将计算后的综合匹配度与该施肥期设定的判定阈值进行比较。若大于判定阈值则输出确定为该施肥期,可以准备施肥作业;若不大于判定阈值则输出未到该施肥期。

### 3 试验与结果分析

#### 3.1 试验设置

##### 3.1.1 试验环境

本试验硬件环境为:CPU 为 Intel Core i7-13700 k, GPU 为 NVIDIA GeForce RTX 4070 Ti,内存容量为 32 GB,操作系统为 Windows 11,运行环境为 Pytorch 2.0.0 + Cuda 11.7。

##### 3.1.2 数据集设置

将大田采集到的两地共 3 099 幅图像使用数据增强操作,通过图像几何变换:翻转、缩放、裁剪、平移中的一种或多种组合来增加数据量与提升系统鲁棒性<sup>[35-36]</sup>后,共得到 6 198 幅图像数据,数据分组

标签见表 2。

表 2 水稻施肥期图像数据集

| 标签名称 | 图像数量 | 扩充后图像数量 |
|------|------|---------|
| 返青肥前 | 200  | 400     |
| 返青肥  | 218  | 436     |
| 返青肥后 | 206  | 412     |
| 分蘖肥前 | 337  | 674     |
| 分蘖肥  | 308  | 616     |
| 分蘖肥后 | 343  | 686     |
| 穗肥前  | 312  | 624     |
| 穗肥   | 320  | 640     |
| 穗肥后  | 350  | 700     |
| 粒肥前  | 166  | 332     |
| 粒肥   | 166  | 332     |
| 粒肥后  | 173  | 346     |

在数据集中分别抽取各标签 50 幅图像总计 600 幅图像作为测试数据集,剩余 5 598 幅图像作为训练数据集。其中包含施肥当日及前后 5 d 的各施肥期:返青肥期、分蘖肥期、穗肥期、粒肥期数量图像数量分别为 1 098、1 826、1 814、860。

训练数据集从各施肥期中随机抽取 20% 为一组,共划分为 5 组,各组数据是独立同分布的,用于 5 折交叉验证;测试数据集用于评估方法的性能。数据集划分及使用示意图 9。

3.1.3 评价指标

选择准确率 (Accuracy)、召回率 (Recall) 和 F1 值作为评价指标<sup>[37]</sup>。

对于各施肥时期的数据:施肥前、后 5d 图像为负例,施肥当日图像为正例。

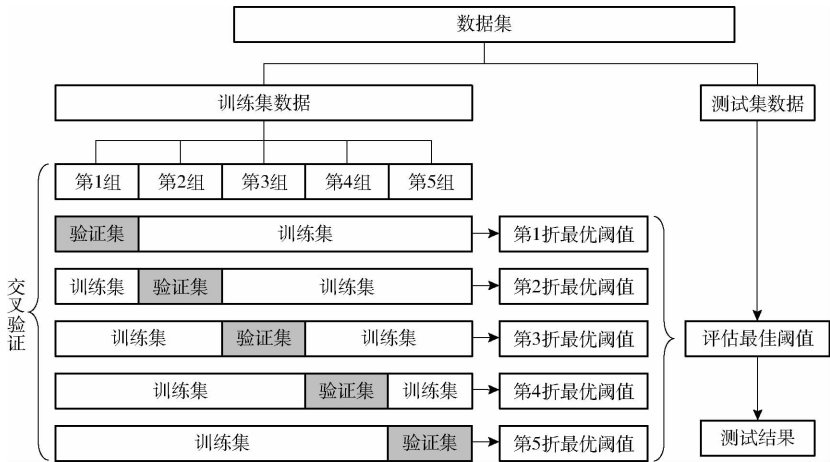


图 9 数据集设置  
Fig. 9 Dataset settings

3.1.4 多模态知识图谱

按照 2.2.1 节步骤从文本数据源中抽取与水稻施肥相关的案例与知识,以三元组形式存储在 Neo4j 中。构建完成的单模态水稻施肥知识图谱共包含 7 558 个实体节点信息、9 593 条关系信息,其中包含 360 个品种与 245 个种植试验。然后按照 2.2.2 节方法,分别为返青肥、分蘖肥、穗肥和粒肥 4 个时期提取了 37、42、35、31 个跨模态特征短语。最后按照 2.2.3 节的方法,每个施肥期优先从 153 幅文本数据源的配图中选择对应的图像,然后再以拍摄图像补齐至 50 幅,经过 3.1.2 节中相同的数据增强操作扩充至 100 幅。本文使用的 Chinese CLIP 模型视觉编码器为 VIT - H/14<sup>[38]</sup>,文本编码器为 RoBERTa - wwm - Large<sup>[39]</sup>,利用 Chinese CLIP 模型分别计算每个跨模态特征短语的归一化权重,得到跨模态三元组存入 Neo4j 数据库。

3.2 施肥期判定阈值对判别结果的影响

利用构建完成的多模态水稻施肥知识图谱可

以计算输入信息的综合匹配度。为判别输入信息是否符合某施肥期特征,需要确定各施肥期综合匹配度的判定阈值。在现有数据样本基础上,分析阈值变化对判别准确程度的影响,从而确定最佳阈值。

为保证该方法的可靠性与泛化能力,避免欠拟合与过拟合情况的发生,本文在训练数据集中采用 5 折交叉验证的方式<sup>[40]</sup>。使用每一折中的训练集得到最优阈值后再使用验证集对该阈值进行性能评估,重复 5 次后再综合考虑每一折得到的阈值对判别方法准确性和稳定性的影响,从而得到最佳判定阈值。

具体的每一折试验中,利用 Chinese CLIP 模型计算训练集中每幅图像与对应施肥期的综合匹配度,设置阈值从 0 到 1,步长为 0.01,统计各施肥期判定阈值对应的评价指标值,图 10 为第 1 折中返青肥训练集数据的结果。

在每一折中,各施肥期判定阈值寻优过程的召

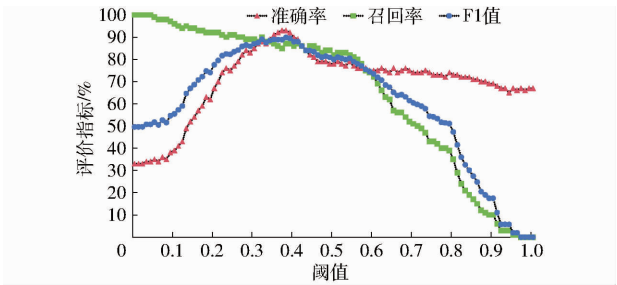


图 10 第 1 折返青肥判定阈值寻优过程  
Fig. 10 Process for determining threshold in the first fold of cross-validation for re-greening

回率和准确率的整体变化趋势相近。当阈值接近于 0 时,判别模型倾向于将大多数甚至所有图像均判断为正例,因此召回率接近或等于 100%,但此时存在较多假正例,导致准确率较低;当阈值增加时,判别模型变得更加保守,判断为正例的数量减少,从而导致召回率下降,但因为假正例数量减少,准确率逐渐提升;当阈值接近于 1 时,判别模型仅判定高度确定的图像为正例,导致召回率极低,接近于 0,同时过高的阈值可能导致判别模型错过真正的正例,影响准确率。综合考虑召回率与准确率,每一折中选取 F1 值最高的阈值作为 4 个施肥期的最优判定阈值。

每一折中的验证集可以评估该折最优阈值的性能。本文统计了每一折最优阈值在验证集中的判别准确率,同时为确定一组兼顾准确率和稳定性的最佳判定阈值,还统计了每个阈值在其他验证集上的表现。

图 11 左侧为每折各施肥期最优阈值在本折验证集及其他折验证集的准确率热力图,其中横坐标是第 1 折到第 5 折的最优阈值,对角线为每折最优阈值在该折验证集上的准确率;右侧是每折最优阈值在所有验证集上的平均准确率,误差线为准确率的标准差。最佳判定阈值应该在保证一定准确率的同时也能在大多数验证集上表现稳定。

由图 11 可以看出,返青肥判定阈值为 0.37 时平均准确率最高,当判定阈值为 0.41 时准确率下降 0.55% 但标准差仅为前者的 40%,因此返青肥的最佳判定阈值选择 0.41 更合适。分蘖肥选择高准确率和低标准差的 0.53 作为最佳判定阈值。穗肥判定阈值为 0.63 和 0.65 时平均准确率较高,前者高于后者 1%,标准差后者低于前者 7%,在准确率相差不大时,本文更倾向于稳定性好的阈值,因此选择 0.65 作为穗肥的最佳判定阈值。粒肥判定阈值选择准确率和稳定性均较好的 0.32 作为最佳判定阈值,该施肥期出现了阈值相差较大的情况,可能是该施肥期阈值寻优过程中有两个峰,分别在 0.32 和

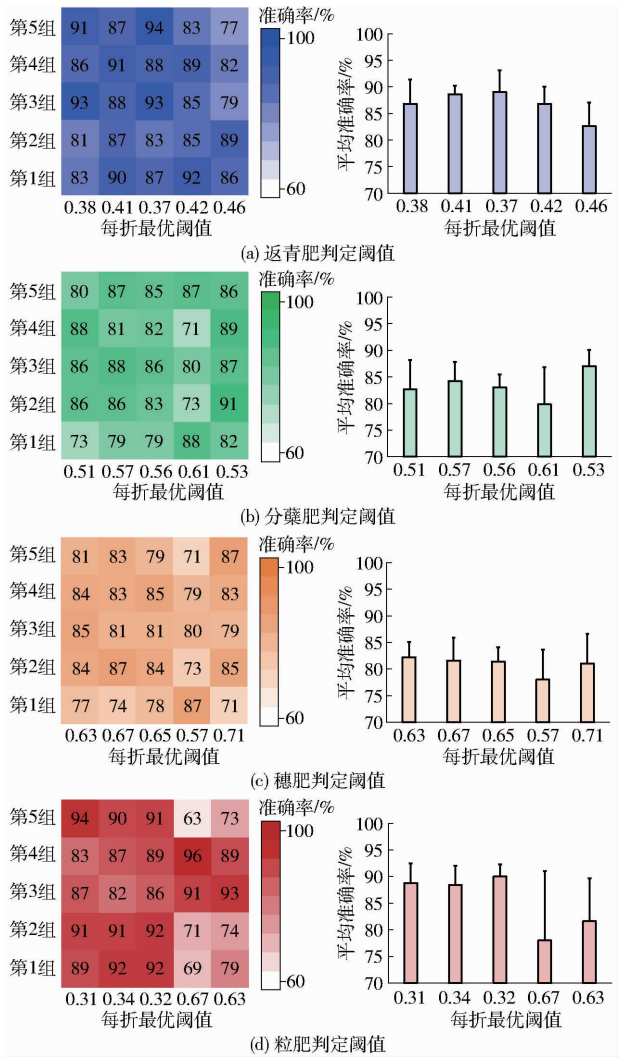


图 11 各施肥期最优阈值在验证集的判别准确率  
Fig. 11 Discrimination accuracy of optimal threshold for each fertilization period on validation set

0.65 左右,但 0.32 作为阈值具有更好的泛化性,稳定性更好。

### 3.3 水稻各施肥期判别结果测试

#### 3.3.1 测试试验

使用测试数据集进行测试。依据多模态水稻施肥知识图谱可以判断数据集的日期是临近施肥期,因为数据集是人为在各施肥期采集的,故这一过程省略,得到该数据集中的图像均临近各施肥期的结论。该数据集中的 600 幅图像分为 4 组对应 4 个施肥期,每一组都包含 50 幅施肥当日图像、50 幅施肥前 5 d 图像和 50 幅施肥后 5 d 的图像。然后对这些图像进行施肥期判别,结果如表 3 所示。4 个施肥期判别准确率为 86.2%,其中返青肥与粒肥时期的判别准确率较其他两个施肥期更高,可能因为这两个时期水稻生长形态变化较快,与施肥前后有较大差异。返青肥时期秧苗在 10 d 内生长迅速,变化较大;粒肥在抽穗时施用,抽穗前后水稻形态特征变化也很明显。

表 3 各施肥期判别结果

Tab.3 Discrimination results for each fertilization period

| 临近施肥期 | 样本数 | 判别结果 |    | 准确率/% |
|-------|-----|------|----|-------|
|       |     | 正确   | 错误 |       |
| 返青肥   | 150 | 135  | 15 | 90.0  |
| 分蘖肥   | 150 | 126  | 24 | 84.0  |
| 穗肥    | 150 | 121  | 29 | 80.7  |
| 粒肥    | 150 | 136  | 14 | 90.1  |

3.3.2 消融试验

本文的施肥期判别方法综合利用了多模态水稻施肥知识图谱中的部分文本节点的信息和跨模态特征短语节点信息。如果受知识图谱案例数量限制无法得到施肥区间,则直接计算与各个施肥期的综合匹配度后,选择匹配度最高的施肥期进行阈值判定后再确定是否为某一施肥期。因此本文设计了不使用文本节点的试验,仅使用跨模态特征短语节点进行施肥期判别。使用与 3.3.1 节相同的验证集,直接使用图像进行 4 个施肥期的综合匹配度计算,取综合匹配度最高的施肥期判定阈值进行比较,试验结果如表 4 所示。

表 4 消融试验结果

Tab.4 Ablation experiment results

| 临近施肥期 | 样本数 | 判别结果 |    | 准确率/% |
|-------|-----|------|----|-------|
|       |     | 正确   | 错误 |       |
| 返青肥   | 150 | 135  | 15 | 90.0  |
| 分蘖肥   | 150 | 115  | 35 | 76.7  |
| 穗肥    | 150 | 109  | 41 | 72.7  |
| 粒肥    | 150 | 134  | 16 | 89.3  |

从表 4 可以看出,在仅使用一种模态信息时,判别方法准确率为 82.2%。其中对于返青肥和粒肥

时期判别准确率下降幅度较小,可能因为这两个时期形态特征与其他时期差异较明显,判别方法基本可以正确判断。对于分蘖肥和穗肥会导致判别方法准确率分别下降 7.3、8.0 个百分点,可能是因为在确定权重时,模型更关注同一施肥期前后的特征变化,从而倾向于给细粒度差异更高的权重,而粗粒度权重较低。这两个水稻在施肥期前后可能存在部分细致形态特征的交叉重叠,导致判别方法混淆了这两个施肥期,在判断临近施肥期时就已经错误。

4 结论

(1)提出了一种基于多模态知识图谱的水稻施肥期判别方法。该方法利用图像和文本两个模态的信息结合 Chinese CLIP 模型构建了多模态水稻施肥知识图谱,并使用它对水稻施肥期进行判别。结果表明,该方法总体准确率达到 86.2%,在返青肥和粒肥时期判别准确率更好,达到 90.0% 和 90.1%;消融试验结果表明,本方法相较于使用单一图像信息时准确率提升 4.0 个百分点,其中分蘖肥和穗肥时期分别提升 7.3、8.0 个百分点。

(2)基于农业应用场景的复杂性,综合使用多种模态数据有助于提高信息利用率。本文将大语言模型引入多模态知识图谱的构建与应用,利用大模型对通识知识的理解能力和知识图谱对专业领域的概念表达优势,是大模型跨领域应用的探索。试验结果表明该方法在农业场景中有实际应用的可能。使用现有大模型作为多类型数据的跨模态映射工具,在保证良好性能的基础上避免了耗时耗力的训练与数据标注过程。

参 考 文 献

[1] 严如玉,甘国渝,赵希梅,等.我国水稻优势产区生产格局及施肥现状研究[J].中国稻米,2023,29(3):1-8.  
YAN Ruyu, GAN Guoyu, ZHAO Ximei, et al. Study on the production pattern and fertilization status of rice dominant production areas in China[J]. China Rice,2023,29(3):1-8. (in Chinese)

[2] 梁玉刚,李静怡,周晶,等.中国水稻栽培技术的演变与展望[J].作物研究,2022,36(2):180-188.  
LIANG Yugang, LI Jingyi, ZHOU Jing, et al. Evolution and prospect of rice cultivation technology in China[J]. Crop Research,2022,36(2):180-188. (in Chinese)

[3] 赵春江.智慧农业的发展现状与未来展望[J].华南农业大学学报,2021,42(6):1-7.  
ZHAO Chunjiang. Current situations and prospects of smart agriculture[J]. Journal of South China Agricultural University, 2021,42(6):1-7. (in Chinese)

[4] 李道亮,李震.无人农场系统分析与发展展望[J].农业机械学报,2020,51(7):1-12.  
LI Daoliang, LI Zhen. System analysis and development prospect of unmanned farming[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(7):1-12. (in Chinese)

[5] 田仓,虞轶俊,吴龙龙,等.不同灌溉和施肥模式对稻田磷形态转化和有效性的影响[J].农业工程学报,2021,37(24):112-122.  
TIAN Cang, YU Yijun, WU Longlong, et al. Effects of various irrigation and fertilization schedules on the transformation and availability of phosphorus in paddy fields[J]. Transactions of the CSAE,2021,37(24):112-122. (in Chinese)

[6] 郭全全.不同时期施肥对水稻生长及产量的影响探析[J].南方农业,2021,15(24):49-50.

- GUO Quanquan. Analysis of the effect of fertilization at different stages on the growth and yield of rice[J]. South China Agriculture, 2021, 15(24): 49–50. (in Chinese)
- [7] 杨振忠, 方圣辉, 彭漪, 等. 基于机器学习结合植被指数阈值的水稻关键生育期识别[J]. 中国农业大学学报, 2020, 25(1): 76–85.
- YANG Zhenzhong, FANG Shenghui, PENG Yi, et al. Recognition of the rice growth stage by machine learning combined with vegetation index threshold[J]. Journal of China Agricultural University, 2020, 25(1): 76–85. (in Chinese)
- [8] IKASARI I H, AYUMI V, FANANY M I, et al. Multiple regularizations deep learning for paddy growth stages classification from LANDSAT-8[C]//2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS). IEEE, 2016: 512–517.
- [9] 徐建鹏, 王杰, 徐祥, 等. 基于 RAdam 卷积神经网络的水稻生育期图像识别[J]. 农业工程学报, 2021, 37(8): 143–150.
- XU Jianpeng, WANG Jie, XU Xiang, et al. Image recognition for different developmental stages of rice by RAdam deep convolutional neural networks[J]. Transactions of the CSAE, 2021, 37(8): 143–150. (in Chinese)
- [10] 白晓东. 基于图像的水稻关键发育期自动观测技术研究[D]. 武汉: 华中科技大学, 2014.
- BAI Xiaodong. Image-based automatic observation technology for critical growth stages of rice[D]. Wuhan: Huazhong University of Science & Technology, 2014. (in Chinese)
- [11] ZOU Yidong, PAN Shouhui, YANG Feng, et al. Precise recommendation method of suitable planting areas of maize varieties based on knowledge graph[J]. Agriculture, 2023, 13(3): 526.
- [12] 郑泳智, 吴惠萍, 朱定局, 等. 基于荔枝和龙眼病虫害知识图谱的问答系统[J]. 计算机与数字工程, 2021, 49(12): 2618–2622.
- ZHENG Yongzhi, WU Huilin, ZHU Dingju, et al. Question and answer system based on the knowledge graphs of litchi and longan diseases and insect pests[J]. Computer & Digital Engineering, 2021, 49(12): 2618–2622. (in Chinese)
- [13] 徐畅, 张领先, 乔岩. 基于电子病历的作物病虫害关联挖掘及智能诊断[J]. 智能化农业装备学报(中英文), 2023, 4(4): 1–10.
- XU Chang, ZHANG Lingxian, QIAO Yan. Association mining of crop diseases and intelligent diagnosis based on electronic medical records[J]. Journal of Intelligent Agricultural Mechanization, 2023, 4(4): 1–10. (in Chinese)
- [14] 张博凯, 李想. 基于知识图谱的 Android 端农技智能问答系统研究[J]. 农业机械学报, 2021, 52(增刊): 164–171.
- ZHANG Bokai, LI Xiang. Design of agricultural question answering system based on knowledge graph[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 164–171. (in Chinese)
- [15] 周俊, 郑彭元, 袁立存, 等. 基于改进 CASREL 的水稻施肥知识图谱信息抽取研究[J]. 农业机械学报, 2022, 53(11): 314–322.
- ZHOU Jun, ZHENG Pengyuan, YUAN Licun, et al. Knowledge graph information extraction for rice fertilization based on improved CASREL[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 314–322. (in Chinese)
- [16] 许多, 鲁旺平, 许瑞清, 等. 基于农业时空多模态知识图谱的水稻精准施肥决策方法[J]. 华中农业大学学报, 2023, 42(3): 281–292.
- XU Duo, LU Wangping, XU Ruiqing, et al. A method of deciding precision fertilization of rice based on spatiotemporal multi-modal knowledge graph of agriculture[J]. Journal of Huazhong Agricultural University, 2023, 42(3): 281–292. (in Chinese)
- [17] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//International Conference on Machine Learning. PMLR, 2021: 8748–8763.
- [18] WANG M, WANG H, QI G, et al. Richpedia: a large-scale, comprehensive multi-modal knowledge graph[J]. Big Data Research, 2020, 22: 100159.
- [19] 肖桂阳, 王立松, 江国华. 一种文本-图像增强的多模态知识图谱嵌入方法[J]. 计算机科学, 2023, 50(8): 163–169.
- XIAO Guiyang, WANG Lisong, JIANG Guohua. Multimodal knowledge graph embedding with text–image enhancement[J]. Computer Science, 2023, 50(8): 163–169. (in Chinese)
- [20] 赵明, 杜亚茹, 杜会芳, 等. 植物领域知识图谱构建中本体非分类关系提取方法[J]. 农业机械学报, 2016, 47(9): 278–284.
- ZHAO Ming, DU Yaru, DU Huifang, et al. Research on ontology non-taxonomic relations extraction in plant domain knowledge graph construction[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 278–284. (in Chinese)
- [21] 张海瑜, 陈庆龙, 张斯静, 等. 基于语义知识图谱的农业知识智能检索方法[J]. 农业机械学报, 2021, 52(增刊): 156–163.
- ZHANG Haiyu, CHEN Qinglong, ZHANG Sijing, et al. Intelligent retrieval method of agricultural knowledge based on semantic knowledge graph[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 156–163. (in Chinese)
- [22] 于合龙, 沈金梦, 毕春光, 等. 基于知识图谱的水稻病虫害智能诊断系统[J]. 华南农业大学学报, 2021, 42(5): 105–116.
- YU Helong, SHEN Jinmeng, BI Chunguang, et al. Intelligent diagnostic system for rice diseases and pests based on knowledge graph[J]. Journal of South China Agricultural University, 2021, 42(5): 105–116. (in Chinese)

[23] 侯琛,牛培宇. 农业知识图谱技术研究现状与展望[J]. 农业机械学报,2024,55(6):1-17.  
HOU Chen, NIU Peiyu. Review of research status and prospects of agricultural knowledge graphs[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(6):1-17. (in Chinese)

[24] 陈宝发,任妮. 面向农业学者领域的本体构建及可视化研究[J]. 江苏农业科学,2023,51(18):191-200.

[25] 郑毅,陆进. 水稻不同种植、施肥方式对产量及施肥成本的影响[J]. 农业装备技术,2022,48(1):28-29,32.  
ZHENG Yi, LU Jin. The impact of different planting and fertilization methods on rice yield and fertilization cost[J]. Agricultural Equipment & Technology,2022,48(1):28-29,32. (in Chinese)

[26] 王新,徐照. 基于依存句法分析的建筑设计规范条文自动结构化方法[J]. 土木工程与管理学报,2023,40(4):133-140.  
WANG Xin, XU Zhao. Automatic structural method of architectural design rules based on dependency parsing[J]. Journal of Civil Engineering and Management,2023,40(4):133-140. (in Chinese)

[27] LI Xiang, WEN Congcong, HU Yuan, et al. RS-CLIP: zero shot remote sensing scene classification via contrastive vision-language supervision[J]. International Journal of Applied Earth Observation and Geoinformation, 2023,124:103497.

[28] CAO Yiyi, CHEN Lei, YUAN Yuan, et al. Cucumber disease recognition with small samples using image-text-label-based multi-modal language model[J]. Computers and Electronics in Agriculture, 2023,211:107993.

[29] 李田芳,普园媛,赵征鹏,等. 基于 CLIP 和双空间自适应归一化的图像翻译[J]. 计算机工程,2024,50(5):229-240.  
LI Tianfang, PU Yuanyuan, ZHAO Zhengpeng, et al. Image-to-image translation based on CLIP and dual-spatially adaptive normalization[J]. Computer Engineering, 2024,50(5):229-240. (in Chinese)

[30] 韩江洪,袁稼轩,卫星,等. 基于深度学习的井下巷道行人视觉定位算法[J]. 计算机应用,2019,39(3):688-694.  
HAN Jianghong, YUAN Jiaxuan, WEI Xing, et al. Pedestrian visual positioning algorithm for underground roadway based on deep learning[J]. Journal of Computer Applications,2019,39(3):688-694. (in Chinese)

[31] YUAN X, LIN Z, KUEN J, et al. Multimodal contrastive training for visual representation learning[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2021:19-25.

[32] YANG A, PAN J, LIN J, et al. Chinese CLIP: contrastive vision-language pretraining in Chinese[J]. arXiv:2211.01335, 2022.

[33] OpenAI. CLIP:connecting text and images[EB/OL]. [2022-06-30]. <https://openai.com/blog/clip/>.

[34] ZHU Xiangru, LI Zhixu, WANG Xiaodan, et al. Multi-modal knowledge graph construction and application: a survey[J]. IEEE Transactions on Knowledge and Data Engineering,2022,36(2):715-735.

[35] 岑科廷,沈华伟,曹漪,等. 图对比学习综述[J]. 中文信息学报,2023,37(5):1-21.  
CEN Keting, SHEN Huawei, CAO Qi, et al. A survey on graph contrastive learning[J]. Journal of Chinese Information Processing,2023,37(5):1-21. (in Chinese)

[36] 王东方,汪军. 基于迁移学习和残差网络的农作物病害分类[J]. 农业工程学报,2021,37(4):199-207.  
WANG Dongfang, WANG Jun. Crop disease classification with transfer learning and residual networks[J]. Transactions of the CSAE,2021,37(4):199-207. (in Chinese)

[37] 奚雪峰,周国栋. 面向自然语言处理的深度学习研究[J]. 自动化学报,2016,42(10):1445-1465.  
XI Xuefeng, ZHOU Guodong. A survey on deep learning for natural language processing[J]. Acta Automatica Sinica,2016,42(10):1445-1465. (in Chinese)

[38] ARNAB A, DEHGhani M, HEIGOLD G, et al. ViViT: a video vision transformer[C]//Proceedings of IEEE/CVF International Conference on Computer Vision. Piscataway, NJ:IEEE Press, 2021:6816-6826.

[39] LIU Y, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[J]. arXiv: 1907.11692, 2019.

[40] 沈维政,张哲,戴百生,等. 基于 Cow-DETR 与深度图像的非接触式奶牛体质量评估[J]. 农业机械学报,2023,54(8):277-285,319.  
SHEN Weizheng, ZHANG Zhe, DAI Baisheng, et al. Non-contact predicting method of dairy cow weight based on Cow-DETR and deep image[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(8):277-285,319. (in Chinese)