

作物病害智能诊断与处方推荐技术研究进展

张领先^{1,2} 韩梦瑶¹ 丁俊琦¹ 李凯雨¹

(1. 中国农业大学信息与电气工程学院, 北京 100083;
2. 农业农村部农业信息化标准化重点实验室, 北京 100083)

摘要: 由“植物诊所”形成的电子病历为作物病害处方推荐提供了新的思路。如何高效地挖掘电子病历数据并辅助作物病害处方推荐, 目前还是亟待解决的研究热点问题。在总结和整理现有国内外研究文献的基础上, 对基于显微图像的作物病害病菌孢子识别、基于光谱的作物病害诊断、基于电子病历的作物病害处方推荐等作物病害诊断与处方推荐关键技术进行了系统分析与讨论。综述结果表明, 围绕作物病害病菌侵染过程, 以智能化处方推荐需求为导向, 开展基于电子病历数据挖掘的作物病害处方推荐研究, 将成为一个研究重点。针对作物病害处方推荐过程中, 存在由于作物病害致病机理复杂、作物品种及病害种类多、病害病症动态变化且特征多等特点和难点, 研究基于电子病历数据挖掘的作物病害致病机理理解析、诊断推理、智能化处方推荐及其应用策略, 将是研究的重大方向; 探索基于知识图谱分析、大数据挖掘和机器学习算法推理等关键技术的作物病害电子病历数据挖掘分析研究, 从区域宏观视角可视觉解析作物病害致病机理及其与特征间的关联关系, 面向实际应用场景实现基于诊断推理的单一作物病害处方推荐、基于语义匹配的多种作物多种病害处方推荐, 具有更大的实际意义。

关键词: 作物病害; 病菌孢子识别; 病害检测; 病害诊断; 处方推荐; 电子病历

中图分类号: S431.9; S-1 文献标识码: A 文章编号: 1000-1298(2023)06-0001-18 OSID: 

Research Progress in Intelligent Diagnosis and Prescription Recommendation of Crop Diseases

ZHANG Lingxian^{1,2} HAN Mengyao¹ DING Junqi¹ LI Kaiyu¹

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. Key Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs, Beijing 100083, China)

Abstract: The plant electronic medical records formed by the “plant clinic” provide new ideas for the prescription recommendation of crop diseases. How to efficiently mine electronic medical record data and assist crop disease prescription recommendation is still a hot research issue, and needs to be solved urgently at home and abroad. On the basis of summarizing and sorting out the existing domestic and foreign research literature, the key technologies of crop disease diagnosis and prescription recommendation, such as spores recognition based on microscopic image, crop disease diagnosis based on spectrum, crop disease prescription recommendation based on electronic medical records, were systematically analyzed and discussed. The results showed that centering on the infection process of crop disease pathogens, the research on crop disease prescription recommendation based on electronic medical record data mining would become a research focus, guided by intelligent prescription recommendation demand. In the process of crop disease prescription recommendation, due to the characteristics and difficulties of crop disease pathogenesis complex, crop varieties and disease types, disease dynamic changes and characteristics, it would be an important direction to research on the analysis of crop disease pathogenesis, diagnostic reasoning, intelligent prescription recommendation and its application strategy based on electronic medical record data mining. It was of greater practical significance to explore the data mining analysis and research of crop disease electronic medical record based on key technologies such as knowledge graph analysis, big data mining and machine learning algorithm reasoning, and visually analyze the pathogenic mechanism of crop disease, and the correlation between characteristics from the

收稿日期: 2022-12-22 修回日期: 2023-03-01
基金项目: 国家自然科学基金项目(62176261)和全国农业专业学位研究生教育指导委员会2021年研究生教育研究重点项目(2021-NYZD-07)
作者简介: 张领先(1970—),男,教授,博士生导师,主要从事农业信息技术研究,E-mail: zhanglx@cau.edu.cn

regional macro perspective in order to realize single crop disease prescription recommendation based on diagnostic reasoning, and multiple crop disease prescription recommendation based on semantic matching for practical application scenarios.

Key words: crop disease; pathogen spore recognition; disease detection; disease diagnosis; prescription recommendation; plant electronic medical records

0 引言

作物病害是制约农业可持续发展的主要因素之一。在种植过程中,作物会受到其他生物的伤害或不适宜环境的影响而引发病害,造成作物品质下降和产量减损,进而影响生产者的效益。研究和掌握不同品种作物多种病害发生的规律和特点,及时帮助生产者对病害及时诊断、对症防治、科学用药和辅助决策,在未来农业生产中具有重要意义^[1]。

传统的作物病害防治方案主要依赖于人工经验,基本处于定性阶段,受人为主观性判断影响较大。随着计算机技术发展,专家系统实现了自动推荐作物病害防治方案。主要是利用计算机技术和人工智能技术,根据作物病害领域的专家知识和经验,进行推理和判断,模拟人类专家的决策过程,能够根据受害作物的症状等信息逐步推断,最终得到包含诊断结果以及农药的防治方案,即病害处方。但是专家系统存在不足:①系统构建成本较高,需要收集并整理作物病害领域的专家知识和经验,并据此编写推理程序。由于受害作物属性包括作物种类、发育阶段、受害部位等,编写详细的推理程序费时费力。②普适性不足,大多数专家系统只能对个别种类的作物进行推理和判断,对于不同种类的作物往往需要多个系统,在实际应用中受到限制^[2]。

由“植物诊所”形成的电子病历(Plant electronic medical records, PEMRs)为作物病害处方推荐提供了新的思路^[3]。现有的作物处方数据包括作物、环境和病害信息以及诊断知识,为作物病害诊断提供了新的分析视角:通过已有的处方数据挖掘出有效信息,辅助植物医生开具作物病害处方,缓解当前作物病害处方的困境。在生物医学研究领域,多项研究证明电子病历数据具有回溯性和可预测性,以及辅助构建临床决策支持系统的能力^[4-5]。基于此,本文对作物病害诊断与处方推荐技术国内外的研究进展进行综述,分析作物病害诊断与处方推荐研究中面临的关键问题,并对作物病害诊断与处方推荐技术的未来发展加以展望。

1 作物病害诊断与处方推荐技术概述

1.1 作物病害三角关系原理

作物病害的产生原因可以由植物病理学中的病

害三角原理解释为环境、病原物和作物三者相互作用^[6]。病害大多数是由真菌、病毒、细菌等病原物引起的,加之合适的土壤环境、气候环境和栽培条件等。病原体的毒力、宿主的遗传易感性和有利于感染的非生物环境决定了作物病害的表现形式^[7]。对于侵染性病害,当条件有利于病原物生长时,病原物就会侵染寄主植物。病原物侵入寄主植物到表现病症的连续过程称为病程,具体分为接触期、侵入期、潜育期和发病期4个时期。病菌孢子发育过程能够表示病原物侵染过程,通过作物病菌孢子侵染特征识别与行为分析,能够为作物病害早期预警和防控提供理论支撑。

1.2 作物病害诊断方法

作物在遭受病害侵袭时,外部形态特征和内部生理特征均会发生细微的变化。外部表现出诸如退绿、变色、变形、卷曲、枯萎等特征,而作物内部的水分、色素含量、光合作用、呼吸作用、防御酶系统等也会发生多种生理变化^[8]。通过检测病害发生后作物的外部形态特征和内部生理特征变化,可以获取作物的染病情况。传统的病害症状观察法,结合病原菌的形态特征以及过往经验进行识别,这种方法主观性强,且对专家的依赖性较大;20世纪70年代兴起的酶联免疫法,可以灵敏地检测作物中病毒蛋白的含量,但价格昂贵,在细菌和真菌病害检测方面应用较少。随着信息技术的快速发展以及各种仪器设备的不断出现,多种传感器应用于作物病害的识别诊断中。

从病原物侵染过程和病害诊断数据获取的角度可以将作物病害诊断方法归纳为:基于显微图像的作物病害病菌孢子识别和基于光谱成像的作物病害诊断,前者主要是病原物侵染过程接触期、侵入期和潜育期前3个阶段对病菌孢子的个体和群体特征识别及其定量表达,后者是发病期对作物内外部表现的病症进行识别、定量表达与诊断。

1.2.1 基于显微图像的作物病害病菌孢子识别

借助显微设备获取显微图像,实现作物病害病菌孢子的识别。可以搭建病菌孢子显微图像采集平台,平台一般由体视显微镜、光源、CCD彩色相机和计算机组成(图1)。平台能够实时采集病菌孢子侵染过程图像,并通过数据转换传到计算机中,通过计算机来保存孢子图像并用于实时查验,进一步通过

软件系统进行病菌孢子形态特征识别和动态特征定量表征分析。

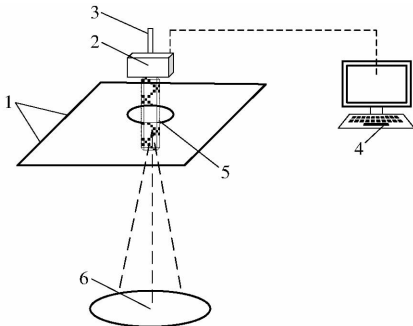


图 1 病菌孢子显微图像采集平台

Fig. 1 Pathogen spore microscopic image acquisition platform

1. 光源 2. CCD 相机 3. 相机固定杆 4. 计算机 5. 显微镜 6. 载物台

1.2.2 基于光谱成像的作物病害诊断

作物在遭受病菌侵袭后,作物色素、水分等内部物质的浓度或分布发生了改变,表现出不同的病斑^[9]。研究表明,作物内部特性改变后,对于光谱的反射特性亦会随之改变,从而为作物病害的光谱特性定量分析提供了理论基础^[10],如多光谱和高光谱传感器、热成像或叶绿素荧光成像能够检测到内部生理变化,已被应用于病害的早期检测和定量识别中,RGB 传感器能够根据病斑图像的颜色、形状和纹理等特征,结合机器视觉方法进行病害识别和定量诊断,基于光谱成像的作物病害诊断基本步骤如图 2 所示。

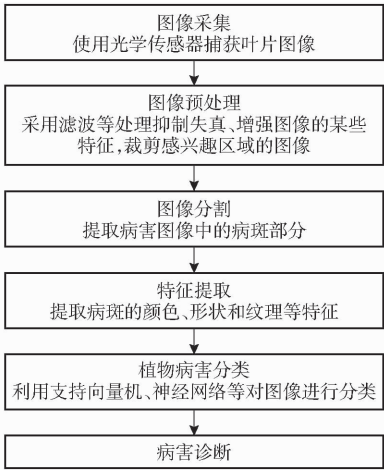


图 2 光谱成像检测植物病害流程图

Fig. 2 Flowchart for detection of plant diseases using an imaging technique

机器视觉技术是在数字图像处理、人工智能、模式识别等技术基础上逐渐发展形成的一种新的技术,为作物病害识别与诊断提供一种快速且有效的方法。可以利用数字图像处理技术,分析叶片的这

些症状来诊断作物病害并进一步估算病害发生的严重程度。基于机器视觉的设施蔬菜诊断系统流程如图 3 所示。

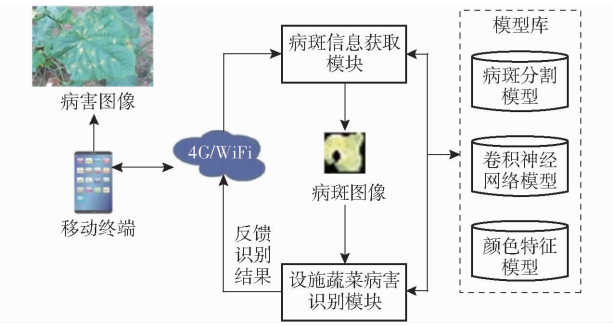


图 3 系统流程图

Fig. 3 System flowchart

1.3 作物病害处方推荐技术原理

1.3.1 “植物诊所”形成的电子病历

针对生产中面临的病害识别诊断预警相对滞后,绿色植保技术落地难,公共植保服务难以全覆盖等问题,北京市植物保护站联合中国农业大学等 4 家单位,开展了基于生产实际需求的绿色智慧关键植保技术研究及应用。2012 年北京市首次引入国际先进的植物诊所理念,开始在全市范围内建立市区乡(村)三级植物健康服务体系。先后建立植物诊所 115 个,区级二级植物医院 4 个,北京市植物总医院 1 个,植物医生及培训师 665 名,服务范围覆盖全市 13 个区,161 个乡镇,1 744 个村,还辐射到河北省廊坊市、张家口市、邢台市以及天津市武清区等地区^[11]。

植物医生遵循有害生物综合防治(Integrated pest management, IPM)原则,以开处方的形式,为农民提供病害诊断和防治技术咨询^[12],问诊完成后的电子病历都被备份在系统中(图 4),具体包括农户、植物医院、作物、病害性状、诊断结果、处方等信息^[13]。

1.3.2 处方数据预处理与扩充

处方数据预处理是将原始数据转换为可理解的格式的过程,这也是数据挖掘的重要一步。处方数据预处理的一般步骤是:对源数据文件整理、转换,数据清洗(删除重复值、缺失值处理、一致化处理和异常值处理),数据统计,最后对输出数据进行编码(标签编码和 One-hot 编码)(图 5)。

处方数据扩充是在原有数据的基础上进行修改,最终获得相似但不相同的数据的方法,被广泛应用于机器学习中^[14-15]。对于作物病害处方数据,可以使用简单数据增强(Easy data augmentation, EDA)^[16],包括以下 4 种数据扩充方法:

(1)同义词替换(Synonyms replace, SR):设句

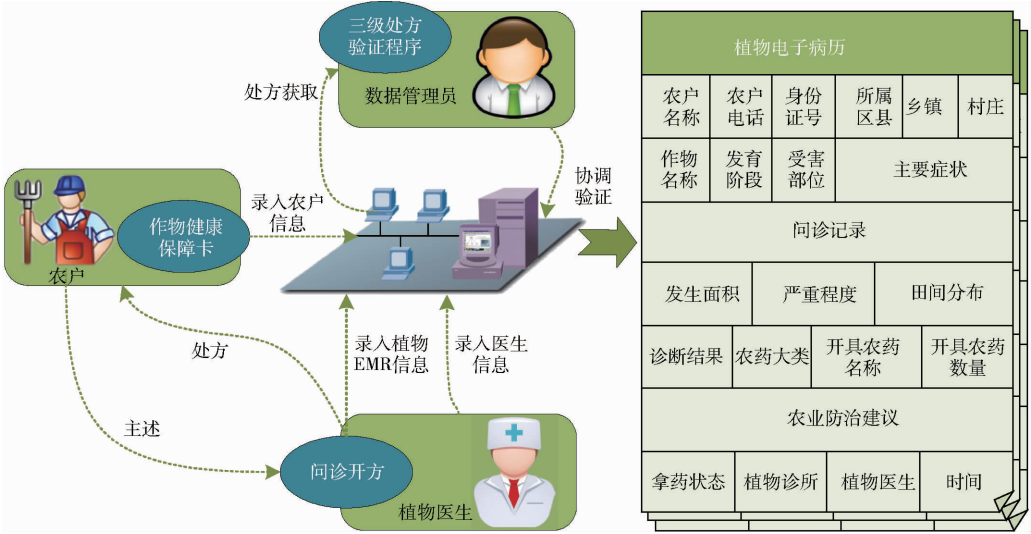


图 4 植物诊所病历填写流程图^[13]

Fig. 4 Plant clinic medical record filling process

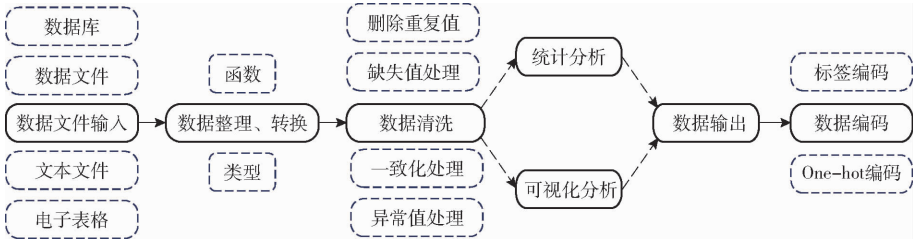


图 5 处方数据预处理流程图

Fig. 5 Prescription data preprocessing process

长为 l 个单词,替换比例为 α 。不考虑停用词,在句子中随机选择 $l \times \alpha$ 个词,然后在同义词词典中找到对应的同义词,最后随机选择同义词将原本的词汇替换。

(2)随机插入(Randomly insert, RI):将随机抽取的某个单词的同义词插入到句子中任意位置,重复 $l \times \alpha$ 次。

(3)随机交换(Randomly swap, RS):将句子中 $l \times \alpha$ 个单词位置互换。

(4)随机删除(Randomly delete, RD):剔除句子中 $l \times \alpha$ 个单词。

此外,变分自动编码器(Variational autoencoder, VAE)等文本生成模型^[17]也可用于处方数据扩充,以学习文本中的潜在性解释,生成具有特定语义的文本。

1.3.3 电子病历挖掘与处方推荐

作为最重要的临床数据类型,电子病历以结构化和非结构化结合的形式记录了大量关于疾病症状、统计数据、诊疗决策、药物处方以及环境特征的信息,能够提供完整准确的诊疗数据以及具备构建临床辅助决策支持系统的能力^[18-19]。国内外相关研究表明,对电子病历数据进行相关分析具有一定的合理性和必要性,从而可以进一步揭示特征与病

害间的深层联系^[20]。

通过处方数据分析可以获取处方数据中有价值的信息,辅助人们开展处方推荐相关研究,实现智能化诊疗。有关处方数据挖掘的研究主要有病害诊断、数据检索与管理,以及智能化处方推荐3个角度。其中处方推荐是解决信息超载问题的有效工具^[21],即通过对历史数据进行分析,发现处方数据中的规律,从而预测问诊对象可能需要的处方。

处方推荐的思路比较如表1所示。

2 作物病害诊断与处方推荐关键技术

2.1 基于显微图像的作物病害病菌孢子识别关键技术

基于显微图像的作物病害病菌孢子识别涉及的关键技术包括作物病菌孢子个体目标识别技术、作物病害病菌孢子群体目标识别技术和作物病害病菌侵染行为分析技术。

2.1.1 作物病菌孢子个体目标识别技术

作物病害病菌多以有序的状态进行生长繁殖,不同时期孢子形态特点明显,但也可能受外界因素的影响发生形态变异和部分残缺^[22]。通过检测病菌孢子,提取病菌孢子动态特征是病害早期诊断的

表 1 不同处方推荐思路特点

Tab.1 Comparison of different prescription recommendation ideas

处方推荐思路	思路描述	优势和局限
基于诊断推理	通过挖掘电子处方中的数据规律,提取处方数据中的症状-病害关联信息,通过智能病害诊断配合既定的用药知识推理实现处方推荐	对不同种病害的区分性更强,难以区分同种病害的不同用药方式
基于知识关联	基于病害-处方知识图谱,通过构建实体,分析各个实体间的关联进行药剂处方推荐	包含大量实体及关系,可扩展性强、支持智能应用,但构建成本高,需要足够的处方推荐相关知识
基于相似检索	通过处方匹配,度量处方症状距离,在检索库中进行列表检索和推荐诊疗建议	训练所需数据量较小,推荐结果更具多样性,可扩展性更强,但推荐的准确性受检索库规模和质量制约

重要环节。传统的显微镜观察主要依赖于人眼观察识别,效率低下,耗时费力,且要求专业人员持续观察。随着计算机技术、图像处理技术、模式识别技术的发展,将机器视觉技术引入到病菌孢子的识别中,提高了病菌孢子检测效率。基于机器视觉技术的病菌孢子识别算法主要通过对病菌孢子图像进行图像分割、特征提取与构建分类器模型完成对病菌孢子的识别^[23]。其中在病菌孢子图像分割中,学者常用基于阈值、边缘检测^[24]、区域生长和聚类分析等图像分割方法,获得病菌孢子图像,进而提取病菌孢子的周长、面积、圆形度、半径和弧长等形态特征^[25],纹理特征,HOG 特征,SIFT 特征^[26-27],Haar 算子,Harris 角点等特征,并结合决策树、支持向量机(SVM)、基于规则的粗糙集、LDA (Latent dirichlet allocation)主题模型、K-means、贝叶斯分类以及人工神经网络等机器学习方法进行病菌孢子的识别,均取得了良好的识别效果^[28-29]。但是,随着数据量的剧增,上述方法在特征提取方面存在计算复杂和特征不可迁移性等不足,并且需要人为提取特征和普适性不强等问题。近年来,深度学习中的卷积神经网络的出现已经彻底改变了图像识别在相关领域的应用,陆续实现了病菌孢子识别,识别率也有很大提高^[30-31]。Li 等^[32]提出使用多头注意力优化 YOLO v5 检测黄瓜灰霉病菌孢子,对模糊、多形态的孢子有较好的检测效果。但是实际采集的显微图像中也存在复杂噪声,且病原目标物比较小等系列问题给实际应用带来巨大挑战。而且病菌孢子是一种生物,本身发育过程中存在形态变异,且由于外界因子的影响,也会发生形态变异和部分残缺,因此有必要结合病菌孢子发育过程,对不同侵染期状态的病菌孢子展开深入研究,为病害早期预警提供理论支撑。

2.1.2 作物病害病菌孢子群体目标识别技术

在病菌孢子计数的研究中,目前大多采用显微镜观察法、分子生物学方法和基于显微图像处理法等。通过孢子捕捉仪捕捉到病菌孢子之后,光学显微镜下通过肉眼观测以确定孢子个数,存在工作量

大、效率低且随工作时间延长而准确性降低等缺点^[33-34]。利用分子生物学检测方法(PCR)鉴定 DNA 序列来定量检测具有客观、准确和高通量等优点^[35-37],但是,基于 PCR 技术的孢子计数方法操作复杂,成本较高,也耗时间^[38-40]。基于显微图像处理的孢子计数方法是在传统显微镜计数方法的基础上,利用计算机技术实现孢子的自动计数。图像处理首先对孢子显微图像进行灰度化、中值滤波去噪等预处理;其次使用阈值分割、边缘检测、分水岭分割和 K-means 聚类分割处理获取孢子目标区域^[41],然后常用形态学处理消除孢子区域的背景噪声和孔洞;最后通过标记计数法、平均面积法和角点检测法等实现孢子的自动计数。上述方法对未粘连的孢子能很好的计数,具有快捷、高效等特点。对于粘连孢子的情况,也有相关改进研究,如基于符号对数高斯混合模型相似度(SLGS)的水平集法、基于距离变换的改进分水岭算法、改进 Harris 角点检测法和循环标记腐蚀法^[42],但对复杂的多粘连情况下的鲁棒性和准确性不高,导致计数不准确是一个亟待解决的问题。研究表明深度学习方法相比于传统的手工提取特征的方法在图像识别领域具有巨大优势,逐渐应用到病菌孢子定量分析^[43],但是现有的通用深度学习模型在多形态、粘连和小目标孢子显微图像中并不能取得很好的识别效果,需要构建一个适合孢子显微图像的深度学习方法。作物病害致病过程与叶片上病菌孢子密度相关^[44],而上述开展的研究大多是针对孢子捕捉仪捕捉到的病菌孢子进行定量计数,文献少有探究病菌孢子侵染过程各个时期动态变化和时序演化规律。

2.1.3 作物病害病菌侵染行为分析技术

作物病害是病菌、环境和寄主作物 3 方面的统一体,当环境条件有利于病菌生长时,病菌进入细胞,通过病菌分泌的毒素和细胞壁降解酶致病^[45],进而引起叶绿素含量、气孔导度、叶表温度和孔隙结构等发生变化^[46],作物病菌的相关研究主要集中在生物学特性^[47-48]、抗病机制^[49]以及侵染特性^[50]等

方面。遵循病害三角关系,在研究作物病害发病的预测过程中,应该利用环境条件与致病真菌生长发育的关系,综合考虑影响病害的主导因素(温度、湿度和结露时间),其次还有一些其他的因子(病情指数、作物是否具有抗病性、菌源数或病菌孢子浓度和栽培条件等)^[51]。分析病情指数等病情预测模型大致可分为3类:①经验模型。基于生产经验、多年观察或从已有文献中归纳总结适宜的发病条件,通过定性、定量或数理统计构建模型表达式。使用最大空气湿度、最大空气温度、活动积温、活动积湿、累积相对湿度与气温的比值、昼夜温差等因子^[52-53]。②机理模型。能够详细地描述病害发展的各个阶段,从而更好地了解寄主与病原物之间的关系。③数理统计模型。通过与现代信息技术相结合,提高模型的准确率,并尝试自我学习来对病害进行模拟,如构建多元线性回归、Logistic 回归等模型。近年来,BP 神经网络、决策树和马尔科夫链等机器学习方法在病害预测的应用中也取得了阶段性研究成果。但是缺乏综合考虑作物病害三角关系及病菌孢子侵染过程动态演化规律的研究,无法满足当前作物绿色生产对病害时序化、数字化、精准化早期预警和防控的需求。

2.2 基于内部光谱成像的作物病害诊断技术

基于光谱成像的作物病害诊断涉及关键技术包括:基于热红外成像的作物病害检测技术、基于多光谱成像的作物病害检测技术、基于病症可见光图像的作物病害识别技术和基于病症可见光图像的作物病害严重度估算技术。

2.2.1 基于热红外成像的作物病害检测技术

热红外成像技术利用作物染病后的温度变化差异来对病害进行识别区分,该技术已开始应用于农作物病害的检测中,并取得了良好的效果。KIM 等^[54]利用数字红外热像仪研究了紫薇感染烟煤病后叶片温度场的空间分布规律,发现在热红外图像中,健康区域和染病区的平均温度分别为 26.98℃ 和 28.44℃,表明染病区的平均温度明显高于健康区域。LÓPEZ-LÓPEZ 等^[55]通过热成像和高光谱成像计算得出冠层温度和植被指数,并分析了它们在早期发现疾病的能力。结果显示,线性模型显示出更高的区分无症状树和红叶斑块发展后期树的能力,而非线性模型则更好地将无症状植物与红叶斑块发展的早期植物区分开。MASTRODIMOS 等^[56]为了评估空间温度的异质性,该研究利用热红外成像技术,计算了浆果表面的平均温度以及浆果表面受感染区域和未感染区域之间的最大温度差。研究发现,浆果中的真菌菌丝体发育期间的葡萄叶片平

均温度明显低于健康的葡萄,而在真菌定殖过程中的最大温度差却增加了。最后将热成像的温度数据分部进行拟合得出病害感染估计因子,实现了葡萄生理状态的无损监测。FAROKHZAD 等^[57]使用热像仪和加热箱获取热图像,研究处于不同阶段(感染后 1~7 d)的健康马铃薯块茎和被真菌污染的块茎温度,通过线性和二次判别分析方法提取并分类了一些温度统计特征。最终建立了一种基于主动热成像的可靠、无损、快速的方法来检测马铃薯块茎中的真菌。

由于欧美国家对我国的技术进出口限制,我国的红外热成像技术起步较晚。李小龙等^[58]通过连续采集小麦不同生理健康状态的植株热红外图像,分析叶片温度随锈病病害接种天数的变化趋势,实现了对小麦条锈病潜伏期叶片的检测与识别。朱文静等^[59]以感染叶锈病的小麦叶片为研究对象,分别采集健康组、潜伏期组和发病组的红外热图像,并利用边缘检测算法提取病斑的区域,根据病斑面积占比实现对小麦叶锈病的病害严重度分级。陈欣欣等^[60]利用热红外成像技术检测受菌核病侵染的油菜,发现利用热红外图像可在接种病害 24 h 后,观察到微小的病斑,且随着侵染时间的增加,病斑面积逐渐变大;但直到第 3 天肉眼才可以清晰地识别出病斑,表明热红外图像可以更早、更直观、更清晰地识别出作物染病早期的病害情况。温冬梅等^[61]通过热红外成像技术,记录了不同湿润持续时间下黄瓜霜霉病显症后叶片温度的变化,并分析了其温度变化规律,建立了黄瓜霜霉病流行趋势模型。姚志凤等^[62]进行了将热红外成像技术用于小麦条锈病早期检测的可行性研究。实验发现,随着接种时间的增加,接种病害的小麦植株冠层的平均温度会逐渐降低,叶片间的最大温差会不断加大。结果显示,热红外成像技术可观测到小麦条锈病病斑,较肉眼观察时间提前,可实现基于热红外成像技术的小麦条锈病早期检测。

热红外成像能更容易地观察到被病害侵染叶片的温度变化,将其作用于农作物病害检测,有着广阔的应用前景。但由于热红外成像受到光照、环境干扰较大,且由于热红外图像的像素质量限制,图像存在边缘模糊、信噪比较低等缺点,因此,基于热红外图像技术的作物病害诊断研究还需要进一步深入开展。

2.2.2 基于多光谱成像的作物病害检测技术

光谱成像技术是基于成像学和光谱学发展起来的一种技术,光谱成像技术可以同时从光谱维和空间维获取被测目标的信息等。一幅多光谱图像是由

一系列灰度图像组成的三维数据立方体,二维图像记录了样本的形态信息,三维坐标则记录光谱信息,映射出叶片每个像素点的组分含量和内部特性,有利于病害的精准定位以及早期诊断。

刘鑫等^[63]用波段指数法提取多光谱图像的特

征波段进行彩色合成,能快速获取马铃薯叶片的最佳波段。近年来更多的学者将多光谱相机与无人机关结合^[64-65],大面积诊断病害,相关文献表明将该技术应用在病害检测方面取得了较好的效果(表 2)。

表 2 基于多光谱成像的作物病害检测研究成果

Tab. 2 Research results of crop disease detection based on multispectral imaging				
研究对象	研究方法	技术原理	性能指标	文献序号
小麦黄锈病	改进的随机森林、贝叶斯分类器	采用随机森林作为分类器,利用贝叶斯优化对分类器的超参数进行微调,以保证分类器的性能	开发的黄锈病病期监测(接种后 45 d)的平均精度、召回率和准确率分别为 89.2%、89.4% 和 89.3%	[66]
	HLS 颜色特征提取	将不同水稻纹枯病严重程度的数字图像转换为色相、明度和饱和度(HLS),然后提取 HLS 的平均值	利用多光谱图像提取的基于图像的 NDVIs 可以量化田间图中不同水平的 ShB,准确度为 63%	
西瓜炭疽病、枯萎病、疫霉果腐病、链孢叶斑病、葫芦叶皱病	边际同质性 Bhapkar 检验	利用无人机采集的多光谱传感器数据,通过计算 NDVIs 和胁迫指数,检测了西瓜田间最常见的病害	比传统侦察方法更快速地识别疾病焦点和关注区域的能力增强,使用无人机辅助侦察的疾病早期检出率提高 20%	[68]
油菜光叶斑病	ND - SVM (改进的支持向量机)	通过定义正常的类边界来描述未见的测试样本,然后将决策函数的输出校准为类概率	在接种叶斑病后 12 d 和出现明显症状前 13 d 对整个油菜植株进行成像,能够以 92% 的准确率检测出光叶斑感染	[69]

2.2.3 基于病症可见光图像的作物病害识别技术
按照特征提取的方法可以将以往的基于病症可见光图像的作物病害识别技术研究划分为机器学习方法和深度学习方法。
基于机器学习方法的研究多是分割病斑、提取病斑特征、构建病害识别分类器的一个流程,目前文

献研究已表明此类方法已经取得了较好的识别效果(表 3)。首先通过条件随机场^[76]、Otsu 分割^[77]等分割方法获得病斑图像,进而提取病斑图像的颜色、纹理、形状等特征^[70-71,78],基于支持向量机、BP 神经网络、决策树等分类模型识别病害类别^[72,76-77]。所有上述用于病害识别的方法都是基于从病斑图像

表 3 基于机器学习的病害识别研究成果

Tab.3 Research on disease recognition based on machine learning				
研究对象	研究方法	技术原理	性能指标	文献序号
茄子褐纹病	基于费歇尔判别函数	提取 <i>H</i> 分量上病斑的纹理、形状与颜色特征,并建立费歇尔判别函数	识别的精准度达到 90%	[70]
黄瓜白粉病、霜霉病和灰霉病	基于 BP 神经网络动态集成	采用不一致度量方法对构建的单分类器进行差异性度量并进行排序,动态选择差异性大的分类器子集对图像进行集成识别	识别错误率为 3.32%	[71]
黄瓜霜霉病	基于支持向量机	基于条件随机场的方法分割黄瓜霜霉病病斑,提取病斑的 25 个底层特征,经粗糙集优化后选用 12 个特征构建支持向量机分类器	识别准确率为 90%	[72]
黄瓜角斑病、棒孢霉叶斑病、白粉病、霜霉病和炭疽病	基于词袋特征 PCA 多子空间自适应融合	建立多种病害类别相关词袋模型,提取病害图像的高维词袋特征;通过主成分分析法将病害高维特征降维到多个不同维数子空间;在各子空间上分别训练 BP 神经网络,通过设置自适应阈值对待分类图像在各子空间上的分类得分进行融合得到识别结果	在 2 个主成分子空间融合分类的平均准确率为 90.38%	[73]
马铃薯晚疫病	基于颜色纹理形状特征结合的识别方法	提取晚疫病斑的颜色、纹理和形状特征参数,建立马铃薯晚疫病的无病和患病模型,利用灰度共生矩阵,采用纹理统计参数进行病害等级评价,用熵值和能量值描述晚疫病的严重程度	对患病中期与后期的晚疫病识别率分别为 90% 和 92.5%	[74]
番茄晚疫病、黑斑病、细菌性斑病、细菌性溃疡病、番茄卷叶病	基于决策树	结合 Otsu 分割番茄病害病斑,提取病斑颜色、形状和纹理特征,构建决策树分类识别病害	对 6 种番茄病害图像的识别准确率为 97.3%	[75]

中提取的手工设计的特征,而人工设计的病斑特征难以完整的描述病害类别间的差异,容易出现图像语义鸿沟问题。这些局限性直接导致了该方法很难满足实际场景中病害识别的要求。

深度学习卷积神经网络的主要思想是通过深度神经网络的层层映射,来自自主学习图像像素特征、底层特征、高层抽象特征直至最终类别间的隐式表达关系,更加有利于捕获数据本身的丰富内涵信息,同时也避免了复杂的人工设计过程。卷积神经网络的发展为图像处理技术提供了新的契机。现今,卷积神经网络已在农业各领域得到了广泛应用^[73],如植物病虫害识别分类^[74-75,79]、植物器官计数^[75,80]、杂草识别^[81]等农业领域,并取得了令人欣喜的成果。在病害识别问题中,基于 AlexNet、VGGNet、GoogleNet 和 ResNet 等架构,结合迁移学习方法训练病害识别模型^[82-86],实验证明,迁移学习能够提高模型的准确率。除了现有 CNN 架构的应用之外,还提出了几种定制架构用于作物叶部病害检测,如三重损失的 FSL 网络^[87]、多尺度特征融合网络^[88]、无监督卷积自动编码器^[89]、注意力机制优化的网络^[90],在简单背景下的病害图像中均取得了较高的识别准确率。采集自建的数据集应用于特定作物类型病害的研究也很常见,翟肇裕等^[91]也做了相关研究和综述。但是实际环境下的图像背景复杂、光照条件多样、病斑小且不明显、病斑与背景对比度不大,两者很难区分。现有方法在面向实际场景复杂背景和噪声条件下的作物病害识别时,识别准确率往往会大大降低,识别速度也会变慢,难以满足实际应用需求。

2.2.4 基于病症可见光图像的作物病害严重度估算技术

一般在衡量病害发生程度时主要有两个指标:发病率和严重度,发病率是指同类被侵染的单位(叶片、植株、茎、果实)占同类总测量单位的百分比(0~100%),严重度则指病害的严重程度,对叶部病害来说,通常使用定性量表和定量量表进行评估。其中,定性量表使用描述性术语将病害严重程度描述为几种类别,如轻度、中度和重度。定量量表通常以百分比表示,即病斑面积与整个叶片面积的比值来表示。

随着计算机视觉技术的发展,许多研究者通过图像处理和机器学习方法进行作物病害的严重度评估,该方法具有相同的评估程序^[92]。李井祝等^[93]利用扫描仪扫描黄瓜霜霉病叶片得到扫描图像,采用线性运算得到病情指数,平均识别正确率达到 98.3%。鲍文霞等^[94]提出一种滑窗最大值特征提

取方法,对分割后的感染小麦白粉病的叶片图像采用滑窗法提取 HSV 颜色特征和 LBP 纹理特征,以此来识别叶部病害的严重度,准确度显著高于传统方法。GALLEGO-SANCHEZ 等^[95]开发了一个开源且用户友好型的脚本工具 RUST,基于颜色特征半自动评估叶锈病。通过以上研究可以发现,严重程度的计算结果依赖于图像分割技术,且已有的研究中大多基于简单背景的叶片,有的方法只能适用于单一的危害种类,难以应用到实际农业场景下采集的多噪声、复杂背景的危害叶片中。

深度学习在病害识别方面已经取得重大进展,在病害严重度估算方面也有应用。将定性量表和定量量表估算严重度转化为计算机学科问题,可以将严重度估算方法划分为基于分类、基于回归和基于深度分割的严重度估算方法,严重度估算研究成果如表 4 所示。基于分类的方法是指通过定义严重度的类别或区间将其转化为分类问题,采用卷积神经网络(CNN)建立输入图像与严重度类别的关联关系。文献[96-100]将病害严重程度划分为不同等级进行识别,取得了准确的结果,但是病害的分级难以具体量化病害严重度。基于回归和深度分割的严重度估算方法可以得到百分比的病害严重度,更具有说服力。基于回归的严重度估算方法是将输入的危害图像直接与严重度百分比对应起来。张领先等^[101]构建一个 CNN 模型估计黄瓜霜霉病的严重度,以手动去除背景的危害图像作为输入,证明了 CNN 的准确性优于浅层机器学习模型,决定系数 R^2 达到 0.919 0。然而,这种方法对背景噪声比较敏感。基于深度分割的严重度估算方法是指通过语义分割或实例分割方法为每个像素分配适当的标签,实现病斑、健康叶片的自动化分割,以获得百分比的严重度。常用的分割网络包括 DeepLab V3 +^[102]、U-Net^[103]、PSPNet 和 Mask R-CNN。相关研究表明语义分割模型在病害严重度估算中的应用是可行的,然而当数据量较小、图像存在大量复杂背景干扰时,这仍然是一个挑战。

基于深度学习严重度估算的思路比较如表 5 所示。

2.3 基于电子病历的作物病害处方推荐技术

作物病害处方推荐涉及的关键技术包括基于实体关联的病害机理解析、基于诊断推理的作物病害处方推荐、基于交互式语义匹配的作物病害处方推荐以及面向农户的作物病害在线问诊。

2.3.1 基于实体关联的病害机理解析

作物病害积累的基础数据可以提供病害的发病症状、发病阶段、用药方案等重要信息。作物病害处

表 4 基于深度学习的严重度估算研究成果

Tab.4 Research on severity estimation based on deep learning

严重度估算思路	研究对象	研究方法	技术原理	性能指标	文献序号
基于分类	苹果黑腐病	迁移学习的 VGG16	利用迁移学习训练 VGG16 模型,实现严重度的分类识别	对健康、轻度、中度和重度病害严重程度识别准确率为 90.4%	[96]
	9 种不同植物 27 个病害类别	改进 ResNet50	以 ResNet50 为基础,组合 shuffle 单元,并构建 3 个分支的多任务学习	对叶片种类、叶片病害、叶片病害程度分类的准确率分别达到 0.98、0.99、0.91	[97]
	木薯花叶病、褐条病、绿蚜病	CNN	使用针对移动设备优化的 CNN 架构的单个检测器模型,评估了模型在检测疾病类别的性能	自动检测两个严重级别,低严重程度检测的准确率为 29.4% ± 12.2%	[99]
	咖啡潜叶虫、锈病、褐叶斑病、尾孢叶斑病	改进 CNN	增加一个与现有层并行的新的全连接层,得到个性化的分类块	对 5 个严重程度识别的准确率为 84.13%	[100]
基于回归	黄瓜霜霉病	CNN	构建一个 CNN 模型,使用回归层得到严重度	证明了 CNN 的准确性优于浅层机器学习模型,决定系数 R^2 为 0.919 0	[101]
基于深度分割	黄瓜白粉病、霜霉病	改进 DeepLab V3 +	利用迁移学习和混合注意力优化 DeepLab V3 + 构建病害分割模型,基于分割图估算严重度	有效提升了复杂背景下病害严重度估算精度,估算误差为 1.138 5	[102]
	黄瓜霜霉病、白粉病、病毒病	DeepLab V3 + 和 U-Net	选择两个模型两步依次分割出叶片和病斑	平均病情严重程度分类准确率达到 92.85%	[103]

表 5 基于深度学习的严重度估算思路特点

Tab.5 Characteristics of severity estimation based on deep learning

严重度估算思路	思路描述	优势和局限
基于分类	采用深度学习技术自动提取病害图像特征,通过分类层输出不同严重度等级	能够估算定性的病害严重度,不能精准估算定量的严重度
基于回归	采用深度学习技术自动提取病害图像特征,利用回归层直接关联病害严重度百分比	能够精准估算定量的病害严重度,但是面对复杂背景图像估算精度不高,鲁棒性低
基于深度分割	采用基于语义分割、实例分割的深度学习技术自动分割病斑和叶片,通过像素比例得到严重度	能够精准估算定量的病害严重度,缺乏复杂背景下,快速、实时的严重度估算方法

方数据包含的作物信息、环境信息、病害信息对于作物病害机理解析间接提供了全方位真实数据源,同时基于宿主、病原体和环境的传统流行病学和植物病理学知识为处方数据分析提供了新的研究视角。

近年来,知识图谱作为一种语义网络,具有可扩展性强、支持智能应用等优点,因此在自然语言处理、智能问答系统、智能推荐系统等领域得到了广泛的应用。知识图谱(Knowledge graph)的本质是一个由大量实体及其之间的关系组成的大规模知识库。知识图谱包含了丰富的语义信息,作为一个庞大的基于知识系统,它相比于结构化数据库可以敏锐地获取领域实体间的复杂关联关系,并将其可视化展示,同时还可将分布于不同信息系统中的零碎知识连接起来。知识图谱基于图模型将知识抽象,可以为各领域提供简洁和直观的知识展示,其中边和路径可以捕获实体之间不同的、潜在的复杂关系^[104],解决了碎片化数据存储和关联关系挖掘的问题。

现有的研究已经从各种数据源中确定了药物和

疾病之间的实体^[105-107]和关系^[108-109],如图 6 所示。在下游任务中,知识图谱可以与机器学习等算法相结合,实现处方推荐^[110-112]。同时,基于知识图谱的推荐面临着高计算复杂性、缺乏长尾实体、规则冲突、扩展困难和在非结构化 EMR 中应用的局限性等

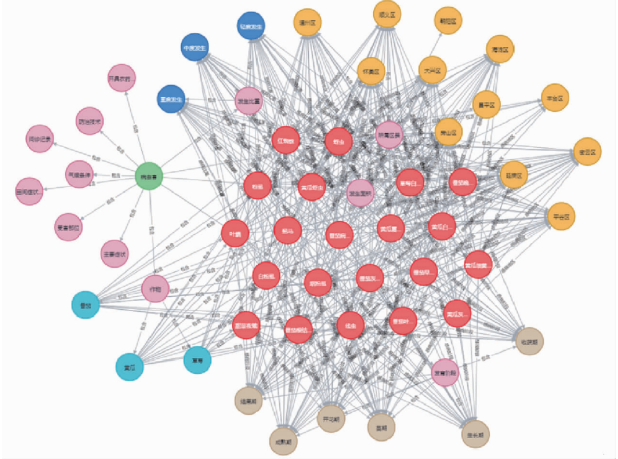


图 6 作物病害知识图谱部分展示

Fig.6 Part of crop disease knowledge map

挑战^[113]。

2.3.2 基于诊断推理的作物病害处方推荐

处方推荐与作物病害的诊断息息相关,一些研究通过机器学习、深度学习或者基于特征融合的多输入多输出方法挖掘电子病历信息,实现作物病害的准确诊断,最后结合规范的病害治疗方案来实现有效的作物病害处方推荐。

(1) 机器学习方法

机器学习可以从大量数据中挖掘出能够代表一类事物的规律,从而对事物进行预测、分类和推荐,是挖掘处方数据中有效信息的有力工具。机器学习算法具有计算时间短、精度高、可移植性强的优点,各种有监督和无监督的机器学习方法已经被应用于疾病诊断的研究。例如,VENKATESH等^[114]使用大数据预测分析模型,基于朴素贝叶斯(BPA-NB),对不同的诊断结果概率分类,进而给出治疗建议,对于UCI机器学习库中的疾病数据预测准确率为97.12%。WANG等^[115]基于处方数据开发的智能处方系统能够从药物信息中提取特征,根据问诊对象的症状预测药剂,对于同时确诊多种病的问诊对象减少重复药剂,给出适当的处方,能够减少14%潜在的重复处方。GALVEIA等^[116]提出了基于随机森林的分类器模型,用于推荐诊疗建议。

但是随着现实应用场景中数据量的激增和多元化,尤其是面对作物处方等具有复杂性和专业性的数据,传统的分类算法已经不能契合现存实际需求的需求。集成方法被认为是增强机器学习效果的高级解决方案^[117],尤其对于分类问题具有较强的优势^[118]。集成学习通过利用基础算法的多样性提高模型的分类准确度、泛化能力和鲁棒性^[119]。机器学习中提出了各种集成学习算法,其中最具代表性的方法是Bagging、Boosting和Stacking。Bagging算法生成并行基学习器,并使用随机抽样(bootstrapping)训练模型^[120-121]。Boosting方法依次训练一系列分类器,将弱分类器提升为强分类器,使错分的样本得到更多的关注。其代表性算法有Adaboost、梯度上升决策树(GDBT)、极限梯度提升(XGBoost)和轻量级梯度提升机(LightGBM)。在以上集成方法中,Stacking模型在分类问题上表现良好,特别是对不平衡数据分类。Stacking模型主要目的是减少泛化误差。由于单一分类器种类复杂且各具优势^[122],基于不同分类器的Stacking集成备受国内外学者的关注,经研究证明它能够在不同的应用场景下提高模型分类精度^[117-123]。但是机器学习模型仍然没有解决EMR中的自由文本语义理解问题。

(2) 深度学习方法

一些研究将疾病诊断问题转化为病历文本的分类问题,通过自然语言处理(NLP)方法挖掘电子病历信息,实现对疾病的诊断或风险评级。许多研究使用了深度学习方法,如神经网络(CNN)、循环神经网络(RNN)和自动编码器(AE),帮助计算机更好地理解电子医疗记录的语义^[124-126]。例如,ZHANG等^[127]提出的无监督深度学习框架能够注释电子病历中的表型异常数据,并使用不同的先验分布学习文本数据的语义潜在表示,预测诊断结果与处方内容。程铭等^[128]基于电子病历数据,构建混合注意力机制模型,分析病历文本之间的语义关系,展开处方推荐,同时采用自注意力机制从病历文本中识别特定病种的病历表示,将二者进行有机地融合,生成最终的病历表示,最后构建多标签分类器进行处方推荐。

深度学习方法通过训练大量带有标签的电子病历数据,在医学领域取得了良好的效果。但是在基于植物电子病历的作物疾病诊断中使用深度学习方法的缺点是缺乏足够的训练数据。原因在于CEMRs需要由专业的植物医生进行标注和记录,导致样本量小。变换器和预训练语言模型^[129]的提出为解决训练数据的局限性提供了一个突破口。预训练语言模型可以从大量的语料库中学习通用的语言表征,而不需要人工标注^[130]。一些研究在任务领域的数据集上对语言表示模型进行了领域适应性预训练^[131]。例如,DING等^[132]提出基于作物疾病领域BERT和RCNN(CdsBERT-RCNN)的作物疾病诊断模型,为进一步实现基于诊断推理的作物病害处方推荐打下基础。

(3) 多输入多输出模型

植物电子病历不是简单的文本描述,而是经过科学设计的、符合植物病理学中病害诊断基本原理的规范结构,包含结构化的地理、时间、环境、分布等特征。研究证明,病害发生的环境特征、时空分布等信息对病害的准确识别意义重大,但是这些信息在病害智能诊断的研究中尚未得到有效利用^[133-134]。如果仅聚焦于植物电子病历中的单一类型数据,仅对问诊记录文本或者结构化数据进行特征抽取,将会造成大量的信息损失。丁俊琦等^[13]提出基于多类型数据融合的病害诊断模型用于解决这个问题。

得到诊断结果后,可以进一步实现处方的推荐,即用药名称和数量的确定,一些研究使用多输出(Multi-output)方法结合机器学习模型实现此功能。以多输出结合机器学习进行预测的方法在声学、力学以及通信领域被广泛应用。ZHOU等^[135]将

多输出支持向量机(M-SVM)和多任务学习(MTL)算法相结合,通过解决区域预测中常见的误差积累问题,有效提高区域多步提前预测的准确性。应启帆等^[136]通过对单种粒径预测的梯度提升决策树算法进行组合构建多输出回归算法对粒径分布进行预测。

2.3.3 基于语义匹配的作物病害处方推荐

语义匹配是 NLP 领域的基础问题之一,被广泛应用于信息检索、推荐系统和问答系统等下游任务。基于处方内容语义匹配的处方推荐方法是通过对处方文本展开分析,根据历史处方数据生成推荐列表,推荐结果更具多样性,可扩展性更强。语义匹配包括交互型和表示型两种匹配方式。

(1)交互型语义匹配

基于文本相似度的处方推荐方法是通过分析处方文本中的语义信息,计算向量得到语义相似度,生成推荐列表。ZHANG 等^[137]提出的电子病历相似度计算方法,根据检查项目将电子病历划分为不同部分,筛选有效部分后运用词向量与词移距离(Word mover's distance, WMD)计算相似度,最后利用 KNN 聚类对电子病历间的相似性进行评价,与 LDA 和 LSI 等传统的疾病分类方法相比,该方法具有较高的召回值,能够改进处方推荐效果。赵明等^[138]基于双向门控循环单元神经网络(BiGRU)构建病虫害问句分类模型,利用问句的语义信息,辅助实现番茄病虫害智能诊疗。YE 等^[139]使用词嵌入将处方文本语义上相似的词投射到向量空间中的邻近点,提升了诊疗系统的检索与决策支持功能,证明使用语义相似的术语,可以更快速地检索和推荐诊疗建议。邱硕等^[140]使用聚类的方法挖掘电子病历中的处方关联,依据问诊对象相似度实现处方的多样化,同时程序执行时间有所提升。对于文本相似度计算,深度语义匹配模型(DSSM)通过多层次的语义分析表现更好。XIE 等^[141]提出的主题增强的语义匹配模型在有关语义匹配的问答库任务中获得了 21 个系统中的第 3 名,表现出较强的语义分析能力。LARIONOVA 等^[142]基于推荐系统构建 DSSM,学习推荐目标之间的相似性,对不同类别内的推荐对象进行排序,结果表明,DSSM 相比传统相似度推荐方法显著提高了推荐的总体质量。交互计算更好地把握了语义焦点和上下文重要性,但是计算成本很高。

(2)表示型语义匹配

基于表示型语义匹配的作物病害处方推荐方法核心是句嵌入,在表示层将文本转换成整体的表示向量之后再进行匹配。在推荐系统中,基于表示的模型可以通过句子嵌入对文本预处理,构建索引,大

幅度降低在线计算耗时。基于 BERT, REIMERS 提出了 Sentence-BERT^[143],它是目前最常用的 BERT 式双塔模型,效果较好,提供方便的开源工具,可以有效缓解处方推荐中的在线计算耗时问题。GAO 等^[144]提出了一个简单的句子嵌入的对比学习框架(SimCSE),包括无监督和有监督的版本,实现了基于对比学习和辍学数据增强的句子级语义表示的 SOTA 性能。

2.3.4 面向农户的作物病害在线问诊

问答系统的应用涉及诸多领域,如医药、电力、交通等各方面^[145]。问答系统技术在农业领域发展迅速,并已经形成了一些相对完整的体系。传统的农业信息服务多为上网搜索、电话咨询和专家现场指导等方式,张博凯等^[146]基于网络爬虫得到的大量农业问答知识数据形成的语料库,结合命名体识别和知识图谱查询推荐算法,设计实现 Android 端的智能问答机器人,为农业领域智能信息服务提供了一种新的解决方案。张领先等^[147]开发了面向移动终端的作物病害处方推荐系统。用户输入受害作物的症状描述后,系统输出诊断结果及相应处方,实现了面向实际应用场景的作物病害处方推荐。

3 作物病害诊断与处方推荐研究难点与发展趋势

3.1 作物病害诊断与处方推荐研究难点

国内外学者在作物病害诊断与处方推荐方面开展了广泛的研究,既取得了较多的研究成果,也面临着一些亟需解决的难点。

(1)目前,计算机视觉技术可以实现病原物的持续监测。但是在实际应用中,病菌侵染作物是一个动态的过程,病菌孢子形态和数量在侵染过程中会受到作物抗病性以及环境温度和湿度等因素影响,使得基于机器视觉技术准确提取与分析病菌孢子形态特征、动态变化规律及其病害三角关系等成为研究的关键科学问题和难点。尤其是病菌孢子交叉、遮挡、动态变化等特点导致病菌孢子量化识别困难等。

(2)热成像和多光谱成像技术能够根据内部生理变化检测发病期之前的早期侵染。但是热成像受环境影响较大,检测植物病害时需要严格控制环境温湿度,而对于多光谱成像,许多学者采用光谱指数或者需要选取感兴趣区域、图像分割等处理,过程复杂且受限于人工选取特征。文献[148-149]证明通过结合各种传感器系统中包含的丰富光谱,空间、结构和热信息的优势来改善植物性状估计。因此,研究基于多源图像的病害早期检测方法,提高病害

侵入期的检测效果。

(3)卷积网络有强大的特征学习能力,基于卷积神经网络的作物病害识别方法可以快速、准确地识别病害种类。但是现有研究大都针对公开数据集,部分自己采集的数据也都是简单背景,在实际应用时受环境等因素影响导致识别精度不够,因此,针对农业领域复杂背景,高精度、泛化性强的病害识别方法有待于进一步研究。

(4)作物病害严重度定量估算效果受病斑分割和特征提取的影响,分割操作繁琐,易受光照影响,提取特征又有一定的主观性,会影响模型的泛化能力。因此,研究基于深度学习的自动化作物病害分割方法,可以提高分割精度并计算作物病害严重度。

(5)目前关于作物病害的研究大多以设施温室(小气候)环境为基础,多停留在单一数据源的获取或基于单一作物的小尺度分析,而缺乏从宏观角度基于数据挖掘解析病害三角原理的研究。而作物病害处方数据几乎未被应用于辅助处方推荐,其中包含了大量区域性作物信息、环境信息和病害信息及其防治知识,可以解决多源数据采集难的问题。

(6)基于诊断推理的作物病害处方推荐鲜有研究。与常规推荐算法使用的场景不同,处方数据大多为结构化数据,且为多变量数据。Multi-output结合机器学习模型已广泛应用于声学、力学以及通信领域,但是在作物病害治疗方案推荐方面鲜有研究。深度学习算法可以根据采集的环境信息及作物生长信息辅助病害诊断,即对应计算机领域的多分类问题,其中集成学习算法对于不平衡数据集的处理具有一定的优势。同时,还未有研究从多尺度角度利用数据和深度学习算法根据病害发生机理进行病害诊断的研究。

(7)基于语义匹配的处方推荐方法的推荐结果更具多样性,可扩展性更强。但是,目前相关研究大多是生物医学领域,农业领域的应用偏少,实现深度语义匹配在农业领域的处方推荐应用将有助于提高

病害治理效果。因此,基于语义匹配,尤其是表示型语义匹配的处方推荐是重要的研究方向。

(8)对于我国区域作物生产和小农户分散种植国情,由于受到数据获取困难和物联网技术实施成本高以及作物病害发生态势复杂和传播途径多样等因素的限制,多应用场景、时空迁移和多目标决策的作物病害早期诊断、预测与主动防控成为难点。因此,基于电子病历多模态数据的作物病害关联挖掘与多目标决策研究,将对农业病害防治领域具有更大的实际意义,为实际应用复杂生产场景作物病害早期预警与主动防控提供决策支持和参考。

3.2 作物病害诊断与处方推荐发展趋势

(1)开展作物病害早期检测以及定量识别诊断方法的研究是必要的。在病害发病之前,深入挖掘可见光、热成像、多光谱图像数据对病害早期特征的解析能力,同时探索多源图像对侵入期病害检测的新思路;在病害发病期,提高复杂背景下病害的识别精度,准确量化病害严重度,为精准施药提供依据,对提高作物病害精细化管理水平,提升作物品质有重要意义。

(2)针对作物病害处方推荐过程中,存在由于作物病害致病机理复杂、作物品种及病害种类多、病害病症动态变化等特点导致缺乏可行的数据挖掘技术等问题,以作物病害处方为研究对象,针对电子病历数据特点,开展基于机器学习和知识图谱的作物病害致病机理解析、诊断推理、处方智能化推荐及其应用策略研究;攻克基于知识图谱分析、大数据挖掘和机器学习算法推理等关键技术,在作物病害处方数据挖掘分析研究;可视化分析作物病害病症形态特征、时空变化及其与种植环境和作物品种的病害三角关系,从区域宏观视角解析作物病害致病机理及其与特征间的关联关系,面向不同实际应用场景需求实现作物病害精准诊断与处方推荐。研究成果可为作物种植智能诊断提供科学依据和方法支撑,推进农业科技服务新模式、新业态。

参 考 文 献

[1] 杨森, 冯全, 张建华, 等. 基于深度学习与复合字典的马铃薯病害识别方法[J]. 农业机械学报, 2020, 51(7): 22-29. YANG Sen, FENG Quan, ZHANG Jianhua, et al. Identification method for potato disease based on deep learning and composite dictionary[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 22-29. (in Chinese)

[2] 王春山, 赵春江, 吴华瑞, 等. 采用双模态联合表征学习方法识别作物病害[J]. 农业工程学报, 2021, 37(11): 180-188. WANG Chunshan, ZHAO Chunjiang, WU Huarui, et al. Recognizing crop diseases using bimodal joint representation learning[J]. Transactions of the CSAE, 2021, 37(11): 180-188. (in Chinese)

[3] 徐畅, 丁俊琦, 赵聃桐, 等. 基于 LightGBM 和处方数据的番茄病害诊断方法[J]. 农业机械学报, 2022, 53(9): 286-294. XU Chang, DING Junqi, ZHAO Dantong, et al. Tomato disease diagnosis method based on LightGBM and prescription data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(9): 286-294. (in Chinese)

[4] ACZON M, LEDBETTER D, HO L, et al. Dynamic mortality risk predictions in pediatric critical care using recurrent neural networks[J]. arXiv Preprint, arXiv: 1701.06675, 2017.

[5] BEN MILED Z, HAAS K, BLACK C M, et al. Predicting dementia with routine care EMR data[J]. Artificial Intelligence in

- Medicine, 2020, 102: 101771.
- [6] JOHNSON T W. Review of plant pathology, by G. N. Agrios[J]. The Bryologist, 1970, 73(4): 736.
- [7] BUSBY P E, RIDOUT M, NEWCOMBE G. Fungal endophytes: modifiers of plant disease[J]. Plant Molecular Biology, 2016, 90(6): 645–655.
- [8] MAHLEIN A K. Plant disease detection by imaging sensors-parallels and specific demands for precision agriculture and plant phenotyping[J]. Plant Disease, 2016, 100(2):241–251.
- [9] 李凯雨,朱昕怡,马浚诚,等. 基于混合扩张卷积和注意力的黄瓜病害严重度估算方法[J]. 农业机械学报, 2023,54(2): 231–239.
- LI Kaiyu, ZHU Xinyi, MA Juncheng, et al. Estimation method of leaf disease severity of cucumber based on mixed dilated convolution and attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(2):231–239. (in Chinese)
- [10] RAY M, RAY A, DASH S, et al. Fungal disease detection in plants: traditional assays, novel diagnostic techniques and biosensors[J]. Biosensors & Bioelectronics, 2017, 87: 708–723.
- [11] 肖长坤,周春江,王克武,等. 北京植物诊所建设的探索与实践[J]. 中国植保导刊, 2012, 32(11):57–60.
- XIAO Changkun, ZHOU Chunjiang, WANG Kewu, et al. Exploration and practice of the construction of the Beijing plant clinics[J]. China Plant Protection, 2012,32(11):57–60. (in Chinese)
- [12] 赵磊,张涛,郑书恒,等. 运行植物诊所破解绿色防控技术推广难题[J]. 中国植保导刊, 2018, 38(2):84–85.
- ZHAO Lei, ZHANG Tao, ZHENG Shuheng, et al. Operation of plant clinics to solve the problem of promoting green prevention and control technology[J]. China Plant Protection, 2018, 38(2):84–85. (in Chinese)
- [13] 丁俊琦,李博,乔岩,等. 基于植物电子病历多类型数据融合的作物病害诊断方法[J]. 农业机械学报, 2023, 54(1): 196–204,223.
- DING Junqi, LI Bo, QIAO Yan, et al. Crop disease diagnosis method based on fusion of multiple types of data from plant EMRS[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1):196–204,223. (in Chinese)
- [14] 王蓬辉,李明正,李思. 基于数据增强的中文医疗命名实体识别[J]. 北京邮电大学学报, 2020, 43(5): 84–90.
- WANG Penghui, LI Mingzheng, LI Si. Data augmentation for Chinese clinical named entity recognition[J]. Journal of Beijing University of Posts and Telecommunications, 2020, 43(5): 84–90. (in Chinese)
- [15] 余朝阳,严馨,徐广义,等. 融合数据增强与半监督学习的药物不良反应检测[J]. 计算机工程, 2022,48(6):314–320.
- SHE Zhaoyang, YAN Xin, XU Guangyi, et al. Adverse drug reaction detection combined with data augmentation and semi-supervised learning[J]. Computer Engineering, 2022, 48(6): 314–320. (in Chinese)
- [16] WEI J, ZOU K. EDA: easy data augmentation techniques for boosting performance on text classification tasks[C] // 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP 2019), 2019: 6382–6388.
- [17] 徐聪,李擎,张德政,等. 文本生成领域的深度强化学习研究进展[J]. 工程科学学报, 2020, 42(4): 399–411.
- XU Cong, LI Qing, ZHANG Dezheng, et al. Research progress of deep reinforcement learning applied to text generation[J]. Chinese Journal of Engineering, 2020, 42(4):399–411. (in Chinese)
- [18] HANAUER D A, MEI Q, LAW J, et al. Supporting information retrieval from electronic health records: a report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE)[J]. Journal of Biomedical Informatics, 2015, 55: 290–300.
- [19] YU Kaiye, YANG Zhongliang, WU Chuhan, et al. In-hospital resource utilization prediction from electronic medical records with deep learning[J]. Knowledge-Based Systems, 2021, 223:107052.
- [20] LIU Y, YU Z, YANG Y. Diabetes risk data mining method based on electronic medical record analysis[J]. Journal of Healthcare Engineering, 2021, 2021(6):1–11.
- [21] 刘怡. 数据挖掘技术在电子病历系统中的应用研究[J]. 科技创新导报, 2019,16(6):136–137.
- LIU Yi. Research on the application of data mining technology in electronic medical record system[J]. Science and Technology Innovation Herald, 2019,16(6):136–137. (in Chinese)
- [22] ANDRE S, CHARTON A, PONS A, et al. Viability of bacterial spores surviving heat-treatment is lost by further incubation at temperature and pH not suitable for growth[J]. Food Microbiology, 2021, 95: 103690.
- [23] LEI Yu, YAO Zhifeng, HE Dongjian. Automatic detection and counting of urediniospores of Puccinia striiformis f. sp. tritici using spore traps and image processing[J]. Scientific Reports, 2018, 8: 13647.
- [24] DHAL K G, DAS A, RAY S, et al. Nature-inspired optimization algorithms and their application in multi-thresholding image segmentation[J]. Archives of Computational Methods in Engineering, 2020, 27(3): 855–888.
- [25] LOUTFI H, FAYAD N, PELLEN F, et al. Morphological study of bacillus thuringiensis crystals and spores[J]. Applied Sciences-Basel, 2021, 11(1): 155.
- [26] ZHANG Dongyan, WANG Zhicun, JIN Ning, et al. Evaluation of efficacy of fungicides for control of wheat fusarium head blight based on digital imaging[J]. IEEE Access, 2020, 8:109876–109890.
- [27] LI Xiaolong, MA Zhanhong, BIENVENIDO F, et al. Development of automatic counting system for urediospores of wheat stripe rust based on image processing[J]. International Journal of Agricultural and Biological Engineering, 2017, 10(5):134–143.
- [28] ZHAO Yaochi, LIU Shiguang, HU Zhuhua, et al. Separate degree based Otsu and signed similarity driven level set for segmenting and counting anthrax spores[J]. Computers and Electronics in Agriculture, 2020, 169:105230.

- [29] IPPOLITI P, NARGI F, HAN J, et al. Highly accurate classification of biological spores by culture medium for forensic attribution using multiple chemical signature types and machine learning[J]. *Analytical and Bioanalytical Chemistry*, 2020, 412(18):4287–4299.
- [30] KURODA S, OKUDA H, ISHIDA W, et al. Modeling growth limits of *Bacillus* spp. spores by using deep-learning algorithm[J]. *Food Microbiology*, 2019, 78: 38–45.
- [31] 李鑫铭, 赵磊, 邵宝民, 等. 一种新的基于形态学与 Yolov3 算法的显微镜图像中孢子识别方法[J]. *智能计算机与应用*, 2020, 10(5): 30–34.
- LI Xinming, ZHAO Lei, SHAO Baomin, et al. A new method for spore identification in microscope images based on morphology and Yolov3 algorithm[J]. *Intelligent Computer and Applications*, 2020, 10(5): 30–34. (in Chinese)
- [32] LI K, ZHU X, QIAO C, et al. The gray mold spore detection of cucumber based on microscopic image and deep learning[J]. *Plant Phenomics*, 2023, 5: 0011.
- [33] WALLACE E C, D'ARCANGELO K N, QUESADA-OCAMPO L M. Population analyses reveal two host-adapted clades of *pseudoperonospora cubensis*, the causal agent of cucurbit downy mildew, on commercial and wild cucurbits[J]. *Phytopathology*, 2020, 110(9): 1578–1587.
- [34] ANDERSSON M A, SALO J, KEDVES O, et al. Bioreactivity, guttation and agents influencing surface tension of water emitted by actively growing indoor mould isolates[J]. *Microorganisms*, 2020, 8(12): 1940.
- [35] DHAR N, MAMO B E, SUBBARAO K V, et al. Measurements of aerial spore load by qPCR facilitates lettuce downy mildew risk advisement[J]. *Plant Disease*, 2020, 104(1): 82–93.
- [36] WITHERS S, GONGORA-CASTILLO E, GENT D, et al. Using next-generation sequencing to develop molecular diagnostics for *pseudoperonospora cubensis*, the cucurbit downy mildew pathogen[J]. *Phytopathology*, 2016, 106(10):1105–1116.
- [37] BRACHACZEK A, KACZMAREK J, JEDRYCZKA M. Monitoring blackleg (*leptosphaeria* spp.) ascospore release timing and quantity enables optimal fungicide application to improved oilseed rape yield and seed quality[J]. *European Journal of Plant Pathology*, 2016, 145(3): 643–657.
- [38] RAY M, RAY A, DASH S, et al. Fungal disease detection in plants: traditional assays, novel diagnostic techniques and biosensors[J]. *Biosensors & Bioelectronics*, 2017, 87: 708–723.
- [39] LEE J H, PARK M H, LEE S. Identification of *pseudoperonospora cubensis* using real-time PCR and high resolution melting (HRM) analysis[J]. *Journal of General Plant Pathology*, 2016, 82(2): 110–115.
- [40] SUBBARAO C S, ANCHIETA A, OCHOA L, et al. Detection of latent *peronospora effusa* infections in spinach[J]. *Plant Disease*, 2018, 102(9): 1766–1771.
- [41] LI Zhenbo, GUO Ruohao, LI Meng, et al. A review of computer vision technologies for plant phenotyping[J]. *Computers and Electronics in Agriculture*, 2020, 176: 105672.
- [42] GHARIPOUR A, LIEW A W C. Segmentation of cell nuclei in fluorescence microscopy images: an integrated framework using level set segmentation and touching-cell splitting[J]. *Pattern Recognition*, 2016, 58:1–11.
- [43] ZHANG Ying, LI Jiangtao, TANG Fang, et al. An automatic detector for fungal spores in microscopic images based on deep learning[J]. *Applied Engineering in Agriculture*, 2021, 37(1):85–94.
- [44] PETRIACQ P, STASSEN J H M, TON J. Spore density determines infection strategy by the plant pathogenic fungus *plectosphaerella cucumerina*[J]. *Plant Physiology*, 2016,170(4): 2325–2339.
- [45] KANKAM F, QIU Huizhen, PU Lumei, et al. Isolation, purification and characterization of phytotoxins produced by *rhizoctonia solani* AG-3, the cause agent of potato stem canker[J]. *American Journal of Potato Research*, 2016, 93(4): 321–330.
- [46] 张妍, 王蕊, 李天来, 等. 黄瓜霜霉病菌侵染黄瓜叶片过程中内部组织变化的研究[J]. *植物生理学报*, 2019, 55(12): 1797–1805.
- ZHANG Yan, WANG Rui, LI Tianlai, et al. Changes in cucumber leaf internal tissues during the infection of cucumber downy mildew[J]. *Plant Physiology Journal*, 2019, 55(12): 1797–1805. (in Chinese)
- [47] WANG Yafei, MA Guoxin, DU Xiaoxue, et al. Effects of nutrient solution irrigation quantity and downy mildew infection on growth and physiological traits of greenhouse cucumber[J]. *Agronomy*, 2020, 10(12): 1921.
- [48] WEN Dongmei, CHEN Meixiang, ZHAO Li, et al. Use of thermal imaging and fourier transform infrared spectroscopy for the pre-symptomatic detection of cucumber downy mildew[J]. *European Journal of Plant Pathology*, 2019, 155(2): 405–416.
- [49] D'ARCANGELO K N, ADAMS M L, KERNS J P, et al. Assessment of fungicide product applications and program approaches for control of downy mildew on pickling cucumber in North Carolina[J]. *Crop Protection*, 2021, 140: 105412.
- [50] XI Yadong, XIANG YunJia, HAN Shuai, et al. Advances in *Pseudoperonospora cubensis* and its resistance in cucumber[J]. *Journal of Southern Agriculture*, 2019, 47(10): 1709–1714.
- [51] NEUFELD K N, KEINATH A P, OJIAMBO P S. A model to predict the risk of infection of cucumber by *Pseudoperonospora cubensis*[J]. *Microbial Risk Analysis*, 2017, 6:21–30.
- [52] THIESSEN L D, NEILL T M, MAHAFFEE W F. Assessment of *erysiphe necator* ascospore release models for use in the mediterranean climate of western Oregon[J]. *Plant Disease*, 2018, 102(8): 1500–1508.
- [53] 杨银娟, 鞠中安, 施颖红, 等. 黄瓜霜霉病春秋发生的气象要素预警模型[J]. *中国植保导刊*, 2018, 38(8): 43–47, 70.
- YANG Yinjuan, JU Zhongan, SHI Yinghong, et al. Early warning model of meteorological elements of cucumber downy mildew

- in spring and autumn[J]. *China Plant Protection*, 2018, 38(8): 43–47, 70. (in Chinese)
- [54] KIM J, KWEON S, PARK J, et al. Digital infrared thermal imaging of crape myrtle leaves infested with sooty mold[J]. *The Plant Pathology Journal*, 2016, 32(6): 563–569.
- [55] LÓPEZ-LÓPEZ M, CALDERON R, GONZALEZ-DUGO V, et al. Early detection and quantification of almond red leaf blotch using high-resolution hyperspectral and thermal imagery[J]. *Remote Sensing*, 2016, 8(4): 276.
- [56] MASTRODIMOS N, LENTZOU D, TEMPLALEXIS C, et al. Development of thermography methodology for early diagnosis of fungal infection in table grapes: the case of *aspergillus carbonarius*[J]. *Computers and Electronics in Agriculture*, 2019, 165: 104972.
- [57] FAROKHZAD S, MOTLAGH A M, MOGHADAM P A, et al. Application of infrared thermal imaging technique and discriminant analysis methods for non-destructive identification of fungal infection of potato tubers [J]. *Journal of Food Measurement and Characterization*, 2020, 14(1): 88–94.
- [58] 李小龙,王库,马占鸿,等. 基于热红外成像技术的小麦病害早期检测[J]. *农业工程学报*, 2014, 30(18): 183–189.
LI Xiaolong, WANG Ku, MA Zhanhong, et al. Early detection of wheat disease based on thermal infrared imaging[J]. *Transactions of the CSAE*, 2014, 30(18): 183–189. (in Chinese)
- [59] 朱文静,陈华,李林,等. 基于红外热成像边缘检测算法的小麦叶锈病分级研究[J]. *农业机械学报*, 2019, 50(4): 36–41.
ZHU Wenjing, CHEN Hua, LI Lin, et al. Grading of wheat leaf rust based on edge detection of infrared thermal imaging[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(4): 36–41. (in Chinese)
- [60] 陈欣欣,刘子毅,吕美巧,等. 基于热红外成像技术的油菜菌核病早期检测研究[J]. *光谱学与光谱分析*, 2019, 39(3): 730–737.
CHEN Xinxin, LIU Ziyi, LÜ Meiqiao, et al. Diagnosis and monitoring of sclerotinia stem rot of oilseed rape using thermal infrared imaging[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(3): 730–737. (in Chinese)
- [61] 温冬梅,纪涛,陈梅香,等. 不同湿润持续时间及叶片温度对温室黄瓜霜霉病发生的影响[J]. *植物保护学报*, 2019, 46(3): 663–669.
WEN Dongmei, JI Tao, CHEN Meixiang, et al. Effects of the duration of leaf wetness and leaf temperature on downy mildew in cucumber[J]. *Journal of Plant Protection*, 2019, 46(3): 663–669. (in Chinese)
- [62] 姚志凤,雷雨,何东健. 基于高光谱成像的小麦白粉病与条锈病识别[J]. *光谱学与光谱分析*, 2019, 39(3): 969–976.
YAO Zhifeng, LEI Yu, HE Dongjian. Identification of powdery mildew and stripe rust in wheat using hyperspectral imaging [J]. *Spectroscopy and Spectral Analysis*, 2019, 39(3): 969–976. (in Chinese)
- [63] 刘鑫,冯洁,杨舒明. 基于多光谱图像选取马铃薯叶片的最佳彩色波段组合[J]. *光学仪器*, 2017, 39(5): 35–40.
LIU Xin, FENG Jie, YANG Shuming. Selecting best color band combinations for potato leaves based on multi-spectral images [J]. *Optical Instruments*, 2017, 39(5): 35–40. (in Chinese)
- [64] JAY S, COMAR A, BENICIO R, et al. Scoring cercospora leaf spot on sugar beet: comparison of UGV and UAV phenotyping systems[J]. *Plant Phenomics*, 2020, 2020(3): 1–18.
- [65] HATTON N M, MENKE E, SHARDA A, et al. Assessment of sudden death syndrome in soybean through multispectral broadband remote sensing aboard small unmanned aerial systems[J]. *Computers and Electronics in Agriculture*, 2019, 167: 105094.
- [66] SU J Y, LIU C J, COOMBES M, et al. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery[J]. *Computers and Electronics in Agriculture*, 2018, 155: 157–166.
- [67] ZHANG D, ZHOU X, ZHANG J, et al. Detection of rice sheath blight using an unmanned aerial system with high-resolution color and multispectral imaging[J]. *Plos One*, 2018, 13(5): e0187470.
- [68] KALISCHUK M, PARET M L, FREEMAN J H, et al. An improved crop scouting technique incorporating unmanned aerial vehicle-assisted multispectral crop imaging into conventional scouting practice for gummy stem blight in watermelon[J]. *Plant Disease*, 2019, 103(7): 1642–1650.
- [69] VEYS C, CHATZIAVGERINOS F, ALSUWAIDI A, et al. Multispectral imaging for presymptomatic analysis of light leaf spot in oilseed rape[J]. *Plant Methods*, 2019, 15: 4.
- [70] 秦立峰,何东健,宋怀波. 词袋特征 PCA 多子空间自适应融合的黄瓜病害识别[J]. *农业工程学报*, 2018, 34(8): 200–205.
QIN Lifeng, HE Dongjian, SONG Huaibo. Bag of words feature multi-PCA subspace adaptive fusion for cucumber diseases identification[J]. *Transactions of the CSAE*, 2018, 34(8): 200–205. (in Chinese)
- [71] 党满意,孟庆魁,谷芳,等. 基于机器视觉的马铃薯晚疫病快速识别[J]. *农业工程学报*, 2020, 36(2): 193–200.
DANG Manyi, MENG Qingkui, GU Fang, et al. Rapid recognition of potato late blight based on machine vision [J]. *Transactions of the CSAE*, 2020, 36(2): 193–200. (in Chinese)
- [72] 王志彬,王开义,王书锋,等. 基于动态集成的黄瓜叶部病害识别方法[J]. *农业机械学报*, 2017, 48(9): 46–52.
WANG Zhibin, WANG Kaiyi, WANG Shufeng, et al. Recognition method of cucumber leaf diseases with dynamic ensemble learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(9): 46–52. (in Chinese)
- [73] KAMILARIS A, PRENAFETA-BOLDU F X. Deep learning in agriculture: a survey [J]. *Computers and Electronics in Agriculture*, 2018, 147: 70–90.
- [74] PICON A, ALVAREZ-GILA A, SEITZ M, et al. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild[J]. *Computers and Electronics in Agriculture*, 2019, 161: 280–290.

- [75] LU H, CAO Z, XIAO Y, et al. TasselNet: counting maize tassels in the wild via local counts regression network[J]. *Plant Methods*, 2017, 13: 79.
- [76] 马浚诚, 温皓杰, 李鑫星, 等. 基于图像处理的温室黄瓜霜霉病诊断系统[J]. *农业机械学报*, 2017, 48(2): 195–202.
MA Juncheng, WEN Haojie, LI Xinxing, et al. Downy mildew diagnosis system for greenhouse cucumbers based on image processing[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(2): 195–202. (in Chinese)
- [77] SABROL H, SATISH K. Tomato plant disease classification in digital images using classification tree[C] // *International Conference on Communication & Signal Processing*. IEEE, 2016: 1242–1246.
- [78] 田凯, 张连宽, 熊美东, 等. 基于叶片病斑特征的茄子褐纹病识别方法[J]. *农业工程学报*, 2016, 32(增刊1): 184–189.
TIAN Kai, ZHANG Liankuan, XIONG Meidong, et al. Recognition of phomopsis vexans in solanum melongena based on leaf disease spot features[J]. *Transactions of the CSAE*, 2016, 32(Supp.1): 184–189. (in Chinese)
- [79] LU Y, YI S J, ZENG N Y, et al. Identification of rice diseases using deep convolutional neural networks[J]. *Neurocomputing*, 2017, 267: 378–384.
- [80] UBBENS J, CIESLAK M, PRUSINKIEWICZ P, et al. The use of plant models in deep learning: an application to leaf counting in rosette plants[J]. *Plant Methods*, 2018, 14(6): 29375647.
- [81] BAH M D, HAFIANE A, CANALS R. Deep learning with unsupervised data labeling for weed detection in line crops in UAV images[J]. *Remote Sensing*, 2018, 10(11): 1690.
- [82] JIA Shijie, JIA Peiyi, HU Siping, et al. Automatic detection of tomato diseases and pests based on leaf images[C] // *2017 Chinese Automation Congress, CAC*, 2017: 3507–3510.
- [83] SURYAWATI E, SUSTIKA R, YUWANA R S, et al. Deep structured convolutional neural network for tomato diseases detection[C] // *2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2018.
- [84] RANGARAJAN A K, PURUSHOTHAMAN R, RAMESH A. Tomato crop disease classification using pre-trained deep learning algorithm[J]. *Procedia Computer Science*, 2018, 133: 1040–1047.
- [85] ZHANG K, WU Q, LIU A, et al. Can deep learning identify tomato leaf disease? [J]. *Advances in Multimedia*, 2018, 6710865: 1–10.
- [86] BRAHIMI M, BOUKHALFA K, MOUSSAOUI A. Deep learning for tomato diseases: classification and symptoms visualization[J]. *Applied Artificial Intelligence*, 2017, 31(4): 299–315.
- [87] ARGUESO D, PICON A, IRUSTA U, et al. Few-shot learning approach for plant disease classification using images taken in the field[J]. *Computers and Electronics in Agriculture*, 2020, 175: 105542.
- [88] 宋余庆, 谢熹, 刘哲, 等. 基于多层 EESP 深度学习模型的农作物病虫害识别方法[J]. *农业机械学报*, 2020, 51(8): 196–202.
SONG Yuqing, XIE Xi, LIU Zhe, et al. Crop pests and diseases recognition method based on multi-level EESP model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(8): 196–202. (in Chinese)
- [89] PARDEDE H F, SURYAWATI E, SUSTIKA R, et al. Unsupervised convolutional autoencoder-based feature learning for automatic detection of plant diseases[C] // *2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, 2018: 158–162.
- [90] KARTHIK R, HARIHARAN M, ANAND S, et al. Attention embedded residual CNN for disease detection in tomato leaves[J]. *Applied Soft Computing*, 2019, 86: 105933.
- [91] 翟肇裕, 曹益飞, 徐焕良, 等. 农作物病虫害识别关键技术研究综述[J]. *农业机械学报*, 2021, 52(7): 1–18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(7): 1–18. (in Chinese)
- [92] BARBEDO J G A. A new automatic method for disease symptom segmentation in digital photographs of plant leaves[J]. *European Journal of Plant Pathology*, 2017, 147(2): 349–364.
- [93] 李井祝, 王鹏, 耿长兴. 基于图像处理的黄瓜霜霉病情评估[J]. *中国农机化学报*, 2017, 38(6): 67–71.
LI Jingzhu, WANG Peng, GENG Changxing. Disease assessment of cucumber downy mildew based on image processing[J]. *Journal of Chinese Agricultural Mechanization*, 2017, 38(6): 67–71. (in Chinese)
- [94] 鲍文霞, 赵健, 张东彦, 等. 基于椭圆型度量学习的小麦叶部病害识别[J]. *农业机械学报*, 2018, 49(12): 20–26.
BAO Wenxia, ZHAO Jian, ZHANG Dongyan, et al. Recognition of wheat leaf diseases based on elliptic metric learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(12): 20–26. (in Chinese)
- [95] GALLEGOS-SANCHEZ L M, CANALES F J, MONTILLA-BASCON G, et al. RUST: a robust, user-friendly script tool for rapid measurement of rust disease on cereal leaves[J]. *Biosystems Engineering*, 2020, 9(9): 1182.
- [96] WANG G, SUN Y, WANG J X. Automatic image-based plant disease severity estimation using deep learning[J]. *Computational Intelligence and Neuroscience*, 2017, 2017: 2917536.
- [97] LIANG Q, XIANG S, HU Y, et al. PD²SE-Net: computer-assisted plant disease diagnosis and severity estimation network[J]. *Computers Electronics Agriculture*, 2019, 157: 518–529.
- [98] 刘媛, 冯全. 葡萄病害的计算机识别方法[J]. *中国农机化学报*, 2017, 38(4): 99–104.
LIU Yuan, FENG Quan. Identification method of grape diseases based on computer technology[J]. *Journal of Chinese Agricultural Mechanization*, 2017, 38(4): 99–104. (in Chinese)
- [99] RAMCHARAN A, MCCLOSKEY P, BARONOWSKI K, et al. A mobile-based deep learning model for cassava disease diagnosis[J]. *Frontiers in Plant Science*, 2019, 10: 272.

- [100] ESGARIO J G M, KROHLING R A, VENTURA J A. Deep learning for classification and severity estimation of coffee leaf biotic stress[J]. *Computers and Electronics in Agriculture*, 2020, 169: 105162.
- [101] 张领先, 田潇, 李云霞, 等. 可见光光谱和机器学习的温室黄瓜霜霉病严重度定量估算[J]. *光谱学与光谱分析*, 2020, 40(1): 227–232.
ZHANG Lingxian, TIAN Xiao, LI Yunxia, et al. Estimation of disease severity for downy mildew of greenhouse cucumber based on visible spectral and machine learning[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(1): 227–232. (in Chinese)
- [102] LI Kaiyu, ZHANG Lingxian, LI Bo, et al. Attention-optimized DeepLab V3+ for automatic estimation of cucumber disease severity[J]. *Plant Methods*, 2022, 18(1): 1–16.
- [103] WANG C, DU P, WU H, et al. A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net[J]. *Computers and Electronics in Agriculture*, 2021, 189: 106373.
- [104] HOGAN A, BLOMQUIST E, COCHEZ M, et al. Knowledge graphs[J]. *ACM Computing Surveys*, 2022, 54(4): 1–37.
- [105] GUO X, ZHOU H, SU J, et al. Chinese agricultural diseases and pests named entity recognition with multi-scale local context features and self-attention mechanism[J]. *Computers and Electronics in Agriculture*, 2020, 179: 105830.
- [106] ZHANG J, GUO M, GENG Y, et al. Chinese named entity recognition for apple diseases and pests based on character augmentation[J]. *Computers and Electronics in Agriculture*, 2021, 190: 106464.
- [107] GUO X, LU S, TANG Z, et al. CG-ANER: enhanced contextual embeddings and glyph features-based agricultural named entity recognition[J]. *Computers and Electronics in Agriculture*, 2022, 194: 106776.
- [108] ZHANG Z, ZHOU T, ZHANG Y, et al. Attention-based deep residual learning network for entity relation extraction in Chinese EMRs[J]. *BMC Medical Informatics and Decision Making*, 2019, 19(2): 55.
- [109] ZHAO Q, XU D, LI J, et al. Knowledge guided distance supervision for biomedical relation extraction in Chinese Electronic Medical Records[J]. *Expert Systems with Applications*, 2022, 204: 117606.
- [110] YANG Y, RAO Y, YU M, et al. Multi-layer information fusion based on graph convolutional network for knowledge-driven herb recommendation[J]. *Neural Networks*, 2022, 146: 1–10.
- [111] YU G, TABATABAEI M, MEZEI J, et al. Improving chronic disease management for children with knowledge graphs and artificial intelligence[J]. *Expert Systems with Applications*, 2022, 201: 117026.
- [112] ZHOU Gengxian, E Haihong, KUANG Zemin, et al. Clinical decision support system for hypertension medication based on knowledge graph[J]. *Computer Methods and Programs in Biomedicine*, 2022, 227: 107220.
- [113] WANG N, CAI X, YANG L, et al. Safe medicine recommendation via star interactive enhanced-based transformer model[J]. *Computers in Biology and Medicine*, 2022, 141: 105159.
- [114] VENKATESH R, BALASUBRAMANIAN C, KALIAPPAN M. Development of big data predictive analytics model for disease prediction using machine learning technique[J]. *Journal of Medical Systems*, 2019, 43(8): 1–8.
- [115] WANG M, LU T, HUANG W, et al. Intelligent prescription system combined with a national pharmacloud for geriatrics care[J]. *Expert Review of Pharmacoeconomics & Outcomes Research*, 2018, 18(5): 559–564.
- [116] GALVEIA J A T, CRUZ L. An ophthalmology clinical decision support system based on clinical annotations, ontologies and images[C]//2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 2018: 94–99.
- [117] SAGI O, ROKACH L. Ensemble learning: a survey[J]. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018, 8(4): e1249.
- [118] MATLOCK K, DE NIZ C, RAHMAN R, et al. Investigation of model stacking for drug sensitivity prediction[J]. *BMC Bioinformatics*, 2018, 19(Suppl. 3): 71.
- [119] DONG X, YU Z, CAO W, et al. A survey on ensemble learning[J]. *Frontiers of Computer Science*, 2020, 14(2): 241–258.
- [120] AGARWAL S, CHOWDARY C R. A-stacking and a-bagging: adaptive versions of ensemble learning algorithms for spoof fingerprint detection[J]. *Expert Systems with Applications*, 2020, 146: 113160.
- [121] HONG H, LIU J, ZHU A. Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble[J]. *Science of the Total Environment*, 2020, 718: 137231.
- [122] ZIAN S, KAREEM S A, VARATHAN K D. An empirical evaluation of stacked ensembles with different meta-learners in imbalanced classification[J]. *IEEE Access*, 2021, 9: 87434–87452.
- [123] NEJATIAN S, PARVIN H, FARAJI E. Using sub-sampling and ensemble clustering techniques to improve performance of imbalanced classification[J]. *Neurocomputing*, 2018, 276: 55–66.
- [124] SEGURA-BEDMAR I, COLÓN-RUÍZ C, TEJEDOR-ALONSO M Á, et al. Predicting of anaphylaxis in big data EMR by exploring machine learning approaches[J]. *Journal of Biomedical Informatics*, 2018, 87: 50–59.
- [125] SONG K, ZENG X, ZHANG Y, et al. An interpretable knowledge-based decision support system and its applications in pregnancy diagnosis[J]. *Knowledge-Based Systems*, 2021, 221: 106835.
- [126] LENG J, WANG D, MA X, et al. Bi-level artificial intelligence model for risk classification of acute respiratory diseases based on Chinese clinical data[J]. *Applied Intelligence*, 2022, 52(11): 13114–13131.
- [127] ZHANG J, ZHANG X, SUN K, et al. Unsupervised annotation of phenotypic abnormalities via semantic latent representations on electronic health records[C]//2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2019: 598–603.
- [128] 程铭, 熊蜀峰, 李霏, 等. 基于混合注意力机制的脑卒中康复方案推荐[J]. *武汉大学学报(理学版)*, 2021, 67(6):

569–577.

CHENG Ming, XIONG Shufeng, LI Fei, et al. Rehabilitation program recommendation for stroke based on hybrid attention mechanism[J]. Journal of Wuhan University (Natural Science Edition), 2021, 67(6): 569–577. (in Chinese)

- [129] QIU X P, SUN T X, XU Y G, et al. Pre-trained models for natural language processing: a survey[J]. Science China Technological Sciences, 2020, 63(10): 1872–1897.
- [130] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C] // Conference of the North-American-Chapter of the Association-for-Computational-Linguistics-Human Language Technologies, arxiv:1810.04805, 2018.
- [131] WANG J, ZHANG X, CHEN L. How well do pre-trained contextual language representations recommend labels for gitHub issues? [J]. Knowledge-Based Systems, 2021, 232: 107476.
- [132] DING J, LI B, XU C, et al. Diagnosing crop diseases based on domain-adaptive pre-training bert of electronic medical records[J/OL]. Applied Intelligence, 2022. <http://doi.org/10.1007/s10489-022-04346-x>.
- [133] 杨森, 冯全, 张建华, 等. 基于深度学习与复合字典的马铃薯病害识别方法[J]. 农业机械学报, 2020, 51(7): 22–29.
YANG Sen, FENG Quan, ZHANG Jianhua, et al. Identification method for potato disease based on deep learning and composite dictionary[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 22–29. (in Chinese)
- [134] 孔建磊, 金学波, 陶冶, 等. 基于多流高斯概率融合网络的病虫害细粒度识别[J]. 农业工程学报, 2020, 36(13): 148–157.
KONG Jianlei, JIN Xuebo, TAO Zhi, et al. Fine-grained recognition of diseases and pests based on multi-stream Gaussian probability fusion network[J]. Transactions of the CSAE, 2020, 36(13): 148–157. (in Chinese)
- [135] ZHOU Y, CHANG F, CHANG L, et al. Multi-output support vector machine for regional multi-step-ahead PM_{2.5} forecasting [J]. Science of the Total Environment, 2019, 651: 230–240.
- [136] 应启帆, 谢代梁, 徐志鹏, 等. 多输出回归算法在超声粒径分布反演中的研究[J]. 声学技术, 2022, 41(1): 137–143.
YING Qifan, XIE Dailiang, XU Zhipeng, et al. Research on multiple output regression algorithm for inversion of ultrasonic particle size distribution[J]. Technical Acoustics, 2022, 41(1): 137–143. (in Chinese)
- [137] ZHANG Y, HE Z, YANG J, et al. Re-structuring and specific similarity computation of electronic medical records[C] // 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), 2017: 230–235.
- [138] 赵明, 董翠翠, 董乔雪, 等. 基于 BIGRU 的番茄病虫害问答系统问句分类研究[J]. 农业机械学报, 2018, 49(5): 271–276.
ZHAO Ming, DONG Cuicui, DONG Qiaoxue, et al. Question classification of tomato pests and diseases question answering system based on BIGRU[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 271–276. (in Chinese)
- [139] YE C, FABBRI D. Extracting similar terms from multiple EMR-based semantic embeddings to support chart reviews[J]. Journal of Biomedical Informatics, 2018, 83: 63–72.
- [140] 邱硕, 刘佳欣, 李启康, 等. 数据聚合与共享技术在电子医疗病历中的应用[J]. 南京理工大学学报, 2021, 45(6): 672–679.
QIU Shuo, LIU Jiaxin, LI Qikang, et al. Application of data aggregation and sharing technology in electronic medical records [J]. Journal of Nanjing University of Science and Technology, 2021, 45(6): 672–679. (in Chinese)
- [141] XIE Z, ZENG Z, ZHOU G, et al. Topic enhanced deep structured semantic models for knowledge base question answering [J]. Science China Information Sciences, 2017, 60(11): 1–15.
- [142] LARIONOVA A, KAZAKOVA P, NIKITINSKY N. Deep structured semantic model for recommendations in e-commerce[C] // International Conference on Hybrid Artificial Intelligence Systems, 2019: 85–96.
- [143] REIMERS N, GUREVYCH I. Sentence-bert: sentence embeddings using siamese bert-networks[J]. arXiv Preprint, arXiv: 1908.10084, 2019: 3982–3992.
- [144] GAO T Y, YAO X C, CHEN D Q. SimCSE: simple contrastive learning of sentence embeddings[C] // EMNLP 2021—2021 Conference on Empirical Methods in Natural Language Processing, 2021: 6894–6910.
- [145] 袁博, 施运梅, 张乐. 基于知识图谱的问答系统研究与应用[J]. 计算机技术与发展, 2021, 31(10): 134–140.
YUAN Bo, SHI Yunmei, ZHANG Le. Research and application of question answering system based on knowledge graph[J]. Computer Technology and Development, 2021, 31(10): 134–140. (in Chinese)
- [146] 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(Sup.): 164–171. (in Chinese)
- [147] 张领先, 赵聘桐, 丁俊琦, 等. 基于 CDSSM 的作物病害处方推荐方法[J]. 农业机械学报, 2023, 54(3): 308–317.
ZHANG Lingxian, ZHAO Dantong, DING Junqi, et al. Recommendation method of crop disease prescription based on CDSSM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 308–317. (in Chinese)
- [148] MAITINIAZI M, GHULAM A, SIDIKE P, et al. Unmanned aerial system (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 134: 43–58.
- [149] MAIMAITIJIAN M, SAGAN V, SIDIKE P, et al. Soybean yield prediction from UAV using multimodal data fusion and deep learning[J]. Remote Sensing of Environment, 2019, 237: 111599.