

• AiC 2026 • 算法创新赛道 — AI + 软件创新

深瞳蔬检

基于视觉小模型与 ReAct 智能体的蔬菜病害诊断与防治决策系统

• 核心作物：马铃薯·番茄（已上线专有模型）

• 主场景：设施蔬菜大棚病害诊断防治

让农户拍一张照片，就得到一份
「看得准、开得对、信得过」的防治方案



番茄叶片这是什么病？

17:42

🔍 专有模型识别：番茄晚疫病 · 置信度 96%

📖 知识库：番茄晚疫病防治方案 等3篇

- ① 诊断：番茄晚疫病（中度发生）。
- ② 防治（绿色防控版）：
- 1 控湿排湿：通风降湿、避免叶面结露；
 - 2 清除病株：拔除深埋，病穴撒石灰；
 - 3 保护预防：雨前喷代森锰锌 150-200 g/亩；
 - 4 治疗用药：氟菌·霜霉威 60-75 mL/亩；
- ⚠️ 预防为主、发现即治，采收前注意停药。

👨‍⚕️ 一键转专家会诊

输入文字描述问题，或点击 + 拍照识别病虫害...

CONTENTS

四步讲清一件事

从产业里的真问题出发，讲我们怎么把它做成一个能跑、能测、能持续变准的软件产品。

01

02

03

04

需求与痛点 P04 – P05

01

15.7 万个大棚对 2.6 万名持证农技员；通用大模型「看得懂图、认不准病」。



方案与架构 P07 – P08

02

三端一体五步闭环；小程序前端、后端接口、模型服务、知识库、数据分层解耦。



核心技术与创新 P10 – P12

03

专有小模型真实验证集指标；三路混合 RAG；智能体工具调用做到一地一策。



工程落地与成效 P14 – P17

04

11 个端到端脚本全通过；样本反哺通道；成效、短板与可考核路线。



CHAPTER 01

需求与痛点

AI 要解决的，首先是产业里真实存在的矛盾。

三重矛盾

病害频发 · 需求量大 · 力量不足

通用模型短板

看得懂图，认不准病

15.7^{万个}

蔬菜大棚

60^{万亩}

设施蔬菜种植面积

450^{万吨}

蔬菜年产量

12^{万人}

设施蔬菜从业人员

诊断需求侧

15.7 万个大棚的日常诊断

VS

专业力量侧

2.6 万名持证农民技术员

棚内高温高湿使晚疫病、灰霉病、细菌性斑点病易发多发——约 6 个大棚才能摊到 1 名持证农技员。

「看病难、看病慢、看病不准」是制约产业提质增效的现实痛点。

场景：农户拍下一张番茄叶照片——通用模型与专有模型的回答差别在哪？

用户痛点

通用大模型认不准病

能说出「叶片有褐色病斑」，但难以区分番茄早疫病、晚疫病与细菌性斑点病这类同表型病害；未接入植保知识库时，防治建议只能凭语言先验拼凑，剂量模糊。

解决方式

专有模型 + 专业知识库

马铃薯 / 番茄专有病害模型只认训练过的类，mAP@0.5 均为 99.50%（权重文件验证集指标）；认不出时诚实提示「存疑」，不编造病名。知识库负责给出有依据的防治方案。

软件价值

一地一策的可信方案

结合地块品种、生育期、棚内小气候与农事记录，输出带剂量、安全间隔期的方案；禁用农药自动拦截，建议「可用、可信、可控」。

植保诊断必须「专用模型 + 专业知识」双轮驱动——小模型看得准，知识库开得对。

CHAPTER 02

方案与架构

不是一个算法脚本，而是三端一体、分层解耦、可持续演进的软件系统。

三端一体

小程序 / 电脑端 / 后端

四层体系

小模型 + 知识库 + 场景数据 + 专家协同

五步首尾相接，最后一环回灌第一环——软件用得越久，模型越准

农户 微信小程序

农技专家 + 管理员 电脑端 Vue3

后端 FastAPI · 157 个接口

1 拍照识别

专有小模型给出病名、置信度与病灶框

2 智能问答

智能体多轮调用工具检索知识库

3 方案生成

结合地块农事与棚内小气候，一地一策

4 专家会诊

图片 + AI 结论 + 地块数据打包成工单

5 样本反哺

专家校正标注，导出 YOLO 训练数据集

第 5 环回灌第 1 环



标注样本进入样本库 → 触发增量训练任务 → 新版本模型上线切换 → 识别准确率与覆盖类别同步更新
(导出训练数据集接口已就绪，真实增量训练尚未执行，见 P15)

接口就绪

✓ 完整的业务闭环，让项目从「单点算法」变成「可持续运营的软件产品」。

前端、接口、模型、知识库、数据分层解耦——每层可独立替换，接口契约不变

前端层 CLIENT	微信小程序（农户）	电脑端 Vue3（专家 / 管理员）
接口层 API	157 个接口路径	JWT + RBAC 三角色
	审计日志 682 条	
模型服务层 MODEL	专有小模型识别	智能体工具调用（SSE）
	三路混合 RAG	禁用农药拦截
知识库层 KNOWLEDGE	35 份病害档案	55 套防控方案
	ChromaDB 向量库	16 种合规农药
数据层 DATA	SQLite 业务库	传感读数 276 条
	地块 / 农事记录	样本库 37 份

外部服务

DeepSeek 大模型 API（Provider 可在后台切换，密钥经 .env 配置）

已接入

核心业务流程

- 1

用户提交图片或问题

小程序拍照 / 文字提问
- 2

作物模型识别

按作物路由到专有小模型
- 3

知识检索与安全检查

三路 RAG + 禁用农药拦截
- 4

生成诊治建议

结合地块上下文，一地一策
- 5

历史记录或反馈

诊治记录入档，可转专家会诊

接口有契约、状态机有断言、权限有审计——工程质量可以被逐条验证。

CHAPTER 03

核心技术与创新

三项技术选择，分别解决「看得准」「开得对」「用得活」。

专有小模型

真实验证集指标

三路混合 RAG

有效 + 安全

智能体工具调用

一地一策

马铃薯 / 番茄 2 个专有 YOLO 模型已上线（指标读目权重 checkpoint）；茄子暂以通用检测 + 知识库推演支撑，非病害专用模型

mAP@0.5 验证集指标 (checkpoint 未轮)

99.50 %

马铃薯 / 番茄一致（权重 checkpoint 未轮验证）

Precision

99.78 %

马铃薯 99.78% · 番茄 99.87%

Recall

100.0 %

两个模型一致

CPU 推理延迟 (实测)

47 / 433

ms · 512px ms · 640px

本机

✓ 已验证

权重文件 checkpoint 内 train_metrics (split=val)；CPU 实测延迟

指标 / 模型

马铃薯

番茄

骨干网络

YOLOv8 自定义

YOLO11n 预训练

类别数

3 类

10 类

输入尺寸

640

512

Precision

99.78%

99.87%

Recall

100.0%

100.0%

mAP@0.5

99.50%

99.50%

mAP@0.5:0.95

97.07%

94.83%

训练配置

110ep·bs8·SGD

100ep·bs8·Adam

🕒 待补充

基线对照实验、独立测试集、消融实验——暂不声称相对基线有提升

⚠ 待核实

参考文档「平均准确率 94.1%」「样本 10,800 张」口径不明，不作核心卖点

推理流程

- 1 小程序压缩上传
> 1280px 压缩
- 2 YOLO 推理
按作物路由模型
- 3 NMS + 置信度过滤
conf 0.25 / IoU 0.45
- 4 类别映射病害档案
映射不上则回落知识库推演

mAP@0.5 为 IoU=0.5 下的平均精度，不等同于「识别准确率」；指标来自权重 checkpoint 未轮验证，尚无论文级独立测试集。

大模型负责「说人话」，知识库负责「说得对」——检索增强是方案可信的前提



三路召回及当前权重配置 当前配置

- 0.45 中文 2-gram 词面匹配**
缓解中文无空格分词与同义表述差异
- 0.45 向量语义召回**
ChromaDB 内嵌向量，捕捉语义相近的提问
- 0.10 名称命中**
病害名直接命中时优先提升排序

以上权重为当前工程配置，尚未通过检索效果实验验证为最优组合。

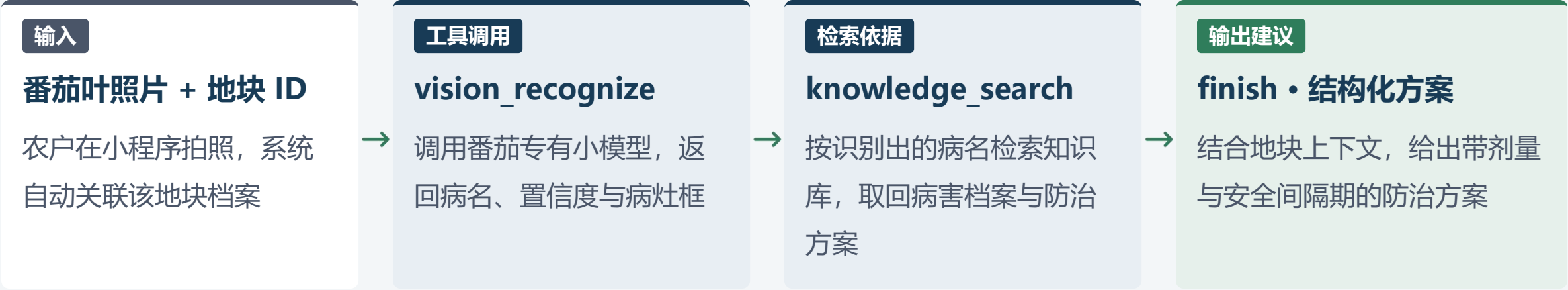
安全规则检查（独立环节）

命中禁用 / 限用农药的方案条目被直接拦截；专家撰写/审核 → 管理员发布，两级把关。

安全规则是风险控制手段之一，不等同于对全部用药场景的绝对安全保证；知识库内容持续更新。

知识是权威的，红线是硬编码的——每一条建议都「可用、可信、可控」。

按「调用工具 → 读取结果 → 再决定下一步」循环推进，而不是一次性黑箱输出



📍 诊断时自动带入的地块上下文字段（均来自项目数据库）

地块与品种 大棚编号 · 作物品种	生育期 定植 / 开花 / 采收	棚内小气候 温湿度 · 276 条读数	农事操作 施肥 · 用药 · 整枝	历史诊治 该地块此前的方案	天气预报 未来 7 天预报
-----------------------------	----------------------------	-------------------------------	-----------------------------	-------------------------	-------------------------

小程序展示的是**工具调用状态与检索过程**，不是模型完整内部推理；同一病害在不同地块给出不同用药时机与剂量。

CHAPTER 04

工程落地与成效

能不能跑通、能不能测住、能不能持续变准——这是「AI+软件」与「算法 Demo」的分界线。

测试验证

11 个脚本全通过

样本反哺

通道已打通

落地成效

三项分级

测试概况
OVERVIEW

11 个端到端验证脚本
tests/ 目录

157 个接口路径
/openapi.json 实测

205 个接口方法
按 @router 统计

15 个业务模块
app/api/v1/

二、测试覆盖 脚本名与实际项目目录一致

覆盖域	验证脚本
智能体与诊治	verify_agent_react.py · _smoke_plot_advice.py
专有模型接入	verify_vision_models.py
工单状态机	verify_ticket_sync.py · _smoke_ticket_chat.py
农场与地块	verify_farm_manage.py
样本库与导出	verify_sample_library.py · verify_ticket_sample_oneclick.py
AI 自动回复	_smoke_auto_reply.py
三角色接口	smoke_test.py

三、执行结果 已验证

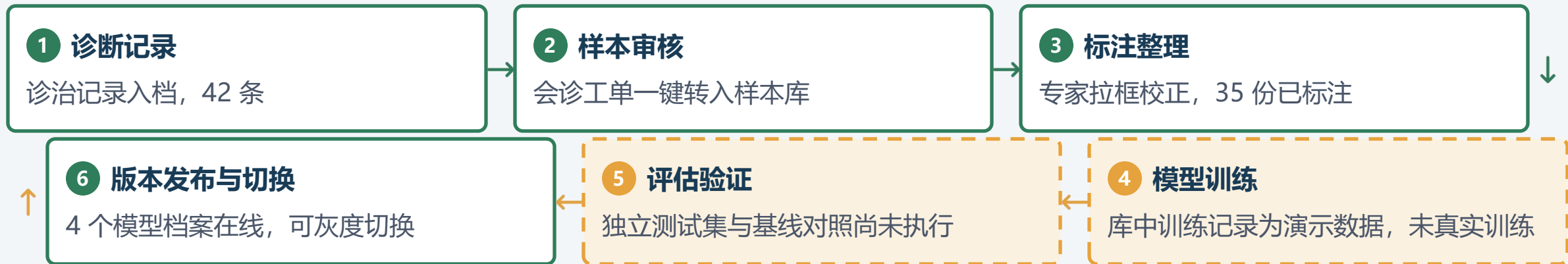
- ✓ verify_ticket_sync.py ALL_PASS
工单状态机 pending→closed 全链路
 - ✓ verify_farm_manage.py ALL_PASS
农场 / 分区 / 地块 / 农事全链路
- 其余脚本按目录内既有执行记录标注通过状态；复跑命令见 README「tests/」说明。

四、证据来源（可追溯）

- tests/ 目录
11 个验证脚本本体
- Swagger 接口文档
启动后 /docs 可查全部接口
- 接口数统计口径
205 个方法 / 去重 157 个路径

✓ 接口有契约、状态机有断言、权限有审计——工程质量可以被逐条验证，而非口头承诺。

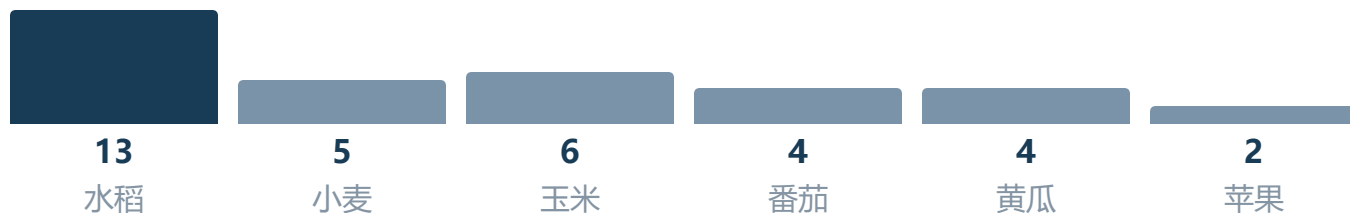
实线为已实现环节，虚线为尚未真正执行的环节——我们如实标注这条链路的真实进度



37 份样本在库 (35 份已标注)

来源: 诊断 18 · 田间 14 · 会诊工单 5

样本库作物分布



需正视的差距

样本库以水稻、小麦、玉米为主, 尚无马铃薯、番茄样本, 与已上线 (马铃薯 / 番茄) 专有模型的作物不匹配; 真实增量训练尚未执行。

闭环的「软件能力」已经跑通, 缺的是按已上线作物补采的真实标注数据。

改变的不是「有没有 AI」，而是农户获取专业植保服务的方式

识别能力有据可查

已验证

2 个专有模型上线运行，mAP@0.5 均为 99.50%（权重文件验证集指标）；CPU 实测延迟 47 / 433 ms。

服务可及性提高

当前能力

三端齐备，农户在棚口即可获得识别结果与防治方案；疑难病害一键转专家会诊兜底。

绿色防控与减药增效

未来规划

55 套绿色防控方案 + 16 种合规农药已入库，禁用农药自动拦截——安全机制已就绪，规模成效待真实使用数据验证。



输入（真实）：番茄叶片照片



输出（真实）：识别结果 + 知识库命中 3 篇防治方案

AI 把「专家经验」变成农户手机里可复用的服务——这是算法工程化落地的价值。

把问题说清楚，把路线定下来——每条短板都对应一个可考核的动作

当前限制

- ① 覆盖作物有限**
已上线 2 个作物专有模型；黄瓜、辣椒等暂无权重。
- ② 缺基线与消融实验**
无法证明相对基线有提升，指标口径也需统一。
- ③ 样本与模型不匹配**
样本库以水稻 / 小麦 / 玉米为主，无马铃薯、茄子样本。
- ④ 物联与数据底座薄**
设备接入覆盖不足，「数据孤岛」制约决策深度。
- ⑤ 经费与运维压力**
训练算力、大模型 API、服务器与知识更新需持续投入。

下一步工作

- ① 补基线实验** 最高优先
同数据划分重训，输出 P / R / mAP 对照表与独立测试集结果
- ② 补茄子权重** 高
替换现有通用检测权重为真实茄子病害权重并出评估报告
- ③ 补样本数据** 高
按已上线作物补采标注，让反哺闭环有真实数据支撑
- ④ 扩作物覆盖**
分步推出黄瓜、辣椒等专有模型
- ⑤ 接物联数据**
对接智能大棚环境数据，做「农事—环境—病害」关联分析

我们不做「万能 AI」的承诺，而是把**准确率、覆盖率、证据链**写成可以逐年考核的指标。

• AiC 2026 · 算法创新赛道 — AI + 软件创新

深瞳蔬检

让 AI 成为农户身边「会看病的植保助手」



恳请各位评委批评指正