

香蕉成熟度 在线检测系统

连续评估与批次分拣

项目答辩 / 2026 年 10 月

0.9525

mAP@50
智视先锋队

0.9295

指数 ρ

53.6 ms

本地处理均值 / CPU

面向分拣与周转的软件产品

目标用户

果品分拣人员 批次管理与仓储人员

使用普通图像输入，辅助判断成熟度、
安排出库优先序并交接批次报告。

判断

逐目标类别、连续指数与待复核标记

分流

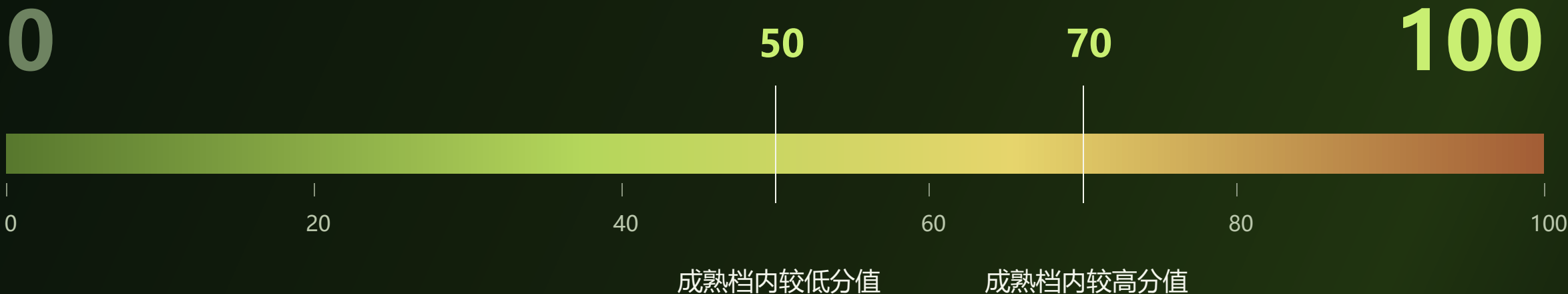
按等级与估计货架期组织批次

交付

生成 PDF，保留检测记录与导出结果

同一成熟档，也需要出库优先序

类别给出等级，连续指数补充档内分值。



成熟程度

0-100 连续指数

周转时间

仓温相关的货架期估计

批次处置

等级结构与出库优先序

软件差异化设计与可交付结果

用户任务	仅输出类别的功能基线	本系统的扩展
同档目标如何排序	显示未熟、成熟或过熟	连续成熟度指数与目标明细
低把握结果如何处理	显示模型原始分值	校准分值与待复核标记
一批货如何安排周转	需另行整理单图结果	批次汇总与出库优先序
结果如何交接留存	逐张查看检测图	PDF 报告、历史查询与导出

产品贡献 把分散的检测结果组织为可复核、可汇总、可交接的业务流程。

完整操作流程与人员分工

01	选择输入	图片、视频或摄像头；批次图片单批最多 60 张	操作人员
02	运行检测	模型给出类别、指数、分值与待复核标记	软件
03	复核异常	人工检查低把握与边界目标，再作业务处置	操作人员
04	形成批次建议	汇总等级、估计货架期、估算货值与优先序	软件
05	交接与追溯	导出 PDF 或数据，查询检测历史与明细	操作人员

系统架构与数据流

交互层

图片 / 视频 / 摄像头

检测工作台 批次分拣 统计与历史 知识与后台

服务层

Django 请求与业务处理

输入校验 权限控制 批次聚合 PDF 与数据导出

算法层

YOLO26 与后处理

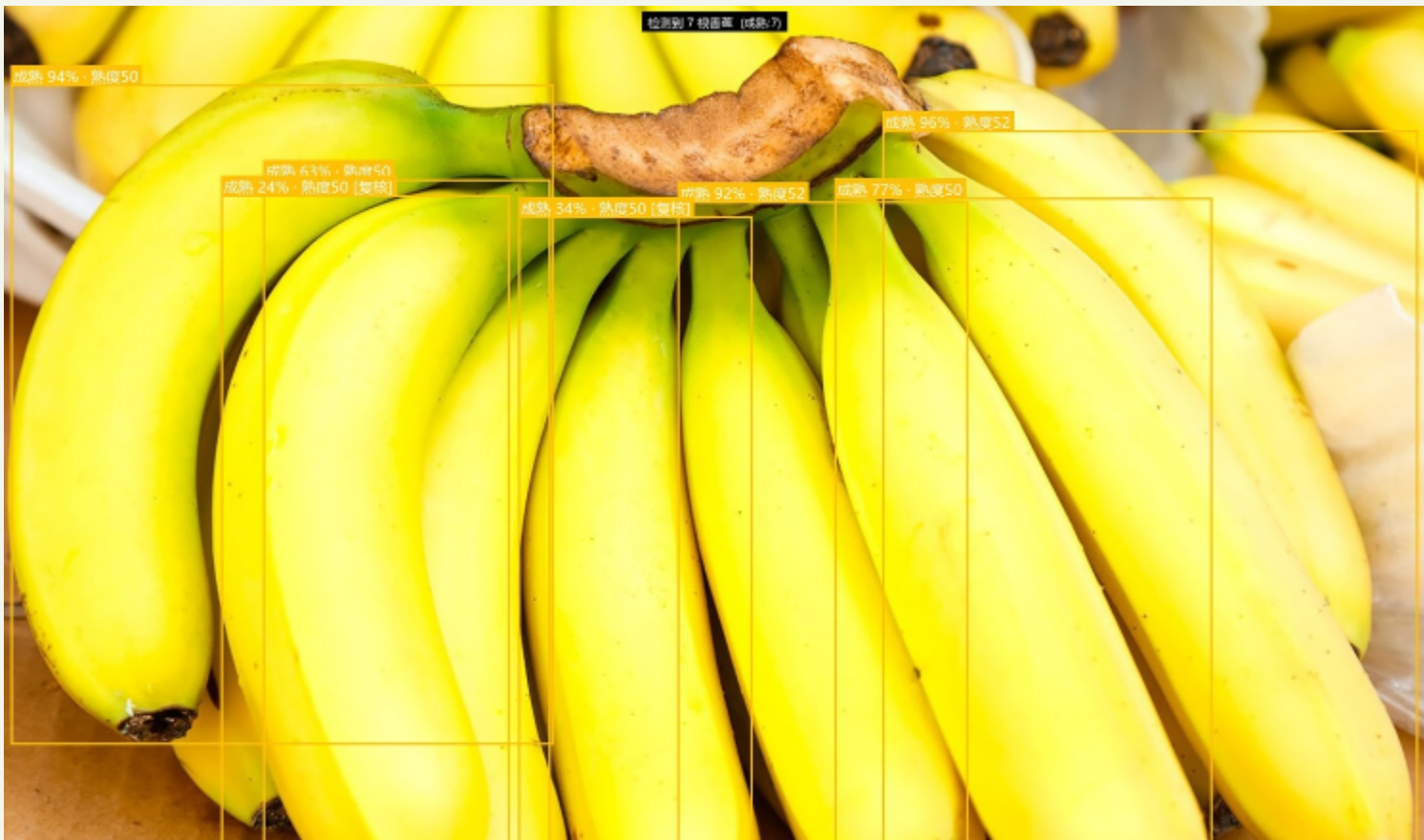
类别检测 颜色融合 置信度校准 ByteTrack

数据层

结果、模型与评估文件

检测记录 模型权重 知识文章 reports/*.json

检测工作台与逐目标输出



01 上传与检测

图片、视频与摄像头
统一进入后端推理流程

02 逐目标解释

类别与校准分值
成熟度指数、货架期
五级评级与处置建议

03 异常反馈

推理失败时明确提示

一批图片，一份可交接的报告



单批最多 60 张图片

设置仓温、单价与单果重量
汇总目标状态与批次建议

报告结构	输出内容
分级	等级数量与占比
周转	估计货架期与优先序
估值	规则损耗与估算货值
交接	逐图明细与处置建议

PDF 报告用于交接与留存

检测数据汇总，服务日常查看



统计来源明确

从检测记录聚合次数、目标数、平均分值与推理耗时。

查看批次状态

结合近 14 天趋势、档位结构与最近检测明细查看变化。

与评估指标区分

这里展示业务记录。检测精度见固定测试集评估。

检测留痕与运行追溯



检测结果有记录

保存检测来源、模型与时间，以及类别统计、耗时和明细。

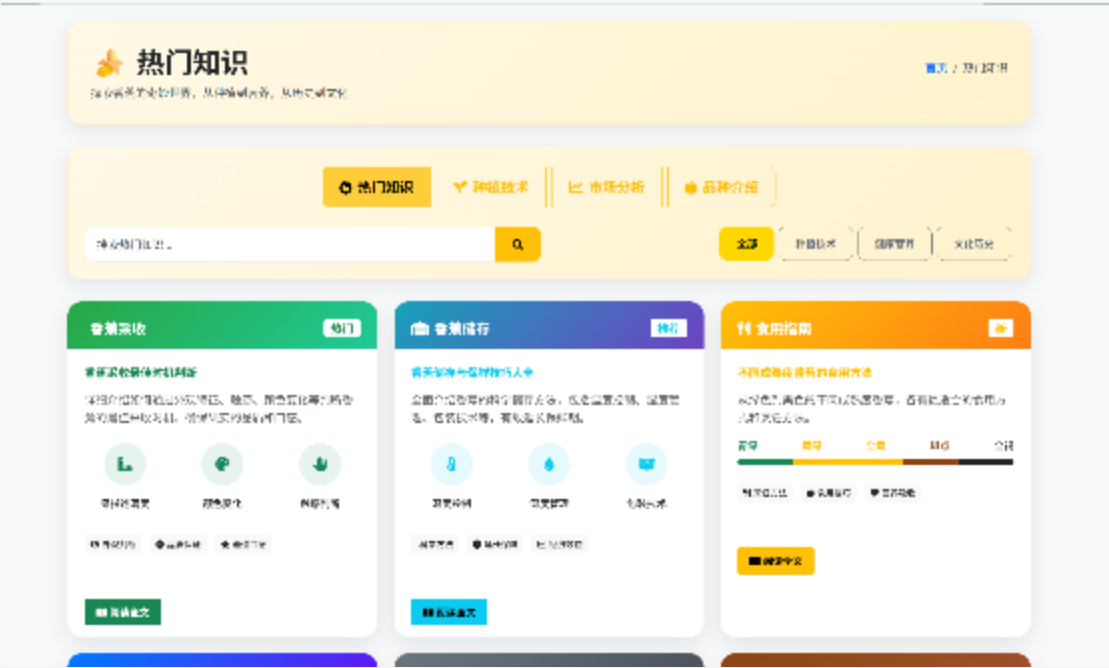
评估结果有来源

指标页读取评估文件，评估脚本与结果相互对应。

交付结果可留存

批次 PDF 用于交接，检测结果可按接口导出。

知识服务与后台管理



知识资料查询

热门知识、种植技术、市场分析与品种介绍
支持分类浏览和关键词搜索。



后台日常维护

提供统计、日志、设置等管理入口。
重点展示有记录支撑的统计与追溯。

三个标签，补充连续成熟度刻度

模型确定类别区间，果皮颜色补充档内分值。



未熟 0–36

成熟 38–78

过熟 72–94

$$I = \text{clip}(0.35 \times P + 0.65 \times C, L + 1, U - 1)$$

P: 类别先验 C: 色相与暗斑得分 [L, U]: 类别区间

19.4 ± 4.5

62.0 ± 7.3

84.7 ± 7.5

未熟 / 51 个目标

成熟 / 65 个目标

过熟 / 84 个目标

同一测试划分下的权重表现

0.9525

主模型 mAP@50

三分类输出与业务档位对应。
旧权重按三档归并后进行评估。



每组左点为旧权重，右点为主模型

主模型成熟类 Precision 为 0.7857，仍需补充相邻阶段与新增场景验证。

软件测试与现有验证证据

验证内容	覆盖范围	目前证据
检测精度	218 张固定测试图	已保存评估结果
成像扰动	7 类输入扰动	已保存评估结果
本地运行性能	120 张图，CPU	已保存延迟基准
货架期计算规则	12 项锚点与边界核验	规则核验 12/12

交付前回归范围

图片与视频检测、批次报告、历史导出、权限与异常输入。

当前验收边界

完整 HTTP 回归、浏览器兼容性和并发压测尚待形成通过报告。

CPU 运行延迟与部署条件

51.0 ms

平均模型推理

53.6 ms

本地处理平均耗时

19.6

推理吞吐量 / 帧每秒

长尾延迟

推理 P95: 57.4 ms

推理 P99: 65.4 ms

适用条件

当前结果支持本地交互演示。
高速产线与多路并发需另行压测。

本地部署路径与运行条件

浏览器访问本机服务

图像与模型在本机处理
CPU 可运行，数据库可选 SQLite
模型、页面资源与报告文件随项目管理

关键运行条件

匹配的依赖环境与模型权重
可写的上传、结果与报告目录

01 安装依赖

按 requirements.txt 准备运行环境

02 准备模型与数据

放置主模型，设置 DB_ENGINE=sqlite

03 初始化并启动

执行数据库迁移，启动本地服务

04 检查业务页面

检测、批次分拣、模型指标与报告

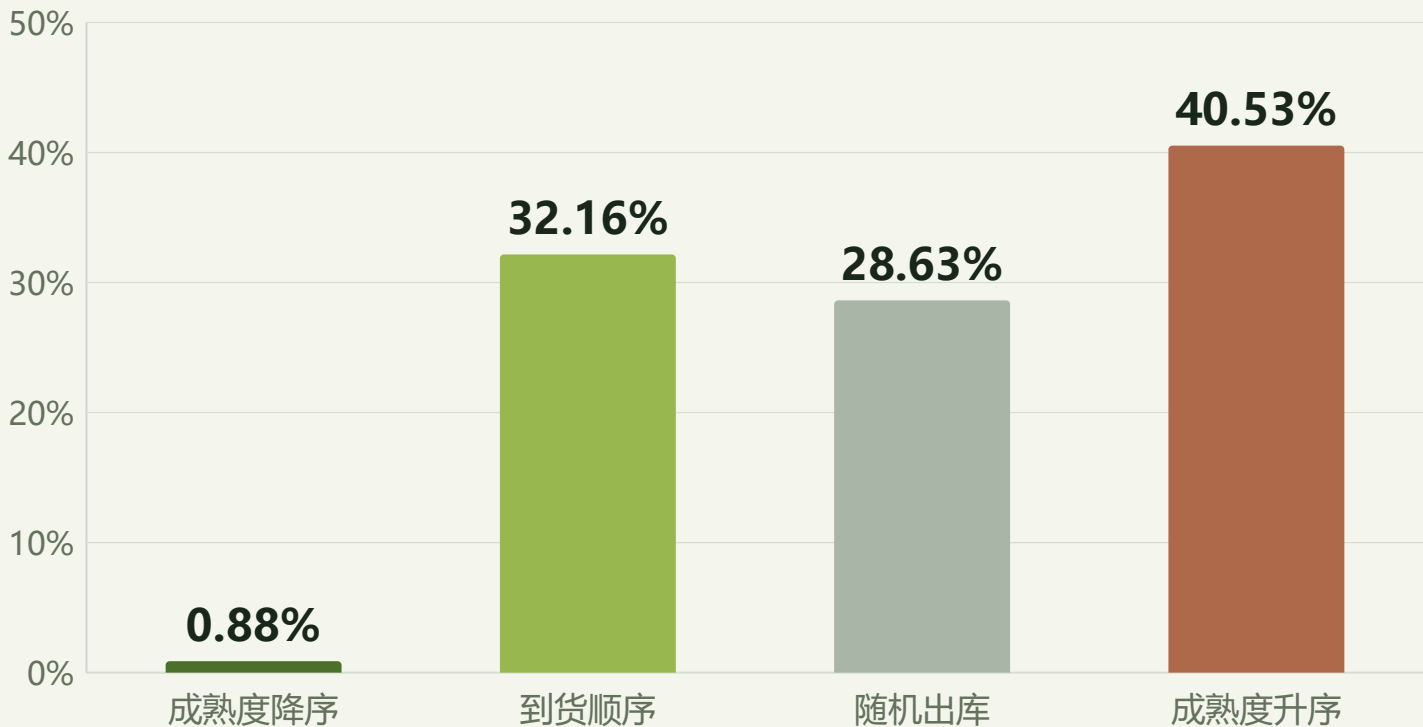
出库顺序的策略仿真

仿真条件

20℃ 恒温
7 天，无补货
每日出库能力为初始批量的 20%
成熟度达到 100 视为损失

0.88%

成熟度降序策略的仿真损耗率



该结果展示规则模型下的策略差异，真实损耗收益仍需现场对照验证。

每项局限，对应一个验证动作

当前问题	已有证据与限制	下一步验证动作
跨场景适用性	同源划分与输入扰动测试	冻结参数，新增封存场景测试
连续分值精度	与人工三档序数相关	分级员盲评，检验档内排序
货架期与降损	规则模型与出库策略仿真	真实储藏跟踪与批次对照
软件交付稳定性	已有功能与部分评估结果	完整回归、兼容性与并发测试
实际使用价值	尚缺正式用户试用证据	果店或仓储试用，记录任务表现

试点记录内容 **完成时间、人工复核数量、操作问题与真实用户反馈。**

项目成果与核心贡献

01 类别与连续指数

为目标提供档位与档内分值

$\rho = 0.9295$, 仅对照人工三档序数

02 分值与人工复核

显式标记低把握与边界目标

校准改善有限, 结果仍须人工判断

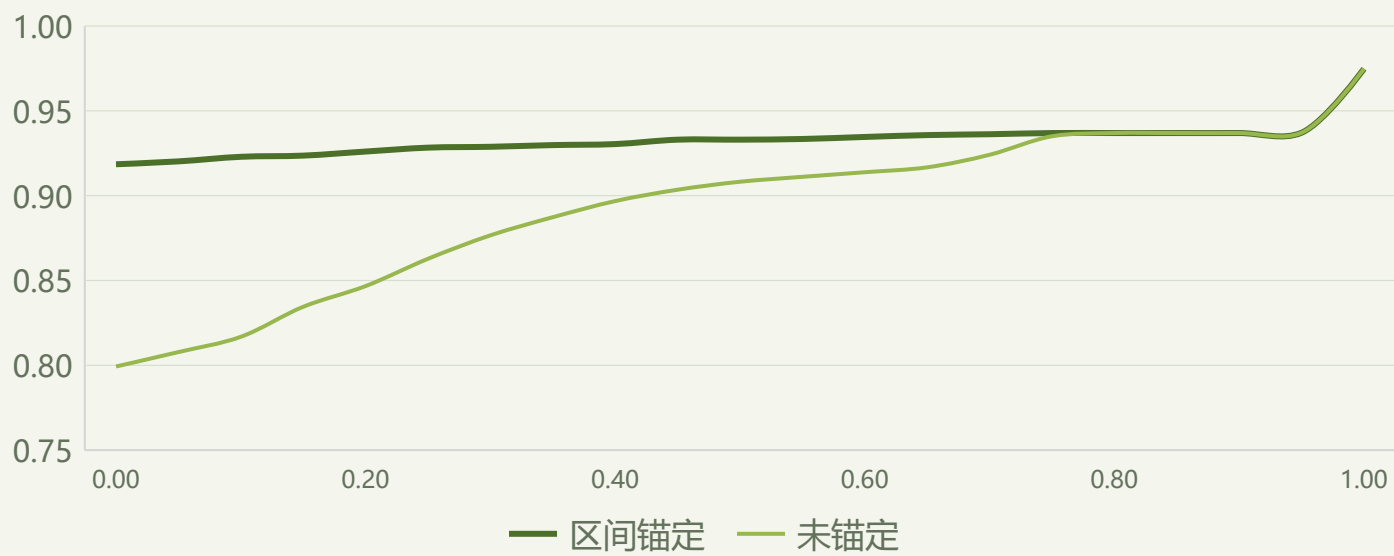
03 检测与批次交付

汇总等级、周转建议与报告

货架期和货值按模型与业务规则估计

融合权重与区间锚定的作用

类别先验权重 w 与 Spearman ρ



0.0109

w 在 0.2–0.8 内的 ρ 全距

单靠三档先验可取得高相关，但仅有 4 个分值，无法体现同档内的连续变化。

方法	ρ	不同分值数
仅颜色	0.7993	144
仅类别先验	0.9748	4
融合并锚定	0.9298	124

置信度校准与人工复核

温度缩放参数

$$T = 0.856$$

在验证集上拟合，再对测试集复算。
目标是改善分值与正确频率的一致性。

$$p_{cal} = \text{sigmoid}(\text{logit}(p) / T)$$

校准改善有限，分值仍是估计，不能保证单次判断正确。

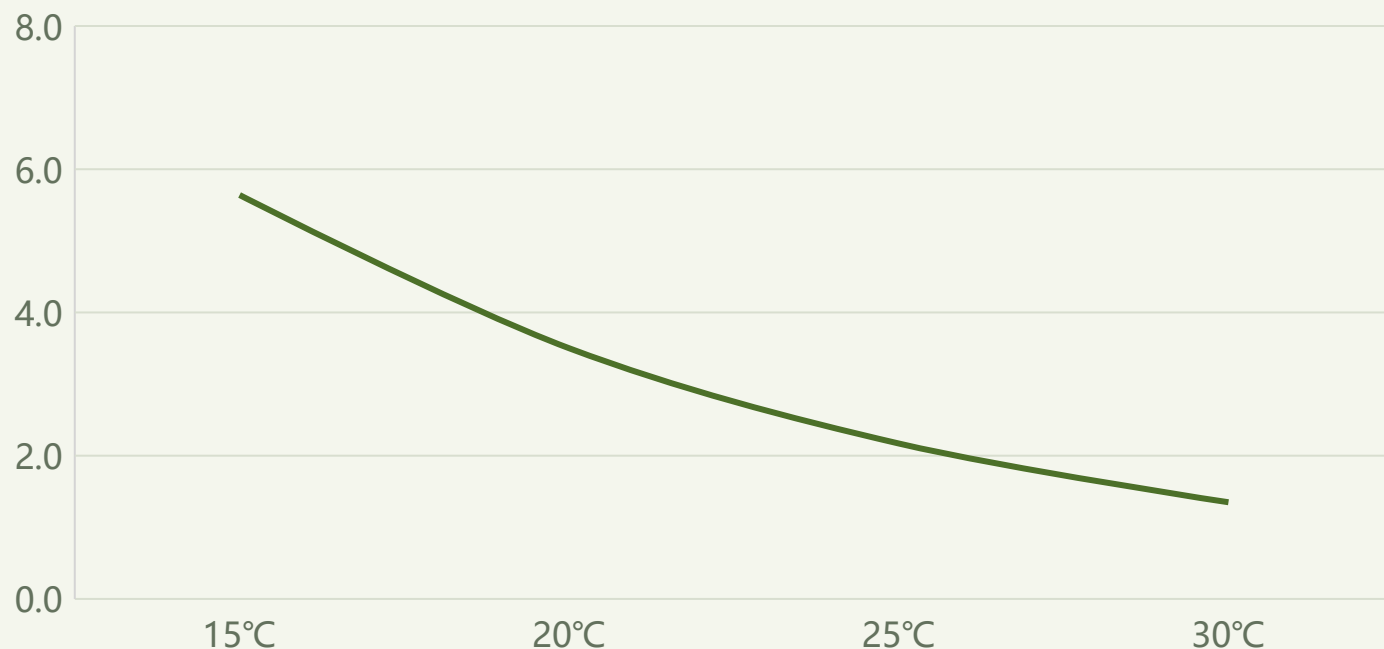
ECE，越低越好	校准前	校准后
验证集	0.0613	0.0567
测试集复算	0.0726	0.0718

低把握目标标记待人工复核

校准分值低于 0.45，或成熟度指数距档位边界不足 3 分，标记待复核。

温度自适应货架期估计

$$L(T) = L_{20} \times 2.6^{((20 - T) / 10)}$$



同一成熟度，随温度调整

示例 $I = 60$

20°C: 3.50 天 30°C: 1.35 天

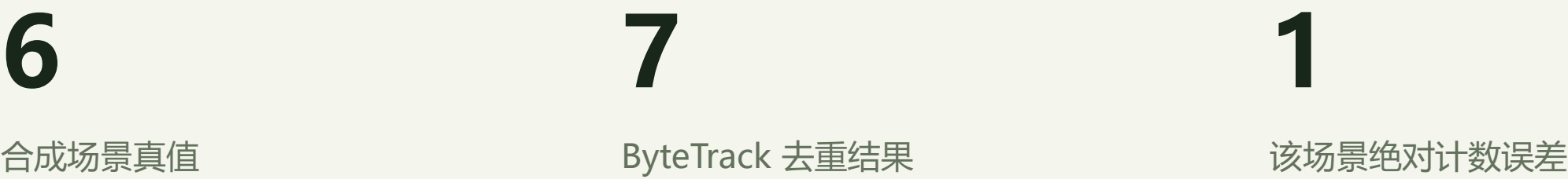
输出范围与风险提示

Q10 参数区间为 2.2–3.0

低于 13°C、高于 32°C 触发风险提示

模型换算示例，尚未通过真实储藏实验验证天数误差。

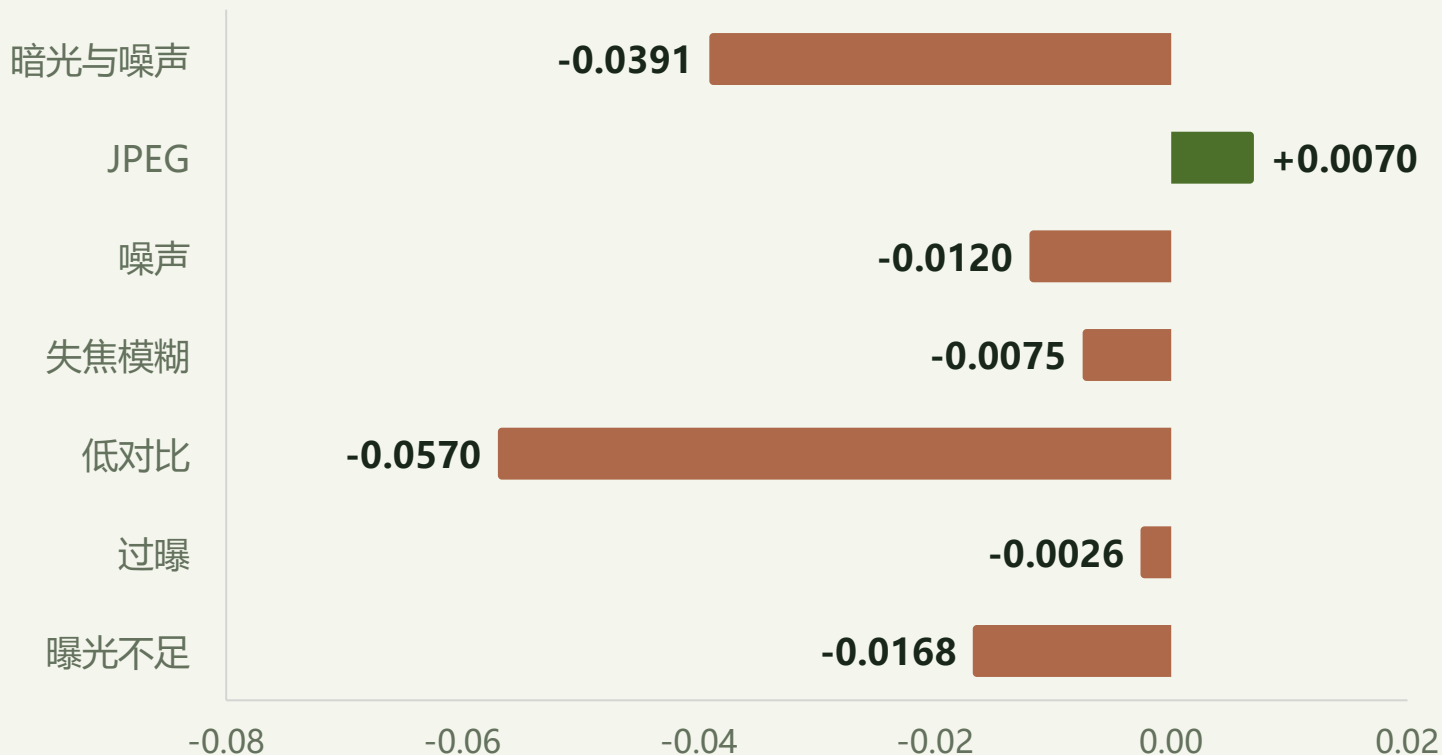
视频跟踪与去重计数



计数方式	计数结果	绝对误差
逐帧累加	364	358
单帧最大值	3	3
ByteTrack 轨迹去重	7	1

同一轨迹只计一次；当前结果来自单个合成视频，真实遮挡与 ID 切换仍需验证。

成像扰动下的性能变化



-0.057

低对比条件下的 F1 降幅

72.3%

噪声下档位一致率

基线 F1 = 0.8864，扰动后的表现需与检测精度页区分评估路径。

感谢各位评委 敬请批评指正

让检测结果，服务每一批货的分拣与周转

香蕉成熟度在线检测系统

智能检测与批次分拣平台 / V1.0

刘玉浩 万世豪 张兴桦