Journal of Image and Signal Processing 图像与信号处理, 2026, 15(4), 570-581
Published Online October 2026 in Hans. https://www.hanspub.org/journal/jisp
https://doi.org/10.12677/jisp.2026.154050

Hans汉斯

# 文化遗产场景下花山岩画智能检测与可视化分析系统研究与实现

**胡亚祺，唐靖华，熊圣昊，韦玉荣，刘小满***

广西民族师范学院数学与计算机科学学院，广西 崇左

收稿日期：2026年9月26日；录用日期：2026年10月17日；发布日期：2026年10月28日

## 摘　要

**针对花山岩画姿态识别依赖人工、效率偏低且易受主观判断影响的问题，本文提出了一种基于YOLOv8目标检测的可视化分析系统。系统采用CustomTkinter构建交互界面，集成模型推理、结果渲染、批量处理、数据统计与场景识别功能。通过引入自适应标签布局策略，根据目标框位置自动调整标签显示区域，有效避免密集人像区域的标签重叠遮挡；提供置信度阈值实时调节功能，用户可根据图像质量与检测需求动态调整灵敏度，实现检出数量与准确率的灵活平衡；基于考古学研究成果构建场景识别规则体系，系统支持普通正身、首领正身、巫师正身、普通侧身、持鼓侧身、群聚正身六类姿态检测，可根据检测结果自动判定祭祀、集会、狩猎等场景类型，辅助参观人员快速理解岩画文化内涵。系统在常规硬件环境下能够实现快速检测，界面交互流畅，批量处理流程稳定。本系统将高精度检测模型转化为考古人员可直接操作的桌面工具，为花山岩画数字化保护提供了可用的技术实现方案。**

## 关键词

**花山岩画，目标检测，YOLOv8，文化遗产数字化，场景识别**

# Research and Implementation of an Intelligent Detection and Visual Analysis System for Huashan Rock Art in Cultural Heritage Contexts

**Yaqi Hu, Jinghua Tang, Shenghao Xiong, Yurong Wei, Xiaoman Liu***

School of Mathematics and Computer Science, Guangxi Minzu Normal University, Chongzuo Guangxi

Received: September 26, 2026; accepted: October 17, 2026; published: October 28, 2026

*通讯作者。

文章引用: 胡亚祺, 唐靖华, 熊圣昊, 韦玉荣, 刘小满. 文化遗产场景下花山岩画智能检测与可视化分析系统研究与实现[J]. 图像与信号处理, 2026, 15(4): 570-581. DOI: 10.12677/jisp.2026.154050

## Abstract

**Addressing the issues of manual reliance, low efficiency, and susceptibility to subjective judgment in the pose recognition of Huashan Rock Art, this paper proposes a visual analysis system based on YOLOv8 object detection. The system utilizes CustomTkinter to build an interactive interface and integrates functions such as model inference, result rendering, batch processing, data statistics, and scene recognition. By introducing an adaptive label layout strategy, the system automatically adjusts label display areas based on bounding box positions, effectively preventing label overlap and occlusion in areas with dense figures. It offers real-time adjustment of confidence thresholds, allowing users to dynamically tune sensitivity according to image quality and detection requirements, thereby achieving a flexible balance between the number of detections and accuracy. A scene recognition rule system, constructed based on archaeological research findings, enables the detection of six pose categories—ordinary frontal, leader frontal, shaman frontal, ordinary profile, profile holding a drum, and clustered frontal—and automatically classifies scenes such as rituals, gatherings, and hunting, helping visitors quickly grasp the cultural significance of the rock art. The system achieves rapid detection on standard hardware, features a smooth interactive interface, and ensures stable batch processing workflows. By transforming a high-precision detection model into a desktop tool directly operable by archaeologists, this system provides a practical technical solution for the digital preservation of Huashan Rock Art.**

## Keywords

**Huashan Rock Art, Object Detection, YOLOv8, Digital Preservation of Cultural Heritage, Scene Recognition**

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

## 1. 引言

### 1.1. 研究背景与意义

花山岩画地处广西左江流域，是古代骆越部族在崖壁上绘制的图像遗迹，2016 年入选世界文化遗产[1]。遗产区域内岩画点数量多、分布广，可辨识人像与图案达数千个，是研究古代部族社会生活与宗教信仰的重要实物资料。

岩画中人物姿态丰富，涵盖正身、侧身、持器等多种类型。考古学研究表明，不同姿态对应祭祀、集会、劳作等不同活动场景。例如，“巫师正身”与“首领正身”多与祭祀活动相关，大量“普通正身”聚集可能反映集会场景，“群聚正身”则体现了群体性活动的特征。准确识别各类姿态并推断场景类型，直接关系到对岩画文化内涵的理解。

但在实际研究中，人工识别耗时较长，且岩画经长期风化导致部分图像清晰度有限，不同人员识别结果存在差异。开发智能检测与可视化分析系统，将检测模型转化为考古人员可直接操作的桌面工具，对提升岩画数字化保护效率具有重要意义[2]。

### 1.2. 现有研究的不足

近年来，计算机视觉在文化遗产数字化保护中得到日益广泛的应用，壁画、石刻、岩画等考古图像

的分析、修复与语义理解均取得了较多进展。就岩画而言，现有研究主要集中于图像增强与去噪、风化区域检测、图案分割、符号识别，以及基于深度学习的分类与检索等任务。这些工作为岩画的数字化保存与初步解读提供了重要支撑，但多数方法仍停留在单个符号、局部图案或整图分类的层面，较少从“场景”层面建模岩画中人物与环境元素之间的空间关系与语义关联。有鉴于此，本文尝试从岩画图像中识别具有完整语义的场景单元，并借助场景识别规则体系对人物分布、空间邻近关系与基本构图模式进行初步建模；与已有研究相比，本文的贡献不在于提出一套通用的考古图像识别框架，而在于探索一种面向岩画场景理解的轻量化方法，为后续引入更复杂的语义关系建模与场景级分析奠定基础。

当前花山岩画数字化研究主要面临以下问题：交互工具方面，现有检测方法多以命令行形式存在，考古人员使用门槛较高；可视化方面，多数工具缺乏中文标签与自适应布局功能，密集区域标签易重叠遮挡；场景分析方面，现有方法停留在单一姿态分类层面，未能结合考古学成果提供场景判断功能。

针对上述不足，本文设计并实现了一套面向花山岩画的智能检测与可视化分析系统，采用CustomTkinter 构建图形化界面，集成模型推理、批量处理、数据统计与场景识别功能。系统对普通正身、首领正身、巫师正身、普通侧身、持鼓侧身、群聚正身六类姿态采用不同颜色标注，支持置信度阈值调节，自动排布标签避免遮挡，并根据检测结果自动判断场景类型，为花山岩画数字化保护提供可用的实现方案。

## 2. 系统总体设计

### 2.1. YOLOv8 目标检测技术

YOLOv8 是一种高效的端到端目标检测方法，相比 Faster R-CNN 等两阶段方法，在保持较高检测精度的同时显著提升了检测速度[3]。该模型采用 Anchor-Free 检测头设计，直接预测目标中心点与宽高，并引入 C2f 模块通过更丰富的梯度流提升特征提取能力。

本文采用的 best.pt 模型以 YOLOv8n 为基础架构，模型参数量约 3.2 M，在训练阶段结合了 Mosaic 数据增强、Copy-Paste 增强策略与 Focal Loss 损失函数，专门针对花山岩画低对比度、严重风化的图像特点进行了优化[4]。该模型支持 6 类姿态检测：普通正身、首领正身、巫师正身、普通侧身、持鼓侧身、群聚正身，在岩画人物姿态检测任务上取得了较好的检测效果。

### 2.2. 技术架构

本文所实现的系统按照功能不同进行了模块化划分：检测引擎、可视化渲染、场景识别、统计分析、数据管理五大模块。模块之间仅通过数据进行关联，没有过度的逻辑依赖，这样的结构也方便后续对功能进行扩展和完善。核心技术栈如表 1 所示。

**Table 1.** System technology stack
**表 1.** 系统技术栈

| 组件 | 技术/版本 | 用途 |
|---|---|---|
| 图形界面 | CustomTkinter 5.2.0 | 现代化主题窗口与控件 |
| 检测引擎 | Ultralytics YOLOv8 8.4.48 | 目标检测推理 |
| 图像处理 | OpenCV 4.8.0 | 图像读取、缩放、格式转换 |
| 图像标注 | Pillow 10.0.0 | 边界框绘制、中文标签渲染 |
| 数据可视化 | Matplotlib 3.7.0 | 统计图表生成 |
| 数据存储 | Pandas 2.0.0 | CSV 报告导出 |
| 并发处理 | threading | 批量检测异步执行 |

系统界面采用左右分栏设计，左侧为导航栏，含 Logo 区域、五个功能导航按钮和底部信息区域；右侧为主内容区，根据用户选择动态切换功能页面。系统主界面如图 1 所示。




**Figure 1.** System main interface
**图 1.** 系统主界面

### 2.3. 功能模块

系统由五大模块组成。

检测引擎模块负责模型调用与任务执行，界面上提供模型加载、图像选择、参数调整以及启动检测等常用操作按钮，后台会通过接口调用 YOLOv8 模型完成推理。用户能够通过滑动条调整置信度大小，取值范围为 0.1 到 0.9，调整后的数值会实时显示在界面上，并在下次检测时生效。

可视化渲染模块在检测结果图像上绘制彩色边界框和中文标签，实现标签位置自适应调整。

统计分析模块读取历史记录数据，以饼图、柱状图、折线图展示类别数量与占比，界面支持一键刷新，并提供导出 CSV 按钮。

数据管理模块保存标注后的图像(JPG 格式)、检测历史(JSON 格式)和批量报告(CSV 格式)。

场景识别模块根据检测出来的姿态组合，按照提前设定好的规则给出场景判断。如果图片中出现巫师正身，系统会判定为祭祀场景；普通姿态数量较多时判定为集会场景；出现多名首领正身则判定为狩猎场景；其余情况统一归为普通场景。

模块数据流如图 2：用户输入图像→检测引擎执行推理→输出检测结果→同时送入可视化模块(生成标注图像)和场景识别模块(判断场景类型)→统计模块汇总历史数据→数据管理模块保存结果。

## 3. 核心功能实现

### 3.1. 检测引擎实现

本系统在界面中调用了 YOLOv8 模型的预训练权重文件 huashanyan.pt，该模型以轻量化的 YOLOv8n 为基础架构，在训练阶段结合了 Mosaic 数据增强、Copy-Paste 增强策略以及 Focal Loss 损失函数进行优化，最终在花山岩画专用数据集上达到了较好的检测效果[5]。

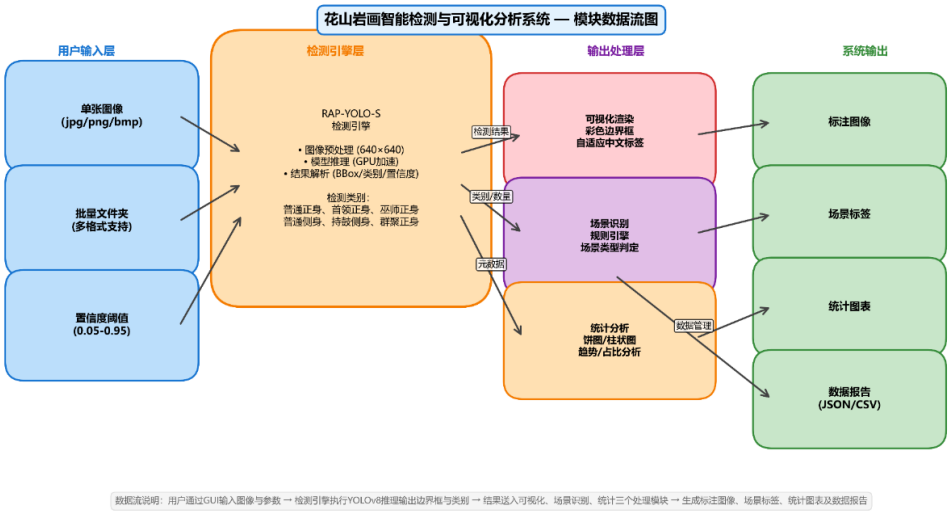



**Figure 2.** Module data flow diagram
**图 2.** 模块数据流图

在实际运行流程中，系统首先通过 OpenCV 读取图片并保留原始色彩信息，随后将图像统一缩放到 640 × 640 的尺寸，并把像素值归一化到 0 到 1 之间，再送入模型完成推理。模型输出结果后，程序会对边界框坐标、类别编号以及置信度分数进行解析，最终在界面上呈现标注信息。

为了适应清晰度不同的岩画图片，系统使用滑动条来调节置信度大小，范围从 0.1 到 0.9，每次最小调整 0.01。参数改完马上就能用，不用重启软件。风化比较严重、轮廓不清楚的图片，可以把阈值调低一些，更容易找到目标；图像质量较好、目标清晰时，可以适当提高阈值，减少不可靠的结果。

考虑到权重文件可能出现路径丢失或移动的情况，系统在启动加载模型时会进行判断。如果找不到指定的.pt 文件，会自动弹出提示窗口，让用户手动选择文件位置，保证界面能够正常运行。

单张检测界面如图 3 所示，左侧控制面板包含图片选择、参数调节和检测触发功能，右侧为图像显示区域。

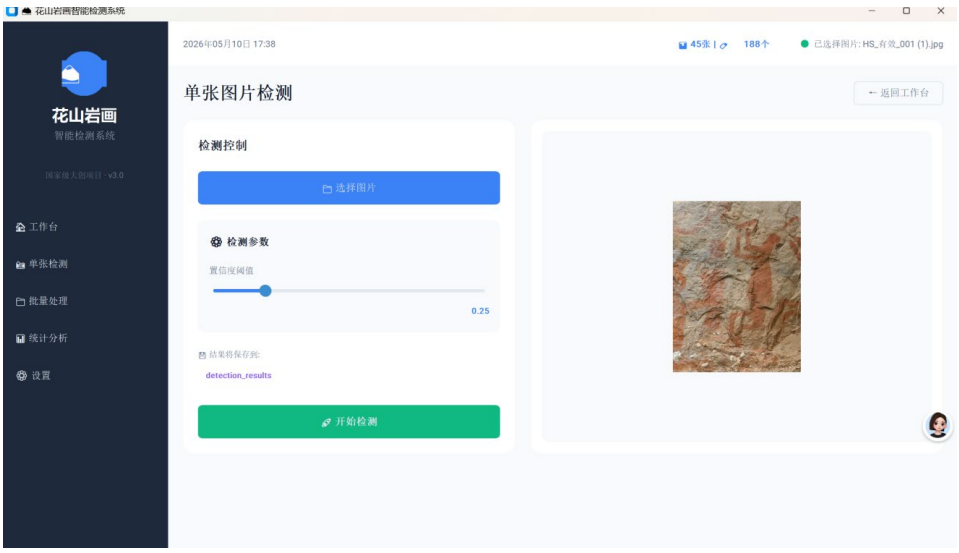



**Figure 3.** Single-image detection interface
**图 3.** 单张检测界面

## 3.2. 可视化渲染实现

### 3.2.1. 颜色方案

六类姿态分配固定颜色，确保在暗红色岩画背景上的可视性和区分度(表 2)：

**Table 2.** Color scheme for pose categories
**表 2.** 姿态类别颜色方案

| 姿态类别 | RGB 值 | 色样 |
|---|---|---|
| 普通正身 | (102, 204, 0) | 柔和绿色 |
| 首领正身 | (255, 153, 51) | 柔和橙色 |
| 巫师正身 | (102, 178, 255) | 柔和蓝色 |
| 普通侧身 | (255, 102, 178) | 柔和粉色 |
| 持鼓侧身 | (153, 102, 255) | 柔和紫色 |
| 群聚正身 | (0, 206, 209) | 深青色 |

### 3.2.2. 标签自适应布局算法

系统在显示标签时，会统一使用“姿态名称 置信度”这样的格式，比如“普通正身 95%”。文字优先使用微软雅黑，大小设为 20 px，如果当前系统不支持这种字体，会自动换成黑体或者系统默认字体，确保中文能够正常显示出来。

为了避免密集人像区域出现标签重叠遮挡，程序设计了一套自适应布局逻辑：先通过 draw.textbbox 计算标签文本所占像素大小，优先将标签放在目标框正上方。如果上方空间不足，则放在目标框正下方；如果上下空间都不够，则放在框内右侧位置。最后检查标签是否超出图片左右边界，若超出则自动左移或右移，保证完整显示。

标签背景色与边框颜色保持一致，文字使用白色，提高对比度。在人像密集、排列紧凑的岩画区域，这套策略可以明显减少标签重叠，提升可读性。

检测结果可视化效果如图 4 所示，系统在岩画图像上叠加彩色边界框和中文标签，并实时显示场景识别结果。

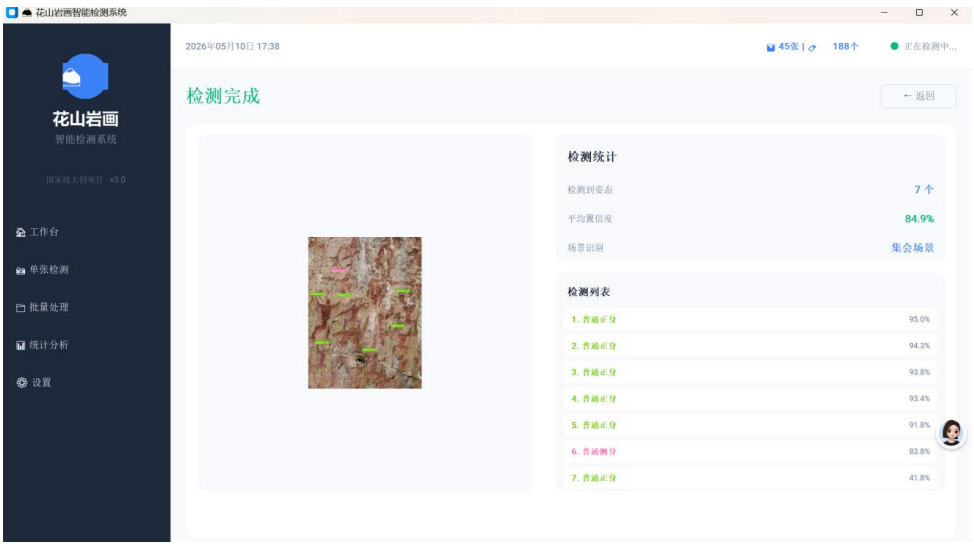



**Figure 4.** Visualization of detection results and scene recognition
**图 4.** 检测结果可视化与场景识别

### 3.3. 批量处理实现

批量处理功能支持用户直接选择整个文件夹，系统会自动遍历目录下的图片(支持.jpg、.jpeg、.png、.bmp、.webp 格式)。

批量处理界面如图 5 所示，用户选择文件夹后系统遍历图片文件并弹出确认对话框。

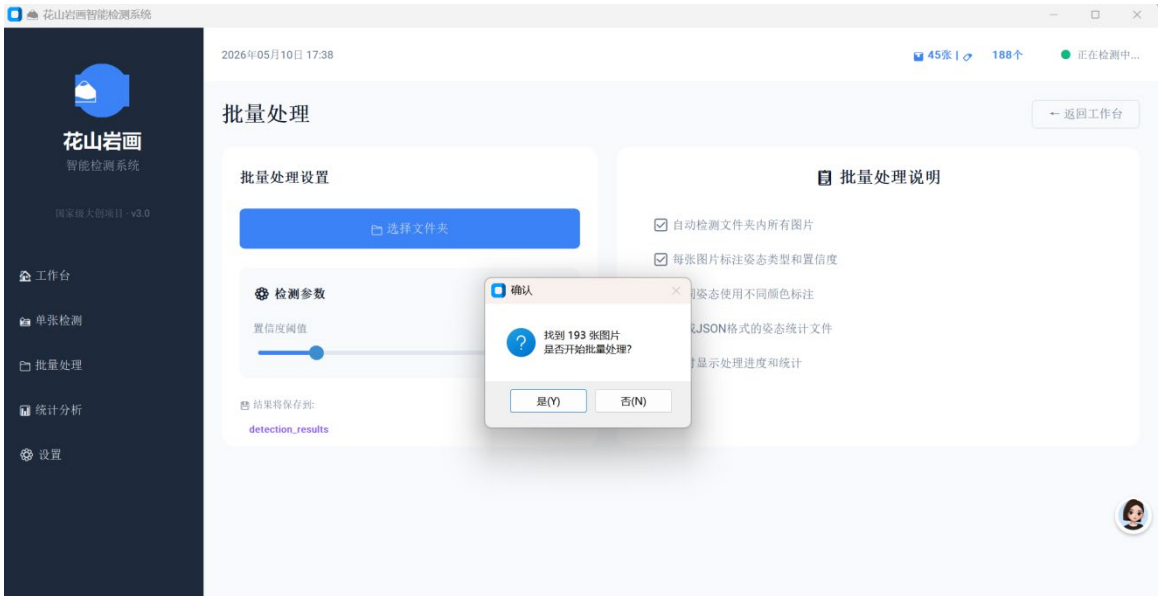



**Figure 5.** Batch processing confirmation screen
**图 5.** 批量处理确认界面

批量窗口包含以下组件：进度条、状态提示标签、实时统计面板、六类姿态计数面板、日志滚动区域以及停止按钮。

为了避免处理大量图片时出现界面卡死，批量任务放在独立线程中运行，界面更新通过 after 方法回到主线程执行。用户点击停止按钮后，系统会设置中断标记，线程在处理下一张图片前检查标记并安全退出，不会出现崩溃或卡顿。

批量处理完成后自动生成：

姿态统计.json：各类别数量统计；检测报告.csv：包含文件名、输出文件、检测数量、姿态类型四列。

### 3.4. 统计分析实现

统计指标包括：处理图片总数、识别姿态总数、各姿态类别数量及占比。

可视化嵌入 CustomTkinter，绘制饼图、柱状图、趋势图、水平条图等，直观展示分析的数据对比。图表使用 FigureCanvasTkAgg 渲染，支持实时刷新。用户可导出统计结果文件，包含姿态类别、识别数量、占比，以及处理图片总数和识别姿态总数。

统计分析界面如图 6 所示，系统汇总历史检测数据并生成多维度可视化图表。

### 3.5. 场景识别实现

基于考古学研究成果，系统根据检测结果自动判断场景类型(表 3) [6]：

识别逻辑按优先级执行：巫师正身(最高)→ 普通姿态数量 → 首领正身数量 → 普通场景。该优先级基于考古学意义的重要性设计。




**Figure 6.** Statistical analysis interface
**图 6.** 统计分析界面

**Table 3.** Scene recognition rules
**表 3.** 场景识别规则

| 条件 | 场景判定 | 考古学依据 |
|---|---|---|
| 检测到巫师正身 | 祭祀场景 | 巫师正身为祭司形象，标志祭祀仪式 |
| 普通正身 + 普通侧身 > 5 | 集会场景 | 大量普通姿态反映集体活动 |
| 首领正身 > 1 | 狩猎场景 | 多个首领共现可能反映狩猎或战争 |
| 其他 | 普通场景 | 不符合以上特征的日常活动 |

场景识别结果如图 7 所示，系统根据检测结果自动判断岩画场景类型。




**Figure 7.** Detection results for a scenario with co-existing multiple categories
**图 7.** 多类别共存场景检测结果

## 4. 系统测试

### 4.1. 测试环境

测试配置如表 4。

**Table 4.** Test configuration
**表 4.** 测试配置

| 项目 | 配置 |
|---|---|
| 操作系统 | Windows 11 |
| CPU | Intel Core i7-12700K |
| 内存 | 32 GB DDR4 |
| GPU | NVIDIA RTX 3060 12GB |
| Python | 3.10.0 |
| CUDA | 11.8 |

测试图像为花山岩画人物姿态样本集，包含 831 张图像、4345 个标注实例。

### 4.2. 功能测试

**单张检测测试：**选取一张包含多类人物姿态的岩画图片，导入系统后点击检测按钮。系统能够正常读取模型文件，并顺利完成推理过程，在界面右侧显示出标注完成的效果图，如图 3 所示。经过观察，边框颜色与姿态类别能够正确对应，标签位置排布合理没有重叠，置信度格式显示正常，该项测试可以通过。

系统在常规 GPU 环境下能够实时处理，单张图片检测与可视化渲染总耗时在数十毫秒级别，批量处理流程稳定流畅，满足日常使用需求。

**批量处理测试：**批量处理测试选用了包含 193 张图片的文件夹，系统成功识别图片数量并弹出确认提示。点击开始处理后，进度条能够从 0 逐步增长到 100%，日志面板同步显示处理详情。在处理过程中点击停止按钮，系统能够在短时间内安全终止任务，处理结束后正常输出统计信息，整体流程稳定，测试通过。

**置信度调节测试：**初始阈值 0.25，检测结果显示若干目标。调至 0.1，检出目标数量增加；调至 0.7，检出数量减少。调节实时生效，无需重启。测试通过。

**场景识别测试：**选择包含多个首领正身的图像(如图 7)，系统显示“狩猎场景”；选择包含 5 个以上普通姿态的图像(如图 4)，显示“集会场景”。规则执行符合设计预期。测试通过。

**统计分析测试：**加载检测历史数据(如图 5)，饼图、柱状图、趋势图显示与数据一致。点击导出报告按钮，生成 CSV 文件，Excel 打开中文正常。测试通过。

### 4.3. 性能测试

系统在测试环境下的性能表现如表 5：

**Table 5.** Test performance
**表 5.** 测试性能

| 测试场景 | 平均检测时间 | 端到端时间 | 内存占用 |
|---|---|---|---|
| 单张图片(GPU 模式) | 8 ms | 20 ms | 1.5 GB |
| 批量 100 张(GPU 模式) | 8 ms/张 | 18 ms/张 | 1.8 GB |
| 批量 500 张(GPU 模式) | 7 ms/张 | 18 ms/张 | 2.0 GB |

### 4.4. 可视化效果测试

**密集排列场景：**图像包含超过 20 个层叠排列人像。系统检出 16 个可辨认人像(检出率与模型性能一致)。边界框分配清晰，标签经自适应调整后无重叠(如图 8)。

**多类别共存场景：**图像同时包含四类人像。系统准确区分体型偏大的首领正身与周围普通正身。各类别边界框颜色区别明显，便于快速识别。

**低对比度退化图像：**风化严重图像中，系统检出 80%以上人像，未出现虚检。对严重剥落仅剩局部轮廓的目标未检出，符合预期。

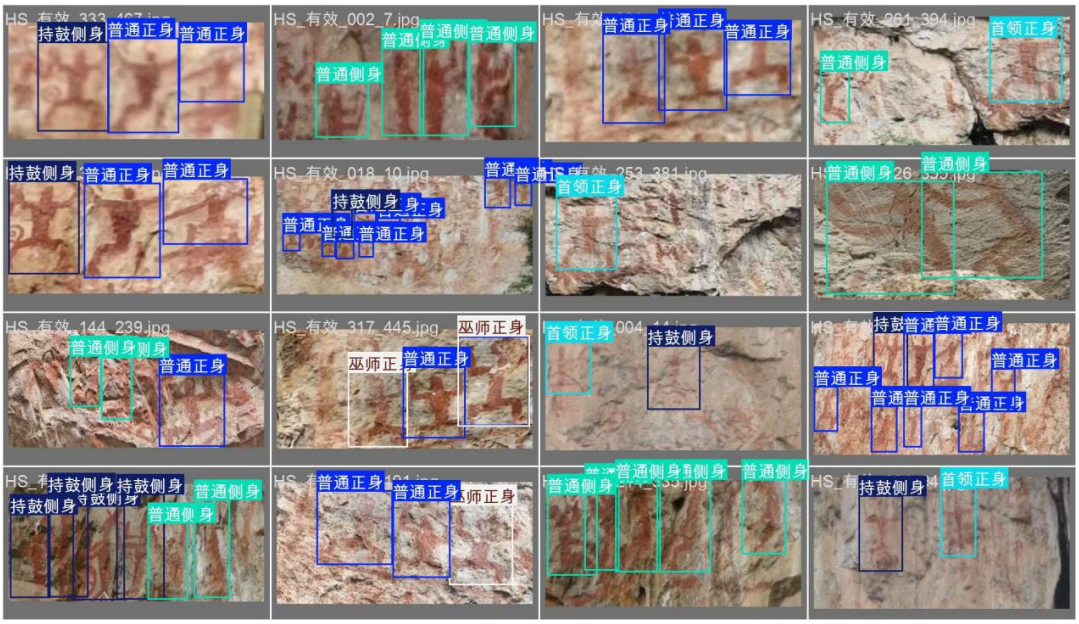



**Figure 8.** System operation results
**图 8.** 系统运行结果

## 5. 结论与展望

### 5.1. 结论

考虑到花山岩画姿态识别对效率与准确性的双重要求，本文基于 RAP-YOLO-S 检测模型，设计并实现了一套智能检测与可视化分析系统。将整个流程整合于统一的图形化界面中，采用 CustomTkinter 构建现代化交互环境，集成模型推理、结果渲染、批量处理、数据统计与场景识别功能，有效解决了传统命令行工具使用门槛高、可视化效果欠佳、缺乏场景分析能力的问题。

在可视化方面，系统采用自适应标签布局策略，根据目标框位置自动调整标签显示区域，确保标签在密集人像区域也能完整显示，避免了标签重叠遮挡，显著提升了可读性。对普通正身、首领正身、巫师正身、普通侧身、持鼓侧身、群聚正身六类姿态采用不同颜色标注(绿色、橙色、蓝色、粉色、紫色、深青色)，提高视觉区分度。在交互设计上，提供置信度阈值实时调节功能(范围 0.1~0.9)，用户可根据图像风化程度与检测需求动态调整灵敏度，调节后即时刷新检测结果，实现检出数量与准确率的灵活平衡。在场景分析方面，构建基于考古学研究成果的场景识别规则体系，依据检测结果自动判定祭祀(出现巫师正身)、集会(普通正身 + 侧身 > 5)、狩猎(首领正身 > 1)等场景类型，使检测结果从单一姿态分类进一步延伸至场景层面的辅助分析。

本系统采用统计卡片式界面设计，在工作台直观展示已处理图像数量、识别姿态总数与保存位置信息，提供单张检测、批量处理、统计分析三大快捷功能入口。系统支持多线程批量处理，自动生成统计图表(饼图、柱状图)与 CSV 报告，为考古人员提供了完整的桌面工具。与现有工具相比，本系统相较于命令行式检测工具提供了完整的 GUI 界面，无需配置环境或修改代码；相较于 LabelImg、CVAT 等通用标注工具，本系统针对花山岩画低对比度图像和六类姿态进行了专门优化，更贴合考古研究需求[7]。

## 5.2. 展望

本文采用自底向上、符号化的场景识别规则体系：先检测出画面中的人物，再依据人工设定的空间关系规则，判断整幅画面所表达的活动类别。该体系不依赖大规模标注，每一步判断均可回溯解释，但其局限性同样明显，因此本文将其定位为初步探索与基线方法，而非最终方案。

泛化能力方面，场景识别规则体系受限于规则的完备性，而规则的完备性又受限于专家经验。岩画的构图方式千差万别，人工规则只能覆盖研究者事先预设的典型空间模式，当画面出现罕见、长尾或多种要素复合的构图时，泛化能力明显不足。

鲁棒性方面，规则的判断以检测结果为唯一输入，对检测误差缺乏缓冲。一旦关键人物因风化、遮挡而被漏检或误检，高层的规则判断也会随之出错，错误沿规则链逐层传递。

语义表达方面，离散的布尔规则难以刻画空间语义的连续性与多义性。空间语义本身是连续而模糊的，例如“靠近”与“围绕”之间并没有清晰的界限，规则既无法刻画这种程度上的差异，也无法给出带置信度的综合判断。

除方法本身外，规则体系的可靠性还受图像质量制约。风化退化方面，岩画因长期暴露常出现刻痕变浅、轮廓模糊、地衣与矿物结壳覆盖、岩面开裂剥落等现象，人像轮廓一旦残缺、与背景对比度骤降，前端检测器的召回率即明显下降，而规则体系对这类漏检尤为敏感，关键人像一旦丢失，整条规则链便随之失效；遮挡与叠加方面，同一岩面往往叠加着不同时期凿刻的图案，人像在二维投影中彼此交叠，不仅降低检测精度，也扰乱了规则赖以判断的空间关系；光照与对比度方面，现场拍摄条件导致图像对比度剧烈波动，虽可通过预处理加以缓解，但仍是影响结果稳定性的系统性因素。

针对上述局限，后续可尝试将岩画场景显式表示为场景图，并利用图神经网络(GNN)端到端地学习实体间的空间关系，以替代依赖人工枚举的判别方式。具体构建上，可以检测到的每一个实体为节点，节点特征由检测器提取的外观或姿态特征与类别嵌入拼接而成；在存在空间互动的实体之间连接有向或无向边，边的属性编码相对位置、欧氏或图上距离、朝向夹角、接触或重叠等几何信息，可为带类型或带权重的边；此后利用 GCN、GAT 或 Graph Transformer 进行多层消息传递，使每个节点的表示逐步融合邻域的空间上下文，最终经图级池化得到整个场景的表示，用于活动分类。

相较场景识别规则体系，GNN 的关系是可学习的，网络能在训练中自行发现对识别起关键作用的空间模式，无需依赖人工枚举；其对检测噪声也更为鲁棒，图结构的归纳偏置可在一定程度上缓解个别节点的漏检与误检；同时便于可解释性改造，借助注意力权重能够回溯判断主要来自哪些实体与空间关系。此外，场景识别规则体系的语义规则还可作为先验反哺网络，用于初始化网络结构或约束边的类型，形成“符号先验 + 可微学习”的混合范式。

除上述局限外，本文还受单标签假设制约，即默认一幅场景只属于一种活动。然而，许多构图复杂的画面往往同时包含多种活动，例如同一画面既有狩猎又有舞蹈，或在祭祀场景中叠加了其它元素；将这类场景强行归入单一类别，既损失了信息，也造成了标注歧义。为此，后续有必要引入多标签分类框架，将顶层的 softmax 替换为 sigmoid 加二分类交叉熵，使同一场景能够同时命中多个活动标签；对于多标签场景中常见的类别不平衡，采用加权损失或 Focal Loss 加以缓解；标签之间的相关性可借助 Label

Powerset、Binary Relevance 或基于 GNN 的标签共现图建模；评估指标也应由单一准确率转向宏/微 F1、mAP 与 precision@k，以更全面地反映多标签识别效果。将 GNN 的场景图表示与多标签分类头相衔接，即可构成“场景图编码→多标签活动预测”的完整可学习框架，这也是本文场景识别规则体系最自然的演进方向。

在数据与评价层面，未来还可从以下几个方向继续完善。鲁棒检测方面，通过模拟刻痕变浅、地衣覆盖、岩面开裂等退化形态进行合成数据增强，并结合域自适应/域泛化技术，缩小不同遗址、不同风化程度之间的域间差异，提升前端检测在退化图像上的召回率；不确定性量化方面，为场景识别引入概率输出或置信度校准，使系统在低对比度、强遮挡等高风险样本上主动降低置信度，而非强行输出某一类别；基准数据方面，构建更大规模、具备结构化空间关系标注的岩画场景数据集与统一评测基准，探索不同文化区域、不同年代岩画之间的跨域迁移学习，推动该方向由单点方法验证走向可比较、可复现的系统研究。

未来可从跨平台打包、交互优化、功能扩展等方向继续完善系统。具体包括使用 PyInstaller 或 Nuitka 将项目打包为独立 exe 实现开箱即用；邀请考古领域用户进行实际测试，根据反馈调整界面布局与操作流程；接入三维模型接口实现二维检测与三维岩画模型的联动显示[8]，引入知识图谱提升场景识别的深度与准确性；增加导出功能，支持检测结果与标注图像的批量导出与分享。

## 基金项目

2025 年自治区级大学生创新创业训练计划项目创新训练项目《基于改进 YOLOv8 的花山岩画多姿态人像检测与数字化保护系统研究》(编号：S202510604109)成果。

## 参考文献

[1] 肖波. 左江花山岩画申遗——实现文化遗产保护与利用的有效途径[J]. 齐鲁艺苑, 2015(2): 10-14.

[2] 张超. 数字化技术在岩画保护中的应用——以中宁石马湾岩画数字化保护项目为例[J]. 文物鉴定与鉴赏, 2020(3): 72-74.

[3] 郭辉, 姚志强, 何多政, 等. 基于 YOLOv8 的目标检测技术应用研究[J]. 机械工程师, 2026(7): 126-130.

[4] 陈翠琴, 范亚臣, 王林. 基于改进 Mosaic 数据增强和特征融合的 Logo 检测[J]. 计算机测量与控制, 2022, 30(10): 188-194, 201.

[5] 关雪敏, 何皓怡. 基于 AKAZE 特征匹配的花山岩画图像识别研究[J]. 信息与电脑(理论版), 2023, 35(19): 146-148.

[6] 王璐, 陆筱霞, 蔡自兴. 基于局部显著区域的自然场景识别[J]. 中国图象图形学报, 2008(8): 1594-1600.

[7] 王景鑫, 潘欣. 一种基于 Labelimg 的辅助标注方法[J]. 科技创新与应用, 2023, 13(29): 145-148.

[8] 雷国洪. 岩画形象的三维数据采集与应用[J]. 三峡论坛(三峡文学理论版), 2014(1): 92-95.