

个人信息 OCR 比对工具 使用说明

一、工具简介

本工具用于自动识别身份证照片，并与 Excel 人员名单进行比对校验。

核心功能：

- 自动识别身份证上的姓名和证件号码
- 与 Excel 表格数据进行匹配比对
- 生成比对结果报告（标注匹配/不匹配）

二、准备工作

1. Excel 文件要求

- 格式：.xlsx 或 .xls
- 必须包含列：**姓名**、**身份证号**（列名需完全一致）
- 示例：

姓名		身份证号
张三		110101199001011234
李四		110101199001011234

2. 附件文件夹结构

附件需按以下结构整理：

附件文件夹/

```
|—— 张三/  
|   |—— 身份证正面.jpg  
|   |—— 其他照片.png  
|—— 李四/  
|   |—— 身份证.jpg  
|—— ...
```

要求：

- 每个人一个文件夹，文件夹名称为姓名
- 文件夹内放置该人员的身份证照片
- 支持格式：jpg、jpeg、png、bmp

三、操作步骤

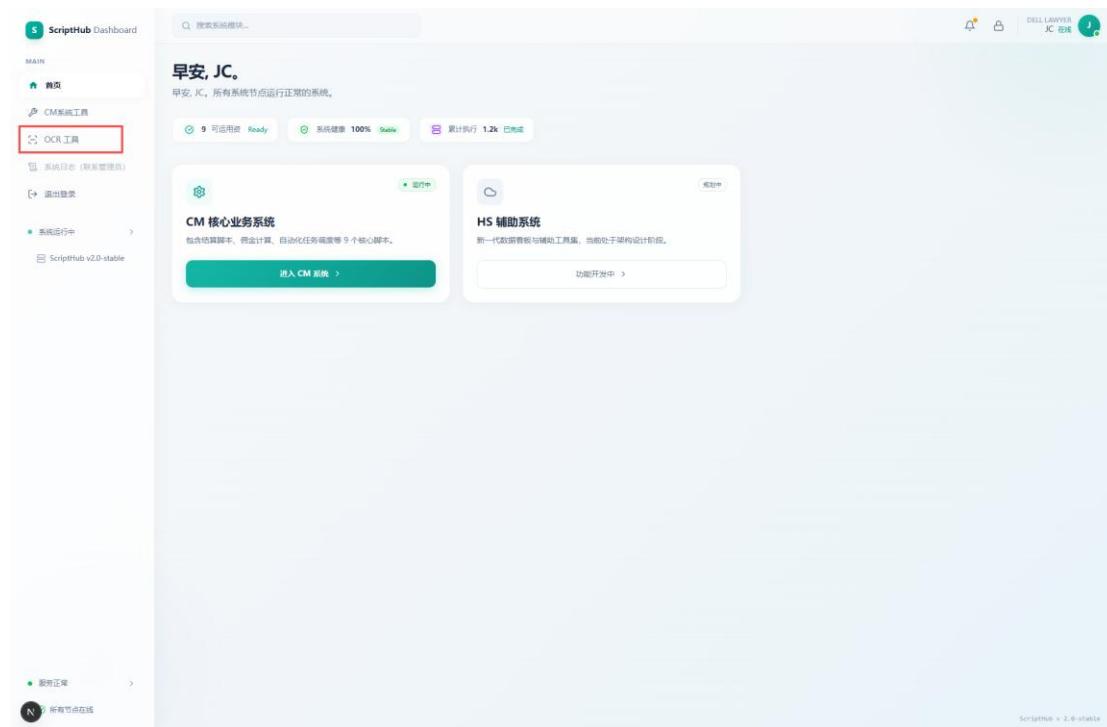
步骤 1：访问工具页面

【系统访问时间在 9:30-18:30 之间，如遇到不可访问的问题请及时反馈】

打开浏览器，访问工具地址（<https://scripthub.tail150593.ts.net/>）

登录秘钥：`U5*mO~yW6%LxJ3dQ0nHaD8~L0bV&cXoM9uA2j`

登录成功后，点击 **【OCR 工具】** 去使用



步骤 2：上传文件

1. 点击「个人信息表」，选择准备好的 Excel 文件
2. 点击「附件文件夹」，选择整个附件文件夹（**不是单个文件**）



步骤 3：选择运行模式

- **模式 1 (推荐): Excel 优先** (按 Excel 顺序匹配附件)
 - 适用于: 附件齐全, 文件夹命名规范
- **模式 2: 附件优先** (按附件识别后反查 Excel)
 - 适用于: 附件可能有多余或命名不规范的情况

步骤 4：执行比对

点击「立即开始比对」按钮，等待处理完成。

处理过程中可在日志区域查看进度，每个人员处理时会显示：

- 识别到的姓名和身份证号
- 与 Excel 的比对结果

在按钮下方/右上角可以看到整体执行进度：2/10（代表总数 10，执行到了第二个）

The screenshot shows the 'OCR Intelligent Comparison' interface. At the top right, it says '正在处理: 2 / 3' (Processing: 2 / 3) and has a red button labeled '终止处理' (Stop Processing). The main area is divided into two sections: 'Data Source Configuration' (数据资源配置) on the left and a log window on the right.

Data Source Configuration:

- 上传文件: '个人信息表.xlsx' (9.3 KB)
- 附件图片库 (包含以姓名命名的子文件夹): '已选择 4 个文件' (总大小: 4.90 MB)

运行模式:

- Excel 优先: 按 Excel 名单顺序去查找并匹配附件 (推荐)
- 附件 优先: 遍历附件文件夹识别, 反查匹配 Excel

正在智能处理中...

处理进度: 2 / 3
预计剩余 1 人待处理

Log Window (task_output.log):

```
开始上传: Excel + 4 个图片 (共 4.90 MB)
上传中, 请耐心等待...
上传完成, 开始 OCR 处理...
已接收 Excel: 个人信息表.xlsx
mode 参数值: 1
原始路径: 附件信息/王京川/毕业.png → 映射为: 附件信息/王京川/毕业.png
已接收图片文件: 4 个
[14:29:30] =====
[14:29:30] 【模式】按 Excel 顺序匹配附件 开始...
[14:29:30] 正在获取 PaddleOCR 引擎 (CPU + oneDNN)...
[14:29:46] PaddleOCR 引擎就绪!
[14:29:46] 共需处理 3 人
[14:29:46] -----
[14:29:46] [1/3] 处理人员: 张三 (文件夹 key=张三)
[14:29:46] ▲ 未找到附件文件夹: 张三, 标记为未识别
[14:29:46] -----
[14:29:46] [2/3] 处理人员: 王京川 (文件夹 key=王京川)
[14:29:46] [阶段] 文件名包含身份证/证件的优先图片: ['证件.jpg']
[14:29:46] -> 阶段识别: 证件.jpg
```

执行过程中可以点击右上角【终止处理】进行中断执行操作

步骤 5：下载结果

处理完成后, 点击「**下载 Excel 报表**」获取结果文件。



四、结果说明

生成的 Excel 文件会新增以下列:

列名	说明
OCR_姓名	从身份证识别出的姓名
OCR_身份证号	从身份证识别出的证件号
OCR_比对结果	比对状态 (见下表)

比对结果状态: | 状态 | 含义

全部匹配 | 姓名和身份证号均一致 ||

姓名不匹配 | 姓名不一致, 证件号一致 ||

证件号不匹配 | 证件号不一致, 姓名一致 ||

都不匹配 | 姓名和证件号均不一致 ||

未识别 | 未找到附件或识别失败 |

五、常见问题

Q1：上传后一直没反应？

- 检查文件大小，大量图片上传需要时间
- 查看日志区域是否有错误提示

Q2：识别不准确？

- 确保身份证照片清晰、无遮挡
- 避免过度倾斜或模糊的照片

Q3：找不到附件？

- 确认文件夹名称与 Excel 中的姓名完全一致
- 注意空格、特殊字符等

文件过大可能会遇到上传失败的情况，如遇到该问题，请及时反馈！