

铁路货运单信息提取 · 使用文档

批量上传 PDF 铁路货运单，自动提取 24 列信息汇总成一张预览表，一键导出 Excel / CSV。多页 PDF 按页逐张提取。

一、功能介绍

「信息提取」是铁路货运单处理工具的第一个功能页签，解决“货运单信息要手工录入 Excel”的问题：

- 批量解析：单次最多 300 个 PDF 货运单（超过 50 个自动分批，边处理边显示），并发解析，逐个汇总成一张预览表
- 24 列信息：运单号、制单日期/制单人、需求号、公司、托运/收货双方公司、发站/到站、经办人及手机号、受票方名称/识别号、总费用、费用类目、货物名称/件数/价格/重量/箱号
- 多页逐张提取：多页 PDF（一页一张货运单）按页独立提取，页码列标注页序
- 货物明细展开：一页含多条货物明细时按货物展开成多行，基础信息每行重复，方便 Excel 筛选透视
- 失败明示：扫描件（图片式 PDF）、损坏文件单独列出失败原因，不影响其他文件
- 双格式导出：预览确认后一键导出 .xlsx 或 .csv
- 隐私安全：网站不保留任何形式的用户数据，用完即删

适用人群：铁路运输企业货运台账登记、物流公司运单核对、贸易企业运费结算对账。

二、详细操作步骤

第 1 步：选择文件

三种方式任选：

- 点击上传：点击上传区 → 弹出文件选择框 → 多选 PDF 文件
- 拖拽上传：从资源管理器拖一批 PDF 到上传区
- 整个文件夹：点「选择整个文件夹」→ 文件夹内全部 PDF 自动加入（含子文件夹）

要求：PDF 电子铁路货运单（有文字层），单文件 ≤ 20MB，单次 ≤ 300 个。混入的非 PDF 文件自动跳过。

第 2 步：等待解析

松开鼠标后系统自动开始解析，进度条显示“正在解析...”与完成计数。超过 50 个时自动分批处理（每批 50 个），每批完成的结果即时追加到预览表。解析在服务器独立进程进行，完成后自动展示结果。

第 3 步：查看结果

解析完成后，下方显示：

- 统计栏：共 N 个文件 / 成功 N / 失败 N / 货运单 N 张（按运单号去重统计）
- 24 列预览表：每行一条货物记录（一张货运单多条货物 = 多行），左右滚动查看全部列；行数多时表格内滚动，页面不无限拉长
- 失败清单：解析失败的文件逐个列出原因（扫描件暂不支持等）

第 4 步：导出表格

点「导出Excel」或「导出 CSV」→ 下载文件（文件名含时间戳，如`货运单提取结果\_20261009-153000.xlsx`）。失败文件不进表格，请在失败清单中单独处理。

重新上传

点「重新上传」收起结果，回到初始上传状态。

三、常见问题解答

Q1：支持哪些货运单？

铁路货运单（电子版 PDF），共提取 24 列信息：运单号、制单信息、托运/收货双方公司、经办人及手机号、受票方、费用明细、货物明细等。扫描件（图片式 PDF）暂不支持，会明确提示。

Q2：一个 PDF 里有多张货运单？

可以。多页 PDF 会按页逐张提取，每页一张货运单独立成行；一页含多条货物明细时按货物展开成多行，「页码」列标注页序，导出表格自动全部包含。

Q3：批量上传有限制吗？

单文件 ≤ 20MB，单次 ≤ 300 个。超过 50 个时自动分批处理（每批 50 个），结果边处理边显示，无需手动分批。

Q4：导出的表格包含什么？

24 列货运单要素：运单号、制单日期/制单人、需求号、公司、托运/收货公司、发站/到站、经办人及手机号、受票方名称/识别号、总费用、费用类目、货物名称/件数/价格/重量/箱号，可导出 .xlsx 或 CSV。

Q5：一张货运单多个货物怎么显示？

每个货物一行（货物名称、件数、货物价格、重量、箱号各占一行），表头的基础信息（运单号、制单日期等）每行重复，方便 Excel 筛选透视。

Q6：扫描件怎么办？

图片式扫描 PDF 无法提取文字，会在失败清单中标注“未提取到货运单数据（扫描件暂不支持）”。需要 OCR 的场景请后续关注工具升级。

Q7：文件会被保存吗？

不会。网站不保留任何形式的用户数据，用完即删。

Q8：下载的文件在哪里？

在浏览器的默认下载文件夹（通常是「下载」目录）。点击浏览器右上角的下载图标可直接找到，也可按 Ctrl + J 打开下载列表定位。

如果有其他问题或需求，可以联系技术处理，技术微信：ljq101450