

KdanAuto 使用帮助

一、为什么要开发KdanAuto?

QA 做为 ComPDFKit_Conversion(以下统称 SDK) 交付品质把关最后一环，人力投入不足，无法满足SDK 内容 **随测**、以及每个版本的 **通测**，更加没有人力去做每次提测版本的全文档、各种情况覆盖测试，SDK 交付品质难以得到保障。同时，如果品质出现问题，整个团队又需要消耗掉更多的时间和精力去应对各种突发状况。

SDK 测试大致工作量如下：

- 转档SDK支持14种格式，每种格式又分别涉及注释、签名、表单等各种情况测试；
- 命令行工具各种参数组合保守估计多达300-400种，每一种参数组合按一个 PDF 文档（一页）算
- 工作量（不算转档实际耗时，）只算对比结果，每轮通测所需要的人力消耗大致为（文档转换 14种格式 x 100份文档 / 10小时工作时间 / 60分钟 = 2.3 人.天）
- 同时纯人工识别对照，还存在一定的失误率（例如一篇文章中，错了几个字，缺了几个字是无法肉眼察觉的）。

Note：随测和功能通测都是最耗费人力的

二、KdanAuto 能干什么？

KdanAuto 里面直接集成转换 SDK，能够对指定目录内的文档，进行一键式自动转档，并同步将转换出来的档案与前一个（SDK）版本进行预览结果对照、生成对照报告。给 QA 后续将要执行的新功能测试做粗过滤、做参考。

对于字符串文档格式，如 RTF、TXT，由于要对比的内容为纯字符，人工无法快速进行字符准确度精准对比，可以使用KdanAuto 工具提供的字符对比工具，进行逐个字符对比，并在测试报告中，给出字符对比效果。以利观察字符对照差异。

同时可以对【命令行工具 注释、高亮、加水印、提取等编辑功能自动化执行生成的结果文档】做对照，粗过滤出存在差异的内容，供 QA 精准测试。

具体功能有：

- KdaAuto对照工具 -> 导入文档到不同格式转档目录 -> 执行一键转档 -> 自动对照转档结果（与之前版本结果的差别） ->生成对照报告 (第三条，第3功能点中详细介绍)

需要将原PDF文档 导入到KdanAuto；（参考第三条，第3点）

- KdanAuto字符对比工具 -> 导入文档到不同格式转档目录 -> 执行一键转档/OCR(RTF、TXT) -> 自动对比转档结果（与对照文档的字符差别） ->生成字符对比报告 (第三条，第5功能点中详细介绍)

需要将原PDF文档、原TXT/RTF 对照文档导入到KdanAuto；
需要注意这2个目录（Origin Files、Check Files）内的文档结构、文档名字要一致

- 命令行对照工具->任意平台，命令行批量编辑（注释、加水印、高亮等） ->将结果文档导入 KdanAuto -> 一键（与之前版本结果进行）快照对照->生成对照报告 (第三条，第4功能点中详细介绍)

需要将原PDF文档、以及不同平台转档后的文档导入到KdanAuto；
需要注意这2个目录（Origin Files、Result Files）内的文档结构、文档名字要一致

- 跨平台字符对比工具->任意平台，批量转换/OCR RTF/TXT->将转档结果文档导入 KdanAuto -> 自动对比转档结果（与对照文档的字符差别） ->生成字符对比报告 (第三条，第6功能点中详细介绍)

需要将原PDF文档、原TXT/RTF 对照文档、以及不同平台转档/OCR后的文档导入到KdanAuto；
需要注意这3个目录（Origin Files、Check Files、Result Files）内的文档结构、文档名字要一致

- 可考虑作为工程师自测工具，确保开发的功能块，不会对其它的功能模块产生影响

三、具体有哪些功能？

1、新建测试项

以添加“TXT”格式测试项为例，右击左侧列表中的格式行，点击弹出菜单中的“新建转【TXT】测试项”，在弹出来的窗口中，设置以下属性：

- 测试项名称：用于测试项目的名称，仅用于阅读识别分类区分
- 测试项文件夹：对应于“(功能) 3”中，设置测试文件夹目录 中的子目录名称
- 测试类型：测试类型目前支持4种，快照对照测试，PDF 对照测试，字符对比测试、跨平台对比测试

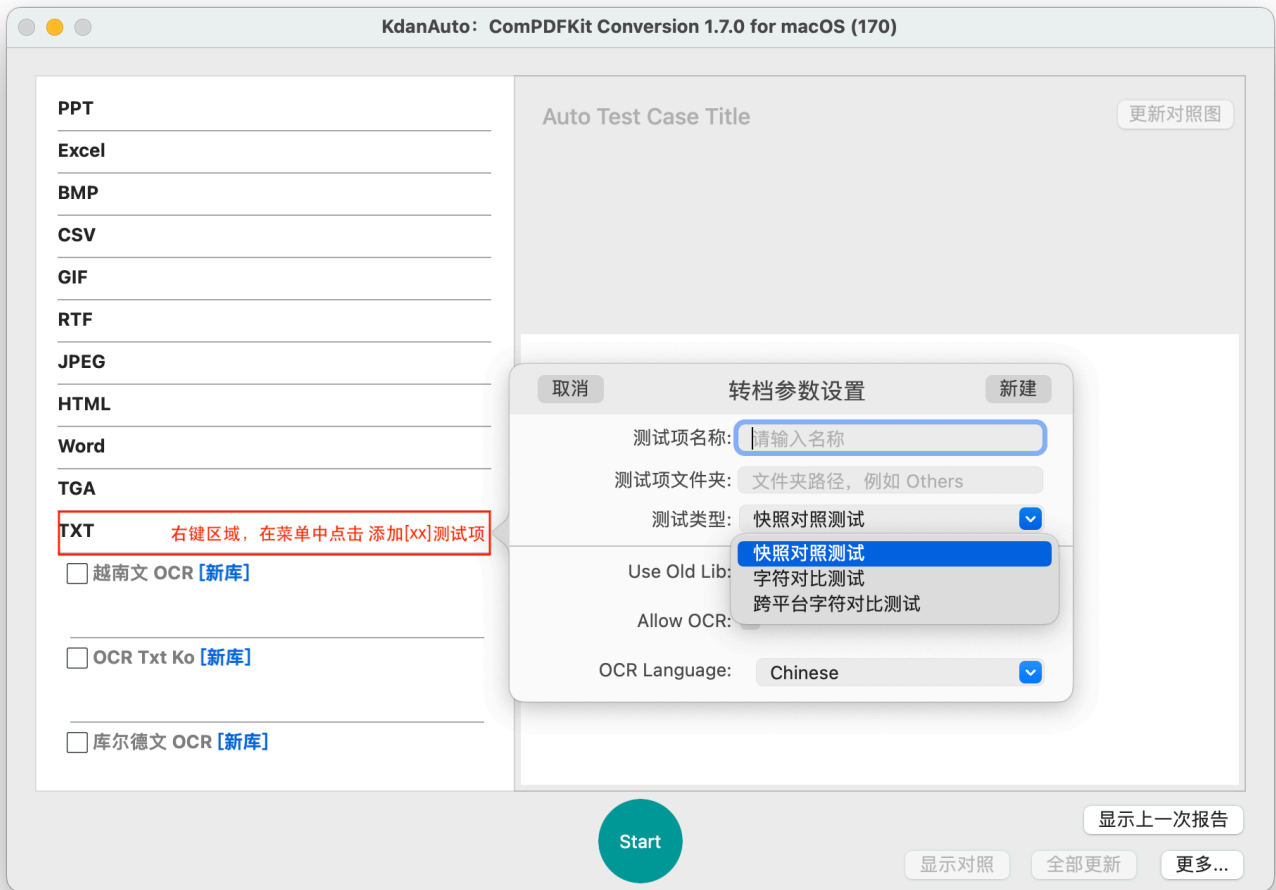
// 其它参数

- Use Old Lib： 是否是使用老库，用于切换 ComPDFKit 新库和之前徐建伟那套库
- Allow OCR： PDF转 TXT 格式时，转档SDK 中提供的参数设置，设置是否开启 OCR
开启 OCR 后，图片中的文字也会被识别出
- OCR Language： PDF转TXT 格式时，转档SDK 中对应指定的 OCR 识别语言

Note： 测试类型说明

- 快照对照测试
 - 对新（如1.7.0）/旧（如1.6.0）发布的转档SDK版本，转档/OCR出来的文档，进行快照，生成 JPG 图片（除JPG、PNG 等图片格式外，其他格式均只有一页），进行图片相素级对比，对生成文档差异进行粗过滤
 - 支持左侧列表中，除 PDF 格式外的所有格式类型
- PDF对照测试
 - 用于对其他平台（如 Windows、Android、Saas、命令行工具），以任意形式（如编辑、添加注释等）操作后的文档，导入，跟 前一次（做同样操作）处理后的文档，进行快照，生成 JPG 图片（全部页面），进行图片相素级对比，对生成文档差异进行粗过滤
 - 仅支持左侧列表 PDF 格式类型
- 字符对比测试
 - 用于对PDF 文档，进行 OCR/转档 生成 RTF/TXT 格式，再将生成的文档，与原始（RTF/TXT）文档，进行逐个字符对比，以检测出在转档/OCR 过程中转档/OCR生成字符的准确度
 - 仅支持 TXT、RTF 格式
- 跨平台对比测试
 - 用于其他平台（如 Windows、Saas 等），对PDF 文档进行 OCR/转档生成 RTF/TXT 格式，再将生成的文档，与原始测试范例素材资料（RTF/TXT）文档，进行逐个字符对比，以检测出在转档/OCR 过程中，对字符处理的准确度
 - 仅支持 TXT、RTF 格式

相关属性设置好后，点击右上角【新建】。（如下图）



2、编辑测试项

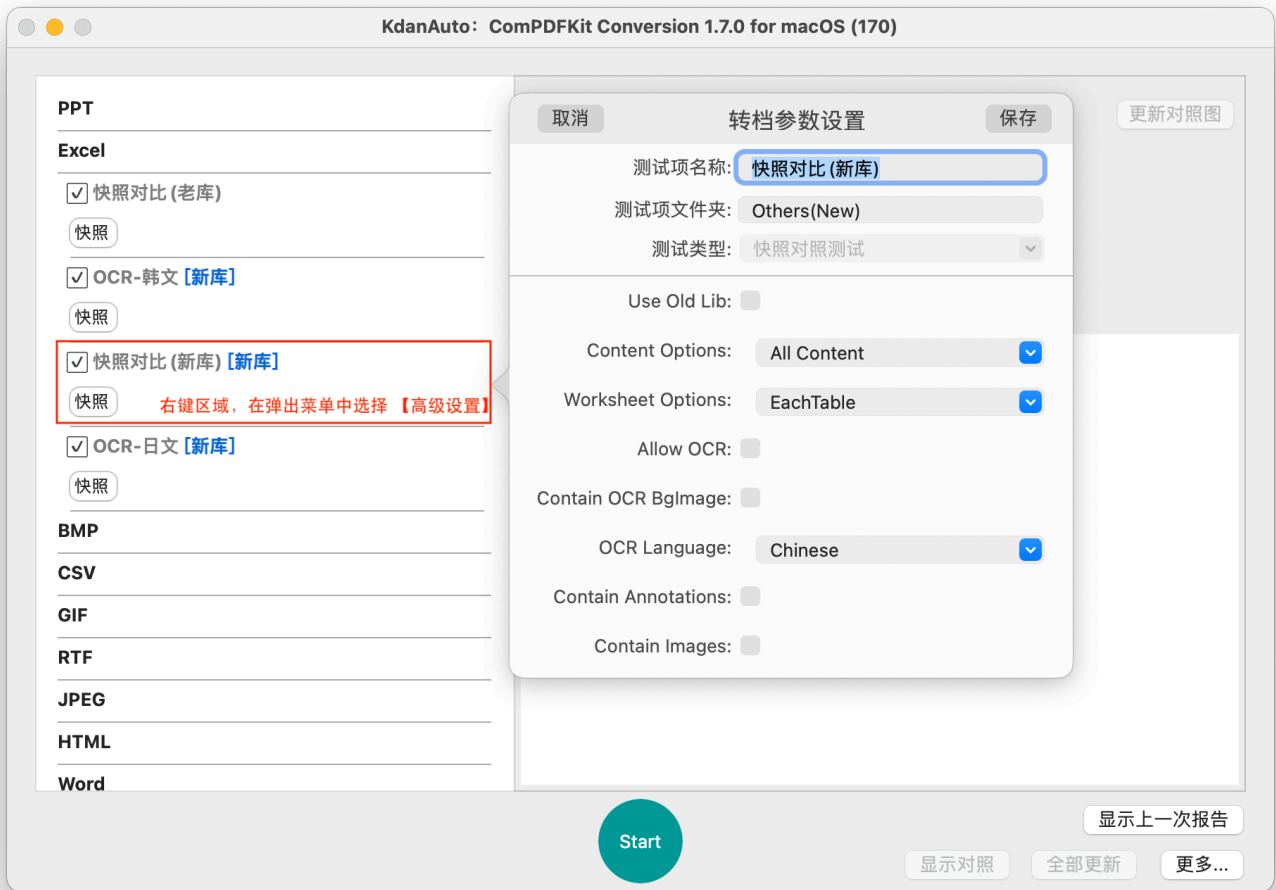
以修改“Excel”格式“快照对比(新库)”测试项为例，右击左侧列表中的 Excel 格式中的 '快照对比(新库)', 弹出菜单中的“高级设置”，在弹出来的窗口中，设置以下属性：

- 测试项名称：用于测试项目的名称，仅用于阅读识别分类区分，可修改
- 测试项文件夹：对应于“(功能) 3”中，设置测试文件夹目录 中的子目录名称，不可修改
- 测试类型：测试类型目前支持4种，快照对照测试，PDF 对照测试，字符对比测试、跨平台对比测试
 - 但 Excel 格式只支持 快照对照测试，所以当前示例不可修改

// 其它参数

- Use Old Lib: 是否是使用老库，用于切换 ComPDFKit 新库和之前徐建伟那套库
- Content Options: PDF转 Excel 格式时，SDK 中提供的参数，
参数具体解释及用法参考转档 SDK 文档
- Worksheet Options: PDF转 Excel 格式时，SDK 中提供的参数，
参数具体解释及用法参考转档 SDK 文档
- Allow OCR: PDF转 Excel 格式时，SDK 中提供的参数，是否开启 OCR识别，
参数更详细用法参考转档 SDK 文档
- Contain OCR BGIImage: PDF转 Excel 格式时，SDK 中提供的参数，OCR 是否包含背景图片，
参数更详细用法参考转档 SDK 文档
- OCR Language: PDF转 Excel 格式时，SDK 中对应指定的 OCR 识别语言，
参数更详细用法参考转档 SDK 文档
- Contain Annotations: PDF转 Excel 格式时，SDK 中提供的参数，是否包含 Annotations，
参数更详细用法参考转档 SDK 文档
- Contain Images: PDF转 Excel 格式时，SDK 中提供的参数，是否 包含图片，
参数更详细用法参考转档 SDK 文档

修改相关属性设置好后，点击右上角【保存】。（如下图）



3、自动化（转档）对照测试

自动化对照测试，用于将 PDF 自动化转换成对应目标格式，并将转出结果文件，同此前（版本）SDK转出结果进行截图对照，以确认当前版本，所修改/增加的内容，与此前（SDK）版本的差异。

目前ComPDFKit支持 PDF 转 PPT、Excel、BMP、CSV、GIF、RTF、JPEG、HTML、Word、TGA、TXT、JPEG2000、PNG、TIFF 14种。

新库支持PDF 转 PPT、Excel、BMP、CSV、JPEG、HTML、Word、TXT、PNG 9种格式。

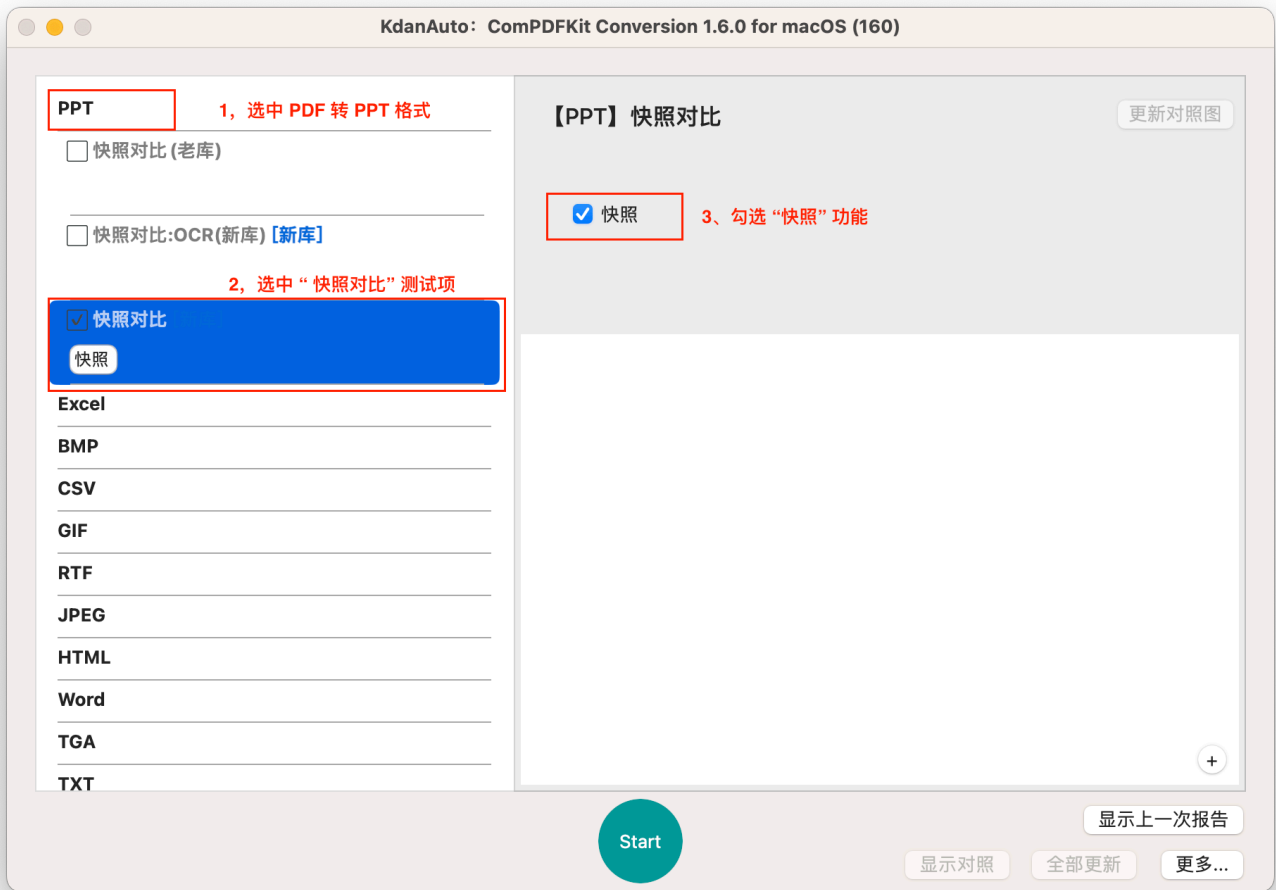
对于快照，PPT、Excel、BMP、CSV、HTML、Word、TXT 这7种格式只会截取第一页进行快照对比。BMP、JPEG、PNG、CSV、TGA、JPEG2000、TIFF 这7种 格式，会对每一页进行快照对照测试。

具体使用流程如下：

- S1：设置测试根目录（，执行任何测试功能前，均需要先设定好测试根目录；）例子中，以PDF 转【PPT】格式，【快照对比】测试项进行举例，新建【PPT】格式【快照对比】测试项



- S2: 选中对应测试项，勾选该测试项 【快照】 测试功能 勾选项



- S3: 点击 在右侧详情面板 右下角“+”按钮，进行 Origin Files 导入，建议（默认）以 '超链接' 形式导入（，能减小文档在 Mac 中的磁盘空间消耗）

2023-06-01 15.13.43

Favorites

- Recents
- Documents
- zhudongy...
- Applicati...
- Music
- Downloads
- Movies
- Desktop
- iCloud
- iCloud Dri...
- Shared
- Locations
- 朱东勇's...

Folder shared with File Sharing

- 表格
- 常规文档
- 大文档
- 动态表单
- 法文
- 韩文
- 交叉引用流
- 签名
- 权限问题
- 日文

01-精銳雲...515-封面.pdf

004-原有...施工方案.pdf

乱码_瑞慈...曹大林.pdf

企业综合信息表.pdf

人事異動單 (1).pdf

人員基本資...181016.pdf

外国居住者..._0327.pdf

AAB Loan A...lication.pdf

AF Lifestyle...tionaire.pdf

AF Nutritio...hecklist.pdf

2、选中要导入的文件

3、展开导入配置，并设置导入超链接，默认为超链接

4、点击 Open 添加

Hide Options

添加超链接

添加超链接

导入文件

Cancel

Open

PPT

- 快照对比 (老库)
- 快照对比:OCR(新库) [新库]
- 快照对比 [新库]

快照

Excel

BMP

CSV

GIF

RTF

JPEG

HTML

Word

TGA

TXT

【PPT】快照对比

更新对照图

快照

1、点击导入文件

Start

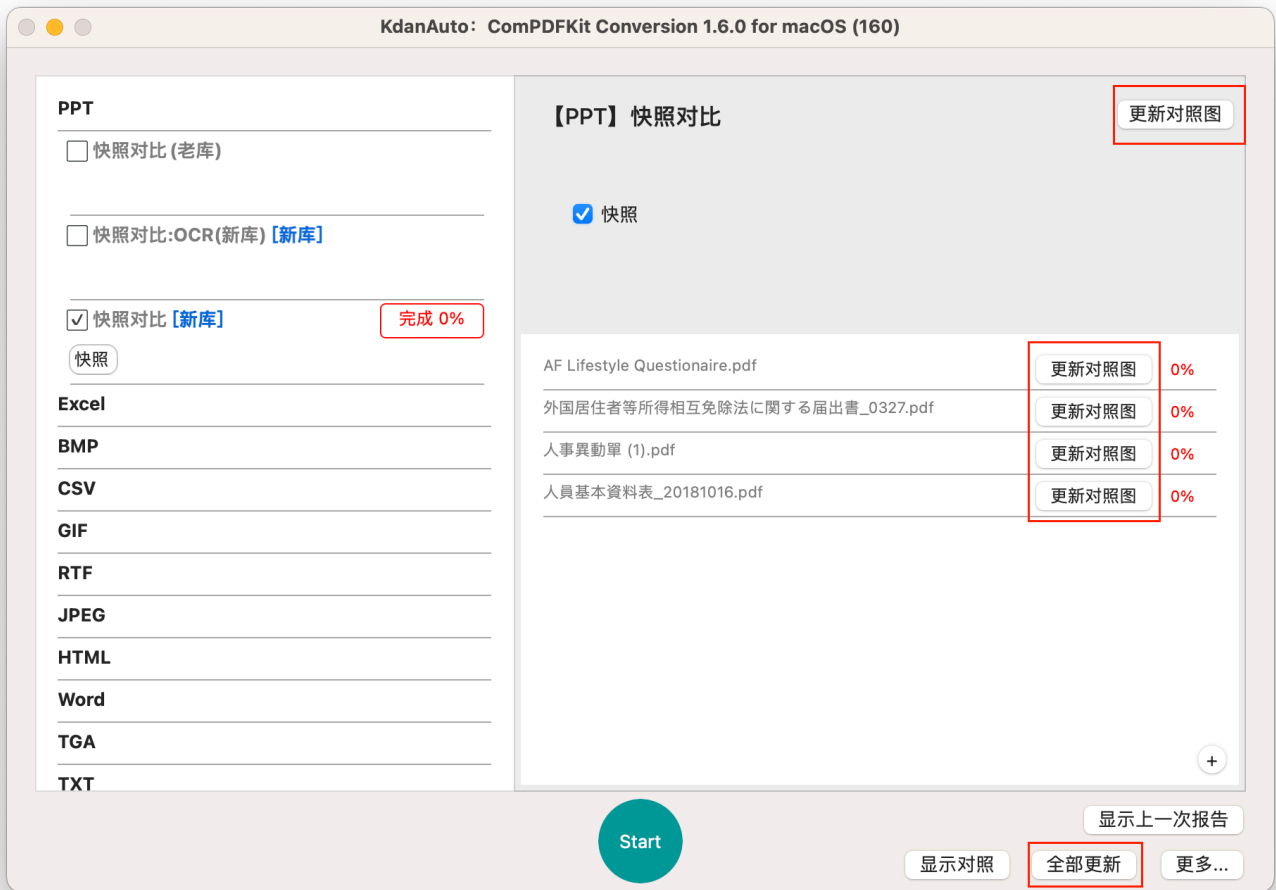
显示上一次报告

显示对照

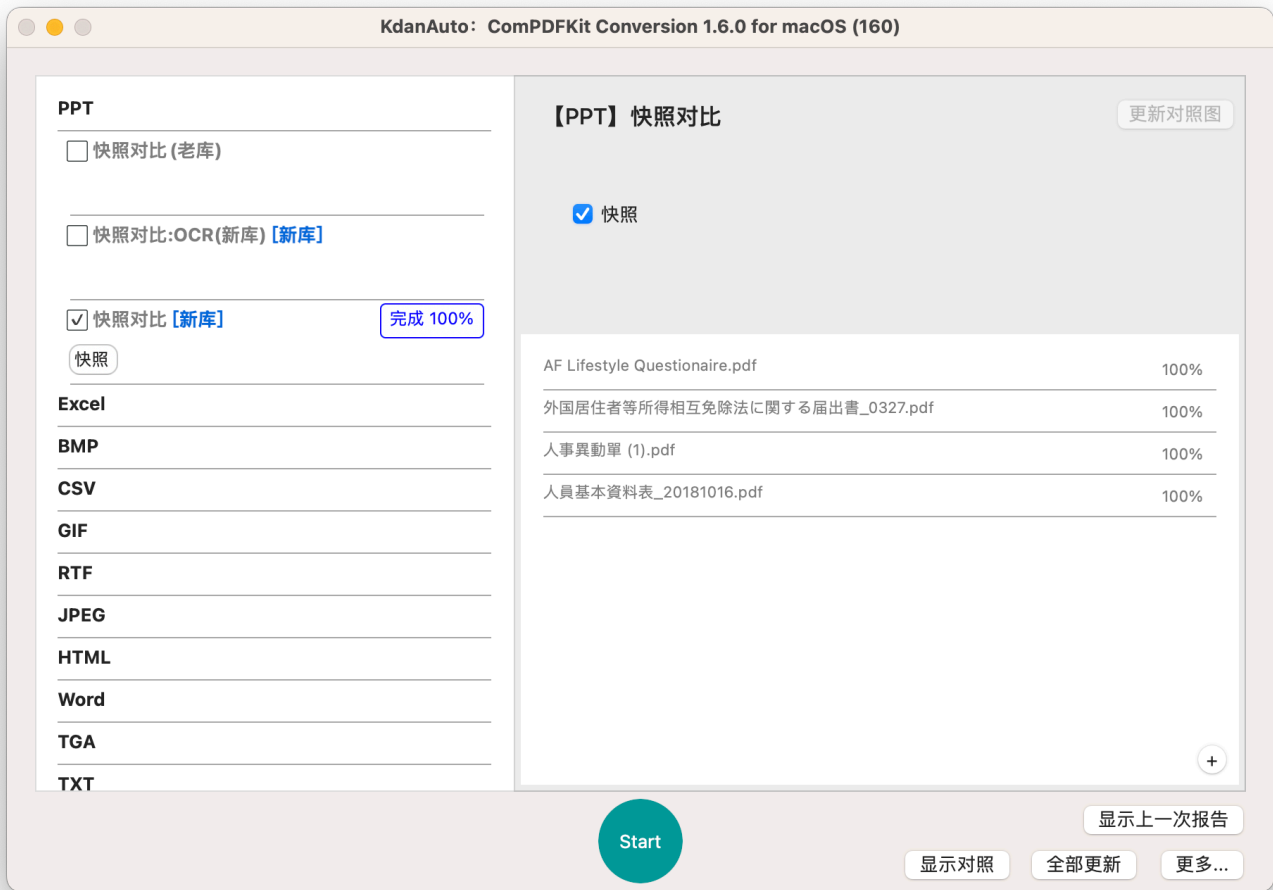
全部更新

更多...

- S4: 导入文件后，点击下方“Start”，启动转档 + 快照生成（+ 快照对照，由于是第一测试，没有对照图，所以对照结果为 0%）

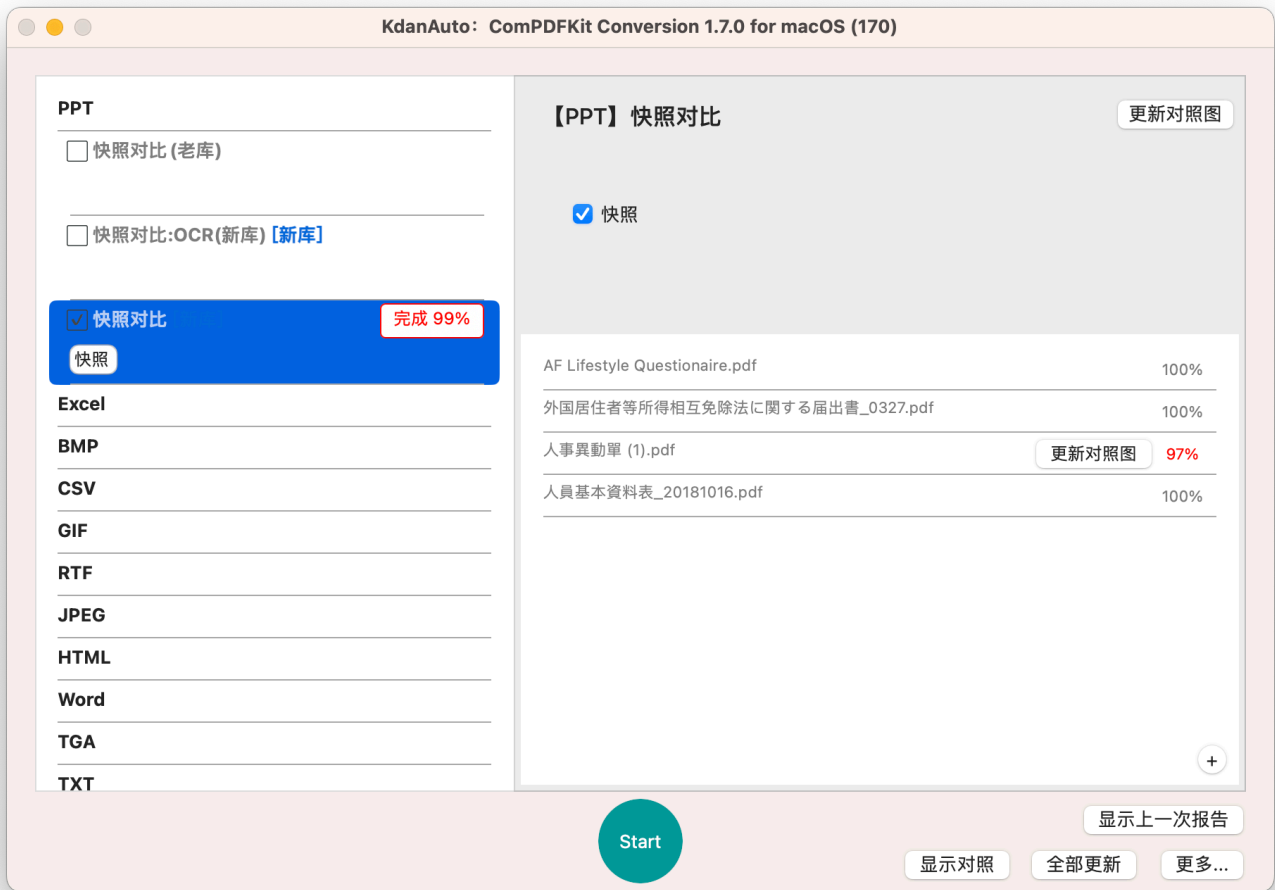


- S5: 点击详情面板 右上角“更新对照图”，将当前版本生成的快照作为对照图（，以后 SDK 有更新，会拿此次版本转档效果作为对照）
- S6: 再次点击下方“Start”，验证看 相同的文档（进行转档）、相同的转档 SDK，对照验证结果是否为 100%



- S7: (SDK 当前版本为 1.6.0,) 更新 SDK 版本到 1.7.0, 再次点击“Start”按钮, 启动转档 + 快照生成 + 快照对照, 验证 转档 SDK, V1.7.0 和 V1.6.0 版本 将相同的 PDF 文档, 转换成 PPT 文档, 结果差异。

需要 替换 SDK版本, 参见第三条, 功能7中, 切换 SDK 功能



- S8: 点击 测试项 右侧 或 详情面板中文件列表 右侧 “完成 xx%”， 按钮， 查看 自动化对照测试 图片 结果报告； 在 'S1' 中设置的根目录下面， 会产生一份 rtf 字符串报告

测试 图片 报告

人事異動申請單

申請人: _____ 申請日期: 年 月 日

異動種類	任用	升職	職調職	免職	復職	調薪	
	單位	單位					
職稱	職稱						
到任日期	年月日		到任日期	年月日			
區分	職等	級	本薪	生活津貼	職務加給	技術加給	合計
原薪							
調整薪							
...							
批示		複核	敬會單位	單位主管			

人事異動申請單

申請人: _____ 申請日期: 年 月 日

異動種類	任用	升職	職調職	免職	復職	調薪	留職停薪 異議
	單位	單位					
職稱	職稱						
到任日期	年月日		到任日期	年月日			
區分	職等	級	本薪	生活津貼	職務加給	技術加給	合計
原薪							
調整薪							
...							
批示		複核	敬會單位	單位主管			

更新对照图

100%

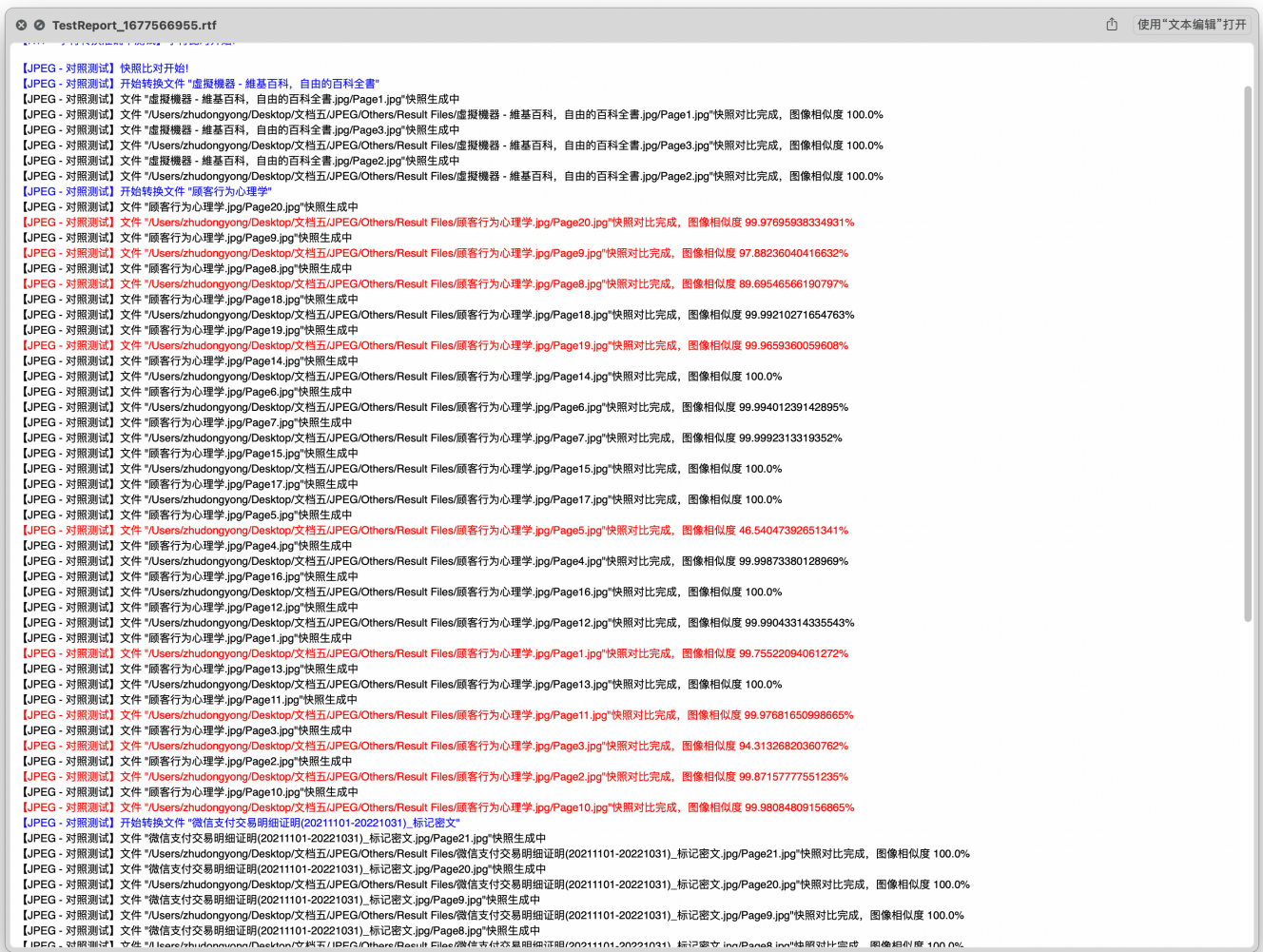
100%

97%

100%

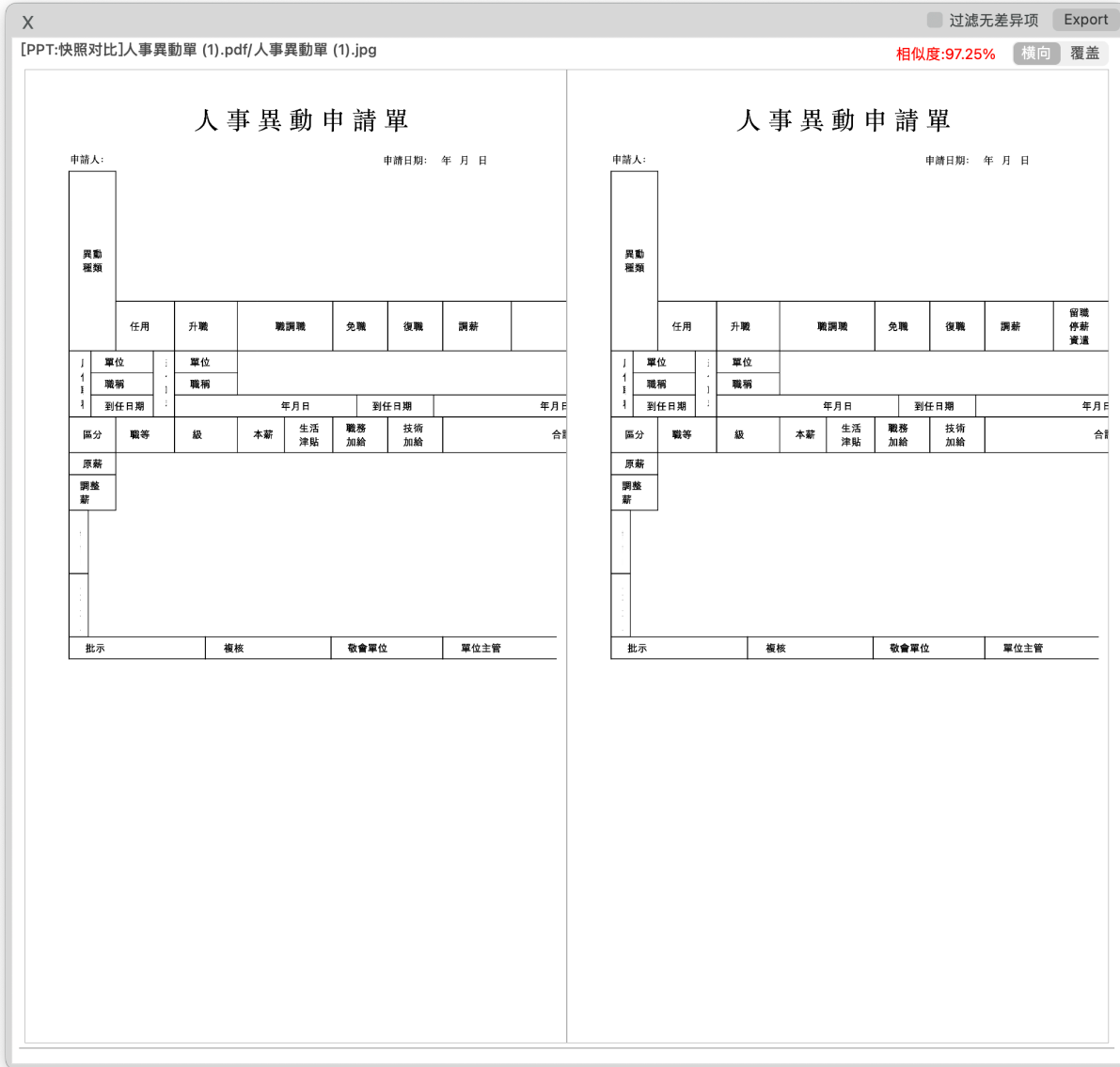
+

更多...

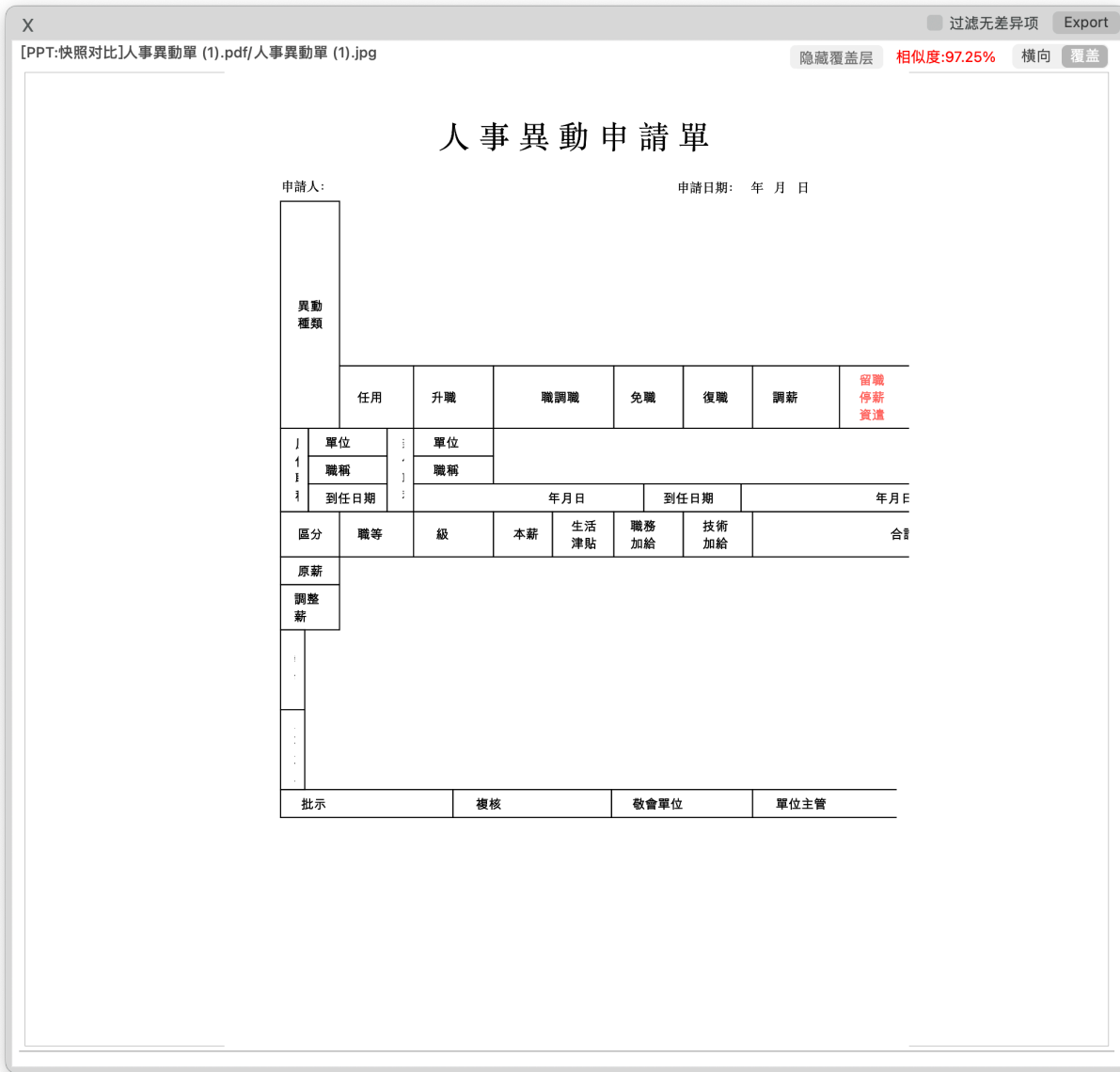


- S9: 在弹出的图片执行窗口中, 可分别进行 横向结果对比, 以及覆盖结果对比, 并可点击窗口右上角“Export”按钮, 导出 PDF 转 PPT 自动化对照测试结果;

测试结果横向对照



测试结果覆盖对照



4、PDF 格式对照测试（命令行编辑工具对照测试）

PDF 格式对照测试，用于将 其它平台或工具（如命令行编辑，Saas 在线编辑）生成的 PDF 文档，同之前（版本）SDK编辑导出的结果文档 进行逐页截图对照，以确认当前（命令行/Saas SDK）版本，所增加/修改的功能效果 和 之前（命令行/Saas SDK）版本功能效果的差异。

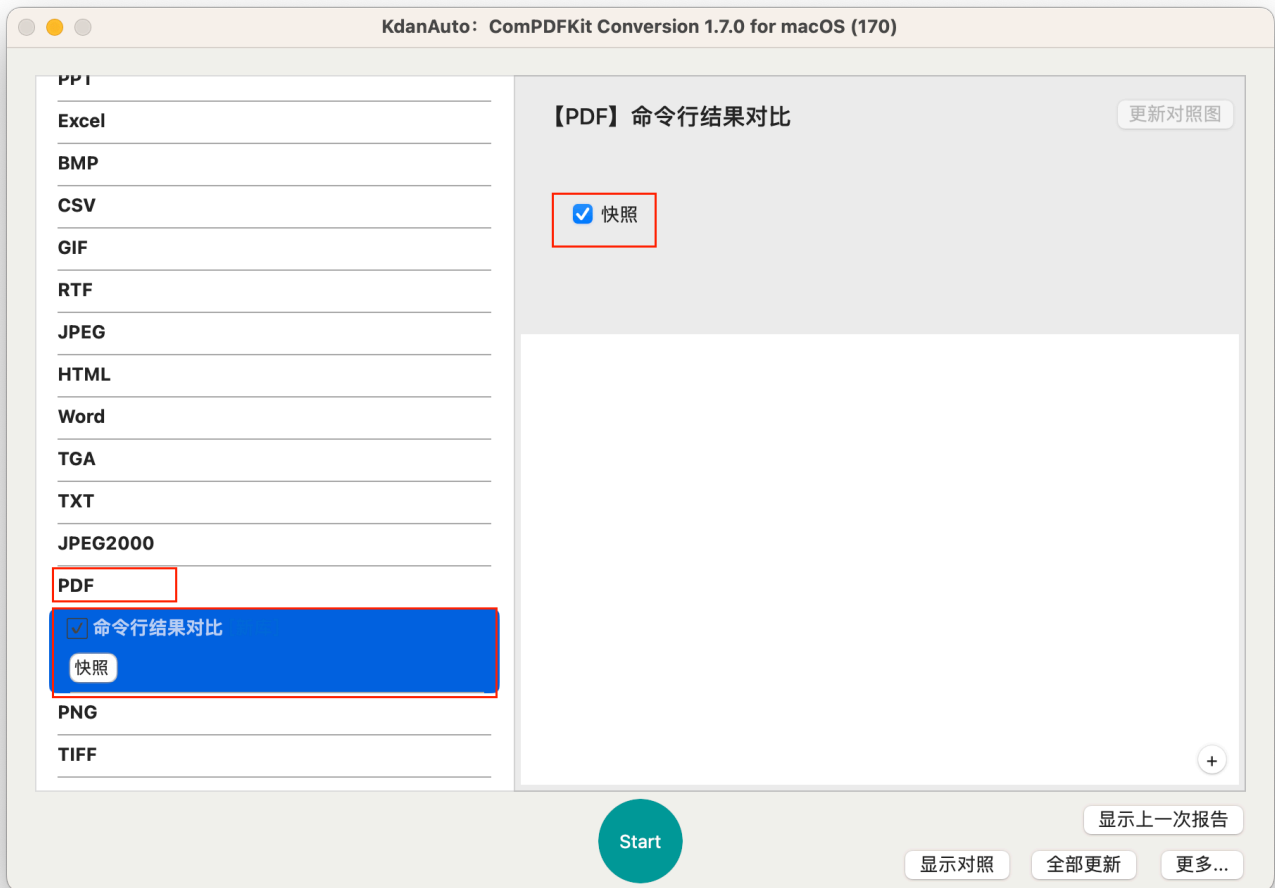
目前 PDF 格式对照测试，原文档 *Origin Files* 和 *Result Files* 都需要从外部导入，会对 PDF 文档每一页进行快照对照测试。

具体使用流程如下：

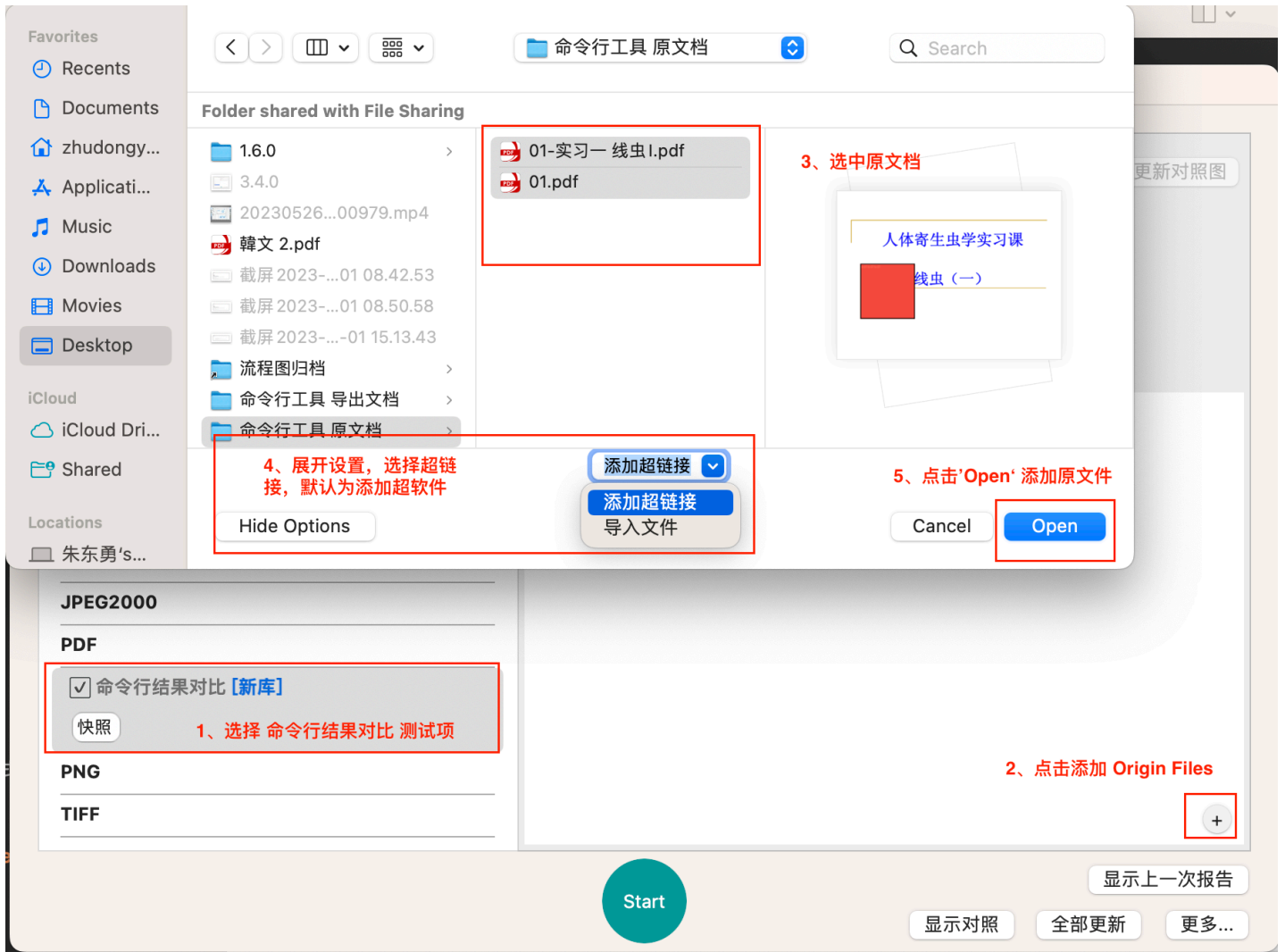
- S1：设置测试根目录（，执行任何测试功能前，均需要先设定好测试根目录；）例子中，以PDF 类型【命令行结果对比】测试项进行举例，新建【PDF】格式【命令行结果对比】测试项

- 设置根目录参考 【自动化对照测试】流程中的 S1 步骤；
- 新建 【PDF】格式【命令行结果对比】测试项，参考 第三条，功能1 【新建测试项】

- S2: 选中对应测试项，勾选该测试项【快照】测试功能 勾选项

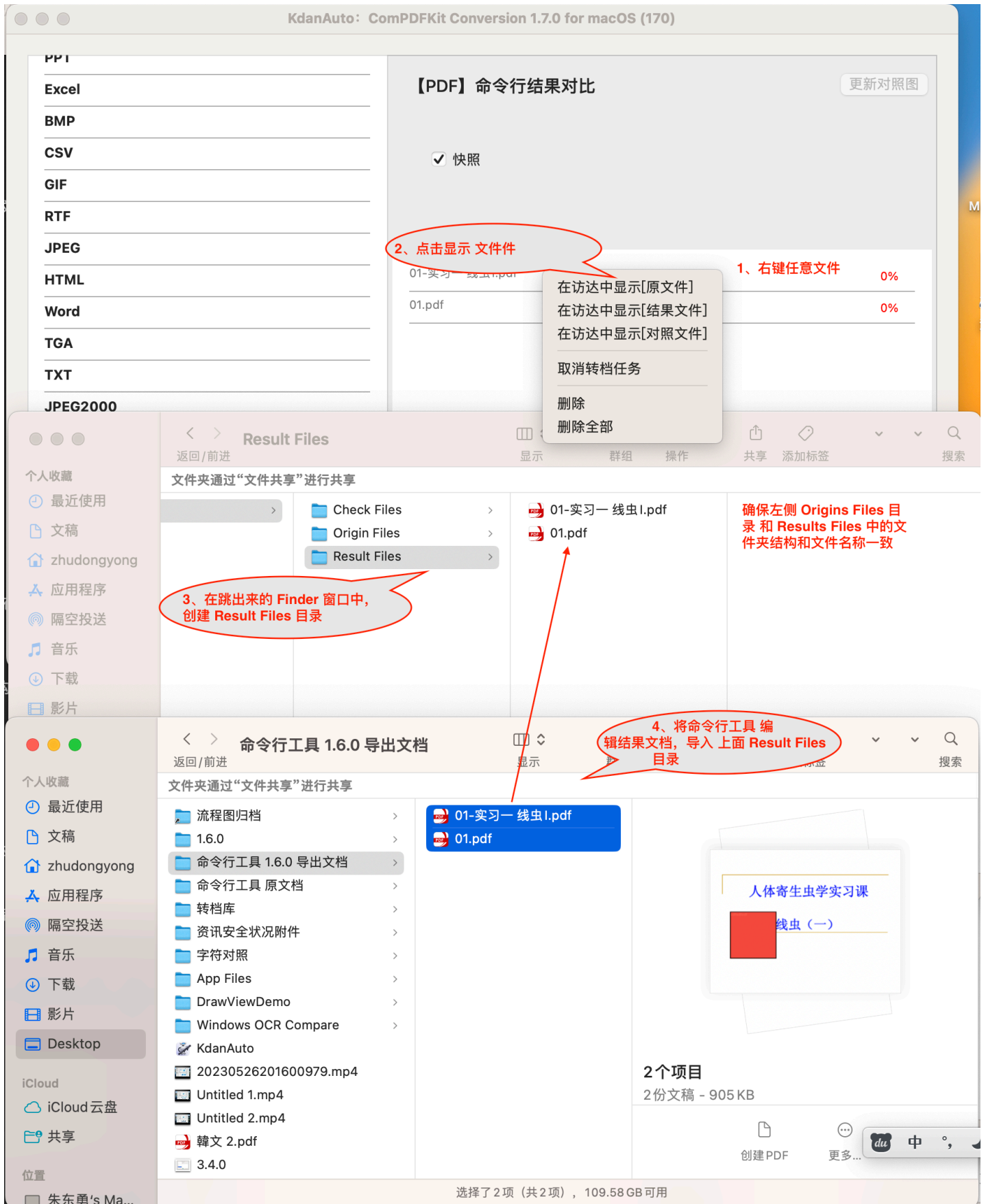


- S3: 点击 在右侧详情面板 右下角“+”按钮，进行 **Origin Files** 导入，建议（默认）以 **超链接** 形式导入（，能减小文档在 Mac 中的磁盘空间消耗）



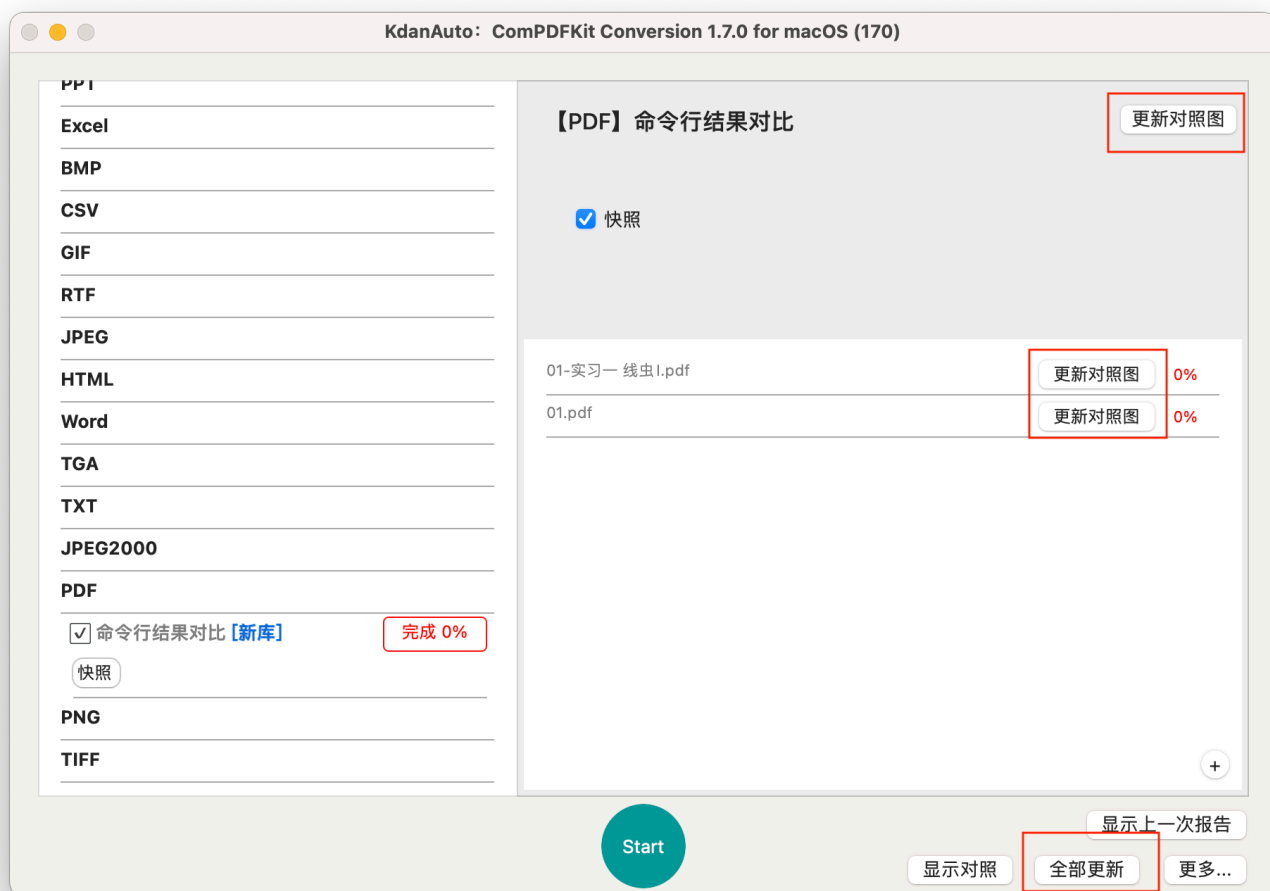
- S4: 导入 **命令行工具** 版本 (如 1.6.0) 编辑生成的结果文档，到 **Result Files** 目录中，留意跟 **Origin Files** 目录中的文件结构和文件名保持一致

在右侧详情面板中，选择一个文件，然后右键，点击菜单中的 '在访达中显示[原文件]'，会打开对应的文件目录，在 Origin Files 同级目录中，创建 Result Files，再将 '命令行工具' 编辑结果文件导入 Result Files 目录

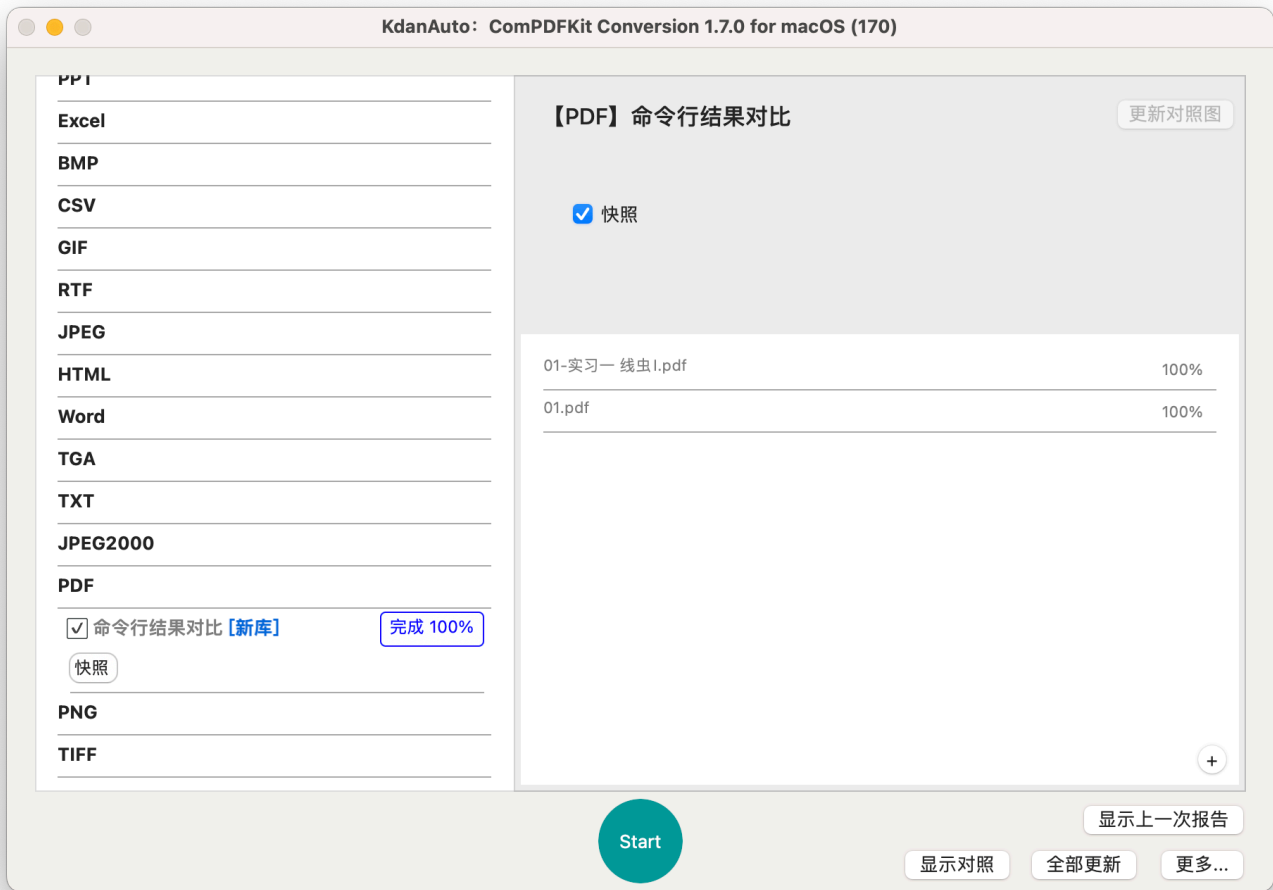


- S5: 导入文件后, 点击下方“Start”, 启动转档成 JPG + 快照生成 (+ 快照对照, 由于是第一测试, 没有对照图, 所以对照结果 均为 0%)
- S6: 点击详情面板 右上角“更新对照图”, 将当前版本生成的快照作为对照图 (以后 替换 命令行工具 新

版本<如 1.7.0> 编辑生成的结果文档，再次进行转档成 JPG + 快照生成 + 快照对照，将能获得 相同 PDF 文档、相同转档 SDK、不同命令行工具 编辑产生的结果文件，与对照文档进行对照测试的结果情况)



- S7: 再次点击下方“Start”，验证看 (同一份 命令行工具编辑生成的)相同的文档进行逐页快照对照验证，看结果是否均为 100%



- S8: 参考 S4 步骤, 将**Result Files** 目录中的文件, 替换成新版本 (如1.7.0) 命令行工具编辑导出的 PDF 文档, 再次点击“Start”按钮, 启动转JPG + 快照生成 + 快照对照, 验证不同命令行工具版本, 编辑生成的文档, 用同一个转档 SDK 转成 JPG, 再进行逐页对比得出的结果差异。
- S9: 点击 测试项 右侧 或 详情面板中文件列表 右侧 “完成 xx%”, 按钮, 查看 PDF对照测试 图片 结果报告; 在 'S1' 中设置的根目录下面, 会产生一份 rtf 字符串报告

具体报告效果及报告位置参考 【自动化对照测试】流程中的s8步骤;

- S10: 在弹出的图片结果中, 可分别进行 横向结果对比, 以及覆盖结果对比, 并可点击窗口右上角 “Export” 按钮, 导出 PDF对照测试结果;

具体报告效果及报告位置参考 【自动化对照测试】流程中的s9步骤;

5、(PDF转RTF/TXT 文档)字符对比测试

PDF转RTF/TXT 文档)字符对比测试, 简称 字符对比测试, 用于将 PDF 自动化转换/OCR识别成 RTF/TXT 格式, 再用转换(/OCR识别)出的 RTF/TXT 结果文档, 同标准 (无差错) 的 RTF/TXT 对照文档 进行逐字符对比, 以确认当前转档/OCR SDK版本, 对于PDF 文档中的字符 进行转档/OCR识别处理的准确度。

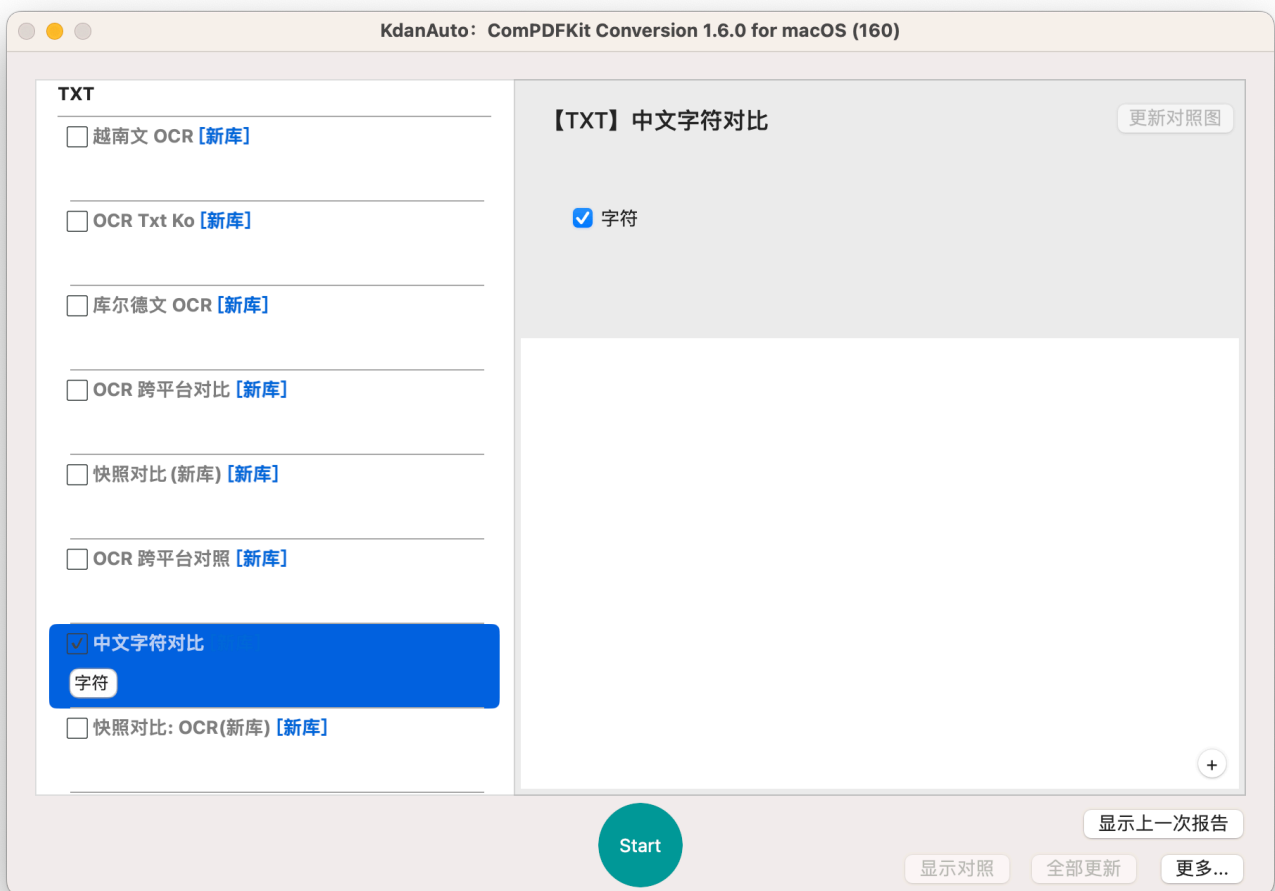
目前字符对比支持 PDF 转 RTF、TXT 2种，其中 RTF（新版本 SDK 导出的）框排 和（徐建伟那套老版本 SDK 导出的）流排模式均支持。同时，结果文档中的空格和换行将过滤不做识别，但（目前）部分空格可能会影响到准确率计算；在导出的对比报告中，会完整呈现字符排版；

具体使用流程如下：

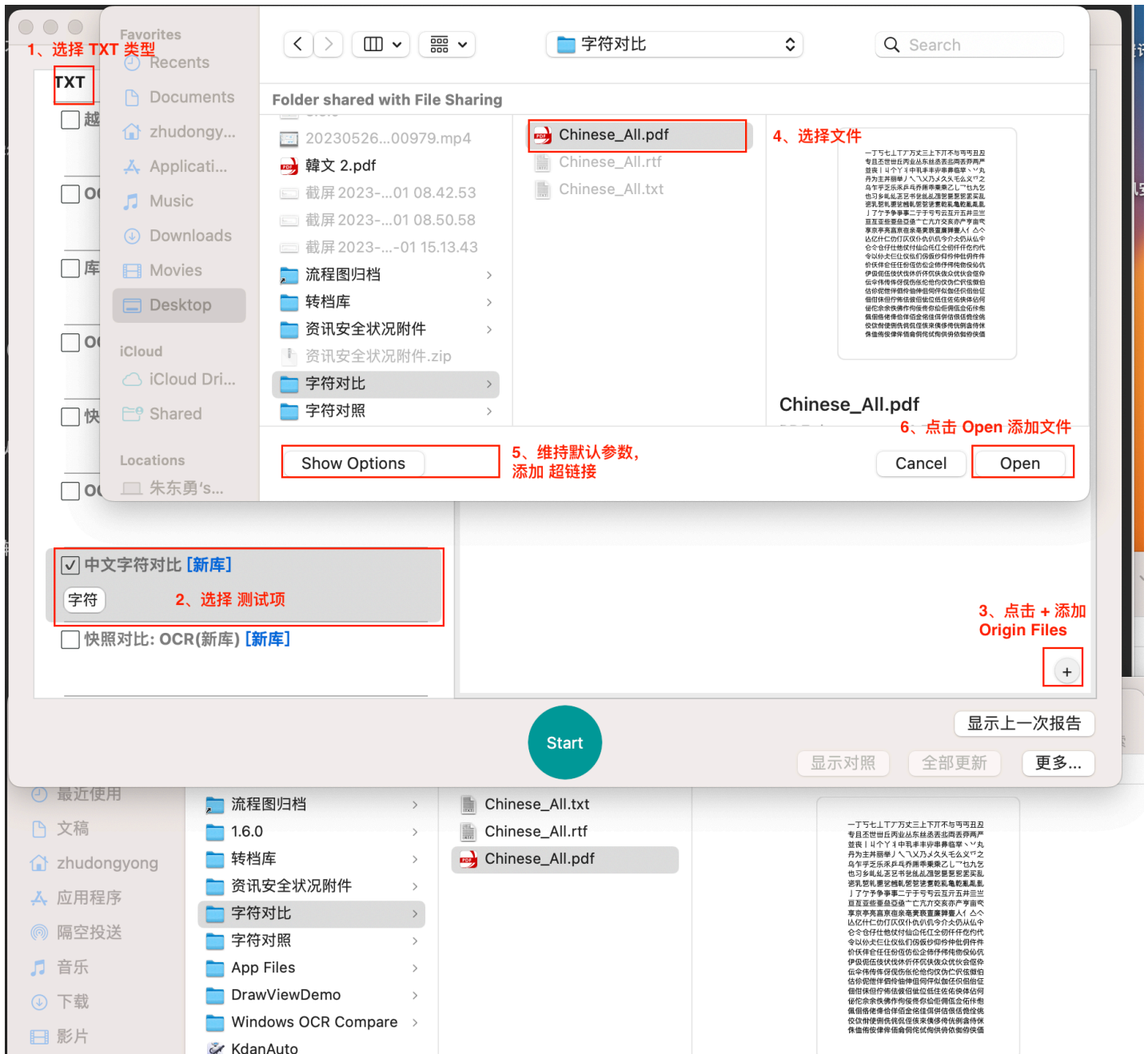
- S1：设置测试根目录（，执行任何测试功能前，均需要先设定好测试根目录；）例子中，以PDF 转【TXT】格式，【中文字符对比】测试项进行举例，新建【TXT】格式【中文字符对比】测试项

测试项参数设置中，测试类型选择【字符对比测试】，相关设置参考 第三条，功能1、功能2

- S2：选中对应测试项，勾选该测试项【字符】测试功能 勾选项



- S3：点击 在右侧详情面板 右下角“+”按钮，进行 Origin Files 导入，建议（默认）以 '超链接' 形式导入（，能减小文档在 Mac 中的磁盘空间消耗）



- S4: 导入原始精准的 对照 RTF/TXT文档（示例中是 TXT 文档），到 **Check Files** 目录中，留意跟 **Origin Files** 目录中的文件结构和文件名保持一致

在右侧详情面板中，选择一个文件，然后右键，点击菜单中的 ‘在访达中显示[原文件]’，会打开对应的文件目录，在 **Origin Files** 同级目录中，创建 **Check Files**，再对照 RTF/TXT文档（示例中是 TXT 文档）导入 **Check Files** 目录



- S5: 导入文件后, 点击下方“Start”, 启动转档/OCR识别 + 字符对比, 将得到相应文件, 转换/OCR识别出来的文档 (示例中是 TXT 文档) 中字符处理精准度, 以及具体哪些字符有出错状况
- S6: 点击测试项右侧或详情面板中文件列表右侧“完成 xx%”, 按钮, 预览查看字符对比结果报告; 在 'S1' 中设置的根目录下面, 会产生一份 rtf 字符串报告

[TXT:中文字符对比]Chinese_All.pdf
一丁弓七上丁万文三上下开不与丐丐丑丑且专且世世出丘丙业从生丝丞去北两舌丙两并严並...
卍之卍之互乐承互互承互乘互乘互七乙七乙七乙也习山乱七乙七乙也习山乱七乙七乙也习山乱七乙七乙也习山乱...



90%

上一次报告

更多...



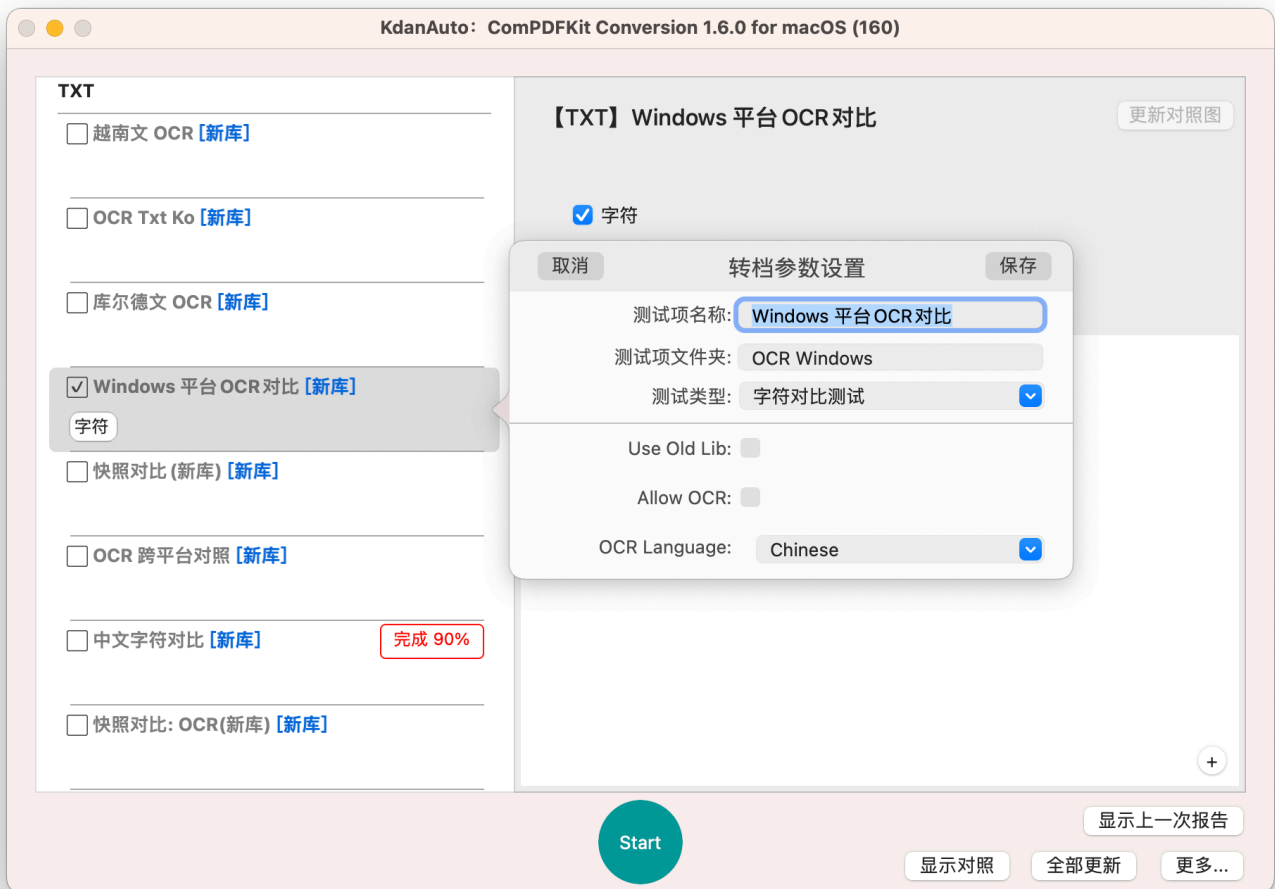
- S7: 在弹出的对比结果预览窗口中，并可点击窗口右上角“Export”按钮，导出 字符对比测试报告；

具体使用流程如下：

- S1：设置测试根目录（，执行任何测试功能前，均需要先设定好测试根目录；）例子中，以PDF 转【TXT】格式，【Windows平台 OCR 对比】测试项进行举例，新建【TXT】格式【Windows平台 OCR 对比】测试项

测试项参数设置中，测试类型选择【跨平台字符对比测试】，相关设置参考 第三条，功能1、功能2 其它参数勾选不起作用，因为不走本地转档

- S2：选中对应测试项，勾选该测试项【字符】测试功能 勾选项

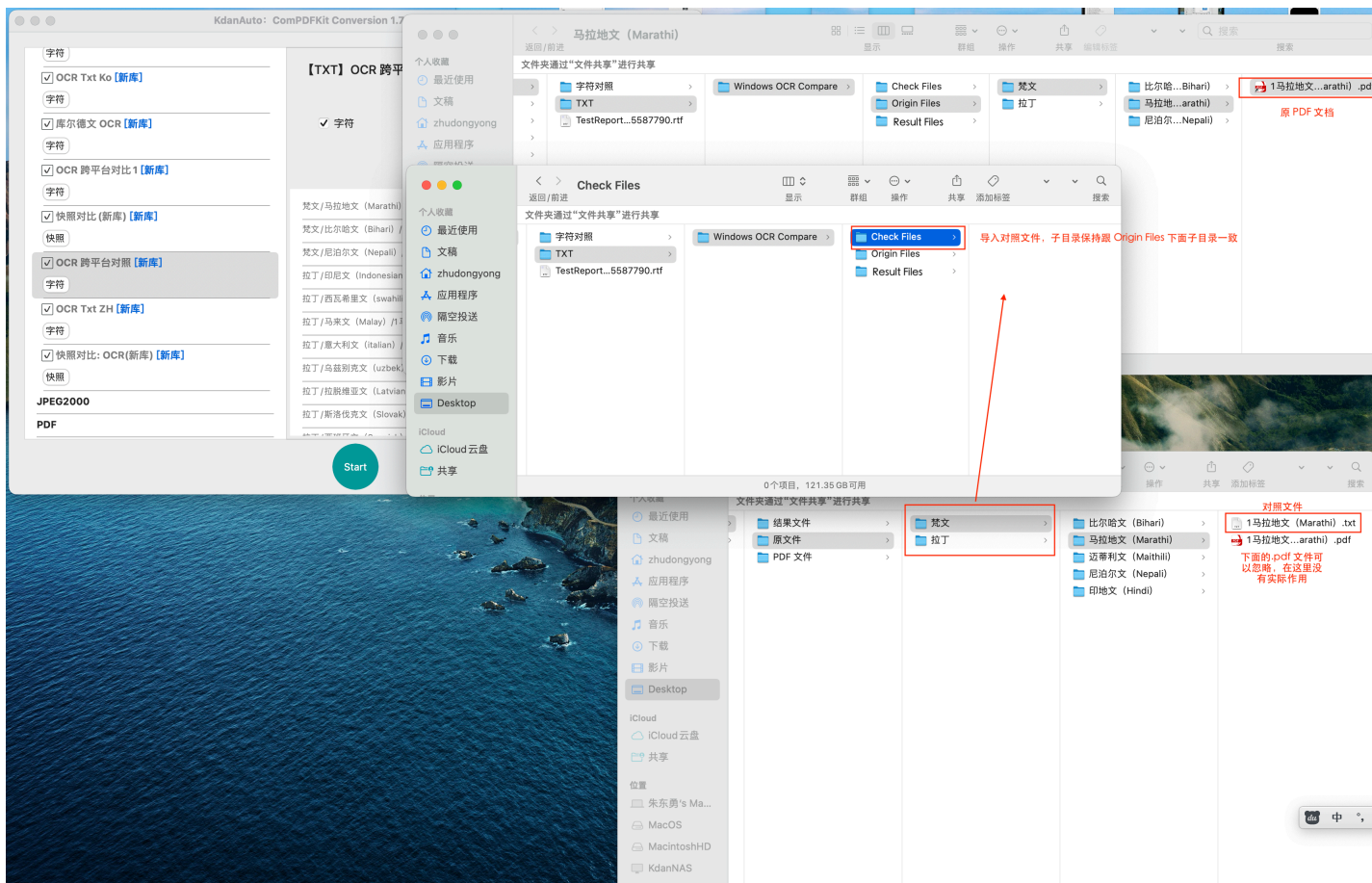


- S3：点击 在右侧详情面板 右下角“+”按钮，进行 Origin Files 导入，建议（默认）以 '超链接' 形式导入（，能减小文档在 Mac 中的磁盘空间消耗）



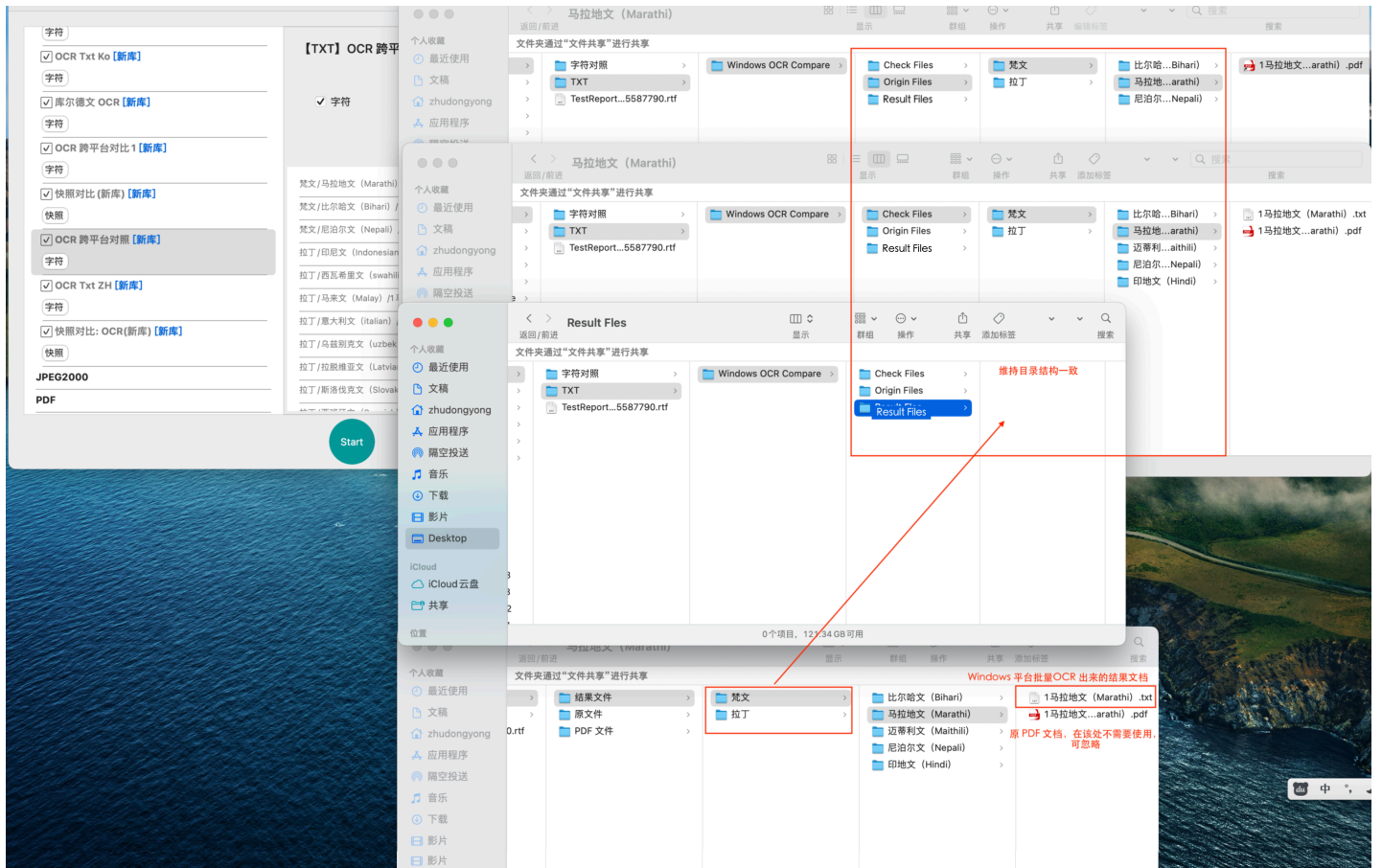
- S4: 导入原始精准的 对照 RTF/TXT文档（示例中是 TXT 文档），到 **Check Files** 目录中，留意跟 **Origin Files** 目录中的文件结构和文件名保持一致

在右侧详情面板中，选择一个文件，然后右键，点击菜单中的‘在访达中显示[原文件]’，会打开对应的文件目录，在 Origin Files 同级目录中，创建 Check Files，再对照 RTF/TXT文档（示例中是 TXT 文档）导入 Check Files 目录



- S5: 在每次对比测试前, 导入从其它平台批量转换/OCR 识别得到的 RTF/TXT 文档 (此处是 TXT 文档), 到 **Result Files** 目录中, 留意跟 **Origin Files**、**Check Files** 目录中的文件结构和文件名保持一致

在右侧详情面板中, 选择一个文件, 然后右键, 点击菜单中的 '在访达中显示[原文件]', 会打开对应的文件目录, 在 Origin Files 同级目录中, 创建 Result Files, 再对照 RTF/TXT 文档 (示例中是 TXT 文档) 导入 Result Files 目录 (如果之前旧版本的文件存在, 则先删除 Result Files 目录中的文件)



- S5: 导入文件后, 点击下方“Start”, 启动 字符对比, 将得到 Windows平台 导入的文档, 字符转换/OCR识别处理的精准度, 以及具体哪些字符有出错状况
- S6: 点击 测试项 右侧 或 详情面板中文件列表 右侧“完成 xx%”, 按钮, 预览查看 跨平台字符对比 结果报告; 在 'S1' 中设置的根目录下面, 会产生一份 rtf 跨平台字符串报告



- S7: 在弹出的对比结果预览窗口中，并可点击窗口右上角“Export”按钮，导出跨平台字符对比测试报告；

Med dette i tankene er livsmening forskjellig for forskjellige mennesker og avhenger av hva som er viktig for dem som individer og hva de søker etter. I denne konteksten kan forskjellige personer gjenkjenne forskjellige formål og engasjementer og bygge livet sitt rundt disse ideene.

【TXT - Windows 平台OCR对比】文件 "拉丁/挪威文 (norwegian) /1挪威文 (norwegian) " 比对完成, 准确率100.0%(645/645)

【TXT - Windows 平台OCR对比】开始转换文件 "拉丁/罗马尼亚文 (Romanian) /1罗马尼亚文 (Romanian) "
"Wu Song lovește tigru!" este una dintre cele mai faimoase povești din romanul chinezesc din dinastia Ming, "Povestea din Suho". Iată principalele detalii ale poveștii:

Contextul poveștii: La începutul dinastiei Ming, în orașul Jiangzhou (în prezent orașul Nanchang din provincia Jiangxi), exista un tigru fioros, cunoscut pentru răufăcătorile sale care bântuia prin munții și pădurile din afara orașului, rânind mulți oameni, populația era terorizată.

Introducerea personajului principal - Wu Song: Wu Song era un mare luptător din Jiangzhou, el fiind un bărbat inteligent și curajos cu o forță fizică deosebită. El a auzit unul dintre zvonurile despre tigrul și a decis să se ducă să lupte cu el.

Procesul de lovire a tigrului: Wu Song a pregătit câteva coarde și o groapă cu un animal (o scroafă) legat de un stâlp de lemn înăuntru, pentru a atrage tigrul într-o capcană. Când tigrul a intrat în capcană și a fost prins cu corzi, Wu Song a profitat de ocazie să se apropie, să sară peste gard, să-i dea trei pumni, lovindu-l până l-a omorât.

Urmările luptei cu tigrul: Wu Song a transportat tigrul mort înapoi în oraș, a raportat autorității și a primit recompense și faimă. Acest eveniment a devenit subiectul discuțiilor din Jiangzhou, iar Wu Song a devenit celebru în întreaga lume.

În general, "Wu Song lovește tigru!" este o parte din literatura clasică chineză și este una dintre poveștile tradiționale ale poporului chinez despre eroi, reflectând valorile chinezești de curaj, dreptate și înțelepciune.

【TXT - Windows 平台OCR对比】文件 "拉丁/罗马尼亚文 (Romanian) /1罗马尼亚文 (Romanian) " 比对完成, 准确率 97.79179382324219%(1470/1516)

【TXT - Windows 平台OCR对比】开始转换文件 "拉丁/法文(french)/1法文 (french) "
Testing requiert une préparation dans les domaines suivants:

Plan de test et stratégie: Le plan de test et la stratégie sont la première étape dans le processus de test et contrôlent l'ensemble du processus de test, y compris l'objectif du test, les méthodes de test, la sélection, l'exécution et l'évaluation du test suite et l'utilisation d'outils de test, etc. Une planification de test complète et claire contribue à un test efficace et précis.

Environnement de test: L'environnement de test est le cadre et l'équipement pour effectuer le test et doit correspondre à l'environnement de production réel, y compris le matériel, les logiciels, le réseau, les données, etc. Pour minimiser les résultats de test imprécis en raison de différences environnementales ou de tests destructifs, il est nécessaire de fournir une simulation ou un environnement de test.

Données de test: Les données de test sont la base et la condition pour le processus de test. Les données de test doivent être actuelles, valides et conformes à l'objectif de test. Par conséquent, avant de commencer le test, une quantité et une qualité de données de test appropriées doivent être préparées pour garantir la précision et la largeur du test.

Équipe de test et personnel: L'équipe de test et le personnel comprennent des gestionnaires de test, des ingénieurs de test, des ingénieurs d'automatisation de test, des analystes de test, etc. Ils doivent effectuer le test conformément au plan de test et à la stratégie et avoir des compétences techniques et une expérience professionnelle.

Outils de test: Différents outils de test sont utilisés dans les tests pour aider le personnel de test à améliorer l'efficacité et la précision. Par exemple, un outil de test de charge doit être utilisé pour les tests de performance et un outil de test automatisé doit être utilisé pour les tests d'automatisation.

En préparant soigneusement le plan de test, l'environnement de test, les données de test, l'équipe de test et les outils de test, la précision et l'efficacité du test peuvent être maximisées, et les résultats de test peuvent donc être plus fiables et efficaces.

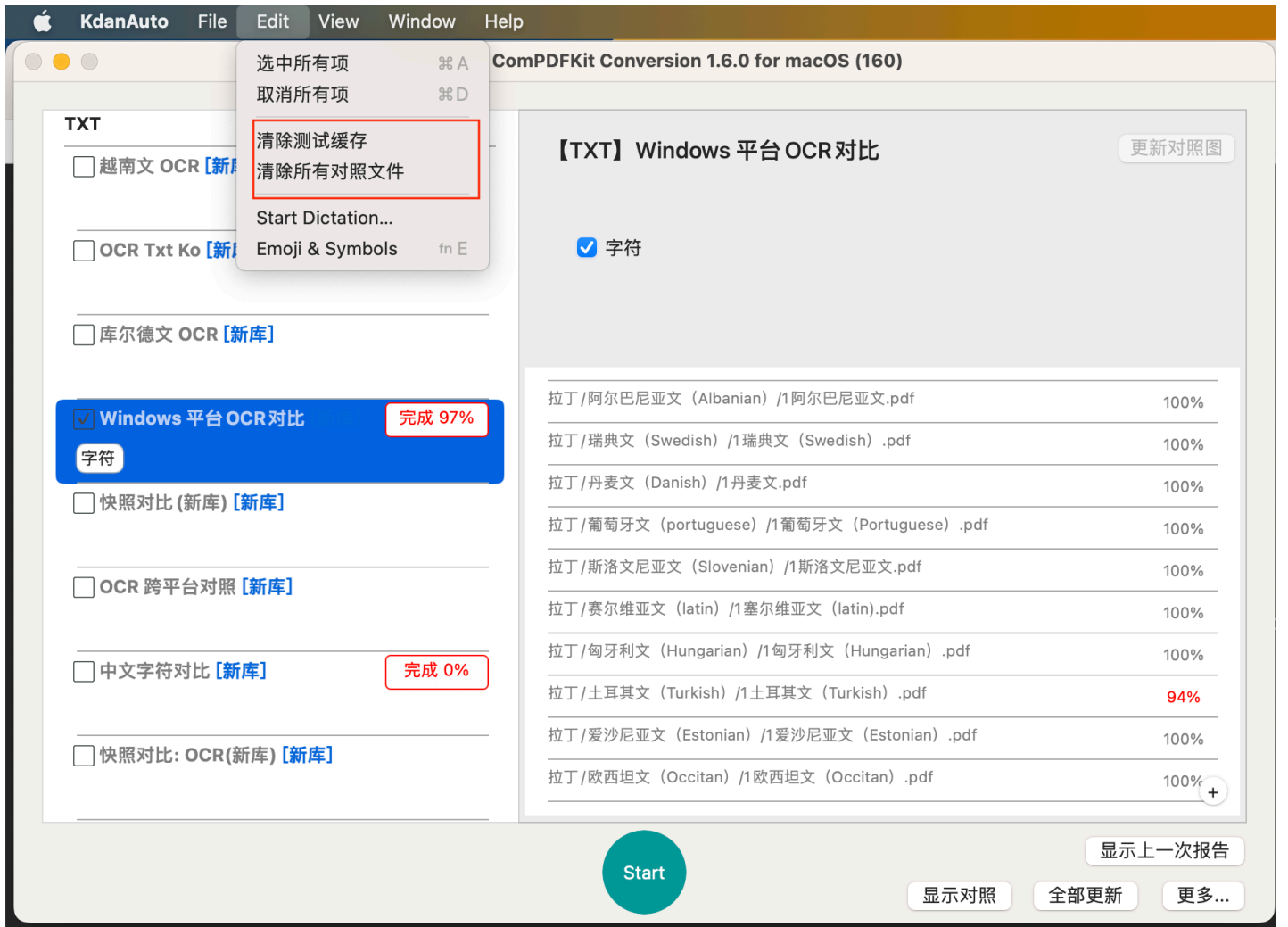
【TXT - Windows 平台OCR对比】文件 "拉丁/法文(french)/1法文 (french) " 比对完成, 准确率100.0%(2141/2141)

7、其它附加功能

【a】测试项批量勾选和取消勾选



【b】 Origin Files、Result Files、Check Files 清除操作



TXT

越南文 OCR [新库]

OCR Txt Ko [新库]

库尔德文 OCR [新库]

Windows 平台 OCR对比

完成 97%

字符

快照对比 (新库) [新库]

OCR 跨平台对照 [新库]

中文字符对比 [新库]

快照对比: OCR(新库) [新库]

高级设置

清除原文件
清除结果文件
清除对照文件

删除测试项

完成 0%

【TXT】Windows 平台 OCR对比

更新对照图

字符

拉丁/阿尔巴尼亚文 (Albanian) /1阿尔巴尼亚文.pdf	100%
拉丁/瑞典文 (Swedish) /1瑞典文 (Swedish) .pdf	100%
拉丁/丹麦文 (Danish) /1丹麦文.pdf	100%
拉丁/葡萄牙文 (portuguese) /1葡萄牙文 (Portuguese) .pdf	100%
拉丁/斯洛文尼亚文 (Slovenian) /1斯洛文尼亚文.pdf	100%
拉丁/塞尔维亚文 (latin) /1塞尔维亚文 (latin).pdf	100%
拉丁/匈牙利文 (Hungarian) /1匈牙利文 (Hungarian) .pdf	100%
拉丁/土耳其文 (Turkish) /1土耳其文 (Turkish) .pdf	94%
拉丁/爱沙尼亚文 (Estonian) /1爱沙尼亚文 (Estonian) .pdf	100%
拉丁/欧西坦文 (Occitan) /1欧西坦文 (Occitan) .pdf	100%

Start

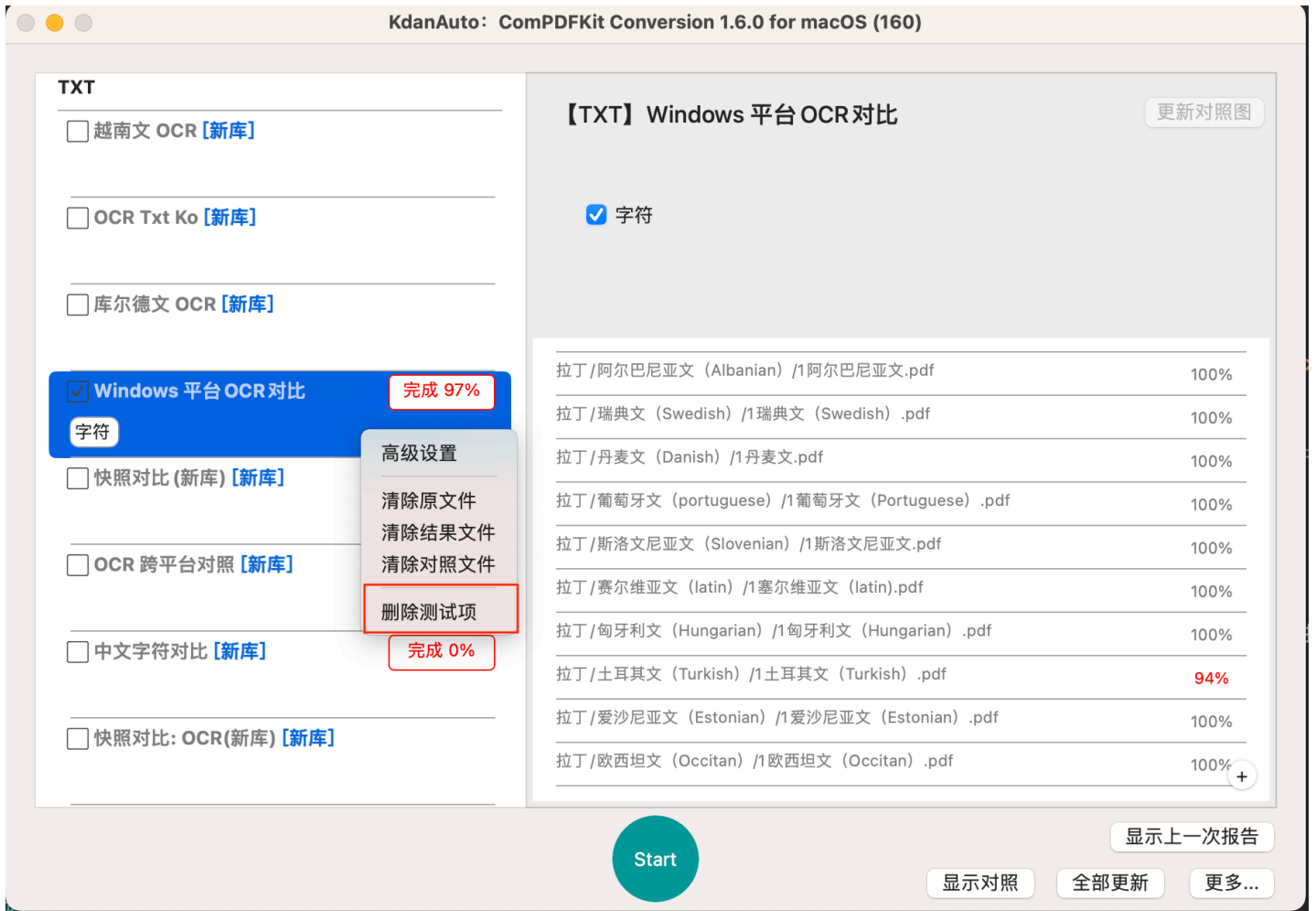
显示上一次报告

显示对照

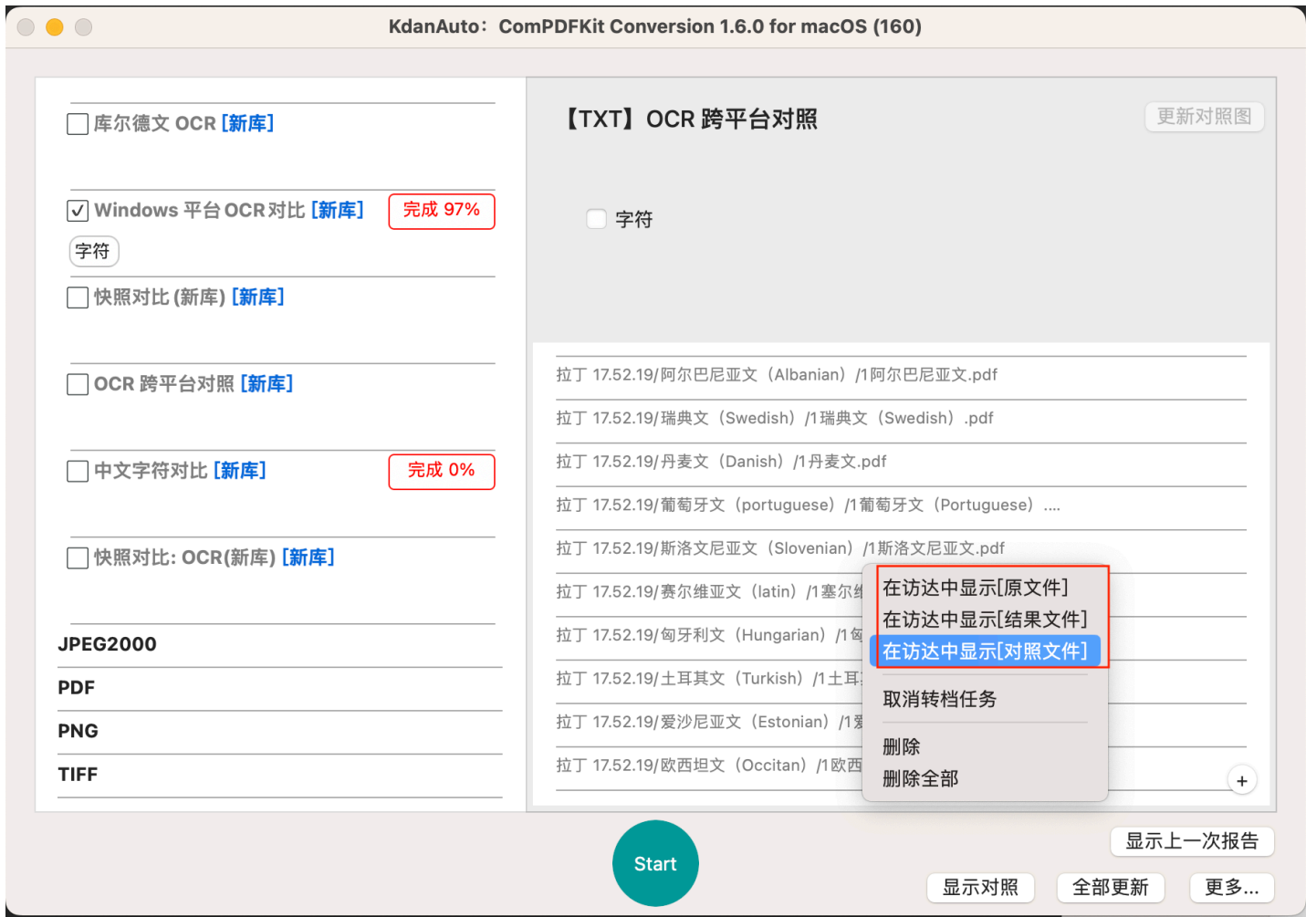
全部更新

更多...

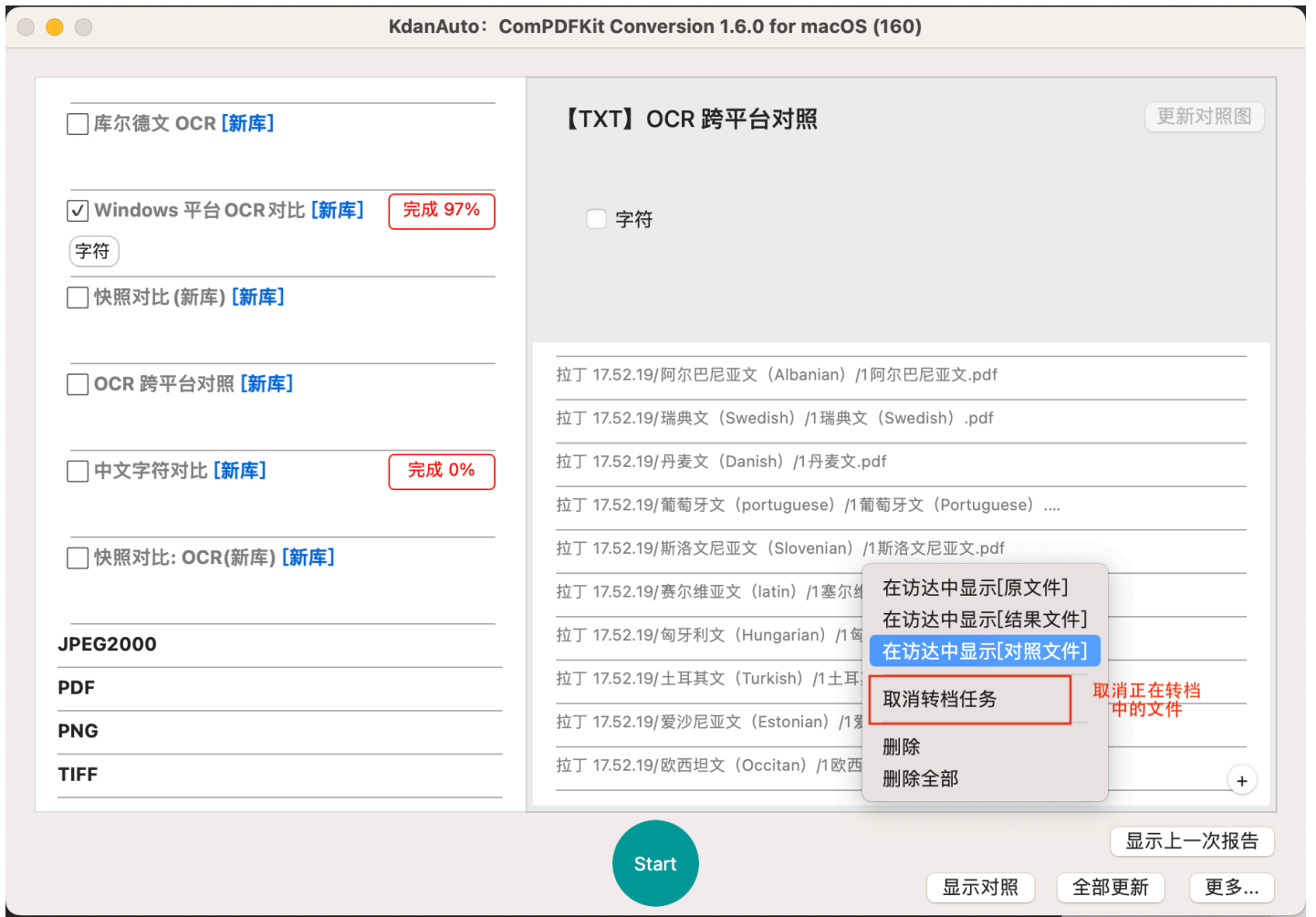
【c】删除测试项



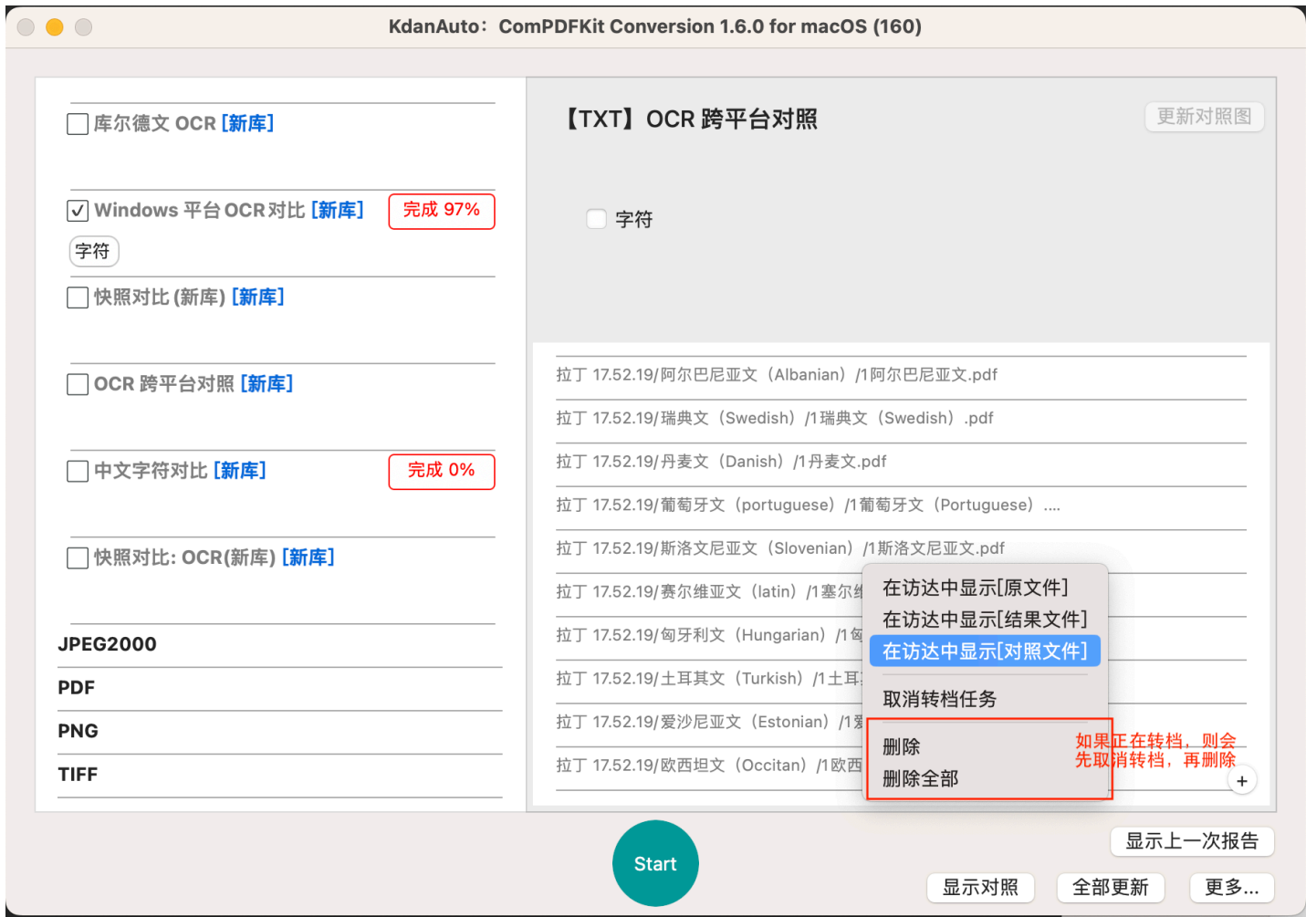
【d】 显示 原文件、显示对照文件、显示结果文件



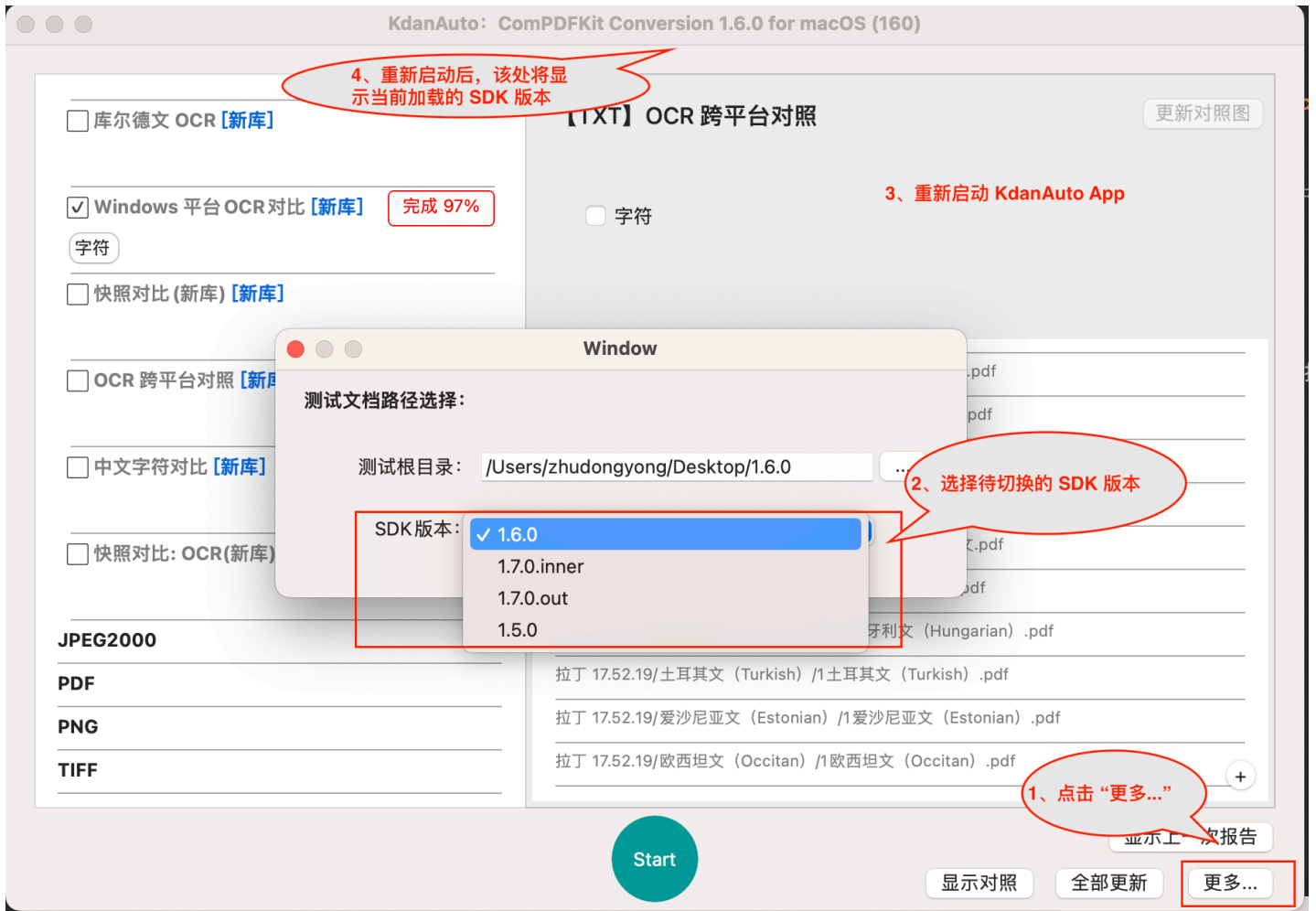
【e】 取消正在转档中的文件



【f】 删除测试项中测试原文档



【g】切换转档 SDK 版本，留意切换后，KdanAuto App 重启后生效



四、类型及参数配置（以下参数不再需要手动配置，通过高级设置界面进行设置）

配置表中，目前有预置 *PDF*、*RTF*、*Word*、*PPT*、*Excel*、*CSV*、*HTML*、*TXT*、*JPEG*、*PNG*、*GIF*、*TIFF*、*TGA*、*BMP*、*JPEG2000*格式类型，其中*PDF* 暂时考虑用作命令行注释编辑工具，批量执行文档结果对照测试项；而*RTF*、*Word*、*PPT*、*Excel*、*CSV*、*HTML*、*TXT*、*JPEG*、*PNG*、*GIF*、*TIFF*、*TGA*、*BMP*、*JPEG2000*为自动化转档测试对照测试项；

- *PDF* 用于 命令行工具，注释、编辑等操作，批量操作文档自动化校对选项
- *Word* 用于 *PDF to Word*，转档自动化测试及结果自动化对照测试项
- *Excel* 用于 *PDF to Excel*，转档自动化测试及结果自动化对照测试项

Note:

- 只能对照每份文档的左上角部分；
- 交互式表单无法识别（不确定是不是所有）

- *PPT* 用于 *PDF to PPT*，转档自动化测试及结果自动化对照测试项，

Note: 只能对照每份文档的第一页

- CSV 用于 *PDF to CSV*, 转档自动化测试及结果自动化对照测试项

Note:

- 只能对照每份文档的左上角部分;

- HTML 用于 *PDF to HTML*, 转档自动化测试及结果自动化对照测试项

Note:

- 新库转档出来是 zip

- RTF 用于 *PDF to RTF*, 转档自动化测试及结果自动化对照测试项 还可用于 *PDF to RTF*, 本地 (或其他平台) 转档/OCR 结果字符对比测试

Note:

- 新库 RTF 格式, 不支持预览, 需要用 WPS 等软件才能打开看效果;

- TXT 用于 *PDF to TXT*, 转档自动化测试及结果自动化对照测试项 还可用于 *PDF to TXT*, 本地 (或其他平台) 转档/OCR 结果字符对比测试

- PNG 用于 *PDF 转 PNG *图片格式, 转档自动化测试及结果自动化对照测试项

Note:

- 新库转档出来是 zip

- JPEG 用于 *PDF 转 JPEG *图片格式, 转档自动化测试及结果自动化对照测试项

Note:

- 新库转档出来是 zip

- GIF 用于 *PDF 转 GIF *图片格式, 转档自动化测试及结果自动化对照测试项

- TIFF 用于 *PDF 转 TIFF *图片格式, 转档自动化测试及结果自动化对照测试项

- TGA 用于 *PDF 转 TGA *图片格式, 转档自动化测试及结果自动化对照测试项

- BMP 用于 *PDF 转 BMP *图片格式, 转档自动化测试及结果自动化对照测试项

- JPEG2000 用于 *PDF 转 JPEG2000 *图片格式, 转档自动化测试及结果自动化对照测试项

支持的参数有 **Type (必选)**、**Name (必选)**、**Extention (必选)**、**Class (必选)**、**Params (可选)**, 对应参数介绍如下:

【Type】 Type为对应格式选项下, 识别类型, 同时对应于当前格式下, 测试类型目录, 例如 【JPEG】 格式

下面的【DPI-300】测试类型

【Name】 *Name*为当前 *Type*, 在 App 中显示的别名, 例如“对照识别”“DPI-300对照识别”

【Extention】 *Extention*为当前 *Type*文件处理后, 所得到的文件的后缀

【Class】 *Class*为当前格式、当前类型 (*Type*) , 用来做自动化处理的 *TestCase* 类, 与测试代码进行关联, 其中通用对照测试使用 *AutoTest*, PDF 命令行工具使用 *PDFCompareTest*

【Params】 *Params*为各自动化测试文件格式中, 可以再从外部 (不需要重新编译App) 来控制的参数配置, 其中各种格式对应的控制参数如下表:

```
- password: PDF 文档密码, String 类型
- useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新库, NSNumber类型 (0/1) 或 String 类型
- paneOptions: HTML 格式, 导出 HTML 页面及导航样式
  - CPDFConvertHtmlSinglePage = 0, // Single page with no na
    - CPDFConvertHtmlSinglePageNavigationByBookmarks = 1, // Navigate single pa
    - CPDFConvertHtmlMultiplePages = 2, // Multiple pages wit
    - CPDFConvertHtmlMultiplePagesSplitByBookmarks = 3 // Split multiple pag
- KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600,
- CPDFConvertOptionsKeyAllInOneSheet: 老库 PDF 转CSV, 多页合并为一个文档, Bool
```

不同格式对应控制参数

- PDF: PDF 文档快照对照
 - password, 所导入PDF 文档的密码, String 类型
- RTF: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
- Word: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
- PPT: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
- Excel: 转档及结果对照

- password, 所导入文档的密码
- useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
- CSV: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
 - CPDFConvertOptionsKeyAllInOneSheet: 老库 PDF 转CSV, 多页合并为一个文档, Bool
- HTML: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
 - paneOptions: HTML 格式, 导出 HTML 页面及导航样式
 - CPDFConvertHtmlSinglePage = 0, // Single page with no navigation.
 - CPDFConvertHtmlSinglePageNavigationByBookmarks = 1, // Navigate single page using bookmarks from PDF.
 - CPDFConvertHtmlMultiplePages = 2, // Multiple pages with no split.
 - CPDFConvertHtmlMultiplePagesSplitByBookmarks = 3 // Split multiple pages using bookmarks from PDF.
- TXT: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
- JPEG: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72
- PNG: 转档及结果对照
 - password, 所导入文档的密码
 - useOldLib: 转档类型中, 确认是用旧版本转档库 或是 用新转档库, NSNumber类型 (0/1) 或 String 类型 (TRUE, FALSE), 默认缺失时使用新转档库
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72

- GIF: 转档及结果对照
 - password, 所导入文档的密码
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72
- TIFF: 转档及结果对照
 - password, 所导入文档的密码
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72
- TGA: 转档及结果对照
 - password, 所导入文档的密码
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72
- BMP: 转档及结果对照
 - password, 所导入文档的密码
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72
- JPEG2000: 转档及结果对照
 - password, 所导入文档的密码
 - KMPDFConvertOptionsKeyImageDPI: 老转档库 DPI 参数, 常规设置有50、72、96、144、300、600, String类型, 默认72

五、问题备注

- PPT、Word、HTML、TXT、RTF 只有第一页能展示
- 交互式表单, 新库解析看起来有问题
- 老库注释会丢失
- 古代法制史 (打印版) .pdf 等比较多文档乱码