1、转换 PDF中的表格 文件到 Microsoft Excel 文档

2、程序自动识别pdf中的表格

3、电子工程师用excel建原理图的库（orcad）会需要，特别是引脚多的原理图库，如FPGA等，厂家没有给excel的引脚表，然后分了很多页的pdf,每页都有表头，如果全部复制入EXCEL，需要后期做很多处理，而且有换行和空格等很麻烦，费时间，本程序完美解决，其他行业也应该有用到

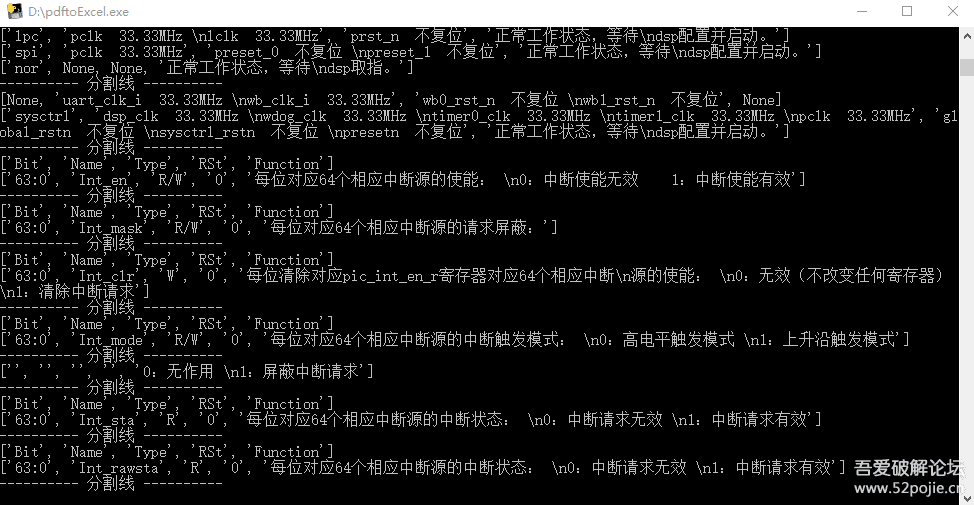
4、网上也有在线的转换工具和离线的转换工具，试了下不好用，上传后你的文件等于公开了，下载的工具好多不安全，有的要注册会员或有文件大小限制

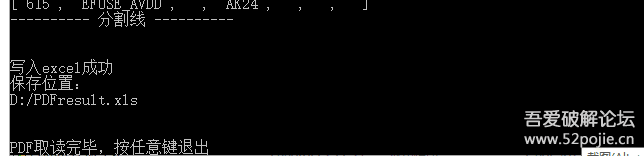
5、将exe文件放到D盘根目录下（程序里写死了），把你的pdf文件改名为test.pdf(程序里写死了),后面有源码你可以自己改成输入路径的，图方便

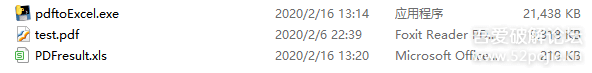
注意事项：

1、需不需要安装Microsoft Excel没有试过，测试时候是装着的

2、执行exe后需要一会时间估计几十秒，会出现cmd控制台打印信息，pdf文件越大时间越长，实测9M多的pdf表格都可以







源程序如下，有python环境条件的也可以自己写：

import pdfplumber

import xlwt

# 定义保存Excel的位置

workbook = xlwt.Workbook() #定义workbook

sheet = workbook.add\_sheet('Sheet1') #添加sheet

i = 0 # Excel起始位置

#path = input("E:/MyProject/python/test.pdf")

path = "D:/test.pdf" # 导入PDF路径

pdf = pdfplumber.open(path)

print('\n')

print('开始读取数据')

print('\n')

for page in pdf.pages:

# 获取当前页面的全部文本信息，包括表格中的文字

# print(page.extract\_text())

for table in page.extract\_tables():

# print(table)

for row in table:

print(row)

for j in range(len(row)):

sheet.write(i, j, row[j])

i += 1

print('---------- 分割线 ----------')

pdf.close()

# 保存Excel表

workbook.save('D:/PDFresult.xls')

print('\n')

print('写入excel成功')

print('保存位置：')

print('D:/PDFresult.xls')

print('\n')

input('PDF取读完毕，按任意键退出')