

提高OCR软件工作效率两法计算机等级考试 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/644/2021_2022__E6_8F_90_E9_AB_98OCR_E8_c98_644860.htm OCR软件就是光学输入识别软件，它可以把图片中的文字转换为字符，使用OCR处理印刷文档可以提高录入工作效率。但处理的长达几十页，甚至几千页的文档也是一件非常累人的活儿。笔者这里就和大家分享一些使用OCR软件的经验。

一、减少时间等待，提高扫描识别工效 使用OCR软件必经的程序一般是先用OCR自带的扫描功能扫描图片，再进行识别。而当处理较长的文档时等待扫描的时间会很长，如果要充分利用好时间我们可以利用Windows的多任务特性，用一个软件专门负责扫描，同时用一个OCR软件进行识别。www.Examda.CoM考试就到百考试题 扫描软件推荐使用《丹青中英文辨识系统》4.0以上或ACDSEE4.0以上版本，它们的特点是能自动按编号存储图片，换稿后只需按“扫描”按钮继续扫描。OCR软件推荐使用《尚书六号》，它能支持包括JPG图片的多种图片格式，用灰度级和真彩图片时还能自动调节对比度，或者手动调节对比度来提高识别率。尚书六号还能准确识别中文引号、用软回车存档、支持追加存档，可以有效的减少了后续排版处理的工作量。

二、弥补软件的不足，快速进行后续排版处理 要快速进行后续文档处理，选择合适的OCR软件十分重要，前面也提到，《尚书六号》还能准确识别中文引号、用软回车存档、支持追加存档，有效的减少了后续的工作量。但不管什么软件都有其缺陷，就如《尚书六号》，它在识别“加”字时经常会识别成“力口”，追加存档时，括号、冒号、

分号被识别为半角标点，跨页段落中会多出一个硬回车。采集者退散 因此，当熟悉使用一个OCR软件之后，可针对其缺陷，用排版软件的功能来弥补其不足，就像“加”字识别为“力口”。因“力口”在文档中出现的几率极小，可以在排版时直接批量替换掉。半角标点则可全部替换为全角标点，对于跨页段落中间的硬回车，可在碰到跨页段落时时，在追加存档时在页间插入一个特殊记号，在排版处理时连同硬回车一并替换为空。编辑特别推荐: 全国计算机等级一级笔试练习习题 计算机一级考试MS模拟试题及答案解析 100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com