

# Nachkalkulation mit OCR und Excel

Übersicht der [EVU-Seiten im VKU](#)

Die OCR in Tabellen stellt eine spezielle Herausforderung dar, die von Standard-OCR nicht gemeistert wird.

□

## Inhaltsverzeichnis

- [1 Online-Konvertierung](#)
- [2 Offline-Konvertierung](#)
- [3 Mit Microsoft Office 365](#)
- [4 Unter Python](#)
- [5 Tesseract](#)

## Online-Konvertierung

- <https://pdftables.com>
- <https://www.adobe.com/de/acrobat/how-to/pdf-to-excel-xlsx-converter.html>

## Offline-Konvertierung

- <https://pdf.wondershare.net>
- <https://www.adobe.com/de/acrobat/pricing.html>

Ältere Versionen von *Acrobat Pro* sind noch vereinzelt als Dauerlizenz erhältlich.

## Mit Microsoft Office 365

- <https://www.youtube.com/watch?v=EaS2Ooe9BNc>

## Unter Python

## Tesseract

*Googles Tesseract* ist in diesem Fall zwar in diesem speziellen Fall nicht die Lösung, kann bei anderen (größeren) Projekten jedoch eine große Hilfe sein. (Sie kann z.B. in [ffmpeg](#) eingebunden werden, um Texte in Videos zu erkennen.) Die kompilierte Windowsversion (läuft im CMD-Fenster) wird u.a. von der Universität Mannheim angeboten, siehe

[https://github.com/UB-Mannheim/Tesseract\\_Dokumentation/blob/main/Tesseract\\_Doku\\_Windows.md](https://github.com/UB-Mannheim/Tesseract_Dokumentation/blob/main/Tesseract_Doku_Windows.md).