

Table recognition is an important tool for digitalizing documents that contain tabular data, which often occur in areas of administration, finances and education. This thesis re-uses existing optical character recognition software to construct a new table recognition algorithm that aims to simplify the digitalization of diverse document types. The resulting algorithm achieves comparable or better results than currently available open-source software. Thesis additionally reviews common methods of OCR software implementation, and measures the influence of image preprocessing quality on the outcome of the table recognition.