

GImageReader: An open-source PDF app with OCR capability

By *Roy Schestowitz*

Created *04/03/2020 - 2:04am*

Submitted by Roy Schestowitz on Wednesday 4th of March 2020 02:04:26 AM Filed under [Software](#) [1]



GImageReader is a front-end application for the Tesseract OCR engine. For those new to Tesseract, it is an Optical Character Recognition Engine (OCR) that makes use of artificial intelligence to search and recognize printed text on images. It's an opensource library and one of the most popular OCR engines in the market.

Every day, be it in offices, home, etc., we find ourselves in situations where we need to extract text from an image. It could be a scanned document in image format, a piece of paper, or old research work. The outright option is to type the whole text with a text editor. But this process is time-consuming. Why not use an OCR to extract the text automatically?

[2]

Source URL: <http://www.tuxmachines.org/node/134768>

Links:

[1] <http://www.tuxmachines.org/taxonomy/term/38>

[2] <https://www.fosslinux.com/27198/gimagereader-an-open-source-pdf-app-with-ocr-capability.htm>