

tesseract

OCR (Optical Character Recognition) engine.

- Recognize text in an image and save it to `output.txt` (the '.txt' extension is added automatically):

```
tesseract {{image.png}} {{output}}
```

- Specify a custom language (default is English) with an ISO 639-2 code (e.g. `deu` = Deutsch = German):

```
tesseract -l deu {{image.png}} {{output}}
```

- List the ISO 639-2 codes of available languages:

```
tesseract --list-langs
```

- Specify a custom page segmentation mode (default is 3):

```
tesseract -psm {{0_to_10}} {{image.png}} {{output}}
```

- List page segmentation modes and their descriptions:

```
tesseract --help-psm
```