**OCR Tool Comparison + Research for Japanese Scripts (Katakana, Kanji, Hiragana)**

# OCR Tool Comparison for Japanese Scripts (Katakana, Kanji, Hiragana)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tool Name | Script Support | Accuracy | Cost | Ease of Use | Pros | Cons |
| Tesseract OCR | Katakana, Kanji, Hiragana (jpn, jpn\_vert) | Good (80–90%) | Free | Moderate (CLI or Python) | Open-source, customizable, strong community support | Struggles with complex layouts or faded documents |
| EasyOCR | Katakana, Kanji, Hiragana | Very Good (85–92%) | Free | Easy (Python package install) | Deep learning based, easy setup | Heavy installation size (~1GB), slower for large batches |
| Google Cloud Vision OCR | Katakana, Kanji, Hiragana | Excellent (>95%) | Paid (after free tier) | Very Easy (API-based service) | Very high accuracy, scalable for large projects | Paid service after a limited free tier, API key management |
| Kraken OCR | Katakana, Kanji, Hiragana | Good (~80–88%) | Free | Hard (Linux CLI, requires model tuning) | Good for historical/degraded text, customizable | Harder to setup, sometimes needs manual training |
| Microsoft Azure OCR | Katakana, Kanji, Hiragana | Excellent (>95%) | Paid (after free tier) | Very Easy (API-based service) | Strong Japanese and multilingual support, handles vertical text well | Azure account and billing setup needed |