

2023年3月27日

世界初^(注1) 近代書籍対応の市販 AI-OCR ソフト 「FROG AI-OCR」新機能追加 ～透明テキスト付 PDF 出力機能・多段組対応機能～

【概要】

モルフォグループにおいて AI の事業化を担う、株式会社モルフォ AI ソリューションズ（所在地：東京都千代田区、代表取締役：神田武、以下 モルフォ AIS）は、AI-OCR^(注2) ソフトウェア「FROG AI-OCR」の新機能搭載版を 2023 年 3 月 27 日（月）よりリリースいたします。



モルフォ AIS は、令和 3 年度から国立国会図書館の委託事業として図書館向け AI-OCR（NDLOCR）の開発に従事してきました。特徴としては、既存の OCR では読み取りが不可であった、近代書籍の旧字旧仮名や複雑なレイアウトへの対応、既存 OCR の 2 倍以上の読み取り精度^(注3)を達成いたしました。開発した AI-OCR プログラムは、NDLOCR として国立国会図書館が CC BY 4.0 ライセンスで公開しています。^(注4)

この度、図書館向け AI-OCR（NDLOCR）を活用した市販ソフト「FROG AI-OCR（2022 年 7 月 1 日発売）」に新機能の追加をいたします。

【新機能について】

①透明テキスト付 PDF 出力機能

現状テキストファイルでの出力に加え、透明テキスト付 PDF での出力も可能になります。こちらにより、出力の PDF 内でテキスト検索を行うことができるようになりました。

紙をスキャンしたPDF ※テキストを選択できない

透明テキスト付PDF ※テキストを選択できる！

透明テキスト化

キーワード検索も可能！
(例)「老婆」

※PDF画像上に透明なテキストデータを付与することで、「コピー&ペースト」や「検索」などが可能となります。

②多段組対応機能

本改善では読み取り結果を表示する際に、多段組であっても、適切に推定し表示するように変更いたしました。

また、従来版は3段以上の段組がある場合も、一律で2段での表示となっておりますが、上記変更を元に出力されたテキストデータを1段で表示するように変更いたしました。

従来版

多段組対応版

多段組対応

【お申込み・問い合わせ窓口】

<https://www.morphoai.com/ai-ocr>

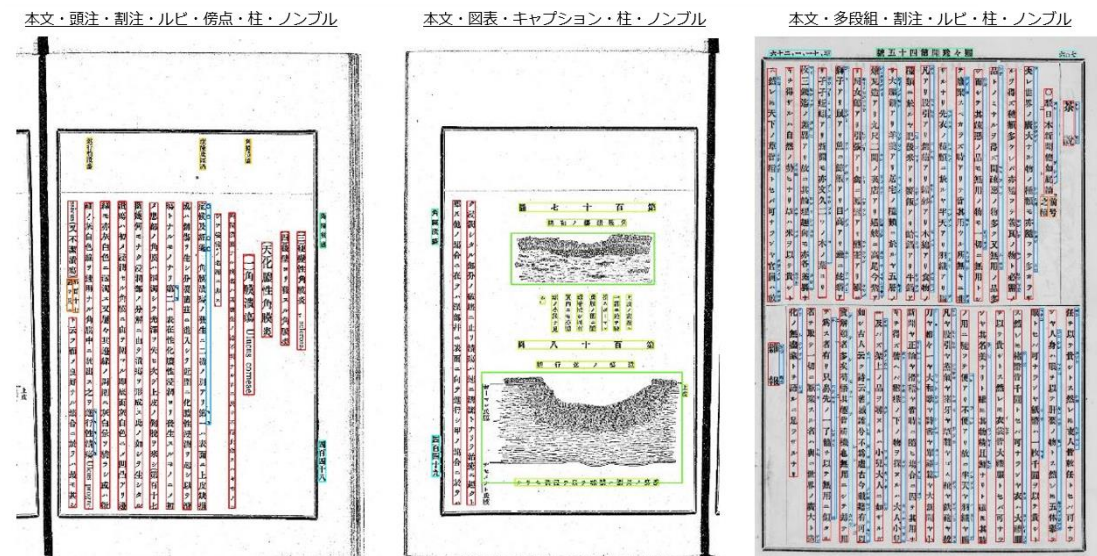
こちらより無償トライアル頂く事が可能です

【FROG AI-OCR 紹介ムービー】

<https://www.youtube.com/watch?v=MxAAHdQka4I>

【NDLOCR の特徴】

①複雑なレイアウトへの対応



②多様な文字種（旧字旧仮名）への対応

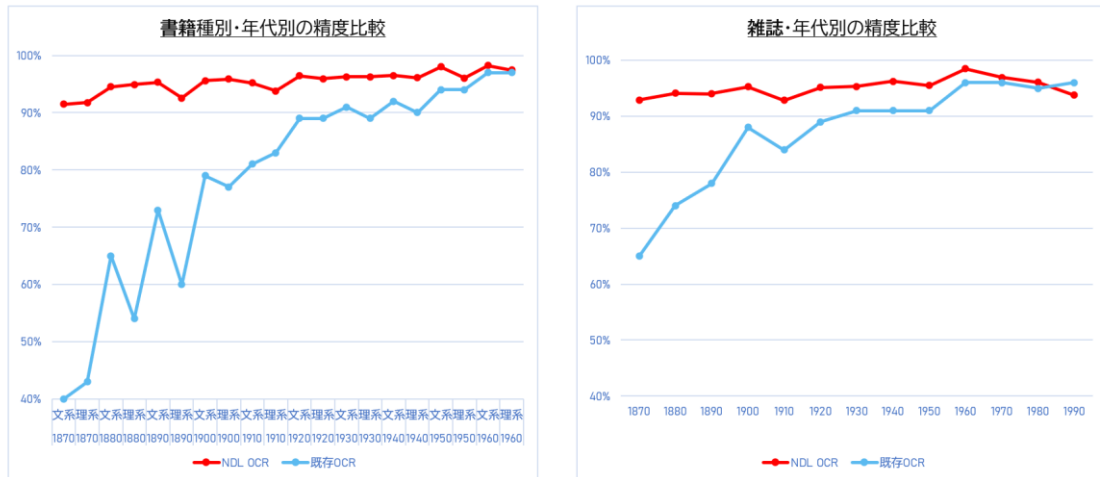


③ OCR 処理プログラムの精度向上

1860 年代以降の書籍・雑誌において市販 OCR より高い 90%以上の精度で判定可能となりました。

(注 5) 特に明治期～昭和初期の近代書籍・雑誌においては、市販 OCR より約 2 倍 (約 40%→90%以上) の読み取り精度となりました。

NDL OCRと既存OCRの精度比較(文字単位での一致度(F値))



(出典) 国立国会図書館 NDLラボ「令和3年度OCR処理プログラム研究開発」https://lab.ndl.go.jp/data_set/ocr/r3_morpho/

【新機能版リリース日】

2023年3月27日(月)

【注釈】

- (注 1) 弊社調べ、明治～昭和初期の書籍(文系・理系)及び雑誌の認識精度 90%以上を保有する市販の AI-OCR 製品が対象
- (注 2) OCR (Optical Character Recognition) 光学式文字認識
- (注 3) 令和 3 年度 OCR 処理プログラムの研究開発リリース参照
https://www.morphoinc.com/news/20220428-jpr-mais_ndl
- (注 4) NDLOCR のアプリケーション (GitHub 上で公開)
https://github.com/ndl-lab/ndlocr_cli
- (注 5) 国立国会図書館 NDL ラボ「令和 3 年度 OCR 処理プログラム研究開発」
https://lab.ndl.go.jp/data_set/ocr/r3_morpho/

【関連プレスリリース】

2022年4月28日

モルフォ AI ソリューションズが、国立国会図書館の最新 AI 技術を活用した OCR 処理プログラムの開発を完了

https://www.morphoinc.com/news/20220428-jpr-mais_ndl

2022年5月9日

モルフォ AI ソリューションズが、国立国会図書館から視覚障害者等用の OCR 開発を受託

https://www.morphoinc.com/news/20220509-jpr-mais_ndl

2022年6月14日

世界初、近代書籍対応の市販 AI-OCR ソフト「FROG AI-OCR」新発売

https://www.morphoinc.com/news/20220614-jpr-mais_frog_aiocr

【株式会社モルフォ AI ソリューションズについて】

モルフォ AI ソリューションズは、AI（人工知能）の事業化に取り組む企業です。行政、電力、交通、製造といった社会インフラの領域で、AI-OCR をはじめとする最先端の AI 技術の導入と実運用を推進しております。

所在地：東京都千代田区神田錦町 2-2-1 KANDA SQUARE 11 階 WeWork 内

代表者：代表取締役 神田 武

設立：2019年12月

事業内容：AI コンサルティング、システムインテグレーション、SW・HW 販売など

ホームページ：<https://www.morphoai.com>

【株式会社モルフォについて】

モルフォは「画像処理／AI（人工知能）」の研究開発型企業です。高度な画像処理技術を組み込みソフトウェアとして、国内外のスマートフォン、半導体メーカーを中心にグローバルに展開しています。また、カメラで捉えた画像情報をエッジデバイスやクラウドで解析する、AI を駆使した画像認識技術を車載や産業 IoT 分野へ提供し、様々なイノベーションを先進のイメージング・テクノロジーで実現しています。

所在地：東京都千代田区神田錦町 2-2-1 KANDA SQUARE 11 階 WeWork 内

代表者：代表取締役社長 平賀 督基（まさき）、【博士（理学）】

設立：2004年5月26日

資本金：1,783,958 千円（2022年10月31日現在）

事業内容：画像処理および AI（人工知能）技術の研究・製品開発。スマートフォン・半導体・車載・産業 IoT 向けソフトウェア事業をグローバルに展開。

ホームページ : <https://www.morphoinc.com/>

Facebook : <https://www.facebook.com/morphoinc>

Twitter : https://twitter.com/morpho_inc

【お問合せ先】

株式会社モルフォ 広報担当 大野

TEL : 080-8433-3415

お問い合わせフォーム : <http://www.morphoinc.com/contact>

*モルフォ、Morpho およびモルフォロゴは株式会社モルフォの登録商標または商標です。