厚生労働科学研究費補助金(長寿科学政策研究事業)
分 担 研 究 報 告 書
「機械学習を用いた介護認定審査会の審査判定プロセス等を補助するシステムを開発するための研究」

OCR の読み込み技術に関する比較

研究分担者　蜂屋 孝太郎（帝京平成大学 人文社会学部）
研究代表者　高橋 秀人（帝京平成大学 薬学部）

研究要旨

【目的】

手書き文字を含む文書を対象とした機械学習システムの精度向上には、OCR の性能が重要である。本研究では、OCR 技術の変遷を整理するとともに、最新の Vision Language Model (VLM) を用いた手書き文字認識の有効性を評価することを目的とした。

【方法】

Google Scholar を用いて OCR に関する総説論文を調査し、主要技術の変遷を整理した。また、主治医意見書を模したダミーデータを用い、複数の VLM による手書きチェックおよび手書き文章の読み取り精度を比較評価した。

【結果】

手書きチェックの読み取りでは多くのモデルが高い精度を示し、特に一部のクローズドモデルは完全一致を達成した。一方、手書き文章の読み取りではモデル間で性能差が見られ、クローズドモデルが低い誤り率と高い精度を示した。

【結論】

VLM は手書き文書に対して高い性能を有するが、機密性の高い医療文書への適用においてはクラウド利用の制約がある。実運用では、セキュリティと性能のバランスを考慮したモデル選択および環境構築が重要である。

A．研究目的

開発するシステムでは機械学習を活用する。学習で使用するデータセットには、手書き文字など、文字がテキストデータではなく画像データとして記録されている文書が含まれる。これらの文書を機械学習モデルに入力する場合、光学文字認識(Optical Character Recognition: OCR)により画像情報をテキスト情報に変換してから入力する（一般的にはさらに数値やカテゴリ値に変換されてから入力される)。この OCR の精度は、最終的な機械学習モデルの性能にも影響を与える。したがって、適切な OCR 技術を採用することが重要である。本研究では、そのための技術調査および予備的な評価を実施し、その結果を報告する。

B．研究方法

OCR 技術の調査では、Google Scholar を用いてキーワード「OCR」により総説論文の検索を行い、関連度上位 20 件のオープンアクセス論文を抽出し[1]-[20]、主要技術の変遷を整理した。

OCR の予備的評価では、実際の主治医意見書を模したダミーデータ（1 名分・2 ページ、付録の縮小図参照）を用い、以下の 2 種類の情報について、主要な Vision Language Model (VLM) による読み取り精度を評価した。

- チェックボックス上の手書きチェック
- 手書き文章（診断名、傷病経過、特記事項）

評価対象の VLM として、以下のモデル下を選定

表 1. OCR 技術の変遷

時代	主流認識方式	前世代の限界・移行背景	主な改善点
黎明期 (～1940年代)	アナログテンプレートマッチング	電信・ 視覚障害者支援ニーズ	機械的な文字変換の実現
第1世代 (1950～60年代)	専用フォントテンプレートマッチング	コンピュータ登場による大量データ入力需要	デジタルデータへの変換
第2世代 (1960～70年代中)	エッジ・輪郭パターン認識	専用フォントしか読めない制約が実用上の障壁	汎用フォント、手書き数字対応
第3世代 (1970～80年代)	多次元特徴抽出+統計的分類	通常フォントや漢字の手書き認識に限界	漢字認識、任意手書き対応
デジタル化時代 (1990～2000年代)	HMM＋ソフトウェア化	専用機コストの高さ、 多様化する文書形態	一般化、低価格化、商業展開
ディープラーニング (2010年代)	CNN+LSTM+CTC	手動特徴設計の限界、 精度95%の壁	認識精度99%、End-to-End学習
Transformer/VLM (2017年～)	Vision Transformer + LLM	文字は認識できても意味や構造が理解できない	構造理解、内容理解、多言語対応

略語: OCR, Optical Character Recognition(光学文字認識); HMM, Hidden Markov Model(隠れマルコフモデル); CNN, Convolutional Neural Network(畳み込みニューラルネットワーク); CTC, Connectionist Temporal Classification(コネクショニスト時系列分類); VLM, Vision Language Model(視覚言語モデル); LLM, Large Language Model(大規模言語モデル)

した。

- Qwen3.6（パラメータ数 35B）
- GLM-4.6V（パラメータ数 106B）
- GPT-5.5
- Gemini 3.1 Pro Preview
- Claude Opus 4.7

前者 2 モデルは、モデルファイルが公開されているオープンウェイトモデルのうち、OCR 性能に定評のある系列の最新バージョンである。後者 3 モデルは、モデルファイルが公開されていないクローズドモデルのうち、各種ベンチマークにおいて高い性能を示す商用モデルである。いずれのモデルも OpenRouter 社の API サービスを利用して推論処理を実行した。

手書きチェックの読み取りでは以下の指示をプロンプトとして入力した。

「この主治医意見書において、チェックボックスにチェックが入っている箇所の文字を読み取ってください。」

1 ページ目の手書き文字の読み取りでは以下の指示をプロンプトとして与えた。

「この主治医意見書に記載されている、『診断名及び発症年月日』と『生活機能低下の直接の原因となっている傷病または特定疾病の経過及び投薬内容を含む治療内容』を読み取ってください。誤字は修正してください。しかし、意訳せずにできるだけ原文のままとしてください。」

2 ページ目の手書き文字の読み取りでは、上記のプロンプトにおける読み取り箇所を「特記すべき事項」に置き換えた。推論結果のばらつきを抑え、再現性を確保するため、確率的生成の自由度を適度に制限する目的で temperature = 0.5 を設定した。

C．研究結果

1. OCR の歴史と技術変遷

OCR（光学文字認識）は、画像として取得された文字情報をテキストデータに変換する技術であり、文書電子化や業務自動化などに広く用いられている。本技術は約 100 年にわたり発展し、「文字認識」から「文書理解」へと進化してきた（表 1）。

初期（1910 年代～1940 年代）は、光電管を用いたアナログ的なパターンマッチングが主流であり、対応文字や精度に大きな制約があった。1950～1960 年代前半には、OCR 専用フォント（OCR-A/B）を前提としたテンプレートマッチング方式が実用化され、商用利用が始まったが、一般文書や手書き文字には対応できなかった。

1960 年代後半～1970 年代には、輪郭や線分といった特徴を抽出するパターン認識技術が導入され、複数フォントへの対応が進んだ。また、日本では郵便番号認識などの実用化が進展したが、手書き文字やノイズへの耐性には課題が残った。

1970 年代後半～1980 年代には、特徴量を数値化し統計的に分類する手法が主流となり、認識精度が向上した。しかし、特徴設計が人手に依存し、複雑な文書への対応には限界があった。

1990 年代～2000 年代初頭には、PC 普及に伴いソフトウェア OCR が一般化し、HMM（隠れマ

ルコフモデル）や辞書補正により精度が改善された。印刷文字では高精度を達成した一方、非定型帳票や手書き文字の認識は依然として困難であった。

2010 年代にはディープラーニングが導入され、CNN (Convolutional Neural Network)・LSTM (Long Short-Term Memory)・CTC (Connectionist Temporal Classification)を組み合わせた手法により、特徴抽出と系列認識が自動化され、特に手書き文字認識の精度が飛躍的に向上した。しかし、文書全体の構造理解には課題が残った。

近年（2017 年以降）は Transformer を基盤としたモデルが登場し、OCR は VLM（Vision-Language Model）へと発展している。これにより、画像から直接情報を抽出し、表構造や意味内容を含めて理解・出力することが可能となった。現在では、OCR は単なる文字認識ではなく、文書構造解析や意味理解を含む統合的な技術へと進化している。

2. OCR ツールの予備的評価

手書きチェックの読み取り精度を表 2 に示す。ここで、チェックがある項目を正しく抽出できた数を TP (True Positive)、チェックがあるのに見落とした数を FN (False Negative)、チェックがない項目、もしくはチェック項目でないものをチェックがあるものとして抽出した数を FP (False Positive)としてカウントし、再現率と適合率、F1 スコアを算出した。

表 2. 手書きチェックの読み取り精度

	再現率	適合率	F1 ｽｺｱ
GLM-4.6V	0.91	0.94	0.92
Qwen3.6	0.99	0.94	0.96
Claude Opus 4.7	1.00	0.92	0.96
GPT-5.5	1.00	1.00	1.00
Gemini 3.1 Pro Prev.	1.00	1.00	1.00

手書き文字の読み取り精度を表 3 に示す。ここで、最短編集距離は正解文字列を推定結果に変換するために必要な挿入・削除・置換の最小回数を表す指標である。CER (Character Error Rate) は、この最短編集距離を正解文字数で正規化した文字誤り率であり、値が小さいほど精度が高いことを示す。また、精度は CER を基に算出した正解率（1 − CER）として定義した。

表 3. 手書き文字の読み取り精度

	最短編集距離	CER	精度
GLM-4.6V	228	0.60	0.40
Qwen3.6	82	0.22	0.78
Claude Opus 4.7	33	0.09	0.91
GPT-5.5	28	0.07	0.93
Gemini 3.1 Pro Prev.	23	0.06	0.94

D. 考察

本研究では、OCR 技術の歴史的変遷の整理とともに、最新の VLM を用いた手書き文字認識の性能を予備的に評価した。その結果、従来の OCR と比較して、VLM は手書き文字を含む文書に対して高い認識精度を有することが確認された。

手書きチェックの読み取りにおいては、すべてのモデルが高い F1 スコアを示し、とくに GPT-5.5 および Gemini 3.1 Pro Preview は完全一致を達成した。このことから、単純な視覚的パターンの抽出に関しては、現代の VLM が極めて高い性能を有することが示唆される。一方で、手書き文章の読み取りでは、モデル間で顕著な性能差が見られ、GPT-5.5 および Gemini 3.1 Pro Preview が最も低い CER と高い精度を示した。これに対し、GLM-4.6V および Qwen3.6 では誤認識が多く、長文になるほど誤りが蓄積する傾向が確認された。これらの結果は、モデルの規模や学習データの違いが認識性能に大きく影響することを示している。

また、手書きチェックと手書き文章の結果の差から、単純な記号認識に比べ、文脈を伴う文字列認識の難易度が高いことが明らかとなった。これは、後者が文字認識に加えて文脈理解や言語生成能力にも依存するためであり、VLM の特性がより強く反映される領域であると考えられる。

一方で、本研究の対象である要介護認定審査会資料は、個人の医療・生活情報を含む機密性の高い文書である。このため、クラウド上でのみ提供されるクローズドモデル（GPT-5.5 や Gemini など）を実運用で利用することには、情報セキュリティおよび法令遵守の観点から制約が生じる可能性が高い。すなわち、本研究で高い性能を示したモデルが必ずしも実運用に適用可能とは限らないという課題が存在する。

この点において、オープンウェイトモデルである Qwen3.6 および GLM-4.6V は、ローカル環境での運用が可能であり、機密情報を外部に送信せずに処理できるという利点を有する。ただし、本評価ではこれらのモデルはクローズドモ

デルと比較して認識精度が劣る結果となっており、セキュリティと性能のトレードオフが明確に示された。

以上より、実運用を見据えた OCR システムの設計においては、単純な認識性能だけでなく、データの機密性や運用環境を含めた要件を総合的に考慮する必要がある。今後は、ローカル実行可能なモデルの性能向上や、オンプレミス環境での安全な推論基盤の整備、さらには匿名化や部分的クラウド利用といったハイブリッドな運用方法の検討が重要である。

また、本研究は 1 サンプルに基づく予備的評価であるため、今後はデータ数を拡充し、より多様な帳票・筆跡に対する性能評価を行うことで、実用化に向けた信頼性の検証を進める必要がある。

E．結論

本研究では、OCR 技術の変遷を整理するとともに、最新の VLM を用いた手書き文字認識の性能を予備的に評価した。その結果、VLM は従来の OCR と比較して高い認識精度を示し、特に手書き文書に対して有効であることが確認された。一方で、クラウド上で提供される高性能モデルは機密性の高い医療文書への適用に制約があり、実運用においてはオープンウェイトモデルとの性能およびセキュリティのトレードオフを考慮する必要がある。今後は、安全性を確保した運用環境の構築とローカルモデルの性能向上を通じて、実用的な OCR システムの実現を目指す。

F．健康危険情報

なし

G．研究発表

1．論文発表

なし

2．学会発表

なし

H．知的財産権の出願・登録状況

1．特許取得

なし

2．実用新案登録

なし

3．その他

なし

参考文献

[1] Thi Tuyet Hai Nguyen, Adam Jatowt, Mickael Coustaty, Antoine Doucet, “Survey of Post-OCR Processing Approaches,” ACM Comput. Surv. 54(6), 2022.

[2] Safiullah Faizullah, Muhammad Sohaib Ayub, Sajid Hussain, Muhammad Asad Khan, "A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges" Applied Sciences 13(7), 2023.

[3] Jamshed Memon, Maira Sami, Rizwan Ahmed Khan, Mueen Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," IEEE Access 8, 2020.

[4] Nishant Subramani, Alexandre Matton, Malcolm Greaves, Adrian Lam, “A Survey of Deep Learning Approaches for OCR and Document Understanding,” arXiv:2011.13534 [cs.CL], 2021.

[5] Xujun Peng, Huaigu Cao, Srirangaraj Setlur, Venu Govindaraju, Prem Natarajan, “Multilingual OCR research and applications: an overview,” Proceedings of the 4th International Workshop on Multilingual OCR (MOCR '13), 2013.

[6] Rayyan Najam, Safiullah Faizullah, "Analysis of Recent Deep Learning Techniques for Arabic Handwritten-Text OCR and Post-OCR Correction," Applied Sciences 13(13), 2023.

[7] K.Karthick, K.B.Ravindrakumar, R.Francis, S.Ilankannan, “Steps Involved in Text Recognition and Recent Research in OCR; A Study,” International Journal of Recent Technology and Engineering 8(1), 2019.

[8] Everistus Zeluwa Orji, Ali Haydar, İbrahim Erşan, Othmar Othmar Mwambe, "Advancing OCR Accuracy in Image-to-LaTeX Conversion—A Critical and Creative Exploration," Applied Sciences 13(22), 2023.

[9] Jay Dilipbhai Thanki, Priyank Dineshbhai Davda, Dr. Priya Swaminarayan, “A Review on OCR Technology,” Journal of Emerging Technologies and Innovative Research 8(4), 2021.

[10] B. Anuradha Srinivas, Arun Agarwal, C. Raghavendra Rao, “An Overview of OCR

Research in Indian Scripts," International Journal of Computer Sciences and Engineering Systems 2(2), 2008.
[11] Amarjot Singh, Ketan Bacchuwar, Akshay Bhasin, "A Survey of OCR Applications," International Journal of Machine Learning and Computing 2(3), 2012.
[12] Clemens Neudecker, Konstantin Baierer, Mike Gerber, Christian Clausner, Apostolos Antonacopoulos, Stefan Pletschacher, "A survey of OCR evaluation tools and metrics," Proceedings of the 6th International Workshop on Historical Document Imaging and Processing (HIP '21), 2021.
[13] Noman Islam, Zeeshan Islam, Nazia Noor, "A Survey on Optical Character Recognition System," Journal of Information & Communication Technology 10(2), 2016.
[14] Rohit Verma, Jahid Ali, "A-Survey of Feature Extraction and Classification Techniques in OCR Systems," International Journal of Computer Applications & Information Technology, 1(3), 2012.
[15] Sivashanth Suthakar, Sarveswaran Kengatharaiyer, Andrew Charles Eugene Yugarajah, "Comprehensive Evaluation of Tamil OCR Systems: A Survey, Dataset Creation, Benchmarking, and Error Analysis," International Journal on Advances in ICT for Emerging Regions 18 (3), 2025.
[16] U. Pal, B.B. Chaudhuri, "Indian script character recognition: a survey, Pattern Recognition 37(9), 2004.
[17] Manoj Sonkusare, Narendra Sahu, "A survey on handwritten character recognition (HCR) techniques for English alphabets," Advances in Vision Computing: An International Journal 3(1), 2016.
[18] Ayush Purohit, Shardul Singh Chauhan, "A Literature Survey on Handwritten Character Recognition," International Journal of Computer Science and Information Technologies 7(1), 2016.
[19] Xiaoxue Chen, Lianwen Jin, Yuanzhi Zhu, Canjie Luo, Tianwei Wang, "Text Recognition in the Wild: A Survey," ACM Comput. Surv. 54(2), 2022.
[20] Sukhjinder Singh, Naresh Kumar Garg, Munish Kumar, "Feature extraction and classification techniques for handwritten Devanagari text recognition: a survey," Multimed Tools Appl 82, 2023.

付録：

OCR の予備的評価で使用した主治医意見書の模擬データ

（別添２）

主治医意見書

記入日　平成　年　月　日

申請者	（ふりがな） 明・大・昭　年　月　日生（84歳）	男・**女**	〒　－ 連絡先　（　）

上記の申請者に関する意見は以下の通りです。
主治医として、本意見書が介護サービス計画作成に利用されることに　☑同意する。　□同意しない。

医師氏名
医療機関名　　　　電話　（　）
医療機関所在地　　　FAX　（　）

（１）最終診察日	~~平成~~R３年　11月　10日
（２）意見書作成回数	☑初回　□２回目以上
（３）他科受診の有無	☑有　□無 （有の場合）→☑内科　☑精神科　□外科　□整形外科　□脳神経外科　□皮膚科　□泌尿器科　□婦人科　□眼科　□耳鼻咽喉科　□リハビリテーション科　□歯科　□その他（　）

１．傷病に関する意見

（１）診断名（特定疾病または生活機能低下の直接の原因となっている傷病名については１．に記入）及び発症年月日

1. アルツハイマー型認知症　　発症年月日　（~~昭和・平成~~ R1 年 頃 月 日頃）
2. 　　発症年月日　（昭和・平成　年　月　日頃）
3. 　　発症年月日　（昭和・平成　年　月　日頃）

（２）症状としての安定性　□安定　☑不安定　□不明
（「不安定」とした場合、具体的な状況を記入）

（３）生活機能低下の直接の原因となっている傷病または特定疾病の経過及び投薬内容を含む治療内容
〔最近（概ね６ヶ月以内）介護に影響のあったもの　及び　特定疾病についてはその診断の根拠等について記入〕

2〜3年前頃より物忘れに家族が気づくようになったが日常生活はかろうじて保たれていた。R2年夏に夫が入院、施設入居し、本人は独居生活となった。そのころより、認知機能の低下が著しくなり、定期的に様子を見に来る義娘に対して物盗られ妄想を抱くようになる。R2年12月10日、当院精神科を初診、メマンチン、少量のリスペリドン投与を開始。BPSDは軽減した一方、歩行障害~~[illegible]~~が残っており、独居では転倒の危険が高い。娘は近いうちに施設入居を希望している。

２．特別な医療　（過去１４日間以内に受けた医療のすべてにチェック）

処置内容　□点滴の管理　□中心静脈栄養　□透析　□ストーマの処置　□酸素療法　□レスピレーター　□気管切開の処置　□疼痛の看護　□経管栄養
特別な対応　□モニター測定（血圧、心拍、酸素飽和度等）　□褥瘡の処置
失禁への対応　□カテーテル（コンドームカテーテル、留置カテーテル 等）

３．心身の状態に関する意見

（１）日常生活の自立度等について
・障害高齢者の日常生活自立度（寝たきり度）　□自立　□J1　☑J2　□A1　□A2　□B1　□B2　□C1　□C2
・認知症高齢者の日常生活自立度　□自立　□Ⅰ　□Ⅱa　□Ⅱb　□Ⅲa　□Ⅲb　☑Ⅳ　□M

（２）認知症の中核症状（認知症以外の疾患で同様の症状を認める場合を含む）
・短期記憶　□問題なし　☑問題あり
・日常の意思決定を行うための認知能力　□自立　□いくらか困難　□見守りが必要　☑判断できない
・自分の意思の伝達能力　□伝えられる　□いくらか困難　☑具体的要求に限られる　□伝えられない

（３）認知症の周辺症状　（該当する項目全てチェック：認知症以外の疾患で同様の症状を認める場合を含む）
□無　☑有 → {□幻視・幻聴　☑妄想　☑昼夜逆転　☑暴言　☑暴行　☑介護への抵抗　□徘徊　☑火の不始末　□不潔行為　□異食行動　□性的問題行動　□その他（　）}

（４）その他の精神・神経症状
□無　□有　〔症状名：　　専門医受診の有無　□有　（　）□無〕

（５）身体の状態
利き腕（☑右　□左）　身長＝152 cm　体重＝40 kg（過去６ヶ月の体重の変化　□増加　☑維持　□減少）
□四肢欠損　（部位：　）
□麻痺　□右上肢（程度：□軽　□中　□重）　□左上肢（程度：□軽　□中　□重）
　□右下肢（程度：□軽　□中　□重）　□左下肢（程度：□軽　□中　□重）
　□その他（部位：　程度：□軽　□中　□重）
□筋力の低下　（部位：　程度：□軽　□中　□重）
□関節の拘縮　（部位：　程度：□軽　□中　□重）
□関節の痛み　（部位：　程度：□軽　□中　□重）
□失調・不随意運動　・上肢　□右　□左　・下肢　□右　□左　・体幹　□右　□左
□褥瘡　（部位：　程度：□軽　□中　□重）
□その他の皮膚疾患　（部位：　程度：□軽　□中　□重）

４．生活機能とサービスに関する意見

（１）移動
屋外歩行　□自立　☑介助があればしている　□していない
車いすの使用　☑用いていない　□主に自分で操作している　□主に他人が操作している
歩行補助具・装具の使用（複数選択可）　☑用いていない　□屋外で使用　□屋内で使用

（２）栄養・食生活
食事行為　☑自立ないし何とか自分で食べられる　□全面介助
現在の栄養状態　☑良好　□不良
→　栄養・食生活上の留意点（　）

（３）現在あるかまたは今後発生の可能性の高い状態とその対処方針
□尿失禁　☑転倒・骨折　☑移動能力の低下　□褥瘡　□心肺機能の低下　☑閉じこもり　☑意欲低下　☑徘徊　□低栄養　□摂食・嚥下機能低下　□脱水　□易感染性　□がん等による疼痛　□その他（　）
→　対処方針（　）

（４）サービス利用による生活機能の維持・改善の見通し
□期待できる　□期待できない　☑不明

（５）医学的管理の必要性（特に必要性の高いものには下線を引いて下さい。予防給付により提供されるサービスを含みます。）
□訪問診療　☑訪問看護　□看護職員の訪問による相談・支援　□訪問歯科診療　□訪問薬剤管理指導　□訪問リハビリテーション　□短期入所療養介護　□訪問歯科衛生指導　□訪問栄養食事指導　□通所リハビリテーション　□その他の医療系サービス（　）

（６）サービス提供時における医学的観点からの留意事項
・血圧　□特になし　□あり（　）・移動　□特になし　□あり（　）
・摂食　□特になし　□あり（　）・運動　□特になし　☑あり（　）
・嚥下　□特になし　☑あり（　）・その他（　）

（７）感染症の有無（有の場合は具体的に記入して下さい）
□無　□有（　）　☑不明

５．特記すべき事項

要介護認定及び介護サービス計画作成時に必要な医学的なご意見等を記載して下さい。なお、専門医等に別途意見を求めた場合はその内容、結果も記載して下さい。（情報提供書や身体障害者申請診断書の写し等を添付して頂いても結構です。）

長年夫以外との交流が少なかったためか、支援者を拒否する傾向がつよい。3〜4年前に胃ガンの手術を受けているようですが、その後の経過観察受診は中断しています。在宅生活を維持するためには週4〜5日、いずれかの支援者が本人宅に様子を見にいくくらいの支援が必要となります。