

2025 年 6 月 10 日

リコー、GENIAC でマルチモーダル LLM の基本モデルを開発完了 ～日本企業の知の結晶であるドキュメント活用を目指し、7 月に無償公開～

株式会社リコー（社長執行役員：大山 晃）は、経済産業省と国立研究開発法人新エネルギー・産業技術総合開発機構（NEDO）が実施する、国内における生成 AI の開発力強化を目的としたプロジェクト「GENIAC (Generative AI Accelerator Challenge) ※1」において、マルチモーダル大規模言語モデル（以下「LMM」）の開発に取り組んできました。このたび、リコーは日本企業の図表を含むドキュメントの読み取りに対応した LMM の基本モデルの開発を完了したことをお知らせします。視覚とテキストの両方の情報を活用する日本語の質問応答データセットである JDocQA および独自ベンチマークツール※2 による検証の結果、他のモデルと比較しても優れた性能を示すことが確認されました（2025 年 4 月 24 日時点）。リコーは、7 月 29 日から開催される画像の認識・理解シンポジウム「MIRU2025」にて論文を発表し、本基本モデルおよび独自に開発したベンチマークツールを無償で公開します。

1. 取り組みの背景と社会課題

LMM とは、テキスト・画像・音声・動画など複数の種類のデータを同時に処理できる AI 技術のことです。スクリーンショットからのテキスト要約や、図を用いた質問への適切な回答など、さまざまなタスクにおいて高い性能を発揮し、幅広いデータ形式を効果的に処理する適応性に期待が集まっています。

企業内で蓄積されるドキュメントには、請求書や領収書などのトランザクションデータ、事業戦略や計画などの経営資料、サービスマニュアルや社内で定めた技術標準、品質管理基準などの技術文書など、さまざまな形式があります。これらのドキュメントには、テキストだけでなく、図や表組、画像なども含まれています。企業内文書の効率的な利用や新たな価値・イノベーションを生み出すための活用が期待されていますが、「テキスト検索でも意図した結果が得られない」、「検索のみの利用では効果的な活用ができていない」などの課題があります。

また、近年の労働力の減少に対応した効率的な働き方や、ベテラン社員の退職に伴う技能伝承、外国人労働者の増加に伴う社内文書の多言語化への対応などの経営課題に対して、性能が著しく向上している AI を活用し、より効率的で、付加価値のある企業知の活用のニーズが高まっています。既存の LLM/LMM では、きめ細かな画像認識を必要とするビジネス文書の読解精度に課題があり、リコーは GENIAC において、日本企業で活用できる LMM の開発に取り組んでいました。

2. 今回の成果

リコーは、本基本モデルの開発に際して、文字/円グラフ/棒グラフ/フローチャートなどの視覚データ合計 600 万枚以上を人工生成しました。学習用データの人工生成手法を確立することで、大量の学習用データの整備が可能となり、LMM の性能向上に寄与しています。

株式会社リコー <https://jp.ricoh.com/>

報道関係のお問い合わせ先 広報室 TEL：050-3814-2806（直通） E-mail：koho@ricoh.co.jp

開発パートナーの1社であるFastLabel株式会社(代表取締役 CEO:鈴木健史、以下「FastLabel」)からは、学習用および評価用の実データの収集とアノテーション※3 業務の支援を受けました。FastLabel とは、リコーが主催するアクセラレータープログラム「TRIBUS(トライバス)2022」での採択がきっかけで、両社の協業が始まりました。

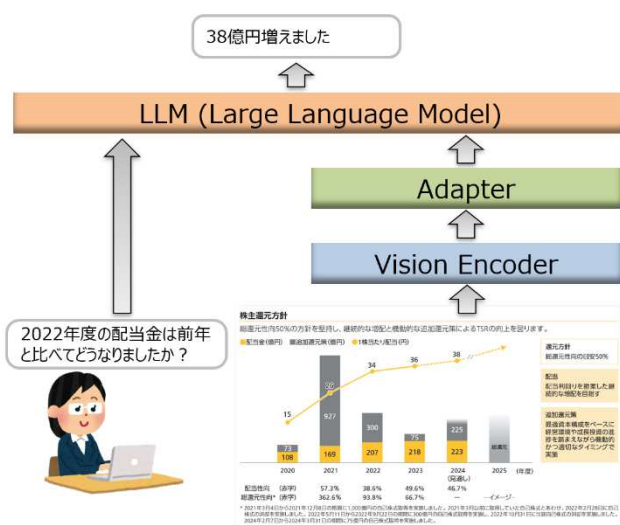
また、セキュリティやプライバシー、ガバナンスなどの観点から、オンプレミスや自社データセンターなどの社内専用環境で AI を利用したいと考える企業も多く、省リソースでの AI 活用のニーズが高まっています。リコーが開発したLMMは、アーキテクチャも改良され、オンプレミス環境において、お客様情報を用いた追加学習が可能なコンパクトなモデルサイズを実現しています。具体的には、本モデルは、図表を処理する Vision Encoder (ビジョンエンコーダー)※4 という第1階層と、第1階層からの出力を後段の LLM が理解できる形式に変換する Adapter(アダプター)という第2階層、そして第2階層で変換された情報と文字情報を統合処理する LLM の第3階層の3層構造になっています。第1階層においては複数の Vision Encoder を評価し、「Qwen2-VL-7B-Instruct※5」に採用されている手法を選定しました。また、第3階層においてもオンプレミス環境で実装可能な 70B クラスのモデルを複数評価し、「Llama-3.1-Swallow-70B-Instruct-v0.3※6」を選定しました。さらに、第2階層である Adapter に独自の工夫※7を追加することにより、別々のモデルが由来であるため本来は接続できない第1・第3階層を、精度を維持しながら接続することに成功しました。

このように、アーキテクチャを改良する独自技術と大量の人工データによる学習を組み合わせることで、同規模のオープンソースモデルを凌ぐ性能を確保しながら、省コスト・省リソースでの運用を実現します。

評価手法としては、JDocQA に加えて、日本企業の文書に特有な複雑な図表をテストデータの中心に据えた、マルチモーダル性を評価する評価手法も独自に開発し、本モデルの性能を確認しました。

リコーは、本開発にあたり、アマゾン ウェブ サービス(以下、AWS)※8 から、計算資源の提供及び技術サポートを受けています。「採択事業者が計算リソース提供事業者と個別に調整し直接確保」するスキームを通じて、NVIDIA H200 Tensor Core GPU を搭載する Amazon EC2 P5e インスタンスが提供されています。開発終盤での総合的な学習ならびに顧客向けファインチューニングにおいて、安定的で大規模な資源提供を迅速に受けることで、今回の成果につながっています。

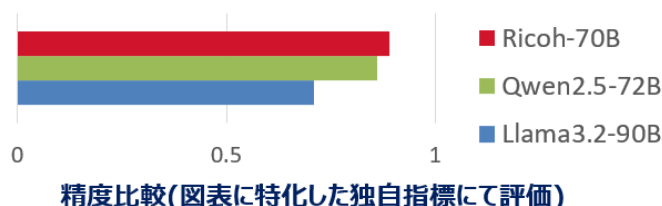
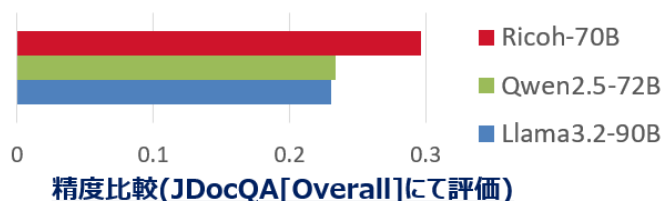
リコーは、これまでの LLM の開発においても、「AWS LLM 開発支援プログラム」と「AWS 生成 AI イノベーションセンター(AWS Generative AI Innovation Center)」による支援のもと、効率的な開発を実現しています。



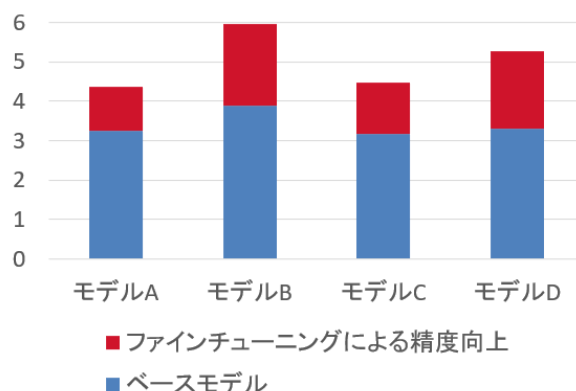
なお、今回の取り組みにおいて、基本モデルをチューニングして個社の業務に合わせて精度を向上させる手法も確立しました。具体的には、損害保険ジャパン株式会社(代表取締役社長:石川 耕治、以下「損保ジャパン」)が保有する保険の引受規定が記載された図表などを含むマニュアルを用い、同社の保険業務に対応するように学習(ファインチューニング)を行ったプライベートなモデルは、基本モデルに比べて顕著に性能が向上しました。今後、同社がトライアル運用する保険業務に関する規定、マニュアル、Q&A データなどを学習させ、社内外からの照会内容に対して最適な回答案を自動生成するシステム「おしそん LLM※9」に適用を検討していく予定です。

3. 今後の予定

リコーは、基本モデルと独自に開発したマルチモーダル性能を評価する環境を無償で公開します。リコーは、GENIAC で得られた成果を広く社会に還元することで、日本企業の知の結晶ともいえるドキュメントの利活用を促進し、業務革新と効率的で付加価値の高い働き方を支援することで、日本企業の企業価値の向上に貢献することを目指して取り組みを進めてまいります。



ベンチマークツールにおける他モデルとの比較結果
(上段は一般的な指標、下段はリコーが独自開発した指標による評価)



顧客データによるチューニングの効果

チューニングの成果

(モデルが異なってもファインチューニングによりグラフの赤い部分の精度向上が見られる)

アマゾン ウェブ サービス ジャパン合同会社 常務執行役員 サービス & テクノロジー統括本部 統括本部長 安田 俊彦様からのコメント

リコー様が推進する高性能なマルチモーダル LLM の開発に、AWS がご支援できたことを嬉しく思います。AWS LLM 開発支援プログラムならびに AWS 生成 AI イノベーションセンターを通じたリコー様への包括的な支援が企業内文書の高度活用という課題解決に寄与し、リコー様のデジタルサービス会社への変革が加速することを期待しています。

FastLabel 株式会社 代表取締役 CEO 鈴木 健史様からのコメント

リコー様における GENIAC のマルチモーダル LLM 開発プロジェクトにおいて、パートナーとして参画できたことを大変光栄に思います。今回開発された基本モデルは LLM の性能向上に直結し、国内企業において業務効率化や生産性向上に大きく貢献する可能性を持った取り組みになるであろうと強く感じています。FastLabel では、リコー様をはじめとした LLM の開発を行う企業に、AI 開発のデータ作成という観点で質の高い支援を引き続き行なってまいります。

株式会社リコー リコーデジタルサービスBU AIサービス事業本部 本部長 梅津 良昭からのコメント
FastLabel 様には学習・評価に使えるデータの大量収集とアノテーションを実施いただきました。これら質の高いデータにより、モデル開発が加速し、ベンチマークの公開も可能になりました。誠にありがとうございます。

また、AWS 様には先進的な AI チップ「Trainium」を 512 台規模で活用した LLM 開発などでお世話になってきましたが、今回、新たな挑戦である国産 LLM (大規模マルチモーダルモデル) の開発においても、高速で安定した GPU 環境をご提供いただくだけでなく、先進的な AI 技術のご支援や多大なサポートを賜り、心より感謝申し上げます。

今後も両社様との連携により、AI 研究開発をさらに加速させていきたいと考えておりますので、引き続きよろしくお願いいたします。

※1 GENIAC (Generative AI Accelerator Challenge) :リコーは、NEDO が、2024 年 7 月に実施した公募「ポスト 5G 情報通信システムの開発／競争力ある生成 AI 基盤モデルの開発(助成)」(以下、「本事業」)に採択されました。本事業では、主に生成 AI のコア技術である基盤モデルの開発に対する計算資源の提供や、データや AI の利活用に向けた実証調査の支援等が行われます。

※2 ベンチマークツールについて:図表を含む日本語文書をもとにして、視覚情報とテキスト情報の両方を参照する質問応答データセット JDocQA とリコーで独自開発したデータセットの 2 種類のベンチマークツールを活用して評価しました。

https://www.anlp.jp/proceedings/annual_meeting/2024/pdf_dir/C3-5.pdf

※3 アノテーション(annotation): データにタグを付与して整理し、AI が正しく認識できるように加工するプロセス。AI が正しく学習できるよう、データに「正解」や「特徴」などを注釈のように書き込むことで、AI がデータから何を学習すべきかを明確にします。

※4 Vision Encoder(ビジョンエンコーダー): 図表などの視覚情報を言語モデルが理解できる形に変換するモジュール。

※5 Qwen2-VL-7B-Instruct: <https://huggingface.co/Qwen/Qwen2-VL-7B-Instruct>

※6 Llama-3.1-Swallow-70B-Instruct-v0.3: <https://huggingface.co/tokiyotech-llm/Llama-3.1-Swallow-70B-Instruct->

[v0.3](#)

※7 佐藤諒,木下彰,中田乙一,金箱裕介,麻場直喜. 大規模マルチモーダルモデルにおけるビジョンエンコーダーの付け替えと、日本語に強いモデル作成へ向けて. 言語処理学会 第31回年次大会 発表論文集. pp.954-959,2025.

※8 AWS による支援について: <https://aws.amazon.com/jp/blogs/news/geniac-cycle2-kick-off/>

※9 おしそん LLM について: 損保ジャパンのニュースリリース「大規模言語モデルを活用した照会回答業務の効率化トリアルの実施」 https://www.sompo-japan.co.jp/-/media/SJNK/files/news/2024/20241022_1.pdf

■リコーの AI 開発について

リコーは、1980 年代に AI 開発を開始し、2015 年からは画像認識技術を活かした深層学習 AI の開発を進め、外観検査や振動モニタリングなど、製造分野への適用を行ってきました。2021 年からは自然言語処理技術を活用し、オフィス内の文書やコールセンターに寄せられた顧客の声(VOC)などを分析することで、業務効率化や顧客対応を支援する「仕事の AI」の提供を開始しました。

さらに、2022 年からは大規模言語モデル(LLM)の研究・開発にもいち早く着手し、2023 年 3 月にはリコー独自の LLM を発表。その後も、700 億パラメータという大規模ながら、オンプレミス環境でも導入可能な日英中 3 言語対応の LLM を開発するなど、お客様のニーズに応じて提供可能なさまざまな AI の基盤開発を行っています。また、画像認識や自然言語処理に加え、音声認識 AI の研究開発も推進し、音声対話機能を備えた AI エージェントの提供も開始しています。

■関連ニュース

リコー、経済産業省の国内生成 AI 開発力強化プロジェクト「GENIAC」に採択

https://jp.ricoh.com/release/2024/1017_1

リコーと損保ジャパン、マルチモーダル LLM の共同開発を開始

https://jp.ricoh.com/release/2025/0321_1

■関連リンク

リコーの AI 開発の最前線:企業の AI 活用を広げる「マルチモーダル LLM」

<https://jp.ricoh.com/news/stories/articles/multimodal-llm>

TRIBUS オウンドメディア:【Challenger's Interview】日本を再び「世界レベル」へ。独自の AI インフラで新たな価値を創造する FastLabel

<https://accelerator.ricoh/2024/01/12/fastlabel>

※社名、製品名は、各社の商標または登録商標です。

｜リコーグループについて｜

リコーグループは、お客様のDXを支援し、そのビジネスを成功に導くデジタルサービス、印刷および画像ソリューションなどを世界約200の国と地域で提供しています(2025年3月期グループ連結売上高2兆5,278億円)。

”はたらく”に歓びを 創業以来85年以上にわたり、お客様の“はたらく”に寄り添ってきた私たちは、これからもリーディングカンパニーとして、“はたらく”の未来を想像し、ワークプレイスの変革を通じて、人ならではの創造力の発揮を支え、さらには持続可能な社会の実現に貢献してまいります。

詳しい情報は、こちらをご覧ください。

<https://jp.ricoh.com/>