

補助事業成果報告書

補助事業の名称	視覚障害者・ディスレクシアのための音声を使った読書方法の研究開発
補助事業の概要	以下に記す

【研究開発の実施内容と成果】

1. 本年度研究の概要

YourEyes (<https://youreyes.jp>) は、老眼や弱視を含む広義の視覚障害者、さらには識字障害の一種であるディスレクシア症例を持つ人への読書支援サービスである。2021 年 12 月にプレスリリースが株式会社ポニーキャニオンより行われ、2022 年 4 月に一般に公開された。その後、2022 年、ポニーキャニオンから株式会社スプリュームへ運営会社の変更があったが、弊社想隆社は引き続き開発チームとして本プロジェクトのコア技術の研究開発に関わっている。

令和 5 年度は、令和 3 年度、4 年度のこれまでの 2 年間の研究を踏まえ、全体的な精度の向上と、実用上の利便性の向上を目指した。また、昨今 ChatGPT を始めとする大規模言語モデル (Large Language Model、以下 LLM) の発展が目覚ましく、LLM を使ってこれまで光学的なアプローチだけでは解決できなかった問題を、言語処理して解決し、文章の読取りから音声化までの精度の向上や利便性向上を図った。

特に音声読み上げの修正は、ボランティアツールで、人によってしか修正できない部分があり、これを LLM を使って自動化することは大きな効率化の期待があった。

以上のことから、2 年間の研究期間で、視覚障害者に実際に YourEyes を利用してもらい、いただいた意見をもとに、本年は以下をサブテーマとして研究した。

- ① OCR 誤りを含むテキストの整形
- ② 書籍画像内の挿絵や写真の検出とキャプションの生成
- ③ テキスト内容の要約

2. 各サブテーマの結果

① OCR 誤りを含むテキストの整形

これまでの利用者アンケートから、ユーザは、テキストのタイプミスや OCR の読み取りミスに頻繁に直面している現状があったが、読み取りを光学的に解決しようと考えても限界があった。これを LLM の自然言語理解能力及び推論能力によってテキスト内の誤りを推定し、修正したテキストを生成するという研究を行った。このアプローチでは、従来形が似ている文字、例えば、「待」と「侍」、「島」と「鳥」、「の」と「@」などのこれまでの光学的なアプローチでは難しかった修正が可能になる。以下がその一例であるが、正しい文字が置き換わることがあるものの、文意は保たれるという効果が認められた。まだ実用上の効果は少ないかもしれないが今後 LLM の発展によって大きく精度が向上すると考えられる。

修正例 1: 芥川龍之介「トロッコ」より

原文

往きと返りと変るせいか、景色の違うのも不安だった。すると今度は着物までも、汗の濡れ通ったのが気になったから、やはり必死に駆け続けたなり、羽織を路側へ脱いで捨てた。蜜柑畑へ来る頃には、あたりは暗くなる一方だった。「命さえ助かれば——」良平はそう思いながら、這ってもつまずいても走って行った。

誤りを含む OCR 結果

往きと返りと変るせいか、景色の違うのも不安だった。すると今度は着物までも、汗の濡れ通ったのか気になったか ら やはり必死に飪け続けたなり羽織を路側へ脱いで捨てた。
蜜柑畑へ来る頃には、あたりは暗くなる一方だった。`命さえ助かればー」良平はそう思いながら、しつてもつまずいても走って行った。

LangChain を導入し ChatGPT-3.5 を使った修正結果

往きと返りと変わるせいか、景色の違うのも不安だった。すると今度は着物までも、汗の濡れ通ったのか気になったかもしれないが必死に飪け続けたなり羽織を路側へ脱いで捨てた。
蜜柑畑へ来る頃には、あたりは暗くなる一方だった。「命さえ助かればいい」と良平はそう思いながら、走って行った。

② 書籍画像内挿絵写真の検出とキャプションの生成

昨年まで行ったレイアウト解析の手法を発展させ、版面の中の挿絵や画像、図表の検出を行い、検出した部分に対して AI によって図の説明文を生成した。

その方法は、まず、昨年の研究同様、学習データとして文庫本の小説や新書のエッセイ、技術書やハードカバーなど様々な形式の書籍の画像データを学習させて YOLOv8m という物体検出モデルをファインチューニングする。そして、これにより生成された機械学習モデルを使って自動で挿絵、写真、図表等の検出とその説明文の生成を可能とするというものである。

以下はその一例であるが、画像の矩形が検出できており、さらにそのパーツがどのようなもの（図表、イラスト）かのラベル付けがされていることがわかる。

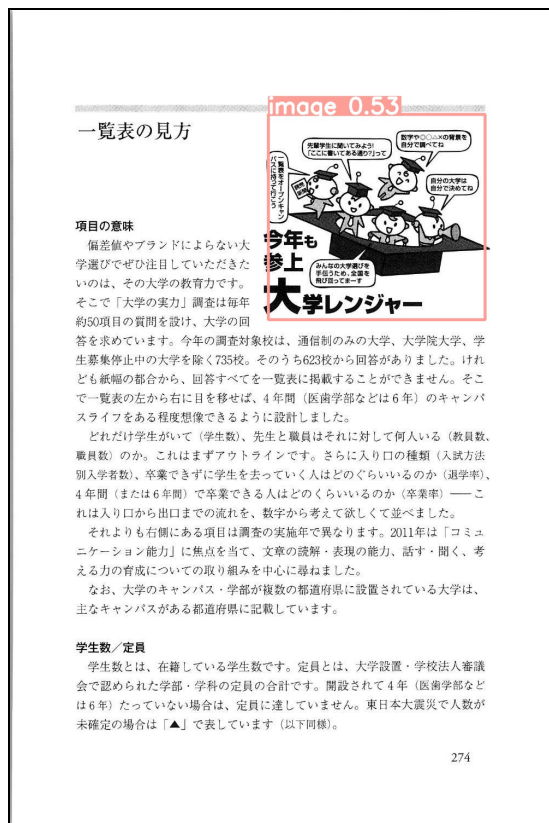


図 1 物体検出による図版検出



図 2 キャプションの自動生成

出典：「中小企業の事業継承」牧口晴一、 斎藤孝一

さらに、これらの物体検出の結果を商用 API などのモーダル AI を用いると、次のように検出したイラストに対してキャプションを得ることができるようになった。この例では、AI によって作成された 「車のハンドルを握っている男性」という音訳用説明文が自動作成された。

商用 API は様々なものを比較したが、今後そのコストや性能は大きく発展すると考えられ、本機能もより正確に、人間を満足させるものに近づくと考えている。

③ テキスト内容の要約

晴眼者では、文庫の裏表紙にその本のあらすじなどが書かれていて、読書前に書籍の概要やあらすじを知ることができる。YourEyes では、書籍のメタデータにそのようなデータがあればその情報を得ることができる。章ごとや節などの書籍の部分ごとに内容の要約を取得し、もう少し細かいサマリーが欲しいという意見をもとに、本機能を実装した。

YourEyes は 1 ページごとにページをスキャンする逐次処理であるが、一定量をバッファリングし、内容から LLM を使ってあらすじ等を生成する機能として実装した。

下記は太宰治「走れメロス」の要約で、本文は 1300 字程度である。ページをめくるとに要約を作成し、書籍の内容を的確に表現できるサマリーが自動生成されることを確認した。

要約例：太宰治「走れメロス」

メロスは王を倒すために行動することを決意し、市に向かう途中で様々な困難に直面する。彼は友人のセリヌンティウスを救うために走り続け、最終的には間に合う。セリヌンティウスを解放し、二人は抱擁し喜び合う。暴君ディオニスも感銘を受け、メロスとセリヌンティウスを仲間に加えることを願う。群衆も彼らを称え、メロスには少女からマントが贈られる。メロスは恥ずかしさを感じながらも、マントを受け取る。

3. 被験者による評価

前述の 3 つの機能に対して YourEyes を利用したことのあるユーザ 6 名を対象に被験者実験を行い、評価を行った。

3.1 評価実験① OCR 誤りを含むテキストの整形の評価

書籍画像を写真撮影し OCR 処理を行って誤りを含むテキストと、誤りを LLM の操作用モジュールである LangChain を用い修正したテキストの、それぞれに対して評価を集めた。評価は「耳で聞いたときの理解のしやすさ」を 5 段階評価で行った。選択肢 1 非常に読みにくい、選択肢 2 読みにくい、選択肢 3 普通、選択肢 4 読みやすい、選択肢 5 非常に読みやすい、である。使用したテキストは、芥川龍之介「トロッコ」、江戸川乱歩「怪人二十面相」の一部である。LLM による修正により、読みやすさの改善がされたことがわかる。

	評価の平均値
OCR 誤りテキスト	2.75
プロンプトでの修正テキスト	3.5
LangChain を用いた修正テキスト	3.83

表 1 ユーザによる修正前と修正方法ごとの平均評価値

3.2 評価実験② 書籍画像内の挿絵や写真の検出とキャプションの生成

キャプションを含まないテキスト、キャプションを挿入したテキストを用意し、評価を行った。使用した書籍は、美馬のゆり「理系女子的生き方のススメ」、安部公房「壁」で、評価はキャプションを含むものと含まないもの、情景が理解しやすいと感じた方を選択してもらった。キャプション付きの文が選択された割合は 66.6%であった。ユーザ評価の結果から、画像キャプションの有効性が確認できた。

3.3 評価実験③ テキスト内容の要約

テキストとその要約テキストを読んでもらい、要約の妥当性を 5 段階で評価した。テキストは星新一「表と裏」「健康な犬」、太宰治「走れメロス」の全文を用いた。評価は 4.0 に近いものもあり、有効であると考えられる。

	テキスト長さ	評価の平均値
「表と裏」	496	3.83
「健康な犬」	1260	4.0
「走れメロス」	10073	2.67

表2 ユーザによる AI による要約文の平均評価値

4. 成果

本年度研究では、視覚障害者読書支援システム YourEyes の利便性と性能向上を目的に、LLM の特性を考慮した新しい読書支援機能の開発と検証を行った。これらの要素技術は、API として YourEyes から利用できる形で実装を行った。またこの API を使って作成したテストデータによるユーザ評価アンケートを実施し、その有効性が確認された。

3 年を通じて、YourEyes に応用可能な読書支援の研究を行ってきたが、おおむね機能は精度の向上と利便性の向上が定量的にも定性的にも認められた。今後サービスに実装していきたいと考えている。

5. 研究業績

5.1 受賞

日本電子出版協会の JEPA アワードの部門賞、全部門から年の1位を決定する大賞を受賞した。大賞受賞時の他のノミネート（部門賞）はマイクロソフト社、Adobe 社、小学館らのサービスであり、YourEyes がこれら有名企業のサービスを押しつけて大賞を受賞したことは出版界自体のアクセシビリティへの期待がうかがえる。

5.2 研究発表

- ① 山本幸太郎 . AI2oT 講座 『AI プログラミング概論 - 実務で使う機械学習』 . エレクトロニクス実装学会
- ② 山本幸太郎 . 特別講演(異分野ECE連携プログラム) 『AI プログラミング概論-実務で使う機械学習』 . 計測自動制御学会
- ③ 小林 大晟 . 読書支援システムにおけるレイアウト解析及び LLM を用いた機能開発 . 画像電子学会 VMA 研究会（研究補助者の学生発表）