

漫画の高精度な機械翻訳技術および多言語漫画配信サービスの実現

—世界の言葉でマンガを届ける—

1. 背景

日本漫画の市場は国内外合わせて6,000億円と大規模であり、文化的・産業的な価値が高い。一方、米国では1年間の海賊版漫画被害額が1兆円を超えるなど、日本国外における漫画の海賊版被害は甚大である。この被害額は国外における日本漫画の売上1,500億円のおよそ10倍と大きく、世界中の漫画読者を海賊版から正規版へ誘導することは、出版業界にとって喫緊の課題である。

本プロジェクトで日本国外の漫画ファンを対象に行なった調査においては、海賊版ユーザーのうち6割は「正規に翻訳されていないから」「正規版翻訳よりも早く入手できるから」海賊版を利用していると回答している。すなわち、海外の海賊版を駆逐するためには、法的な規制を強めることよりも、ユーザーが利用しやすい(すなわち、安価で更新頻度が高く、作品の種類が多い)コンテンツ配信サービスを提供することの方が有効である。しかし、漫画の翻訳は単行本一冊で20万円程度と極めて高コストであり、かつ長い時間がかかる。

2. 目的

本プロジェクトの目的は、海賊版による漫画産業・文化の破壊の防止と、海外における新たな市場の創出である。そのために、魅力的な漫画を多言語配信するサービスを展開し、海賊版漫画の読者を正規版コンテンツへ誘導することを目指す。多言語配信において課題となる翻訳コストを削減するため、高精度な漫画自動翻訳技術を実現する。

3. 製品・サービスの内容

3.1 高精度な漫画翻訳技術

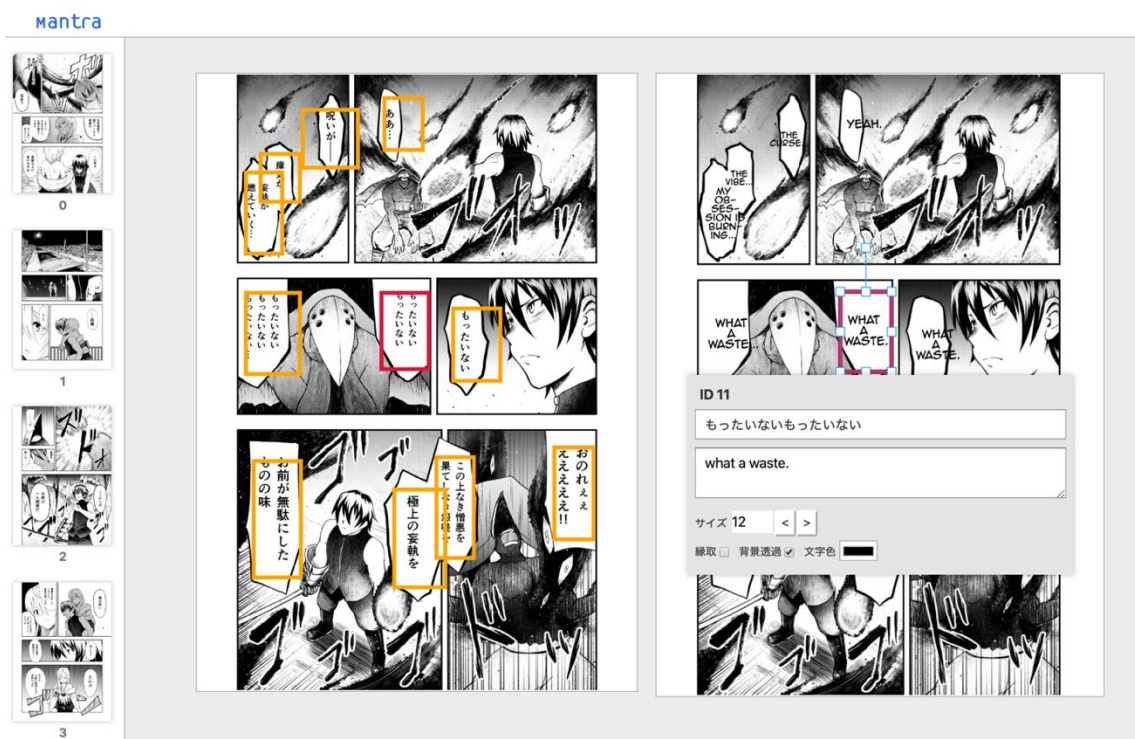
文脈を考慮した翻訳技術のプロトタイプを開発した。従来の単文翻訳では主語を適切に補完できない、一つの文が複数の吹き出しに分割されている場合にうまく翻訳できない等のエラーがあった。文脈を考慮した翻訳により、このような文脈を考慮できていないことに起因するエラーのうち55%を改善することができた。また、漫画中の文字認識についても、既存のシステムでは全く認識することができない「書き文字」の認識技術の開発を行い、半分以上のテキストについて完璧に認識できる精度を達成した。(下図参照。)



3.2 翻訳修正ツール

現状では、自動翻訳技術で100%正確に漫画を翻訳することはできない。そのため、

実際に自動翻訳された漫画を出版・配信するためには、自動翻訳の結果を手で修正する工程が不可欠である。通常、漫画の翻訳は非常に複雑な作業工程で行われている。まずエクセル等を用いて翻訳者が翻訳を行い、その後、写植の担当者がそのテキストをフォトショップ等のソフトウェアを用いて漫画上に埋め込み、最後に校閲者が全体を見て翻訳の修正を行う。このような煩雑な作業工程による翻訳コストを削減するために、簡単に人手で翻訳結果を修正することのできるツールを試作した。(下図参照。)

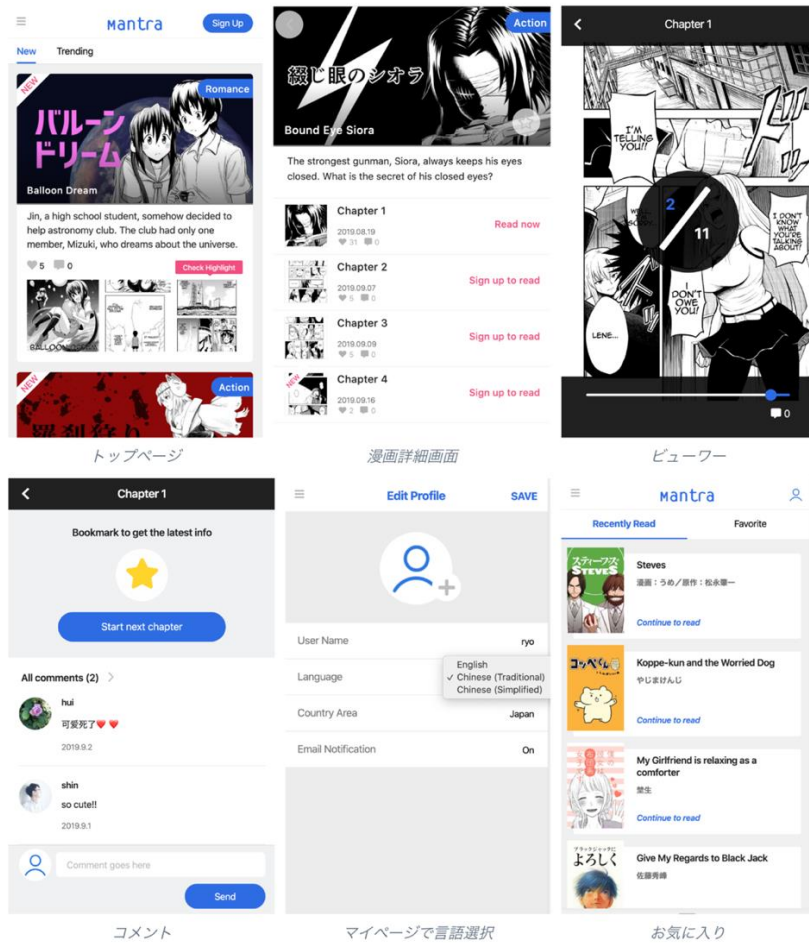


3.3 多言語漫画配信サービス

機械翻訳エンジンおよび上記の試作ツールを用いて翻訳した漫画を海外の読者に届けるため、多言語漫画配信サービスを開発した。一般的な漫画配信アプリと同様、以下のようなページ・機能を持つ。

- ・ トップページ: 新しく更新された「新着」の漫画、直近一週間で最もいいねが集まった「急上昇」の漫画のリストが表示される。ユーザーはリストから読みたい漫画を探して選択する。
- ・ 各漫画の詳細ページ: 各漫画の概要とチャプター一覧を表示。
- ・ お気に入り: ユーザーは好きな漫画をお気に入り登録すると、新しい話が更新されたタイミングでメール通知を受け取ることができる。
- ・ ビューワー: 一般的な漫画アプリのビューワーと同様、横スワイプでページめくりするなどして直感的に漫画を読むことができる。
- ・ いいね、コメント: 一話ごとにユーザーはいいね・コメントをすることができる。
- ・ SNS でシェア: Twitter や Facebook、Weibo などを読んだ漫画をシェアすることができる。

配信サービスのサンプル画面を下図に示す。



4. 新規性・優位性

開発した漫画の文字認識エンジンおよび機械翻訳エンジンは、ともに 2020 年 2 月現在、世界最高精度を達成している。文字認識は、Google Cloud Vision API の精度 21.5% に対し、提案技術は 60.1%；機械翻訳では、Google Translate の BLEU スコア 4.37 に対し、提案技術は BLEU スコア 8.00 の性能を達成した。本プロジェクトの翻訳エンジンは、コマ割りや吹き出しの位置といった画像情報を認識することで、複数の文にまたがる文脈の認識を行う点が従来の機械翻訳技術と本質的に異なる。また、本プロジェクトで作成した多言語配信サービスは、日本漫画を英語・中国語で同時に配信することのできる、世界初のプラットフォームである。

5. 事業普及（または活用）の見通し

前節で述べた本プロジェクトの成果物を活用し、(1)「漫画翻訳エンジン」「漫画 OCR エンジン」の技術提供、(2) 安価で高速な「定額翻訳」の提供、(3)「多言語漫画配信サービス」のリリースを予定している。この3つの事業それぞれについて、普及の見通しを以下に述べる。

(1)「漫画翻訳エンジン」「漫画 OCR エンジン」の技術提供

現在、3社の出版社への導入が検討されており、うち1社に対しては見積もりを発行済である。

(2) 「定額翻訳」の提供

これまでに、国内出版社1社に対しての有償での翻訳代行サービスの提供を行った。また、中国の大手漫画配信サービスとのPoCも予定している。

(3) 「多言語漫画配信サービス」

16の作品を対象とし、ベータ版 (<https://pf.mntr.jp>) の公開を行った。当該サービスのユーザー登録者数はリリースから5ヶ月で713人に達し、このうち22名のユーザーへのヒアリングを通して今後の事業展開の方針検討を行うことができた。

6. 期待される波及効果

6.1 漫画の多言語出版・配信方式の転換

漫画の人手翻訳は高コストであるがゆえに、コスト回収が確実に見込める人気漫画以外に対しては、翻訳版の制作や海外での翻訳権・出版権のライセンスは行われにくいという課題があった（出版社、印刷会社、翻訳会社へのヒアリングより）。本提案の成果は翻訳のコストを大幅に削減するため、漫画の多言語出版は「国内で売れた極少数の作品にのみ翻訳コストをかけて海外に出す」方式から「最初から国内外問わず全世界に向けて配信する」方式へのパラダイム転換が起こることが期待される。

6.2 学術的意義および関連産業へのインパクト

「画像から文脈を読み取りながら翻訳する」手法は斬新であり、学術的観点からも未踏性が高い。当該技術は今後、漫画以外の機械翻訳（例：リアルタイム音声翻訳技術や動画コンテンツの翻訳）へ応用される可能性もあり、関連産業へのインパクトが期待される。

7. イノベータ名（所属）

石渡 祥之佑（Mantra 株式会社）

日並 遼太（Mantra 株式会社）

（参考）関連 URL

漫画翻訳エンジンデモ：<https://mntr.jp/#demo>

多言語漫画配信サービスデモ：<https://pf.mntr.jp>