
報告書

1 今週の進捗

- 翻訳タスク
- llama.cpp の動作確認
- コマの抽出

2 翻訳タスク

- PDF から画像を抽出して, その画像にある日本語を他言語に翻訳, テキストを貼り付けた.
- 画像解析および翻訳に使用したモデルは Gemini の "gemini-2.5-flash-preview-04-17".
- 入力文とシステムプロンプトは以下の通り:

2.1 入力文

Translate into English.

2.2 システムプロンプト

Translate the Japanese text inside the boxes in the image into {language}.

This is reference material for anime production, derived from original anime drawings.

If there is no text in the specified area, return "None".

Additionally, display only the translated content.

Be mindful of expressions unique to anime.

If there are any misrecognitions, please ignore them.

Break lines every three words in the translated result.

2.3 結果画像

2.4 問題点

- PDF が多くなると (40 枚程度) Google のサーバとの接続が切れ, 処理が途中で止まる問題が発生.

3 llama.cpp

- コマンドラインで llama3 の動作確認ができた.
- これからコード内でマルチモーダル入力に対応したモデルを動かせるようにする.

4 コマの抽出

5 今後

- llama.cpp のマルチモデル対応の実装
- Google サーバ接続の安定化方法の検討
- コマ抽出の精度向上

参考文献