

ザの言語設定に応じた適切な案内文を生成し、アプリ側に返送する。

3.2 認識モジュール

現在試作中の認識モジュールは、Open AI 社の GPT-4o API を用いたプロンプト設計により実装されている。プロンプトの冒頭では、本タスクが「切符認識システムのスキナー」であることを明示し、入力画像に基づいて内容を判定するよう指示している。

また、プロンプト内で認識すべき対象を明確化するため、補助的な内容として、以下の情報を指定している：

1. 判別対象となる 7 種の鉄道切符種別
2. 領収書・利用票など 2 種の類似票類
3. 出力に用いるフォーマット (JSON テキスト形式、出力例付き)
4. 情報が記載されていない場合の処理方針 (「None」として出力する)

3.3 判断モジュール

判断モジュールは、「情報整形器」と「行程判定器」の二つから構成される。情報整形器は、認識モジュールから出力されたテキスト情報を構造化し、内部的なオブジェクトとして整形する役割を担う。認識モジュールでは、あらかじめ定義された出力フォーマットを用いて情報抽出を行う設計となっているが、GPT-4o などの大規模言語モデルの特性上、すべての出力が厳密にフォーマット通りとは限らない。そのため、整形器では出力の読み取りに失敗した場合、自動的に再度認識モジュールを呼び出して情報を再生成するリトライ機構を備えている。

一方、正しく情報が抽出・整形された場合は、その内容がオブジェクトに格納され、続く行程判定処理に渡される。行程判定器では、主要鉄道事業者の旅客営業規則および切符の使用規則に基づき、券種の組み合わせや利用可否を判定し、最終的にユーザに対する適切な案内文を生成し出力する。

4. 認識実験と考察

4.1 実験方法

本研究では、筆者らが作成した切符画像データセットの拡張版を用いて認識実験を行った。

拡張データセットには、元のデータに私鉄各社から発行された切符や領収書などの画像 7 枚を追加しており、データ作成手法は主データセットと同一である。これにより、異なる鉄道事業者に対しても本システムが有効に機能するか、という汎用性を検証することを目的とした。

本実験では、データセット全体に対する 1 回の評価を 1 エポックと定義し、計 10 エポックを実行した上で、各評価指標の平均値を算出した。

4.2 評価基準

本実験は情報項目の認識正答率を用いた。ここでいう情報項目とは、本研究において指定した 6 つの情報要素「券種」「出発駅」「到着駅」「席種」「席番号」「電車番号」を指す。これらの要素は各認識処理において、必ず出力される仕様となっている。

$$\text{認識正答率} = \frac{\text{情報を正しく認識した要素数}}{\text{総要素数}}$$

本システムでは認識モジュールの非標準的な出力が得られた場合に再認識を実行するため、認識正答率はフォーマットが正しい最終出力に基づいて評価を行っている。これは、実際の利用者が最終的に得る結果と一致するため、現実的な指標といえる。

一方で、再認識の発生はユーザの待機時間の増加につながるため、その影響を測る補助指標として、再認識の発生率もあわせて計測する。

$$\text{再認識の発生率} = \frac{\text{再認識の総回数}}{\text{画像認識の総回数}}$$

4.3 実験結果と考察

実験の結果、試作バックエンドにおける情報抽出の認識正答率は、10 エポックの平均正答率で 83.0%となった。正答率の分散は 0.062 である。

また、10 エポックの合計 320 件の認識処理において、再認識が発生した回数は 33 回であり、再認識の発生率は 10.3%であった。

以上から、本システムは基本的な券面情報の認識能力を有している一方で、さらなる精度向上の余地も残されていると言える。特にデータの安定性に関しては、分散が小さいとは言えず、認識結果には一定のばらつきが見られた。

5. おわりに

本研究では、画像間の Cosine 類似度を算出する手法により、日本の鉄道において発行される複数種の切符および類似票類の間に高い視覚的類似性が存在することを示した。これは、訪日外国人が切符を誤って使用する背景に、客観的な困難要因が存在することを示唆するものである。

本研究では、外国人観光客が鉄道切符を円滑に使用するための支援システムの一部である画像認識と券種処理を担うバックエンド処理系のプロトタイプを作製し、技術的な実現可能性を検証した。また、認識精度に関する初期的な評価から、今後の改善に向けた課題と方向性を明らかにした。

今後の研究として、まず、支援システムに求められる機能要件を明確化するため、駅係員および訪日外国人観光客へのインタビュー調査を実施し、多角的な視点からのニーズ分析を進める。次に、プロンプト設計の最適化や画像伝送処理の改善を行い、認識精度の向上および再認識率の低減を目指す。最後に、スマートフォンアプリとバックエンドを接続した実用レベルのシステムの構築に取り組み、提案手法の社会実装を見据えた取り組みを進めていく。

参考文献

- [1] 藤原 あゆみ, 鶴心治, 小林剛士, 多田村 克己, “外国人観光客向け公共交通機関利用支援アプリの開発,” 画像電子学会第 292 回研究会講演予稿, pp. 18-22, 2020.
- [2] He, K., Zhang, X., Ren, S., and Sun, J., “Deep residual learning for image recognition,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 770-778, 2016.