

深層学習を用いた店舗のジャンル推定 に関する検討

東畑 湧登

† 龍谷大学先端理工学部

1. はじめに

近年、スマートフォンなどの端末の普及により、地図アプリによる現在地推定やナビゲーションが利用されている。一方で、屋内施設、特に、ショッピングモールなどの大型商業施設における現在地推定やナビゲーションは、多層で店舗が密集している複雑性や店舗の入れ替わりやレイアウトの変更によるインフラの変更などの原因により普及が進んでいない。この問題を解決するための手法として、文字検出やロゴ、外観を用いた現在地推定システムおよびナビゲーションシステムが提案されている。[1]

そこで、本研究では大型商業施設における画像認識を用いた現在地推定システムの精度を上げる付加機能として、ショッピングモール内における店舗の外見画像から深層学習による画像分類を用いて店舗のジャンルを推定することを提案する。

2. 提案手法

画像認識モデルの ResNet50 を転移学習させ、ショッピングモールの店舗のジャンルを識別できるようにする。

2.1 データセットの作成

ショッピングモール内の店舗の外見（正面）画像を集める。識別したい各クラスに対応するフォルダを作成しておく。本研究では、“accessory”、“bag”、“book”、“clothes”、“goods”、“restaurant”、“shoes”の7つを作成する。集めた画像をそれぞれ対応するフォルダに分類する。これによって、フォルダ名が自動的にラベルとして扱われる。

2.2 データ画像から人物除去

YOLOv5 を用いて、画像から人物を検出する。[2][3]検出した人物はバウンディングボックスで囲まれ、その範囲を黒く塗りつぶすことで人物を除去する。

2.3 ResNet-50 を転移学習

ImageNet で学習済みの ResNet-50 を作成した画

像データセットを用いて転移学習させる。[4]この時、データが少なく、偏りがあったため、5分割の k 分割交差検証で評価を行う。エポックごとに訓練時の損失と正解率、テスト時の損失と正解率を出力し、分類器の精度を見る。損失にはクロスエントロピー損失を用いる。

3. おわりに

ショッピングモール内の店舗のジャンル推定を行う手法を提案し、その方法としてデータセットの作成と画像認識モデルを転移学習させるものである。

現段階での課題は、全体としてのデータ数が少なく、クラスごとにデータ数に偏りがあったため、学習時の精度が異様に高くなってしまったりデータの少ないクラスの分類が正しくできていなかったりすることである。また、今回の研究では、1枚の画像に1店舗だけ写っているものを使用したため、1枚の画像に複数の店舗が写っている場合どのように処理をするのかも課題である。

参考文献

- [1] S. Wang, et al., “Lost Shopping! Monocular Localization in Large Indoor Spaces,” IEEE International Conference on Computer Vision, 2015, pp. 2695–2703.
- [2] J. Redmon, et al., “You only look once: Unified, real-time object detection,” IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.
- [3] Ultralytics, <https://github.com/ultralytics/yolov5>.
- [4] K. He, “Deep Residual Learning for Image Recognition,” IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.