

2-16 学内ナビゲーションに向けた特定物体画像からの撮影地点推定とその精度の向上

小林研究室
2119039 宮谷 翔

1. はじめに

本研究では、学内において、ユーザがスマートフォンによる現在地の撮影を行うことでその地点から目的地までの案内図を取得できるシステムを想定し、学内の表札の画像からその部屋名を推定するモデルを機械学習によって作成することを目的とする。

先行研究 [1] では、ある繁華街においてユーザが携帯端末で店舗の看板を撮影することで、リアルタイムにその店舗情報を見ることができるシステムを考案し、撮影された看板画像と店舗情報を結びつけるにあたり、機械学習を用いて看板領域の抽出、抽出された領域の分類をおこなった。本研究はこの手法を用いて、学内の表札領域を抽出し、その領域を分類することで現在地を推定する。

2. 機械学習

物体検出アルゴリズム YOLOv8[2] を用いた検出モデル、CNN のひとつである VGG16[3] を用いた分類モデルをそれぞれ学習する。検出モデルの学習データには画像と表札位置の組、分類モデルの学習データには表札の画像とその部屋名を用いた。学習後に検出モデルによって抽出された画像を分類モデルに入力することで、表札を含む画像から位置を推定するモデルを作成した。検出モデルと分類モデルの学習条件を変更しながら、200 個のテストデータを入力し、検出の過不足数と推定の正解率を取得した。それぞれのモデルの条件を表 1 に示す。

表 1: 各推定モデルの学習条件

推定モデル番号	検出モデル学習データ数 (訓練, 検証)	分類モデル学習データ数 (訓練, 検証)	データオーグメンテーション	分類モデル学習反復回数
1	32, 15	682, 175	水平反転	100
2	32, 15	682, 175	水平反転	300
3	32, 15	1277, 147	水平反転	100
4	123, 21	1277, 147	水平反転	100
5	123, 21	1277, 147	ぼかし	100

ここで推定モデル 4 の分類器についての学習回数に対する、損失 (loss) と複数データに対する正解率 (acc) を図 1 に示す。図 1 より、分類の精度は高い値で収束したことがわかる。

また、テストデータに対する各推定モデルの出力結果を表 2 に示す。

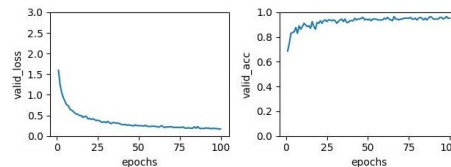


図 1: 分類器の学習回数に対する損失と正解率

表 2: 各推定モデルの結果

推定モデル番号	検出過不足数	正解率
1	+3	0.778
2	+3	0.778
3	+3	0.990
4	0	0.990
5	0	0.985

訓練データ数について、検出モデルは 91 枚増やしたときに検出過不足数は 3 減少して 0 となり、分類モデルは 595 枚増やしたときに正解率は約 1.3 倍に上昇した。検出モデルは本研究のテストデータに対してはこれ以上訓練データを増やす必要はないが、正解率が 1.0 とはならなかった分類モデルは、訓練データを増やしつつ、正解率を測定していく必要がある。

3. おわりに

本研究において、推定モデルは十分な精度を持ったといえる。

学内ナビゲーションシステムを作成するにあたり、学内位置推定モデルを組み込んだ目的地までの経路導出システムおよびユーザインターフェースの作成、これらをサーバ上として構築し、リアルタイムにアクセスできるようにする必要がある。

参考文献

- [1] 北村茂生, 三溝俊介, 松下光範, オンサイト検索: 携帯端末を用いたリアルタイムでの看板画像からの店舗情報アクセス手法, 電子情報通信学会論文誌, 2020.
- [2] Glenn Jocher, Ayush Chaurasia, Jing Qiu, Ultralytics YOLOv8, <https://docs.ultralytics.com/ja/models/yolov8>, 2024.
- [3] PyTorch, <https://pytorch.org>, 2017.