

P10 地域特性推定のための地物カテゴリを利用した自己教師あり学習

福地湧¹ 牛尼剛聡² 田中克己³ 角谷和俊¹ (関西学院大学¹ 九州大学² 福知山公立大学³)

概要

- 目的: 地物カテゴリを基に地域特性の推定
- 手法: 地物カテゴリを地図上に可視化した地図画像(256px*256px)から地域特性を推定するCNNモデルを構築

(1) データセットの作成

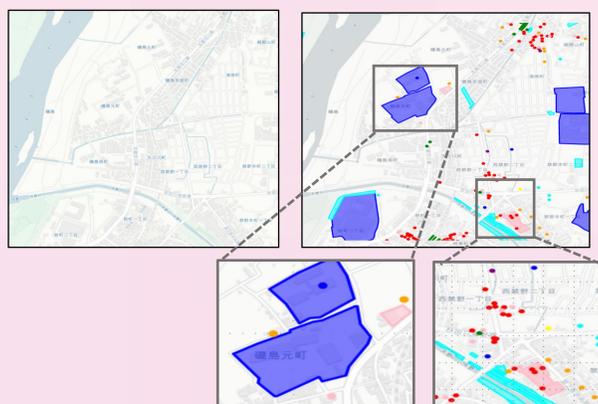
- ・ 学習データ: 地物のカテゴリをプロットした地図画像 (OSMの地理情報を基に地物を地図上に可視化)
- ・ ground truth: 口コミから抽出した地域特性 (単語の出現率を地域特性と仮定, レビューはLIFULL HOME'Sのまちむすびデータセットを利用)

(2) 地域特性を予測するための学習

地物カテゴリの地理的な意味に基づいて予測するために自己教師あり学習をチューニング

(1) データセットの作成

学習データ



ground truth

車に乗れないと買い物が不便。

子育てしやすい環境。

子育てのしやすさラベル 0.5

- ・ 全投稿数: 2件
- ・ 子育てを含むセンテンス数: 1件

(2) 地域特性を予測するための学習



自己教師あり学習①

■ 地物カテゴリ推定タスク

- 目的: カテゴリ特有の量や位置関係の学習
- データセット: 各カテゴリ200枚の1,000枚

学習データ

学校 駅 飲食店

公園 小売店

描画地物のカテゴリを推定?

自己教師あり学習②

■ 駅の有無による地域の分類タスク

- 目的: 周辺地物の学習
- データセット: 1,500枚

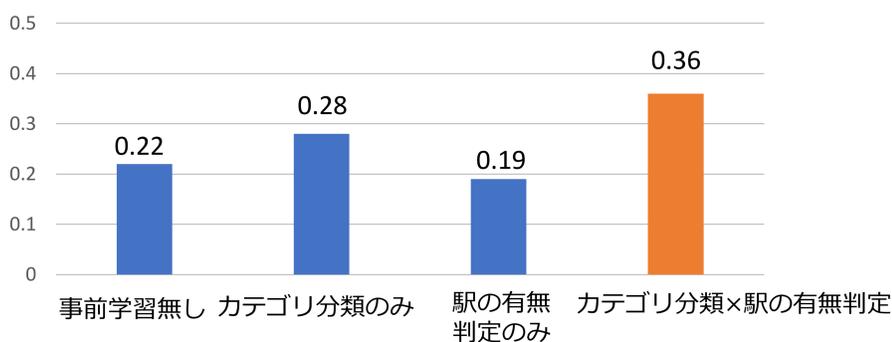
駅の有無を推定

駅がある地域 駅が無い地域

評価実験

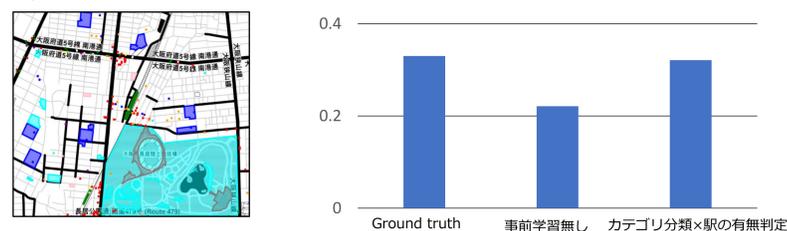
■ 自己教師あり学習による予測精度の向上を検証

- train data: 720 validation data: 180 test data: 100
- evaluation metrics: ground truthと予測の相関係数
- 結果: 自己教師あり学習により相関が強くなる



■ 考察

(自己教師あり学習により予測が向上した地域)
ground truthの値が大きい地域 (子育てしやすい地域)



(自己教師あり学習でも予測が向上しなかった地域)
地理的な観点を含めないレビューが含まれる地域

幼稚園や保育園に入りにくいので生活しにくい。

まとめと今後の課題

■ 地図とレビューを用いた地域特性推定

- 地物のカテゴリを可視化した地図画像データ
- 自己教師あり学習によりカテゴリの意味を学習

■ 今後の課題

- 道路環境を利用した自己教師あり学習
- 地物の多様な意味を読み取るための自己教師あり学習