

# 地域に対するレビューと地図画像を用いた地域特性学習モデルの提案

福地湧<sup>1</sup>, 芝原拓人<sup>1</sup>, 牛尼剛聡<sup>2</sup>, 田中克己<sup>1</sup>, 角谷和俊<sup>1</sup> (関西学院大学<sup>1</sup> 九州大学<sup>2</sup>)

## 研究背景と目的

背景：街の雰囲気予測することの困難さと地理データの存在

- 馴染み無い地域の雰囲気の把握は困難 (例：子育てしやすい地域が分からず物件探しが難航)
- 地図や口コミなど地域の特徴を表すデータが存在

目的：地図画像と口コミから街の雰囲気を学習する“地域特性学習”モデルの提案

【入力】



日常生活に  
困らない地域

地域特性  
学習モデル

【出力】

1位	〇〇県△△市□□
2位	〇〇府△△市□□
3位	〇〇府△△町□□

## 地図画像

■ 地理情報を色で表現

- 地理オブジェクトのカテゴリを色で区別
- “Colored Category Map画像(CCM画像)”



## 地域に対するレビューと地図画像

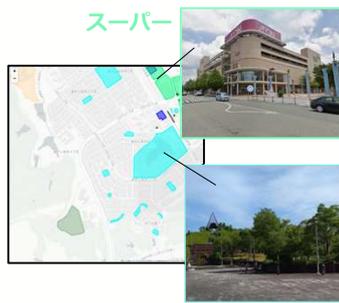
■ 兵庫県三田市けやき台



田舎やけどショッピング  
モールが充実している。



自然環境も良く治安も  
良い。閑静な住宅街。



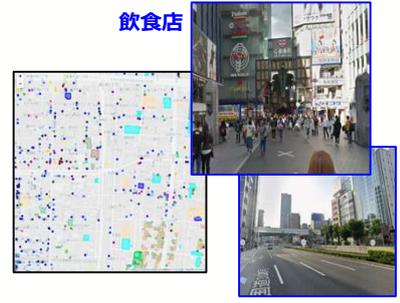
■ 大阪府大阪市中央区



飲食店の数も多いので、  
住むには便利。



子育て環境は良くない。



## 提案手法：レビューと地図画像を用いた地域特性学習

地域レビューとCCM画像のペアを学習する“地域特性学習モデル”

- 学習フェーズ：“地域に対するレビューの要約文”と“CCM画像”の正しいペアの類似度が高くなるようにTextEncoderとImageEncoderを学習

地域に対するレビューの要約文

T1：教育機関も充実し、閑静な住宅街  
T2：非常に便利で都会的だが治安が悪い地域

Text Encoder

T<sub>1</sub> T<sub>2</sub> ... T<sub>N</sub>

CCM画像 (色付きカテゴリ地図画像)

地図  
データ



Image Encoder

I <sub>1</sub>	I <sub>1</sub> ·T <sub>1</sub>	I <sub>1</sub> ·T <sub>2</sub>	I <sub>1</sub> ·T <sub>3</sub>	...	I <sub>1</sub> ·T <sub>N</sub>
I <sub>2</sub>	I <sub>2</sub> ·T <sub>1</sub>	I <sub>2</sub> ·T <sub>2</sub>	I <sub>2</sub> ·T <sub>3</sub>	...	I <sub>2</sub> ·T <sub>N</sub>
⋮	⋮	⋮	⋮	⋮	⋮
I <sub>N</sub>	I <sub>N</sub> ·T <sub>1</sub>	I <sub>N</sub> ·T <sub>2</sub>	I <sub>N</sub> ·T <sub>3</sub>	...	I <sub>N</sub> ·T <sub>N</sub>

- 推論フェーズ：ユーザが入力したテキストとCCM画像をの類似度を予測

都会と自然の融合エリア

Text Encoder

T



Image Encoder

I<sub>1</sub>·T  
I<sub>2</sub>·T  
⋮  
I<sub>N</sub>·T

## 評価実験

実験の内容と目的

- 内容) “テキスト”と“CCM画像”のペアの類似度を予測
- 目的) “地域特性予測モデル”により、CCM画像の特徴を言語化できるか検証

結果

- 正しいペアの類似度が相対的に高い傾向にはあるが、高い値を出力することはできていない
- 都会と田舎など大きな区別は可能

	(博多) 駅前で多くの人が	(中州) 歓楽街で	(春日) 住みやすい	(丹波篠山) 暮らしやすい	(佐賀県川副町) 買い物に不便な
(博多)	0.18	0.19	0.22	0.04	0.10
(中州)	0.25	0.24	0.27	0.13	0.13
(春日市)	0.15	0.18	0.19	0.08	0.11
(丹波篠山)	0.14	0.15	0.13	0.11	0.21
(佐賀県川副町)	0.14	0.15	0.16	0.13	0.25

テキストと地図画像の類似度

## まとめと今後の課題

(まとめ)

- 地域に対するレビューと地図画像から地理的な特徴を把握可能
- レビューと地図画像のコサイン類似度を学習しペアを予測するモデルを構築

(今後の課題)

- 最適な学習率やバッチサイズ, epoch数の検討
- 学習に最適なレビューや地図画像の検討

## 参考文献

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual mod-els from natural language supervision. In International conference on machine learning, pp. 8748-8763. PMLR, 2021

## 謝辞

本研究では、株式会社LIFULL様からご提供いただいた「LIFULL HOME'S まちむすびデータセット」を利用しました。また、本研究はJSPS科研費 JP23K25159, JP23K21845の助成を受けたものです。