Detecting Neighborhood Gentrification at Scale via Street Views and POIs (Student Abstract)

Tianyuan Huang
Stanford University
tianyuah@stanford.edu

Abstract

Neighborhood gentrification plays a significant role in shaping the social and economic status of both individuals and communities. While some efforts have been made to detect gentrification in cities, existing approaches mainly relies on estimated measures from survey data and requires substantial work of human labeling yet fails to characterize the physical appearance of neighborhoods. To this end, we introduce a novel approach to incorporate data like street view images and POI features to represent urban neighborhoods comprehensively at each timestamp. We show the effectiveness of the proposed methods with previous research on gentrification measures: each neighborhood representation we trained not only indicates its gentrification status, but also could become supplementary parts for the current measures and valid resource for researchers and policy makers.

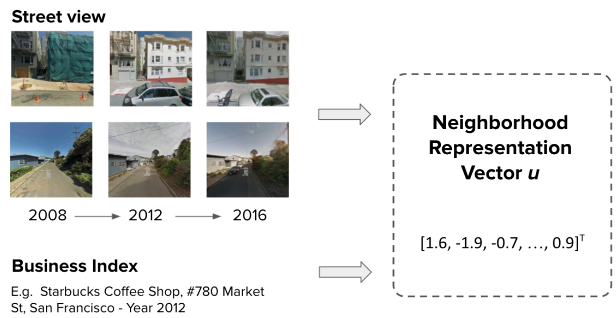

Figure 1: Vector representation of urban neighborhoods. Each neighborhood is a *container* of spatial-temporal multi-modal inputs, e.g., time-stamped street views and POIs.

Introduction

Gentrification is the process of changing the character of urban neighborhood through reinvestment, renewal, and the influx of middle- and upper-middle-class residents in previously disinvested and declined neighborhoods. Detecting gentrifying neighborhoods is crucial for investigating the dynamics of urban change, which consequently contribute to economic development as well as inequality in cities. Traditional approaches to detect gentrification rely on collecting demographic information, like the Decennial Census conducted by U.S. Census Bureau. However, the data produced by such survey-based methods are often restrained by their spatial and temporal granularity. Recent efforts have been made to identify urban change with time-stamped like street view images. However, existing methods rely on hand-labeled data and are hard to scale up to multiple locations. In this project, we will explore widely-available urban-associated time-stamped data like Google historical street views and business index, and use data-driven approaches, including both supervised and self-supervised learning to detect gentrification at a large scale. We conclude with an assessment of using this approach for measuring physical neighborhoods across ten of the largest metropolitan areas in the US.

Related Work

Spatio-temporal Representation Learning Spatio-temporal representation learning aims to produce region embedding using geo or temporal-tagged data. Street view images and POI textual data (Wang, Li, and Rajagopal 2020) were used in previous works to generate the neighborhood embedding. However, most of them are not able to utilize time-stamped data and learn neighborhood representation in the temporal dimension. In this work, we will introduce a systematic approach for spatial-temporal representation learning of urban neighborhoods.

Neighborhood Gentrification In order to identify neighborhood gentrification, (Hwang and Sampson 2014) relied on public data from the national survey and administrative records and examined its uneven evolution across time and space. Since physical conditions of neighborhood built environment such as property appearance and the level of disorder are important contexts for neighborhood trajectories, (Naik et al. 2017) and (Ilic, Sawada, and Zarzelli 2019) measured changes of neighborhood characteristics using historical street view images and human labeling. Fueled by multi-modal time-stamped data like street view images and POI features with a self-supervised approach, we hope our proposed framework will provide a more convenient and comprehensive tool to detect gentrification on a large scale.

Model and Implementation

Data Description

The time-stamped street view images are obtained from Google Static Street view API[1] between the year of 2007 and 2021. Historical Business data are licensed from Data Axle's ReferenceUSA[2]. We consider the largest 10 metropolitan areas to implement our model, and we start the preliminary experiment in the San Francisco Bay Area. Specifically, we sample around 3 million street view images in $1,198$ census tracts from the Bay Area. And we select commercial POIs such as restaurants and coffee shops from the business dataset.

Street View Image Segmentation

We start by implementing a pretrained semantic segmentation model for the street view images, each street view image is segmented into 65 classes including building, sky, road, vegetation, etc. We then aggregate street view images into the neighborhood level and calculate the average number of pixels for each class, thus generating a 65 dimensional vector for each neighborhood. In addition, we calculate each class's change in the number of pixels between the year of 2007 and 2021 for each location we have sampled, and we then formulate a change vector for each neighborhood by aggregating the change vectors of all locations within the neighborhood. Finally, we concatenate the segmentation vector together with the change vector for each neighborhood.

Self-supervised Neighborhood Representation

Each neighborhood can be modeled as a "container" with a batch of street view images and POIs. We assume street view images with smaller geographical distances are more likely to share semantic similarities compared with those with larger distances according the First Law of Geography[3]. We adopt a contrastive learning scheme where the positive samples are the closest k images of each anchor image while the negative samples are those further away images. Then we generate the neighborhood representation by aggregating and averaging the street view images embeddings within each neighborhood. To incorporate POI textual data into the neighborhood representation, we use word embedding to encode POI's category and name, then we use another contrastive learning approach to minimize the neighborhood embeddings with the POI's embeddings which are within its context, while maximizing with those out of it. After training on both modalities of data, we concatenate the self-supervised neighborhood embedding with the representation we get from the segregation vector to formulate our final neighborhood representation.

[1] https://developers.google.com/maps/documentation/streetview

[2] Available at http://www.referenceusa.com/

[3] "Everything is related to everything else, but near things are more related than distant things."

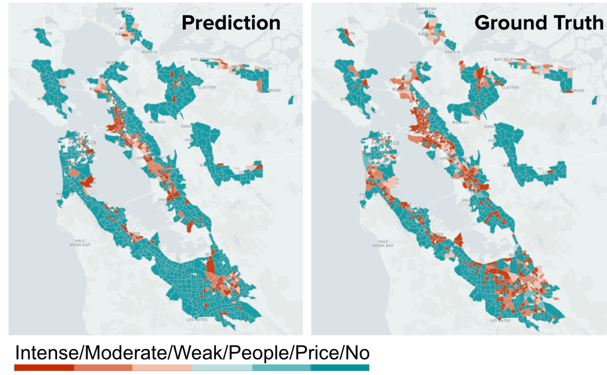

Figure 2: Preliminary prediction results on neighborhood gentrification status

Experiment and Future Work

To evaluate, we treat neighborhood embeddings as input features to predict the gentrification status for each census tract. Here we refer the gentrification status of neighborhood as the stage score similar to (Hwang and Sampson 2014), which includes 6 levels of gentrification: intense/-moderate/weak/people/price/no gentrification. In this work, we try random forest classifier and we use accuracy as the major metrics and 5-fold cross-validation to verify the prediction results. As is shown in Figure 2, our embeddings achieve 60% accuracy in predicting the gentrification status, which indicates the ability to represent the neighborhood and detect its gentrification status. Overall, our framework shows promising results in the preliminary experiment, and we plan to explore an end-to-end approach to integrate existing components and scale the experiment up to other major metropolitan areas of the US in future work.

Acknowledgments

This project was supported by the Google Cloud Grant from the Stanford Institute for Human-Centered Artificial Intelligence. The author would like to thank Prof. Ram Rajagopal and Prof. Jackelyn Hwang for their extensive guidance.

References

Hwang, J.; and Sampson, R. J. 2014. Divergent Pathways of Gentrification: Racial Inequality and the Social Order of Renewal in Chicago Neighborhoods. *American Sociological Review*, 79(4): 726–751.

Ilic, L.; Sawada, M.; and Zarzelli, A. 2019. Deep mapping gentrification in a large Canadian city using deep learning and Google Street View. *PLOS ONE*, 14: e0212814.

Naik, N.; Kominers, S. D.; Raskar, R.; Glaeser, E. L.; and Hidalgo, C. A. 2017. Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, 114(29): 7571–7576.

Wang, Z.; Li, H.; and Rajagopal, R. 2020. Urban2Vec: Incorporating Street View Imagery and POIs for Multi-Modal Urban Neighborhood Embedding. arXiv:2001.11101.