Streetplay: Assessing Boston Neighborhood Playability with Street View Images

Kruse and Liu, Geog 573

05/04/2020

Introduction

Our project used deep learning to automatically classify or label the playability of each neighborhood using Street View images that we have collected in different neighborhoods in the Boston metropolitan area.

Assessing playability in this way would be a scalable and an easily updatable method for giving policy makers and citizens a metric for evaluating how their neighborhoods can be improved

Street View image classified as Very Playable by Amazon Mechanical Turk

Data

Images:

- 46,000 Street View images for Boston
- Street network for Boston from OpenStreetMap
- 3,000 images labelled nine times each by Amazon Mechanical Turk

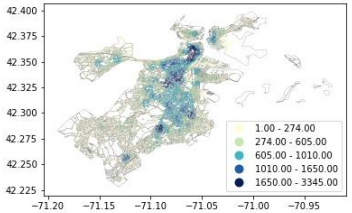
Features that potentially have an impact on the value of playability:

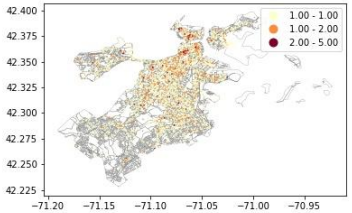
1. Safety -- Traffic and crime
   a. Street speed limit https://data.boston.gov/dataset/boston-street-segments
   b. Street Light https://data.boston.gov/dataset/streetlight-locations
   c. Crime incident reports https://data.boston.gov/group/public-safety
2. Health -- Sustainability and greenness
   a. Walkability Index, 2019 https://catalog.data.gov/dataset/walkability-index
   b. Tree, 2019 https://data.boston.gov/dataset/trees

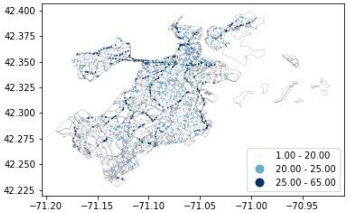
3. Diversity -- Public space types
   a. Open Space, 2018 https://data.boston.gov/dataset/open-space

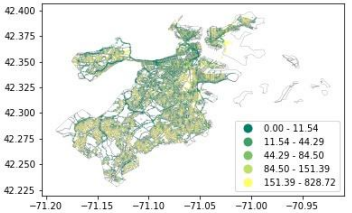
Data -analyzed and visualized using geopandas in python

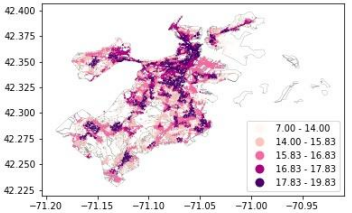
Crime	Street Light	Street speed limit	Open Space distance	Walkability index
count	count	mean	-	mean
400 m Buffer zone	10m buffer zone	-	Shortest path	-



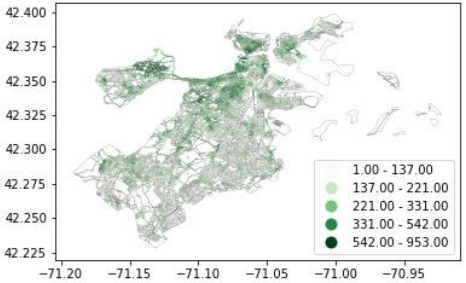


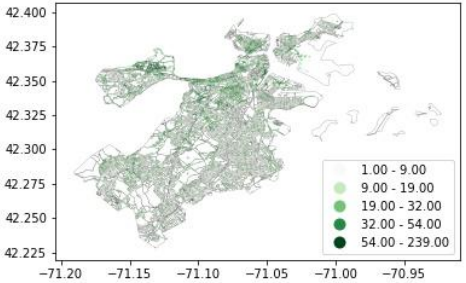


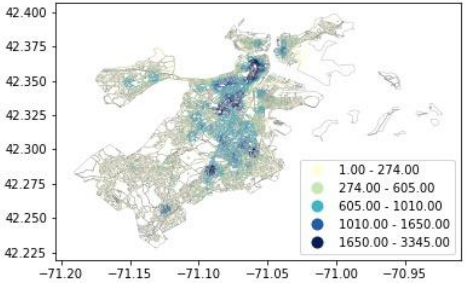




Tree
count
10/100/200 m Buffer zone







Methods

- Pre-trained deep convolutional neural network (Resnet 18), through Jupyter Notebook
- Mgwr.gwr package in Python
  - Global OLS and multiscale GWR
- Covariates:
  - 'crim_count_400m','light_count_10m','tree_count_200m','dis_OpenSpace_m','WalkIndex','SPEEDLIMIT'
  - Aggregated into 558 census blocks
- 3,000 labelled image (mode of the nine labels used)

Results and Findings

OLS

- Using 558 census blocks with seven covariates as the OLS model input, there are 5 explanatory variables that show significant, except for the tree_counts_400m
- R^2 of 0.407

MGWR

- Strong fit, R^2 of 0.752,
  - R^2 of 0.416 with a global model
- Bandwidth is calculated for each covariate, which here ranged from 143 to 557, with four variables having a bandwidth of 557
- dis_OpenSpace_m and SPEEDLIMIT variables had bandwidths of 143 and 157
  - Relatively local level compared to other variables
- Speed limit twice as impactful as the other variables: coefficient of -2.869
  - higher speed limits associated with lower playability
- Crime, lights and walkability index have also negatively correlated with playability,
- Distance to open space is positively correlated to it.

Conclusions and Discussion

- Trying several aggregation levels necessary to see how spatial scale influences output
- Identify which cities can use the same training data as Boston
- Compare model performance to random forest classification

Acknowledgements and References

Thanks to Yuhao Kang and Professor Song Gao for making this project possible and for guiding us through this project. This work was made possible by a generous grant from Microsoft's AI for Earth, and by a grant for image labelling from the Geography department at UW Madison.

References:

Oshan, T., Li, Z., Kang, W., Wolf, L. and Fotheringham, A. (2019). mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information*, [online] 8(6), p.269. Available at: https://www.mdpi.com/2220-9964/8/6/269/htm [Accessed 4 May 2020].