Shu Xingxing

LARP 750 Final Project Proposal

Professor Dana Tomlin

How Green are Neighborhoods in Philly?

1.Background

Urban thermal comfort would influence human outdoor activities and the utilization of urban space. With the rise in global warming, many cities are expected to experience more severe heat waves. Carbon pollution is already changing weather patterns in Philadelphia, in 2014, more than 53,138 sensitive Philadelphians lived in hotter than average neighborhoods. But we can make sure that these shifts are as limited as possible by reducing our contributions to climate change. The city's canopy can reduce the temperature of the city by blocking short-wave radiation and increasing the evaporation of water, while also creating a more comfortable microclimate and reducing air pollution caused by daily urban activities. Their absorptive root systems also help to avoid heavy rains and floods during storms. So, in general, trees are important for both cities and communities.

Accurate assessment of the tree canopy in our communities is the very first step to improve living conditions of residents. Calculating NDVI by remote sensing to get a picture of tree canopy cover has been widely used to evaluate the amount of street trees, but this method has several limitations. First, traditional remote-sensing techniques for NDVI are based on moderate-resolution imagery (e.g., 30 m of LANDSAT 8 open source data) which has limited utility at the scale of cities for it cannot fully reflect the shading condition of street trees like details of green density and vertical information. More recent works are using high-resolution satellite images like LiDAR to do a better job, but however the request of data is of high costs and need very specialized software which is not kind to most researchers.

We used Google Street View images to represent a perspective view of a city's streets and is able to cover the entire city communities, so we developed and tested a new method to quickly quantify and characterize urban vegetation, especially trees. Our goal is to prove that we can perceive the existence of urban street trees and quantify the coverage of trees from the perspective of pedestrians by: i) sampling a series of continuous streetscape image scenes; ii) quantifying the amount of tree coverage present therein; and iii) Model the relationship between the tree coverage of these sample points and the characteristics of community they are in. To evaluate the accuracy of this method, we compared our results with traditional remote sensing techniques (i.e., 30-m LANDSAT 8 data) used to estimate urban canopy coverage.

2.Abstract

In view of the city's complex spatial structure, traditional remote sensing methods cannot allow ordinary researchers to map and study the distribution of urban green

space due to high costs or poor spatial resolution. Moreover, these tools cannot see the urban landscape from the perspective of the residents. We tested a new type of computer vision application to quantify urban tree coverage on the street level in Philadelphia. We do this by leveraging the existing rich open source image data of city street views (Google Street View). We show a multi-step computer vision algorithm that can subdivide and quantify the tree coverage in streetscape images to achieve high accuracy. Then, by modeling the characteristic relationships of the communities where the sampling points are located, we can more accurately assess the perceptible tree cover in the urban street landscape. Although this method cannot replace high-resolution remote sensing technology (for example, LiDAR), this method provides a new urban tree cover metric that can quantify the existence and distribution of trees from the same perspective as citizens experience and observe trees, giving cities Planners, policy makers, and community residents understand the city in a new light.

3.Dataset

Our study area is the City of Philadelphia, PA and the datasets used in this study include building boundary data, land use data, LANDSAT 8 data, tree canopy data, and the Google street maps of Philadelphia. The LANDSAT 8 data was downloaded from USGS (https://earthexplorer.usgs.gov/). The building footprints map (https://www.opendataphilly.org/dataset/buildings) and land use map (https://www.opendataphilly.org/dataset/land-use) in the study area were collected from OpenDataPhilly. The original land use map includes 31 land use types in the study area. In order to facilitate further analysis, we aggregated similar land use types in the original land use map into several major land use types: commercial land, industrial land, recreational land, and residential land. The tree canopy data was also from OpenDataPhilly (https://www.opendataphilly.org/dataset/pa-high-resolution-tree-canopy) which is at 1m resolution and was developed from publicly available LiDAR data. The Google street data was collected by Google API and to get a more accurate result, we choose street views between June and September for more leaves are on the trees.

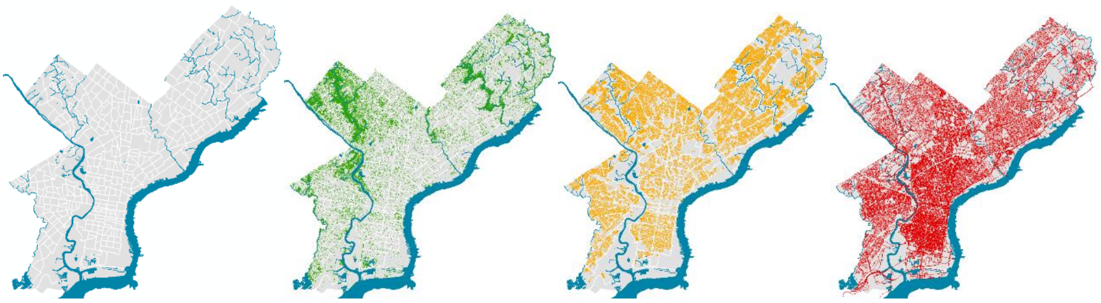

Fig.1. Neighborhoods & Tree Canopy Map & Residential Land Map & Road Map of Philadelphia

4.Methods

4.1 Google Street View Image Analysis.

4.1.1 Choose appropriate neighborhoods for research.

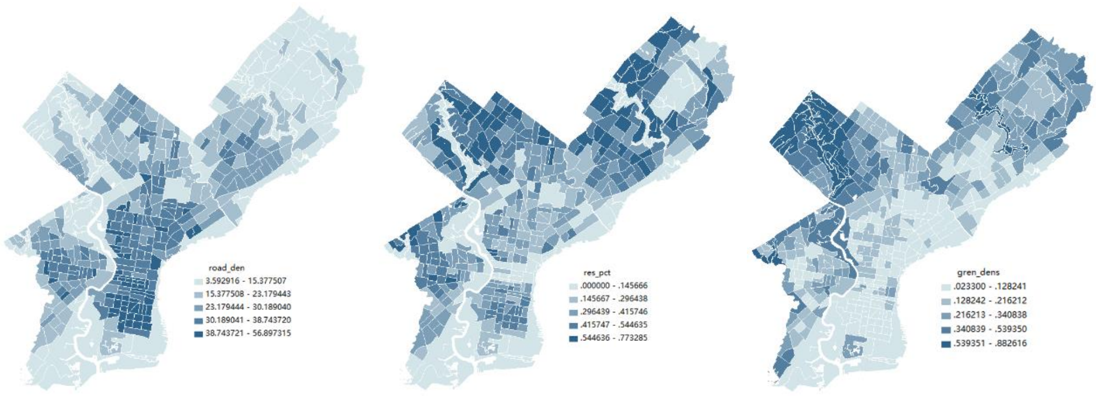

Fig.2. Road Density & Residential Building Density & Green Density Map of Philly

I choose two residential neighborhoods with distinct Greenery Density calculated based on canopy data (Fig.4) and details shows in Table.1, which intends to answer the two questions: 1) How do sentiments and emotions differ in green dominated community compared to community with very few greenery? 2) How does the proximity of green spaces affect these sentiments and emotions?

Table.1:

	Canopy Density (%)	PCT of Residential Land (%)	Road Density (km/km²)
CT 81.02	6.1	58.7	36.85
CT 231	72.7	68.9	22.7

4.1.2 Create sample points of streets of selected neighborhoods.

With the street network and boundary shapefile of Philadelphia as input, a shapefile containing points every 50m was created to be fed into the Google API to retrieve Google Street View Images (Fig.4).

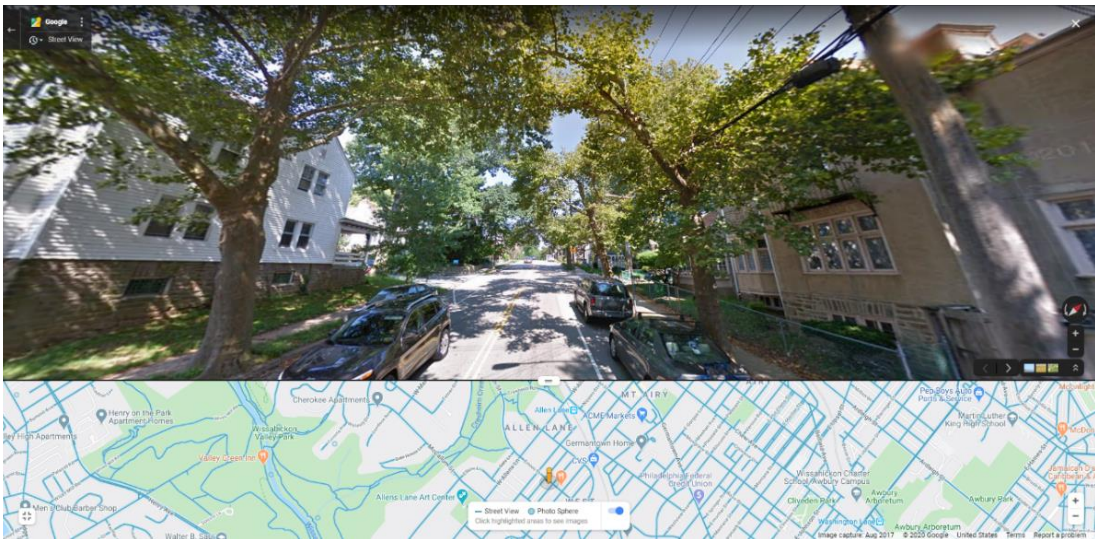

Fig.3. GSV of a sample point in the study area.

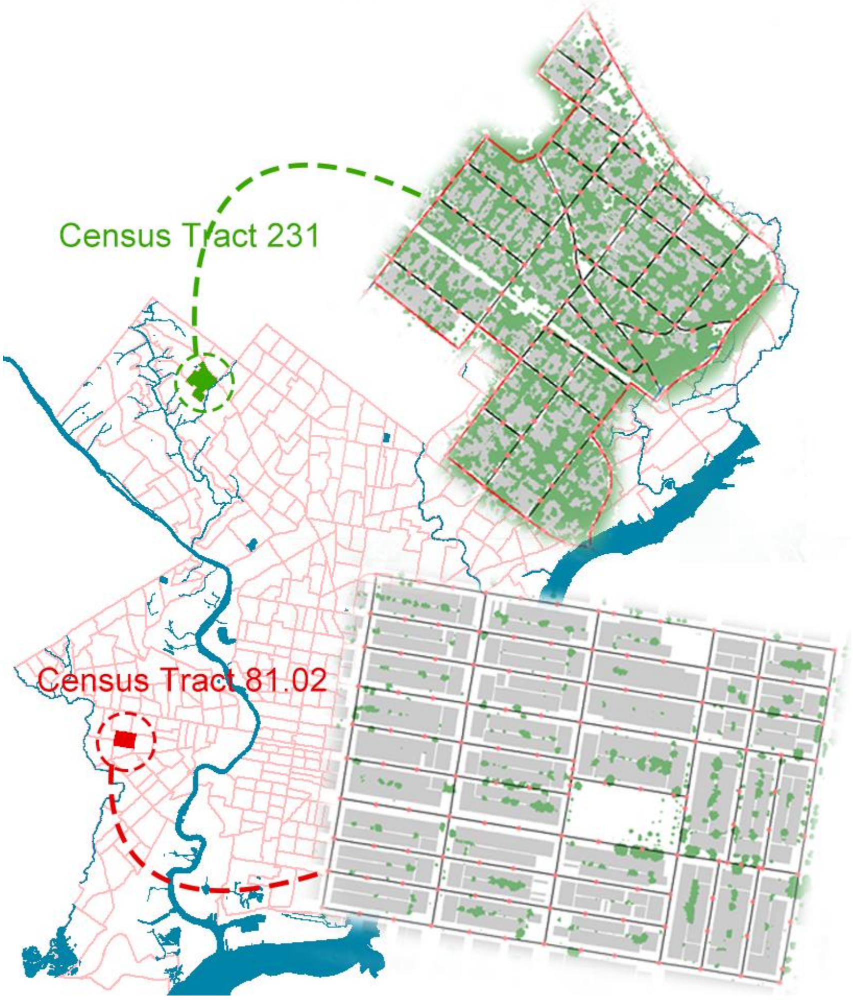

Fig.4. The location, road map, and random points of the study area.

4.1.3 Create Metadata containing GSV panoID.

With the shapefile as input, metadata containing the panoID, panoDate, latitude, longitude and tilt specifications for the image will be stored in text files to be later used to calculate the Green View Index.

```
samplesFeatureClass = '../sample-spatialdata/Philly50m.shp'
num = 100
ouputTextFolder = '../metadata'

metalib.GSVpanoMetadataCollector_fiona(samplesFeatureClass,100,ouputTextFolder)
```

```
The coordinate (-75.180822,39.934928), panoId is: cMPXRurXtp-eRSwaSFluOw, panoDate is: 2018-09
The coordinate (-75.180630,39.935826), panoId is: r3hfVsfMxCb7kd25flUcXA, panoDate is: 2018-09
The coordinate (-75.181214,39.935901), panoId is: 9tgFyBlAMucneJgAGn5DYA, panoDate is: 2018-09
The coordinate (-75.181791,39.935975), panoId is: 7NgFkVM4AmmEKmwkWdOQqw, panoDate is: 2018-09
The coordinate (-75.182284,39.936039), panoId is: NRO-H-AKdYnadaU5pZ1yuA, panoDate is: 2018-09
The coordinate (-75.180847,39.935036), panoId is: _dYhBOKiIF3hvunw6wwOmA, panoDate is: 2018-09
The coordinate (-75.181381,39.935109), panoId is: iedP08c08BOn_ztxGSOJQA, panoDate is: 2018-09
The coordinate (-75.181943,39.935185), panoId is: bLMsIItTyKqG7HZM3fuDEg, panoDate is: 2018-09
The coordinate (-75.182418,39.935250), panoId is: v9DoxdFbxk_buH3IcFuvMQ, panoDate is: 2018-09
The coordinate (-75.180754,39.935233), panoId is: gO3juisC4asH1eDCEOcr1w, panoDate is: 2018-09
The coordinate (-75.180675,39.935598), panoId is: ujE_GOjSGKVzOeM4p12gzQ, panoDate is: 2018-09
The coordinate (-75.180774,39.935410), panoId is: rvGG6-5EZdiaMORcYqx1cg, panoDate is: 2018-09
The coordinate (-75.181260,39.935471), panoId is: FfSSgUT6c-nO8dt7poiCQQ, panoDate is: 2018-09
The coordinate (-75.181838,39.935544), panoId is: qbFRUh_RSj3L_-3u8LLUkQ, panoDate is: 2018-09
The coordinate (-75.182325,39.935606), panoId is: d84-CZZX3lqacEPckU1cWg, panoDate is: 2018-09
The coordinate (-75.182385,39.935568), panoId is: hEIpI6zfWs_GPR5epMWQuA, panoDate is: 2018-09
The coordinate (-75.182315,39.935897), panoId is: ROjZOENxU5oAlVm-wRoy3A, panoDate is: 2018-09
The coordinate (-75.180894,39.934608), panoId is: ERvFB7dRhdaKWsG-OOnggw, panoDate is: 2018-09
The coordinate (-75.181394,39.934673), panoId is: lYTXOtIGmtinF8M1sbLsYg, panoDate is: 2018-09
The coordinate (-75.181957,39.934746), panoId is: gdT7YfHbAeHtgxAKBCEcFw, panoDate is: 2018-09
```

Fig.5. Metadata for GSV calculation.

4.1.4 Collect Google Street View Image.

Based on the metadata we created, use Google Street View Static API to request street images of sample points. GSV panoramas can be requested by inputting coordinates through Google's Application Programming Interface (API). Fig.5 shows a requested GSV static image. In this example, parameter size specifies the output size of the image in pixels, location provides the geo-location of the GSV; heading indicates the compass heading of the camera; fov (default is 90) determines the horizontal field of view of the image; pitch (default is 0) specifies the up or down angle of the camera relative to the Street View vehicle; radius (default is 50) sets a radius, specified in meters, in which to search for a panorama, centered on the given latitude and longitude. In previous studies, the horizontal field view was chosen between 50°and 60°. For our research, we set the fov to 60°; thus, six images can cover the 360° horizontal surroundings.

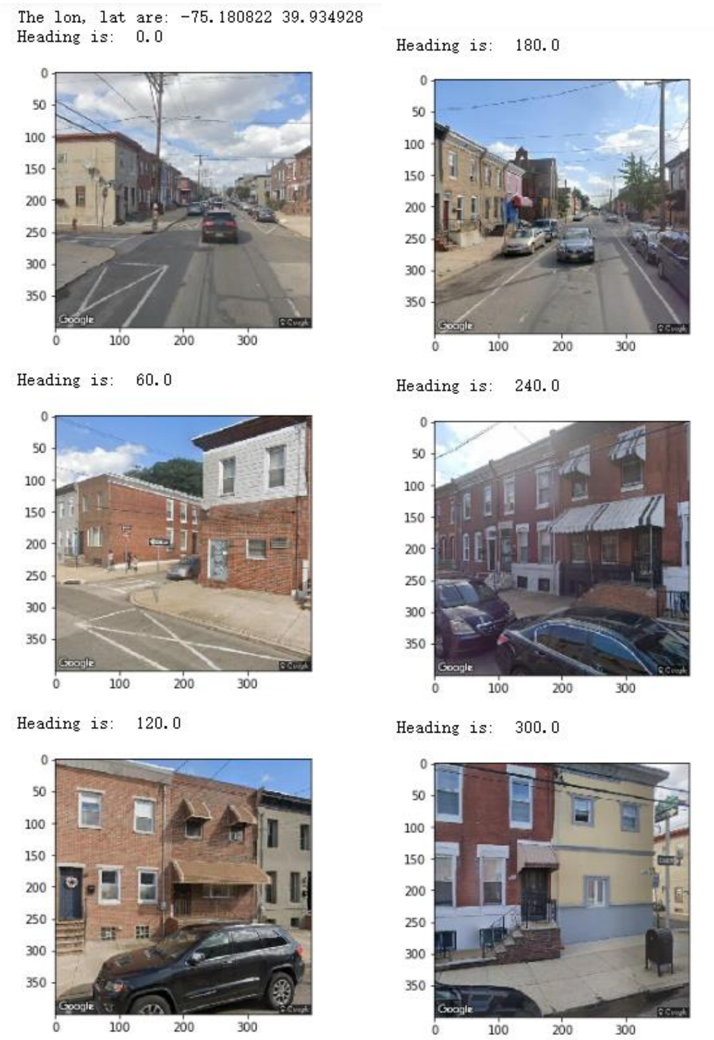

Fig.6. GSV image of one sample point on 6 directions.

4.1.5 Green area extraction from GSV images.

Traditionally we need near infrared band and red band to detect vegetation because vegetation shows high reflectance at near infrared band but shows high absorption at red band. However, GSV images do not have both near infrared band and red band but only red, green and blue bands. So we developed and used a simple automatically unsupervised classification method to extract green vegetation from GSV images.

First, two images Diff 1 and Diff 2 were generated by subtracting respectively red band and blue band from green band. Then the two difference images were multiplied to generate one Diff image. Considering green vegetation normally shows higher reflectance values in the green band than in the other two visible bands, the green vegetation pixels generally have positive values in the Diff image. Those pixels with smaller values in green band than in blue or red band show negative values in the Diff image. If pixel values in green band are smaller than those in both red and blue bands, the corresponding values in the Diff image are still positive. So an additional rule that pixel values in green band must be larger than those in red band was added.

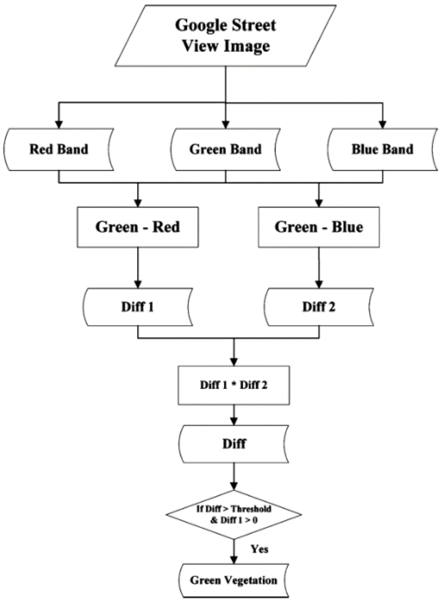

Fig.7. The workflow for green vegetation extraction from GSV images.

4.1.6 GVI Calculation.

GVI was defined as the ratio of the total green area from four pictures taken at a street intersection to the total area of the four pictures, calculated using the following equation:

$$\text{Green View} = \frac{\sum_{i=1}^{6}\sum_{j=1}^{3} Area_{g\_ij}}{\sum_{i=1}^{6}\sum_{j=1}^{3} Area_{t\_ij}} \times 100\%$$

where $Area_{g\text{-}ij}$ is the number of green pixels in one of these images captured in six directions with three vertical view angles (-45°, 0°,45°) for each sample site, and $Area_{t\text{-}ij}$ is the total pixel number in one of the 18 GSV images.

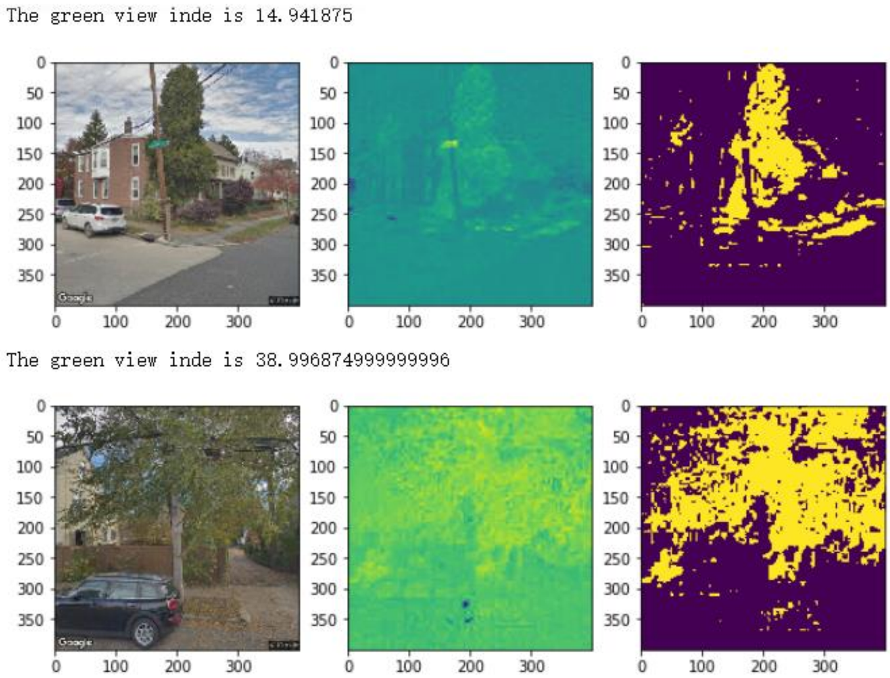

Fig.8. Results of green vegetation extraction.

4.2 Twitter Data Analysis.

4.2.1 Tweets Collection.

We restrict our work to using tweets that are explicitly geo-tagged as such tweets allow us to determine where they are posted from. These steps include the following: 1) Filtering tweets that are explicitly geo-tagged with latitude and longitude coordinates and within 1km of Philly. 2) Selecting tweets that are written in English, based on the "language" field provided by the Twitter API. 3) Tokenizing each tweet into individual words based on separation by spaces. 4) Converting all tweets and tokenized words into lower-case.

4.2.2 Sentiment Analysis.

Once downloaded, tweets were assigned into three categories: positive, negative or neutral. This annotation was based on the presence of emotive words, emoticons or meaning in the tweet text. Subsequently, the positive and negative annotated tweets were further categorized into distinct emotions. The high-level emotions chosen included Ekman's six basic emotions (anger, disgust, fear, sadness, happiness, surprise) plus beauty, in line with previous research which has undertaken sentiment analysis of Twitter data. Beauty was included as a subcategory to the positive tweets but outside of the emotions, to account for the large amount of tweets referencing the beauty of nature and the landscape. Each tweet could only be assigned into one of these emotion categories based on the strongest present emotion.

Each tweet was also categorized by user type: male, female or organization. Male and female categories reflect the gender of the Twitter user, while organizations include the private companies, local council initiatives and local businesses found to be captured in the sample.

5. Outcomes

5.1 Comparison of NDVI & GVI results.

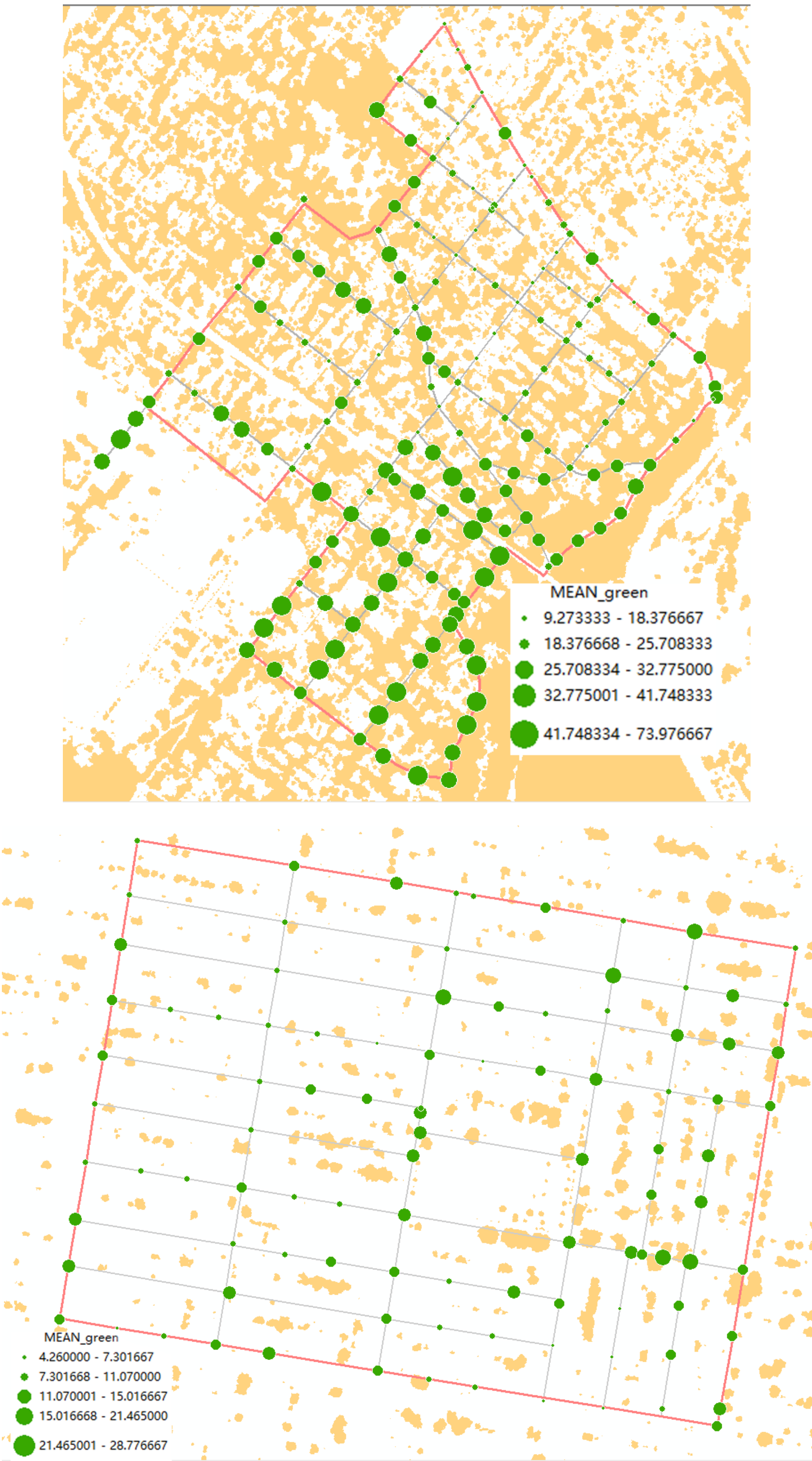

Fig.9. NDVI & GVI results of the two study census tracts.

5.2 Comparison of Tweet Sentiments in Green Spaces Versus Urban areas

5.3 Comparison of Tweet Emotions in Green Spaces Versus Urban areas

5.4 Proximity of Green Spaces and Urban Sentiments

6.Discussion

7.Conclusion

8.References

[1] Xiaojiang Lia, Chuanrong Zhang, Weidong Lia, Robert Ricard, Qingyan Mengc, Weixing Zhang, 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban Forestry & Urban Greening 14 (2015) 675–685

[2] Ian Seiferling, Nikhil Naik, Carlo Ratti, Raphäel Proulx, 2017. Green streets − Quantifying and mapping urban trees with street-level imagery and computer vision. Landscape and Urban Planning 165 (2017) 93-101

[3] Xiaojiang Li, Carlo Ratti 2018. Mapping the spatial distribution of shade provision of street trees in Boston using Google Street View panoramas. Urban Forestry & Urban Greening 31 (2018) 109-119

[4] Almeer, M.H., 2012. Vegetation extraction from Free Google Earth Images ofdeserts using a robust BPNN approach in HSV space. Int. J. Emerg. Technol. Adv.Eng. 2 (5), 1–8.

[5] Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., Wilkinson, P., 2013.Developing and testing a street audit tool using Google Street View to measureenvironmental supportiveness for physical activity. Int. J. Behav. Nutr. Phys.Activity 10 (1), 103.

[6] Yao, Y., Zhu, X., Xu, Y., Yang, H., Wu, X., Li, Y., Zhang, Y., 2012. Assessing the visualquality of green landscaping in rural residential areas: the case study ofChangzhou, China. Environ. Monit. Assess. 184, 951–967.

[7] Xiaojiang Lia, Chuanrong Zhang, Weidong Li, Yulia A. Kuzovkina 2016. Environmental inequities in terms of different types of urbangreenery in Hartford, Connecticut. Urban Forestry & Urban Greening 18 (2016) 163–172

[8] Li, X., Zhang, C., Li, W., Kuzovkina, Y.A., Weiner, D., 2015b. Who lives in greenerneighborhoods? The distribution of street greenery and its association

withresidents' socioeconomic conditions in Hartford, Connecticut, USA. Urban For.Urban Greening 14 (4), 751–759.

[9] Google Street View Image API (https://developers.google.com/maps/documentation/streetview/intro)

[10] Google Maps API (https://cloud.google.com/maps-platform/)

[11] Kwan Hui Lim, Kate E. Lee, Dave Kendal, Lida Rashidi, Elham Naghizade, Stephan Winter,and Maria Vasardani, The Grass is Greener on the Other Side: Understanding the Effects of Green Spaces on Twitter User Sentiments, Cognitive Computing, April 23-27, 2018

[12] Helen Victoria Roberts, Using Twitter data in urban green space research: A case study and critical evaluation, Applied Geography 81 (2017) 13-20

[13] Helen Roberts, Jon Sadler, Lee Chapman, The value of Twitter data for determining the emotional responses of people to urban green spaces: A case study and critical evaluation, Urban Studies 2019, Vol. 56(4) 818–835

[14] Richard A. Plunza, Yijia Zhoua, Maria Isabel Carrasco Vintimillab, Kathleen Mckeownc, Tao Yud, Laura Ugguccionia, Maria Paola Sutto, Twitter sentiment in New York City parks as measure of well-being, Landscape and Urban Planning 189 (2019) 235–246