Identifying Burglary Risks with Google Street View Images

Sam Bodmer, Junyi Li, Xingyu Gong, Shenghan Lyu, Tim Yupeng Sun

NYU MLC
Spring 2024

Introduction

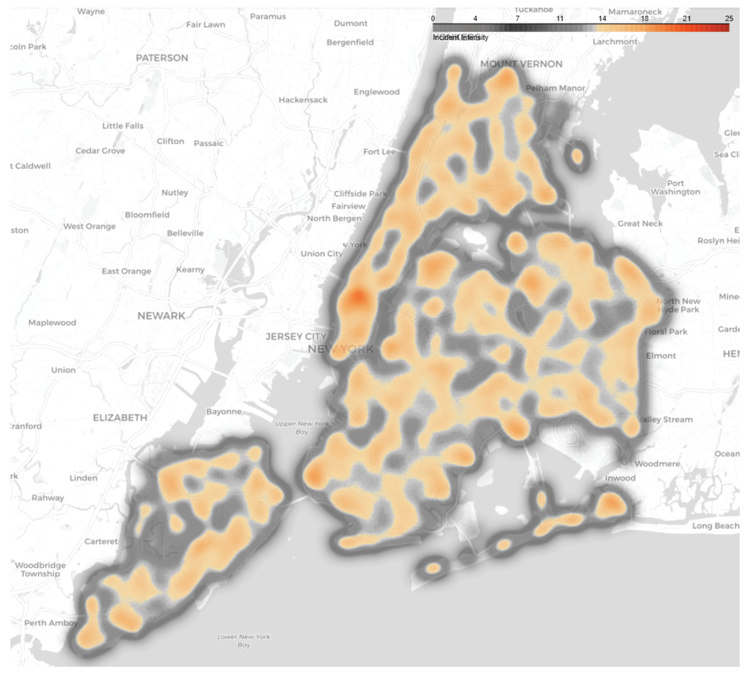
Safety is a fundamental component of quality of life, reflecting the security and well-being of individuals in their environments. Home invasions, one of the more intrusive forms of criminal activity, not only threaten personal safety but also compromise residential security.

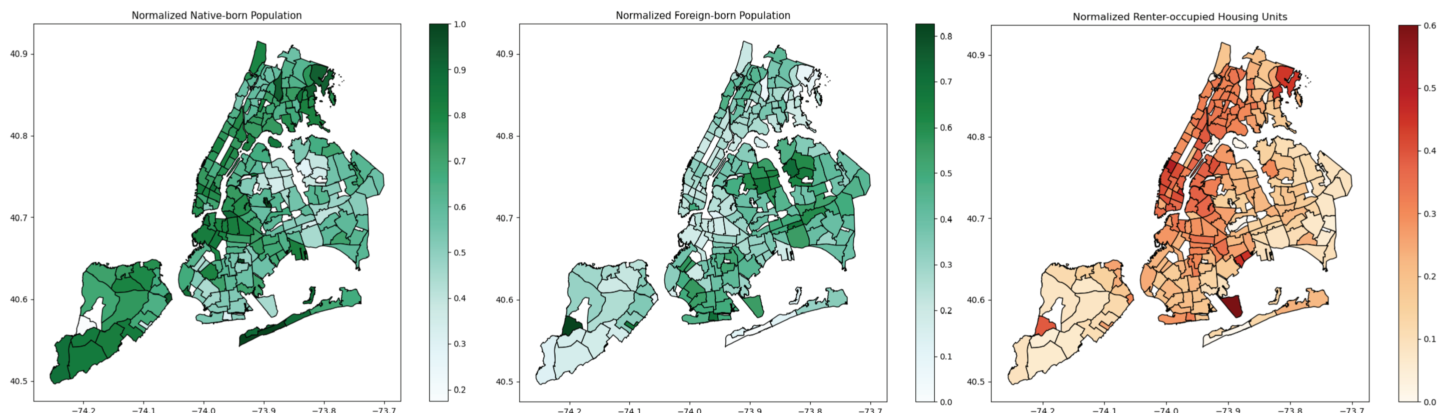
Our research focuses on the relationship between streetscape image features and burglary rates in New York City. By utilizing convolutional neural networks (CNN's) to analyze Google Street View images alongside demographic data, **this study aims to train a model that could predict the burglary risk by street view images** which would be helpful for police to assign their resources effectively, thus enhancing urban safety and quality of life.

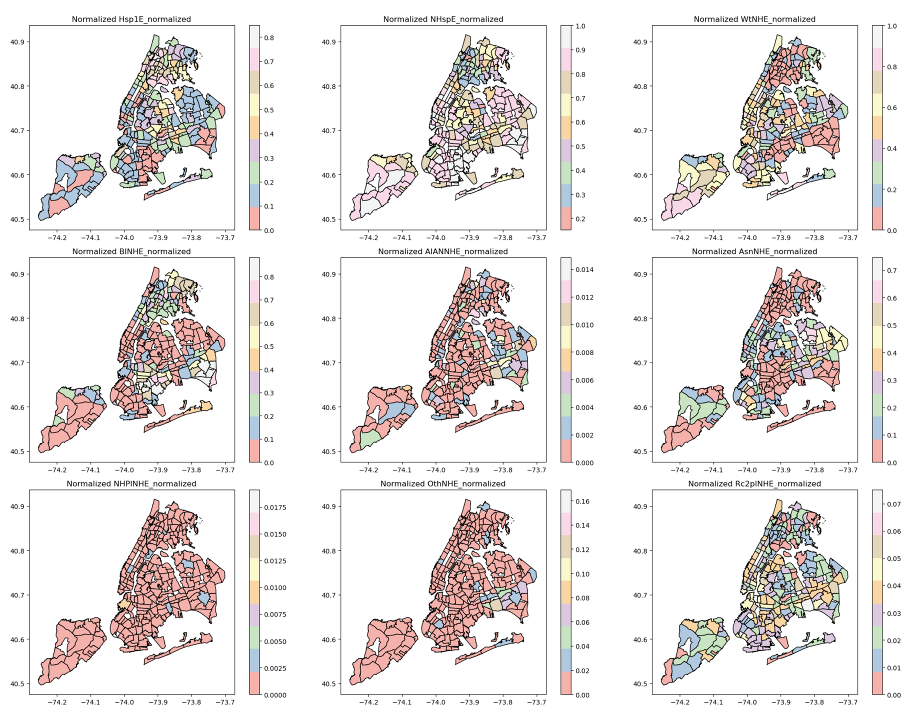
Data

NYPD Complaint Data Historic

This data is from NYC Open data. NYPD Complaint Data-Burglary primarily covers data on complaints about burglary received by nypd between December 2005 and December 2022 across nyc's 5 boroughs. The data includes the location of the incident, geographic information, and a brief description of the incident.







Demographic data

Datasets from the 2020 Census section of the official website of the New York City Department of City Planning.

Google Street View

Street view imagery can be obtained from the GSV Server by utilizing the Google Map Street View Static APIs.

BURGLARY 54914 | NO BURGLARY 49412

Manhattan	Brooklyn	Queens	Bronx	Staten Island
6372	16064	19589	8403	4487

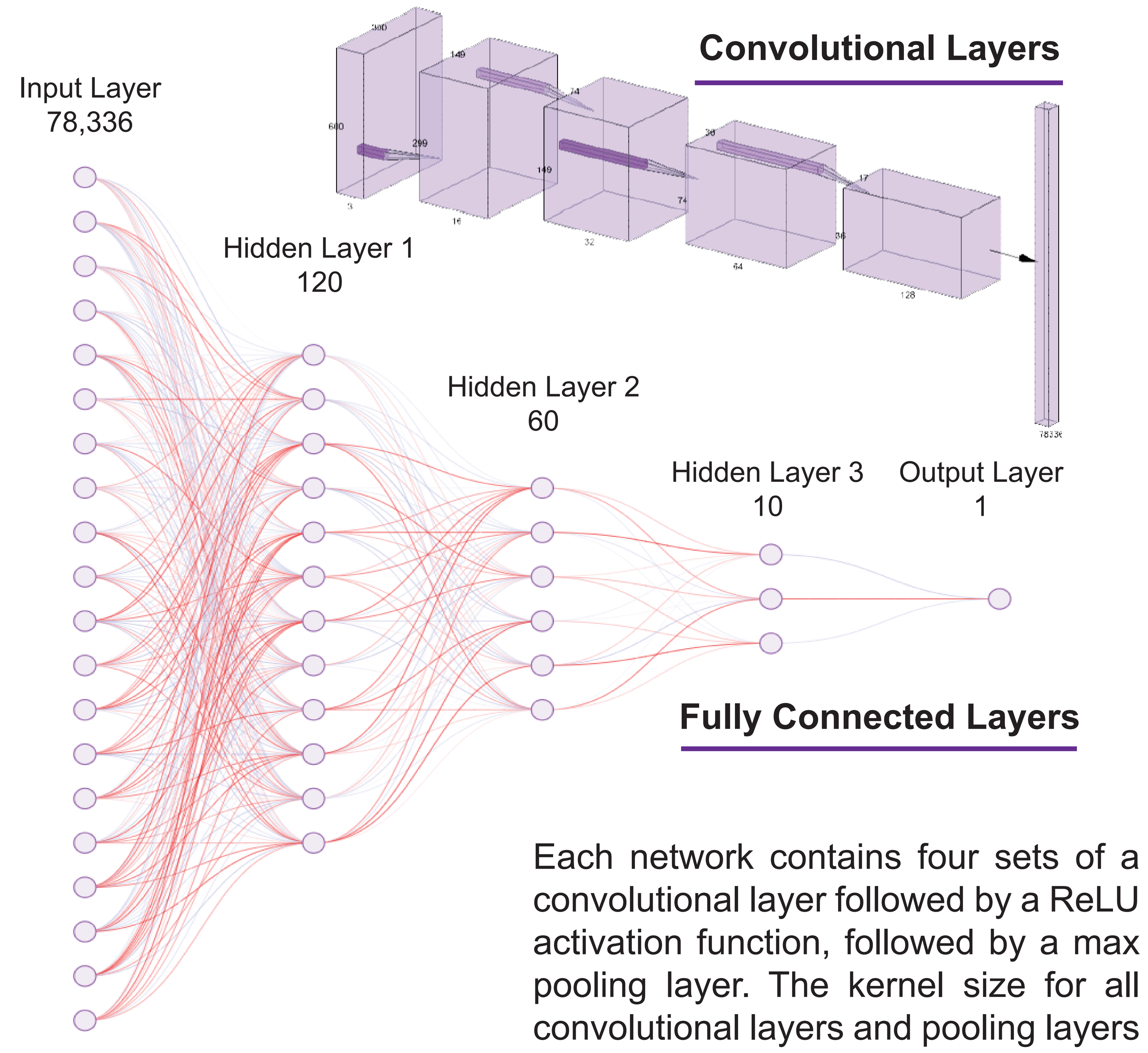
Methods

Data Preprocessing

Name	Size
40.726817_-73.975988.jpg	5 KB
40.740289_-73.974934.jpg	5 KB
40.773335_-73.988123.jpg	5 KB
40.789593_-73.929984.jpg	5 KB
40.708939_-74.013172.jpg	5 KB
nan.jpg	6 KB
40.726354_-74.000407.jpg	7 KB
40.755918_-73.963213.jpg	9 KB
40.75687464_-73.97054389.jpg	11 KB
40.767708_-73.981953.jpg	14 KB
40.753493_-73.990661.jpg	15 KB
40.765078_-73.953597.jpg	16 KB
40.74751483_-73.96879637.jpg	18 KB
40.775297_-73.961816.jpg	19 KB

- Filtering out and deleting files smaller than 6KB, identified by error messages instead of visual content on Google Images.
- Split images randomly for each borough, allocating 90% for training and 10% for testing.

CNN Model Structure



Each network contains four sets of a convolutional layer followed by a ReLU activation function, followed by a max pooling layer. The kernel size for all convolutional layers and pooling layers are each 3, with no stride or padding

The convoltional layers output 16, 32, 64, 128 channels in that order. This is followed by one fully connected layer with a sigmoid activation function to make the model a logistic regressor. The loss function used for training was binary cross entropy, and the optimal model was selected using total test-set accuracy.

Results

Staten Island – Confusion matrix on Testing Set
Total Accuracy: 0.5708
Precision: 0.1982
Recall: 0.6509

TP: 289	FP: 1169
FN: 155	TN: 1472

Brooklyn – Confusion matrix on Testing Set
Total Accuracy: 0.5979
Precision: 0.4758
Recall: 0.6438

TP: 1030	FP: 1135
FN: 570	TN: 1506

Queens – Confusion matrix on Testing Set
Total Accuracy: 0.5837
Precision: 0.5074
Recall: 0.6538

TP: 1273	FP: 1235
FN: 674	TN: 1405

The Bronx – Confusion matrix on Testing Set
Total Accuracy: 0.5718
Precision: 0.3062
Recall: 0.6172

TP: 516	FP: 1169
FN: 320	TN: 1472

Manhattan – Confusion matrix on Testing Set
Total Accuracy: 0.7296
Precision: 0.3884
Recall: 0.6840

TP: 435	FP: 685
FN: 201	TN: 1956

- The Final Results are five models for each borough. Each model takes roughly 6 hours to train.

Conclusion

- This is quite a difficult image classification problem, because intuitively this would be hard for even a human to do. Although the accuracies and other metrics are not amazingly high, we believe this to be a success because most of the models have a recall at or higher than 0.65 and an accuracy at or larger than 0.6. Although those are not high numbers, it means that the model is still abstracting and finding some visual features that may lead to burglary.
- An interesting result is the relatively high accuracy of the model trained on Manhattan data when compared to the other models. This means that there are more visual cues/features that could be used to predict burglary in Manhattan. This could be due to the difference in cars and traffic, the relatively higher percentage of apartment buildings and low percentage of houses, its relative size, and a variety of other factors.
- Staten Island has the lowest accuracy, which we believe could be due partly to the lower standard deviation of median income of Staten Island when compared to the other boroughs. Having a lower standard deviation, mean that many of the neighborhoods should have a similar amount of money and infrastructure upkeep, so there will probably be less or less obvious visual indicators.
- Next Steps include running the model more to try increase accuracy, and looking inside the trained convolutional layers to attempt to identify what features they have learned to detect in images.

Reference

[1] Herbert, D. T., & Hyde, S. W. (1985). Environmental Criminology: Testing Some Area Hypotheses. Transactions of the Institute of British Geographers, 10(3), 259–274. https://doi.org/10.2307/622177
[2] Jabareen, Y., Eizenberg, E., & Hirsh, H. (2019). Urban landscapes of fear and safety: The case of Palestinians and Jews in Jerusalem. Landscape and Urban Planning, 189, 46-57. https://doi.org/10.1016/j.landurbplan.2019.04.010
[3] Vandeviver, C. Applying Google Maps and Google Street View in criminological research. Crime Sci 3, 13 (2014). https://doi.org/10.1186/s40163-014-0013-2
[4] Raskar, Ramesh & Naik, Nikhil & Philipoom, Jade & Hidalgo, Cesar. (2015). Streetscore -- Predicting the Perceived Safety of One Million Streetscapes. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 10.1109/CVPRW.2014.121.