CNNs for Crime Risk Prediction Utilizing Street View & Satellite Imagery

Ricky Toh, Johannes Hui, Jonah Wu

CS 230, Stanford University - https://www.youtube.com/watch?v=IGieFVhDTxk

Stanford Computer Science

Introduction

Problem: Crime risk prediction remains difficult in regions where pertinent data fields are missing. Moreover, those data fields can be subject to historical bias.

Goal: Build a convolutional neural network that can predict relative crime risk of given latitude/longitudinal points in S.F using satellite & street view data.

Model: We initially used a VGG-16 model trained and on street view data, to predict three classes of crime risk.

Street view Examples:

Data and Existing Models

Collection: We used the Google street view and satellite image APIs for data collection. Collecting 9645 images of each (1 per intersection)

Processing: We split our data into 90% training set, 10% validation set.

Crime Prediction using Satellite Imagery with AlexNet : trained on Google Satellite images in Chicago [1]

Stanford University

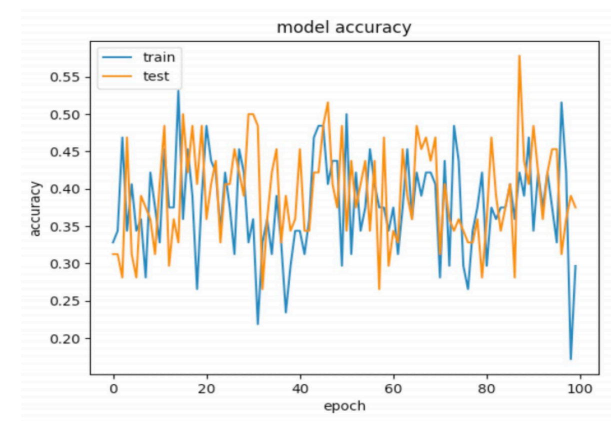
VGG16 Street View Model

Overview: In our VGG16 Street View model, we used street view data labeled into three crime risk classifications – low, medium, and high.

Label Generation for Images:

- Pulled data from SF Crime Dataset since 2018. Enumerating # of felony crimes per intersection.
- Chose # of classes = 3. 'Low' Class was for locations with bottom 40th percentile crime counts, 'Medium' Class was 40th to 70th percentile, 'High' class was > 70th percentile.

Model Details: A standard VGG-16 model with Sparse Categorical Cross Entropy loss, Adam Optimization, minibatch processing @ 256 images, image net weights, softmax prediction layer.



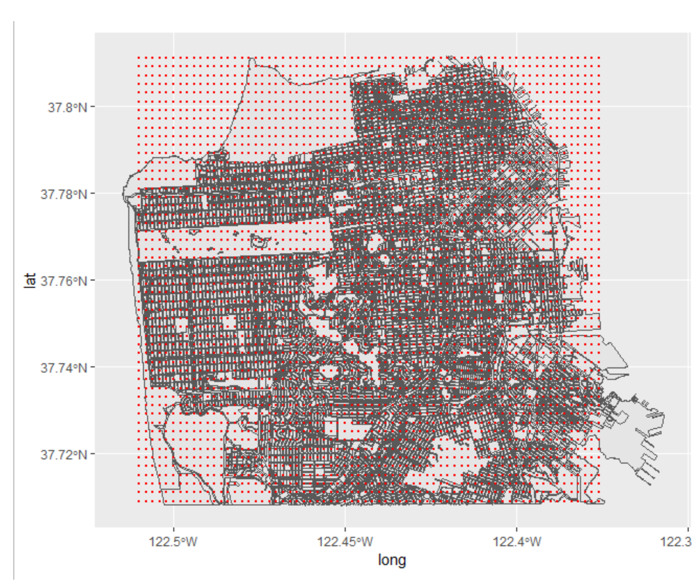
Results: Model was unsuccessful and overfitted.
Accuracy hovered around 40% and the loss function remained erratic.

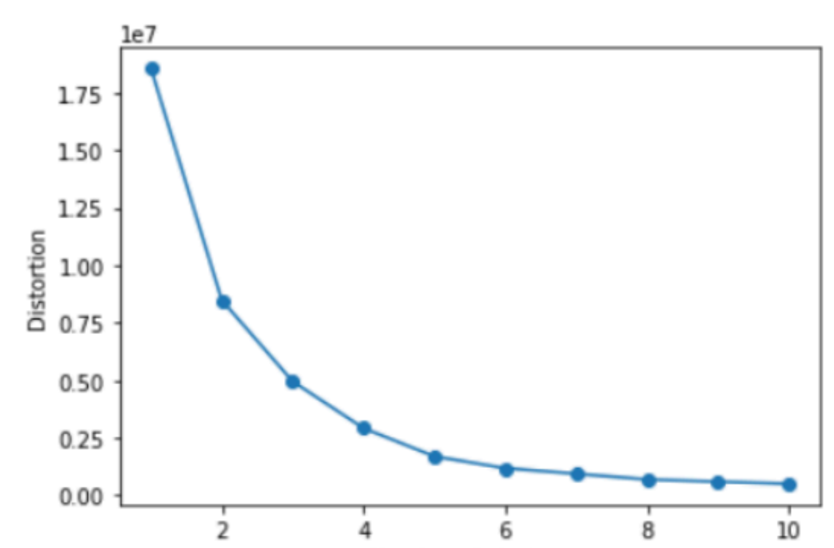
Adjusted VGG16 Street View & Sat Imagery Model

Overview: Our original VGG16 model, with a number of adjustments including k-means clustering for images, discretizing our regions, use of different optimization parameters. Tested on streetview and satellite view data.

Label Generation for Images:

- Pulled data from SF Crime Dataset since 2016 amounting to ~2x more crimes (~330,000 vs. ~160, 000) . Enumerating # of felony crimes per intersection.
- Binned intersections into discretized regional grids to minimize data sparsity problem.
- Enumerated crime counts per grid based on intersection membership .
- Utilized k-means clustering with 3 classes on the discretized regions based on distortion.
- Utilized log transform to compensate for biased distribution toward low crime data.





Model Details: VGG-16 model with Sparse Categorical Cross Entropy loss w/ logits (no softmax activation to improve numerical stability), SGD Optimization with momentum, and minibatch processing @ 128 images. Two additional 512 unit fully-connected layers prior to prediction layer.

Results: Model still unsuccessful and overfitted. Accuracy remained comparable to prior model. When utilized on satellite imagery, accuracy tipped to 50%, but loss function remained just as erratic indicating inability to fit the data.

Dual Input VGG-16 Model

Overview: Two VGG-16 models trained in parallel with one taking in street view and the other taking in satellite imagery.

Label Generation for Images: Same as the second model.
Model Details:

- Multi-input model trained using keras functional API methods.
- Two vgg-16 models trained in parallel one with satellite imagery, the second with street view.
- 512-unit output of each was concatenated into a 1028-unit fully connected layer.
- Fed 512, 256, 128-unit, FC layers prior to 3-unit prediction layer.
- Experimented with dropout layers in between.
- Adam optimizer, same loss function and other hyperparameters as model #2.

Results: No significant improvement.

Challenges & Discussion

Key Challenges:

1. Poor images – Upon analysis of our street view data, it was difficult to see how our model would be able to accurately learn weights given the variability of images. Visible intersections were not reliably picked out by the street view API.
2. Poor clustering – Clustering based on crime count led to a large biased distribution to low crime regions, this led us to use log transforms and oversampling methods to gain a more even distribution, but this may have reduced our model's ability to effectively learn.
3. Data scarcity problem – We may not have pulled enough images, especially given the lack of clear features present for our model to pick out.

Discussion & Further Work

- First and foremost, better quality data and greater training volume along with a richer feature set to cluster around would likely improve our bias problem.
- Use of multi-inputs, including crime-type, time of day, season, etc.
- Include features of neighboring regions into each training sample. [2]

References

[1] Najjar, A., Kaneko, S. I., & Miyanaga, Y. (2018). Crime Mapping from Satellite Imagery via Deep Learning. *arXiv preprint arXiv:1812.06764*.

[2] Duan, L., Hu, T., Cheng, E., Zhu, J., & Gao, C. (2017). Deep convolutional neural networks for spatiotemporal crime prediction. In *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)* (pp. 61-67). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing