CS Capstone Problem Statement

Pedestrian Counting and Privacy Preservation

PREPARED FOR

Oregon State University

DR. FUXIN LI

PREPARED BY

Group 9
Pavement Prometheus

Ian McQuoid *(mcquoidi)*
Mazen Alotaibi *(alotaima)*
Miles Bigelow Davies *(daviesmi)*
Stephanie Allison Hughes *(hugheste)*

Abstract

The City of Portland is updating their data gathering system to better integrate data and technology into the decisions made by the city. One issue that arises is that privacy preservation is often at odds with data gathering. Our task is to provide data on, and hopefully a solution to, this issue. Mainly our concern is manipulation of data so the collected data can be stored and analyzed without drawing the ire of the population or violating privacy portions of the city's social contract. One such solution may arise from using existing tools SSD [1], YOLOv3 [2], Mask R-CNN [3], and Dr. Li's Comumpter Vision model [4] to extract usable information from the data.

CONTENTS

1 PROBLEM

Data collection has always been at odds with privacy. Up until 2018, Portland had rectified this difference by relying on training individuals to stand on city corners and collect census data in lieu of more modern (and invasive) techniques involving cameras. Despite the antiquated and error prone nature of this method, Portland persisted in its use until the Smart City PDX program was initiated by Mayor Ted Wheeler. This program, a broad initiative designed to redress societal inequities and improve under-served communities, relies on updated infrastructure for data collection to inform decisions on enacting policies. To this end, the City of Portland has partnered with Intel and General Electric to procure a wide array of sensors that will serve in gathering transit data on pedestrians and vehicles. This data can be used for many purposes including planning city transit to be more efficient, reducing traffic related accidents, and having general census information for any future community projects. One problem is that using unadulterated sensor information is problematic with regards to protecting civilian privacy.

Portland's primary constraint is that they must make sure technology serves the interests of its city residents. As such, the city requires that identifying-features of any data collected be removed. This would imply that all data operations need to be performed in real time with sensor collection. This is a prohibitive requirement since nearly any form of statistical analysis is a computationally intensive process. Modern limitations being what they are, it is nearly impossible to perform an analysis of any depth under these constraints – especially on a budget. We then must preserve some information so we can perform future metrics against the data. Any realistic solution will need to balance two factors: the amount of processing power available, and the amount of information preserved. The former factor is a definite metric determined entirely by the hardware provided, while the latter is negotiable. What one person deems permissible information another might consider invasive. This preservation of privacy, in turn, needs to be balanced against the amount of information necessary to perform any useful analysis.

2 PROPOSED SOLUTION

There are two key questions we must answer when creating a solution. The first is "How far can we mangle the video while still preserving the important data?" This is key because the further we can alter the video data, the more likely that saving the data will be considered "OK" by the community. Google has implemented ideas in this area for their Streetview software for Google Earth. Their software attempts to blur faces, license plates, and various other information. The issue is that Streetview often over compensates, or misses information. This is a huge problem because the cost for false negatives, not blurring pedestrians' faces, is much greater than the cost for false positives, blurring different objects rather than pedestrians' faces. Four key computer vision models have been introduced to our team by Dr. Li [5]: You Only Look Once (YOLOv3) [2] and Single Shot MultiBox Detector (SSD) [1] (for real-time object detection), Mask Region-Convolution Neural Networks (Mask R-CNN) [3] (for instance segmentation), and Dr. Li's Computer Vision model [4] (for assigning an ID for each pedestrian for continuous identification of assigned masks).

According to Dr. Li [5], our team is required to find a middle point allowing us to preserve the privacy of pedestrians while generating useful traffic data. We will build a computer vision model for detecting pedestrians within each frame. State-of-the-art computer vision models, such as You Only Look Once (YOLOv3) [2] and Single Shot MultiBox Detector (SSD) [1], have poor accuracy in detecting objects. These models average %63 at 78FPS and %74 at 59FPS, respectively, compared to other Computer Vision models that have a higher accuracy, but need more computation time for object detection. Therefore, pedestrian detection using YOLOv3 and SSD increases the chances of having more false negatives,

not blurring pedestrians' faces, due to the low accuracy of real-time analysis. To reduce the chances of having false negatives generated by our model we will need to separate the object detection model into two stages: pedestrian detection and face detection.

We have decided to separate the project into three phases for the development pipeline: pedestrians detection, object mangling, and extracting public opinion. For the first phase, we have three detection methods that we will be testing: pedestrian detection only using YOLOv3 [2], pedestrian detection and face detection using YOLOv3, and pedestrian detection using YOLOv3 and applying instance segmentation using Mask R-CNN [3]. Each pedestrian method has pros and cons. When applying object mangling on pedestrian detection only using YOLOv3 we might secure the privacy of the pedestrian, but we will lose a lot of information about the traffic because more data will be lost to mangling. Therefore, by applying object mangling on face detection after pedestrian detection, we will preserve the privacy of pedestrians while still having some information about the traffic. This solution will require a lot of computation time and power since it means applying YOLOv3 twice in sequence. Lastly, if the general public requires our model to mangle more than their faces, we will need to use Mask R-CNN rather than face detection. Once again, this method require a lot of computation power and time. Another problem is that we will lose a lot of spatial information of the traffic due to frames lost by Mask R-CNN.

For the second phase -object mangling-, we will need to build mangling methods to mangle pedestrians' images without the ability to reverse the image transformations. Therefore, we have three methods for object mangling that we will be testing: mangling pedestrians using the same mask and color, same mask and different color, and different mask and color. Mangling pedestrian images using the same mask is easy to compute and we guarantee the loss of specialty of the detected pedestrian. However, we will lose some information on the pedestrian's behavior and the traffic. In contrast, mangling using the same mask and different color will solve the lost information problem, but it will require us to use Dr. Li's Computer Vision model [4], which will increase computation time. In addition, if we wanted to extract more information about pedestrians' behavior and the traffic we would need to mangle using a different mask and color. Doing so will increase the computation time and may result in an increased security risk since the mask will have some kind of correlation with the detected pedestrian. For the last phase, we will gauge the public opinion of the each of the mangling methods. Then we will use the model with the median privacy level accepted by the general public and the highest level of information preservation.

3   Performance Metrics

In the hopes that our group can deliver tangible traffic efficiency suggestions to the city of Portland, the project will be evaluated in multiple stages with major requirements focused on accuracy of tracking and privacy coverage of pedestrians. The first area of the project actively measured is the accuracy of the program's ability to recognize and track pedestrians. The end goal is achieving at least 70 percent accuracy. Ensuring a high accuracy in pedestrian tracking is important since the data can affect the level of safety that can be achieved through new changes in pedestrian traffic flow. Another important measured aspect is the level of privacy coverage people feel the program has. This measurement will first be recorded through gauging pedestrians' responses on which level of privacy people are most comfortable with when it comes to the video data used. Based on the responses of the people surveyed, our group will develop the software to maintain that level of privacy. Each level of privacy will have a certain amount of hiding of ones identity. People will be surveyed throughout the build cycle so that our group will design the software with the customers needs

always in mind. Consistent feedback and evaluation of privacy standards should meet the pedestrian's expectations by the time we have a final working prototype. The program should also be efficient enough to be used with the amount of processing power the street light provides. Through the accuracy checks, privacy checks, and efficiency checks, our group hopes to have a deliverable solution that provides suggestions for areas of traffic that need to be altered to optimize traffic safety.

REFERENCES

[1] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," *ArXiv e-prints*, Dec. 2015.

[2] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *ArXiv e-prints*, Apr. 2018.

[3] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *ArXiv e-prints*, Mar. 2017.

[4] C. kim, F. li, and J. Rehg, "Multi-object Tracking with Neural Gating Using Bilinear LSTM," Mar. 2018.

[5] F. Li, Group meeting, October 2018.