Developing a real-time foot traffic counter in New York City using publicly available camera feeds

Gregory Dobler[1,2◑], Jordan Vani[2◑], Trang Dam[2,3¤], Name4 Surname[2], Name5 Surname[2‡], Name6 Surname[2‡], Name7 Surname[1,2,3*], with the Lorem Ipsum Consortium[¶]

1 Affiliation Dept/Program/Center, Institution Name, City, State, Country
2 Affiliation Dept/Program/Center, Institution Name, City, State, Country
3 Affiliation Dept/Program/Center, Institution Name, City, State, Country

◑These authors contributed equally to this work.
‡These authors also contributed equally to this work.
¤Current Address: Dept/Program/Center, Institution Name, City, State, Country
†Deceased
¶Membership list can be found in the Acknowledgments section.
* correspondingauthor@institute.edu

Abstract

Filler text

Author summary

Filler text

Introduction

Filler text [1]

Materials and methods

Source data

The New York City Department of Transportation (DOT) maintains a network of real-time traffic cameras across the New York City (NYC) metro area. These camera feeds are publicly available and are served at approximately one frame per second–with each frame being a 352x240 RGB .jpg. All together, the 650 cameras are distributed across the five NYC boroughs (see Table 1 and Fig 1 (map of cameras across boroughs)) and captures a continuous record of the NYC streetscape including roadways, building facades, and pedestrian facilities. With our focus on pedestrian foot traffic, each camera was hand-labelled as including pedestrian facilities or not; Fig 2-3 (pedestrian facility images) illustrate the distinction between cameras with and without pedestrian facilities.

Table 1. DOT cameras by New York City borough.

Borough	Bronx	Brooklyn	Manhattan	Queens	Staten Island	Bridges	Total
Total cameras	44	105	226	195	41	39	650
Cams w ped. facilities	13	30	157	65	11	5	281

Training data

3918 daytime images were scraped from 17 cameras between 11:00-20:00 on April 30th, 2016 and June 19th, 2016. The 3918 images were then labelled by hand for positive and negative examples using 3:4 bounding boxes; the resulting labels were not exhaustive (i.e., not all pedestrians were labelled). Across the 3918 images, 16022 positive examples and 41449 negative examples were labelled–approximately a 2:5 (pos:neg) ratio. XML files were subsequently written for each image, following the PASCAL Visual Object Classes format. Ultimately, this dataset was split using a 70:30 train:test split.

Training the Faster RCNN

A Faster RCNN was trained using Tensorflow following the VGG16 network structure. Training parameters included: (1) a learning rate of 0.0005; (2) a Region Proposal Network (RPN) batch size of 256; (3) an RPN positive overlap of 0.7; and, (4) a minimum RPN size of 2x2. Using this setup, the network was trained for a total of 90,000 iterations on a GeForce GTX 1080 Ti GPU.

Data collection

In order to collect data from all DOT cameras, our Faster RCNN Tensorflow session was initialized. Next, looping over all cameras, an image was downloaded if available and the time at download was recorded. The downloaded image was parsed for available metadata (camera direction and timestamp) overlain on the image. Then, the image was passed to the Faster RCNN which output the number of pedestrian detections. And, lastly, the resulting data (camera id, time at download, camera direction, overlaid timestamp, and the number of pedestrian detections) were saved to a local database. In addition, for each loop over all 650 cameras, 1 image was saved to file at random. Ultimately, each loop over all cameras took approximately 70 seconds, enabling pedetrian detection across all cameras at temporal resolution of nearly 1 minute.

Measuring precision and recall

Filler text

Fig 1. Bold the figure title. Figure caption text here, please use this space for the figure panel descriptions instead of using subfigure commands. A: Lorem ipsum dolor sit amet. B: Consectetur adipiscing elit.

Results

Filler text

Discussion

Filler text

Conclusion

Filler text

Supporting information

S1 Fig. Bold the title sentence. Descriptive text (optional).

Acknowledgments

Filler text

References

1. Conant GC, Wolfe KH. Turning a hobby into a job: how duplicated genes find new functions. Nat Rev Genet. 2008 Dec;9(12):938–950.