Quantifying social distancing compliance and the effects of behavioral interventions using computer vision

Derek Gloudemans
Vanderbilt University
Institute for Software Integrated Systems
Nashville, Tennessee, USA
derek.gloudemans@vanderbilt.edu

Nicole Gloudemans
Vanderbilt University
nicole.k.gloudemans@vanderbilt.edu

Mark Abkowitz
Vanderbilt University
mark.abkowitz@vanderbilt.edu

William Barbour
Vanderbilt University
Institute for Software Integrated Systems
william.w.barbour@vanderbilt.edu

Daniel B. Work
Vanderbilt University
Institute for Software Integrated Systems
dan.work@vanderbilt.edu

ABSTRACT

Social distancing has become a pressing and challenging issue during the Covid-19 pandemic. In a smart cities context, it becomes possible to measure inter-personal distance using networked cameras and computer vision analysis. We deploy a computer vision pipeline based on Retinanet that identifies pedestrians in streaming video frames, then converts their positions to GPS coordinates for distance calculation and further analysis. This processing is applied to nine camera streams at three locations from around Vanderbilt University. We collect 70 hours of baseline distancing data over the course of two weeks, after which time we deploy small behavioral interventions at the three locations aimed at increasing distancing compliance. Another 70 hours of data with the interventions in place will be analyzed against the baseline data to determine if they had an effect on distancing compliance.

KEYWORDS

object detection, social distancing, computer vision, smart cities

ACM Reference Format:
Derek Gloudemans, Nicole Gloudemans, Mark Abkowitz, William Barbour, and Daniel B. Work. 2021. Quantifying social distancing compliance and the effects of behavioral interventions using computer vision. In *The Workshop on Data-Driven and Intelligent Cyber-Physical Systems (DICPS '21), May 18, 2021, Nashville, TN, USA*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3459609.3460523

1 INTRODUCTION AND CONTRIBUTIONS

The use of cameras as a multi-functional sensor is well established in smart cities research and in deployments around the world. Camera installations began as a tool primarily for CCTV security, monitored by humans. Modern installations have evolved into more widespread systems, both in terms of size and diversity of uses, and are heavily supplemented by computer analytics. Using a combination of human monitoring and computer analytics, smart cities applications in traffic control, public transit, safety, urban planning, public health, and more, are now possible.

The onset of the Covid-19 pandemic created a very sudden need for social distancing practices in many contexts. Some companies need to keep their workforce on site and cities need to maintain public services and space, all while preserving health safety. The Centers for Disease Control and Prevention (CDC) maintain that keeping 6 feet (2 meters) of personal space between one's self and others not in one's household is recommended [2]. For situations where the individuals are sedentary or minimally mobile, spaces compliant with social distancing can be delineated (e.g., offices, restaurants). However, contexts with free and less predictable movement of individuals (e.g., hallways, sidewalks, stores) pose a challenge even for individuals with the best intentions.

Herein lies the opportunity for cameras to be used in another smart cities application: social distancing compliance. The expectation is not that cameras and video analytics would be an *enforcement* tool. Rather, they can collect data on areas and circumstances that are problematic for social distancing compliance and give us the capability of studying compliance and strategies to increase voluntary compliance. The resultant data is useful for public health authorities, companies, and organizations to design or modify spaces for pandemic safety, put in place rules for distancing, and nudge people's behavior to encourage distancing.

The contribution of our work is that it is the first to quantitatively evaluate the impact of small behavioral interventions on social distancing compliance. We use computer vision analysis on nine video streams at Vanderbilt University to collect anonymized data on positions of pedestrians, from which we calculate a distance matrix between individuals. Distancing data from behavioral interventions will be compared to baseline data to determine the effectiveness of the interventions.

DICPS '21, May 18, 2021, Nashville, TN, USA

ACM ISBN 978-1-4503-8445-2/21/05…$15.00
https://doi.org/10.1145/3459609.3460523

2 LITERATURE REVIEW

Other works have deployed computer vision algorithms for calculating social distance between people in video, but none have studied the impact of behavioral interventions in this context. Prior work has compared the speed and accuracy of various object detection and tracking algorithms for this problem; they use the centroid of bounding boxes to determine inter-personal distance [9]. A system deployed in [11] focuses on determining distancing in a real-time privacy-preserving environment, while additionally providing warnings and flow control for dense settings. Amazon has introduced a camera-based social distancing detection system and released open-source code with the idea that such a system may prove useful in a warehouse or office setting [8].

Studies have also deployed vision-based social distancing calculation targeted at city-scale analytics. Video analyzed from cameras in a central area of Cucuta, Colombia found that 84% of people didn't comply with two meter distancing guidance [7]. Distancing compliance was integrated into a larger Covid-19 data dashboard in [12], with results gathered during limited periods across 2020 for New York City. A large-scale, worldwide data collection effort was undertaken in [4], where they automatically discovered publicly-accessible network cameras and periodically processed extracted images; this allowed the determination of crowd size and social distancing compliance at many locations over time (April-August 2020).

The distancing study in [1] focuses on more accurate calculation of inter-personal distance using body pose detection and body dimension consideration. Other studies have used more advanced imaging technology to aid in detection and distance calculation. Depth cameras were used in [6] to aid sight-impaired individuals in maintaining distance. Object detection was retrained on thermal cameras by [10] for a social distancing measuring system.

3 METHODOLOGY

3.1 Computer vision analysis

In this section we provide details about the computer vision and data transformation pipeline. The pipeline takes as input the network addresses to available camera streams. It outputs, for each video frame that is analyzed, the timestamp corresponding to the video frame, source camera stream identifier, GPS coordinates of all detected pedestrians, cyclists, and vehicles, number of detected pedestrians, and number of detected social distancing infractions. Video data is not stored, so the data contains no personally identifiable information. When individual annotated video frames are exported for demonstration and validation purposes, detected pedestrian and cyclist faces are automatically pixelated.

Video frames are ingested by separate CPU processes that keep the most recent three frames buffered in memory. Each GPU that is available for processing is managed by a separate worker process. Whenever each process is idled after finishing computing a video frame, it fetches a new frame from the next camera stream in the queue. Camera streams are cycled in a rotating queue, thereby evenly distributing the frequency with which they are queried.

Each video frame is processed by an object detector: a class of deep learning algorithms that takes image pixels as input and outputs a set of bounding boxes, classes, and confidences that correspond to each object it finds in the image. We use a Pytorch implementation of Retinanet with a Resnet 50 backbone [5]. All object detections except for pedestrian, cyclist (counted as pedestrian in final output data), car, bus, and truck are ignored in the output from the detector; low confidence detections are also suppressed.

A homography matrix is used to transform detection locations in image pixel space into real world space. These matrices define the transform for each camera into GPS coordinates. Homography matrices are established for each camera by identifying four matching points in both GPS coordinates and image pixel coordinates and solving a system of linear equations [3]. The bottom center of the bounding box for each detected object is taken as the assumed point location for the object (i.e., the pedestrian's feet). Given pedestrian locations in GPS coordinates, we compute the distance matrix between multiple pedestrians in a video frame using Haversine distance and classify a social distancing infraction as an inter-personal distance of less than six feet. We note that an "infraction" is defined purely by the recommended social distance (6 feet/2 meters) and does not speak to legality due to the variety of legal standards.

Video frames are sometimes annotated for demonstration or validation purposes. Bounding boxes of detected objects are superimposed on the image and ellipses are drawn around the feet of pedestrians to indicate their six-foot distancing zone. The ellipses are colored according to the lowest detected inter-personal distance for each pedestrian: green, > 10ft; yellow, 8-10ft; orange, 6-8ft; red, < 6ft. Two examples of annotated images from the system are shown in Figure 1.

3.2 Camera sites and behavioral interventions

The data pipeline is applied to nine camera streams from three critical areas on the Vanderbilt University campus in Nashville, TN, USA, summarized in Table 1. The camera at each location has three lenses, which provides three streams for different areas around the installed location. Baseline distancing data will be collected for seven hours per day (11am to 6pm, local time), Monday through Friday, for two weeks – a total of 70 hours.

Minor interventions will be targeted at each site in order to promote better social distancing compliance. Around the intersection sites (Broadway and West End), sidewalk stickers and temporary folding signs with social distancing messaging will be deployed. On the Edgehill pedestrian bridge site, ribbon tape will be strung up for a 150ft stretch of the bridge to delineate the directional flow of pedestrians (shown in Figure 2); this makeshift barrier could be easily crossed by pedestrians ducking under the ribbon, but it promotes directionality for walking.

Distancing data will be collected with the interventions deployed for the same amount of time as the baseline data, thereby capturing as similar conditions as possible (conceding the effects of weather and other externalities). Data analysis will be conducted on these two datasets – baseline and intervention – to determine if the frequency of social distancing infractions decreased a statistically significant amount due to the interventions that were deployed. We will also analyze the trends of distancing infractions to determine if they have significant geographical or temporal trends that could be further targeted to increase distancing compliance.

Figure 1: Edgehill (top) and West End (bottom) locations with annotations for detected pedestrians and vehicles. Green shaded circles indicate social distancing being maintained, while red indicate less than 6ft between pedestrians.

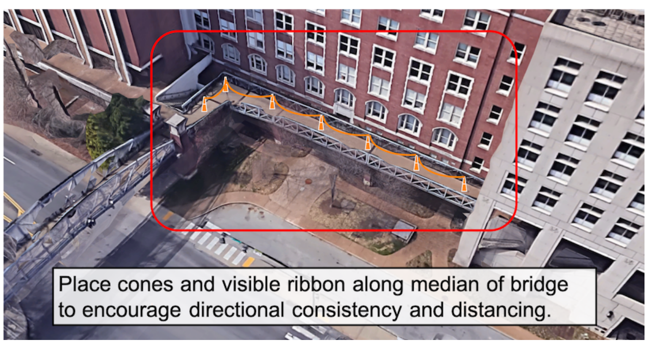

Figure 2: Intervention planned for the Edgehill location on the pedestrian bridge.

Camera	Code	Description
Edgehill & 21st Ave pedestrian bridge	EH21	Bridge crossing a major road that is a known choke point for pedestrians and personal mobility devices.
Broadway & 21st Ave intersection	BW21	Major intersection between Vanderbilt and midtown commercial area.
West End & 21st Ave intersection	WE21	Major intersection for pedestrians crossing the 7-lane West End Ave.

Table 1: Locations where cameras are installed and interventions are deployed.

4 PRELIMINARY RESULTS

Thus far for this work in progress, we have deployed computer vision algorithms running in real time on the camera feeds that are described in Section 3.2. At the time of writing, we have collected nine hours of preliminary data at the Edgehill (EH21) and West End (WE21) locations and one and a half hours of preliminary data at the Broadway (BW21) location.

The time series of pedestrians present and number of infractions present over time is shown in Figure 3. The number of pedestrians is indicated by the solid lines, separated by camera view at the location, and the number of infractions is the dashed lines. One camera view from each location was left off due to very low numbers of observed activity. The time series are presented as rolling average values, taken across the prior 20 minutes. The West End location shows a larger amount of activity overall and a higher frequency of social distancing infractions. The Edgehill location rarely showed infractions at all. Interestingly, we can see the peaks in activity for each location attributed to the time of day: the morning commute and lunch breaks at the West End location, and the morning university class changes at the Edgehill location.

The distribution of inter-personal distance for each location is shown in Figure 4. This distance is computed only for frames containing two or more pedestrians and is cut off in the histograms at twenty feet. The histograms are assembled using data from all video frames, which are only a few seconds apart; they are therefore normalized to better reflect the amount of time that each distance is observed. The observations from the Broadway (BW21) location are included, but are less meaningful due to a low amount of data from this location. We see a distinction between the West End and Edgehill locations: a much larger proportion of observations are non-compliant with social distancing at Edgehill, compared to the flatter distance distribution at West End. This would seem to indicate that pedestrians are walking together at the Edgehill location more often. If distance wasn't being maintained just in opposite direction passing, we would expect to see a far lower frequency. Similarly, if pedestrians walking in the same direction were walking behind each other socially distanced, we would expect a peak in the distribution centered on or extending through the 6-8ft area. From the limited amount of data available at the Broadway location, we do see a sharp disparity between the camera views that invites investigation once we have full datasets collected. It appears that the camera-1 view observes almost exclusively pedestrians inside of 6-foot distance (of those that are inside of twenty feet), whereas camera-2 observes almost none inside of 6 feet but significantly more in the 6-20ft range.

The spatial distribution of infractions that were observed at the Edgehill location (two camera views) are shown in the heatmap in Figure 5. Note that the activity on the elevated bridge appears underneath the bridge due to the three dimensional imagery. We see concentrated distancing issues occurring at the corners of the intersection where pedestrians wait for the crosswalk. There are also some larger areas that indicate pedestrians moving with each other in too close of proximity. These occur on the stretch of the bridge and on the ground level crosswalk parallel to the bridge.

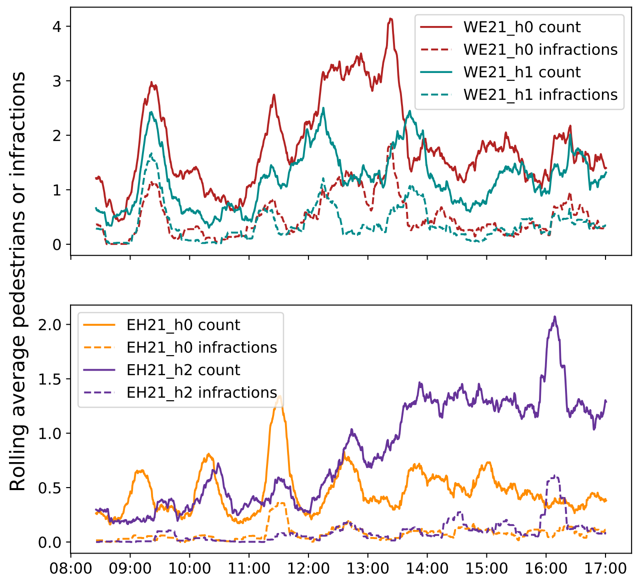

Figure 3: Rolling average number of pedestrians present (solid lines) and number of social distancing infractions (dashed lines) in the scene from each video feed at West End (top) and Edgehill (bottom) locations.

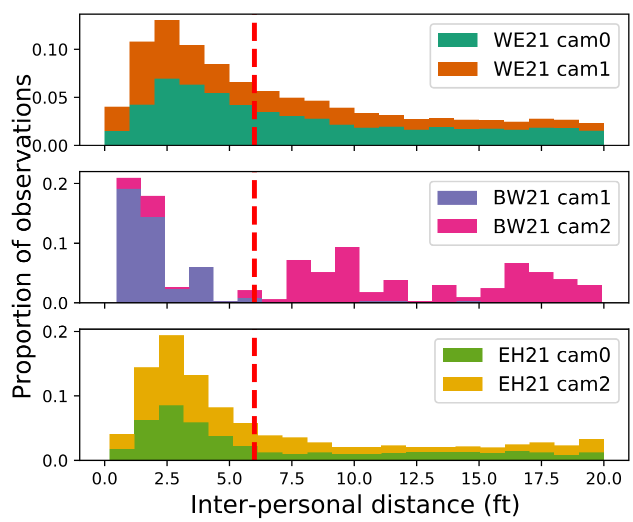

Figure 4: Distribution of inter-personal distance in video frames where greater than two pedestrians are present. Distance is cut off in the histograms at 20 feet.

5 FURTHER WORK

Further work will be focused on the collection of data, as described in Section 3.2. We will perform collection the full baseline dataset, then deploy the interventions at each site and collect the intervention dataset. Analysis of both datasets will be performed in a similar manner as the preliminary results in Section 4, and extended based on observations thus far. Comparison of the baseline and intervention datasets will be performed to determine the effectiveness of these behavioral nudges in promoting social distancing.

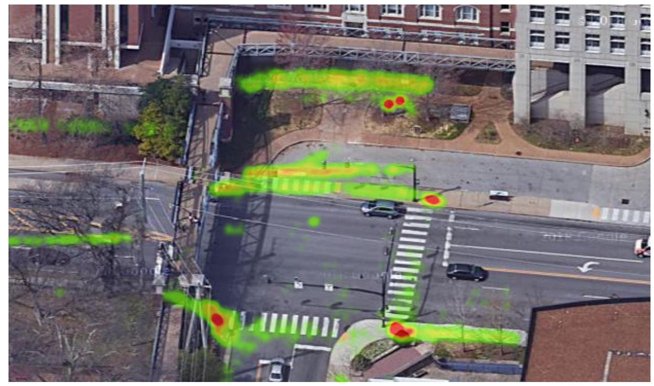

Figure 5: Heatmap of social distancing infractions at the Edgehill location, combined from two camera views.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under the Graduate Research Fellowship Program, grant no. DGE-1937963 (D. Gloudemans, 2019). The authors would also like to thank Vanderbilt University Division of Administration for their ongoing support.

REFERENCES

[1] Maya Aghaei, Matteo Bustreo, Yiming Wang, Gianluca Bailo, Pietro Morerio, and Alessio Del Bue. 2021. Single Image Human Proxemics Estimation for Visual Social Distancing. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2785–2795.

[2] Centers for Disease Control and Prevention. 2020. Social distancing. (2020).

[3] Aubrey Clausse, Salma Benslimane, and Arnaud de La Fortelle. 2019. Large-Scale extraction of accurate vehicle trajectories for driving behavior learning. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2391–2396.

[4] Isha Ghodgaonkar, Subhankar Chakraborty, Vishnu Banna, Shane Allcroft, Mohammed Metwaly, Fischer Bordwell, Kohsuke Kimura, Xinxin Zhao, Abhinav Goel, Caleb Tung, et al. 2020. Analyzing Worldwide Social Distancing through Large-Scale Computer Vision. *arXiv preprint arXiv:2008.12363* (2020).

[5] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*. 2980–2988.

[6] Manuel Martinez, Kailun Yang, Angela Constantinescu, and Rainer Stiefelhagen. 2020. Helping the Blind to Get through COVID-19: Social Distancing Assistant Using Real-Time Semantic Segmentation on RGB-D Video. *Sensors* 20, 18 (2020), 5202.

[7] Carlos Vicente Niño Rondón, Sergio Alexander Castro Casadiego, Byron Medina Delgado, Dinael Guevara Ibarra, and Jorge Gómez Rojas. 2021. Video image processing to verify social distancing during the COVID-19 pandemic. *Revista Logos Ciencia & Tecnología* 13, 1 (2021), 116–127.

[8] Brad Porter. 2020. Amazon introduces 'Distance Assistant'. *About Amazon* (2020).

[9] Narinder Singh Punn, Sanjay Kumar Sonbhadra, and Sonali Agarwal. 2020. Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques. *arXiv preprint arXiv:2005.01385* (2020).

[10] Sergio Saponara, Abdussalam Elhanashi, and Alessio Gagliardi. 2021. Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19. *Journal of Real-Time Image Processing* (2021), 1–11.

[11] Dongfang Yang, Ekim Yurtsever, Vishnu Renganathan, Keith A Redmill, and Ümit Özgüner. 2020. A vision-based social distancing and critical density detection system for covid-19. *arXiv preprint arXiv:2007.03578* (2020), 24–25.

[12] Fan Zuo, Jingxing Wang, Jingqin Gao, Kaan Ozbay, Xuegang Jeff Ban, Yubin Shen, Hong Yang, and Shri Iyer. 2020. An interactive data visualization and analytics tool to evaluate mobility and sociability trends during covid-19. *arXiv preprint arXiv:2006.14882* (2020).