AI-Empowered Decision Support for COVID-19 Social Distancing

Hongchao Jiang[1]**, Wei Yang Bryan Lim**[1]**, Jer Shyuan Ng**[1]**, Harold Ze Chie Teng**[2]**, Han Yu**[2]**,**
Zehui Xiong[1]**, Dusit Niyato**[2]**, Chunyan Miao**[1,2]
[1]Alibaba-NTU Singapore Joint Research Institute (JRI), Nanyang Technological University (NTU), Singapore
[2]School of Computer Science and Engineering (SCSE), NTU, Singapore
{hongchao001, limw0201, s190068, hteng001}@e.ntu.edu.sg, {han.yu, zxiong002, dniyato, ascymiao}@ntu.edu.sg

Abstract

The COVID-19 pandemic is one of the most severe challenges the world faces today. In order to contain the transmission of COVID-19, people around the world have been advised to practise social distancing. However, maintaining social distance is a challenging problem, as we often do not know beforehand how crowded the places we intend to visit are. In this paper, we demonstrate *crowded.sg*, an AI-empowered platform that leverages on Unmanned Aerial Vehicles (UAVs), crowdsourced images, and computer vision techniques to provide social distancing decision support.

Introduction

The Coronavirus Disease 2019 (COVID-19) has infected more than 29 million people worldwide, with number of deaths approaching the 1 million mark. The Centers for Disease Control and Prevention (CDC) has stated that the best way to reduce the spread of the disease is to limit social contact. However, enforcing safe distancing is a challenge.

Epidemiological studies have shown that the virus spreads easily in crowded areas (Rothan and Byrareddy 2020). In Singapore, large COVID-19 clusters were formed in places like worker dormitories and nursing homes with people living in close proximity. Even with various safe distancing measures, places like food eateries and bus stops still experience high traffic flows. An infected but asymptomatic individual visiting such crowded places could cause community transmission of COVID-19, which makes subsequent contact tracing difficult. There is an urgent need for a decision support platform to inform the general public of the crowdedness of places they intend to visit, so that they can make informed decisions when planning their trips.

In this paper, we showcase an Artificial Intelligence (AI)-empowered crowd counting platform - *crowded.sg*[1]. The platform is deployed in Nanyang Technological University (NTU) and has garnered more than 400 users. It leverages on user uploaded images taken at locations of interest (e.g., shops) and aerial images captured with Unmanned Aerial Vehicles (UAVs) (e.g., of sports fields). The images are then analyzed by the platform's AI engine to provide a crowdedness level at each location to help users decide on where to visit.

System Design

The main features of *crowded.sg* are as follows:

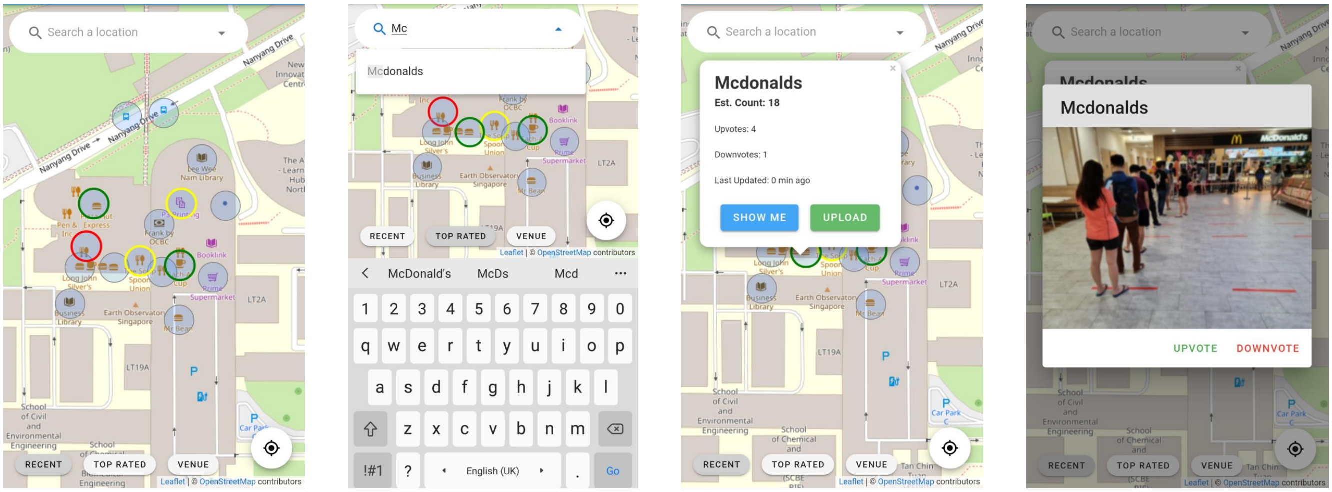
1. *Location markers*: A map interface with locations of interest as circular markers is shown in Figure 1(a). Users can filter between different categories or search for a location by name using the search bar with auto-complete functionality (Figure 1(b)).
2. *Color coded crowd levels*: The color coding scheme at each location is calibrated based on its historical crowd counts. This provides a high level overview of the crowd situation in the vicinity. Details can be obtained by clicking on the markers to get an estimated count (Figure 1(c)) or view the uploaded image which has been algorithmically blurred for privacy reasons (Figure 1(d)).
3. *User contributed photos*: Locations without data or with outdated data are represented by translucent circles. By maintaining the visibility of these locations, the users can be *nudged* (Thaler and Sunstein 2009) to snap a picture with their smartphones and upload it.
4. *User curation of photos*: It is not feasible to manually curate all user uploads. As such, we rely on the community to upvote images that accurately reflect the current situation and downvote those that do not (Figure 1(d)). Images with high net downvotes are automatically removed.

The AI Engine

The AI Engine (Figure 2) combines state of the art computer vision models for crowd counting. According to (Loy et al. 2013), the literature can be generally classified into three categories: detection-based, regression-based, and density estimation-based methods. By considering their strengths and weaknesses, we incorporate the detection-based method, Mask R-CNN (He et al. 2017) and the density estimation method CSRNet (Li, Zhang, and Chen 2018) into a single workflow for accurate crowd density estimation.

Mask R-CNN with ResNet101 (He et al. 2016) backbone and Feature Pyramid Network (Lin et al. 2017) is pre-trained on the MS-COCO dataset (Lin et al. 2014). To ob-

[1]A video demonstration of the platform can be accessed at https://youtu.be/NUoChH3VeOU



(a) (b) (c) (d)

Figure 1: The crowded.sg user interaction design.

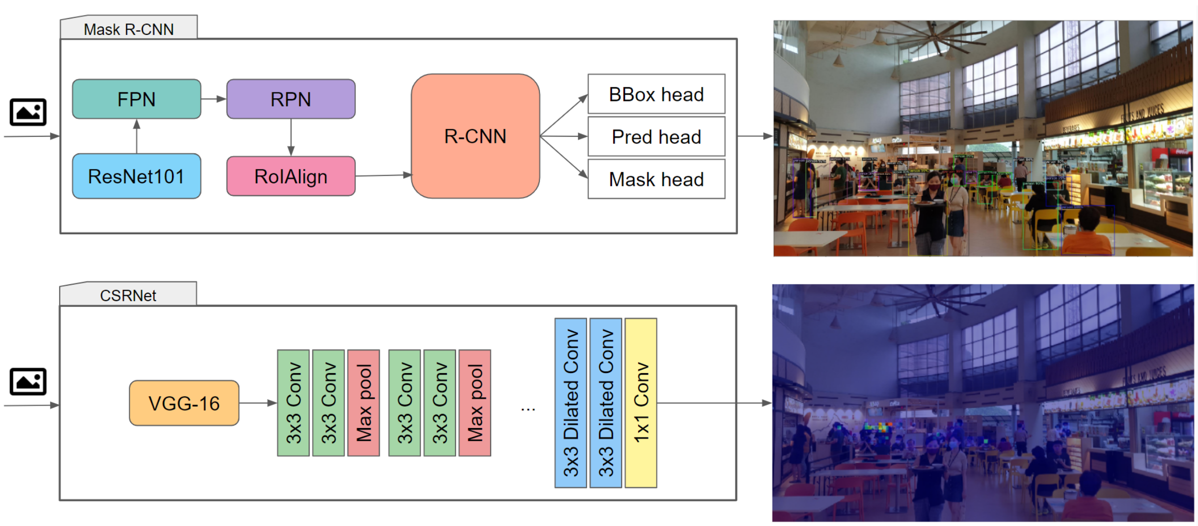


Figure 2: The AI Engine with CSRNet and Mask R-CNN.

tain the crowd count, we sum the number of detection of the "person" class. However, when the image contains a dense crowd, occlusion and cluttered backgrounds degrade the detection accuracy (Gao et al. 2020).

CSRNet is pre-trained on the ShanghaiTech Part A dataset (Liu et al. 2018). Instead of detecting each individual person, it tries to find regions in the image which contain a high concentration of people. A feature map is subsequently produced with a large weight given to such regions. To obtain the crowd count, we aggregate the feature map. Although CSRNet can deal with dense crowds, we observed that it tends to pick up on non-existent patterns and overcount with sparse crowds.

For our AI engine to handle both sparse and dense crowds, we heuristically alternate between the two models. Mask R-CNN is used by default due to its high precision. If it detects a large number of people (i.e., more than 15), CSRNet is used instead.

User Evaluation

We have successfully deployed *crowded.sg* within the NTU campus from 1st August, 2020. Within a month, it has garnered more than 400 users. A user study involving 100 of them has been conducted, with questions designed based on the System Usability Scale (SUS) (Bangor, Kortum, and Miller 2008) to study the usability and usefulness of the system. The results are shown in Figure 3. Among the respondents, 80% of them found our platform easy to learn and intuitive to use, while 81% agreed that the platform makes it easier for them to practice social distancing. 86.7% of respondents reported that their time has been saved, as they could access the platform for updates instead of physically visiting the venue and then making alternative plans due to unexpected high level of crowdedness.

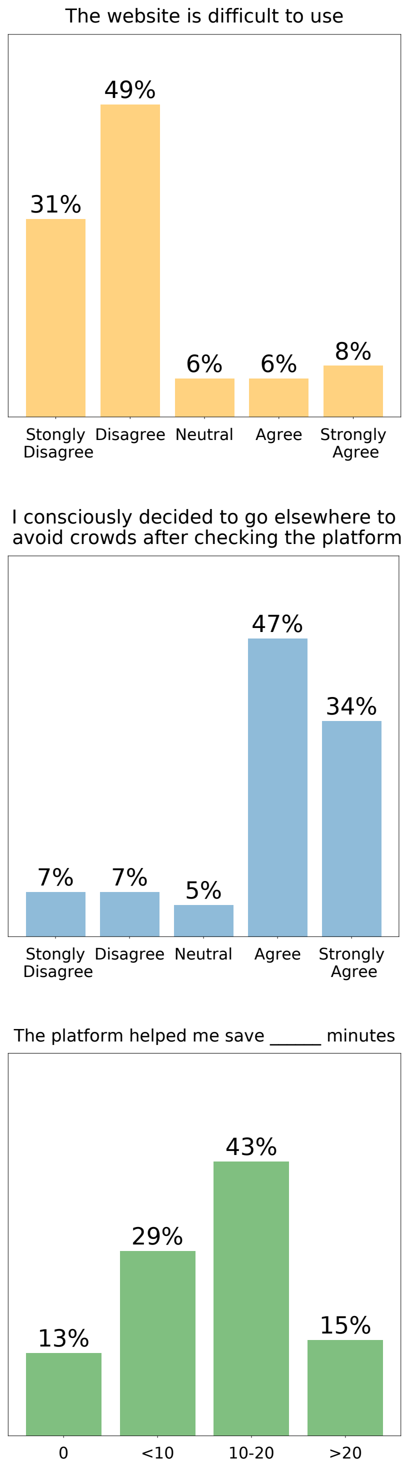


Figure 3: User study questions.

Conclusions and Future Work

crowded.sg is an effective solution to assist users in making better choices of where to visit to facilitate social distancing in times of pandemic. Although there are map applications (e.g., Google Maps) which rely on users' location beaconing information to estimate the crowdedness of an area. Such an approach does not capture non-users or users who disable the location function in their mobile devices. In comparison, *crowded.sg* is application agnostic and accessible to anyone with a smart phone.

In subsequent work, we will incorporate functions to forecast crowd levels and suggest optimal alternatives so as to minimize crowding.

Acknowledgements

This research is supported, in part, by Alibaba Group through Alibaba Innovative Research (AIR) Program and Alibaba-NTU Singapore Joint Research Institute (JRI) (Alibaba-NTU-AIR2019B1); Nanyang Technological University, Singapore, Nanyang Assistant Professorship (NAP); WASP/NTU grant M4082187 (4080), Singapore Ministry of Education (MOE) Tier 1 (RG16/20).

References

Bangor, A.; Kortum, P. T.; and Miller, J. T. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24(6): 574–594.

Gao, G.; Gao, J.; Liu, Q.; Wang, Q.; and Wang, Y. 2020. CNN-based Density Estimation and Crowd Counting: A Survey. *arXiv preprint arXiv:2003.12783* .

He, K.; Gkioxari, G.; Dollár, P.; and Girshick, R. 2017. Mask R-CNN. In *Proceedings of the IEEE international conference on computer vision*, 2961–2969.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, Y.; Zhang, X.; and Chen, D. 2018. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1091–1100.

Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; and Belongie, S. 2017. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2117–2125.

Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; and Zitnick, C. L. 2014. Microsoft coco: Common objects in context. In *European conference on computer vision*, 740–755. Springer.

Liu, W.; Luo, W.; Lian, D.; and Gao, S. 2018. Future Frame Prediction for Anomaly Detection – A New Baseline. In *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Loy, C. C.; Chen, K.; Gong, S.; and Xiang, T. 2013. Crowd counting and profiling: Methodology and evaluation. In

Modeling, simulation and visual analysis of crowds, 347–382. Springer.

Rothan, H. A.; and Byrareddy, S. N. 2020. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of autoimmunity* 102433.

Thaler, R. H.; and Sunstein, C. R. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.