



HAL
open science

An Efficient Method for Addressing COVID-19 Proximity Related Issues in Autonomous Shuttles Public Transportation

Dimitris Tsiktsiris, Antonios Lalas, Minas Dasygenis, Konstantinos Votis,
Dimitrios Tzouvaras

► **To cite this version:**

Dimitris Tsiktsiris, Antonios Lalas, Minas Dasygenis, Konstantinos Votis, Dimitrios Tzouvaras. An Efficient Method for Addressing COVID-19 Proximity Related Issues in Autonomous Shuttles Public Transportation. 18th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Jun 2022, Hersonissos, Greece. pp.170-179, 10.1007/978-3-031-08333-4_14. hal-04317175

HAL Id: hal-04317175

<https://inria.hal.science/hal-04317175v1>

Submitted on 1 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

An Efficient Method for Addressing COVID-19 Proximity Related Issues in Autonomous Shuttles Public Transportation*

Dimitris Tsiktisiris^{1,2}[0000-0001-6475-5865], Antonios Lalas^{1,*}[0000-0002-5337-161X], Minas Dasygenis²[0000-0002-2180-9752], Konstantinos Votis¹[0000-0001-6381-8326], and Dimitrios Tzovaras¹[0000-0001-6915-6722]

¹ Information Technologies Institute, Center for Research and Technology Hellas, 6th km Xarilaou - Thermi, Thessaloniki, 57001, Greece
`{tsiktisiris,lalas,kvotis,dimitrios.tzovaras}@iti.gr`

² Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani, 50100 Greece
`mdasyg@ieee.org`

Abstract. The COVID-19 pandemic has created significant restrictions to passenger mobility through public transportation. Several proximity rules have been applied to ensure sufficient distance between passengers and mitigate contamination. In conventional transportation, abiding by the rules can be ensured by the driver of the vehicle. However, this is not obvious in Autonomous Vehicles (AVs) public transportation systems, since there is no driver to monitor these special circumstances. Since, AVs constitute an emerging mobility infrastructure, it is obvious that creating a system that can provide a sense of safety to the passenger, when the driver is absent, is a challenging task. Several studies employ computer vision and deep learning techniques to increase safety in unsupervised environments. In this work, an image-based approach, supported by novel AI algorithms, is proposed as a service to increase the COVID-19 safety rules adherence of the passengers inside an autonomous shuttle. The proposed real-time service, can detect deviations from proximity rules and notify the authorized personnel, while it is possible to be further extended in other application domains, where automated proximity assessment is critical.

Keywords: Autonomous Vehicles · Public Transportation · Neural Networks · Image Processing

1 Introduction

Public transportation systems progressively adopt the concept of autonomous mobility as a service (MaaS) [2]. However, the transition to fully autonomous

* Supported by the European Union's Horizon 2020 Research and Innovation Programme Autonomous Vehicles to Evolve to a New Urban Experience (AVENUE) under Grant Agreement No 769033.

public vehicles is not seamless and several obstacles arise in the real-world. With the rapid and continuous spreading to this day, COVID-19 is without doubt one of these obstacles. Although urban travel has declined across the globe, public transportation is considered to be the sector that has gotten the hardest hit, as indicated by survey-based data [9]. The phenomenon is caused due to the sometimes-unavoidable closer contact that is observed between people using public transportation vehicles or stations, leading to the impression of public transportation being riskier than any other personal or private means of transport. In addition, the fact that an infected with COVID-19 passenger can contaminate many more before even showing any symptoms [7], [4] is especially troublesome for public spaces. Some of the factors that can lead to a high risk of contamination in public vehicles and stations are listed below:

- Limited space for passengers. The risk of contamination increases in proportion to the number of passengers that are confined in the limited space of the vehicle or the station. Since the global pandemic arised, the level of discomfort associated with public transportation has been increased considering the added risk of becoming infected by COVID-19.
- The inability to recognize passengers or workers who may be infected.
- The ease of transferring germs through multiple surfaces, like ticket machines, doors, handrails, or seats.

Nonetheless, there are ways that can mitigate the associated risk to the aforementioned factors. Furthermore, the level of the contamination risk one takes when traveling versus performing activities in public places, for example, is yet to be determined, as several variables determine these levels in different sets of environments. One of the most universally applied measure to prevent the transmission of COVID-19 is physical or social distancing, which started as a recommendation from the World Health Organization (WHO) so as to keep a distance from another individual of at least 1.0 meter apart [12]. Although other health organizations recommend keeping a physical proximity of two meters, the original suggestion of WHO has been found to significantly reduce the COVID-19 transmission [1]. Even though conflicting with the concept of public transportation, social distancing is considered, to this day, the most important non-pharmaceutical prevention measure for public transportation services, as physical distancing strongly reduces the level of occupancy in vehicles and stations that serve travellers [10].

Up until now, many technology-based solutions have been proposed across the globe to cope with the COVID-19 [8]. A lot of researchers are looking into the fields of computer vision and deep-learning in order to develop solutions able to reduce virus spread. New, but also existing state-of-the-art technologies are employed for this purpose, tailored in specific applications to reduce COVID-19 propagation.

In this paper, we present an image-based approach, supported by novel AI algorithms, as an end-to-end service to increase the COVID-19 safety rules adherence of the passengers inside an autonomous shuttle. The proposed real-time

service, can detect deviations from proximity rules and enable notifications to authorized personnel. Moreover, we implement optimizations tailored to the Autonomous Shuttles for reduced area and power consumption.

The main contributions of this work can be summarized as follows:

- We propose an end-to-end service based on deep learning, for automated distance assessment in Autonomous Shuttles.
- We deploy the service on embedded devices and introduce smart techniques such as adaptive inference to reduce the power consumption.
- We introduce a new dataset for passenger detection from overhead cameras tailored on the environment of the Autonomous Shuttles

2 Related Work

Various countries track the location data of the suspected and infected individuals using GPS technology. An assortment of different emerging technologies that have a prime role in social distancing scenarios, including Wireless networks, smart devices, Global Positioning System (GPS), computer vision, and deep learning, is proposed by Nguyen et al. [11]. Other studies exploit UAVs and camera sensors for detecting large crowds [5]. So far, several progress has been made in detecting the virus [6]. A study by Prem et al. [14] on the impacts of social distancing upon stopping the virus transmission concluded that it could slowly decrease the peak of the virus attack; however, social distancing remains an unpleasant step for many sectors of the economy. On the other hand, researchers provide effective solutions employing surveillance videos in addition to computer vision, machine learning, and deep learning-based approaches for measuring social distancing. A proposed framework by Punn et al. [15] employs firstly the YOLOv3 model for human detection and then the Deepsort [23] approach for tracking the detected individuals, using both assigned ID information and bounding boxes. As for the dataset, the authors used a frontal view dataset by an open image dataset repository and compared their results with faster-RCNN [18] and SSD [24]. Another approach by Ramadass et al. [16] implemented an autonomous drone-based model for social distance monitoring by using a custom training dataset for the YOLOv3 [17] algorithm, comprised of both front and side perspective frames of people. Opposed to the previous approach by Punn, Ramadass’s does monitor face masks too, as the use of a drone camera in conjunction with the YOLOv3 model helps in identifying the physical distance, as well as monitoring individuals with masks in public, from the frontal or side view. Sathyamoorthy et. al. [19] implemented their model for detecting people who do not obey social distancing restrictions in crowded situations, using an autonomous robot with an RGB-D sensor and a lidar in order to navigate among crowded places. Lastly, Pouw’s et. al. [13] proposed framework constitutes a monitoring approach appropriate for social distancing.

Based on the reviewed literature, we have concluded that, even though a considerable progress has been made for monitoring social distancing in public

environments, most of the work done is focused on the side or frontal view perspectives of the camera. Moreover, the aforementioned studies do not take into account the impact of power consumption, as energy efficiency is one of the most critical factors for the Autonomous Vehicles. In contrast, we provide an end-to-end solution as a service, tailored to the Autonomous Shuttle environment, with overhead camera perspective and low power consumption.

3 Methodology

Most of the aforementioned works focus on front and side view input for monitor proximity in public outdoor places. In this work, we present a deep learning-based distance assessment service using an overhead perspective, able to function with high accuracy and low power-consumption into the confined space of an autonomous shuttle. To overcome occlusion issues, we employ fisheye wide-angle cameras with a top-down perspective. In this section, we present the datasets used, followed by the network architecture. Finally, we describe the training process and the post-processing steps.

3.1 Dataset

We collected and labelled a new dataset named CERTH - AVenue Overhead Fisheye (C-AVOF). The new dataset contains frames and human objects in a simulated shuttle environment, and also includes challenging scenarios such as crowded rooms, various body poses, and various-light conditions. The camera used for the data capture process is the D-Link DCS-4625 at 1080p resolution output. During the annotation of this dataset, we used Deepsort [22] for tracking the individual passengers in the shuttle across multiple frames and thus our dataset can be also used for additional vision tasks using overhead, fish-eye images, such as video-object tracking and human re-identification. Also, for evaluation purposes, some cherry-picked samples from the BOSS dataset [21] were included, especially the scenarios from camera 5 and 7 with the top-down overhead perspective.

3.2 Network Architecture

Inspired by RAPiD [3], the passenger detection network consists of three stages: the backbone network, the feature pyramid network (FPN), and the bounding box regression network. The backbone network works as a feature extractor that takes an image as input and outputs a list of features from different parts of the network. In the next stage, we pass those features into the FPN, in order to extract features related to object detection. Finally, at the last stage, a Convolutional Neural Network (CNN) is applied to each feature vector in order to produce a transformed version of the bounding-box predictions (Figure 1).

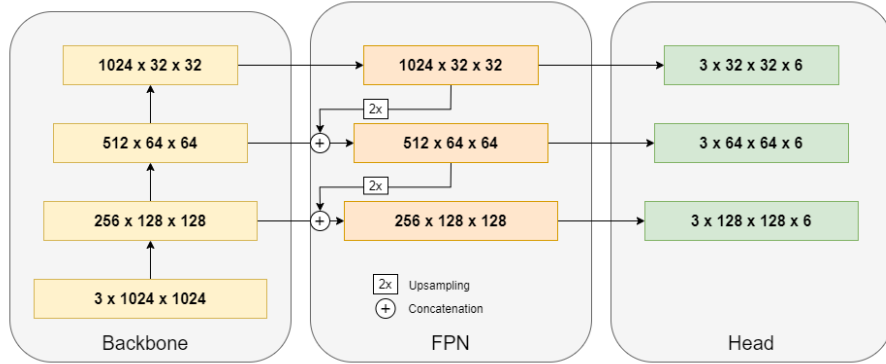


Fig. 1. An illustration of multiple convolutional layers and multi-dimensional matrices such as the feature maps with 1024x1024 input resolution.

3.3 Training

We used the pretrained RAPiD weights on MW-R and HABBOF which we fine-tuned by training on our custom dataset C-AVOF. Rotation, flipping, resizing, and color augmentation are used in the training stage.

3.4 Post-processing

After the detection process, we used the bounding box coordinates in order to compute the centroid. In order to calculate the distance between each bounding box, we used the Euclidean distance formula. Each bounding box is connected with the rest via distance line which represents the real-world distance, when multiplied with a weight factor. The weight value is calculated via the camera calibration process and takes into account various parameters such as the position of the camera and the field of view.

3.5 Experimental Results

The experiments were performed on the Nvidia Jetson AGX Xavier platform that features 512-core Volta GPU with Tensor Cores, 8-core ARM 64bit CPU and 32GB 256-Bit LPDDR4x RAM. The camera system is the D-Link DCS-4625 panoramic fisheye camera with f/2.0 wide-angle panoramic lens and 5MP (2560 x 1920 @30fps) high-resolution video. Results in Table 1 indicate that RAPiD at 608x608 resolution achieved the best performance and the fastest execution speed on our dataset C-AVOF. The averaged FPS value represents the execution speed on the Nvidia Jetson AGX Xavier.

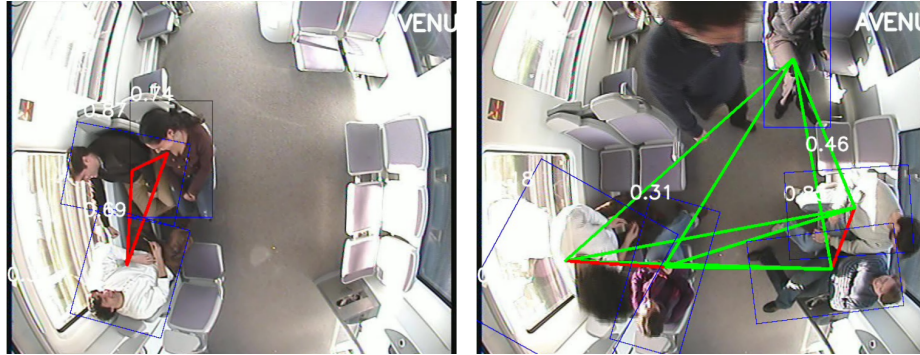
In Figure 2 and Figure 3, we can see the visualized results on unseen scenarios from the BOSS dataset. The blue bounding boxes indicate the detections and the

Table 1. Performance comparison of two state-of-the-art methods on our dataset C-AVOF.

	FPS	MW-R [3]	HABBOF [8]	CEPDOF [3]	C-AVOF (ours)
		AP_{50}	AP_{50}	AP_{50}	AP_{50}
Tamura <i>et al.</i> [20] (608)	5.8	77.4	86.1	59.2	77.6
RAPiD (608)	6.5	96.2	97.6	84.3	97.9

numbers represent the confidence of each detection. Red lines denote an unsafe distance while lines in green a sufficient distance.

Handling reflections is still a challenge on certain scenarios, as also happens in similar approaches. Passenger figures might appear to the windows of the shuttle as reflections, especially when the lighting is low. To mitigate this issue, a custom mask is applied on large reflective surfaces.

**Fig. 2.** Results on unseen scenarios from the BOSS dataset. Green lines represent a safe distance while red lines an unsafe.

3.6 Adaptive Inference

As the service is able to report real-time metrics over the network to the AV-ENUE dashboard, we exploit this bidirectional communication inside the shuttle ecosystem, to enhance its awareness. By receiving useful data about the vehicle’s state, such as the speed, the status of the door and the tires pressure, we can reduce the power consumption and be more energy efficient by decreasing the number of inferences, as Table 2 indicates. Figure 4 illustrates an overview of the dashboard displaying metrics for various services inside the autonomous shuttle ecosystem.



Fig. 3. Results on an unseen crowded scenario from the BOSS dataset, with overlapping detections

In-vehicle security

Event History

- Low distance between passengers in vehicle 2 22/10/2020 14:46:41
- Video modality detected Bag snatch in vehicle 1 10/22/2020 14:46:41
- Audio modality detected Scream in vehicle 1 10/22/2020 14:46:41
- Audio modality detected Scream in vehicle 1 10/22/2020 14:46:41

Automated Passenger Presence		Environmental	
Onboard passengers	6	Temperature	18C
Door status	Open	Humidity	62%
Distance assessment	Unsafe	CO2	360ppm

Fig. 4. An overview of the dashboard displaying metrics for various services inside the autonomous shuttle ecosystem.

Table 2. Estimated power consumption between adaptive modes measured on Jetson AGX Xavier using tegrastats utility.

Inference Frequency	Power Consumption (Wh)
Continuous (real-time)	26.2
Adaptive (crowded-multiple stops)	14.7
Adaptive (average)	11.4

4 Conclusions

In this paper, we present a deep learning approach deployed as a real-time service that can detect deviations from proximity and enable notifications to authorized personnel. As a backend, we are using a pre-trained RAPiD [3] model for human detection and perform fine tuning on our C-AVOF dataset to improve the performance. The model outputs bounding box coordinates that we are using to compute their centroids. By applying the Euclidean distance formula, we are able to compute the pairwise distances for each bounding box. In order to detect proximity violations between passengers, we transform the pixel distance into the real distance by multiplying with a weight value that is defined via camera calibration. Experimental results indicated that the service efficiently identifies passengers with unsafe proximity. The end-to-end service was deployed on the Jetson AGX Xavier with DCS-4625 fisheye camera and the results were visualised via the AVENUE platform. The overhead camera perspective mitigates occlusion issues leading to more robust detections compared to similar approaches with different perspectives. Moreover, as energy efficiency is a critical factor in Autonomous Vehicles, smart techniques based on data deriving from the shuttle ecosystem enabled us to further reduce power consumption by 56%. Finally, we introduce a new dataset for passenger detection from overhead camera perspective. In the future, experiments with an attention-based approach will be conducted to further enhance the accuracy of the model. In addition, an extended version of this approach could be applied in order to calculate the occupancy and weight distribution of the passengers in the shuttle. We believe that our research will be a stepping stone towards increased security in the autonomous public transport and will contribute positively in trust of the passengers to a human driverless transport system during the pandemic, but also beyond with further application-tailored adaptations.

References

1. Chu, H.Y., Englund, J.A., Starita, L.M., Famulare, M., Brandstetter, E., Nickerson, D.A., Rieder, M.J., Adler, A., Lacombe, K., Kim, A.E., et al.: Early detection of covid-19 through a citywide pandemic surveillance platform. *New England Journal of Medicine* **383**(2), 185–187 (2020)

2. Cruz, C.O., Sarmiento, J.M.: “mobility as a service” platforms: A critical path towards increasing the sustainability of transportation systems. *Sustainability* **12**(16), 6368 (2020)
3. Duan, Z., Tezcan, O., Nakamura, H., Ishwar, P., Konrad, J.: Rapid: rotation-aware people detection in overhead fisheye images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. pp. 636–637 (2020)
4. Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., Fraser, C.: Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**(6491), eabb6936 (2020)
5. Harvey, A., LaPlace, J.: Megapixels: Origins, ethics, and privacy implications of publicly available face recognition image datasets. *Megapixels* **1**(2), 6 (2019)
6. Iqbal, M.S., Ahmad, I., Bin, L., Khan, S., Rodrigues, J.J.: Deep learning recognition of diseased and normal cell representation. *Transactions on Emerging Telecommunications Technologies* **32**(7), e4017 (2021)
7. Javid, B., Weekes, M.P., Matheson, N.J.: Covid-19: should the public wear face masks? (2020)
8. Li, S., Tezcan, M.O., Ishwar, P., Konrad, J.: Supervised people counting using an overhead fisheye camera. In: *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. pp. 1–8. IEEE (2019)
9. Molloy, E.J., Bearer, C.F.: Covid-19 in children and altered inflammatory responses (2020)
10. Musselwhite, C., Avineri, E., Susilo, Y.: Editorial jth 16–the coronavirus disease covid-19 and implications for transport and health. *Journal of transport & health* **16**, 100853 (2020)
11. Nguyen, T.T., Nguyen, Q.V.H., Nguyen, D.T., Hsu, E.B., Yang, S., Eklund, P.: Artificial intelligence in the battle against coronavirus (covid-19): a survey and future research directions. *arXiv preprint arXiv:2008.07343* (2020)
12. Olivera-La Rosa, A., Chuquichambi, E.G., Ingram, G.P.: Keep your (social) distance: Pathogen concerns and social perception in the time of covid-19. *Personality and Individual Differences* **166**, 110200 (2020)
13. Pouw, C.A., Toschi, F., van Schadewijk, F., Corbetta, A.: Monitoring physical distancing for crowd management: Real-time trajectory and group analysis. *PloS one* **15**(10), e0240963 (2020)
14. Prem, K., Liu, Y., Russell, T.W., Kucharski, A.J., Eggo, R.M., Davies, N., Flasche, S., Clifford, S., Pearson, C.A., Munday, J.D., et al.: The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study. *The Lancet Public Health* **5**(5), e261–e270 (2020)
15. Punn, N.S., Sonbhadra, S.K., Agarwal, S.: Covid-19 epidemic analysis using machine learning and deep learning algorithms. *MedRxiv* (2020)
16. Ramadass, L., Arunachalam, S., Sagayasree, Z.: Applying deep learning algorithm to maintain social distance in public place through drone technology. *International Journal of Pervasive Computing and Communications* (2020)
17. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018)
18. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* **28** (2015)
19. Sathyamoorthy, A.J., Patel, U., Savle, Y.A., Paul, M., Manocha, D.: Covid-robot: Monitoring social distancing constraints in crowded scenarios. *arXiv preprint arXiv:2008.06585* (2020)

20. Tamura, M., Horiguchi, S., Murakami, T.: Omnidirectional pedestrian detection by rotation invariant training. In: 2019 IEEE winter conference on applications of computer vision (WACV). pp. 1989–1998. IEEE (2019)
21. Velastin, S.A., Gómez-Lira, D.A.: People detection and pose classification inside a moving train using computer vision. In: International Visual Informatics Conference. pp. 319–330. Springer (2017)
22. Wojke, N., Bewley, A.: Deep cosine metric learning for person re-identification. In: 2018 IEEE winter conference on applications of computer vision (WACV). pp. 748–756. IEEE (2018)
23. Wojke, N., Bewley, A., Paulus, D.: Simple online and realtime tracking with a deep association metric. In: 2017 IEEE international conference on image processing (ICIP). pp. 3645–3649. IEEE (2017)
24. Zhang, S., Wen, L., Bian, X., Lei, Z., Li, S.Z.: Single-shot refinement neural network for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4203–4212 (2018)