Health Greeter Kiosk: Tech-Enabled Signage to Encourage Face Mask Use and Social Distancing

Max Hudnell
UNC Chapel Hill
Chapel Hill, North Carolina, USA

Steven King
UNC Chapel Hill
Chapel Hill, North Carolina, USA

(a) (b)

Figure 1: (a) The Health Greeter Kiosk deployed at an entrance to a football stadium. (b) Kiosk front and back view.

ABSTRACT

COVID-19 has been the cause of a global health crisis over the last year. High transmission rates of the virus threaten to cause a wave of infections which have the potential to overwhelm hospitals, leaving infected individuals without treatment. The World Health Organization (WHO) endorses two primary preventative measures for reducing transmission rates: the usage of face masks and adherence to social distancing [World Health Organization 2021]. In order to increase population adherence to these measures, we designed the Health Greeter Kiosk: a form of *digital signage*. Traditional physical signage has been used throughout the pandemic to enforce COVID-19 mandates, but lack population engagement and can easily go unnoticed. We designed this kiosk with the intent to reinforce these COVID-19 prevention mandates while also considering the necessity of population engagement. Our kiosk encourages engagement by providing visual feedback which is based on analysis from our kiosk's computer vision software. This software integrates *real-time* face mask and social distance detection on a low-budget computer, without the need of a GPU. Our kiosk also collects statistics, relevant to the WHO mandates, which can be used to develop well-informed reopening strategies.

SIGGRAPH '21 Emerging Technologies , August 9–13, 2021, Virtual Event, USA

CCS CONCEPTS

• **Computing methodologies** → *Computer vision*; • **Human-centered computing** → **Interaction design process and methods**.

KEYWORDS

COVID-19, digital signage, face mask detection, social distance detection, real-time, edge device

ACM Reference Format:
Max Hudnell and Steven King. 2021. Health Greeter Kiosk: Tech-Enabled Signage to Encourage Face Mask Use and Social Distancing. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Emerging Technologies (SIGGRAPH '21 Emerging Technologies), August 9–13, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3450550.3465339

1 INTRODUCTION

Quickly after the COVID-19 outbreak, governments and corporations responded by closing businesses, parks, airlines and more, in order to prevent in-person interaction and potential spread of the disease. In situations where in-person interactions must take place, the WHO recommends face mask use and social distancing as two primary measures which should be taken to minimize transmission.

Wearing face masks has been proven to reduce the risk the of spreading, and contracting, the virus [Bartoszko et al. 2020; Chu et al. 2020; MacIntyre and Chughtai 2020]. Maintaining at least one meter between yourself and others has been proven to lower transmission rates of airborne viruses, including COVID-19 [Chu et al. 2020; Keskinocak et al. 2020].

It is essential to persuade the public to follow these preventative mandates in order to prevent the spread of disease during a pandemic. Thus it is beneficial to research how to effectively encourage these practices. A primary tool for encouraging behavioral adjustment is the usage of signage [Meis and Kashima 2017]. There is copious COVID-19 related physical signage in public spaces, but they often fail to hold one's attention. A primary goal of this work is to develop more engaging signage which effectively persuades individuals to comply with the preventative mandates. We aim to do this by providing passerby with **engaging visual feedback** based on their detected face mask and social distancing adherence.

In addition, we set the goals of:

- **Developing a statistic collection system** to inform facility-owners during the reopening process.
- **Maintaining a low-budget**. This would allow the production of more kiosks, thus permitting a wider cultural impact and greater mandate adherence.

In this work we describe the software which enables this, which consists of a fully fledged face mask detection system capable of tracking detections as well as a novel implementation of a real-time social distance estimation system. Our software runs at an efficient ~26 fps on a constrained, low-budget system with no GPU.

2 RELATED WORK

2.1 Face Mask Detection

The vast majority of face mask detection techniques are based on supervised learning methods. They differ primarily in their selection of network architecture. [AIZOOTech 2020] uses an SSD structure. [Loey et al. 2021b] uses Resnet50 to perform feature extraction and then an SVM for face mask classification. [Loey et al. 2021a] improves upon this by using YOLOv2 as a classifier instead of an SVM.

We use the model of [AIZOOTech 2020] due to the SSD's run-time efficiency.

2.2 Social Distance Estimation

[Cristani et al. 2020] state that social distance estimation requires solving two distinct tasks, namely, **scene geometry understanding** and **person detection / body pose estimation**. We discuss related work based on their approach to each task.

2.2.1 Scene Geometry Understanding. Understanding scene geometry requires defining a **metric reference** which can be used to establish a distance measure within a scene [Aghaei et al. 2021]. This reference can be established via manual camera calibration, yet we elect not to do this as it requires an extra step during setup, and the camera runs the risk of becoming uncalibrated if disturbed.

It is possible to define a metric reference from a monocular RGB image without the need for calibration. This is commonly done by making assumptions about the size of objects detected in an image. In particular, with social distancing methods, assumptions are commonly made about the average human height [Bhambani et al. 2020; Ghodgaonkar et al. 2020], width [Gupta et al. 2020], and body part size [Aghaei et al. 2021]. These methods also tend to assume detected individuals are standing upright [Gupta et al. 2020]. These assumptions all cause error in relatively common cases, like when presented with children or with a person who is sitting down. For this reason, we elect to obtain a metric reference directly via a pre-calibrated *depth sensor*.

Captured depth provides an inherent 3D understanding of the scene without needing to estimate the ground plane or make any assumptions based on average sizes. The primary downside of using a depth sensor is that the affordable depth sensors can have relatively limited range. Despite this, we found the maximum range of the Intel D415 sensor (10m) suitable, as we do not expect individuals to view the kiosk from further than 10 meters.

2.2.2 Person Detection / Body Pose Estimation. Social distance estimation solutions often perform person detection via bounding box predictors, and use the box's centroid for their positional estimation [Ahmed et al. 2021; Gupta et al. 2020]. This can result in error in the case when an individual is not centered within the detected bounding box. More recent works [Aghaei et al. 2021] have favored the approach of using human pose estimation, as they provide a more detailed localization of detected persons.

In recent years there has been great progress in the run-time efficiency of body pose estimation networks, and they are now capable of running in real-time on edge devices [Osokin 2018], in complex scenarios such as large crowds with lots of occlusion [Cao et al. 2019]. Thus, we perform body pose estimation in our implementation as it fits our real-time performance requirement.

2.3 Deployable Kiosks

A number of COVID-oriented commercial kiosks have been released in the last year. Multiple sign-in kiosks [Digital Touch Systems 2021; Janam 2021] have been developed which test for proper face mask use before allowing individuals to enter an area, however these do not perform social distance detection.

[Indigo Vision 2021] provide an entrance flow management systems which performs face mask detection and encourages social distancing with a 'stoplight' system, but does not perform social distance detection.

[Amazon 2020] has a social distance detection system most similar to ours, implemented with an RGB-D camera and a mirrored visual feedback system. However their model requires a high-performance GPU (NVIDIA GTX 1070+), which significantly increases cost compared to our CPU-only method. Additionally, they do *not* perform mask detection.

As far as we are aware, we are the *first* to produce a form of digital signage which encourages both face mask use and social distancing via real-time visual feedback. In addition, we believe we are also the first to publicly describe the full implementation of a kiosk system with face mask / social distance detection capabilities.

3 METHOD

3.1 Design

In Figure 1b, we show a full portrait image of the physical kiosk. In Figure 2, we show an example of the kiosk display while in use.

The primary item on the screen is an RGB video feed, which streams directly from the RGB-D camera mounted directly on top of the kiosk. This video feed acts as a digital mirror of whomever stands in front of the kiosk. Our motivation behind a mirrored design comes in part from a psychology study by [Beaman et al.

Figure 2: An example of the display while in use (the mannequin is detected as a person by the body pose estimator).

1979], which shows subjects tend to act more responsibly when able to view themselves in a mirror, due to an increase in self-awareness. We aim to increase adherence to the face mask and social distance mandates through this increase in self-awareness.

Overlaid on the video feed we display bounding boxes of the detected faces with labels adjacent which either read "Mask detected" or "Mask needed".

A circle indicator is also displayed over the chest of each person detected. This indicator glows green or red depending on whether a social distance violation is detected.

In order to provide assurance over potential privacy concerns, text is displayed at the bottom which reads: "Video is not recorded or transmitted".

3.2 Face Mask Detection

To perform face and face mask detection, we utilize the open-source model *FaceMaskDetection* developed by [AIZOOTech 2020]. In their implementation they utilize a Single-Shot Detector (SSD) structure with an input size of 260x260 and a total of 24 layers. The model is trained on a total of 7959 images gathered from the WIDER Face [Yang et al. 2016] and MAFA [Ge et al. 2017] datasets.

Tracking The AIZOOTech model outputs detections on a per-frame basis. In order to collect stats on a per-individual basis, we must track and associate detections between frames. To acheive this, we utilize the SORT tracking algorithm [Bewley et al. 2016]. We also use SORT to perform track smoothing, before drawing our graphic overlay, in order to provide better visual feedback.

3.3 Social Distance Detection

We implement social distance estimation via a simple three step procedure:

Acquire 2D body keypoints The first step is to perform human pose estimation, i.e., to estimate 2D keypoints of the detected bodies in image space. To do this, we use Lightweight OpenPose [Osokin 2018], an efficient adaptation of the popular OpenPose algorithm [Cao et al. 2019], capable of running in real-time.

Deproject 2D keypoints to 3D Because we use an RGB-*D* camera with known intrinsic parameters, we can deproject the 2D keypoints to 3D directly using the captured *depth* information. We use the function `rs2_deproject_pixel_to_point`, provided within the Intel Realsense SDK [Intel 2020b], to perform this deprojection.

Estimate social distances Now that each detected person has a 3D coordinate associated with them, we can simply compute the Euclidean distance between them. After associating a 3D coordinate with each detected person, we can simply compute the Euclidean distance between them. For each detected person, the distance to their closest neighbor is stored as their 'social distance'.

3.4 Stat Collection

3.4.1 Face Mask Usage. We collect two stats relevant to face mask usage:

(1) the count of individuals detected wearing masks
(2) the count of individuals detected not wearing masks

In order to compute these statistics, a `mask_status` variable is computed for each track every frame. This variable is computed via mean filtering over a given track's last n confidence scores in order to reduce noise in the model's confidence scores. The score is then thresholded to categorize it as `mask_needed` or `mask_detected`. The final `mask_status` value computed before track loss is used as the track's collected statistic.

3.4.2 Social Distance Adherence. We collect three statistics relevant to social distancing:

(1) The count of total people detected
(2) The count of total social distance violations which occur (we use the WHO guideline of 1m as the violation threshold)
(3) The average 'social distance' maintained by each individual

The first two values are trivial to compute. The third value, average social distance (*ASD*), we maintain as a cumulative moving average which is updated each frame via the step:

$$ASD_{n+1} = ASD_n + \frac{d_{n+1} - ASD_n}{n+1}$$

where d is the social distance value for the current frame and n is the number of frames contributing to the average so far.

Although we compute this value per-individual, the statistic can be aggregated over all individuals over any given period of time, in order to give a more general statistic.

3.5 Implementation

Computer vision application This component is written in C++, for it's efficiency, and it implements all of the image processing required for face mask and social distance detection. We use the Intel OpenVINO Inference Engine to perform model inference for both the face mask detection and human pose estimation networks on the CPU [Intel 2020a]. Optionally, inference of the mask detection network can be offloaded to an Intel Neural Compute Stick 2 for a ~2 fps increase. The inference output is passed on to post-processing code which performs tracking, stat collection, and draws 2D graphic indicators on the color image. The processed image is then sent to the local webserver application over local sockets. Numerical statistics are optionally sent to a remote server via HTTP protocol. Our computer vision component achieves a steady fps of ~26 fps.

Local webserver We use the Flask web framework to host a local webserver. This webserver is responsible for receiving video frames from the C++ application and streaming it to the browser. This local server also hosts an HTML webpage, which acts as our digital sign and is displayed on the kiosk screen.

Hardware We use a Lenovo ThinkCentre M90n 'nano computer', equipped with an Intel Core i5-8265U CPU. The OS used is Ubuntu 18.04. We use the Intel Realsense D415 camera to capture RGB-D video.

4 DISCUSSION

We deployed the Health Greeter Kiosks at multiple UNC Chapel Hill football games during the 2020 season. Six kiosks were set up around the entrance gates of the stadium, where attendees tend to line up. Figure 1a shows one of these entrances. The kiosks were highly noticeable in these locations and many individuals were engaged in the responsive display. We note that in general the kiosks received a positive response from attendees. John Brunner, Associate Athletic Director for Event Management at UNC, stated: "Our staff and fans were really responsive to it, it was [apparent] that there was additional compliance because of these efforts, so it was certainly effective." We leave a more official proof of the kiosk's effectiveness to future work.

5 CONCLUSION

In this paper, we present a form of digital signage built with the intent to effectively persuade individuals to comply with face mask and social distancing mandates. The Health Greeter Kiosk also implements a statistic collection system which can inform facility-owners and enable safe decision-making during the reopening process. We describe the details of our software implementation, which is capable of running in real-time on a low-budget, resource constrained system.

For future work, we propose a potential psychology study which measures the effectiveness of the Health Greeter Kiosk versus traditional physical COVID-19 mandate signage. An ablation study could be performed to measure the effectiveness of the 'mirror image' style of visual feedback versus other visual feedback methods, or even auditory feedback.

ACKNOWLEDGMENTS

We would like to thank Robert Daigle with the Lenovo AI Innovation Center and Monica Livingston with Intel for sponsorship and support of this project.

REFERENCES

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.

Imran Ahmed, Misbah Ahmad, Joel JPC Rodrigues, Gwanggil Jeon, and Sadia Din. 2021. A deep learning-based social distance monitoring framework for COVID-19. *Sustainable Cities and Society* 65 (2021), 102571.

AIZOOTech. 2020. FaceMaskDetection. https://github.com/AIZOOTech/FaceMaskDetection.

Amazon. 2020. Distance Assistant. https://www.aboutamazon.com/news/operations/amazon-introduces-distance-assistant.

Jessica J Bartoszko, Mohammed Abdul Malik Farooqi, Waleed Alhazzani, and Mark Loeb. 2020. Medical masks vs N95 respirators for preventing COVID-19 in healthcare workers: A systematic review and meta-analysis of randomized trials. *Influenza and Other Respiratory Viruses* 14, 4 (2020), 365–373.

Arthur L Beaman, Bonnel Klentz, Edward Diener, and Soren Svanum. 1979. Self-awareness and transgression in children: Two field studies. *Journal of Personality and Social Psychology* 37, 10 (1979), 1835.

Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. 2016. Simple Online and Realtime Tracking. In *IEEE International Conference on Image Processing (ICIP)*. IEEE, 3464–3468.

Krisha Bhambani, Tanmay Jain, and Kavita A Sultanpure. 2020. Real-time Face Mask and Social Distancing Violation Detection System using YOLO. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*. IEEE, 1–6.

Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2019. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 1 (2019), 172–186.

Derek K Chu, Elie A Akl, Stephanie Duda, Karla Solo, Sally Yaacoub, Holger J Schünemann, Amena El-harakeh, Antonio Bognanni, Tamara Lotfi, Mark Loeb, et al. 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* 395, 10242 (2020), 1973–1987.

Marco Cristani, Alessio Del Bue, Vittorio Murino, Francesco Setti, and Alessandro Vinciarelli. 2020. The Visual Social Distancing Problem. *IEEE Access* 8 (2020), 126876–126886.

Digital Touch Systems. 2021. COVID-19 Kiosk. https://digitaltouchsystems.com/covid-19-screening-kiosk/.

Shiming Ge, Jia Li, Qiting Ye, and Zhao Luo. 2017. Detecting Masked Faces in the Wild with LLE-CNNs. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2682–2690.

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).

Savyasachi Gupta, Rudraksh Kapil, Goutham Kanahasabai, Shreyas Srinivas Joshi, and Aniruddha Srinivas Joshi. 2020. SD-Measure: A Social Distancing Detector. In *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 306–311.

Indigo Vision. 2021. IndigoKiosk AI. https://www.indigovision.com/products/covid-19-technologies/indigokiosk-ai/.

Intel. 2020a. *Intel® Distribution of OpenVINO™ Toolkit*. https://software.intel.com/content/www/us/en/develop/tools/openvino-toolkit.html

Intel. 2020b. *Intel® RealSense™ SDK 2.0*. https://www.intelrealsense.com/developers/

Janam. 2021. GT2. https://www.janam.com/products/details/gt2.

Pinar Keskinocak, Buse Eylul Oruc, Arden Baxter, John Asplund, and Nicoleta Serban. 2020. The impact of social distancing on COVID19 spread: State of Georgia case study. *PLOS One* 15, 10 (2020), e0239798.

Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N Taha, and Nour Eldeen M Khalifa. 2021a. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustainable Cities and Society* 65 (2021), 102600.

Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N Taha, and Nour Eldeen M Khalifa. 2021b. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement* 167 (2021), 108288.

C Raina MacIntyre and Abrar Ahmad Chughtai. 2020. A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients. *International Journal of Nursing Studies* 108 (2020), 103629.

Julia Meis and Yoshihisa Kashima. 2017. Signage as a tool for behavioral change: Direct and indirect routes to understanding the meaning of a sign. *PLOS One* 12, 8 (2017), e0182975.

Daniil Osokin. 2018. Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose. *arXiv preprint arXiv:1811.12004* (2018).

World Health Organization. 2021. Coronavirus disease (COVID-19) advice for the public. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.

Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. 2016. WIDER FACE: A Face Detection Benchmark. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.