Mask Recognition in the Covered Safe Entry Scanner

Henry Dietz; University of Kentucky; Lexington, Kentucky

Abstract

SARS-CoV-2 is a highly contagious, airborne-transmission, virus that can be spread by people who do not have obvious symptoms. In 2020, that combination of features forced much of the world to impose a wide variety of forms of social distancing, ranging from simple recommendations restricting how shared spaces can be used to rigidly enforced quarantines. It is unclear how much distancing is enough, but it is clear that the economic and emotional costs of distancing are high. Fortunately, consistent use of simple face masks dramatically reduces the probability of others becoming infected. The catch is that a significant fraction of the US population either is refusing to wear masks or is wearing masks in ways that render them ineffective. For example, it is problematic for a shop owner to prevent potential customers who are not properly masked from entering their store. Thus, we have created the Covered Safe Entry Scanner - an open source system that uses image processing methods to automatically check for proper use of masks and potentially deny entry to those who do not comply. This paper describes the design, algorithms, and performance of the mask recognition system.

Introduction

On July 9, 2020, Andy Beshear, Governor of the Commonwealth of Kentucky, issued Executive Order 2020-586 requiring that face coverings be worn in many public places in the hope of slowing the spread of COVID-19. Later executive orders extended the circumstances in which masks were required. However, no enforcement process nor penalties for non-compliance were specified. In trips to the local supermarket and home supply store, the author noted that not only were a large fraction of the people not wearing masks, but a comparable number were wearing masks in ways that would render them ineffective. Even store employees often were not in full compliance. It was in this context that the Covered Safe Entry Scanner, the topic of this paper, was born.

Severity of the Pandemic

The COVID-19 pandemic is caused by the Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2), a positivesense single-stranded RNA virus. There are many coronaviruses (CoVs) which can infect humans and animals, causing illnesses ranging from the common cold to much more serious diseases. Both the Severe Acute Respiratory Syndrome (SARS) virus detected in 2003 and the Middle East Respiratory Syndrome (MERS) virus identified in 2012 are coronaviruses that have high mortality rates for infected humans, as does SARS-CoV-2.

According to the WHO[1], the origin of the SARS-CoV-2 is not yet certain, but it is very closely related to coronaviruses naturally occurring in Rhinolophus bat populations. Similarity of the virus in samples taken from humans is high enough to suggest there was a single point of introduction to humans – probably around Wuhan, China in the last quarter of 2019. A large

fraction of the earliest COVID-19 cases were linked to the Huanan Wholesale Seafood Market in Wuhan City, which probably played some role in transmitting the virus between species and/or initially amplifying the human outbreak. However, not only did the virus quickly spread beyond Wuhan through humans, but it is also possible that the virus could be re-introduced to humans from animal hosts not only in Wuhan, but perhaps anywhere the bats are found, which spans Asia, Africa, the Middle East, and Europe.

The severity of the COVID-19 disease is disturbingly clear: it is at least an order of magnitude worse than influenza. The CDC COVID-19 site[2] gives many details. As of January 29, 2021, the United States had 25,615,268 confirmed total cases and 431,619 deaths. To put this in perspective, in less than a year from the first United States COVID-19 diagnosis, fatalities exceeded the country's total deaths caused by five years of World War II. Currently, over 1.5% of those becoming sick with COVID-19 die. When patient care was less effective largely due to shortages, in April 2020, fatality rate exceeded 7%. As much as 10-15% of COVID-19 cases are severe. For those who do recover, initial recovery can take as long as 2-6 weeks. The CDC also lists a variety of lesscommon long-term effects involving cardiovascular, respiratory, renal, dermatologic, neurological, and psychiatric complications.

However, the point of the current work is not that the pandemic is a terrible thing, but that it is critical to stop spread of the virus to the other 300 million people in the United States.

Spread of SARS-CoV-2

Unfortunately, the SAR-CoV-2 virus is relatively easily transmitted from one infected person to another. There are three primary modes of transmission:

Contact transmission. The virus can remain viable on contaminated surfaces for periods ranging from hours to days, depending on the surface material and environmental conditions such as ambient temperature. Touching a contaminated surface and then touching your mouth, nose, or possibly eyes can transfer virus into your body potentially causing an infection. Fortunately, the virus does not survive simple cleaning procedures such as washing with soap and water or use of high-alcohol-content hand sanitizer. Thus, contact transmission seems to be the most easily controlled method by which the virus spreads.

Droplet transmission[3]. The virus can become temporarily airborne traveling in droplets of saliva or mucus, such as are expelled when you sneeze, cough, sing, talk, or breath heavily. These droplets can contaminate surfaces, and can directly enter the body through the nose, mouth, or even eyes. Because droplets are relatively large, commonly greater than 5 microns, they typically do not spread much beyond 6 feet before settling out of the air, and this is the main reason for the common rule suggesting social distancing by at least 6 feet. Most COVID-19 cases probably are the result of droplet transmission.

Airborne transmission[3]. Some viruses can survive becoming airborne in an aerosolized form on smaller droplets or dust particles. Aerosolized droplets are emitted mostly from the mouth, but also from the nose. These 2 micron or smaller particles can remain airborne for much longer periods, from minutes to hours, allowing the concentration of particles to increase over time in an enclosed space or for the airborne particles to travel larger distances. The inhalation of these particles also allows them to be carried deeper into the respiratory system, where the virus has a higher probability of causing infection. Fortunately, although there are well-documentated cases of people becoming sick with COVID-19 from passing through areas where an infectious person had been as long as hours ago, or over distances much greater than 6 feet, this does not seem to be common if there is adequate ventilation or air filtration. It is useful to note that HEPA filters in air cleaners typically are designed to remove particles down to 0.3 microns – which is bigger than the virus, but probably smaller than most droplets that carry the virus.

Stopping the spread

The extreme social distancing of a total lock-down could end the pandemic in a matter of weeks, and may be an answer in some countries, but it is not feasible in the United States. In fact, at this writing, Governor Beshear is facing an impeachment petition charging that his actions to limit spread of COVID-19 had violated citizen's rights in limiting certain public gatherings, outof-state travel, and tenant evictions.

Vaccines can be an effective way to reach pandemic-ending "herd immunity." As of January 2021, at least 60 COVID-19 vaccines have reached clinical development and several are in widespread distribution. However, it will take months, perhaps even another year, to get a sufficient fraction of the population vaccinated. It is also significant that vaccines are designed to prevent the illness – they do not make it impossible for an immune person to spread the virus, for example, through contact transmission.

If obvious symptoms were associated with being infectious with COVID-19, it would be relatively easy to quarantine people before they spread the virus. Unfortunately, one can be infectious while showing no symptoms. In fact, even when symptoms do appear, they can be mild and easily mistaken for other issues; for example, many governments have mandated temperature checks, but a fever is neither necessary nor sufficient to indicate that one is infectious with SARS-CoV-2. COVID-19 also has a relatively long incubation time, ranging from 3 to 14 days, making contact tracing more difficult and less effective. Most attempts to improve contact tracing are highly invasive, and personal and medical data are often potentially exposed; in fact, various HIPPA protections were waived effective March 15, 2020[4]. In any case, contact tracing alone is insufficient.

Properly used, masks can allow relatively normal interactions. It takes a mask designed to filter aerosols, such as an N95, to protect the wearer from inhaling airborne virus – and that mask must be tightly sealed against the face. In fact, the seal must be nearly perfect, which is not even possible if, for example, facial hair interferes with the seal. The effectiveness of different types of masks varies significantly, but **almost any mask that seals reasonably well over the nose and mouth will dramatically reduce the droplet quantity escaping and average distance trav-** **eled**. It is useful to understand that good mask materials are not regular meshes with holes small enough to block particles. Good masks are made with complex fibrous structures that force airflow through relatively large paths that twist and turn, so that droplets get trapped as they fail to make the tight maneuvers and collide with the material.

In summary, the goals of the Covered Safe Entry Scanner research project were to:

- Provide a low-cost (sub-\$100) automated system that can operate unattended 24/7 to confirm proper use of a mask
- Allow that system to optionally take non-contact temperature measurements
- Allow that system to implement a form of contact tracing that is inherently non-invasive and secure, by literally having no personally-identifiable data in the database

Related work

At the time this work was initiated, in Summer 2020, literature and patent searches mostly found either research attempting to recognize the person behind a mask or systems designed to confirm various types of protective gear were being worn in hazardous workplaces. In response to the pandemic, many systems have since appeared.

Our problem is unusual in that minimizing system cost was a primary goal. We also wanted the unit to work without a continuous internet connection. Thus, we targeted the cheapest available Android tablet that had front and rear cameras and could be powered 24/7 while allowing USB connection of a thermal imager. Our selection was an off-brand \$65 7" tablet running Android 9 Pie on a quad-core 1.5GHz processor with 2GB of RAM: a very limited computing platform.

Unlike Covered, which seeks to check one aligned face at a time, most pandemic-inspired approaches to recognition of masked faces involve deep learning to create recognizers that can work directly on CCTV and similar surveillance video streams.

SSDMNV2[5] uses the Single Shot Multibox Detector as a face detector and MobilenetV2 to distinguish masked vs. unmasked faces with 92.64% accuracy. However, it uses significantly more substantial computing resources than our target prodives, including a relatively fast processor and GPU. Masked AI[6] also uses deep learning to recognize humans with or without various types of masks (N95, surgical, and cloth-based) in CCTV or drone feeds of crowds. It offloads the relatively expensive processing, e.g., to Azure cloud service. There are many systems with the goal of spotting unmasked individuals in surveillancelike images[7][8].

Those systems do not recognize improper mask use. A YOLOv5-based system[9] discusses distinguishing mask_weared_incorrect, but in the end was not trained for that because too few of the test images were so marked. Facemasknet[10] is a deep learning network that claims 98.6% accuracy for identifying face mask use in CCTV footage, and it does distinguish improperly-worn masks from no mask or a properly-worn mask. However, the method used seems far too computationally heavy for our target platform.

At the finalizing of the current paper for publication, it seems that the combination of requiring face alignment, distinguishing properly used masks from mere presence of a mask, and the abil-



ity to run on a minimal self-contained platform are still unique. The optional low-resolution thermal imager (which is largely viable only because of the face alignment), and contact tracing support without storing personally-identifiable data, also appear to be unique.

Recognizing proper use of a mask

Although the most desirable execution platform for this system would be a sub-\$100 Android tablet, development is easier in a full Linux environment, so initial development and testing was done using a Atom-based Asus Eee PC – an inexpensive notebook computer with hardware performance comparable to an Android tablet. Prototyping was done using C++ code with the OpenCV library, which is easily portable to many platforms.

The basic algorithm is:

- 1. Wait for a person to be detected; this can be as simple as looking for movement
- 2. Show the live camera view and instruct the user to align their eyes with alignment outlines shown on the display, thus making the most out of the poor resolution available; textual prompts were both drawn on screen and rendered as audio speech using Festival lite
- 3. Apply the standard OpenCV HAAR[11] classifiers for left and right eyes (haarcascade_mcs_lefteye.xml and haarcascade_mcs_righteye.xml) to the image areas that should contain the user's eyes; HAAR classifiers execute slowly on these slow systems, but constraining the recognition to a small portion of the low-resolution (640x480 pixel) built-in camera view results in a framerate >5FPS
- 4. If both eyes were found, use the standard OpenCV HAAR classifiers to check for the and nose (haarcascade_mcs_nose.xml) mouth (haarcascade_mcs_mouth.xml) in the appropriate areas of the image; again, restricting the image area

Figure 2. The Masks, positioned incorrectly



Figure 3. Sample Faces, correctly and incorrectly masked

searched by the HAAR classifiers improves the framerate 5. If the timeout has not expired and more samples are needed

- to decide mask status, go to step 3
- 6. Display, announce, record the results, and/or open the door
- 7. After a short delay giving time for the person to react to the results, go to step 1

For simplicity, there are various details omitted from the above description, such as the determination of face area to maximally fit the camera aspect ratio and conversion of the color images to monochrome before applying the classifiers. However, the complete first version required fewer than 200 lines of source code.

This simple processing worked remarkably well, correctly recognizing the mask status about 85% of the time primarily using myself live with various masks and reasonable scene lighting. Performance degraded substantially in poor lighting as camera captures became of very low quality.

However, most of the errors in decent lighting were the system finding the mouth exposed despite it being covered by a mask. This makes sense in that many masks have vaguely mouthlike patterns; even standard rectangular disposable surgical masks have pleats in them that can produce mouth-like shadows. However, it is very difficult to wear most masks properly covering the nose and not covering the mouth, so adding the simple rule that a covered nose implies a covered mouth increased the correct rate well past 90%. The catch is that our testing was nowhere near rigorous enough to be sure the same behavior would be seen in more realistic tests.

When this work was initiated in Summer 2020, we did not find any datasets that contained an appropriate collection of unmasked, properly masked, and improperly masked faces. Most early interest in masked faces seemed to center on being able to identify the person despite a portion of their face being obscured by wearing the mask. In addition, it seems that in many parts of the world, masks nearly always have one of only a few styles, most commonly the light-blue creased rectangle of a disposable surgical mask or a plain black reusable cloth mask. Empirically, this is not true in the United States.

Masks are often treated as a fashion or style statement, or even as a branding opportunity, in the United States. Reusable cloth masks in a wide variety of styles and colors, often with printed patterns or company logos, are very common. A significant number of masks actually have stylized mouths printed on them – and you can order a custom mask with a photo of your mouth printed on it. There are even masks with clear plastic over the mouth that are advertised as helping deaf people by enabling lip reading.

To reasonably represent this variety of mask designs, primarily online ads were searched for mask images – the idea being that advertised masks would have a similar statistical distribution of design variants to masks actually in use. Over 200 images of masks in as-worn orientation were manually extracted from online sources. The images varied wildly in quality, so the collection was paired-down to 100 acceptable-quality images that seemed representative of the variations in the entire set. Each mask image was manually aligned and scaled to a reference composite face image, and then the background was removed. The resulting mask images are shown in Figure 1.

Each of these 100 mask images was then distorted to simulate being improperly worn. The distortion was done manually using GIMP to warp each mask image to expose portions of the nose and/or mouth in a reference composite image averaging all the aligned faces in our dataset. This produced the 100 aligned mask images shown in Figure 2.

These mask images needed to be imposed on scaled and aligned face images. However, existing face databases tended to lack the desired level of diversity, and many faces were oriented in ways not matching the straight-on view in which our scanner would see the user. Thus, we collected approximately 2000 appropriately-aligned base face image that were created by generative adversarial networks (GANs)[12], primarily from online sources including https://thispersondoesnotexist.com/ and https://generated.photos/. Synthesizing faces made it easier to achieve a balanced representation of adult faces across ethnicities, ages, genders, hair, and even eyeglasses. We did not find any generated faces with significant deformities, tattoos, nor jewelry such as nose piercings, so the face data still carries some biases. Unfortunately, redistribution rights are either unclear or not granted for faces synthesized by some tools, so we cannot redistribute the face data set used.

The test and training data was thus created by writing a simple program to impose mask images on randomly-selected



Figure 4. Typical cascade features for mouth vs. nose and mouth

faces. The faces and mask images could also be horizontally flipped. Thus, there could be approximately 4000 unmasked face images and 800,000 possible combinations each of properly and improperly masked faces. Visually insignificant perturbations of the scale, angle, and positioning, as well as smoothing, also could be imposed to avoid training the recognizer to identify the artificially-clipped mask edges. Many training runs were conducted using subsets of this dataset with several different types of recognizers, especially using the OpenCV cascade classifier training support for both HAAR[11] and the LBP (Local Binary Pattern) cascades.

Cascade features selected in two representative training runs (in this case using the cheaper LBP) are shown in Figure 4. The left series was trained to distinguish only if the mouth is covered; the right series was trained to recognize if both the nose and mouth are covered. Not surprisingly, the features sample around the mouth to determine if the mouth is covered, however, only the last feature in the nose and mouth recognizer samples the mouth. This bias toward nose features held across all training runs, and strongly confirms our earlier observation that **it is sufficient to simply check if the nose is visible to confirm a mask is being worn properly**. The system has not yet been tested in a full installation, but a simple nose recognizer can probably deliver better than 90% accuracy despite poor camera quality and framerate.

Temperature check (optional)

Figures 5 and 6 are captures of the summary images displayed by an early version of the system respectively for no mask and a properly-worn mask. Not only does the system recognize proper masking, but it also measures temperatures. The temperature displayed and green coloring of the face indicate the face is within the normal temperature range; the background is blue because it is colder.

Thermal imagers with good resolution are not inexpensive



Figure 5. Nose and mouth not covered – no mask



Figure 6. Nose and mouth properly covered by a mask

and, during this pandemic, have become extremely difficult to obtain. However, because the mask check ensures the user's face is reasonably close and well-aligned, it is possible to use a much lower pixel count thermal imager.

At this conference in 2020, KVIRP (Kentucky's Visual / Infra Red Painter) demonstrated effective use of an inexpensive thermal imager with a pair of fisheye cameras to capture high-resolution 360° visible-light images painted with thermal data[13]. With some repackaging, that same thermal imager is used here.

The thermal imager is shown mounted just above the built-in camera of the Asus laptop in Figure 7, and by itself in Figure 8. Our 3d-printed thermal camera contains a \$3 ATmega32U4 Pro Micro[14] to provide a USB interface and a \$40 Adafruit Grid Eye Thermal Camera board[15]. The 8x8 pixel AMG8833 sensor measures temperatures from 0°C to 80°C with a thermal resolution of approximately 0.25°C and absolute accuracy of +/-2.5°C.



Figure 7. Thermal imager mounted on back of laptop



Figure 8. AMG8833-based thermal imager in 3D-printed mount

Samples from the thermal imager are averaged over periods when eye alignment is detected, and the highest temperature from the face region is reported; it is only because the face is positioned close to the camera that the low 8x8 pixel resolution is usable.

As is the case with most thermal imagers, accuracy is not sufficient to reliably detect fevers. In medical applications, thermal imagers are required to be calibrated by imaging a known temperature source – but that is not feasible here. Instead, the system enters a calibration mode at boot where a reference normal temperature is measured. As the system is operating, it uses an age-weighted average of median values measured to attempt to correct drift of thermal readings. This essentially assumes that the average person measured will have a normal temperature so it can continuously self-calibrate, which is an assumption commonly made by fever-detection scanners.

Contact tracing (optional)

The key to maintaining privacy while contact tracing is the ability to have no personally-identifiable information stored in the database. We propose to do this by issuing each person a unique ID and associating contract, infection, and quarantine status with that ID rather than with the person.

The system does not use face identification to recognize users; in fact, it literally never stores a face image. Instead, we prefer that it recognizes users by either an RFID or QR code, and QR code can be printed with a conventional printer or displayed on a cell phone, and can easily be recognized using the OpenCV QRCodeDetector with the same camera used to check mask status. The key concepts are:

- Each person has one randomly-generated ID and a corresponding password for it. No record is made of which person owns each ID; only the fact that a person has been issued an ID is recorded to ensure one person has only one ID.
- Using their ID and password, a person may create multiple unique, randomly-generated, QR codes to represent that ID. Any data recorded about any of those QR codes is thus connected to the ID that created that QR code.
- Upon entering or exiting an area protected by a Covered Safe Entry Scanner, the user activates the system by showing any one of their QR codes. Thus, each QR code entry/exit is recorded in a common database.
- When the QR code is scanned, Covered checks the database for issues associated with that QR code. An ID that has been flagged as infected, potentially exposed, having a pending request for a meeting with a human contact tracer or a healthcare provider for consultation, etc., causes an appropriate message to be issued. If the ID associated with the scanned QR code has recently passed a check, the system can skip mask and temperature checks to speed entry.
- Because the database literally does not know how to contact the owner of an ID, detailed information about the status of an ID is only available to the person using their ID and password to access their information via an online database interface. Even healthcare workers making authorized updates to the database generally do not need to know IDs and which people are associated with them; all interactions can use a QR code to identify the user.

To provide these services, the computer or tablet running Covered must have at least intermittent internet access to the database, which was prototyped as an SQLite database to be run on a server with an HTTP (web browser compatible) interface.

Conclusion

This paper has discussed the pandemic caused by SARS-CoV-2 and three key mechanisms that can be implemented in a safe entry scanner to slow spread of the virus. Most important is the automatic recognition of a properly-worn mask, which turns out to be remarkably easy to do, essentially by detecting that the nose is covered. Additionally, methods are described for implementing a simple fever check and for semi-automatic contact tracing without exposing any personally-identifiable data.

The system was conceived and prototyped in Summer 2020, but has not yet been tested in a real-world environment. Although a provisional patent was filed by the University of Kentucky, it has always been our intent that the recognition of a properlyworn mask should be an open technology and details will be at http://aggregate.org/DIT/COVERED. The patent reflects our desire to find a partner that could further develop the complete system and put it into widespread use. The do-it-yourself nature of the system, combined with fears about potential liability, have thus far prevented application of the system in venues such as the retail stores that inspired creation of the system.

References

- "Origin of SARS-CoV-2," World Health Organization, WHO/2019nCoV/FAQ/Virus_origin/2020.1, March 26, 2020.
- [2] "COVID-19," Centers for Disease Control and Prevention, https://www.cdc.gov/coronavirus/2019-ncov/index.html, accessed January 29, 2021.
- [3] Mahesh Jayaweera, Hasini Perera, Buddhika Gunawardana, and Jagath Manatungea. "Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy," Environmental research vol. 188 (2020): 109819. doi:10.1016/j.envres.2020.109819 2020.
- [4] "HIPAA and COVID-19 Bulletin: Limited Waiver of HIPAA Sanctions and Penalties During a Nationwide Public Health Emergency," United States Department of Health and Human Services (HHS), March 15, 2020.
- [5] Preeti Nagrath, Rachna Jain, Agam Madan, Rohan Arora, Piyush Kataria, and Jude Hemanthb. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," Elsevier Public Health Emergency Collection PMC7775036, doi: 10.1016/j.scs.2020.102692, December 31, 2020.
- [6] "Masked AI: AI Solution for Face Mask Identification," #Build-ForCovid19 Hackathon, https://buildforcovid19.io/masked-ai-aisolution-for-face-mask-identification/, accessed January 29, 2021.
- [7] Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N. Taha, and Nour Eldeen M. Khalifa. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," Measurement, Volume 167, 2021, 108288, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2020.108288. January 1, 2021.
- [8] Chandrika Deb, et al. "Face Mask Detection system based on computer vision and deep learning using OpenCV and Tensorflow/Keras," https://github.com/chandrikadeb7/Face-Mask-Detection, accessed January 29, 2021.
- [9] Alex Escola' Nixon. "Face Mask Detection using YOLOv5," https://towardsdatascience.com/face-mask-detection-using-yolov5-3734ca0d60d8, November 9, 2020.
- [10] Madhura Inamdar and Ninad Mehendale. "Real-Time Face Mask Identification Using Facemasknet Deep Learning Network," http://dx.doi.org/10.2139/ssrn.3663305, August 6, 2020.
- [11] Paul Viola and Michael Jeffrey Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1:I-511–I-518 vol.1 DOI: 10.1109/CVPR.2001.990517, 2001.
- [12] Tero Karras, Samuli Laine, and Timo Aila. "A Style-Based Generator Architecture for Generative Adversarial Networks," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4401-4410, 2019.
- [13] Henry Dietz and Paul Eberhart. "Senscape: modeling and presentation of uncertainty in fused sensor data live image streams," Electronic Imaging, Computational Imaging XVIII, pp. 392-1-392-7(7), https://doi.org/10.2352/ISSN.2470-1173.2020.14.COIMG-392, 2020.
- [14] Sparkfun, "ATmega32U4 Arduino board," https://github.com/sparkfun/Pro_Micro (accessed 1/14/2020)
- [15] Dean Miller, "Adafruit AMG8833 8x8 Thermal Camera Sensor," August 22, 2018, https://cdnlearn.adafruit.com/downloads/pdf/adafruit-amg8833-8x8-thermalcamera-sensor.pdf (accessed 1/14/2020)