https://www.nist.gov/news-events/news/2020/07/nist-launches-studies-masks-effect-face-recognition-software
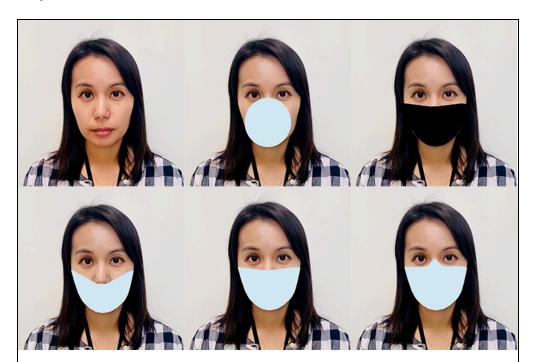
National Institute of Standards and Technology
U.S. Department of Commerce

NEWS (https://www.nist.gov/news-events/news)

NIST Launches Studies into Masks' Effect on Face Recognition Software

Algorithms created before the pandemic generally perform less accurately with digitally masked faces.

July 27, 2020

NIST digitally applied mask shapes to photos and tested the performance of face recognition algorithms developed before COVID appeared. Because real-world masks differ, the team came up with variants that included differences in shape, color and nose coverage.

Credit: B. Hayes/NIST

Now that so many of us are covering our faces to help reduce the spread of ,
how well do face recognition algorithms identify people wearing masks? Th
according to a preliminary study by the National Institute of Standards and
Technology (NIST), is with great difficulty. Even the best of the 89 commercial facial

recognition algorithms tested had error rates between 5% and 50% in matching digitally applied face masks with photos of the same person without a mask.

The results were published today as a NIST Interagency Report (NISTIR 8311 (https://doi.org/10.6028/NIST.IR.8311)), the first in a planned series from NIST's Face Recognition Vendor Test (FRVT) program on the performance of face recognition algorithms on faces partially covered by protective masks.

"With the arrival of the pandemic, we need to understand how face recognition technology deals with masked faces," said Mei Ngan, a NIST computer scientist and an author of the report. "We have begun by focusing on how an algorithm developed before the pandemic might be affected by subjects wearing face masks. Later this summer, we plan to test the accuracy of algorithms that were intentionally developed with masked faces in mind."

The NIST team explored how well each of the algorithms was able to perform "one-to-one" matching, where a photo is compared with a different photo of the same person. The function is commonly used for verification such as unlocking a smartphone or checking a passport. The team tested the algorithms on a set of about 6 million photos used in previous FRVT studies. (The team did not test the algorithms' ability to perform "one-to-many" matching, used to determine whether a person in a photo matches any in a database of known images).

The research team digitally applied mask shapes to the original photos and tested the algorithms' performance. Because real-world masks differ, the team came up with nine mask variants, which included differences in shape, color and nose coverage. The digital masks were black or a light blue that is approximately the same color as a blue surgical mask. The shapes included round masks that cover the nose and mouth and a larger type as wide as the wearer's face. These wider masks had high, medium and low variants that covered the nose to different degrees. The team then compared the results to the performance of the algorithms on unmasked faces.

"We can draw a few broad conclusions from the results, but there are caveats," Ngan said. "None of these algorithms were designed to handle face masks, and the masks we used are digital creations, not the real thing."

If these limitations are kept firmly in mind, Ngan said, the study provides a few general lessons when comparing the performance of the tested algorithms on masked faces versus unmasked ones.

- **Algorithm accuracy with masked faces declined substantially across the board.** Using unmasked images, the most accurate algorithms fail to authenticate a person about 0.3% of the time. Masked images raised even these

top algorithms' failure rate to about 5%, while many otherwise competent algorithms failed between 20% to 50% of the time.

- **Masked images more frequently caused algorithms to be unable to process a face, technically termed "failure to enroll or template" (FTE).** Face recognition algorithms typically work by measuring a face's features — their size and distance from one another, for example — and then comparing these measurements to those from another photo. An FTE means the algorithm could not extract a face's features well enough to make an effective comparison in the first place.

- **The more of the nose a mask covers, the lower the algorithm's accuracy.** The study explored three levels of nose coverage — low, medium and high — finding that accuracy degrades with greater nose coverage.

- **While false negatives increased, false positives remained stable or modestly declined.** Errors in face recognition can take the form of either a "false negative," where the algorithm fails to match two photos of the same person, or a "false positive," where it incorrectly indicates a match between photos of two different people. The modest decline in false positive rates show that occlusion with masks does not undermine this aspect of security.

- **The shape and color of a mask matters.** Algorithm error rates were generally lower with round masks. Black masks also degraded algorithm performance in comparison to surgical blue ones, though because of time and resource constraints the team was not able to test the effect of color completely.

The report, _Ongoing Face Recognition Vendor Test (FRVT) Part 6A: Face recognition accuracy with face masks using pre-COVID-19 algorithms (https://doi.org/10.6028/NIST.IR.8311)_, offers details of each algorithm's performance, and the team has posted additional information online (https://pages.nist.gov/frvt/html/frvt_facemask.html).

Ngan said the next round, planned for later this summer, will test algorithms created with face masks in mind. Future study rounds will test one-to-many searches and add other variations designed to broaden the results further.

"With respect to accuracy with face masks, we expect the technology to continue to improve," she said. "But the data we've taken so far underscores one of the ideas common to previous FRVT tests: Individual algorithms perform differently. Users should get to know the algorithm they are using thoroughly and test its performance in their own work environment."

This work was conducted in collaboration with the Department of Homeland Security's Science and Technology Directorate, Office of Biometric Identity Management, and Customs and Border Protection.

* EDITOR'S NOTE: The headline was updated to clarify the investigation is a scientific study and should not be interpreted as a legal investigation.
[Information technology (https://www.nist.gov/topic-terms/information-technology)](https://www.nist.gov/topic-terms/information-technology), [Biometrics (https://www.nist.gov/topic-terms/biometrics)](https://www.nist.gov/topic-terms/biometrics), [Cybersecurity (https://www.nist.gov/topic-terms/cybersecurity)](https://www.nist.gov/topic-terms/cybersecurity) and [Identity and access management (https://www.nist.gov/topic-terms/identity-and-access-management)](https://www.nist.gov/topic-terms/identity-and-access-management)

Media Contact

- [Chad Boutin (https://www.nist.gov/people/chad-boutin)](https://www.nist.gov/people/chad-boutin)
  [charles.boutin@nist.gov (https://www.nist.govmailto:charles.boutin@nist.gov)](https://www.nist.govmailto:charles.boutin@nist.gov)
  (301) 975-4261

Related Publications

[Ongoing Face Recognition Vendor Test (FRVT) Part 6A: Face recognition accuracy with masks using pre- COVID-19 algorithms (https://www.nist.gov/publications/ongoing-face-recognition-vendor-test-frvt-part-6a-face-recognition-accuracy-masks-using)](https://www.nist.gov/publications/ongoing-face-recognition-vendor-test-frvt-part-6a-face-recognition-accuracy-masks-using)
Released July 27, 2020, Updated August 4, 2020