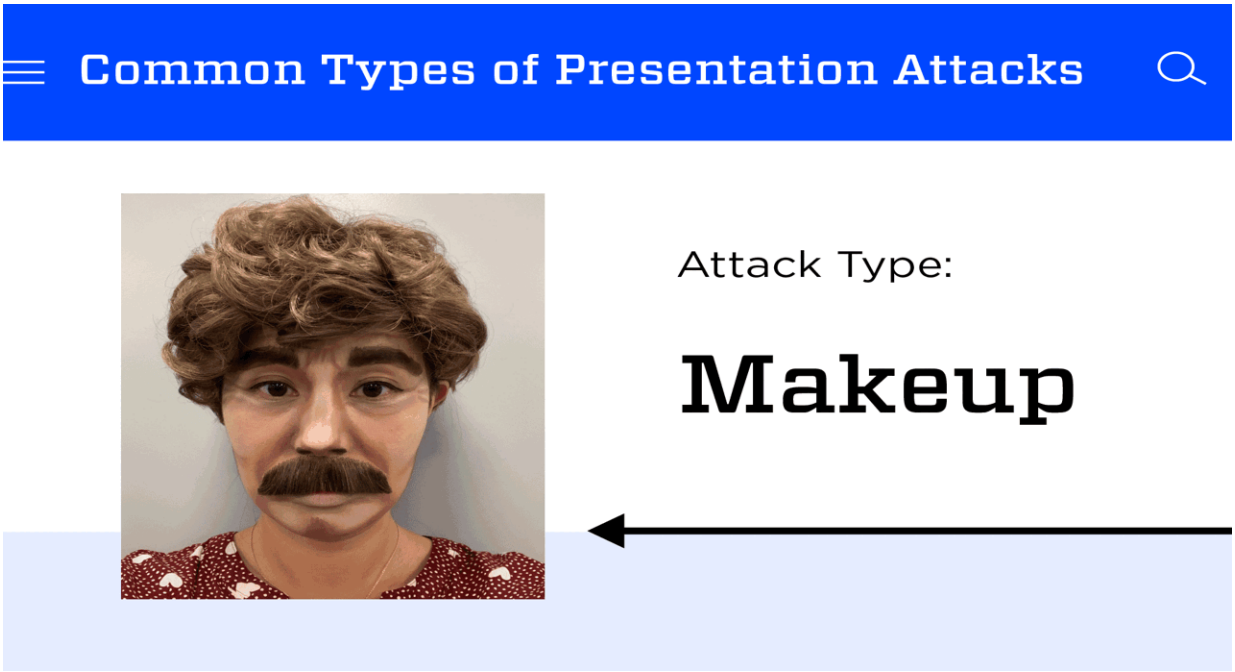
What's wrong with this picture? Face analysis program helps to find answers

September 20 2023



Some NIST face analysis evaluations test whether software can detect presentation attacks, where a spoof image is used in attempts to gain access to a device or space or simply to hide someone's true identity. Presentation attacks can take many forms, such as wearing makeup, holding up a printed photo or displaying a digital photo of another person. Credit: M. Ngan, N. Hanacek/NIST

Face recognition software is commonly used as a gatekeeper for accessing secure websites and electronic devices, but what if someone can defeat it by simply wearing a mask resembling another person's

face? Newly published research from the National Institute of Standards and Technology (NIST) reveals the current state of the art for software designed to detect this sort of spoof attack.

The new study appears together with another that evaluates software's ability to call out potential problems with a photograph or digital face image, such as one captured for use in a passport. Together, the two NIST publications provide insight into how effectively modern image-processing software performs an increasingly significant task: face analysis.

Face analysis is distinct from face recognition, which may be a more familiar term. Broadly speaking, face recognition aims to identify a person based on an image, while face analysis is concerned with image characterization, such as flagging images that are themselves problematic—whether because of nefarious intent or simply due to mistakes in the photo's capture.

The two publications are the first on the subject to appear since NIST divided its Face Recognition Vendor Test (FRVT) program into two tracks, Face Recognition Technology Evaluation (FRTE) and Face Analysis Technology Evaluation (FATE). Efforts involving the processing and analysis of images, as the two new publications do, now are categorized under the FATE track.

Technology tests on both tracks are meant to provide information on the capabilities of algorithms to inform developers, end users, standards processes, and policy and decision makers.

"Can a given software algorithm tell you whether there's something wrong with a face image?" said Mei Ngan, a NIST computer scientist. "For example, are the person's eyes closed? Is the image blurry? Is the image actually a mask that looks like another person's face? These are

the sort of defects that some developers claim their software can detect, and the FATE track is concerned with evaluating these claims."

Ngan is an author of the first study, Face Analysis Technology Evaluation (FATE) Part 10: Performance of Passive, Software-Based Presentation Attack Detection (PAD) Algorithms, which evaluated the ability of face analysis algorithms to detect whether these issues constituted evidence of a spoofing attack, referred to as PAD.

The research team evaluated 82 software algorithms submitted voluntarily by 45 unique developers. The researchers challenged the software with two different scenarios: impersonation, or trying to look like another specific person; and evasion, or trying to avoid looking like oneself.

The team evaluated the algorithms with nine types of presentation attacks, with examples including a person wearing a sophisticated mask designed to mimic another person's face and other simpler attacks such as holding a photo of another person up to the camera or wearing an N95 mask that hid some of the wearer's face.

The results varied widely among PAD algorithms, and Ngan noted one thing: Some developers' algorithms worked well at detecting a given type of presentation attack in the images, but none could detect all attack types tested.

"Only a small percentage of developers could realistically claim to detect certain presentation attacks using software," she said. "Some developers' algorithms could catch two or three types, but none caught them all."

Among the other findings was that even the top-performing PAD algorithms worked better in tandem.

"We asked if it would lower the error rate if you combined the results from different algorithms. It turns out that can be a good idea," Ngan said. "When we chose four of the top performing algorithms on the impersonation test and fused their results, we found the group did better than any one of them alone."

The kinds of algorithms that Ngan and her co-authors evaluated have applications in casinos, for example, where a card counter who has been denied entry tries to sneak in wearing a disguise. But the FATE track also evaluates algorithms that serve more everyday purposes, such as checking whether your new passport photo might be rejected. That's what the second of the new NIST studies, Face Analysis Technology Evaluation (FATE) Part 11: Face Image Quality Vector Assessment: Specific Image Defect Detection, explored.

"If you walk into a drugstore and get a passport photo, you want to make sure your application won't be rejected because there is an issue with the photo," said study author Joyce Yang, a NIST mathematician. "Blurry photos are an obvious problem, but there can also be issues with backlighting or simply wearing glasses. We explored algorithms created to flag issues that make a photo noncompliant with passport requirements."

The evaluation was the first of its kind in the FATE track, and the NIST team received seven algorithms from five developers. The study evaluated the algorithms on 20 different quality measures, such as underexposure and background uniformity, all based on internationally accepted passport standards.

Yang said that all the algorithms showed mixed results. Each had its strengths, doing better on some of the 20 measures than others. The results will inform a standard that NIST is helping to develop—ISO/IEC 29794-5, which lays out guidelines for the quality measures that an

algorithm should check. The Specific Image Defect Detection results show how well algorithms perform those checks.

One thing the study did not evaluate was how "good" a picture is, so don't look for aesthetic judgments from your photo booth.

"We're not deciding if the image itself is of good quality," she said. "We are just looking at whether the analysis of the image is correct."

More information: Mei Ngan, Face Analysis Technology Evaluation (FATE) Part 10:, *NIST* (2023). DOI: 10.6028/NIST.IR.8491

Joyce Yang, Face Analysis Technology Evaluation (FATE) Part 11:, *NIST* (2023). DOI: 10.6028/NIST.IR.8485

Provided by National Institute of Standards and Technology