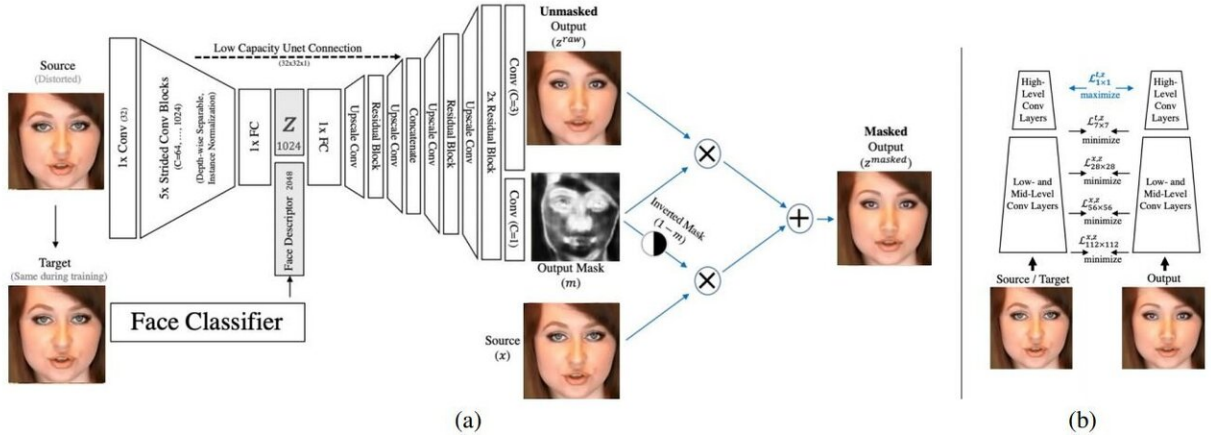
De-identification team explores facial recognition block in videos

October 29 2019, by Nancy Cohen



(a) The architecture of the network. For conditioning, a pre-trained face recognition network is used. (b) An illustration of the multi-image perceptual loss used, which employs two replicas of the same face recognition network. Credit: Live Face De-Identification in Video, Oran Gafni et al.

Facebook has figured out the de-identification of people in videos. Wait, Facebook? Aren't social platforms often criticized over privacy rights? Not this time, at least not over in the halls of Facebook AI Research.

Oran Gafni, a member of the team who worked on protecting faces against recognition systems, posted a video, "De-Identification Video Samples" on October 15. Gafni, a research engineer in Facebook AI Research, holds electrical engineering degrees from Tel-Aviv University,

under the supervision of Prof. Lior Wolf. Gafni's thesis focused on live semantic face editing in video, using deep adversarial autoencoders.

Their paper discussing their work is titled "Live Face De-Identification in Video." Authors are Gafni and Lior Wolf, with stated affiliations of Facebook AI Research and Tel-Aviv University, and Yaniv Tagman, Facebook AI Research.

Khari Johnson in VentureBeat picked up on why the research matters. "Startups like D-ID and a number of previous works have made de-identification technology for still images, but this is the first one that works on video."

Facebook's knowhow is in the realm of research. Specifically, this is Facebook AI Research, and the group has no plans, a spokesperson told *VentureBeat*, "to apply the tech to any part of the Facebook family of apps at this time."

Facebook, though, could benefit from attempts to champion anonymity.

Johnson noted the recent controversy about applications of facial recognition technology. Facebook faces a $35 billion threat of a class action lawsuit, as reported by news sites including TechCrunch.

How it works: The AI is all about automatic video modification. The method maps a slightly distorted version on a person's face such that it becomes difficult for facial recognition technology to identify a person.

Johnson said their method "pairs an adversarial autoencoder with a classifier network."

Johnson went on to clearly describe what is going on: "The AI uses an encoder-decoder architecture to generate both a mask and an image.

During training, the person's face is distorted then fed into the network. Then the system generates distorted and undistorted images of a person's face for output that can be embedded into video."

How well does it work? Researchers tried to fool facial recognition networks and they emerged confident that they have come upon a meaningful technique. "Our contribution is the only one suitable for video, including live video, and presents quality that far surpasses the literature methods. The approach is both elegant and markedly novel, employing an existing face descriptor concatenated to the embedding space, a learned mask for blending, a new type of perceptual loss for getting the desired effect, among a few other contributions."

Check out the figures in the paper, which are examples of their method's identity shifts. The authors pointed out that only minimally changing the image is important for the method to be video-capable. They said that in their work, change is measured using low- and mid-level features and not using norms on the pixels themselves.

They were aware of research that had shown image perturbations caused by adversarial examples distort mid-level features "which we constrain to remain unchanged."

In an interview with *VentureBeat*, Wolf said that "the autoencoder is such that it tries to make life harder for the facial recognition network, and it is actually a general technique that can also be used if you want to generate a way to mask somebody's, say, voice or online behavior or any other type of identifiable information."

The International Conference on Computer Vision (ICCV) in Seoul, South Korea, is a venue where the Facebook researchers were listed to join computer vision experts from around the world to discuss the latest advances. The Seoul schedule was reported to have the team's

presentation on tap, "Live Face De-Identification in Video."

Their paper's abstract reads: "We propose a method for face de-identification that enables fully automatic video modification at high frame rates. The goal is to maximally de-correlate the identity while having the perception (pose, illumination, and expression) fixed. We achieve this by a novel feed-forward encoder-decoder network architecture that is conditioned on the high-level representation of a person's facial image. The network is global, in the sense that it does not need to be retrained for a given video or for a given identity, and it creates natural-looking image sequences with little distortion in time."

Tyler Lee in *Ubergizmo* recognized to the human eye any difference between the before and after photos could be puzzling, but the changes were enough to confuse the system. Lee said "this seems like some kind of reverse deepfake where it distorts the person's face ever so slightly to the point where it can confuse facial recognition systems, but at the same time maintain enough of the original so that you, as a human, will definitely know who it is that you're seeing."

More information: research.fb.com/publications/l … tification-in-video/