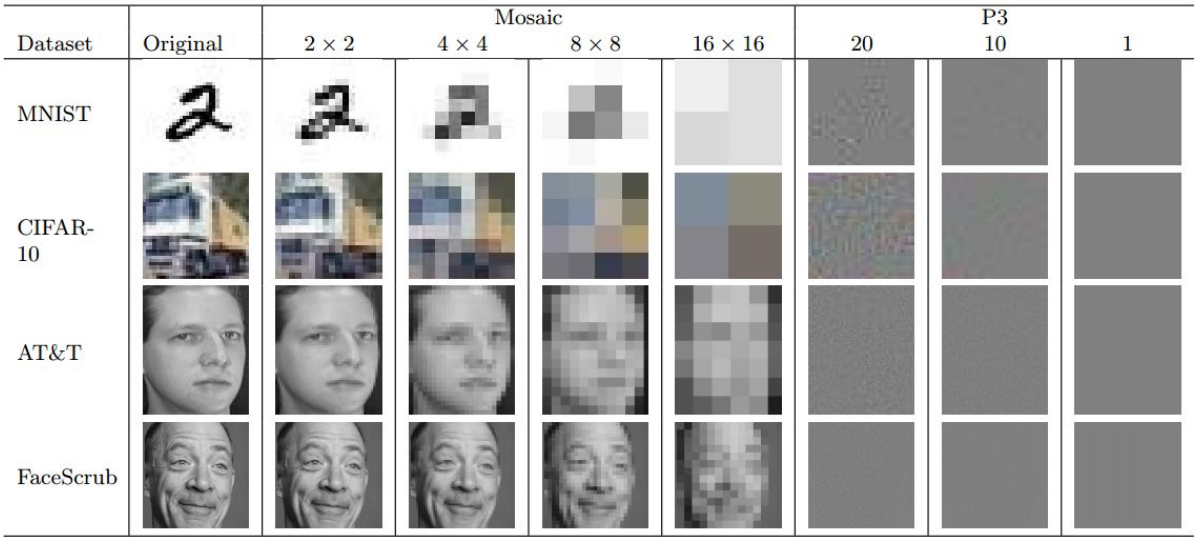
AI software found able to identify people in blurred images

September 16 2016, by Bob Yirka



Examples images from each dataset. The leftmost image is the original image. The remaining columns are the image obfuscated with mosaicing with windows of 2 × 2, 4 × 4, 8 × 8, and 16 × 16 pixels and P3 with thresholds of 20, 10, and 1. Credit: arXiv:1609.00408 [cs.CR]

(Tech Xplore)—A trio of researchers has found off-the-shelf AI software can be used to identify people in blurred or pixilated images. Reza Shokri and Vitaly Shmatikov with Cornell University and Richard McPherson with the University of Texas have uploaded a paper to the *arXiv* preprint server describing the experiments they carried out with AI software identification of people or other items in blurred out

images, what they found and reveal just how accurate they found it could be.

A popular means of retaining privacy in videos or photographs while still maintaining some degree of authenticity is blurring the parts you do not want people to recognize, such as faces that appear at a protest rally, for example. But now, it appears that this technique may not be enough, because computers have become smart enough to recognize them anyway.

The AI software is not able to reconstitute an image, the team notes—rather, it analyzes the image and compares what it finds with other pictures available on Facebook, Instagram or YouTube, for example—places where there are photographs of identifiable people. Their study consisted of obtaining pictures of people from public places on the Internet and then blurring or pixelating them. Both images were then fed to the AI system to teach it how to spot one given the other. Once that was complete, the researchers then fed the system different pictures of the same people and asked it to identify which photos corresponded with blurred images. They found the AI system able to do so with an overall average of 57 percent accuracy. That average accuracy jumped to 85 percent when the system was given four more chances with each image. The accuracy of individual results varied by degree of blurring or pixilating, they note. They also point out that the same type of technology could be used to figure out a street address that has been blurred out or to identify some other object.

What this means, the researchers explain, is that if a person posts a picture on the Internet that has their face in it, but which has been pixilated or blurred out, there is a better than even chance that someone could identify them using similar software, particularly if they have a strong Internet presence. They suggest that users of products such as YouTube's blurring service be made aware that such measures are not

adequate to protect a person's privacy.

More information: Defeating Image Obfuscation with Deep Learning, arXiv:1609.00408 [cs.CR] arxiv.org/abs/1609.00408

Abstract

We demonstrate that modern image recognition methods based on artificial neural networks can recover hidden information from images protected by various forms of obfuscation. The obfuscation techniques considered in this paper are mosaicing (also known as pixelation), blurring (as used by YouTube), and P3, a recently proposed system for privacy-preserving photo sharing that encrypts the significant JPEG coefficients to make images unrecognizable by humans. We empirically show how to train artificial neural networks to successfully identify faces and recognize objects and handwritten digits even if the images are protected using any of the above obfuscation techniques.