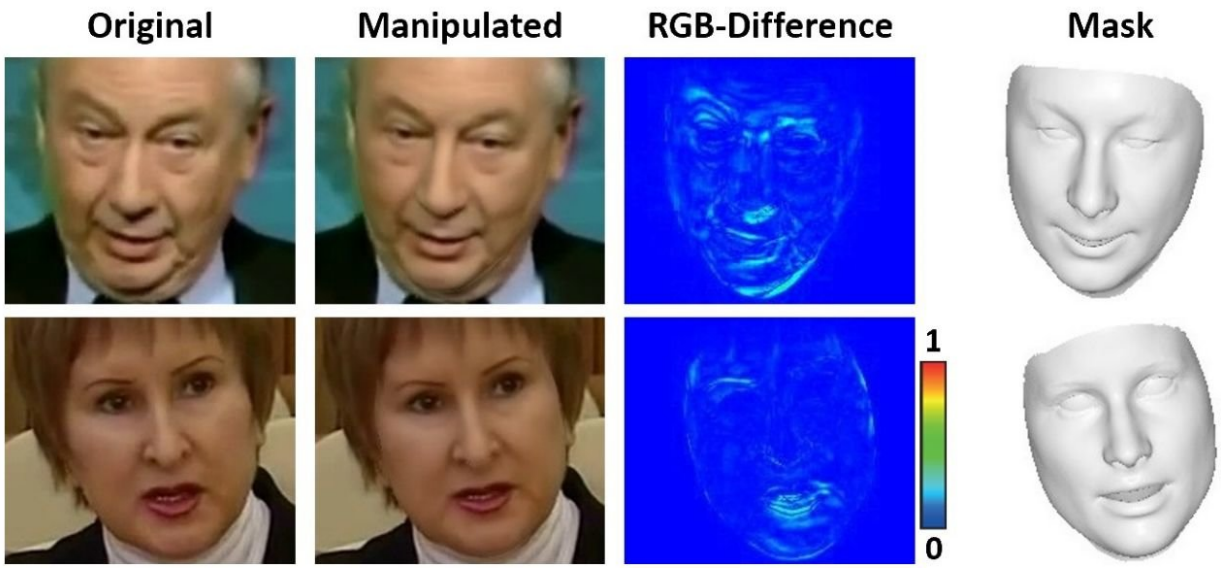
Researchers work on algorithm that reveals face swaps

April 13 2018, by Nancy Owano



Fake or Real? Examples of the FaceForensics Self-Reenactment Dataset. From left to right: original input image, self-reenacted output image, color difference plot and face mask that is used during synthesis of the output image. Credit: arXiv:1803.09179 [cs.CV]

Image manipulation in this advanced stage of the digital age is not as much fun but a dicey weapon, in the shadows of fake news, to sway opinion and spark scandals.

Face-swapping, in particular, sounds like fun if you think of it as a

chuckle at a family table while kids and adults try out different faces on different people. However, it's also a tool for far worse motives. Swapna Krishna in *Engadget* remarked that "People have, of course, taken advantage of this tool for some disturbing uses, including face-swapping people into pornographic videos—the ultimate revenge porn."

In *MIT Technology Review*, the "Emerging Technology from the arXiv" said, "pornographic videos called 'deepfakes' have emerged on websites showing famous individuals' faces superimposed onto bodies of actors."

Researchers, however, interested in exploring the tool and how to tell if it is used, have come up with an algorithm, say observers, that can outdo other techniques available. They figured out a way to detect a face swap via the algorithm, picking up on forged videos as soon as posted.

Analytics Vidhya commented, "Something akin to this algorithm was desperately required to wage the battle against face swaps being used for the wrong reasons. In releasing the research paper to the public, the researchers are hoping others also take up the baton and work on this study to make it more accurate and precise."

Andreas Rossler was team leader of the participants from Technical University of Munich, University Federico II of Naples and the University of Erlangen-Nuremberg.

They trained the algorithm, XceptionNet, using a large set of face swaps, said *Engadget*.

"We set a strong baseline of results for detecting a facial manipulation with modern deep-learning architectures," said Rossler and team in *MIT Technology Review*. Size mattered.

The size of this database was a significant improvement over what had

been previously available. "We introduce a novel data set of manipulated videos that exceeds all existing publicly available forensic data sets by orders of magnitude," said Rossler.

In their paper, the authors said they introduced a face manipulation dataset, FaceForensics, "of about half a million edited images (from over 1000 videos)."

The paper is titled "FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces," on arXiv. Authors are Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies and Matthias Nießner.

The authors called attention to the difficulty—for humans and computers— in trying to distinguish between original and manipulated video, "especially when the videos are compressed or have low resolution, as it often happens on social networks."

They also called attention to the fact that "Research on the detection of face manipulations has been seriously hampered by the lack of adequate datasets."

There is a nuance in their success, though, that also merits attention. The "Emerging Technology from the arXiv" article called it the "sting in the tail." What is it? "The same deep-learning technique that can spot face-swap videos can also be used to improve the quality of face swaps in the first place—and that could make them harder to detect."

Engadget similarly said, "XceptionNet clearly outperforms its rival techniques in detecting this kind of fake video, but it also actually improves the quality of the forgeries. Rossler's team can use the biggest hallmarks of a face swap to make the manipulation more seamless. It doesn't fool XceptionNet, but in the long run, it could make it harder for

other methods to detect faked videos."

Pranav Dar, in *Analytics Vidhya*, also weighed in on what he called "a caveat with this algorithm – it can also potentially be used to improve the quality of the face swaps which will make it harder to detect the fake. Also, as soon as a forgery detection algorithm is launched, the scammers always try to refine their model to stay a step ahead."

Nonetheless, the authors said, "our refiner mainly improves visual quality, but it only slightly encumbers forgery detection for deep learning method trained exactly on the forged output data."

More information: FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces, arXiv:1803.09179 [cs.CV] arxiv.org/abs/1803.09179

Abstract
With recent advances in computer vision and graphics, it is now possible to generate videos with extremely realistic synthetic faces, even in real time. Countless applications are possible, some of which raise a legitimate alarm, calling for reliable detectors of fake videos. In fact, distinguishing between original and manipulated video can be a challenge for humans and computers alike, especially when the videos are compressed or have low resolution, as it often happens on social networks. Research on the detection of face manipulations has been seriously hampered by the lack of adequate datasets. To this end, we introduce a novel face manipulation dataset of about half a million edited images (from over 1000 videos). The manipulations have been generated with a state-of-the-art face editing approach. It exceeds all existing video manipulation datasets by at least an order of magnitude. Using our new dataset, we introduce benchmarks for classical image forensic tasks, including classification and segmentation, considering videos compressed at various quality levels. In addition, we introduce a

benchmark evaluation for creating indistinguishable forgeries with known ground truth; for instance with generative refinement models.