A Review on: FaceForensics ++

FaceForensics++ takes an approach on fake video detection. The inspiration behind this project was the rapid increase of fake face making techniques which causes not only the loss of faith over digital content but also the spread of fake news and misinformation. They have measured the degree to which the detection of such videos is impossible even by humans.

As mentioned in the paper modern editing tools have become widely popular for face editing. Putting one person’s face in another person’s photo. The editing tools have been split in two categories: identity modification and expression modification [1]. Nowadays even lightweight devices can be used to do face swaps such as snapchat filters using mobile phones [1]. Identity modification is the one where one person’s face is replaced into another person’s photo. Expression modification is facial re-enactment which takes one person’s expression and manipulates it.

The dataset consists of around 1.8 million images from 1000 above real and pristine videos [1]. The dataset was later improvised as it avoided as much face occlusions as possible leaving with around 509914 images which were facing the camera (for accuracy). The database was created using two graphics based approaches Face2Face [2] and FaceSwap [3] and two learning based approaches DeepFakes [4] and NeuralTextures [5]. These four methods require source and target actor video pairs as inputs [1]. The video resolution varied from 480p, 720p and 1080p.

After creating the dataset the videos were divided to be used for training, validation and test set and 720, 140 and 140 videos were used in that order for each of the category. Forgery detection was treated as a binary classification problem resulting in each video to be fake or real.

To set the benchmark of fake detection a survey was conducted among computer science students. Where they received 50:50 of real and fake image and were only given a certain amount of time to decide if the image was real or fake. The survey included photos of various quality, which showed that the human accuracy decreased with lower quality photos. From these results it was concluded that humans failed to detect when Face2Face and NeuralTextures were used. So it is safe to say that these two techniques make detection by human eye the most difficult.

For the forgery detection method, a conservative cropping method was used were the photo was enlarged by a scale factor of 1.3 around the center of the face tracked. This is better than other naïve approaches which include the entire picture as oppose to this method where only the face has been used. Methods in [6, 7] are used as well as computer-generated vs natural image detection method [8] and face tampering detection [9] has been used. However the classification using XceptionNet [10] outperforms all other forgery detection techniques.

So we can conclude that XceptionNet brings out the best result when it comes to detecting fake images. As the accuracy was about 96.36% with it.

References:

[1] FaceForensics++

[2] Face2Face Justus Thies, Michael Zollh¨ofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. Face2Face: Real-Time Face Capture and

Reenactment of RGB Videos. In IEEE Conference on Computer Vision and Pattern Recognition, pages 2387–2395, June 2016. 1, 2, 3, 4, 5, 12

[3] Faceswap. https://github.com/MarekKowalski/FaceSwap/. Accessed: 2018-10-29.1, 2

[4] Deepfakes github. https://github.com/deepfakes/faceswap. Accessed: 2018-1029.

1, 2, 4, 14

[5] NeuralTextures: Justus Thies, Michael Zollh¨ofer, and Matthias Nießner. Deferred neural rendering: Image synthesis using neural textures. ACM Transactions on Graphics 2019 (TOG), 2019. 1, 2, 3, 4, 14

[6] Belhassen Bayar and Matthew C. Stamm. A deep learning approach to universal image manipulation detection using a new convolutional layer. In ACM Workshop on Information Hiding and Multimedia Security, pages 5–10, 2016. 3, 6, 7,13, 14

[7] Davide Cozzolino, Giovanni Poggi, and Luisa Verdoliva. Recasting residual-based local descriptors as convolutional neural networks: an application to image

Forgery detection. In ACM Workshop on Information Hiding and Multimedia Security, pages 1–6, 2017. 3, 6, 7, 13, 14

[8] Nicolas Rahmouni, Vincent Nozick, Junichi Yamagishi, and Isao Echizen. Distinguishing computer graphics from natural images using convolution neural networks. In IEEE Workshop on Information Forensics and Security, pages 1–6, 2017. 3, 6, 7, 13, 14

[9] Darius Afchar, Vincent Nozick, Junichi Yamagishi, and Isao Echizen. Mesonet: a compact facial video forgery detection network. arXiv preprint arXiv:1809.00888, 2018. 3, 6, 7, 13, 14

[10] Francois Chollet. Xception: Deep Learning with Depthwise Separable Convolutions. In IEEE Conference on Computer Vision and Pattern Recognition, 2017. 6, 7, 13, 14