6.869 Advances in Computer Vision - Project Proposal

Camilo L. Fosco, Benjamin Lahner

October 25, 2019

We propose to investigate new methods to detect DeepFakes. Facebook recently launched a new Deepfake Detection Challenge ([1]), which comes alongside a new dataset [5] containing tens of thousands of faces with labels indicating if they were manipulated by a Deepfake technique. As this problem is becoming more and more relevant, we propose to explore techniques learned during the course to facilitate Deepfake detection. When looking at Deepfakes, some small perturbations usually become apparent to an attentive observer around the swapped face. These small movements, although close to imperceptible in newer techniques, might form a clear clue that can be used by a neural network to detect the manipulation. We therefore propose to investigate the possibility of using motion magnification to improve Deepfake detection.

We will explore the traditional motion magnification literature ([10, 9]) and focus particularly on the learned motion magnification developments ([7]). We will additionally investigate the literature related to deepfake detection ([6, 4, 2]), and explore in detail the FaceForensics++ publication ([8]), which also comes with a dataset that we will use if time allows it. The main dataset will be the one provided by the Facebook Deepfake Detection Challenge ([1]).

The current alternatives for solving this problem involve different neural networks. Some propose to use biological signals [4] or to detect face warping artifacts [6]. These more complex (although not necessarily more accurate) methods are not readily available. The most accurate technique, however, is available: the XceptionNet proposed in [8] achieves state of the art accuracy at deepfake detection over 4 deepfake generation methods. The implementation of XceptionNet itself is widely available.

We will start by exploring the learning-based video motion magnification algorithm presented by ([7]) to magnify the motion of the videos. To detect the deepfakes, we will use an Xception Network ([3]) trained to detect deepfakes from motion-magnified input, as our hypothesis is that feeding the motion magnified frames will allow our network to pick up on the imperceptible unrealistic face motions generated by the forgery. Using these state of the art motion magnification and deepfake detection models will ensure our results are based on the best current practice. We will evaluate our results in two ways. First, we will submit our approach to Facebook's Deepfake Detection Challenge (DFDC) ([1]) so our methods can be compared and ranked amongst peers. Second, we will calculate the accuracy of our deepfake detection with the use of XceptionNet, the current gold standard in deepfake detection. Based on our two evaluation metrics above, our results will be in the form of (1) a ranking on the DFDC's challenge website and (2) a plot of accuracies on different datasets, including but not limited to the DFDC dataset and FaceForensics++ dataset. By convention, a simple accuracy rating has been the primary way results are compared and evaluated between different deepfake detection techniques. Therefore, we will evaluate and report our results as an accuracy percentage.

References

[1] Deep fake detection challenge, 2019, (accessed October 24, 2019). `https://deepfakedetectionchallenge.ai/`.

[2] Darius Afchar, Vincent Nozick, Junichi Yamagishi, and Isao Echizen. Mesonet: a compact facial video forgery detection network. In *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–7. IEEE, 2018.

[3] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.

[4] Umur Aybars Ciftci and Ilke Demir. Fakecatcher: Detection of synthetic portrait videos using biological signals. *arXiv preprint arXiv:1901.02212*, 2019.

[5] Brian Dolhansky, Russ Howes, Ben Pflaum, Nicole Baram, and Cristian Canton Ferrer. The deepfake detection challenge (dfdc) preview dataset, 2019.

[6] Yuezun Li and Siwei Lyu. Exposing deepfake videos by detecting face warping artifacts. *arXiv preprint arXiv:1811.00656*, 2, 2018.

[7] Tae-Hyun Oh, Ronnachai Jaroensri, Changil Kim, Mohamed Elgharib, Fr'edo Durand, William T Freeman, and Wojciech Matusik. Learning-based video motion magnification. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 633–648, 2018.

[8] Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Nießner. Faceforensics++: Learning to detect manipulated facial images. *arXiv preprint arXiv:1901.08971*, 2019.

[9] Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. Riesz pyramids for fast phase-based video magnification. In *2014 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2014.

[10] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. Eulerian video magnification for revealing subtle changes in the world. 2012.