

Review Paper on Deepfake Video Detection using Neural Networks

Arya Shah¹, Ashwin Thakur², Atharva Kale³, Harsh Bothara⁴, Prof. D. C. Pardeshi⁵

Students, Department of Artificial Intelligence and Machine Learning^{1,2,3,4}

Professor, Department of Artificial Intelligence and Machine Learning⁵

All India Shri Shivaji Memorial Society Polytechnic, Pune, Maharashtra, India

Abstract: *With the increasing computational power, the creation of indistinguishable human synthesized videos, known as deepfakes, has become remarkably easy. These realistic face-swapped deepfakes have raised concerns as they can be utilized for malicious purposes such as causing political unrest, fabricating terrorism events, spreading revenge porn, and blackmailing individuals. In this research, we present a novel deep learning-based method that effectively distinguishes AI-generated fake videos from real ones. Our approach focuses specifically on detecting replacement and reenactment deepfakes. We harness the power of Artificial Intelligence (AI) to combat the challenges posed by AI itself. The core of our system lies in a ResNext Convolutional Neural Network, which extracts frame-level features. These features are then used to train a Long Short-Term Memory (LSTM)-based Recurrent Neural Network (RNN) that classifies videos as either manipulated (deepfake) or authentic (real). To ensure real-time applicability and enhance the model's performance on real-world data, we evaluate our method using a large and balanced dataset.*

Keywords: Deepfake Video Detection, convolutional Neural network (CNN), recurrent neural network (RNN), Long short term memory (LSTM).

REFERENCES

- [1] Yuezun Li, Siwei Lyu, "ExposingDF Videos By Detecting Face Warping Artifacts," in arXiv:1811.00656v3.
- [2] Yuezun Li, Ming-Ching Chang and Siwei Lyu "Exposing AI Created Fake Videos by Detecting Eye Blinking" in arxiv.
- [3] Huy H. Nguyen , Junichi Yamagishi, and Isao Echizen " Using capsule networks to detect forged images and videos".
- [4] Hyeonwoo Kim, Pablo Garrido, Ayush Tewari and Weipeng Xu "Deep Video Portraits" in arXiv:1901.02212v2.
- [5] Umur Aybars Ciftci, İlke Demir, Lijun Yin "Detection of Synthetic Portrait Videos using Biological Signals" in arXiv:1901.02212v2.
- [6] Luisa Verdoliva. Media forensics and deepfakes: an overview. arXiv preprint arXiv:2001.06564, 2020.
- [7] Martyn Jolly. Fake photographs: making truths in photogra phy. 2003.
- [8] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In NIPS, 2014.
- [9] David G`uera and Edward J Delp. Deepfake video detection using recurrent neural networks. In AVSS, 2018.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, 2016.
- [11] An Overview of ResNet and its Variants : <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>
- [12] Long Short-Term Memory: From Zero to Hero with Pytorch: <https://blog.floydhub.com/long-short-term-memory-from-zero-to-hero-with-pytorch/>
- [13] Sequence Models And LSTM Networks https://pytorch.org/tutorials/beginner/nlp/sequence_models_tutorial.html
- [14] <https://discuss.pytorch.org/t/confused-about-the-image-preprocessing-in-classification/3965>
- [15] <https://www.kaggle.com/c/deepfake-detection-challenge/data>
- [16] <https://github.com/ondyari/FaceForensics>

- [15] Y. Qian et al. Recurrent color constancy. Proceedings of the IEEE International Conference on Computer Vision, pages 5459–5467, Oct. 2017. Venice, Italy.
- [16] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros. Image-to- image translation with conditional adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5967–5976, July 2017. Honolulu, HI.
- [17] R. Raghavendra, Kiran B. Raja, Sushma Venkatesh, and Christoph Busch, “Transferable deep-CNN features for detecting digital and print-scanned morphed face images,” in CVPRW. IEEE, 2017.
- [18] D. Guera and E. J. Delp, “Deepfake video detection using recurrent neural networks,” in 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018, pp. 1–6.
- [19] M. A. Younus and T. M. Hasan, “Effective and fast deepfake detection method based on haar wavelet transform,” in 2020 International Conference on Computer Science and Software Engineering (CSASE), 2020, pp. 186–190.
- [20] H. Khalid and S. S. Woo, “Oc-fakedect: Classifying deepfakes using one-class variational autoencoder,” in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 2794–2803.
- [21] U. Ciftci, I. Demir, and L. Yin, “How do the hearts of deep fakes beat? deep fake source detection via interpreting residuals with biological signals,” 08 2020.
- [22] M. Jafar, M. Ababneh, M. Al-Zoube, and A. Elhassan, “Forensics and analysis of deepfake videos,” 04 2020, pp. 053–058.