IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

Robust Deepfake Detection System with Deep Learning Techniques

Yanamala Bhuvaneswari[1], Sanjana Samba[2], Nemani Hiranmayai[3] and Dr Halavath Balaji[4]
Department of Computer Science and Engineering[1,2,3,4]
Sreenidhi Institute of Science and Technology, Hyderabad, India

Abstract: *This comprehensive study investigates the pervasive issue of deep fakes within the context of deep learning applications, focusing on their detection and production. Utilizing a diverse array of deep learning algorithms, including InceptionResnetV2, VGG19, CNN, Xception, InceptionV3, EfficientNetB1, DenseNet121, Hybrid Model, LSTM, ResNext-LSTM, and MRI-GAN, the research systematically evaluates their effectiveness in detecting deep fakes. Results reveal varying levels of accuracy, with Xception emerging as the most precise algorithm, achieving an accuracy of 99.32%. Notably, InceptionResnetV2 and DenseNet121 also demonstrate robust performance, with accuracies surpassing 99%. However, certain models like VGG19 and LSTM exhibit lower accuracy rates, underscoring the need for further refinement. These findings underscore the urgent necessity for robust detection mechanisms amidst the proliferation of malicious deep fakes, safeguarding against potential societal ramifications such as misinformation and privacy breaches.*

Keywords: Deep Learning, Fake Detection, InceptionResnetV2, VGG19, CNN, and Xception

REFERENCES

[1] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.
[2] Y. Bengio, P. Simard, and P. Frasconi, "Long short-term memory," IEEE Trans. Neural Netw, vol. 5, pp. 157–166, 1994.
[3] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.
[4] S. Hochreiter, "Ja1 4 rgenschmidhuber (1997)."long short-term memory"," Neural Computation, vol. 9, no. 8.
[5] M. Schuster and K. Paliwal, "Networks bidirectional reccurent neural," IEEE Trans Signal Proces, vol. 45, pp. 2673–2681, 1997.
[6] J. Hopfield et al., "Rigorous bounds on the storage capacity of the dilute hopfield model," Proceedings of the National Academy of Sciences, vol. 79, pp. 2554–2558, 1982.
[7] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, et al., "Google's neural machine translation system: Bridging the gap between human and machine translation," arXiv preprint arXiv:1609.08144, 2016.
[8] L. Nataraj, T. M. Mohammed, B. Manjunath, S. Chandrasekaran, A. Flenner, J. H. Bappy, and A. K. Roy-Chowdhury, "Detecting gan generated fake images using co-occurrence matrices," Electronic Imaging, vol. 2019, no. 5, pp. 532–1, 2019.
[9] B. Zi, M. Chang, J. Chen, X. Ma, and Y.-G. Jiang, "Wilddeepfake: A challenging real-world dataset for deepfake detection," in Proceedings of the 28th ACM international conference on multimedia, 2020, pp. 2382– 2390.
[10] H. A. Khalil and S. A. Maged, "Deepfakes creation and detection using deep learning," in 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC). IEEE, 2021, pp. 1–4.
[11] J. Luttrell, Z. Zhou, Y. Zhang, C. Zhang, P. Gong, B. Yang, and R. Li, "A deep transfer learning approach to fine-tuning facial recognition models," in 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). IEEE, 2018, pp. 2671–2676.
[12] S. Tariq, S. Lee, H. Kim, Y. Shin, and S. S. Woo, "Detecting both machine and human created fake face images in the wild," in Proceedings of the 2nd international workshop on multimedia privacy and security, 2018, pp. 81–87.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, April 2024

[13] N.-T. Do, I.-S. Na, and S.-H. Kim, "Forensics face detection from gans using convolutional neural network," ISITC, vol. 2018, pp. 376–379, 2018.

[14] X. Xuan, B. Peng, W. Wang, and J. Dong, "On the generalization of gan image forensics," in Chinese conference on biometric recognition. Springer, 2019, pp. 134–141.

[15] P. Yang, R. Ni, and Y. Zhao, "Recapture image forensics based on laplacian convolutional neural networks," in International Workshop on Digital Watermarking. Springer, 2016, pp. 119–128.

[16] B. Bayar and M. C. Stamm, "A deep learning approach to universal image manipulation detection using a new convolutional layer," in Proceedings of the 4th ACM workshop on information hiding and multimedia security, 2016, pp. 5–10.

[17] T. Zhao, X. Xu, M. Xu, H. Ding, Y. Xiong, and W. Xia, "Learning selfconsistency for deepfake detection," in Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 15 023–15 033.

[18] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "Mesonet: a compact facial video forgery detection network," in 2018 IEEE international workshop on information forensics and security (WIFS). IEEE, 2018, pp. 1–7.

[19] E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, and P. Natarajan, "Recurrent convolutional strategies for face manipulation detection in videos," Interfaces (GUI), vol. 3, no. 1, pp. 80–87, 2019.

[20] Y. Li, M.-C. Chang, and S. Lyu, "In ictu oculi: Exposing ai created fake videos by detecting eye blinking," in 2018 IEEE International workshop on information forensics and security (WIFS). IEEE, 2018, pp. 1–7.

[21] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell, "Long-term recurrent convolutional networks for visual recognition and description," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 2625–2634.

[22] Y. Li and S. Lyu, "Exposing deepfake videos by detecting face warping artifacts," arXiv preprint arXiv:1811.00656, 2018.

[23] X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes using inconsistent head poses," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019, pp. 8261–8265.

[24] S. Fernandes, S. Raj, R. Ewetz, J. S. Pannu, S. K. Jha, E. Ortiz, I. Vintila, and M. Salter, "Detecting deepfake videos using attributionbased confidence metric," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 308– 309.

[25] U. A. Ciftci, I. Demir, and L. Yin, "Fakecatcher: Detection of synthetic portrait videos using biological signals," IEEE transactions on pattern analysis and machine intelligence, 2020.

[26]M. S. Rana and A. H. Sung, "Deepfakestack: A deep ensemblebased learning technique for deepfake detection," in 2020 7th IEEE international conference on cyber security and cloud computing (CSCloud)/2020 6th IEEE international conference on edge computing and scalable cloud (EdgeCom). IEEE, 2020, pp. 70–75.

[27] S.-Y. Wang, O. Wang, R. Zhang, A. Owens, and A. A. Efros, "Cnngenerated images are surprisingly easy to spot... for now," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 8695–8704.

[28] A. Gandhi and S. Jain, "Adversarial perturbations fool deepfake detectors," in 2020 International joint conference on neural networks (IJCNN). IEEE, 2020, pp. 1–8.