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 5, May 2024

Exposing Deep Fake Face Detection using LSTM and CNN

Alisha Muskaan[1], Nagarathna S[2], Sandhya C S[3], Viju J[4], B Sumangala[5]
B.E. Students, Department of CSE[1,2,3,4]
Assistant Professor, Department of CSE[5]
Sir M Visvesvaraya Institute of Technology, Bengaluru, India

Abstract: *The rapid advancement of deep learning techniques, creating realistic multimedia content has become increasingly accessible, leading to the proliferation of DeepFake technology. DeepFake utilizes generative deep learning algorithms to produce or modify face features in a highly realistic manner, often making it challenging to differentiate between real and manipulated media. This technology, while beneficial in fields such as entertainment and education, also poses significant threats, including misinformation and identity theft. Consequently, detecting DeepFakes has become a critical area of research. In this paper, we propose a novel approach to DeepFake face detection by integrating Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks. Our method leverages the strengths of CNNs in spatial feature extraction and LSTMs in temporal sequence modeling to enhance detection accuracy. The CNN component captures intricate facial features, while the LSTM analyzes the temporal dynamics of video frames. We evaluate our model on several benchmark datasets, including Celeb-DF (v2), DeepFake Detection Challenge Preview, and FaceForensics++. Experimental results demonstrate that our hybrid CNN-LSTM model achieves state-of-the-art performance, surpassing existing methods in both accuracy and robustness. This study highlights the potential of combining CNN and LSTM architectures for effective DeepFake detection, contributing to the ongoing efforts to safeguard against digital media manipulation*

Keywords: DeepFake, Image Classication, Long Short Term Memory, Convolutional Neural Networks, Residual Networks

I. INTRODUCTION

The pervasive issue of fake documents and media, particularly focusing on the emergence of DeepFakes, which are highly realistic fake images and videos created using deep learning techniques like GANs and autoencoders. These DeepFakes have serious implications, including eroding trust in digital media, promoting disinformation, and even exacerbating social issues like terrorism and violence.

To address the challenges posed by DeepFakes, researchers are actively developing detection techniques. This includes training deep neural network (DNN) models on DeepFake datasets and testing their efficacy in trials. The article aims to delve into various DeepFake detection techniques, tools, and datasets used in both images and videos.

Moreover, it highlights the need for establishing authenticity and provenance in digital content, especially in the era of online misinformation. Traditional methods like certificates of authenticityare insufficient for digital media, leading to the proposal of blockchain-based Proof ofAuthenticity (PoA) systems. These systems utilize blockchain technology to provide immutable and tamper-proof records of digital content, ensuring its authenticity and origin.

Furthermore, the paper discusses the evolution of DeepFake technology, its impact on various industries, and the challenges it poses for society. It also outlines the recent advancements in artificial neural networks (ANNs), particularly in the context of generating and detecting DeepFakes.

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 5, May 2024

II. RELATED WORKS

DeepFake Detection for Human Face Images and Videos

A comprehensive examination of the challenges, methodologies, and advancements in detecting DeepFake videos and images. It begins with an introduction to DeepFake technology, highlighting its profound implications for media manipulation and societal impact. The survey covers various types of DeepFake techniques, emphasizing the rapid evolution of these technologies and the pressing need for robust detection methods. Key challenges discussed include the complexity of distinguishing between real and manipulated media, exacerbated by advancements in DeepFake technology. Detection techniques explored in the survey include image forensics, biometric template protection, and digital watermarking, underscoring the multifaceted approach required to combat DeepFake threats effectively. The document also emphasizes collaborative efforts among researchers and developers to enhance detection technologies and mitigate emerging risks. Beyond media manipulation, the survey discusses practical applications of DeepFake detection in fields such as information security and reversible data hiding within encrypted domains. Overall, the PDF provides valuable insights into the current landscape of DeepFake detection, emphasizing the critical role of ongoing research and cooperation in addressing this evolving challenge.

Comprehensive Review of Deepfake Detection Methods

M. S. Rana et al. presents a detailed analysis of deepfake detection methods based on a systematic review of 112 studies from 2018 to 2020. The study focuses on various aspects, including datasets used, features analyzed, models applied, and measurement metrics used for evaluating deepfake detection techniques. The review highlights the prevalence of deep learning-based methods, particularly CNN models, in detecting deepfakes. The FF++ dataset is noted to be widely used in experiments, and the study emphasizes the importance of setting up a unique framework to enhance the consistency and quality of future research in deepfake detection. Additionally, the paper introduces a taxonomy classifying deepfake detection techniques into four categories and provides insights into the performance evaluation of these methods using different measurement metrics. The conclusion summarizes the key findings and observations, emphasizing the significance of deep learning approaches and the need for more systematic experiments to improve detection outcomes.

A systematic review of Deepfake Detection techniques Models

It offers a comprehensive overview of deepfake detection techniques based on an extensive review of research studies. The review covers a range of topics including the use of deep neural network models, particularly CNN-based models, for effective deepfake detection. Various machine learning models such as SVM and k-MN are also discussed in the context of deepfake detection. The paper introduces a taxonomy for classifying detection algorithms based on media type, features used, detection methods, and clues for detection. It also highlights the importance of combining multiple deep learning methods to enhance detection accuracy. The study emphasizes the significance of setting up a unique framework for consistent and high-quality research in deepfake detection. Additionally, the paper discusses measurement metrics used to evaluate the performance of deepfake detection methods, including the confusion matrix. Overall, the review provides valuable insights into the current state of deepfake detection research and offers guidelines for future studies in this area.

Blockchain Based Provenance Solution for Combating Deepfake Videos

It presents a blockchain-based solution using Ethereum smart contracts to combat deepfake videos by tracing and tracking the provenance of digital content, specifically focusing on video content. The proposed system includes a reputation system for artists, a mechanism to trace videos back to their original publishers, and a security analysis ensuring integrity, accountability, authorization, availability, and non-repudiation. The use of IPFS for storing content and smart contracts for maintaining data integrity is highlighted. The paper also discusses the importance of establishing a trusted data provenance for digital content to prevent the spread of fake content. The solution framework is generic and can be applied to various forms of digital media. The paper provides system architecture, design details, implementation of smart contracts, testing procedures, cost estimation, and security analysis.

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 5, May 2024

DeepVision: Deepfakes Detection Using Human Eye Blinking Pattern

Deepfakes Detection Using Human Eye Blinking Pattern" presents a novel approach to detecting Deepfakes by analyzing changes in the pattern of blinking, a natural and involuntary action. The study conducted experiments to compare the eye blinking patterns in Deepfakes videos with those in real human videos under various conditions such as gender, age, activity, and time of day.

The research found that Deepfakes videos exhibited significantly lower blink frequencies and durations compared to real human videos. By monitoring these differences in eye blinking patterns, the DeepVision algorithm was able to accurately identify Deepfakes videos. The study highlights the importance of integrity verification in the face of increasing misuse of Deepfakes technology for spreading propaganda and fake news.

Overall, it provides valuable insights into a promising method for detecting Deepfakes through the analysis of human eye blinking patterns, offering a potential solution to combat the negative implications of Deepfakes in various domains.

Detecting Partial Spooling in Speech: The PartialSpoof Database and Enhanced Countermeasures

ThePartialSpoof Database and countermeasures for detecting short fake speech segments embedded in an utterance. The research introduces a new spoofing scenario called Partial Spoof (PS), where synthesized or transformed speech segments are inserted into genuine utterances. Existing countermeasures are effective in detecting fully spoofed utterances but need adaptation or extension to address the PS scenario.

The study presents the construction of the PartialSpoof database, which contains spoofed speech for model training. Attackers can create partially-spoofed audio by inserting or replacing segments of spoofed speech generated by Text-to-Speech (TTS) or Voice Conversion (VC) systems into natural utterances. The database includes spoofed speech with different proportions of generated audio segments within a single utterance, termed as the "intra-speech generated segment ratio."

Furthermore, mentions the support received from various organizations for the research, such as the Japanese-French joint national VoicePersonae project, JST CREST, ANR, MEXT KAKENHI Grants, and Google AI for Japan program. The study is partially supported by these entities to encourage research in countermeasure solutions against spoofing attacks in speech applications.

Overall, it provides insights into the development of the PartialSpoof Database, the challenges posed by the PS scenario, and the importance of enhancing countermeasures to detect and combat short fake speech segments embedded in genuine utterances.

III. CONCLUSION

A comprehensive survey on DeepFake detection methods, focusing on the use of Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) for identifying manipulated media. The highlights challenges in creating generalized DeepFake detection models due to the dynamic nature of DeepFake generation techniques and the lack of diverse datasets for training. It also discusses the vulnerability of current detection models to adversarial attacks and the need for robust training strategies to combat these challenges.

It presents various state-of-the-art methods for detecting DeepFakes, emphasizing the effectiveness of deep learning techniques, particularly CNNs, in identifying manipulated content. It also introduces a novel approach using a unified Gabor function to enhance the adaptability of CNNs for DeepFake detection. Additionally, the survey touches on the use of blockchain technology for establishing proof of authenticity of digital content and the development of a method to detect DeepFakes by analyzing eye blinking patterns.

In conclusion, underscores the importance of continuous research and development in DeepFake detection to ensure data integrity and combat the potential misuse of DeepFake technology. It calls for the creation of more robust detection models that cangeneralize to unknown types of manipulations and withstand adversarial attacks. The survey also suggests that future work should focus on improving the generalization ability of models, developing more sophisticated detection techniques, and exploring the use of blockchain and biometric data for authentication and detection purposes

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 5, May 2024

REFERENCES

[1] H.Farid, Image forgery detection, IEEE Signal Process. Mag., vol. 26, no. 2, pp. 1625, Mar. 2009.
[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in Proc. Adv. Neural Inf. Process. Syst., vol. 27, 2014, pp. 19.
[3] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in Proc. ICML Workshop Unsupervised Transf. Learn., 2012, pp. 3749.
[4] T. Karras, S. Laine, and T. Aila, A style-based generator architecture for generative adversarial networks, in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 44014410.
[5] Y. Mirsky and W. Lee, The creation and detection of deep fakes: A survey, ACM Comput. Surv., vol. 54, no. 1, pp. 141, Jan. 2022. [6] S.L. Strunic, F. Rios-Gutierrez, R. Alba-Flores, G. Nordehn, and S. Bums,
[6] Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, Celeb-DF: A large-scale challenging dataset for DeepFake forensics, in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 32043213.
[7] S.Agarwal and H. Farid, Protecting world leaders against deepfakes, in Proc. IEEE/CVF Conf.Comput.Vis.PatternRecognit.(CVPR)Workshops, Jun. 2019, pp. 3845.
[8] D. M. Montserrat, H. Hao, S. K. Yarlagadda, S. Baireddy, R. Shao, J. Horvath, E. Bartusiak, J. Yang, D. Guera, F. Zhu, and E. J. Delp, Deepfakes detection with automatic face weighting, in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2020, pp. 19.
[9] Y. Li, M.-C. Chang, and S. Lyu, Inictu oculi: Exposing AI created fake videos by detecting eye blinking, in Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS), Dec. 2018, pp. 17.
[10] T. Jung, S. Kim, and K. Kim, DeepVision: Deepfakes detection using human eye blinking pattern, IEEE Access, vol. 8, pp. 21693536, 2020.
[11] Y. Mirsky and W. Lee, The creation and detection of deepfakes: A survey, ACM Comput. Surv., vol. 54, no. 1, pp. 141, Jan. 2021.
[11] S. Lawrence, C. L. Giles, A. Chung Tsoi, and A. D. Back, Face recognition: A convolutional neural-network approach, IEEE Trans. Neural Netw., vol. 8, no. 1, pp. 98113, Jan. 1997.
[12] J. G. Lawrenson, R. Birhah, and P. J. Murphy, Tear- lm lipid layer mor phology and corneal sensation in the development of blinking in neonates and infants, J. Anatomy, vol. 206, no. 3, pp. 265270, Mar. 2005.
[13] A. J. Zametkin, J. R. Stevens, and R. Pittman, Ontogeny of spontaneous blinking and of habituation of the blink re ex, Ann. Neurol., vol. 5, no. 5, pp. 453457, May 1979.
[14] P. J. DeJongandH.Merckelbach, Eyeblink frequency, rehearsal activity, and sympathetic arousal, Int. J. Neurosci., vol. 51, nos. 12, pp. 8994, Jan. 1990.
[15] J. Oh and J. Jeong, Potential signicance of eyeblinks as a behavior marker of neuropsychiatric disorders, Korean J. Biol. Psychiatry, vol. 19, no. 1, pp. 920, 2012.
[16] E. Ponder and W. P. Kennedy, On the act of blinking, Quart. J. Exp. Physiol., vol. 18, no. 2, pp. 89110, Jul. 1927.