As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology
Volume 10, Issue 04, pages 365-369.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol10issue04.html
DOI: https://doi.org/10.46501/IJMTST1004056

Unmasking Deep Fakes: Detection of Manipulated Images and Videos through Advanced Deep Learning Techniques

Y.Manas Kumar [1] , Madapati Yoshitha[2], Velampalepu J S D Satya Sai[2], Siddhanathi Surya Sailaja[2], Tallapudi Balaji[2], Chatrathi Sai Surya[2]

[1]Assistant Professor, Department of Computer Science Engineering, Pragati Engineering College, Surampalem , Andhra Pradesh, India.
[2]Department of Computer Science Engineering, Pragati Engineering College, Surampalem , Andhra Pradesh, India.

To Cite this Article
Y.Manas Kumar , Madapati Yoshitha, Velampalepu J S D Satya Sai, Siddhanathi Surya Sailaja, Tallapudi Balaji, Chatrathi Sai Surya, Unmasking Deep Fakes: Detection of Manipulated Images and Videos through Advanced Deep Learning Techniques, International Journal for Modern Trends in Science and Technology, 2024, 10(04), pages. 365-369. https://doi.org/10.46501/IJMTST1004056

Article Info
Received: 06 April 2024; Accepted: 18 April 2024; Published: 26 April 2024.

ABSTRACT

Deepfakes are digitally modified approaches that employ deep learning to create false (misleading) pictures and movies. Identifying deepfake photos is the most challenging aspect of locating the original. Because of the growing notoriety of deep fakes, recognizing authentic photos and videos is becoming increasingly important in detecting modified content. This study investigates and tests with several strategies for detecting fraudulent and authentic photos and videos. Deep fakes were identified using the Convolutional Neural Network (CNN) method known as Inception Net. This paper includes a comparative comparison of several convolutional networks. This study uses the Kaggle dataset with 401 videos of the train samples and 3746 pictures created by the augmentation method. The findings were analysed using measures such as accuracy and confusion matrix. The suggested approach gives improved results in terms of accuracy (93% on spotting deep fake pictures and movies).

Keywords: Deepfake, Inception net, CNN (Convolutional Neural Network), Vision Transformers

1. INTRODUCTION

Deepfake films have become increasingly prevalent as cell phones and social media networks have grown in popularity. These technologies have produced bogus news and films, which are deemed detrimental to society. Terrorist groups also use deceptive photos and films to shame the people and the globe, as well as to threaten the nation. Increased virtualization and globalization shrunk the world while also inviting nonstate dangers to the nation through the use of bogus films, radicalizing individuals from other religions, and disseminating the agenda. Many high-profile persons fell into this trap and endured numerous troubles as a result of phony photographs and films.

The face is the most distinguishing characteristic of humans. With the rapid growth of face blend innovation, the security danger posed by face control is getting more serious. Human faces can frequently shift based on someone's appearance, which can appear as real and authentic human faces due to several computations that rely on profound gaining innovation. It is a rising category of fraudulent insights advancement in which anyone's face may be identified as someone's actual face [3]. Deepfake material is spreading more quickly than previously in the twenty-first century. Because deepfakes are getting more popular, approaches for identifying fake films that appear to be real are becoming increasingly vital. In this publication, we will examine alternative technologies that may be utilized to identify deepfake pictures. Over the last several decades, the emergence of social networking sites and smartphone culture has increased the appeal of digital images and movies.

2. LITERATURE SURVEY

An exhaustive literature review was conducted on related publications concerning deep fake detection models and ways for improving existing systems. A literature review is conducted on various data mining approaches. The following section summarizes the related works that are discussed in this study. Nishat Tasnim Roza et al. (2021) conducted a comparative investigation of the deepfake image identification approach utilizing a convolutional neural network. In their research, they described the following. Human faces have extremely distinguishing traits. Deepfake videos/images employ AI technology to replace a person's facial features. Deepfake pictures were not apparent to normal eye vision owing to the collapse of the pixel, skin tones, and face forms of images, These are manufactured visual abnormalities. Deepfake may be created from photos, videos, and sounds. Advancements in technology have made deepfake photos nearly identical to natural ones. As a result, individuals worldwide face inescapable challenges. In 2016, Kaipeng Zhang et al. introduced MTCN for collaborative face detection and alignment. In their research, they observed the following. Detecting and aligning fraudulent faces is crucial for face systems, including reputation and feature analysis.

However, the large visible versions of faces, which include occlusions, massive posture versions, and excessive lighting, create extremely demanding conditions for those responsibilities in genuine global programs. The cascade face detector used Haar-Like functions and AdaBoost to educate cascaded classifiers, resulting in genuine overall performance and real-time effectiveness. However, several studies indicate that this detector can decline significantly in real-world applications with huge visible versions of human faces, despite more advanced functions and classifiers. Aside from the cascade structure, integrate deformable elements models (DPM) for identifying faces and gain excellent overall performance. However, they need high estimate rates and may necessitate costly annotation throughout the educational stage. Convolutional neural networks have recently made significant advances in a variety of computer creative and prophetic tasks, including image classification and facial recognition. Several CNN-based face identification algorithms have been presented in recent years, motivated by CNN's superior performance in computer creative and prophetic tasks. Yang et al. train deep convolution neural networks for the typical facial repute to acquire excessive reactions in face locations, providing candidate home windows of faces. However, because to its complex CNN structure, this approach is time-consuming in practice. Li et al. employ cascaded CNN for face detection; however, they propose bounding field calibration from face detection with a higher estimation rate and overlook the underlying relationship between facial landmark localization and bounding regression [3]. Christian Szegedy et al. (2015) advocated going further using convolutions. In their research, they described the following. They stated that in the last three years, their item classification and detection abilities have significantly improved as a result of breakthroughs in deep learning about it. Statistics may be maximizing the development isn't necessarily the final result of more effective technology, bigger data files, and greater fashions, but rather a consequence of newest concepts, methods and progressed community architectures. No new statistical reasserts have been employed, for example, by way of the top entries inside the opposition, unless the same class of opposition was used for identification. According to their findings, Google Net must be proposed to employ ten times less parameters than Krizhevsky's successful structure from years ago, while being significantly more accurate. On the item identification front, the highest benefits are no longer from bigger utility and a wider network of deep, but from deep architectures and old laptops, much as the CNN set of rules via Girshick. Another first-rate issue, along with the continuation of cellular and calculated that is incorporated, performance of the method, in particular of the strength & reminiscence profits. The primary concerns are with the highly organized arrangement of the note covered issue rather than the maximum in the trial. Fashions have been made to preserve estimating finance by 1.6 multiplication of billions that gives by right time, so that it is no longer simply a novel educational tool; it will be utilized for real-world, global usage at a fair cost, even on big data sets. In this remark, individuals can pay attention to the well-known they want to develop deeper network meme along with green neural, which is a deeper community structure in laptop vision, in a community article by Lin et al. The term "deep" has a specific connotation in this context: First and foremost, there is the sensation of a whole new stage of business, as well as the more immediate experience of a multiplied community. Thus, the majority of the literature study focuses on approaches for retrieving data from Twitter and news articles, converting that data into the necessary format, and applying operations to determine the user's purpose. However, they have not concentrated on the algorithm for categorizing the topic of news tweets.

3. SYSTEM ANALYSIS

A. EXISTING SYSTEM

Convolutional Neural Networks (CNNs) are a popular component of existing deepfake detection algorithms, which may contain pre-trained Inception Net models. Diverse datasets are required for both testing and training in system development. Certain systems incorporate temporal consistency checks, audio analysis, and facial landmarks to improve detection accuracy. like applications frequently make use of open-source technologies and frameworks like as PyTorch and TensorFlow. Integrating the model into a video processing pipeline enables real-time deepfake detection. Research and development must continue since deepfake production techniques are always developing. Working with subject matter experts and leveraging relevant resources can assist to fine-tune and improve the system's operation.

DISADVANTAGES OF THE EXISTING SYSTEM

1. **Adversarial Attacks**: Deepfake makers constantly modify their tactics to avoid detection, resulting in a cat-and-mouse game. Current systems may struggle to keep up with the increasing deepfake creation techniques.
2. **Generalization:** Deepfake detection methods may not generalize well to new and previously unknown forms of deepfakes, especially if the training dataset is not sufficiently varied.
3. **Computational Intensity**: Deepfake detection is computationally demanding, making it difficult to execute real-time detection on resource-constrained devices.
4. **False Positives and Negatives**: Existing algorithms may generate false positives (misclassifying actual information as deepfakes) and false negatives (failing to detect advanced deepfakes).
5. **Lack of Real-Time Processing**: The rapid spread of deepfake material necessitates real-time detection in video streams, which many systems are not designed for.
6. **Privacy Concerns**: Some deepfake methods for identification may use intrusive tactics, such as analysing biometric characteristics, causing privacy problems and ethical dilemmas.
7. **Limited Data**: The availability of high-quality labelled datasets for training deepfake detection models might be limited, affecting the model's performance.
8. **Resource-Intensive Training**: Large computer resources may be required to train deepfake detection algorithms, which not all businesses or researchers may have access to.
9. **Model Interpretability**: Some deepfake detection models are sophisticated and lack interpretability, making it difficult to grasp how they reach their choices.
10. **Domain Shift**: Models trained on a single dataset may not perform well when applied to another domain or situation, such as changes in lighting or camera quality.
11. **Ethnic and Gender Bias**: Some deepfake detection techniques may be biased in their performance, disproportionately impacting persons from specific ethnic or gender groupings.

B. PROPOSED SYSTEM

The suggested system for "Deepfake Face Detection Using Deep Inception Net Learning Algorithm" seeks to overcome the limitations of existing methods. We will use an upgraded deep learning technique that combines Inception Net with other cutting-edge CNN architectures. To boost generalization, we will compile a broad and large dataset comprising both deepfake and genuine material. To improve detection accuracy, our system will use multimodal analysis, which includes face landmarks and auditory characteristics. Real-time processing skills will be prioritized to enable speedy detection of deepfake information in video streams. To reduce biases, we shall also consider model explainability and fairness. Regular upgrades and close engagement with the research community will assure our system's efficacy against advancing deepfake technologies, while respecting privacy and ethical issues.

ADVANTAGES OF THE PROPOSED SYSTEM

1. **Enhanced Detection Accuracy**: By merging Inception Net with other sophisticated CNN architectures and applying multi-modal analysis, the system enhances accuracy in detecting deepfake material, minimizing both false positives and false negatives.
2. **Real-Time Processing**: The system is built for real-time deepfake detection, therefore it is appropriate for applications that demand quick recognition of altered information, such as live video broadcasts or social media monitoring.
3. **Generalization and Robustness**: With a wide and vast dataset, the system is better suited to generalize to new and growing forms of deepfakes, increasing its robustness and flexibility.
4. **Model Explainability and Fairness**: The method emphasizes model interpretability and fairness, addressing ethical problems and biases in deepfake detection, making it more visible and equitable.
5. **Ongoing Adaptation**: Regular updates and collaboration with the research community guarantee that the system stays successful against the newest deepfake generation techniques, making it a proactive defense against the continuously changing world of deepfake material.

4. SYSTEM DESIGN

SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture.

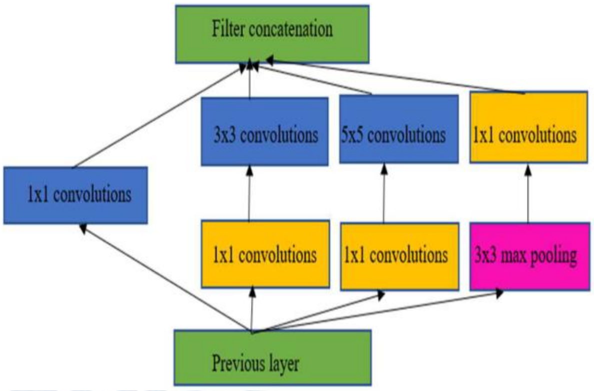


Fig 1. Methodology followed for proposed model

5. SYSTEM IMPLEMENTATION

MODULES

Data Preprocessing: Data gathering and curation of various deepfakes and actual material. Data augmentation increases the dataset's variety. Resizing and normalizing are examples of picture and video preprocessing.

Feature Extraction: Using deep learning architectures such as Inception Net and other CNN models for feature extraction. Extracting face landmarks and auditory characteristics to increase detection accuracy.

Model Training: Using the pre-processed dataset, we trained the deepfake detection model. Fine-tuning and optimizing the chosen CNN architectures. Ensure the model's generalization to diverse deepfake settings.

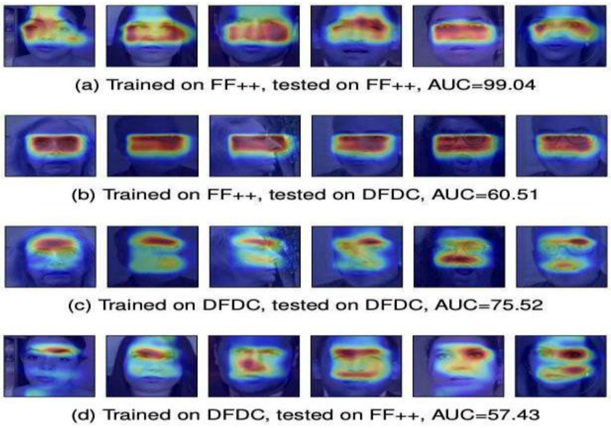
Real-Time Processing: creating a pipeline for real-time video processing to detect deepfakes in real time. Make an intuitive user interface for in-the-moment communication.

Ethical and Fairness Considerations: Implementing fairness and bias detection techniques to ensure that the system operates fairly for all demographic groups.

Monitoring and Reporting: A dashboard or user interface for tracking system performance and the status of the fire detection procedure is included in this module. Additionally, it could provide logs and reports for analysis, evaluation, and potential system updates.

6. RESULTS AND DISCUSSION

The Deepfake Detection Challenge (DFDC) and Face Forensics datasets were used in this study. This dataset contains 30 GB of movies (5000 total) for the face forensic collection and 480 GB of films (124000 total) for the DFDC dataset. For each stage, multiple copies of the data set are made.



(a) Trained on FF++, tested on FF++, AUC=99.04

(b) Trained on FF++, tested on DFDC, AUC=60.51

(c) Trained on DFDC, tested on DFDC, AUC=75.52

(d) Trained on DFDC, tested on FF++, AUC=57.43

Fig : Dataset: Version 0 initial dataset to Version 3 final dataset

Face forensics++ tests were also undertaken to compare various created photos from diverse data sources.

Except for Deep fakes, the architecture employed in face forensic++ outperforms the conventional designs for many sub-datasets. This is most likely the outcome of the network's enhanced capacity to generalize about very specific deepfakes.

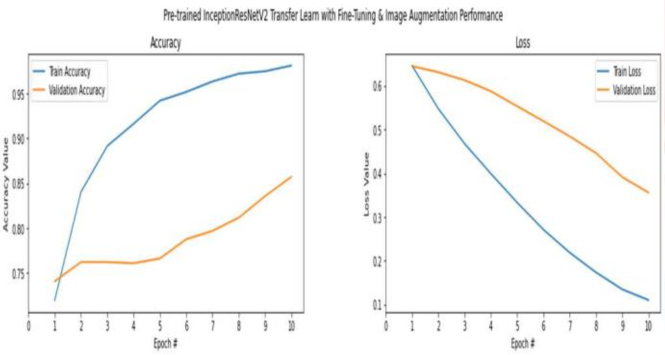


fig 2. Comparison of models in terms of Accuracy

7. CONCLUSION AD FUTURE WORK

In this work, the Inception net architecture was used to identify fake faces. Different types of transitions in real images are used, together with test parameters such as the number of key points in the images, the comparison rate, and the time required for each algorithm to perform. This paper shows that the overall accuracy of the DFDC dataset is 93%. The results shown here can classify deepfake recordings from a variety of sources using different convolutional layers. As a result, the contribution of this paper will inevitably help to reduce the prevalence of fake records and extortion in our society. The proposed work was completed faster than the existing work, and its identification of fake and real images was quite effective For the recommended job, the DFDC dataset has a 93% accuracy rate. It might be expanded in the future to detect deepfake face photographs using different classifiers and distance metrics.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE signal processing letters, 23(10), 1499-1503.

[2] Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. "Inceptionism: Going deeper into neural networks." (2015).

[3] Badale, Anuj, et al. "Deep fake detection using neural networks." 15th IEEE international conference on advanced video and signal-based surveillance (AVSS). 2018.

[4] Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929 (2020).

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

[6] Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448-456). PMLR.

[7] Chen, Chun-Fu Richard, Quanfu Fan, and Rameswar Panda. "Crossvit: Cross-attention multi-scale vision transformer for image classification." Proceedings of the IEEE/CVF international conference on computer vision. 2021.

[8] Heo, Young-Jin, et al. "Deepfake detection scheme based on vision transformer and distillation." arXiv preprint arXiv:2104.01353 (2021).

[9] Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE signal processing letters 23.10 (2016): 1499-1503.,

[10] Kaggle,https://www.kaggle.com/competitions/deepfake-detection challenge/data