Deepfake Video Detection

Team details: Md Mufeez Ur Rahman
N Sai Krishna

Project Guide : Dr. R. Sateesh Kumar

Abstract:

The growing computation power has made the deep learning algorithms so powerful that creating a indistinguishable human synthesized video popularly called as deep fakes have became very simple. Scenarios where these realistic face swapped deep fakes are used to create political distress, fake terrorism events, blackmail peoples are easily envisioned. In this work, we describe a new deep learning-based method that can effectively distinguish AI-generated fake videos from real videos. Our method is capable of automatically detecting the replacement and reenactment deep fakes. We are trying to use Artificial Intelligence(AI) to fight Artificial Intelligence(AI).

Our system uses a Res-Next Convolution neural network to extract the frame-level features and these features and further used to train the Long Short Term Memory(LSTM) based Recurrent Neural Network(RNN) to classify whether the video is subject to any kind of manipulation or not, i.e whether the video is deep fake or real video. To emulate the real time scenarios and make the model perform better on real time data, we evaluate our method on large amount of balanced and mixed data-set prepared by mixing the various available data-set like Face-Forensic++[1], Deepfake detection challenge[2], and Celeb-DF[3]. We also show how our system can achieve competitive result using very simple and robust approach.

Introduction

The advent of deep learning algorithms has revolutionized various domains, including the generation of synthetic media content. However, the proliferation of deepfake technology poses serious challenges, particularly in the context of misinformation, privacy violations, and malicious intent. As the creation of indistinguishable human-synthesized videos becomes increasingly simple, there is a pressing need for robust methods to detect and combat the spread of deepfakes. In response to this challenge, our project aims to leverage AI-based techniques to develop a deepfake detection system capable of distinguishing between authentic and manipulated videos.

Literature Survey

The literature survey conducted in the domain of deepfake detection reveals a multitude of approaches and techniques employed to address the growing threat posed by AI-generated fake videos. Existing systems often utilize variations of deep learning architectures, including CNNs and RNNs, to extract features and classify videos based on their authenticity. However, many of these methods face limitations in terms of scalability, performance, and adaptability to real-world scenarios. Therefore, there is a need for innovative approaches that can overcome these challenges and provide effective solutions for deepfake detection.

Problem statement

The proliferation of deepfake technology has led to a surge in the creation and dissemination of AI-generated fake videos, posing serious threats to various aspects of society, including politics, security, and privacy. The lack of robust methods for detecting deepfakes makes it difficult to combat the spread of misinformation and prevent potential harm to individuals and organizations. Therefore, the primary objective of our project is to develop an AI-based deepfake detection system capable of accurately distinguishing between genuine and manipulated videos in real-time.

Motivation

- The motivation behind choosing the problem statement lies in the urgent need to address the growing threat posed by deepfake technology.

- By developing an effective deepfake detection system, we aim to mitigate the harmful impacts of AI-generated fake videos on society, including political manipulation, misinformation, and privacy violations.

- Moreover, the opportunity to leverage AI to combat AI-driven threats aligns with the broader goal of harnessing technology for the betterment of society and promoting ethical AI practices.

Functions

Feature extraction using Res-Next CNN.

Training LSTM-based RNN for video classification.

Evaluation of the system on diverse datasets.

Real-time deepfake detection.

Integration of the system into existing media verification platforms.

System requirements

- Hardware: GPU-accelerated computing hardware for efficient training and inference.
- Software: Python programming language, TensorFlow or PyTorch for deep learning frameworks, libraries for video processing and analysis, and necessary dependencies.

Conclusion

Our project presents a novel approach to deepfake detection by leveraging AI-based techniques to combat the proliferation of AI-generated fake videos. By combining Res-Next CNN for feature extraction and LSTM-based RNN for classification, we achieve competitive results in distinguishing between authentic and manipulated videos. The system's simplicity and robustness make it well-suited for real-world applications, offering practical solutions to mitigate the harmful impacts of deepfake technology on society.

References

- Face Forensic++
- Deepfake detection challenge
- Celeb-DF

Thank You

Presented by : Md Mufeez Ur Rahman
N Sai Krishna