Deepfake Detection Engine using MTCNN and Facenet_pytorch

Deepfake Detection Using Multi-Task Cascaded Convolutional Networks and InceptionResnetV1

Project giude

Mrs Saranya N

project overview

Abstract

Deepfake technology, which involves the creation of highly realistic synthetic media using machine learning, poses significant risks to information integrity and public trust. This project aims to develop a robust deepfake detection engine leveraging Multi-Task Cascaded Convolutional Networks (MTCNN) for face detection and alignment, and InceptionResnetV1, provided by the facenet_pytorch library, for image classification. Our solution integrates these models into a seamless pipeline to accurately identify manipulated media, enhancing security measures in digital communications and content verification processes.

project overview

Introduction

The proliferation of deepfake technology has led to a surge in digitally manipulated content that can deceive viewers by altering or fabricating visual and audio data. This has serious implications for privacy, security, and the integrity of information disseminated online. Traditional detection methods are becoming increasingly inadequate as deepfake algorithms advance. This project addresses the urgent need for effective detection mechanisms by utilizing state-of-the-art deep learning models for face detection and recognition, ensuring high accuracy in identifying synthetic media.

Base paper details ↘

2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech 2021)

Fake-image detection with Robust Hashing

*

1st Miki Tanaka
Tokyo Metropoliltan University
Tokyo, Japan
tanaka-miki@ed.tmu.ac.jp

2nd Hitoshi Kiya
Tokyo Metropoliltan University
Tokyo, Japan
kiya@tmu.ac.jp

Problem Definition

The primary challenge is to develop an automated system that can accurately and efficiently detect deepfakes amidst the growing sophistication of synthetic media generation techniques. This involves identifying subtle anomalies in facial features and movements that are indicative of manipulation, while minimizing false positives and ensuring robustness across diverse datasets. Additionally, the detection system must be scalable, capable of processing large volumes of media content in real-time or near-real-time to be practical for widespread use in social media, news, and security applications. The model needs to generalize well across different types of deepfakes, including those generated by various algorithms and techniques, ensuring that it can detect new and evolving forms of synthetic media. Achieving high precision in anomaly detection is crucial, as it requires sophisticated feature extraction and analysis techniques to accurately identify the subtle differences in facial features, expressions, and movements that distinguish real content from deepfakes. Furthermore, the system must handle diverse data sources and variations in video quality, lighting conditions, and facial appearances to maintain high detection accuracy in various real-world scenarios.

Objectives

- Objective 1: To create a reliable deepfake detection engine using MTCNN for face detection and alignment.
- Objective 2: To employ InceptionResnetV1 from the facenet_pytorch library for classifying images as real or fake with high accuracy.
- Objective 3: To integrate the detection models into a user-friendly interface for practical deployment and usage.
- Objective 4: To validate the model's performance on various datasets and ensure its adaptability to new types of deepfake content.

Innovation and Methodology Employed in the Project

Innovation: Combining MTCNN and InceptionResnetV1 to leverage their strengths in face detection, alignment, and classification, resulting in a highly accurate and efficient deepfake detection system.

- Data Collection: Curate a comprehensive dataset containing both real and deepfake images.
- Preprocessing: Use MTCNN to detect and align faces in the collected images, ensuring uniformity in input data.
- Model Training: Fine-tune InceptionResnetV1 on the preprocessed dataset to improve its classification performance.
- Integration: Develop a pipeline that seamlessly integrates MTCNN and InceptionResnetV1, facilitating real-time deepfake detection.
- Deployment: Create a user-friendly interface using Jupyter Notebook and VSCode for developers and end-users to easily interact with the detection engine.
- Validation: Test the system on various datasets to assess its accuracy, robustness, and adaptability to different types of synthetic media.

Literature survey

- Deepfake Detection Techniques: Review of various methods including traditional image analysis techniques and modern machine learning approaches.
- MTCNN for Face Detection: Analysis of the Multi-Task Cascaded Convolutional Networks and their effectiveness in detecting and aligning faces in images.
- Facenet_pytorch and InceptionResnetV1: Overview of facenet_pytorch's implementations and the use of InceptionResnetV1 for image classification tasks.
- Challenges in Deepfake Detection: Discussion of the current challenges faced in detecting deepfakes, including the continuous improvement of deepfake generation techniques.

Conclusion

This deepfake detection engine leverages cutting-edge machine learning techniques to address the growing threat of synthetic media. By combining MTCNN for face detection and InceptionResnetV1 for classification, the system ensures high accuracy and efficiency in identifying manipulated content. The project's innovation lies in its integration of these advanced models, providing a powerful tool for maintaining the integrity and security of digital media.