Deepfake Detection Website

A WEB-BASED TOOL FOR DETECTING FAKE VIDEOS

BY

SIDDHARTH GUPTA

MALLIKARJUNE KOLI

APRIL 2024

Introduction

What are Deepfakes?

Deepfakes are AI-generated media, such as videos, images, or audio recordings, that are manipulated to present false or misleading information. They are primarily created using advanced machine learning techniques like Generative Adversarial Networks (GANs) and autoencoders, which allow for highly realistic and convincing alterations of real content. As deepfake technology continues to evolve, it poses significant challenges by fueling misinformation, compromising personal privacy, and eroding public trust in digital media

Problem Statement

The Growing Threat of Deepfakes

Deepfakes are increasingly being exploited for malicious purposes such as financial fraud, identity theft, and political manipulation. Their realistic nature makes them difficult to detect with the naked eye, and currently, there is a significant lack of accessible and reliable tools for identifying deepfakes. As the technology becomes more widespread and sophisticated, there is a growing demand for automated, AI-driven detection methods to help combat their misuse and protect individuals and institutions from potential harm..

Solution Overview

Deepfake Detection Website

This project is a web-based platform designed to detect deepfake videos with ease and efficiency. It leverages powerful technologies such as **TensorFlow** for deep learning and **OpenCV** for video processing to analyze and identify manipulated content. The website features a **user-friendly interface**, making it accessible even to non-technical users, allowing anyone to upload videos and receive reliable detection results within minutes.

Code/Tool Breakdown - Backend

Backend Development Using Flask

The backend of the Deepfake Detection Website is built using **Flask**, a lightweight Python web framework. Flask handles the **web server**, manages **video uploads**, and defines routes for processing and analyzing the uploaded content. The application includes specific endpoints to trigger **deepfake detection algorithms** and return the results to the user interface..

Example:

```
@app.route('/upload', methods=['POST'])
def upload():
   video = request.files['video']
   video.save('static/input.mp4')
   result = detect_deepfake('static/input.mp4')
   return render_template('result.html',
prediction=result)
```

Code/Tool Breakdown - Deepfake Detection

Deepfake Detection Model

The detection system uses **TensorFlow** to load a **pre-trained Convolutional Neural Network (CNN)** model, such as **XceptionNet**, which is known for its high performance in image classification tasks. The video is processed frame by frame, with each frame undergoing a sequence of steps: **face detection**, **image preprocessing**, and finally **classification** to determine the likelihood of manipulation. This layered approach allows for accurate and efficient detection of deepfake content within video files.

Real-World Use Cases - Media

Use Case 1: Journalism

In the field of journalism, deepfake detection plays a critical role in **verifying the authenticity of video news reports** before publication. With the rise of manipulated media, journalists and news organizations face the challenge of distinguishing real footage from fabricated content. Implementing deepfake detection tools helps **prevent the spread of misinformation**, ensuring that the public receives accurate and trustworthy news.

Real-World Use Cases - Legal

Use Case 2: Digital Forensics

In digital forensics, deepfake detection is essential for **verifying the authenticity of video evidence** used in criminal investigations. Law enforcement agencies and forensic analysts rely on accurate video analysis to build credible cases. By using deepfake detection technology, they can **ensure the integrity of digital media presented in courtrooms**, helping to prevent wrongful convictions and uphold justice.

Real-World Use Cases - Social Media

Use Case 3: Content Moderation

Deepfake detection technology is vital for **automatically flagging or removing manipulated content** on social media platforms. As deepfakes can be used to spread harmful, misleading, or abusive material, real-time detection tools help platforms **moderate content more effectively**. This not only reduces the risk of viral misinformation but also **protects users from the negative impact** of deceptive media.

Future Enhancements - Real-Time Deepfake Detection

Real-Time Deepfake Detection

Expanding deepfake detection capabilities to work in **real-time** enables analysis of **live video streams**, such as video calls and livestreams. This feature is crucial for identifying and stopping manipulated content as it happens, rather than after the fact. Additionally, integrating **voice deepfake detection** helps uncover synthetic audio in real-time conversations, offering a more comprehensive solution to detect both visual and auditory deception.

Future Enhancements - Multi-Modal Detection

Multi-Modal Deepfake Detection

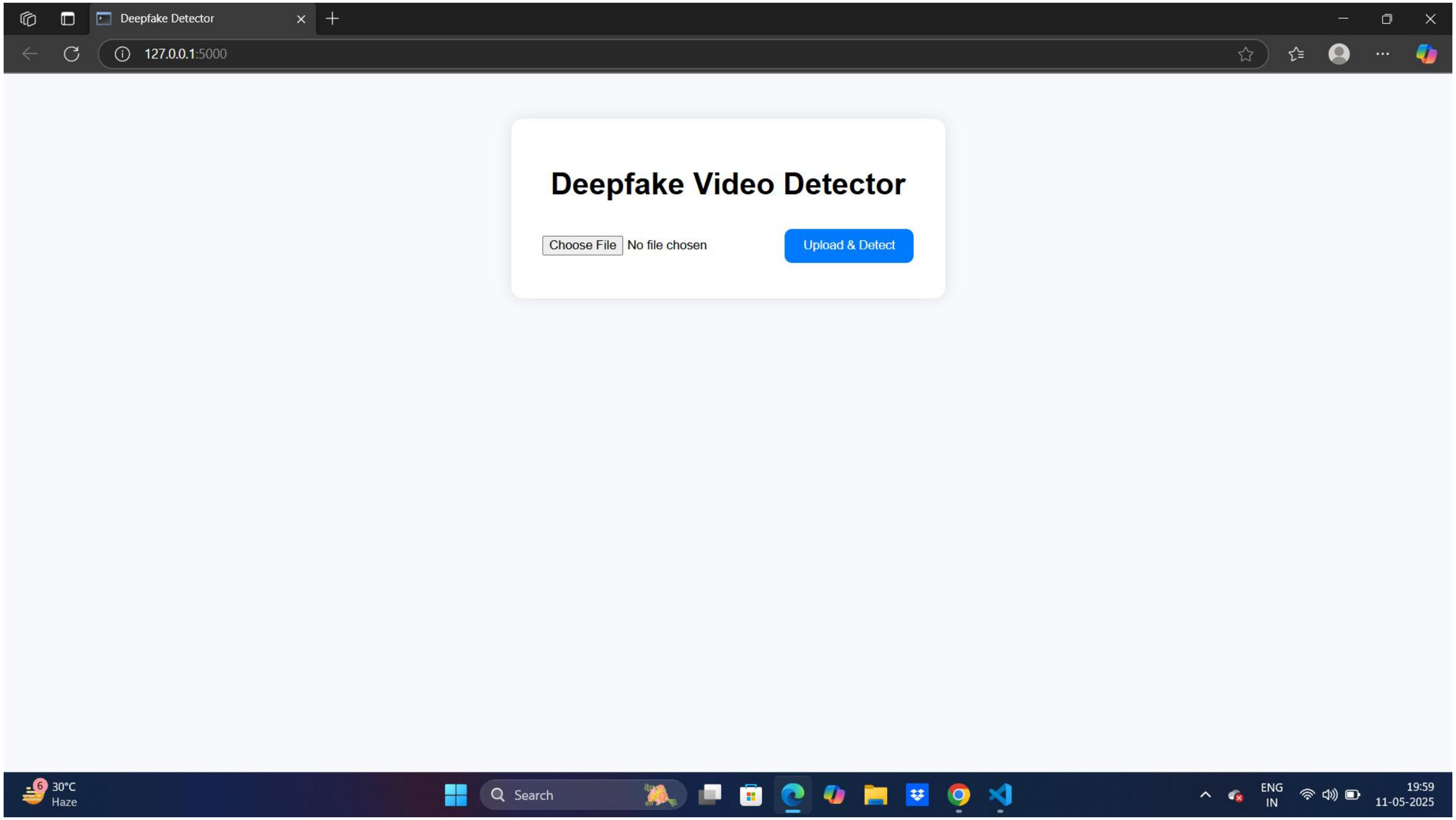
To enhance detection accuracy, **multi-modal deepfake detection** combines both **video and audio analysis**. By integrating advanced techniques such as **facial recognition**, **voice synthesis detection**, and **behavioral analysis**, the system can cross-verify inconsistencies between visual and auditory components. This multi-layered approach provides a more robust and reliable method of identifying deepfakes, reducing the chances of false positives and increasing detection precision.

Future Enhancements - Mobile and Cloud

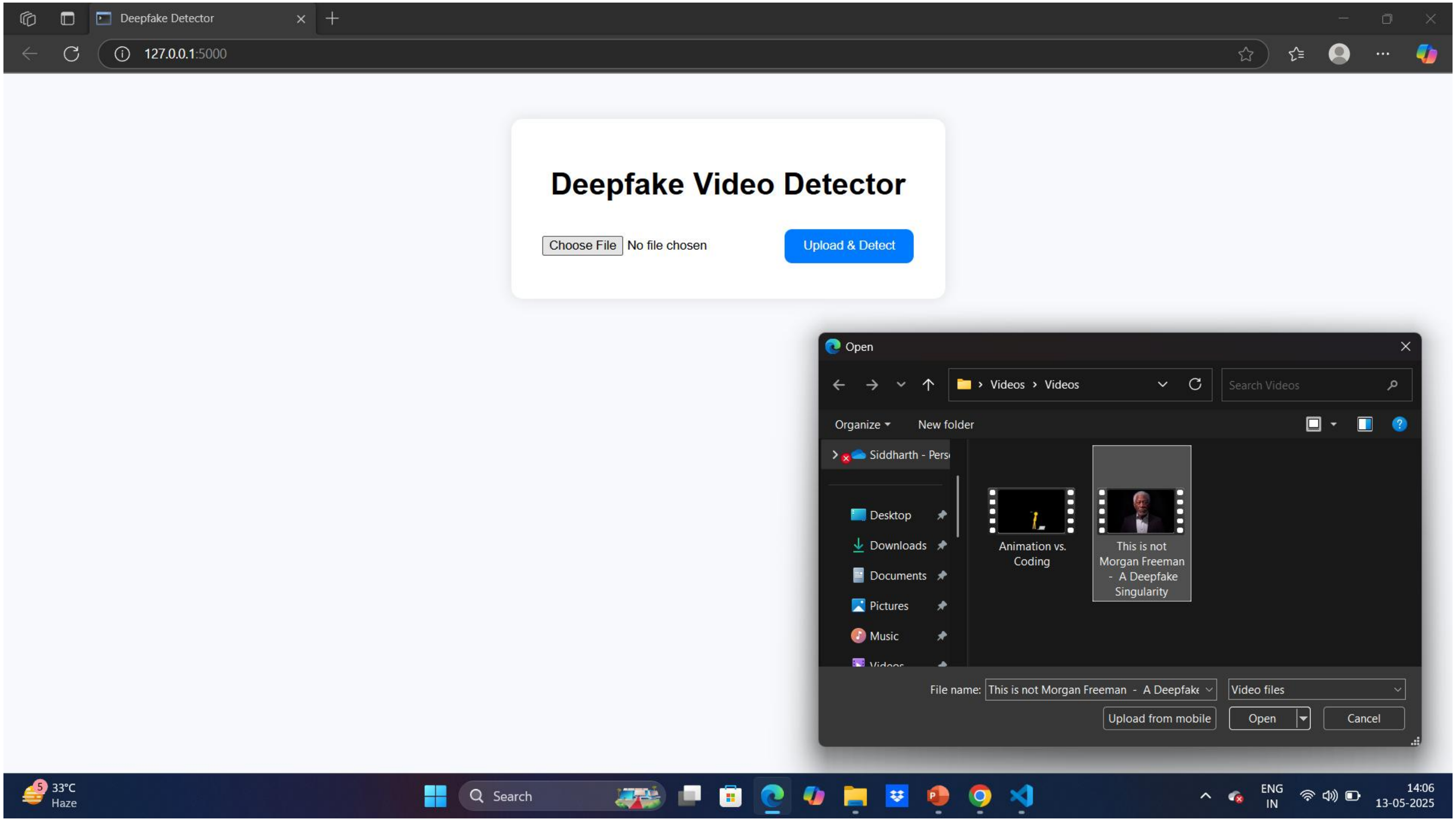
Mobile and Cloud Deployment

To increase accessibility, we are developing **mobile versions** of the deepfake detection tool for both **Android** and **iOS** platforms, enabling users to detect deepfakes on-the-go. Additionally, the application will be **hosted on the cloud**, ensuring **scalability** and **easy access** from anywhere, at any time. Cloud deployment allows for seamless updates, efficient processing power, and global availability, making deepfake detection accessible to a broader audience.

Screenshots



Deepfake Video Detector



Deepfake Detector

127.0.0.1:5000

Deepfake Video Detector

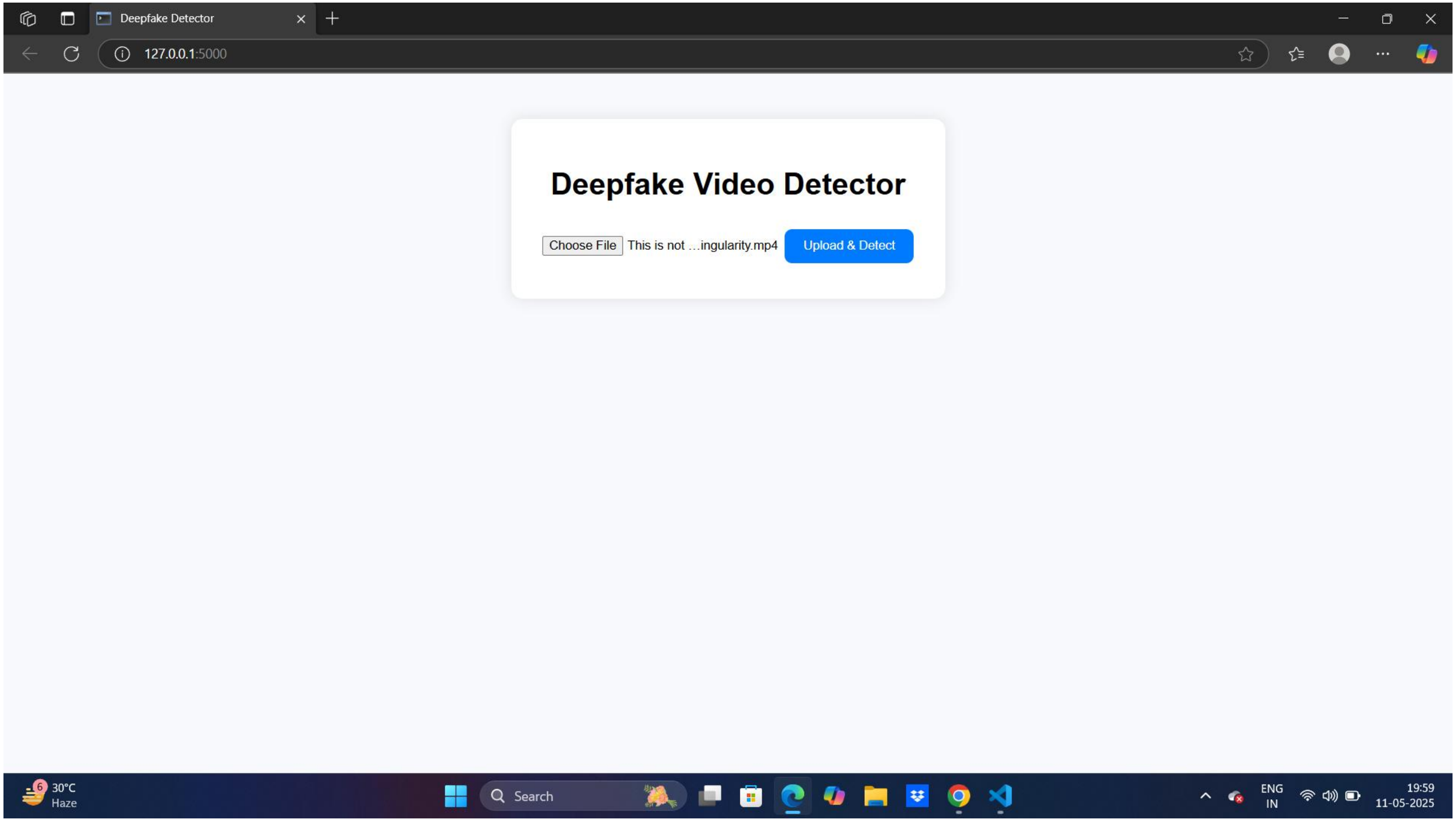
No file chosen

Open

Videos › Videos

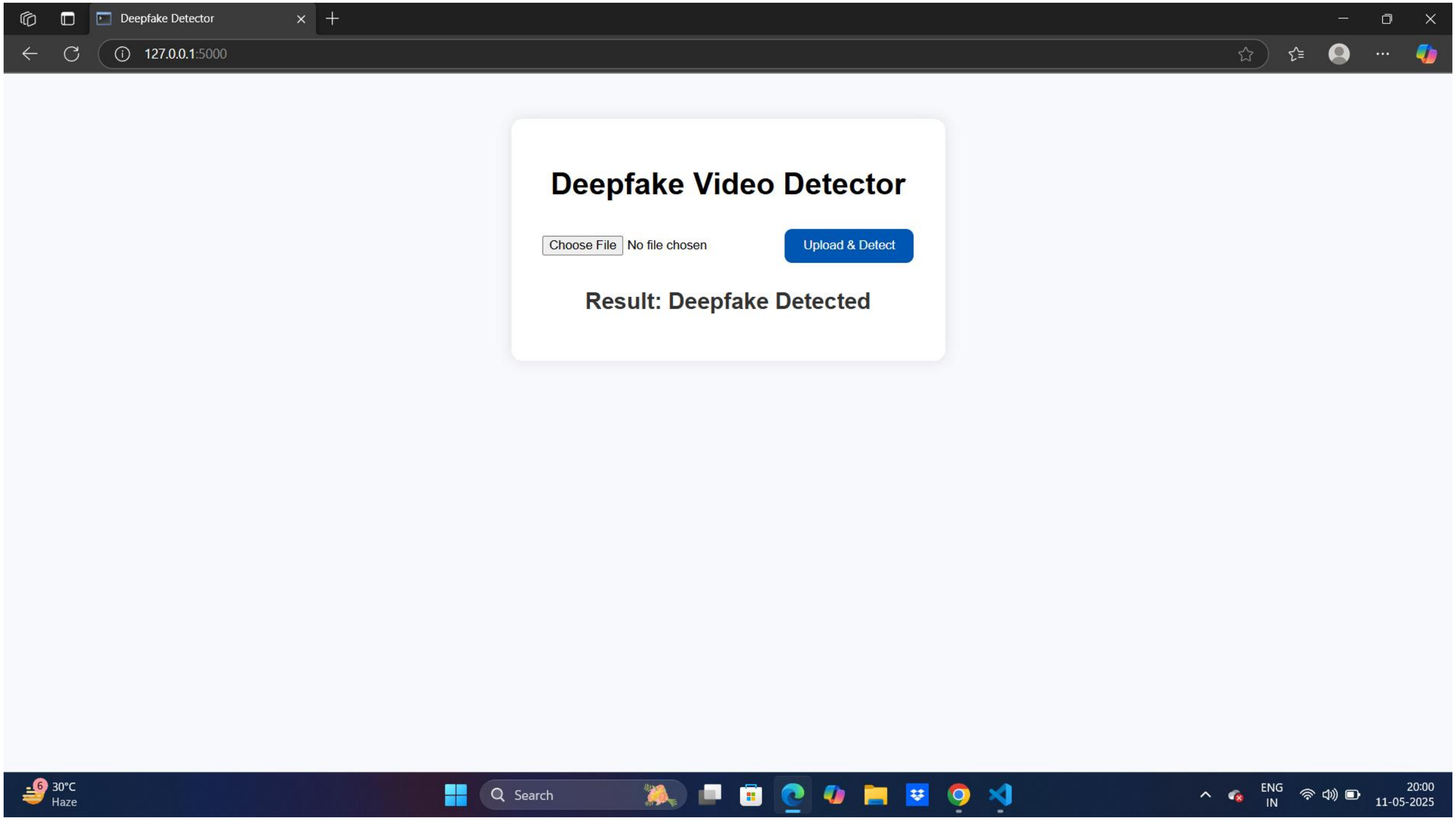
File name: This is not Morgan Freeman - A Deepfake

Video files



Deepfake Video Detector

Choose File This is not …ingularity.mp4



Deepfake Video Detector

No file chosen

Result: Deepfake Detected

Github Link

https://github.com/sid551/Deepfake-video-detector.git

Conclusion

Final Thoughts

Deepfake detection is becoming increasingly important to **maintain trust in digital media**, as the ability to manipulate content becomes more sophisticated. This tool offers an **accessible solution** for identifying deepfake content, empowering users to verify the authenticity of media in real-time. As development continues, new features and improvements will be implemented to **broaden its impact**, ensuring the tool remains effective in combating the rising threat of deepfakes.