TRUTHLENS AI

A STEGANOGRAPHY AND DEEPFAKE DETECTOR

TABLE OF CONTENTS

01 **Problem Statement**
Deepfakes and steganography evading current defenses.

02 **Proposed Solution**
AI-driven multi-modal detection system.

03 **Tech Stack**
Advanced AI, CV, and security tools.

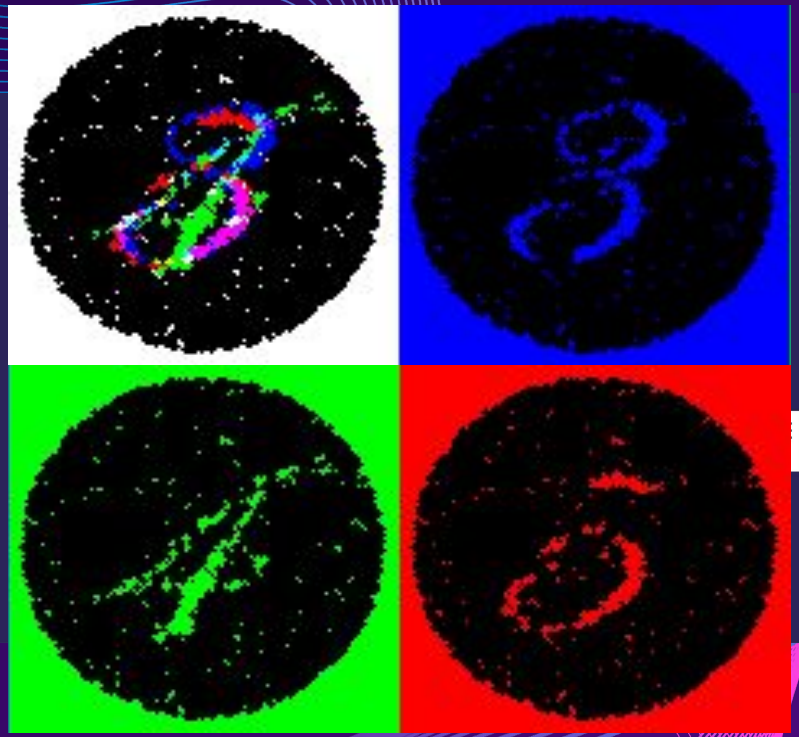
04 **Case Study**
Real incident prevention example.

What is DeepFake ?

Deepfake is a technique that uses artificial intelligence, especially deep learning, to create or alter videos, images, or audio to make them appear real. It can mimic a person's face, voice, or actions with high accuracy, making it difficult to distinguish from genuine content. While it has creative and entertainment uses, it is often exploited for misinformation, fraud, and harassment.

What is Steganography ?

Steganography is the technique of hiding secret data within ordinary files like images, videos, or audio, making the hidden content invisible. It differs from cryptography by concealing the message's existence, often used in both legal and malicious activities.



Problem Statement

WHY ?

Deepfakes and steganography are rapidly growing cyber threats, with deepfake videos increasing by over 900% in three years and 96% being malicious. Steganography is increasingly used to hide malware, stolen data, and covert messages, bypassing traditional detection and exploiting trust in digital media.

WHAT ?

An AI-powered media integrity system that uses Vision Transformers for advanced steganography detection and multi-modal deepfake detection models for verifying video authenticity, ensuring comprehensive protection against evolving media-based threats.

HOW ?

By scanning images, videos, and audio for hidden payloads or manipulations, detecting facial artifacts, audio-visual mismatches, and temporal inconsistencies, and integrating into email gateways, media platforms, and enterprise security tools for real-time analysis, continuously retraining models to adapt to new attack methods.

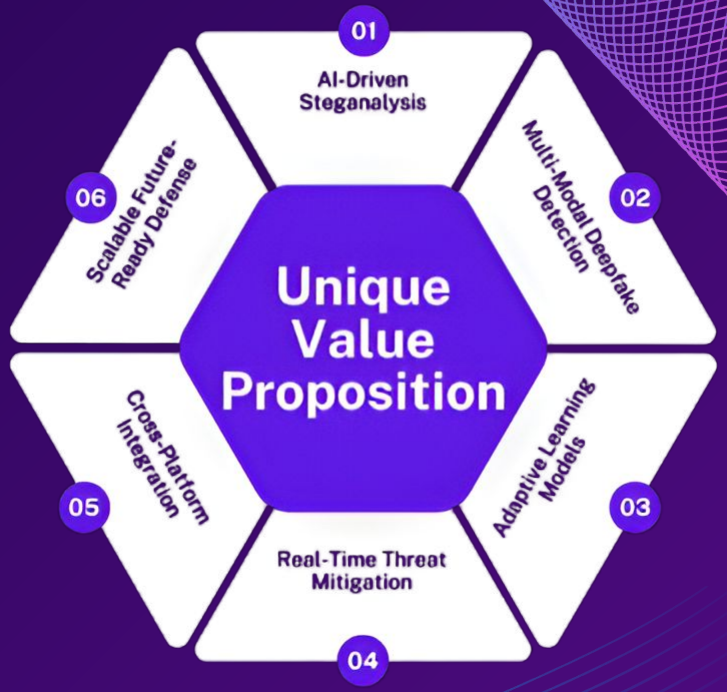
JUSTIFICATION FOR SDG GOAL

TITLE	PRIMARY	SECONDARY	TERTIARY
GOAL	16 (Peace, Justice and Strong Institutions)	9 (Industry, Innovation and Infrastructure)	4 (Quality Education)
TARGET	**Target 16.10** – Ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements.	**Target 9.5** – Enhance scientific research and upgrade technological capabilities to support innovation and secure infrastructure.	**Target 4.4** – Increase the number of people with relevant skills for employment, decent jobs, and entrepreneurship, especially in digital literacy.

PROPOSED SOLUTION

TruthLens AI is an AI-powered media integrity system combining **Vision Transformers** for steganography detection and **multi-modal deepfake detection** to ensure authenticity in digital media across images, videos, and audio.

- **AI-Driven Steganalysis** – Vision Transformers analyze pixel-level patterns to uncover hidden data in images, videos, and audio files.
- **Multi-Modal Deepfake Detection** – Detects facial manipulations, audio-visual mismatches, and temporal inconsistencies in videos.
- **Adaptive Learning Models** – Continuously retrains on new datasets to counter evolving attack techniques.
- **Real-Time Integration** – Easily deployable in email gateways, social media platforms, and enterprise security tools.



Tech Stack

Technology	Deepfake Detection	Steganography Detection
Frameworks	TensorFlow / PyTorch	TensorFlow / PyTorch
Video Processing	FFmpeg	None
Image Processing	OpenCV	OpenCV
Models	ResNeXt CNN + LSTM Networks	Convolutional Neural Networks (CNNs)
Learning	Transfer Learning (ResNeXt Pretrained Models)	Spatial Rich Models (SRM)
Analysis	Trained Dataset	Wavelet Transform
ORM	Prisma ORM	Vision Transformer

Made with Napkin

5. Our Results

Model Name	No of videos	No of Frames	Accuracy
model_84_acc_10_frames_final_data.pt	6000	10	84.21461
model_87_acc_20_frames_final_data.pt	6000	20	87.79160
model_89_acc_40_frames_final_data.pt	6000	40	89.34681
model_90_acc_60_frames_final_data.pt	6000	60	90.59097
model_91_acc_80_frames_final_data.pt	6000	80	91.49818
model_93_acc_100_frames_final_data.pt	6000	100	93.58794

Exposing the Unseen

Detect Deepfakes & Hidden Data

Advanced AI-powered detection of deepfake videos and steganographic content. Protect yourself from digital deception with cutting-edge technology.

Try Deepfake Detection →

Try Steganography Detection →

Deepfake Detection

Upload your video and let our advanced AI analyze it for deepfake manipulation

Upload Video

demo-deepfake-video.mp4

File selected successfully

Start Analysis | Remove File

Try Demo Video (Deepfake)

Detection Tips

- Higher resolution videos provide better detection accuracy
- Videos with clear facial features work best
- Supported formats: MP4, MOV, AVI
- Demo video is a confirmed deepfake for testing

Video loaded successfully. Click "Start Analysis" to begin processing.

Extracted Frames (20)

Upload Image

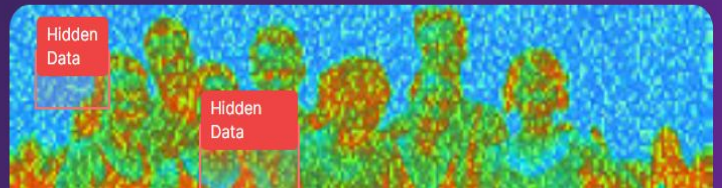
dee11.jpeg

File selected successfully

Remove File

Try Demo Image (Hidden Data)

Original Image



Analysis Results

Hidden Data Found

Confidence 88%

Data Type

Text

Extracted Content

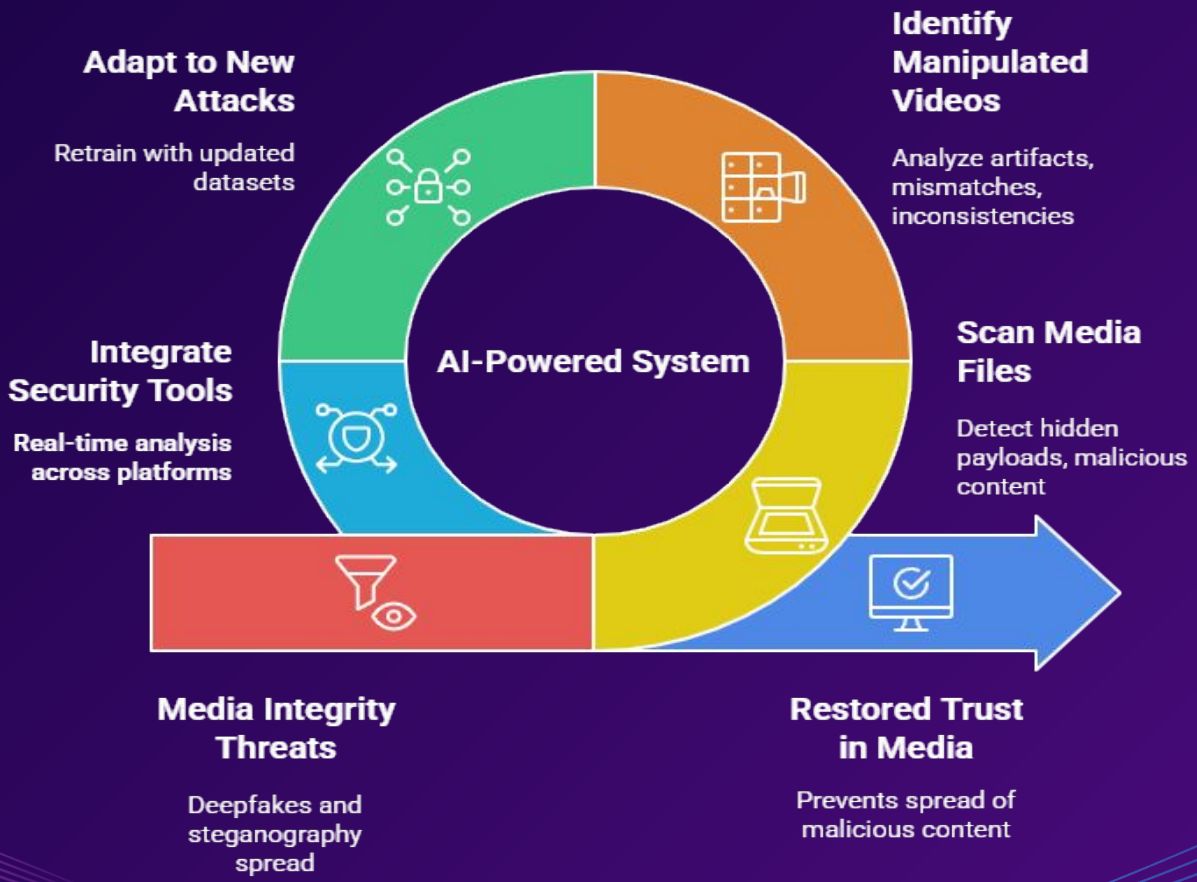
```
Classified data embedded
```

Download Analysis Report

How It Works

- Analyzes least significant bits (LSB) for hidden data

IMPACT



Made with Napkin

CASE STUDY

Arup Deepfake Scam $25.6M Stolen - In early 2024, a finance employee at the UK engineering firm Arup participated in a video call with what appeared to be senior management. All attendees except the employee were AI-generated deepfakes. Convinced the request was legitimate, the employee authorized 15 transfers, totaling $25.6 million.

AI-Generated Deepfake Content in South Korea - On platforms like Telegram, AI-generated deepfake pornography of women is rapidly distributed, with users uploading images and personal information to create explicit content. It's a growing and alarming trend.

2010 Russian Spy Ring (Operation Ghost Stories) – The FBI exposed a Russian spy network in the U.S. where agents hid encrypted instructions inside public online images using steganography, enabling covert communication without direct contact.

2018 Hacking Campaign Against Global Banks – Cybersecurity firm Symantec reported that attackers targeting financial institutions embedded malicious code and stolen data inside seemingly harmless corporate logo images to bypass security systems and exfiltrate sensitive banking information

THANK YOU!!