Baderia Global Institute of Engineering and Management, Jabalpur, Madhya Pradesh 482002

BADERIA GLOBAL THE BEST Institute of Engineering & Management

# BrahmaX 1.0

The Creation of Tomorrow

BrahmaX 1.0

www.codecrax.com

FOUR FIREWALL

## 01 Profile Overview

**Theme :-** Cyber

**Problem Statement Title :-** Develop a real-time AI system to detect deep fake audio, images, and videos, enabling secure, trustworthy communication within enterprise environments.

**Team ID :-**

**Team Name :-** FOUR FIREWALLS

FOUR FIREWALL

# 02 TruthShield: Real-time AI Deepfake Detection

## Solution Overview

Truth Shield is an AI-powered deepfake detection system that identifies fake voices, images, and videos in real-time. By leveraging computer vision, audio forensics, and behavioral analysis, it detects deep fakes across email, video calls, messaging platforms, and social media. Integrated with corporate security systems, it provides automated alerts and reports, ensuring immediate intervention without disrupting workflows.

## Problem-Solving

Our system uses a multi-modal detection engine that combines edge computing for fast analysis and cloud validation for accuracy. It detects deepfakes in real-time and assigns a digital trust score to evaluate the credibility of communications.

## INNOVATION

- Zero-Shot Learning: Detects new deepfakes without retraining
- Contextual Analysis: Assesses communication behavior and history.
- Real-Time Detection: Flags deepfakes instantly.
- Trust Scoring: Rates credibility of interactions.
- Adversarial Resilience: Withstands countermeasures.
- Easy Integration: Compatible with corporate security systems.

FOUR FIREWALL

# 03 Technical Approach

## Technologies Used

**Python, JavaScript**

Core languages for AI and frontend

**TensorFlow, PyTorch**

Deep learning model development

**Transformers, Whisper, OpenCV**

NLP, audio, and vision processing

**FastAPI, Flask**

Real-time backend APIs

**Docker, AWS, Edge Devices**

Scalable, low-latency deployment

**REST APIs, GitHub**

Integration and version control

## Methodology

01 **Data Collection**
Gather real and deepfake media

02 **Preprocessing**
Clean and prepare data for training.

03 **Model Development**
Train AI models (CNN for images, RNN for audio).

04 **Deep Face Detection**
Detect inconsistencies in media

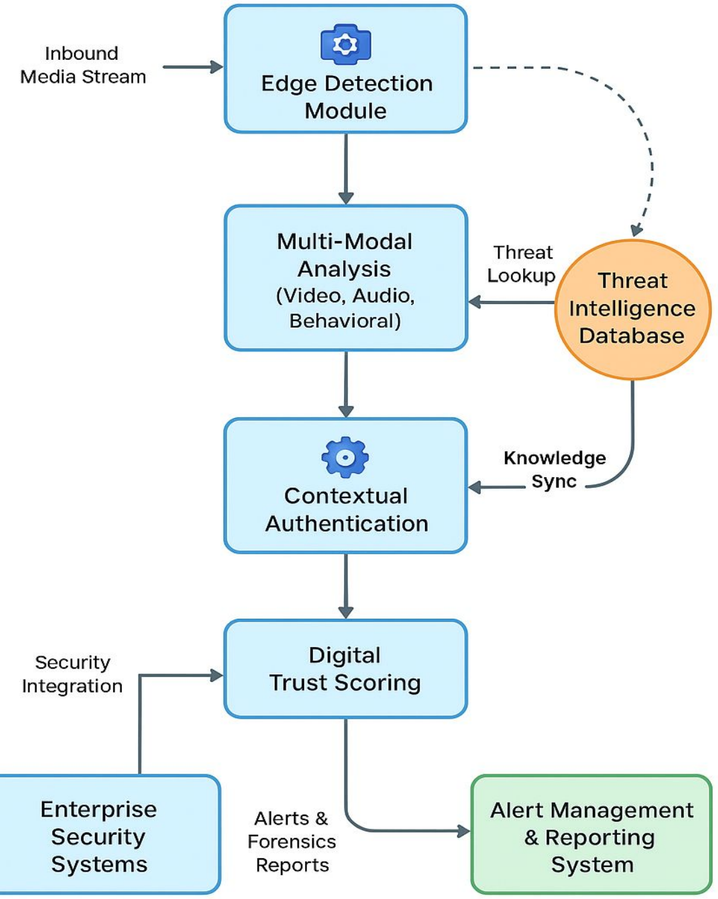
05 **Integration**
Connect with corporate security systems

06 **Real-Time Flagging**
Implement immediate deep face detection.

07 **Testing & Validation**
Assess and optimize model accuracy

08 **Deployment**
Deploy for continuous monitoring.

## TruthShield Architecture:
Real-Time Deepfake Detection Workflow



# 04 FEASIBILITY AND VIABILITY

## Feasibility

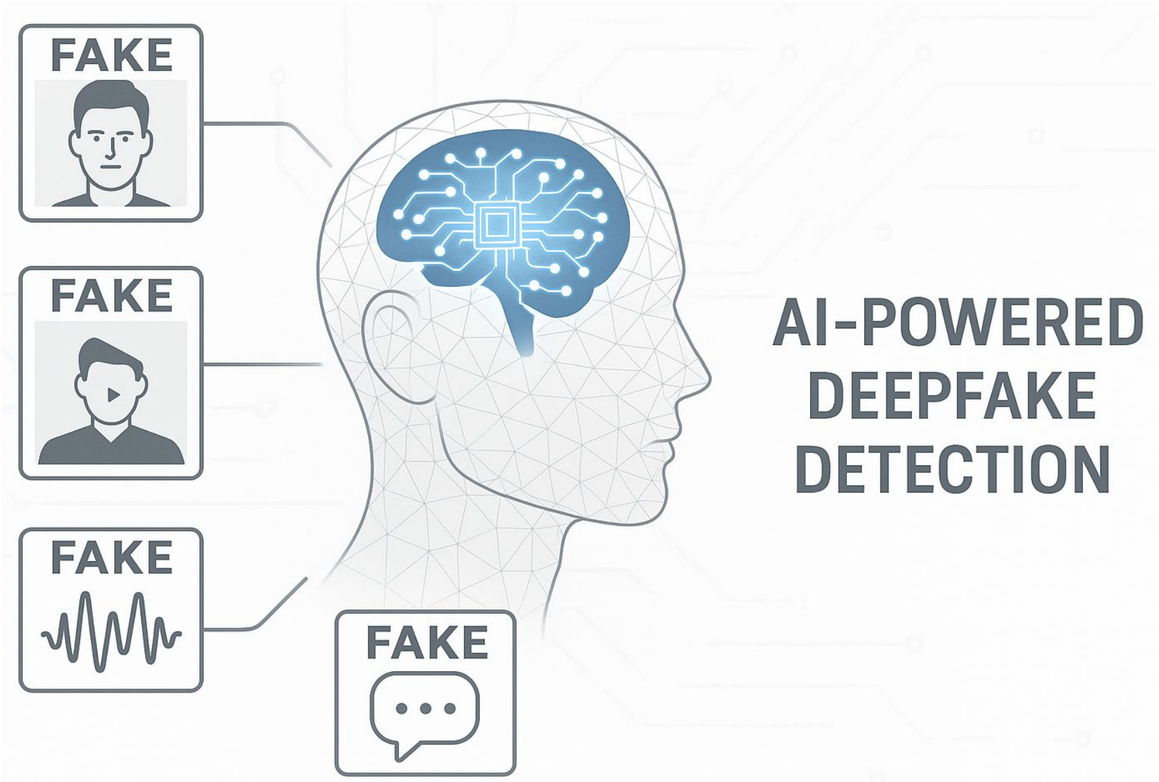
- Proven AI frameworks
- Real-time analysis via edge-cloud setup
- Seamless integration with security tools
- Scalable with Docker and cloud platforms

## Challenges & Risks

- Evolving deepfake techniques
- High compute requirements
- Data privacy concerns
- Risk of model drift

## Mitigation Strategies

- Use zero-shot and adversarial training
- Optimize models for edge devices
- Apply federated learning and encryption
- Continuous model updates and feedback loops



# 05 Impact And Benefits

## Strategic Impact

### Corporate Security Teams

Strengthens defenses against deepfake-based fraud and cyber attacks

### Enterprises & Organizations

Protects brand reputation, sensitive data, and ensures communication integrity

### Compliance & Legal Teams

Provides forensic evidence and audit trails for regulatory compliance

### Executives and Employees

Safeguards internal communications and decision-making processes

### Cybersecurity Industry

Enhances threat intelligence and deepens protection frameworks

## Key Advantages

**Social:** Builds digital trust and combats misinformation

**Economic:** Reduces financial risks from cyber fraud

**Environmental:** Leverages edge AI for lower energy consumption

## Long-Term Value

Scalable across sectors and geographies

Continuously improves through feedback and retraining

Future-ready for integration with national security systems

## References & Research Work


| | |
|---|---|
| 01 | Li, Y., Chang, M. C., & Lyu, S. (2018). *In Ictu Oculi: Exposing AI-Generated Fake Face Videos by Detecting Eye Blinking*. arXiv preprint. https://arxiv.org/pdf/1806.02877 |
| 02 | Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., & Ferrer, C. C. (2020). *The Deep Fake Detection Challenge Dataset*. arXiv preprint. https://arxiv.org/pdf/2006.07397 |
| 03 | Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). *MesoNet: A Compact Facial Video Forgery Detection Network*. In WIGS. https://arxiv.org/pdf/1809.00888 |
| 04 | Agarwal, S., Farid, H., Gu, Y., He, M., Nagano, K., & Li, H. (2019). *Protecting World Leaders Against Deep Fakes*. In *CVPR Workshops*. 🔗 Open Access Paper 🔗 PDF Download |
| 05 | Khalid, H., Woo, W. L., & Sugiarto, A. (2024). *Deepfake Video Detection: Challenges and Opportunities*. *Artificial Intelligence Review*. SpringerLink Article |
| 06 | OpenCV Documentation https://docs.opencv.org |
| 07 | TensorFlow Documentation https://www.tensorflow.org |
| 08 | PyTorch Framework https://pytorch.org |
| 09 | DeepFake Detection Challenge (DFDC) – Kaggle Competition https://www.kaggle.com/c/deepfake-detection-challenge |
| 10 | FaceForensics++ Dataset https://github.com/ondyari/FaceForensics |