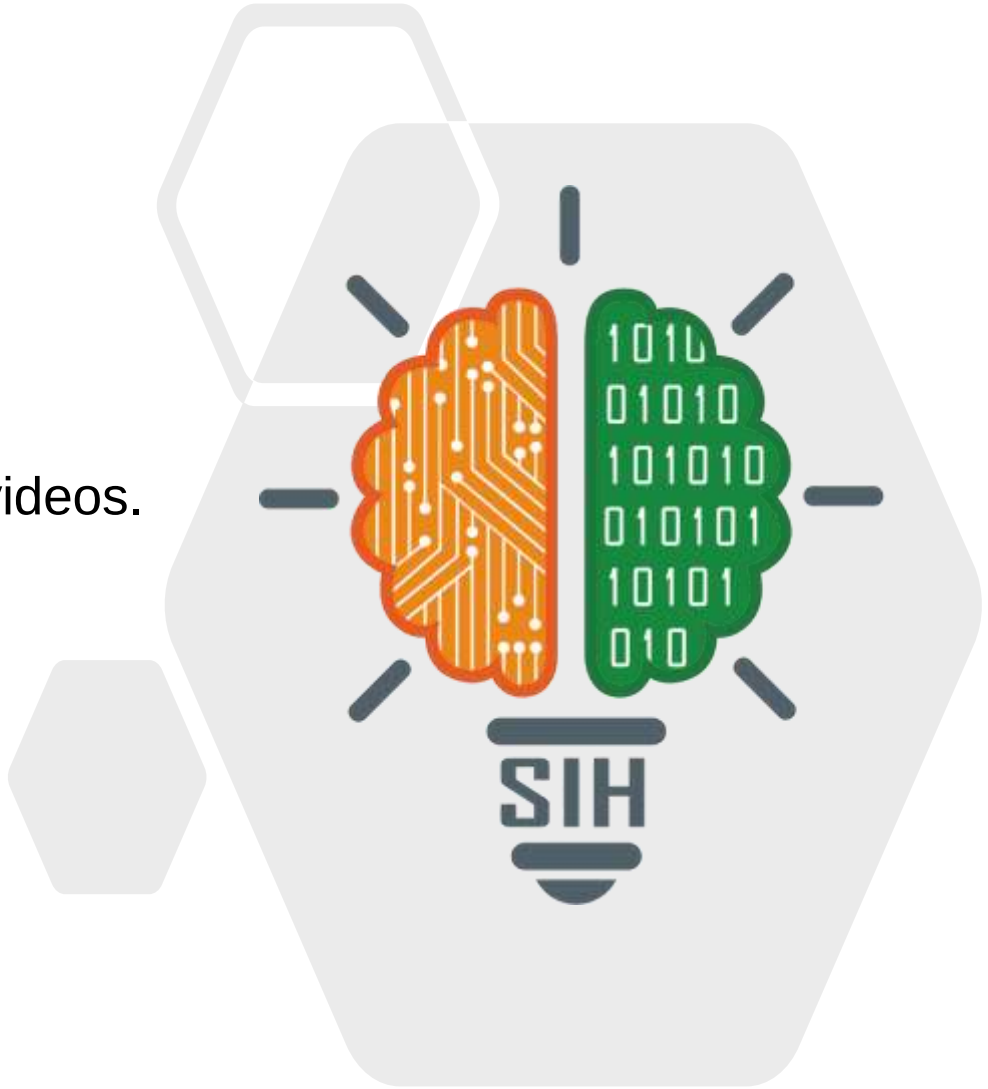
SMART INDIA HACKATHON 2024

- **Problem Statement ID :** SIH 1683
- **Problem Statement Title :**

  Development of AI/ML based solution for detection of face-swap based deep fake videos.
- **Theme :** Miscellaneous
- **PS Category :** Software
- **Team ID :** 33570
- **Team Name :** Wolfenstein



Wolfenstein

OVERVIEW

SMART INDIA HACKATHON 2024

IDEA / PROPOSED SOLUTION

FUSION OF SPATIAL-TEMPORAL DYNAMICS FOR DEEP FAKE DETECTION

- **Spatial Feature Extraction :** Detects pixel-level facial irregularities.
- **Temporal Dynamics Tracking :** Identifies behavioral inconsistencies over time.

ADVANCED 3D POSE AND TEXTURE GRADIENT ANALYSIS

- **3D Pose Estimation :** Captures the natural and unnatural facial shifts.
- **Texture Mapping :** Detects texture deformations in the media.

STRENGTHENING RECOGNITION WITH ADVERSARIAL TRAINING

- **Distortion Detection :** Helps in recognizing subtle deep fake distortions.
- **Model Evolution :** Adapt to evolving deep fake mechanisms via self training on historic data.

CROSS-CHECKING AUDIO AND VISUAL ALIGNMENTS

- **Speech-Lip Sync :** Flag mismatches between speech and lip movement.
- **Cross-Modal Analysis :** Pinpoint inconsistencies between audio and visuals.

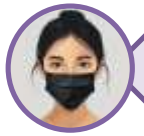
DETAILED REPORT GENERATION FOR ENHANCED USER EXPERIENCE

- **In-Depth Results :** Create reports that summarize outcomes and confidence levels.
- **Smart Responses :** Equip users with insights to respond effectively to detected irregularities.

HOW THE PROBLEM IS ADDRESSED

- **Data Acquisition & Preprocessing:** Source and pre-process real and fake videos for evaluation.
- **Model Development:** Use ResNet for spatial analysis and Long Short-Term Memory (LSTM) for temporal insights.
- **Comprehensive Training & Deployment:** Train and deploy model for detection, generating detailed efficacy reports.

INNOVATION AND UNIQUENESS



Occlusion-Robust Deep fake Recognition: Utilizes an avant-garde algorithm to detect fakes, even with partial facial occlusions.

Blockchain-Backed Immutable Verification: Uses blockchain for an immutable ledger, ensuring security and media authenticity.

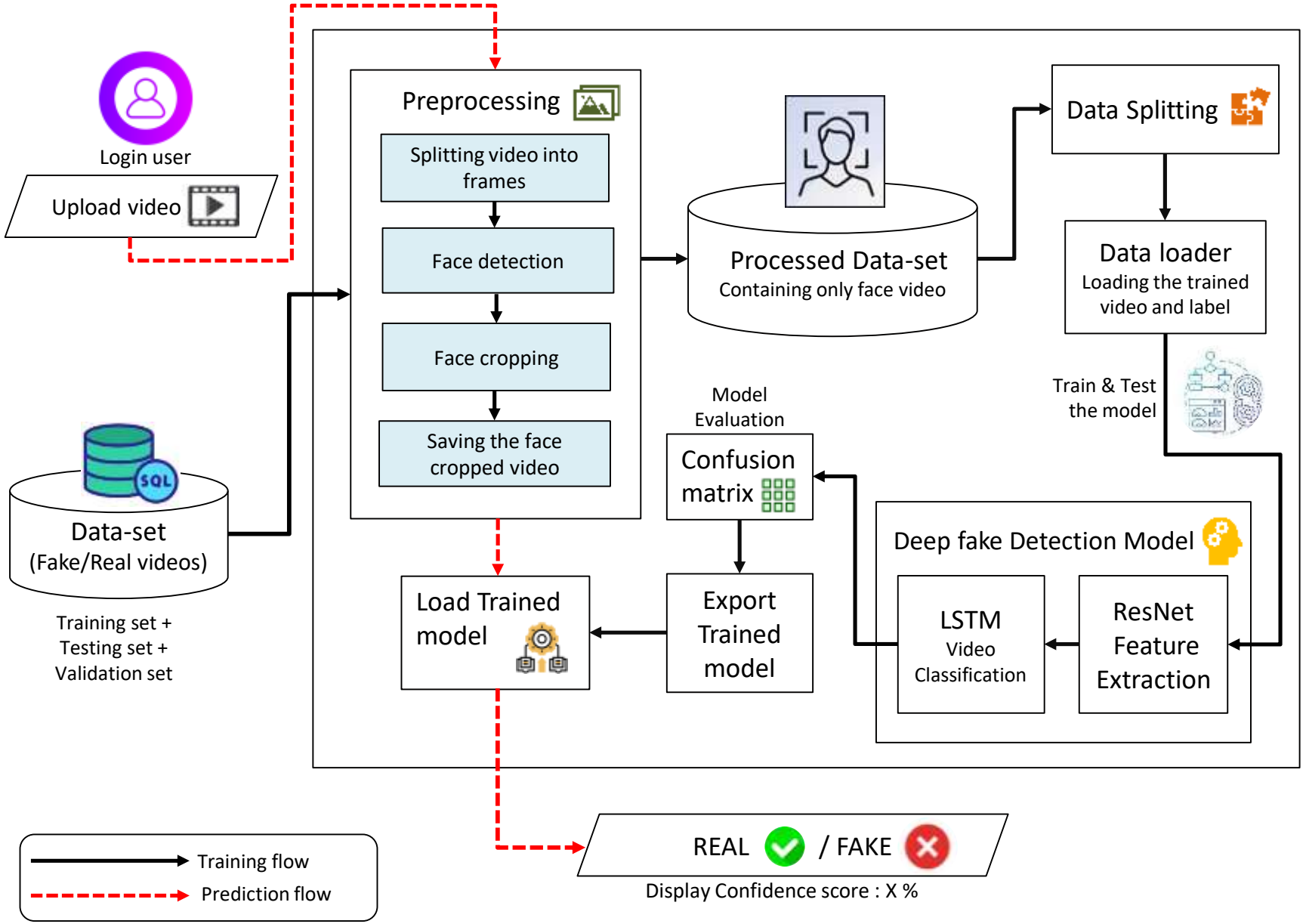
Multifaceted Detection: Integrates spatial, temporal, and 3D analysis to enhance the accuracy of deep fake recognition.

Biometric Examination: Spots visual fabrications by spotting micro-expressions, blinking, hair strands, eyelashes & other cues.

Wolfenstein

TECHNICAL APPROACH

SMART INDIA HACKATHON 2024



TECH STACK

COMPONENT	TECHNOLOGY/ALGORITHM
Frontend	HTML, JavaScript, Flask templates
Backend	Django
Deep Learning Framework	PyTorch
Neural Network Library	Torch (including nn.LSTM)
Image Processing	Torchvision
Automatic Differentiation	Autograd
Video Manipulation	OpenCV
Face Recognition	face_recognition
Lip Reading	LipNet, Amazon Transcribe
Database	ChromaDB, SQL, MongoDB
Visualization	NumPy, Matplotlib
AI/ML Model	PyTorch ResNeXt (resnext50_32x4d), Scikit-learn
Sequential Analysis	Long Short-Term Memory (LSTM)
Probability Conversion	Softmax
Session Management	Flask sessions
File Upload Handling	Flask
Cloud Services	AWS, Google Cloud, or Azure
Containerization	Docker

Wolfenstein

FEASIBILITY AND VIABILITY

SMART INDIA HACKATHON 2024

FEASIBILITY OF OUR PROJECT

TECHNOLOGICAL VIABILITY

Sophisticated algorithms guarantee adaptability to the continually **evolving landscape** of synthetic media methods.

OPERATIONAL EFFICIENCY

Decentralized processing amplifies **real-time responsiveness** and facilitates seamless integration across **diverse platforms**.

FINANCIAL SUSTAINABILITY

Engaging **Indian media** houses and content creators, our solution is essential for **content authenticity**, offered through flexible **subscription models** tailored to their needs.

FUTURE ADVANCEMENTS

- ✓ **Integration of Emotion detection and Action recognition.**
- ✓ **Elevating the model's accuracy from 80% up to 95%.**
- ✓ **Exploration of thermal imaging for anomaly detection.**
- ✓ **Development of browser extensions and APIs for real-time flagging.**

POTENTIAL CHALLENGES & RISKS	STRATEGIES TO OVERCOME THEM
ADVERSARIAL EXPLOITATION Advanced evasion methods undermine detection precision and expose vulnerabilities.	**MODEL OPTIMIZATION** Employ advanced training methodologies, ensemble strategies, and **iterative refinements** to strengthen model resilience.
ACCURACY EQUILIBRIUM Balancing false positives and false negatives presents significant challenges to detecting truthfulness.	**DETECTION RELIABILTY** Implementing **threshold tuning** alongside **precision-recall, F1 metrics** and **Feedback loop** to address temporal inconsistency and elevate detection reliability.
COMPUTATIONAL EXIGENCIES High demand for GPUs and TPUs leads to processing bottlenecks, elevated costs, and increased energy consumption.	**EFFICIENCY ENHANCEMENT** Leveraging **cloud** solutions for **dynamic scaling** and implementing **load balancing** for efficient workload distribution.

Wolfenstein

IMPACT AND BENEFITS

SMART INDIA HACKATHON 2024

POTENTIAL IMPACT ON THE TARGET AUDIENCE

SOCIAL IMPACT

- Detection of manipulated content protect individuals from **mental health issues**, including depression and suicidal thoughts, caused from deep fakes.
- Improving verification efficiency by **50%**, helps **Indian media** and **Content Creators** mitigate revenue drops & enhance trust.

ECONOMICAL IMPACT

- The solution fortifies **brand integrity and trust** by authenticating content and repelling deep fakes.
- Ensuring genuine influencer content, our model enhances **marketing effectiveness**, as 70% of consumers are influenced by authenticity, boosting the **ROI by 10-15%.**

POLITICAL IMPACT

- Deep fake videos distort voting impacting elections. Our model can detect deep fakes of politicians to preserve **Democracy**.
- Altered media undermine **India's foreign policy** by enabling adversaries to exploit misrepresentation, which can be prevented by our model.

LEGAL IMPACT

- Our model aims to cut **average ransom** payments of **₹50,000** from deep fake identity theft by **50%,** easing the **financial burden** on victims.
- It will enhance **digital forensics** up to **95% accuracy**, reduces analysis time by **70%**, and preserves over **85% of evidence** for investigations.

BENEFITS FROM THE SOLUTION

Enhances the credibility of digital content by identifying manipulated videos, helping users differentiate between real and fake media.

Shields women from harmful, non-consensual deep fake content used for abuse, ensuring online safety.

Safeguarding Vulnerable Groups from targeted attacks and harassment facilitated by deep fake technology.

Cultivating Digital Literacy by encouraging users to critically analyze content, fostering a better awareness of digital manipulation techniques.

RESEARCH AND REFERENCES

WEBSITES :

- Indian Ministry of Home Affairs. (2024). *I4C Daily Digest - May 13, 2024*. https://i4c.mha.gov.in/cyber_digest/may_2024/I4C%20Daily%20Digest-%2013.05.2024%20.pdf
- Larger Resolution Face Masked, Weirdly Warped, DeepFake. https://github.com/dfaker/df
- A Denoising Autoencoder + Adversarial Losses and Attention Mechanisms for Face Swapping. https://github.com/shaoanlu/faceswap-GAN
- Data Security Council of India. (n.d.). *Deepfake detection*. https://ccoe.dsci.in/blog/deepfake-detection
- Papers with Code. (n.d.). *Deepfake detection*. https://paperswithcode.com/task/deepfake-detection
- Kumar, S., and A. Gupta. "Deepfake Detection: A Comprehensive Review." *Scientific Research Publishing*, 2024. https://www.scirp.org/journal/paperinformation?paperid=109149

RESEARCH PAPERS AND JOURNELS :

- L. Jiang, R. Li, W. Wu, C. Qian, and C. C. Loy, "DeeperForensics1.0: A large-scale dataset for real-world face forgery detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 2889–2898.
- F. H. Almukhtar, "A robust facemask forgery detection system in video," Periodicals Eng. Natural Sci., vol. 10, no. 3, pp. 212–220, 2022.
- Anjaneyulu, K. S. S. R., and Agarwal, D. P. *Sequence Learning: From Recognition and Prediction to Control*. Wiley, 2020.