



こんにちは

IIT Mandi x Edify Present



ZENTEJ SEASON S 3.0

NEURAL NINJAS
PRESENTS

Deepfake-Proof eKYC Challenge: Building AI for Trustworthy Identity Verification in Real Time

TEAM MEMBERS
SHRUTI & YASH

PROBLEM STATEMENT

The rapid rise of hyper-realistic deepfakes poses serious threats to digital identity verification, enabling identity theft, financial fraud, and misinformation.

Traditional eKYC methods relying on static image checks fail to detect AI-generated manipulations, undermining trust in digital verification systems.

CHALLENGES

- Verify if two facial inputs belong to the same person
- Detect whether the media is authentic or forged
- 01 Ensure real-time liveness and explainable decision-making

EXPECTED

INPUT: Two facial images or a short selfie video

OUTPUT: (1) Identity Match Score
(2) Liveness Score
(3) Authenticity Label



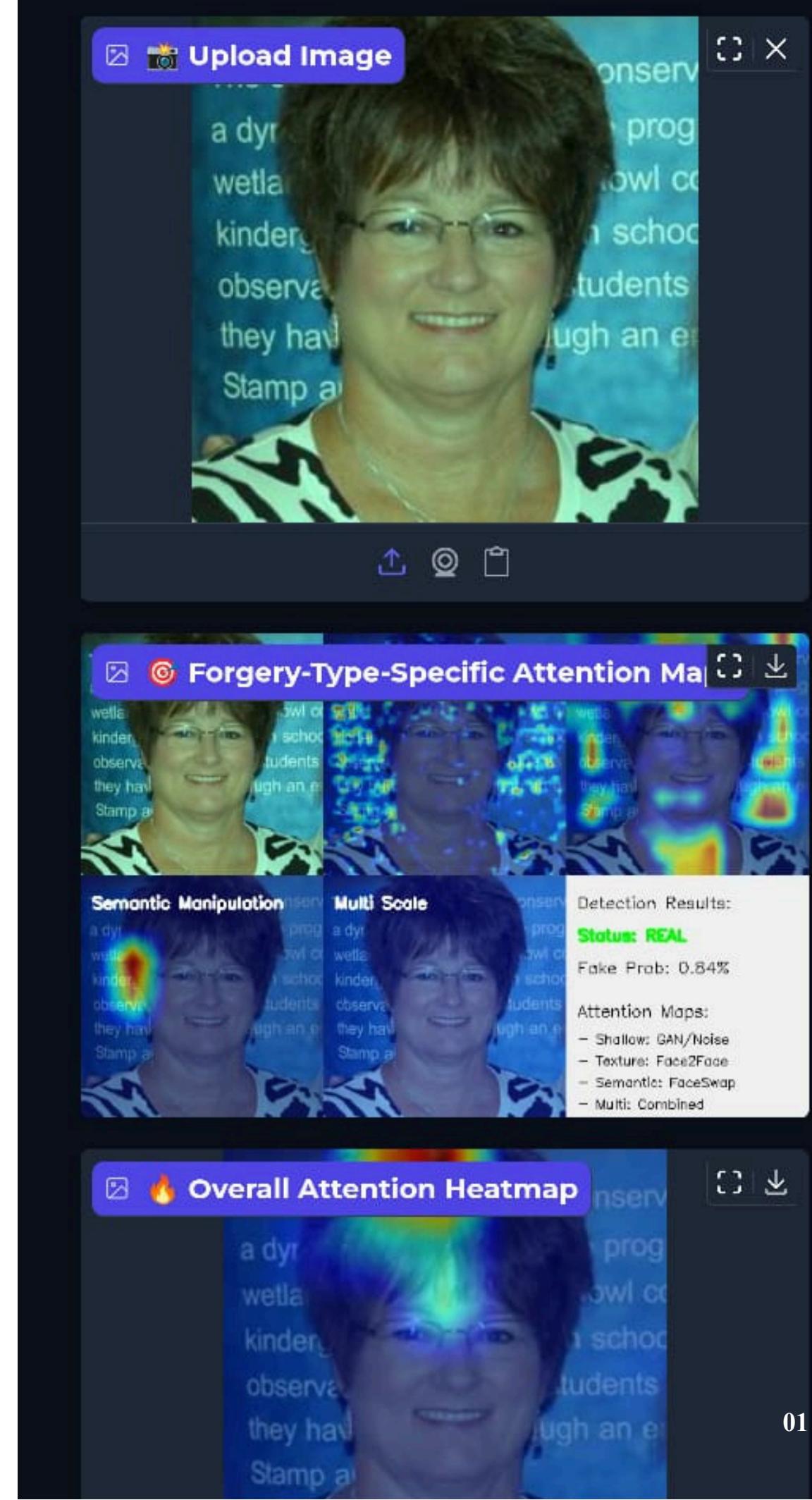


Accuracy & Robustness

Layer	Configuration	Evaluation Metric	Score
Feature Extractor	512 → 256 neurons with heavy dropout (0.6)	Test Accuracy	97.67%
Activation	ReLU with BatchNorm1d	Precision	97.70%
Regularization	Multi-layer dropout (0.5-0.6)	Recall	97.67%
		F1-Score	97.67%

Head	Purpose	Architecture	Output
Authenticity Head	Real/Fake Classification	256→128→2 (classes)	Softmax logits
Match Score Head	Identity Matching Confidence	256→128→1 + Sigmoid	Confidence score
Liveness Head	Live Person Detection	256→128→1 + Sigmoid	Liveness score

The model demonstrates excellent performance with **97.67%** accuracy across all key metrics, indicating strong generalization capability for deepfake detection.



You know it's fake, but not *HOW* or *WHERE*?.

OUR CONTRIBUTION

Forgery-Type-Specific Attention Explainability

Traditional deepfake detectors give you a score like "0.85 fake probability"
but DON'T tell you:

WHERE in the face did they find the manipulation?

WHAT specific features triggered the alarm?

WHY should a human trust this decision?

1. Eye Artifacts (Shallow Layer)

What: Yellow/red hotspots exactly on the eyes

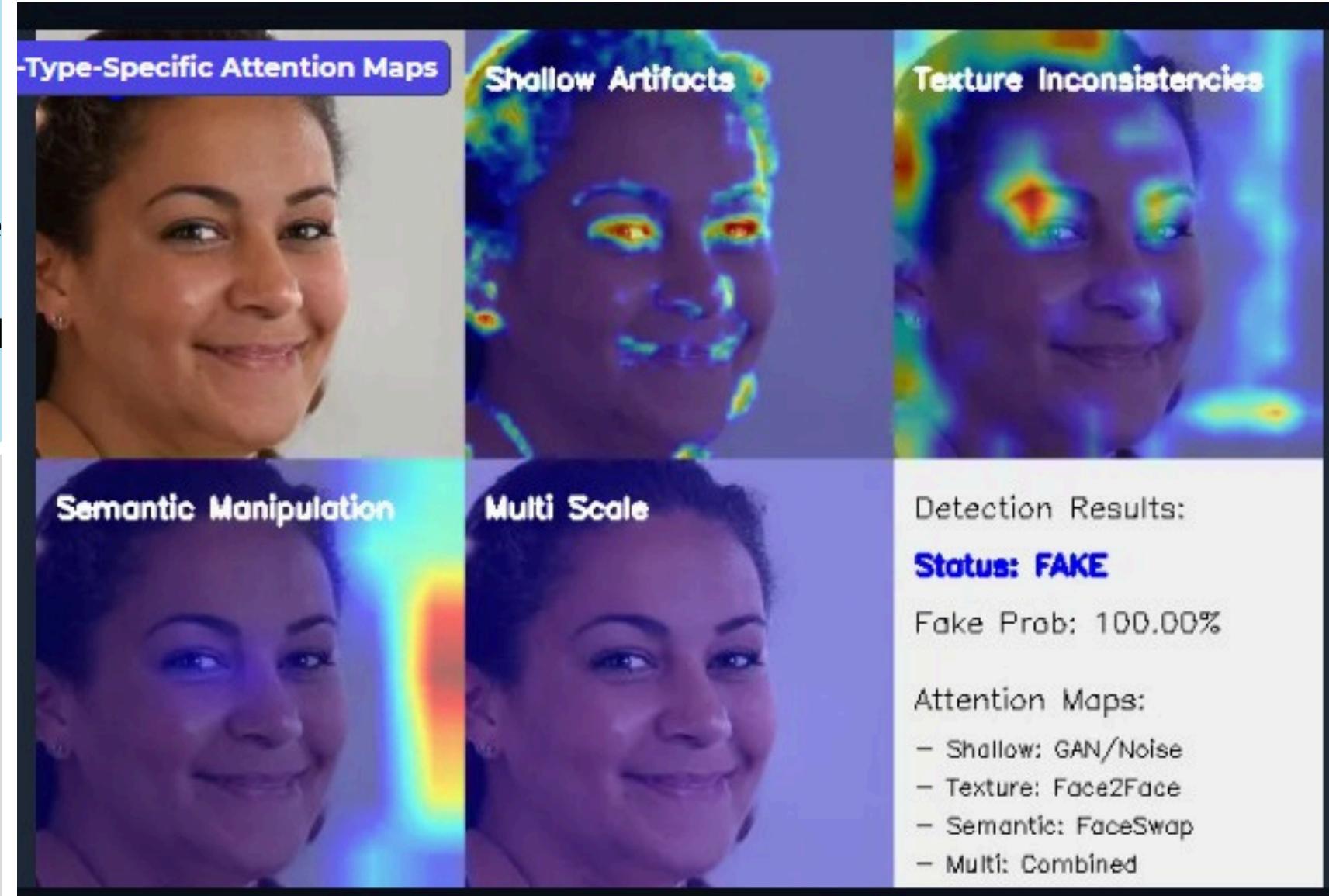
Why it matters: GANs notoriously struggle with eyes - they generate unnatural reflections, pupil inconsistencies, or weird iris patterns

4. Multi-Layer Consensus

What: All 3 network depths flagged the same face regions

Why it matters: When shallow, mid, and deep layers agree, confidence is near 100%

Why This is FAKE ?



2. Texture Discontinuities (Mid Layer)

What: Skin texture doesn't match natural human skin gradients

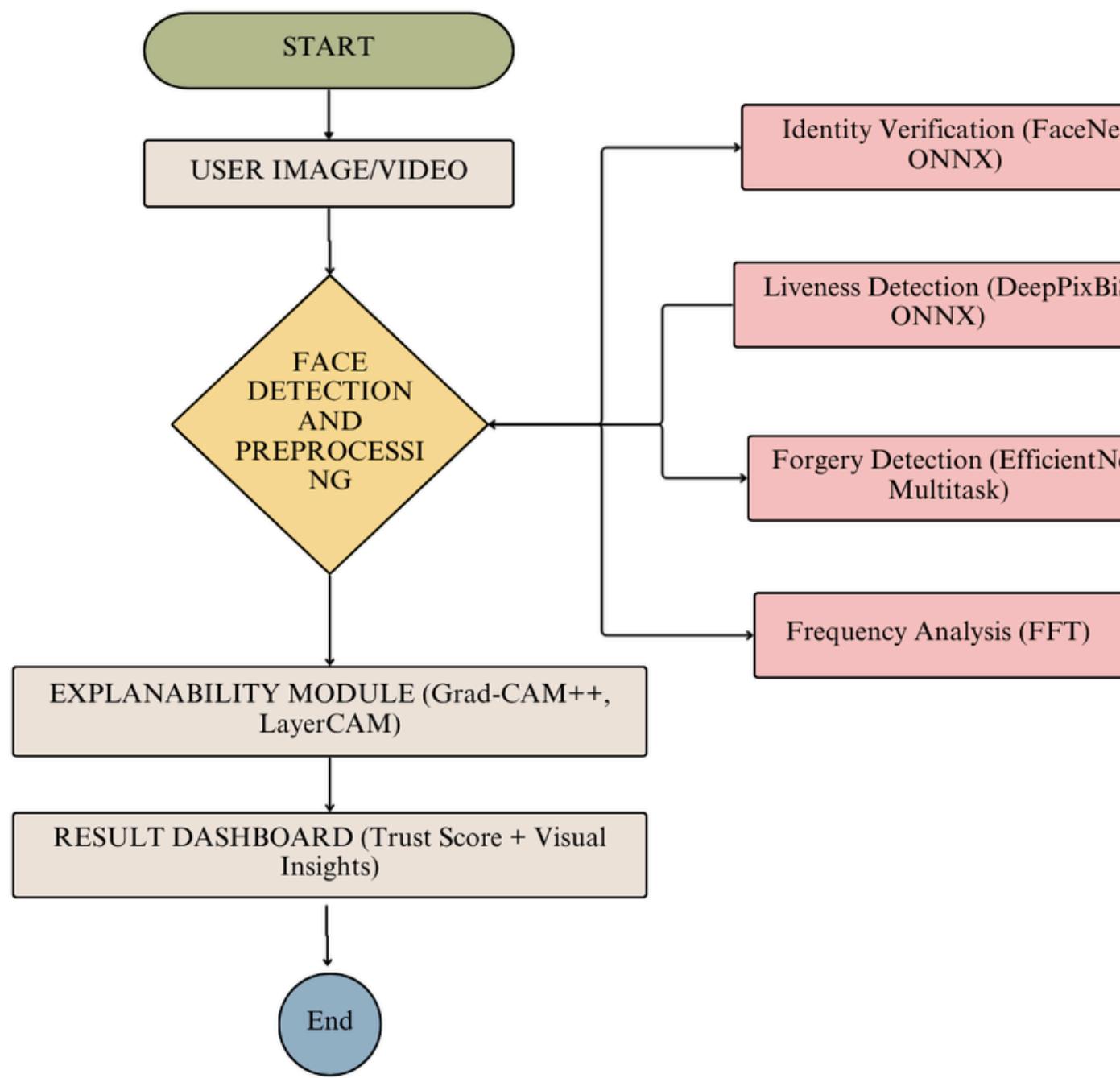
Why it matters: Face manipulation tools leave texture seams where synthetic and real regions meet

3. Asymmetric Attention (Overall Map)

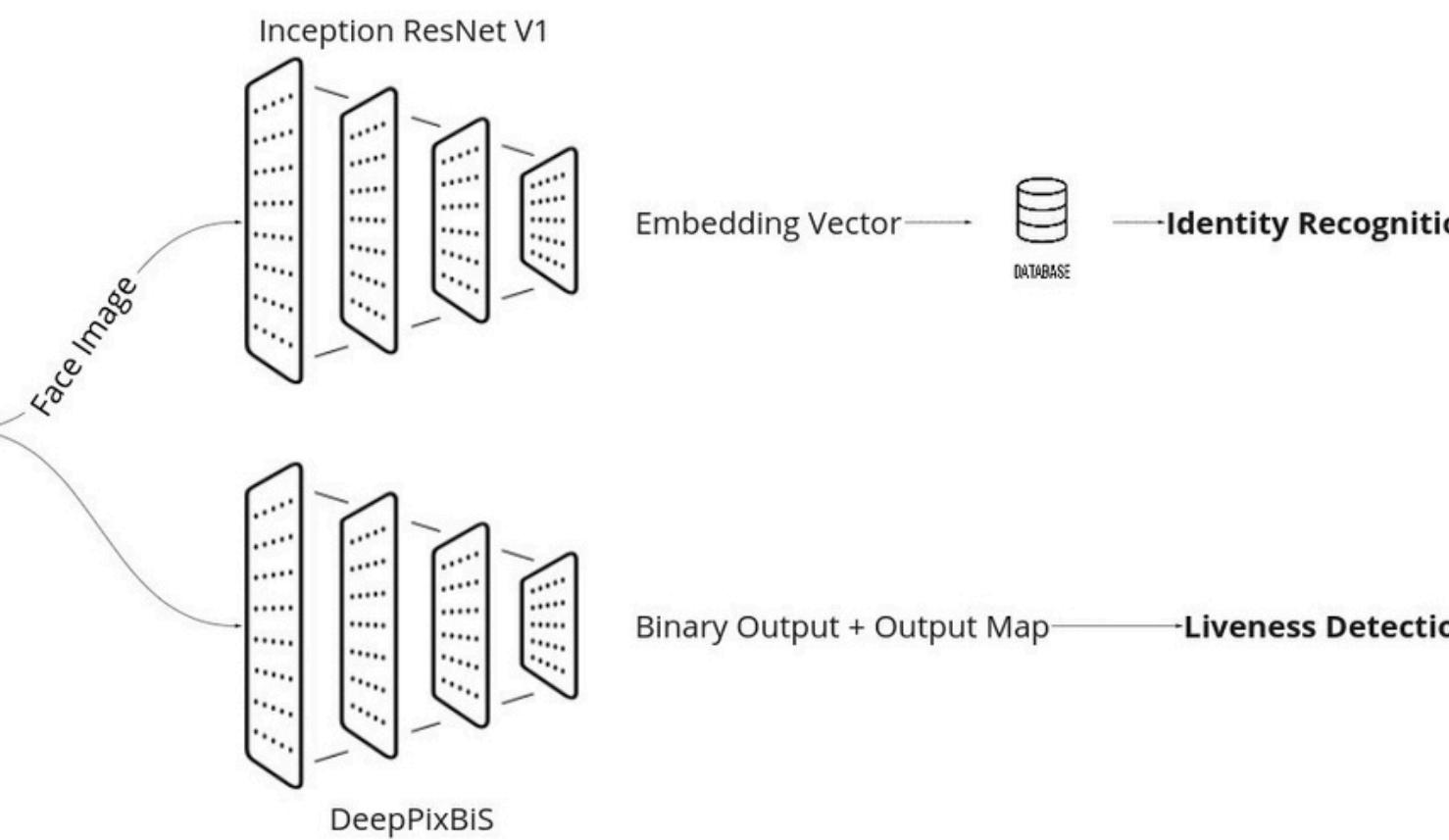
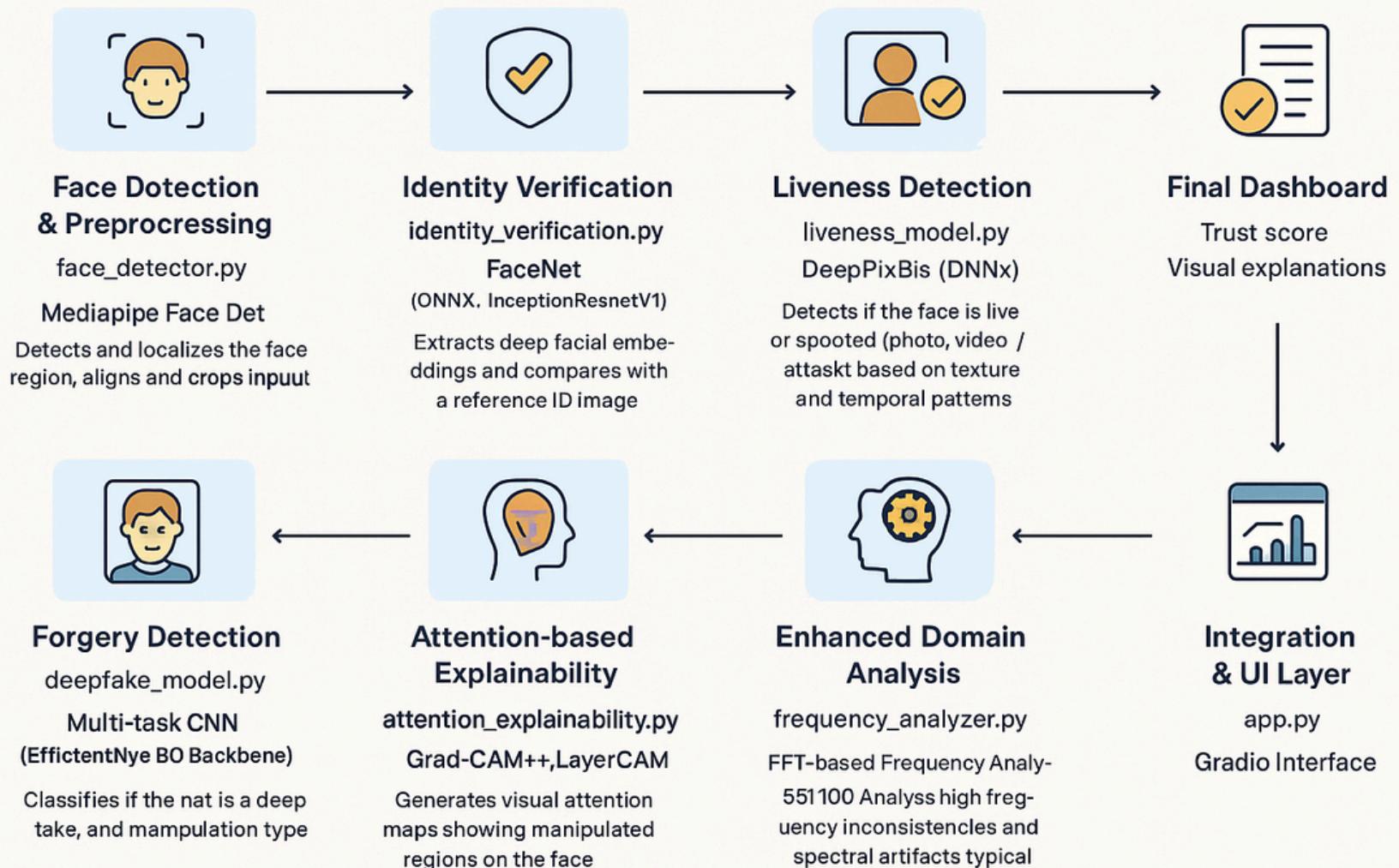
What: Left side of face (red) has way more attention than right side

Why it matters: Real photos have symmetric lighting/textures. Deepfakes often have processing asymmetry

PROPOSED SOLUTION OVERVIEW



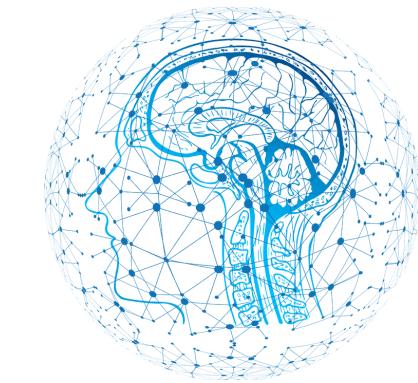
ZenTej AI – Explainable e-KYC Verification System





ありがとうございます

Thank you very much



AIが見ているものを見る – 透明で、信頼でき、解釈可能に

See What AI Sees – Transparent, Trustworthy, Interpretable