Deep Fake Image Detection Using Deep Learning

Team Members :

22BCE7295 - Varshith Eturu

22BCE7224 - Aman Sahu

22MIC7045 - Vasistha Ramana Rao Kv

22BCE9412 - Abhiram Grandhi

Under the Guidance of

Dr. Rajalakshmi Elangovan

Assistant Professor Sr. Grade 1

School of Computer Science and Engineering

VIT-AP University

BRIEF:

Explore the world of deep fake detection. This presentation covers deep learning methods. Learn to identify manipulated images and videos. We address the growing threat. Discover our robust detection system.

Presentation Agenda

VIT-AP UNIVERSITY

1. Introduction
2. Problem Statement
3. Proposed Solutions
4. Literature Survey
5. Existing Methods
6. Challenges
7. Project Objective
8. Architecture Diagram
9. Block Diagram
10. References

Introduction

VIT-AP UNIVERSITY

Deep Fakes

- AI-generated fake media (videos, images, or audio).
- Uses Deep Learning to swap faces or mimic voices.
- Entertainment, but also misinformation or scams.
- Privacy breaches, identity theft, and fraud.
- Tools spot inconsistencies to identify deepfakes

- **Realism & Accuracy**: Advances in GANs now produce highly realistic videos, audio, and AI-generated faces with improved detail.
- **Creative & Practical Uses**: Widely used in movies, gaming, virtual influencers, personalized avatars, and virtual assistants.
- **Medical Applications**: Supports facial reconstruction research and other healthcare innovations.

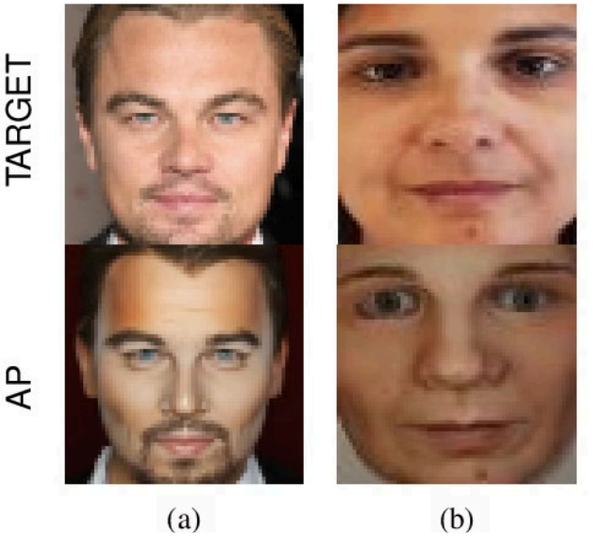
Introduction

VIT-AP UNIVERSITY

- The project focuses on deepfake detection in images and videos using **CNNs (Convolutional Neural Networks)** to **detect spatial inconsistencies** and **RNNs (Recurrent Neural Networks)** to **identify temporal anomalies.**

Deepfake Detection System-

- Spatial Anomaly Detection
- Temporal Anomaly Analysis
- Multi-Model Approach



Problem Statement: The Threat - misuse of technology

1

Misinformation

Deep fakes can spread false information. Manipulating world leaders

2

Reputation Damage

They can damage reputations. They can create false narratives.

3

Erosion of Trust

The authenticity of video's and images are questioned. It is becoming easier to not believe things that are actually true

Solution: Techniques

1

Multi-Model

Leverage multi-model detection. Analyze audio and video consistency.

2

Spatio-Temporal

Train on local artifacts. Use spatio-temporal cues.

3

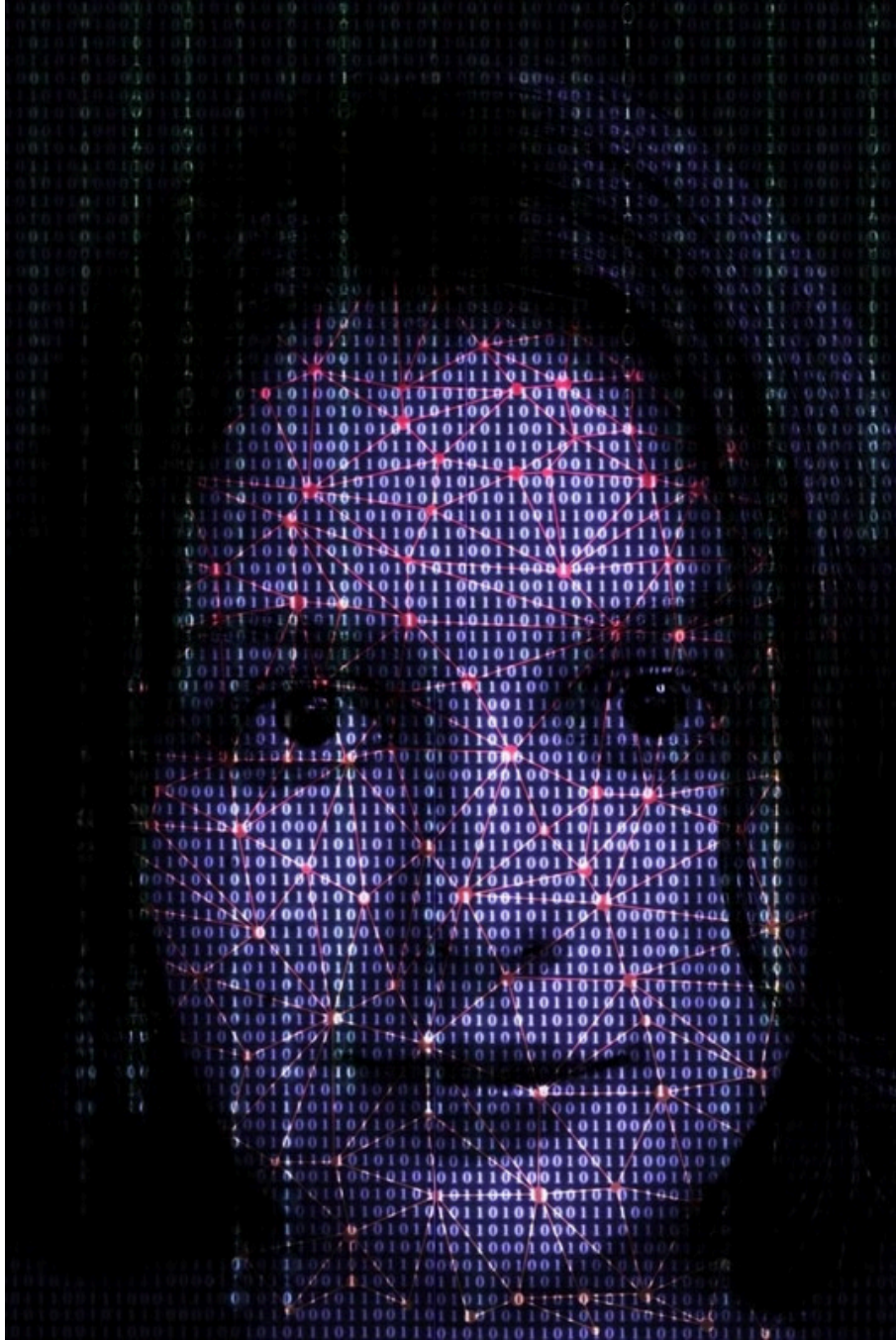
Anomalies

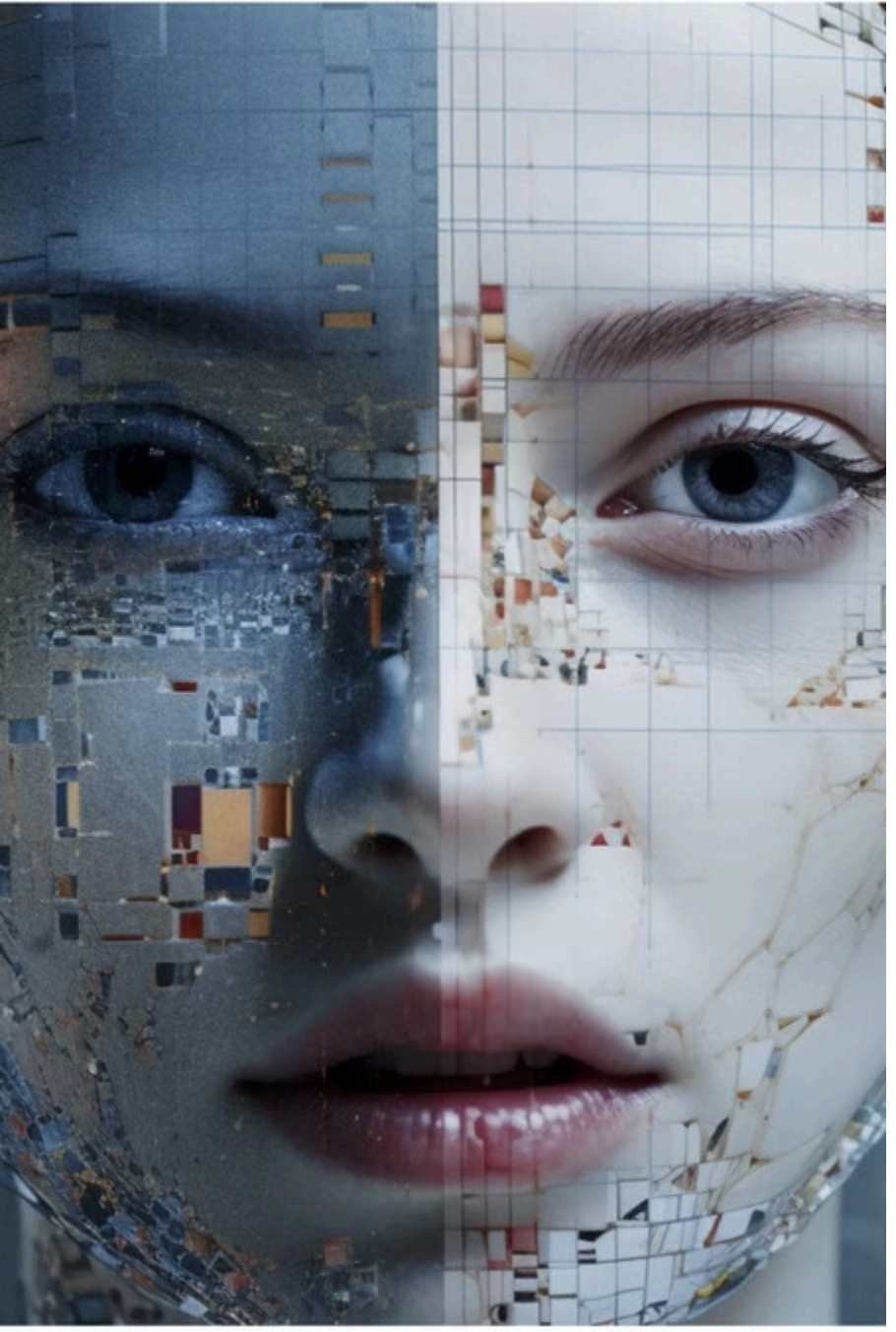
Detect edge, texture, and color anomalies.

4

Temporal Learning

Frame-by-frame temporal learning.





Solution: Frameworks

OpenPose

Human pose estimation framework.

Lip Reading

Analyze lip movements.

LAA-Net

Hierarchical fine-grained models.

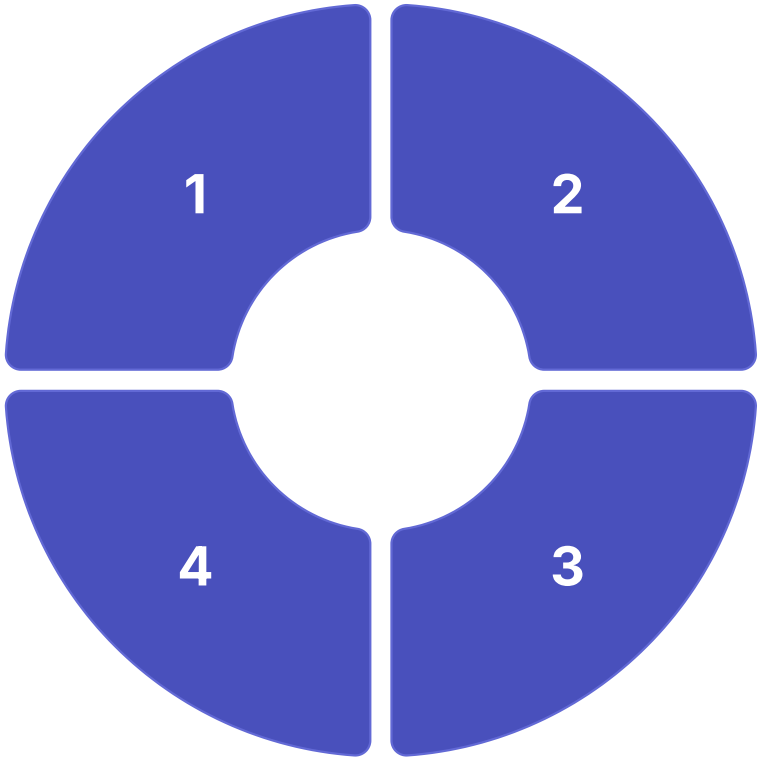
Literature Survey : Temporal Anomaly Analysis in Deepfake Videos

Reference	Methodology	Drawbacks
Rossler, A., et al. (2019) IEEE "Face Forensics: Spatial Deepfake Detection Using CNNs"	• CNN analyzes pixel inconsistencies. • Detects artifacts like blending issues.	• Less accurate with high-quality post-processing. • Noise reduction makes detection harder.
Dang, H., et al. (2020) CVPR "Detecting Manipulated Faces with Attention-Based Convolutional Networks"	• CNN with attention focuses on key facial regions. • Highlights artifacts.	• Struggles with subtle manipulations. • Hard to detect widespread deepfakes.

Existing Methods

Localization-based Detection

Identify manipulated regions.



Multi-Modal Detection

Combine audio and video.

Texture and Artifact Analysis

Detects visual artifacts and texture inconsistencies

Temporal Analysis Detection

Analyze frame consistency.

Challenges

Rapid Evolution

New deepfake techniques emerge fast.

Data Scarcity

Limited high-quality training data.

Generalization

Difficult to detect unseen deepfakes.

Real-time

Difficult to analyse instantly.

Computational Complexity

High resource demand.

Subtle Manipulations

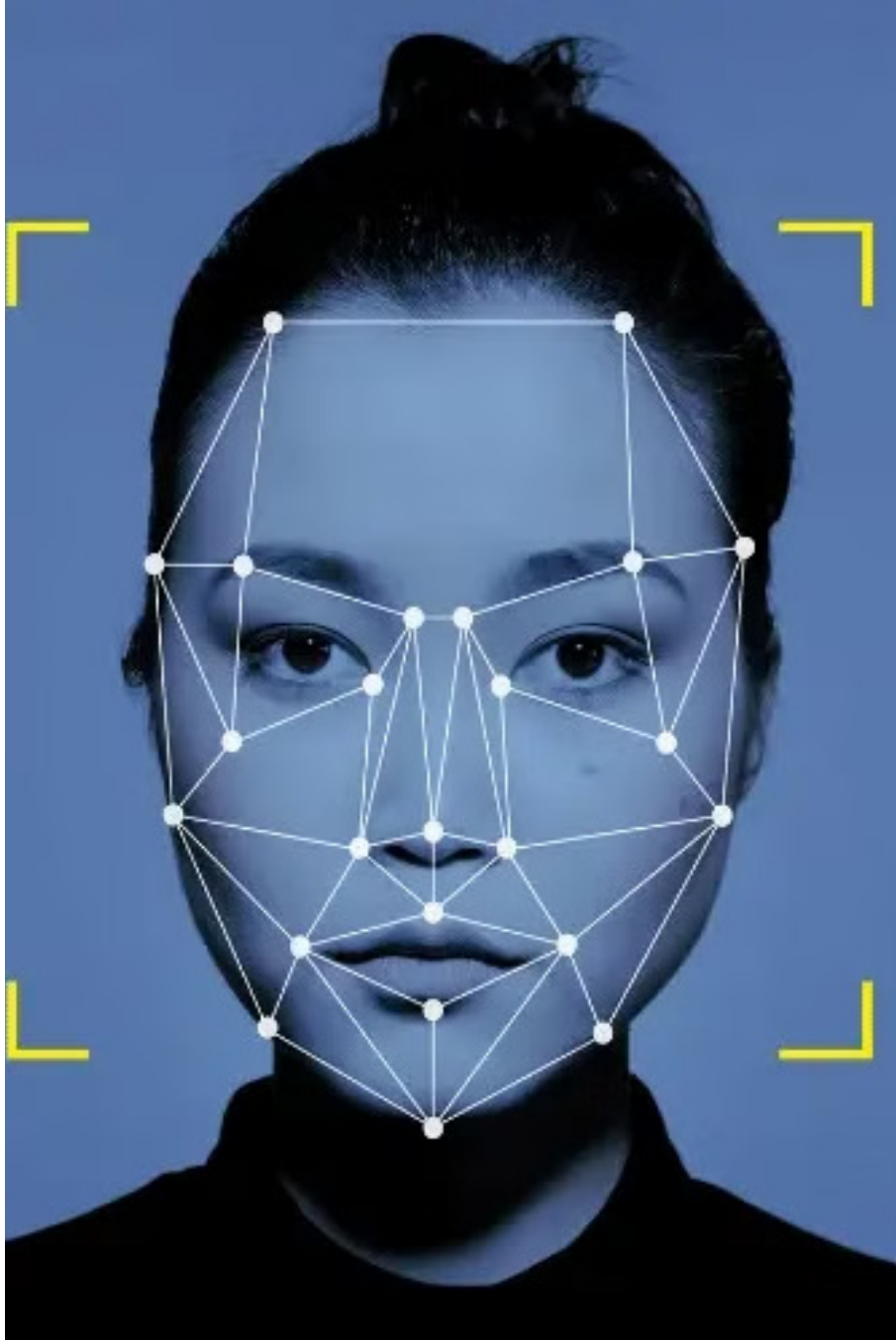
Fine facial details are tough to spot.

Project Objective

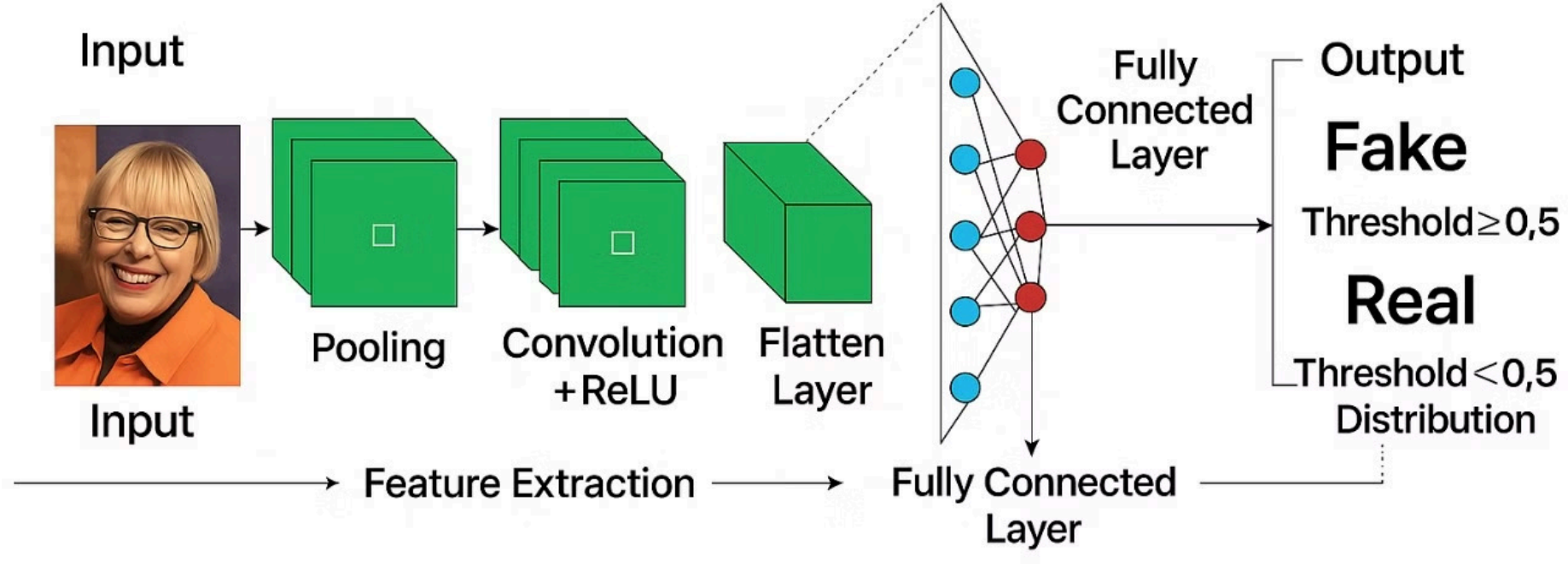
- **Developing a robust deepfake detection system using:**
- **Convolutional Neural Networks (CNNs)** – Detects lighting issues, edge distortions, and blending artifacts.

System Focus:

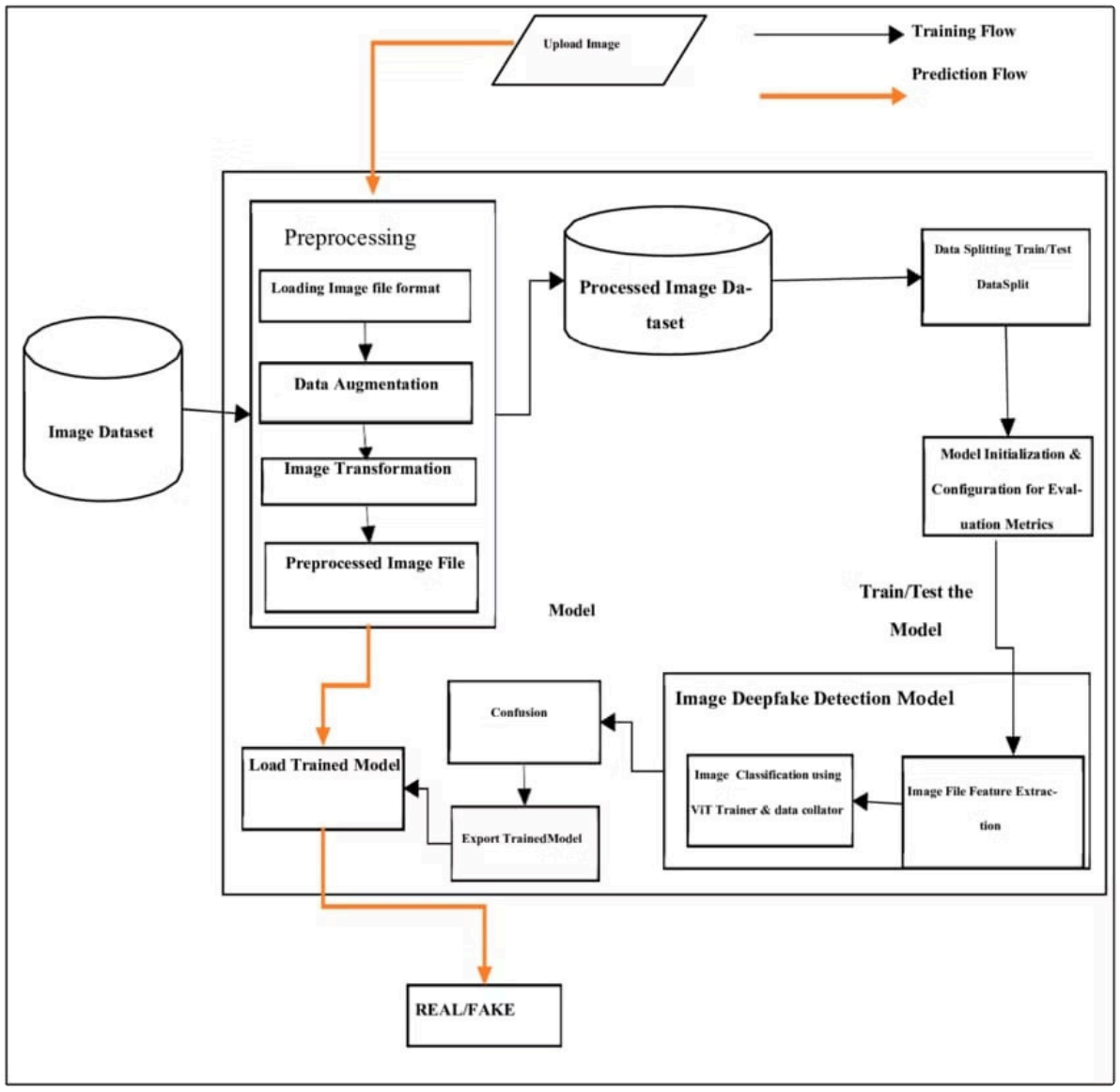
- Detect subtle artifacts and micro-expressions.
- Adapt to evolving deepfake techniques.
- Ensure real-time detection and scalability.



Architecture Diagram



Block Diagram



Feature Extraction using ResNext CNN

The use of ResNext CNN in our project is to extract the features and accurately getting the frame level features also. Our network is finely tuned by addition of extra layers and then choosing the best rate and precisely converge the model gradient descent.

References

1. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. *Deepfakes and beyond: A Survey of face manipulation and fake detection*.
2. Li, Y., Chang, M., Lyu, S., & Jain, A. (2020). *Deepfake Detection: A Survey*.
3. Rossler, A., et al. (2019). *Face Forensics: Spatial Deepfake Detection Using CNNs*. IEEE.
4. Dang, H., et al. (2020). *Detecting Manipulated Faces with Attention-Based Convolutional Networks*. CVPR.
5. Nguyen, H. H., et al. (2021). *Autoencoder-Based Spatial Anomaly Detection for Deepfake Detection*.
6. Sabir, E., et al. (2019). *LSTM-Based Deepfake Detection for Temporal Anomaly Analysis*. ICCV.
7. Guera, D., et al. (2019). *Temporal Inconsistency Detection in Deepfake Videos*. WACV.
8. Daisy-Zhang. *Awesome-Deepfakes-Detection* [GitHub Repository]. Retrieved from **https://github.com/Daisy-Zhang/Awesome-Deepfakes-Detection**.

Thank You