# Deepfake Detection System

## Domain: Deep Learning

Guide By: Prof. Taware R.K.

June 12, 2025

**Presented By:**
Nigade Jaydeep Popatrao
Avatade Rohan Ramdas
Bagal Yashkumar Ramchandra
Shinde Nikhil Kailas

# Problem Statement

To Design and Develop a Deep learning Algorithm to classify videos are either deepfake or genuine (Pristine).

# Introduction

Deepfake is a technique for human image synthesis based on artificial intelligence. Deepfakes are created by combining and superimposing existing images and videos onto source images or videos using a deep learning technique known as generative adversarial network (GAN).

# Motivation

- Fake News
- Malicious Hoaxes
- Financial Fraud
- Celebrity Scandals
- Political Manipulation

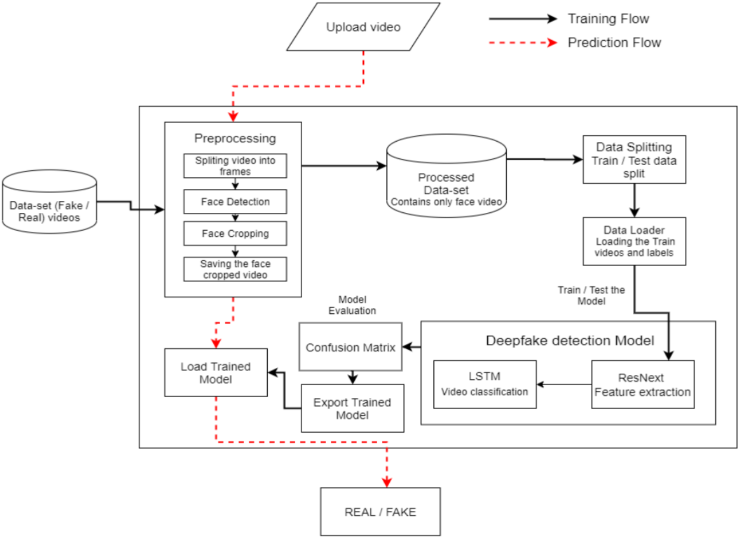
# Objectives

- Create a model to classify videos as deepfake or genuine
- Automatically detect deepfakes with high accuracy
- Improve accuracy and reduce errors in the classification
- Provide a simple interface for users to upload and check videos
- Handle different types of deepfakes effectively.

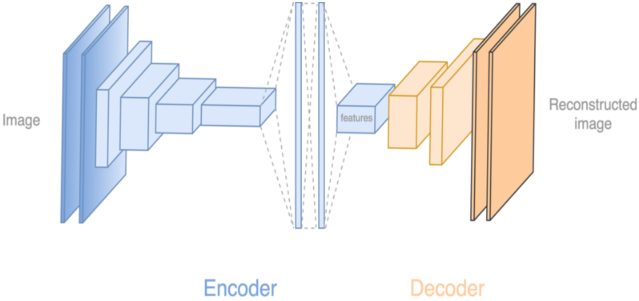
# Technologies

- **Deep Learning and AI:** Utilizing CNNs and LSTMs for pattern recognition in deepfake detection.
- **Data Handling:** Processing datasets like Celeb-DF and FaceForensics++ for training and validation.
- **Web Application Development:** Developing a scalable platform for real-time video analysis and deepfake detection.

# System Architecture



# Deepfake Creation



# Algorithm: Convolutional Neural Networks (CNN)

**Convolutional Neural Networks (CNN):**

- **Purpose:** Used for feature extraction from video frames.
- **Process:**
  1. Extract visual features like edges, textures, and patterns from video frames.
  2. Apply convolution layers to capture spatial hierarchies in the frames.
  3. Use pooling layers to reduce dimensionality and retain important features.
- **Output:** Extracted features used for distinguishing between deepfake and genuine frames.

# Algorithm: Long Short-Term Memory (LSTM)

**Long Short-Term Memory (LSTM):**

- **Purpose:** Designed for sequence modeling of video data.
- **Process:**
  1. Capture temporal dependencies by analyzing sequential video frames.
  2. Use LSTM layers to learn long-term relationships between frames.
  3. Combine sequential data from multiple frames for better context understanding.
- **Output:** Temporal features that help classify videos as deepfake or genuine.

# Algorithm: Generative Adversarial Networks (GAN)

**Generative Adversarial Networks (GAN):**

- **Purpose:** Employed to generate realistic deepfake samples.
- **Process:**
  1. Train a generator to create synthetic deepfake videos.
  2. Use a discriminator to distinguish between real and generated videos.
  3. The generator and discriminator compete to improve their performance.
- **Output:** Generated deepfake samples used for training the detection model.

# Literature Survey

| Reference | Summary | Year |
|---|---|---|
| **Deepfake Detection System** | Hybrid model using ResNet50 for spatial and LSTM for temporal analysis; 92% accuracy on Celeb-DF and FaceForensics++. | 2023 |
| **AI-Based Detection of Deepfakes** | Explores AI-based detection methods to handle more sophisticated deepfakes. | 2022 |
| **Deepfake Detection Techniques** | Latest advances in deepfake detection using improved GAN architectures. | 2021 |
| **Deepfake Video of Mark Zuckerberg** | High-accuracy detection of deepfakes; focuses on preserving biological signals. | 2019 |
| **FaceForensics++** | CNN detects artifacts by comparing generated face areas and their surroundings; lacks temporal analysis. | 2019 |
| **Celeb-DF** | Large-scale dataset for deepfake detection; contains diverse and high-quality data. | 2019 |
| **Face Aging with GANs** | Conditional GANs used to model face aging, tested on controlled datasets. | 2017 |
| **Face2Face** | Real-time face capture and reenactment of RGB videos using face swapping techniques. | 2016 |

# System Requirements

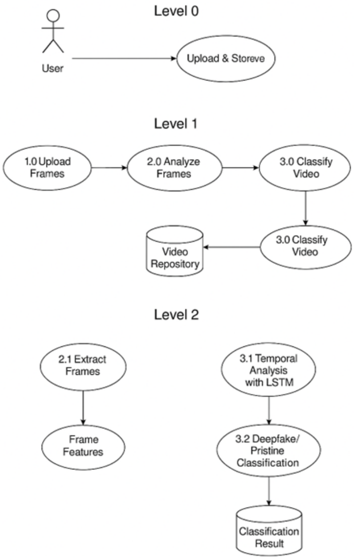
**Hardware:**

- High-performance GPU
- Sufficient storage
- Computer or mobile device

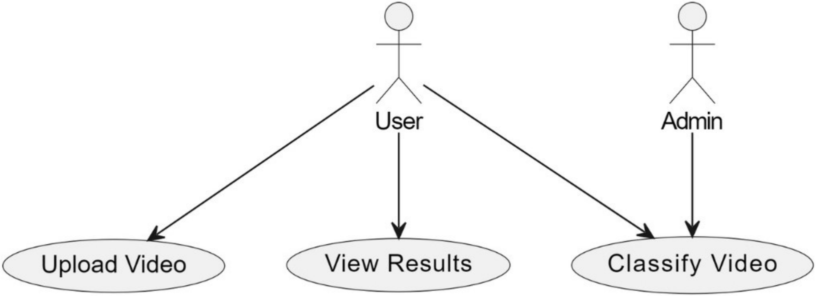
**Software:**

- TensorFlow, Keras, OpenCV, FFmpeg
- Python, Jupyter Notebook, Visual Studio Code
- Windows, macOS, or Linux

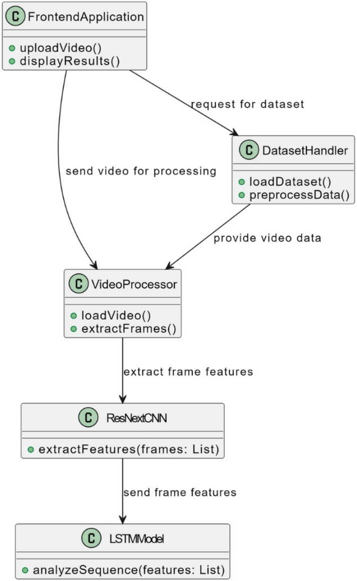
# Data Flow Diagram



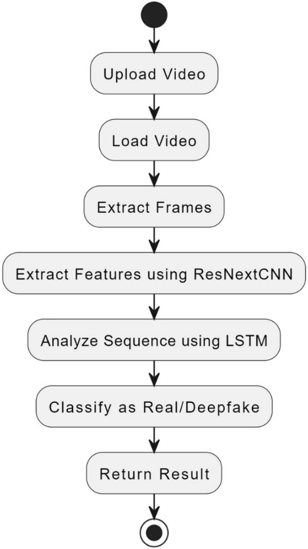
# Use Case Diagram



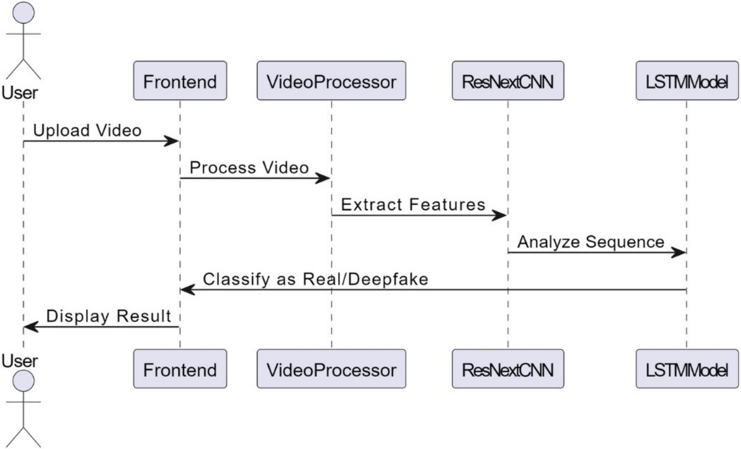
# Class Diagram



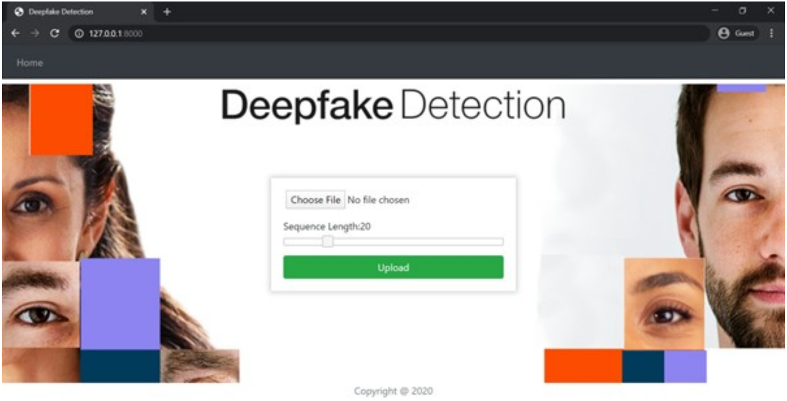
# Activity Diagram


Upload Video
Load Video
Extract Frames
Extract Features using ResNextCNN
Analyze Sequence using LSTM
Classify as Real/Deepfake
Return Result

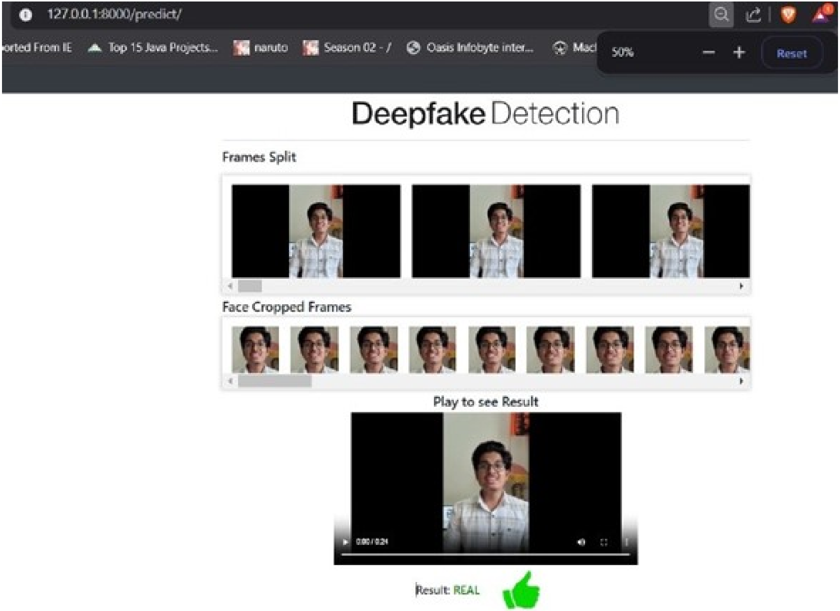
# Sequence Diagram



# Result Diagram



# Result Diagram



Deepfake Detection

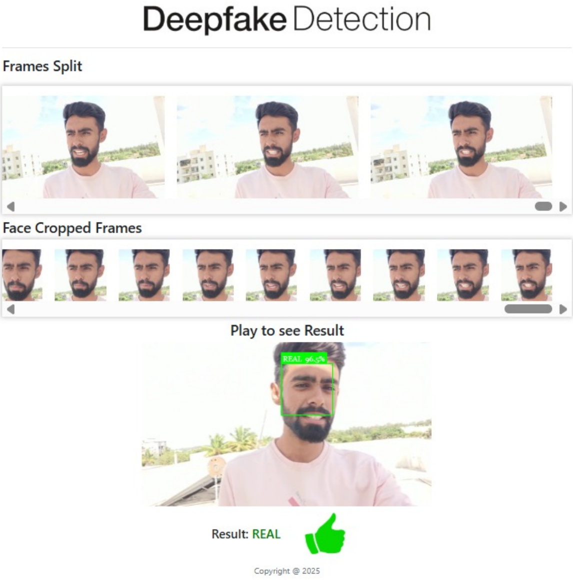
Frames Split

Face Cropped Frames

Play to see Result

Result: REAL

# Result Diagram



Deepfake Detection

Frames Split

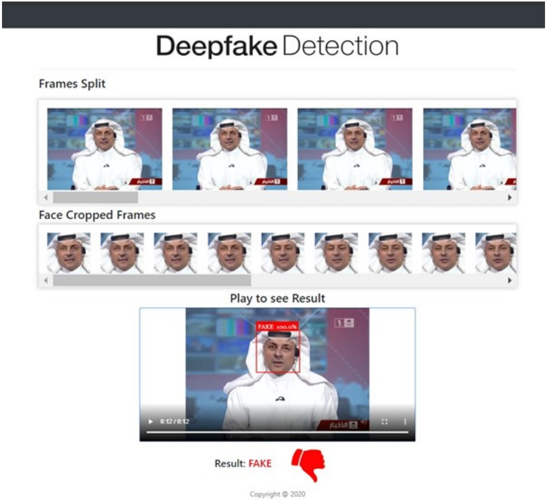
Face Cropped Frames

Play to see Result

Result: REAL

Copyright @ 2025

# Result Diagram



Deepfake Detection

Frames Split

Face Cropped Frames

Play to see Result

Result: FAKE

Copyright © 2020

# Result Accuracy

Results on the FaceForensics++ dataset showed that increasing the sequence length from 10 to 100 frames led to a consistent improvement in accuracy, above 94percentage .

# Conclusion

Deepfake detection is crucial for mitigating digital misinformation. Our system successfully identifies deepfake videos with high accuracy.

# References

- FaceForensics++: Learning to Detect Manipulated Facial Images, arXiv:1901.08971.
- Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics, arXiv:1909.12962.
- Face Aging with Conditional Generative Adversarial Networks, arXiv:1702.01983.
- Face2Face: Real-time Face Capture and Reenactment of RGB Videos, IEEE Conference on Computer Vision and Pattern Recognition.
- Deepfake Video of Mark Zuckerberg Goes Viral on Eve of House A.I. Hearing, accessed on 26 March 2020.

# Thank You!