



**K. Ramakrishnan
College of Technology**

Autonomous

Affiliated to Anna University Chennai, Approved by AICTE New Delhi,

Accredited by NBA and with 'A+' grade by NAAC

Samayapuram, Tiruchirappalli - 621 112, Tamilnadu, India.



**Incubating Minds,
Catalyzing Careers**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

20CS7503 & DESIGN PROJECT 3

END SEMESTER PRACTICAL EXAMINATION - NOV / DEC 2025
REVIEW PRESENTATION

Batch No: 12

Date: 29.11.2025

Session:

AI-POWERED HUMAN DEEPFAKE DETECTION SYSTEM

Guide Name: Mr.P. Matheswaran M.E.,(Ph.D)

AP / CSE

Team Members:

Suvathi R (811722104163)

Varsha RK (811722104173)

Varshita M (811722104176)

OBJECTIVE OF THE PROJECT

1. To develop a system capable of distinguishing real images from deepfakes using AI/ML techniques.
2. To create a reliable model for detecting manipulated images.
3. To enhance trust and security in digital media by reducing misinformation.

INTRODUCTION

- Deepfake technology is one of the most disruptive applications of artificial intelligence, capable of generating highly realistic synthetic images and videos.
- Traditional methods of fake detection are no longer effective due to the sophistication of modern AI models like GANs (Generative Adversarial Networks).
- Our project is designed to address this challenge by applying deep learning, CNN and feature extraction methods to analyze and classify digital content as real or fake.

LITERATURE SURVEY

| TITLE OF PAPER | AUTHOR(S) | PAPER GIST | TECHNOLOGY USED |
|--|--|---|--------------------------------|
| A Survey on Deepfake Detection through Deep Learning (2024) | P. Kamakshi Thai, Sathvik Kalige, Sai Nikhil Ediga, Lokesh Chougoni. | Discusses various deep learning-based approaches and future research challenges. | CNNs, RNNs, GAN-Fingerprinting |
| Analysing the Landscape of Deep Fake Detection: A Survey (2024) | K. Vyas, P. Pareek, R. Jayaswal, S. Patil | Categorizes existing detection techniques and identifies gaps in real-time performance. | CNN + LSTM Hybrid Models |
| A Contemporary Survey on Deepfake Detection: Datasets, Algorithms, and Challenges (2024) | Liang Yu Gong, Xue Jun Li | Highlights datasets, algorithms, and issues in model generalization and domain shift. | CNNs, Vision Transformers |
| Deepfakes Generation and Detection: A Short Survey (2023) | Zahid Akhtar | Summarizes deepfake generation tools and detection approaches with focus on facial media. | CNNs, Feature Extraction |
| Deepfake Detection: A Comparative Analysis (2023) | Sohail A. Khan, Duc-Tien Dang-Nguyen | Compares supervised and self-supervised deepfake detectors on benchmark datasets. | Self-Supervised Transformers |

EXISTING SYSTEM

1. Most current systems rely heavily on visual artifacts.
2. Some systems depend on metadata or compression traces, which can be easily removed by attackers.
3. High computational cost → difficult to run in real-time or on low-resource devices.

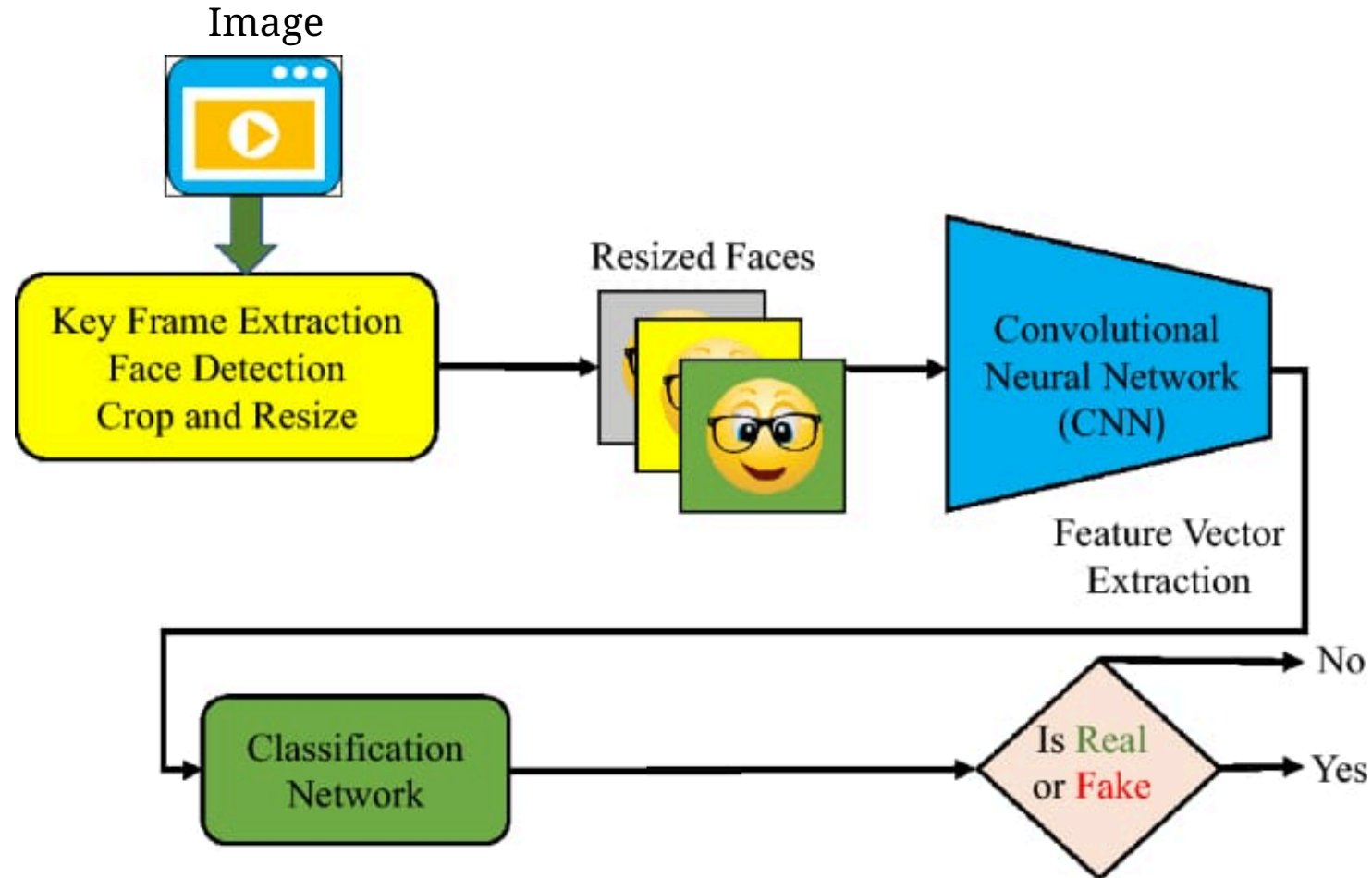
PROBLEM IDENTIFICATION

1. Deepfakes are widely used to spread false news and misinformation, influencing public opinion and creating political or social unrest.
2. They enable misuse of personal identities by cloning faces leading to privacy violations and online harassment.
3. Cybercriminals exploit deepfake technology to conduct financial scams and fraudulent transactions by impersonating trusted individuals.
4. The increasing realism of deepfakes causes people to lose trust in digital media, making it difficult to distinguish between real and manipulated content.

PROPOSED SYSTEM

1. A deep learning-based model is proposed to automatically analyze uploaded images and identify whether they are real or deepfake.
2. The system provides a simple web interface using Streamlit, allowing users to check authenticity without technical knowledge.
3. The model learns distinguishing visual patterns from real and manipulated images, enabling fast and accurate classification.
4. The system outputs a clear authenticity result with confidence score, helping users verify digital media reliability.

PROPOSED SYSTEM ARCHITECTURE



SOFTWARE AND HARDWARE REQUIREMENTS

HARDWARE

1. GPU-enabled system
2. Min. 16 GB RAM.
3. High-speed SSD storage
4. 64-bit operating system

SOFTWARE

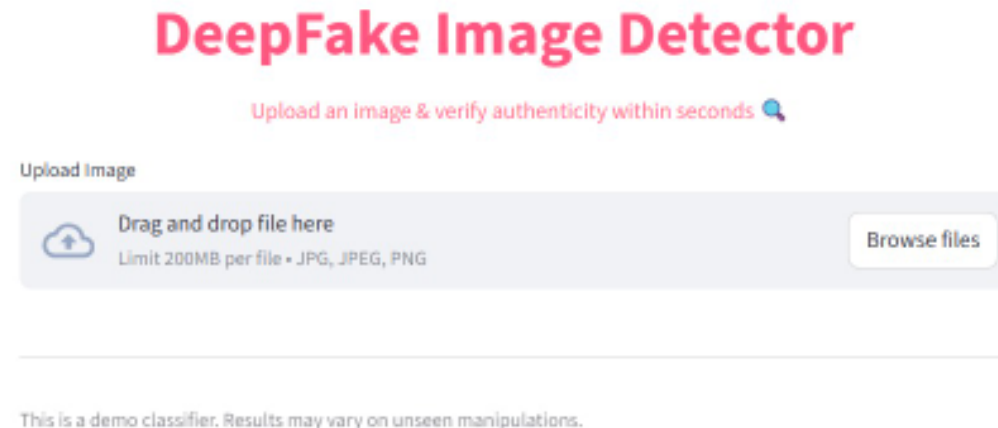
1. Programming Language: Python 3.10+
2. Frameworks: TensorFlow, Keras, OpenCV, NumPy, Streamlit.
5. IDE: VS Code

MODULES

1. Input Acquisition
2. Preprocessing
3. Feature Extraction
4. Classification Model
5. Result & Visualization

IMPLEMENTATION OF INPUT ACQUISITION MODULE

1. The system allows users to upload an image through a Streamlit web interface to start the deepfake detection process.
2. The uploaded file is validated and loaded into the program for further analysis using the trained model.



IMPLEMENTATION OF PREPROCESSING MODULE

1. The input image is converted into a standardized dimension(96 x 96 x 3) required by the deep learning model.
2. Pixel values are normalized to improve classification performance.
3. The image is converted into array form suitable for Machine Learning processing.
4. Basic filtering ensures the input is a clear image with a valid format before prediction.

IMPLEMENTATION OF FEATURE EXTRACTION MODULE

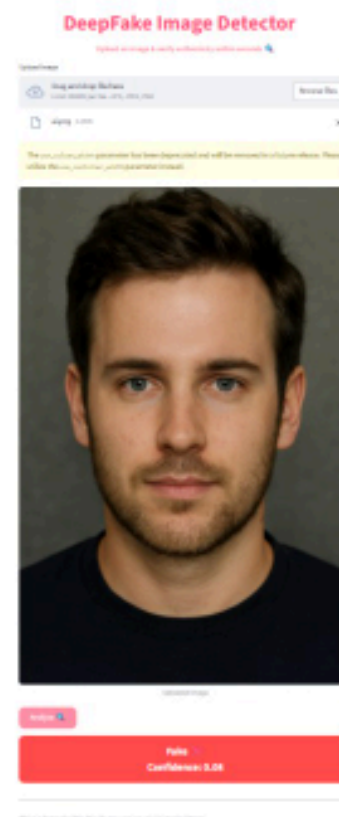
1. A pre-trained Convolutional Neural Network (CNN) extracts meaningful patterns from the image.
2. Hidden layers capture pixel-level variations that may indicate image manipulation.
3. The model learns differences between real and fake facial textures.
4. Extracted features are passed into the classification layer for decision making.

IMPLEMENTATION OF CLASSIFICATION MODEL

1. A deep learning-based classifier analyzes extracted features and predicts whether the image is Real or Fake.
2. The model outputs a probability score indicating confidence of the prediction.
3. Binary classification is used to categorize the media as authentic (real) or tampered (fake).
4. The trained weights enable fast and automated detection without manual evaluation.

IMPLEMENTATION OF RESULT VISUALIZATION MODULE

1. The Streamlit interface displays the final result as “*Real*” or “*Fake*” along with prediction confidence.
2. The detected outcome is shown instantly to the user in a clean and simple UI for practical verification.



CONCLUSION

1. The developed AI-based human deepfake detection system successfully identifies manipulated images by using a CNN-based MobileNet V2 model, ensuring accurate and reliable differentiation between real and fake images.
2. The project enhances digital media security by providing an automated tool that helps prevent misinformation, identity misuse, and online fraud caused by deepfake technology.

FUTURE ENHANCEMENTS

1. The system can be enhanced to detect fake videos by analyzing each video frame for any signs of manipulation.
2. Temporal inconsistencies in facial movements and expressions can be monitored to accurately identify deepfake videos.