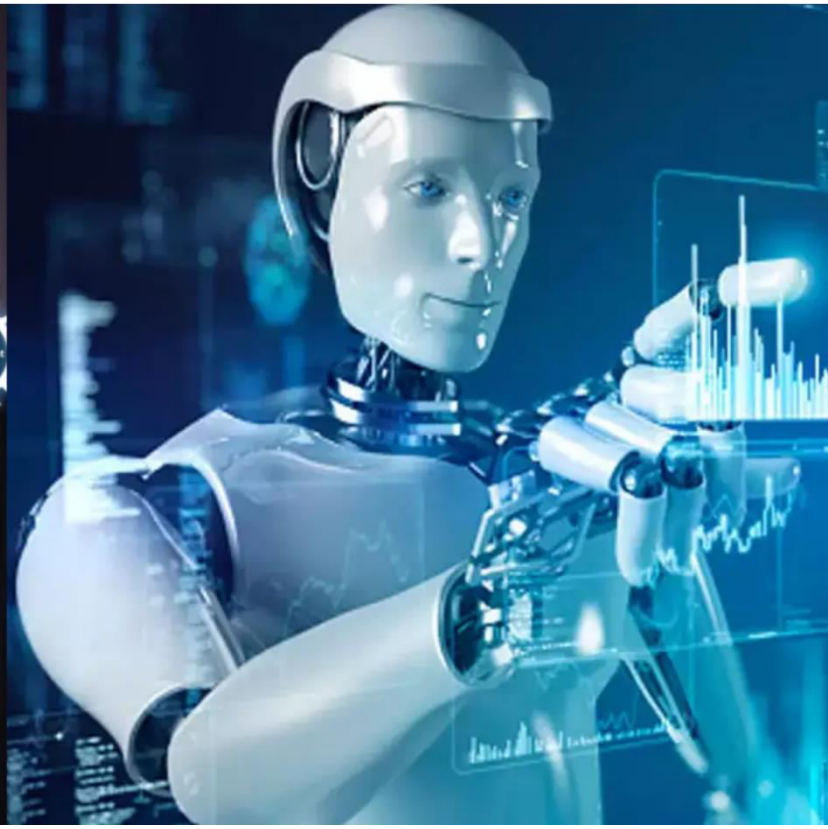
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

Faceswap Detection in Deepfake Videos using AI/ML

Prof. Hamsalatha J, Gaganashree R, Pooja E V, Prakruthi B N, Yashashwini S

Assistant Professor, Department of Information Science and Engineering, S J C Institute of Technology, Chickballapur, Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura, Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura, Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura, Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura, Karnataka, India

ABSTRACT: This project addresses the growing concern of deepfake technology, which manipulates media to create highly realistic but misleading content. To combat this, we propose a comprehensive deepfake detection system capable of analysing both images and videos. For video analysis, we employ Gated Recurrent Units (GRUs) to capture temporal dependencies and identify inconsistencies characteristic of deepfakes. For image analysis, we leverage Convolutional Neural Networks (CNNs), specifically VGG16 and MobileNet, to extract relevant features and classify images as authentic or manipulated. To enhance user accessibility, we develop a web application that allows users to upload images and videos for analysis, receiving accurate classification results in real-time. By combining these techniques, our system aims to contribute significantly to the detection and mitigation of deepfake content, thereby safeguarding the integrity of digital media.

KEYWORDS: GRU, CNN, VGG16, MobileNet, Fake, Real.

I. INTRODUCTION

The primary objective is to develop an AI-powered system to detect deepfake images and videos. By leveraging GRU for videos and CNN, VGG16, and MobileNet for images, the system aims to identify manipulated media. A user-friendly web application will be developed to facilitate easy access and usage. The proliferation of deepfake technology poses a significant threat to the authenticity of digital media. Deepfakes can be used to spread misinformation, manipulate public opinion, and damage reputations. This project aims to address this challenge by developing a robust deepfake detection system. By developing this system, we aim to contribute to the preservation of digital integrity and the fight against the spread of misinformation. By leveraging advanced AI techniques like GRU and CNN, we aim to safeguard the authenticity of online content and mitigate the potential negative impacts of deepfakes. The scope of this project encompasses the development and implementation of a deepfake detection system utilizing GRU for video analysis and CNN, VGG16, and MobileNet for image analysis. The system will be trained on extensive datasets of real and manipulated media to achieve high accuracy. The project will also involve the development of a user-friendly web application to facilitate accessibility and ease of use. Additionally, the system will be evaluated rigorously to assess its performance and robustness. These AI-generated videos and images, often indistinguishable from genuine content, have the potential to mislead, misinform, and cause harm. To combat this growing threat, this project aims to develop a robust deepfake detection system capable of identifying manipulated media with high accuracy. By leveraging advanced deep learning techniques, we will train models to discern subtle inconsistencies and anomalies present in deepfake content. For video analysis, we will employ Gated Recurrent Units (GRUs), a type of recurrent neural

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

network well-suited for processing sequential data. GRUs can capture temporal dependencies within videos, enabling the identification of irregularities in facial expressions, movements, and lighting conditions.

II. PROBLEM STATEMENT

The proliferation of deepfake technology poses a significant threat to the authenticity of digital media. Deepfakes can be used to spread misinformation, manipulate public opinion, and damage reputations. This project aims to address this challenge by developing a robust deepfake detection system. The system will leverage advanced deep learning techniques to analyse both images and videos, identifying subtle manipulations that are often undetectable to the human eye. By developing this system, we aim to contribute to the preservation of digital integrity and the fight against the spread of misinformation.

III. LITERATURE REVIEW

1. Author: Ahmed, M., & Yadav, R. (2023)
Title: *Machine Learning-Based Models for Early Plant Disease Diagnosis*
Outcome: Developed ML models (random forest, k-NN, linear regression, Naive Bayes, neural networks, SVM) using image processing for early detection of plant diseases. The ensemble model showed the highest accuracy using metrics like precision, recall, and F1-score.
Disadvantage: The approach may face limitations in generalizing across vastly different plant species or environmental conditions without large and diverse datasets.

2. Author: Naralasetti, R., & Bodapati, J. (2023)
Title: *Transfer Learning for Enhanced Plant Leaf Disease Prediction*
Outcome: Utilized pre-trained CNNs to improve classification accuracy by extracting deep features, addressing issues of sparse labeled data in agriculture.
Disadvantage: Transfer learning models may still struggle with real-time adaptability and require significant computational resources.

3. **Author:** Sajitha, K., et al. (2023)
Title: *Image-Based Plant Disease Classification Using ML and DL*
Outcome: Reviewed ML and DL techniques with a focus on sophisticated feature extraction and integration into industrial agriculture systems.
Disadvantage: The complexity of integration into legacy farming systems and lack of standardization in data collection remains a barrier.

4. **Author:** Ushadevi, P. (2023)
Title: *Survey on Deep Learning and Machine Learning for Plant Disease Prediction*
Outcome: Emphasized the need for robust architectures and large annotated datasets; detailed preprocessing and model evaluation strategies.
Disadvantage: Heavy reliance on large, high-quality datasets which may not be feasible for all crops or regions.

5. **Author:** Fuentes, A., et al. (2023)
Title: *Real-Field Plant Disease Detection Using CNNs*
Outcome: Designed a robust, real-time CNN-based system for plant disease detection across varied environments, enabling fast farmer response.
Disadvantage: May require high-quality imaging hardware and stable connectivity, which can be challenging in rural settings.

6. **Author**: Tsaftaris, S. A., et al. (2023)
Title: *Image Processing in ML Frameworks for Plant Phenotyping*
Outcome: Highlighted the role of segmentation and enhancement techniques in improving phenotypic predictions through image processing.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

Disadvantage: Accuracy depends heavily on preprocessing quality, and phenotyping tasks may vary widely by crop and region.

7. Author: Lee, H., et al. (2023)
Title: ***Deep Learning-Based Hierarchical Leaf Feature Learning for Classification***
Outcome: Investigated CNN layer-wise learning of plant features, improving interpretability and classification via attention mechanisms**.**
Disadvantage: Attention-based models can be computationally intensive and may require fine-tuning for specific plant types.

8. Author: Buja, A., et al. (2023)
Title: *AI and IoT-Based Early Detection of Plant Diseases and Pests*
Outcome: Combined AI with IoT sensors to enable real-time environmental monitoring and adaptive pest control strategies.
Disadvantage: Sensor network deployment and maintenance costs can be high, and connectivity remains a challenge in remote areas.

9. **Author:** Kamilaris, A., et al. (2023)
Title: *Agri-IoT: Semantic Framework for Smart Farming*
Outcome: Proposed a semantic AI-IoT framework for interoperable smart farming, improving crop yield and decision-making efficiency.
Disadvantage: Integration with legacy equipment and varying data formats can pose interoperability challenges.

10. **Author:** Senthilkumar, P., et al. (2023)
Title: *IoT and Machine Learning-Based Smart Agriculture Solution*
Outcome: Developed a predictive system using sensor data and ML for soil and crop health, with reinforcement learning for irrigation optimization.
Disadvantage: Requires reliable sensor calibration and may not generalize well across different climatic zones without adjustments.

IV. DESIGN AND IMPLEMENTATION

The design of the proposed deepfake detection system emphasizes modularity, scalability, and real-time performance. It integrates deep learning models such as VGG16 and MobileNet for image analysis and GRU for video analysis to capture both spatial and temporal features. A user-friendly web interface allows users to upload media files, which are then pre-processed and classified by the trained models. The system architecture ensures seamless interaction between the frontend, backend, and database, enabling efficient processing and accurate prediction of manipulated content. This design supports future enhancements and real-world deployment across diverse platforms.

System Architecture Overview
The system architecture is structured into three main layers: the presentation layer, the application layer, and the data layer. The presentation layer consists of a web-based user interface that allows users to upload images or videos for analysis. The application layer, built using Python and Flask, handles data preprocessing, model inference, and result generation using deep learning models like VGG16, MobileNet, and GRU. The data layer utilizes a MySQL database to manage user information and log detection results. This layered architecture ensures efficient data flow, maintainability, and scalability for future integration and deployment.

Activity Diagram
The Activity Diagram below illustrates the stepwise flow of user interactions and system processes in deepfake detection:
Stepwise Explanation:

1. Start: The process begins when a user accesses the web application.
2. User Login/Register: The user either registers for a new account or logs in using existing credentials.
3. Upload Media: The user selects and uploads an image or video file through the application interface.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

4. Input Validation: The system validates the uploaded file (e.g., format, size, type) to ensure compatibility.
5. Preprocessing: For images the normalization, resizing, and enhancement and for videos frame extraction and resizing.
6. Feature Extraction: For images: spatial features are extracted using VGG16 or MobileNet and in videos temporal features are extracted using GRU across frames.
7. Model Prediction: The respective model analyzes the features and classifies the input as *Real* or *Fake*.
8. Display Result: The result along with a confidence score is displayed to the user.
9. Logout or New Prediction: The user can either log out or upload another file for analysis.
10. End: The activity ends once the user exits the application.

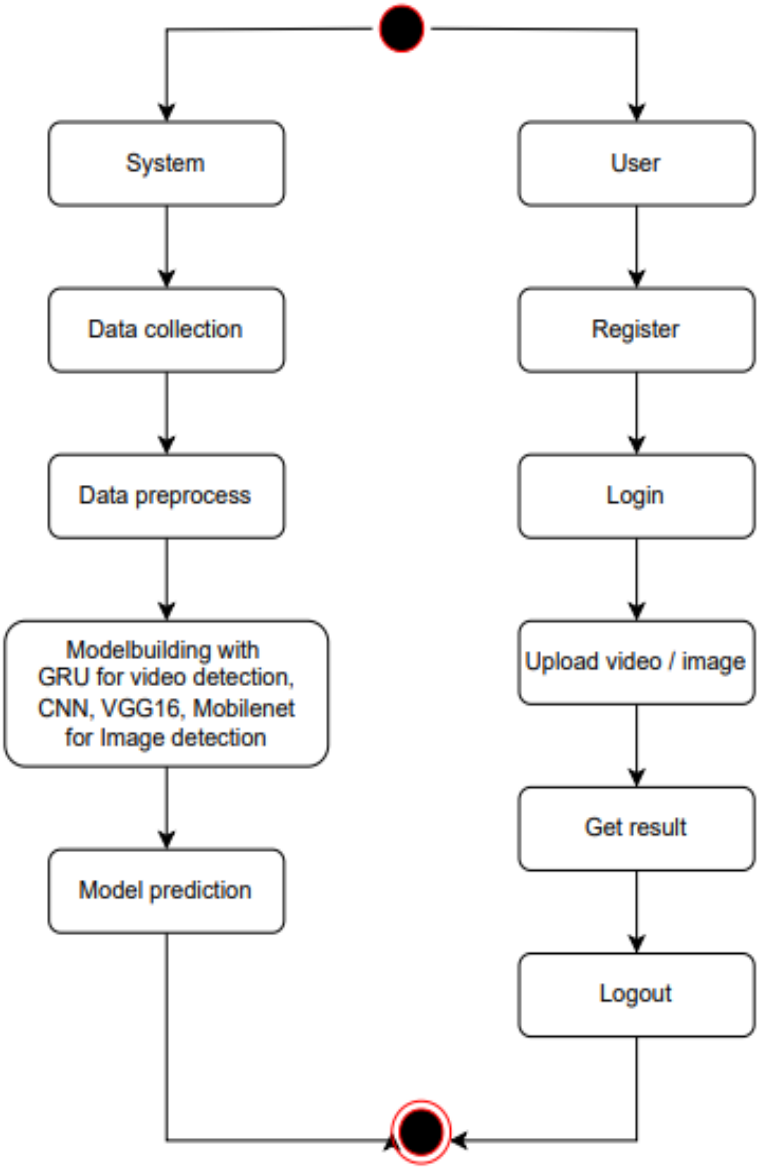


Fig. 1. Activity Diagram of the Deepfake Detection

The activity diagram provides a visual representation of the step-by-step workflow in the deepfake detection system. It outlines the sequence of actions starting from user login or registration, followed by media upload, preprocessing, feature extraction, deep learning model prediction, and result display. This diagram helps in understanding the dynamic

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

behaviour of the system by mapping the flow of control between different activities, ensuring clarity in the execution process from start to end.

Use Case Diagram

The use case diagram illustrates the interactions between users and the deepfake detection system, highlighting roles such as the user and admin, and actions including registration, login, media upload, deepfake analysis, viewing results, and logout.

Role: User

1. **Register** – Create a new account by providing personal details.
2. **Login** – Access the system using valid credentials.
3. **Upload Media** – Upload an image or video for deepfake analysis.
4. **Receive Prediction** – View the result (real or fake) after model processing.
5. **Logout** – Exit the system securely.

Role: System

1. **Validate User Credentials** – Verifies login or registration details during authentication.
2. **Accept Media Upload** – Accepts image or video files uploaded by the user.
3. **Preprocess Input** – Normalizes and resizes media for analysis.
4. **Extract Features** – Applies deep learning models (CNN/GRU) to extract relevant features.
5. **Predict Real or Fake** – Classifies the media as authentic or manipulated using trained models.
6. **Display Results** – Shows the prediction outcome and confidence score to the user.
7. **Logout Session** – Ends the user's session upon logout.

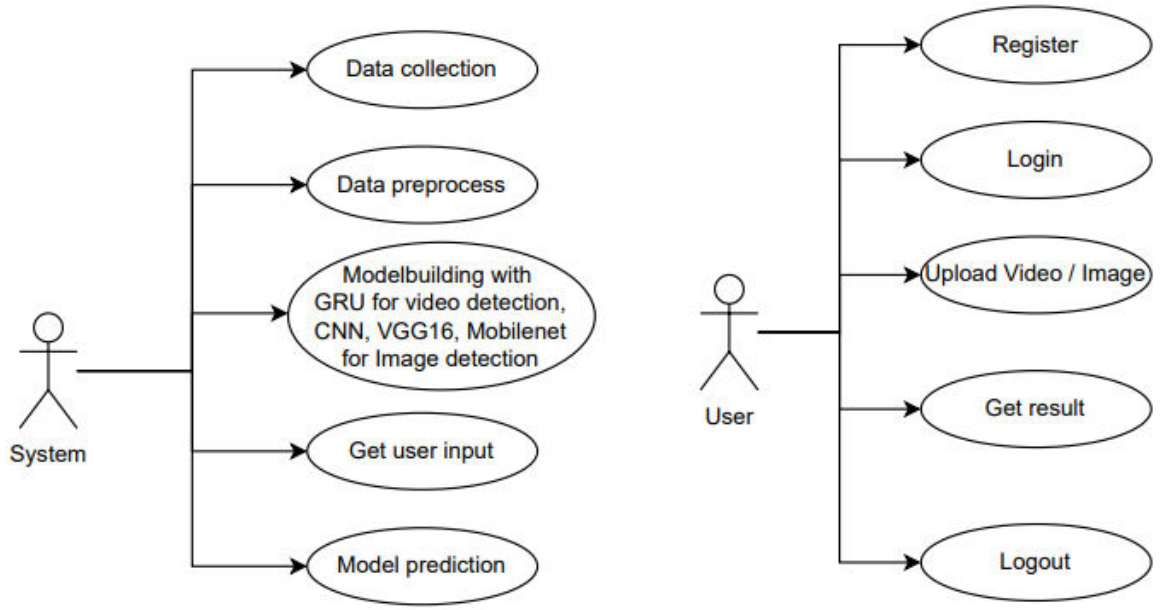


Fig. 2. Use Case Diagram for deepfake detection

Implementation Details

1. Frontend Development: The user interface is built using HTML, CSS, JavaScript, and Jinja2 templating, providing a responsive and intuitive platform for media upload and result display.

2. Backend Development: The backend is implemented using Python with the Flask framework to handle API requests, media preprocessing, and integration with deep learning models.

3. Deep Learning Models: For image detection, VGG16 and MobileNet are fine-tuned to classify images as real or fake; for video detection, GRUs process frame sequences to capture temporal inconsistencies.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405292|

4. Data Acquisition and Preprocessing: Data is collected from public datasets such as FaceForensics++, and preprocessing includes resizing, normalization, and augmentation techniques like rotation and flipping.
5. Database Design: A MySQL database stores user information, media metadata, prediction results, and detection logs for persistent and structured data management.
6. Report Generation: After classification, the system generates a downloadable PDF report detailing the result, confidence score, model used, and timestamp.
7. System Deployment: The application is deployed on a local XAMPP server with Flask and MySQL; MobileNet enables future scalability for real-time detection on resource-constrained edge devices.

V. CONCLUSION

This project aims to address the growing concern of deepfake technology by developing a robust and accurate deepfake detection system. By leveraging advanced deep learning techniques such as GRU and CNN, the system can effectively identify manipulated media, protecting the integrity of digital information. The proposed system, combined with a user-friendly web application, offers a practical solution to combat the spread of deepfakes.

he integration of these models into a user-friendly web application enhances accessibility, allowing users to conveniently upload media and receive real-time analysis results. The modular architecture of the system also supports scalability and easy integration into larger digital security frameworks, such as social media moderation platforms or content verification tools for news organizations.

Future research directions include exploring more sophisticated deep learning architectures, incorporating multimodal analysis (combining audio, text, and visual cues), and developing real-time detection systems. Additionally, addressing ethical considerations and ensuring responsible use of deepfake detection technology will be crucial. By continuously improving the system and staying ahead of evolving deepfake techniques, we can contribute to a safer and more trustworthy digital environment.

REFERENCES

[1] S. Ahmed and V. Yadav, "Machine learning-based models for early plant disease diagnosis using image processing," 2023.
[2] S. Naralasetti and S. Bodapati, "Improved plant leaf disease prediction using transfer learning with CNNs," 2023.
[3] S. Sajitha, P. Kumar, and A. Ramesh, "Image-based plant disease classification using machine learning and deep learning techniques," 2023.
[4] K. Ushadevi, "Survey on plant disease prediction using ML and DL," 2023.
[5] A. Fuentes, M. Hernandez, and A. Molina, "Real-field deep learning model for plant disease detection under varied environmental conditions," 2023.
[6] S. Tsaftaris, D. Minervini, and C. Dupuy, "Role of image processing in plant phenotyping using machine learning frameworks," 2023.
[7] H. Lee, J. Park, and M. Kim, "Hierarchical feature extraction for plant classification using deep learning with attention mechanisms," 2023.
[8] L. Buja, F. Rossi, and G. Moretti, "AI-driven IoT system for early disease and pest detection in agriculture," 2023.
[9] A. Kamilaris, F. Gao, and M. Prenafeta-Boldú, "Agri-IoT: A semantic AI-IoT framework for smart farming," 2023.
[10] R. Senthilkumar, N. Rajkumar, and K. Balamurugan, "IoT-based smart agriculture using machine learning and reinforcement learning for resource optimization," 2023.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details