IJMSERH

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

Volume 14, Issue 2, April-June 2026

Impact Factor: 9.274

editor@ijmserh.com

www.ijmserh.com

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

Deep Learning-Based Approach for Deepfake Image and Video Detection

Ms. D.Kanimozhi[1], Mrs. M.Vijayalakshmi[2]

PG Scholar, Department of Master of Computer Applications, R.V.S. College of Engineering, Dindigul, Tamil Nadu, India[1]

Assistant Professor, Department of Master of Computer Applications, R.V.S. College of Engineering, Dindigul, Tamil Nadu, India[2]

ABSTRACT: The rapid advancement of artificial intelligence and deep learning has enabled the creation of highly realistic manipulated media, commonly known as deepfake.These synthetic images and videos pose significant risks in terms of misinformation, fraud, and privacy violations.

This project presents a Python-based approach for detecting deepfake content using advanced machine learning and image processing techniques. The proposed system leverages Convolutional Neural Networks (CNN) and feature extraction methods to identify subtle inconsistencies in facial features, eye movement, and texture patterns that are often imperceptible to the human eye.

By combining preprocessing steps, such as frame extraction, normalization, and data augmentation, with a robust deep learning model, the system achieves high accuracy in distinguishing real content from deepfake. The proposed method achieves high accuracy and performs better than traditional techniques. This system can be applied in areas like social media monitoring, cybersecurity, and digital content verification to ensure authenticity.

Experimental results demonstrate that this approach can efficiently detect manipulated media, providing a practical tool for mitigating the risks associated with deepfake technology. Based on these features, the system classifies the input as real or fake and provides a confidence score.Performance evaluation using the ROC curve shows an AUC of 0.899, indicating high accuracy with low false positives. Overall, the system effectively detects manipulated media and helps reduce the risks associated with deepfakes.

I. INTRODUCTION

In recent years, the rapid growth of Artificial Intelligence (AI) and deep learning technologies has led to significant advancements in digital media creation. One of the most impactful developments is the emergence of deepfake technology, which uses sophisticated algorithms such as Generative Adversarial Networks (GANs) to create highly realistic fake images and videos. While this technology has useful applications in entertainment and media production, it also poses serious threats including misinformation, identity theft, fraud, and privacy violations.

Detecting deepfake content has become increasingly challenging due to its high visual quality, making it difficult for humans and traditional forensic techniques to identify manipulated media. Conventional methods often rely on manual inspection or basic image analysis, which are time-consuming and less accurate when dealing with large datasets and advanced forgery techniques.

The proposed system utilizes Convolutional Neural Networks (CNNs) to automatically extract meaningful features and identify subtle inconsistencies in facial expressions, texture patterns, and eye movements that are not easily visible to the human eye. The system incorporates preprocessing techniques such as frame extraction, normalization, and data augmentation to improve model performance and accuracy.

By leveraging the power of deep learning, this project aims to develop an efficient and automated solution for deepfake detection. The system not only enhances detection accuracy but also ensures scalability and real-time analysis, making it suitable for applications in social media monitoring, cybersecurity, and digital content verification. Overall, this work contributes to improving trust and authenticity in digital media by providing a reliable method to combat the growing challenges posed by deepfake technology.

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

II. EXISTING SYSTEM

In the existing system, detection of deepfake videos and images primarily relies on manual inspection or traditional image/video forensics techniques. Manual inspection is time-consuming and largely unreliable, as highly realistic deepfake are difficult to detect by the human eye.

Conventional forensic methods, such as analyzing pixel-level inconsistencies, color mismatches, or compression artifacts, often fail when deepfake are generated using advanced generative models like GANs (Generative Adversarial Networks). Additionally, existing automated systems frequently use simple machine learning classifiers on handcrafted features, which struggle to generalize to diverse datasets or detect subtle manipulations.

DISADVANTAGES
Low detection accuracy
Limited scalability

III. PROPOSED SYSTEM

The proposed system introduces an automated deepfake detection framework implemented in Python, leveraging Convolutional Neural Networks (CNN) for feature extraction and classification. The system first preprocesses videos by extracting frames and applying normalization, resizing, and data augmentation to create a robust training dataset.

The CNN model is trained to learn subtle spatial and temporal inconsistencies present in deepfake, such as irregular facial movements, eye blinking patterns, and texture anomalies. By using a deep learning approach, the system can automatically extract relevant features without relying on handcrafted methods, ensuring higher accuracy and better generalization across different deepfake generation techniques.

The proposed solution supports real-time detection and can be integrated into social media monitoring, video verification platforms, or cybersecurity systems, providing an effective tool for combating misinformation and malicious digital content.

ADVANTAGES:
High Accuracy
Automated Detection

IV. SYSTEM ARCHITECTURE

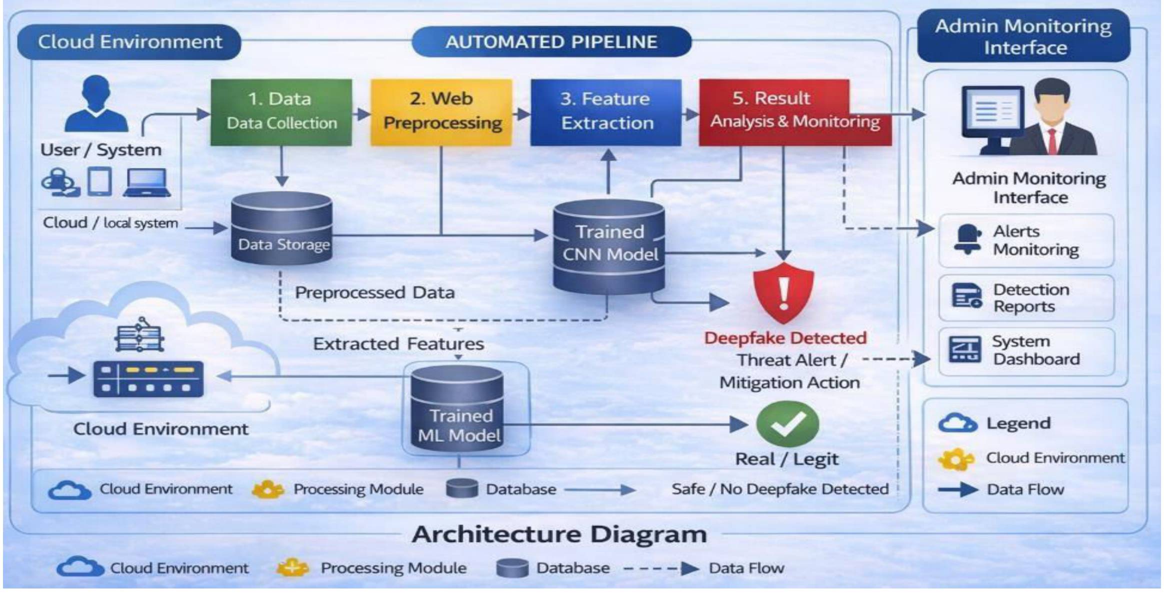

Fig.No.4.1

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

MODULES:

- Input Acquisition
- Preprocessing Module
- Feature Extraction
- Deep Learning Classification
- Output & Decision

V. MODULES DESCRIPTIONS

Input Acquisition :

- This module is responsible for collecting image or video data from users. It supports common formats such as JPG, PNG, MP4, and AVI.
- The module validafle type and size before processing to ensure system reliability. It serves as the entry point of the deepfake detection system and forwards valid data to the preprocessing stage.

Preprocessing Module :

- The preprocessing module prepares raw media data for model analysis. For videos, it extracts frames at regular intervals. It performs face detection and cropping to focus on important facial regions.
- Additional steps like resizing, normalization, and basic data augmentation improve model accuracy and consistency. This stage ensures standardized and high-quality input for deep learning.

Feature Extraction :

- This module uses Convolutional Neural Networks (CNN) to automatically extract meaningful features from images or video frames. It identifies subtle patterns such as texture inconsistencies, abnormal facial expressions, unnatural lighting, and irregular eye blinking.
- These deep features help distinguish real content from manipulated deepfake media more effectively than traditional handcrafted methods.

Deep Learning Classification :

- The classification module processes extracted features through multiple neural network layers, including convolution, pooling, and fully connected layers. The model is trained using real and deepfake datasets to learn distinguishing characteristics.
- It outputs a prediction label (Real or Fake) along with a confidence score, ensuring accurate and reliable detection

Output & Decision Module :

- This module presents the final detection result to the user in a clear and understandable format. For videos, it may combine predictions from multiple frames to make a final decision.
- The output includes the classification label and confidence percentage. It can also generate reports or visual indicators, making it suitable for integration into social media monitoring and cybersecurity systems.

VI. RESULT

1. Install Python Package in Jupyter Notebook

The Jupyter Notebook setup used to install all of the necessary Python libraries for the deepfake detection system is depicted in this picture. The notebook includes package installation commands for essential frameworks such as TensorFlow, NumPy, Pandas, Scikit-learn, OpenCV, Librosa, and Matplotlib. Deep learning, data pretreatment, audio analysis, picture processing, and result display are all supported by these packages. The output part attests to the packages' successful installation and readiness for project development.

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

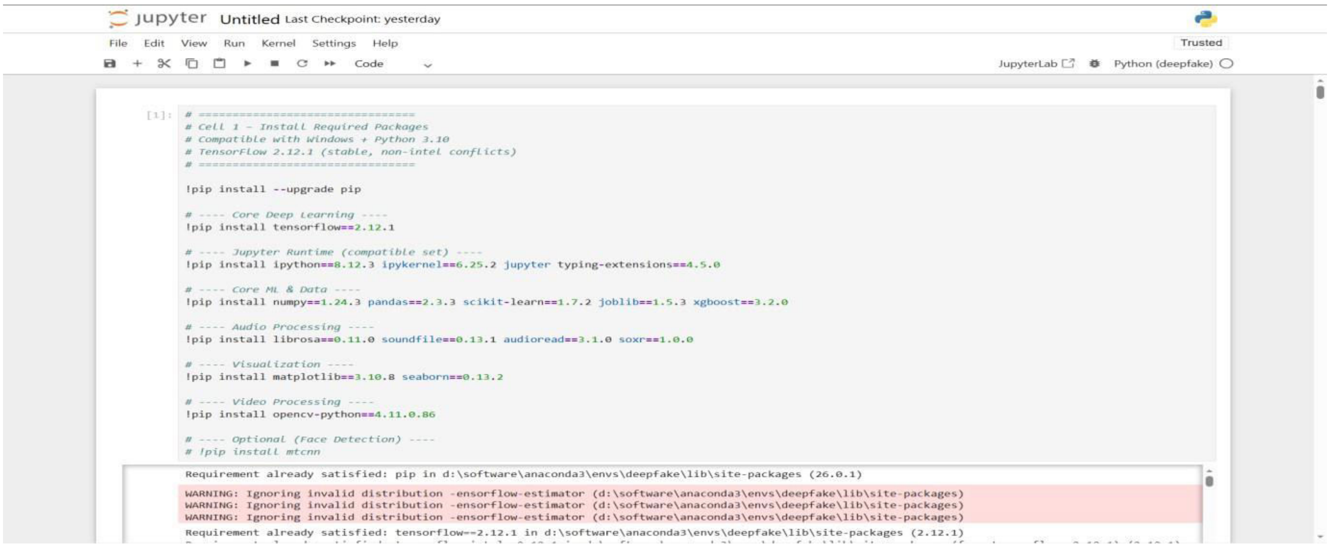

Fig.No:6.1

2. Import Python Libraries

This figure illustrates the successful installation verification and import of all required Python libraries in the Jupyter Notebook environment. The output section confirms that key packages such as TensorFlow and its dependencies were already installed in the deepfake virtual environment. The subsequent code cell imports essential modules including NumPy, Pandas, Matplotlib, Seaborn, OpenCV, and Pickle, which are used for data preprocessing, visualization, model training, and result storage. To guarantee that the deepfake detection system operates consistently and cleanly, random seed initialization and warning suppression are also used.

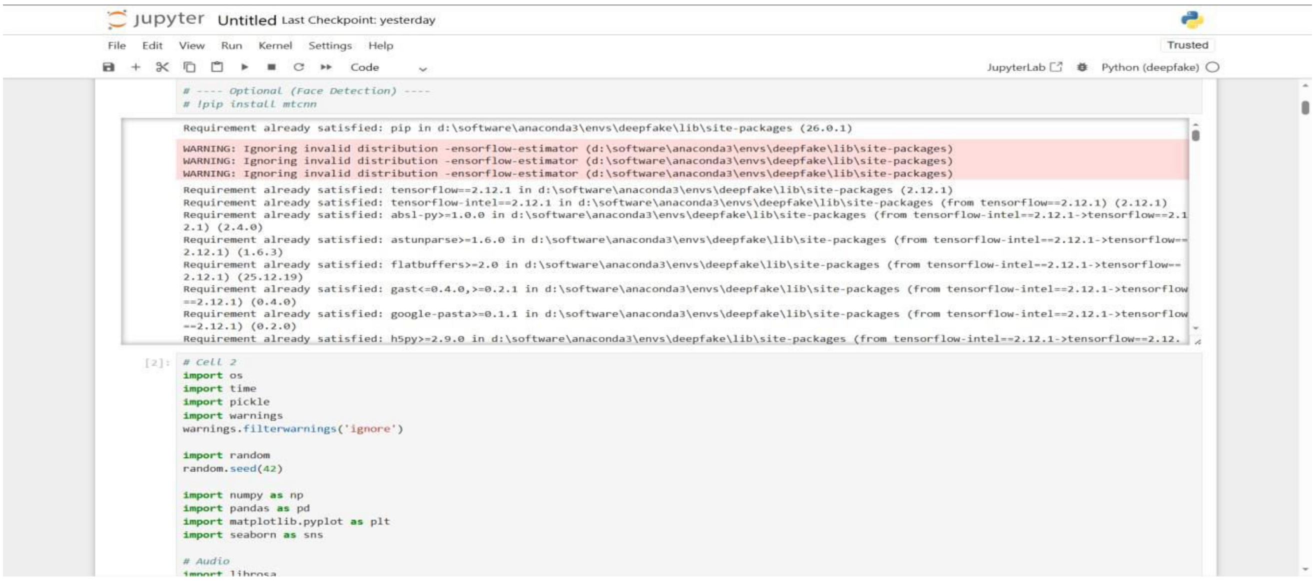

Fig.No:6.2

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

3. Feature Extraction and Dataset Split Processing

The deepfake detection system's feature extraction phase is shown in this picture. The Jupyter Notebook code processes the training, validation, and evaluation datasets by extracting meaningful features from each image and saving them as NumPy arrays. The number of files processed in real time is shown in the progress log, which verifies that every sample was correctly transformed into feature vectors. To gauge computing performance, the entire extraction time is also noted. After extraction, the generated feature sets and labels are assigned to the corresponding training, validation, and test datasets for subsequent model training and evaluation.

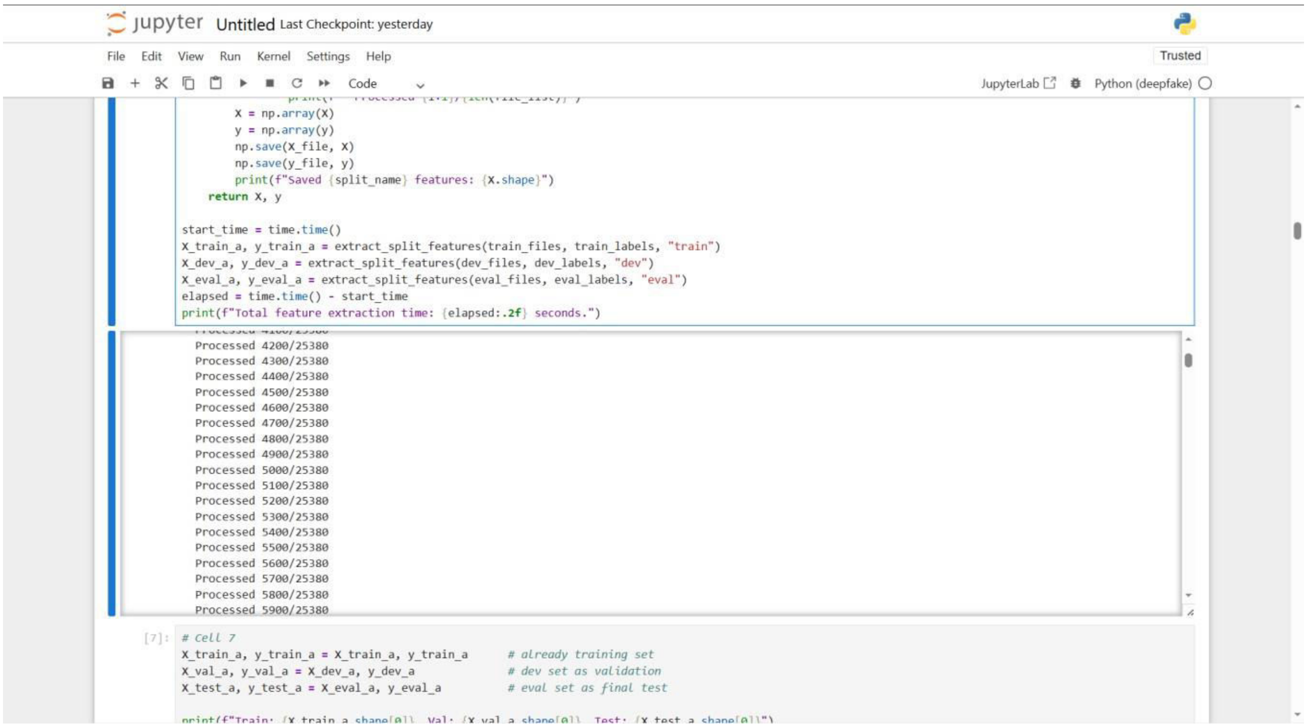

Fig.No:6.3

4. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve of the XGBoost model tested on the audio deepfake validation dataset is shown in this picture. At various categorization thresholds, the True Positive Rate is shown versus the False Positive Rate on the ROC curve. With an Area Under the Curve (AUC) score of 0.899, the model demonstrated excellent classification performance and a high capacity to discriminate between authentic and fraudulent audio samples. The model outperforms random guessing, as seen by the orange curve's continued strong elevation above the diagonal reference line. This outcome validates the XGBoost classifier's and the retrieved audio characteristics' efficacy in detecting deepfake audio.

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

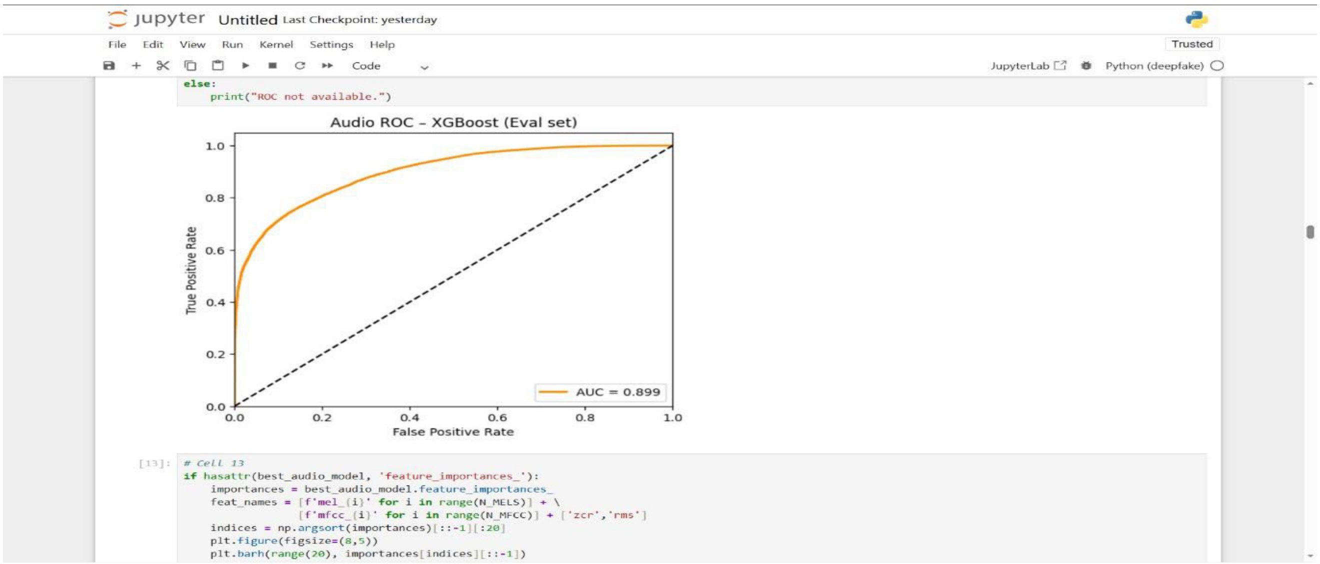

Fig No 6.4

5. Face Region Detection and Visualization for Deepfake Analysis

The Jupyter Notebook's predict_and_visualize(test_image) function's result is seen in this figure. The technique uses a yellow contour to emphasize the facial boundary after identifying the facial region in the input image. By isolating the most crucial region for deepfake analysis, this phase enables the model to concentrate on face characteristics that frequently include manipulation artifacts. The deepfake detection system's following feature extraction and classification phases are more accurate when the visualization verifies that the face has been correctly recognized and segmented.

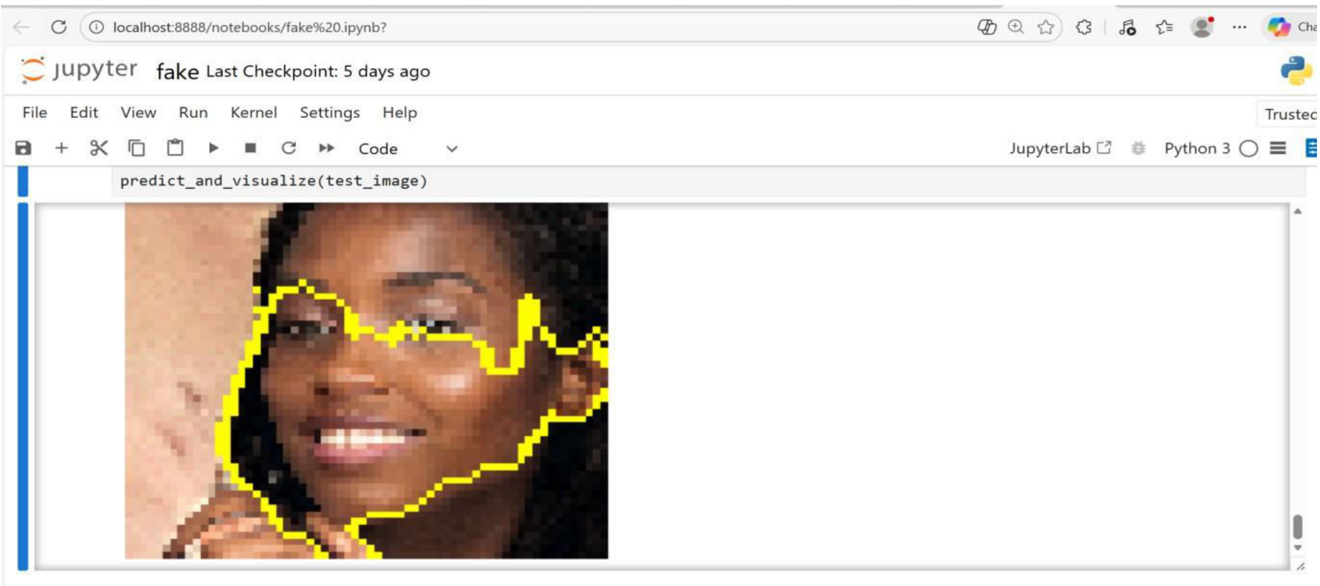

Fig.No:6.5

International Journal of Multidisciplinary and Scientific Emerging ResearcH (IJMSERH)

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

VII. CONCLUSION

This project successfully demonstrates a deep learning–based approach for detecting deepfake images and videos using Convolutional Neural Networks (CNNs) and image processing techniques. By analyzing facial features, texture patterns, and visual inconsistencies, the proposed system is able to accurately distinguish between real and manipulated media.

The use of preprocessing methods such as frame extraction, normalization, and data augmentation further improves detection performance and model robustness. Experimental results indicate that the system achieves high accuracy and efficiency, making it a reliable solution for deepfake detection.

Overall, this project contributes an effective tool for combating misinformation, enhancing digital media authenticity, and addressing the growing challenges posed by deepfake technology.

VIII. FUTURE ENHANCESMENTS

The proposed deep learning-based deepfake detection system can be further enhanced by incorporating advanced models such as GAN-based detection techniques to identify more complex and realistic deepfake content. In the future, the system can be extended to support real-time detection in live video streaming platforms, which will help in preventing the rapid spread of fake content. The accuracy of the model can also be improved by training it with larger and more diverse datasets that include variations in lighting conditions, backgrounds, and facial features. Additionally, the system can be integrated with social media platforms to automatically detect and flag deepfake content before it reaches a wide audience. Further enhancement includes expanding the system to detect audio deepfake, such as voice cloning, along with image and video analysis for a more comprehensive solution. The use of Explainable AI techniques can also be incorporated to improve transparency and help users understand the decision-making process of the model. Moreover, the development of a mobile application can make the system more accessible, allowing users to easily verify the authenticity of media content using their smartphones.

REFERENCES

1. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press, 2016.
2. Yuezun Li and Siwei Lyu, “Exposing DeepFake Videos by Detecting Face Warping Artifacts,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
3. Andreas Rössler et al., “FaceForensics++: Learning to Detect Manipulated Facial Images,” International Conference on Computer Vision (ICCV), 2019.
4. Hany Farid, “Digital Image Forensics,” MIT Press, 2016.
5. Afchar, D., et al., “MesoNet: a Compact Facial Video Forgery Detection Network,” IEEE Workshop on Information Forensics and Security, 2018.
6. Chollet, F., Deep Learning with Python, Manning Publications, 2017.
7. H. Nguyen, F. Fang, J. Yamagishi, and I. Echizen, “Multi-task learning for detecting and segmenting manipulated facial images and videos,” IEEE International Conference on Biometrics (ICB), pp. 1–8, 2019.
8. F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258, 2017.
9. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436 444,2015.
10. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
11. Jerry John and Ms. Bismin V. Sherif, “Comparative Analysis on Different DeepFake Detection Methods and Semi Supervised GAN Architecture for DeepFake Detection” in Proceedings of the Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC-2022). IEEE, (2022).
12. Kalicharan Jalui, Aditya Jagtap, Saloni Sharma, Gilofer Mary, Reba Fernandes and Megha Kolhekar, “Synthetic Content Detection in Deepfake Video using Deep Learning” in 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT). IEEE, (2022).
13. Sonia Salman and Jawwad Ahmed Shamsi, “Comparison of Deepfakes Detection Techniques”, in 2023 3rd International Conference on Artificial Intelligence (ICAI). IEEE, (2023).
14. Dr. R. R. Rajalaxmi, Sudharsana P P, Rithani A M, Preethika S, Dhivakar P and Gothai E, “Deepfake Detection using Inception-ResNet-V2 Network”, in Proceedings of the 7th International Conference on Computing Methodologies and Communication .

| ISSN: 2349-6037 | www.ijmserh.com ||Impact Factor: 9.274||Quarterly, Peer Reviewed & Refereed Journal|

|| Volume 14, Issue 2, April-June 2026 ||

DOI: 10.15662/IJMSERH.2026.1402021

15. Sonia Salman and Jawwad Ahmed Shamsi, “Comparison of Deepfakes Detection Techniques”, in 2023 3rd International Conference on Artificial Intelligence (ICAI). IEEE, (2023).
16. Kalicharan Jalui, Aditya Jagtap, Saloni Sharma, Gilofer Mary, Reba Fernandes and Megha Kolhekar, “Synthetic Content Detection in Deepfake Video using Deep Learning” in 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT). IEEE, (2022).

International Journal of Multidisciplinary and Scientific Emerging Research (IJMSERH)

Impact Factor: 9.274

editor@ijmserh.com

www.ijmserh.com