IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

AI POWERED DEEP FAKE DETECTION SYSTEM

Vishal M[1], Mrs. A. Sathiya Priya[2]

Department of Information Technology, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India[1]

Assistant professor**,** Department of Information Technology, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India[2]

Abstract: The rapid advancement of deep learning technologies has led to the rise of deepfake media, posing serious threats to digital trust and security. This project presents an AI-based Deepfake Detection System that identifies manipulated images using advanced machine learning techniques. The system integrates a React frontend and a Node.js backend for seamless media upload and preprocessing. The processed content is analysis using deep learning models such as Pixel Error level analysis to detect visual and temporal inconsistencies. Based on the extracted features, the system classifies the media as real or deepfake and provides a confidence score along with an explanatory result. The proposed solution aims to enhance digital content verification and strengthen cybersecurity measures against synthetic media threats.

Keywords: Deepfake Detection, Artificial Intelligence, Pixel Error Level Analysis, Image Forensics, Machine Learning, Cybersecurity, Media Authentication

I. INTRODUCTION

People share images widely on social media and digital platforms. With the help of artificial intelligence, it has become easy to create fake or manipulated media called deepfakes. These fake images look real and can be used to spread false information or harm individuals. It is difficult for humans to identify whether a media file is real or fake just by looking at it. This creates a need for a system that can automatically analyse and verify digital content. The proposed project develops an AI-based deepfake detection system that allows users to upload images and check their authenticity. The system is implemented as a web application using modern technologies such as React and Node.js, making it accessible and easy to use. The project aims to raise awareness about deepfakes and provide a technical solution for digital media verification.

II. BACKGROUND AND LITERATURE REVIEW

2.1 Importance of Media Authentication

The rapid advancement of artificial intelligence has enabled the creation of highly realistic synthetic media known as deepfakes, which can manipulate images with minimal technical effort. These technologies have become increasingly accessible, allowing manipulated content to spread quickly across digital platforms. As deepfakes become more convincing, it is becoming difficult for users to distinguish between authentic and altered media. This poses serious challenges to digital trust and information credibility. The misuse of deepfake content can lead to misinformation, identity impersonation, reputational damage, and cybercrime. Hence, there is a growing need for reliable systems that can verify the authenticity of digital media and support safer online environments.

Table 2.1.1 Challenges, Description, Impact

Challenge	Description	Impact
Misinformation	Manipulated media – false information	Loss of public trust
Identity Impersonation	Imitate real individuals without consent	Reputational damage and risk of fraud
Political Manipulation	Influence public opinion or elections.	Threat to democratic processes and social stability.
Cybercrime	Deepfake- scams, phishing, and social engineering attacks.	Financial losses and increased security risks.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

2.2 Challenges in Detecting Deepfake Media

The rapid improvement in deepfake generation techniques has made manipulated images increasingly realistic and difficult to distinguish from authentic content. Traditional manual verification methods are often ineffective, as visual artifacts are subtle and not easily noticeable to the human eye. Additionally, deepfake tools evolve continuously, making detection methods quickly outdated. The large volume of media shared daily across digital platforms further complicates the timely identification of manipulated content. These challenges highlight the need for automated, scalable, and intelligent detection systems capable of adapting to emerging manipulation techniques and ensuring the authenticity of digital media.

Feature	Existing Models	System Capabilities
Detection Approach	ML models and standalone forensic tools	AI-powered analysis integrated into a web application
Scalability	Limited scalability and manual processing	Scalable backend with APIbased AI integration
Analysis Output	Binary result (Real/Fake) with minimal explanation	Verdict with confidence score and detailed explanation
Processing Time	Slower due to offline processing	Faster response using cloudbased AI services
Adaptability	Periodic frequent retraining and manual updates	Easily upgradable via AI service integration

Table 2.2.1 Challenges in Detecting Deepfake Media

2.3 Related Work

Early research focused on identifying visual artifacts and inconsistencies in manipulated images using traditional machine learning and handcrafted features. With the advancement of deep learning, researchers have explored convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer-based models to learn complex spatial and temporal patterns present in deepfake media. Recent works have also investigated frequency-domain analysis and error level analysis to detect subtle manipulation traces. Although these approaches have shown promising results in controlled environments, many existing solutions remain limited to research prototypes and lack userfriendly deployment.

Authors	Year	Algorithm Used	Key Findings
Afchar et al.	2018	CNN (MesoNet)	shallow CNN models can detect visual artifacts in deepfakes.
Rossler et al.	2019	CNN-based classifiers	Introduced large-scale deepfake datasets and benchmarked detection models.
Sabir et al.	2021	CNN + RNN (Temporal modeling)	Highlighted the importance of temporal features in deepfake detection.
Dosovitskiy et al.	2021	Vision Transformer (ViT)	Demonstrated transformer models can capture global visual patterns effectively.
Li and Lyu	2020	CNN with FrequencyAnalysis	Showed frequency-domain patterns can help distinguish real and fake images.

Table 2.3.1 Related Work

III. SYSTEM ARCHITECTURE

The proposed Deepfake Detection System follows a modern full-stack, service-oriented architecture designed for scalability, usability, and efficient AI integration. The system consists of three primary layers: a React-based frontend for user interaction, a Node.js backend for request handling and orchestration, and an AI analysis service powered by Google Gemini AI. Users upload images through the web interface, which are securely transmitted to the backend server. The

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

backend performs validation, preprocessing, and forwards the media to the AI service for forensic analysis. The AI evaluates the content using advanced visual analysis techniques to identify manipulation artifacts and inconsistencies. The processed results, including the authenticity verdict, confidence score, and explanation, are returned to the frontend for display. This layered architecture ensures separation of concerns, improves maintainability, and supports future enhancements such as scaling, logging, and deployment to cloud environments

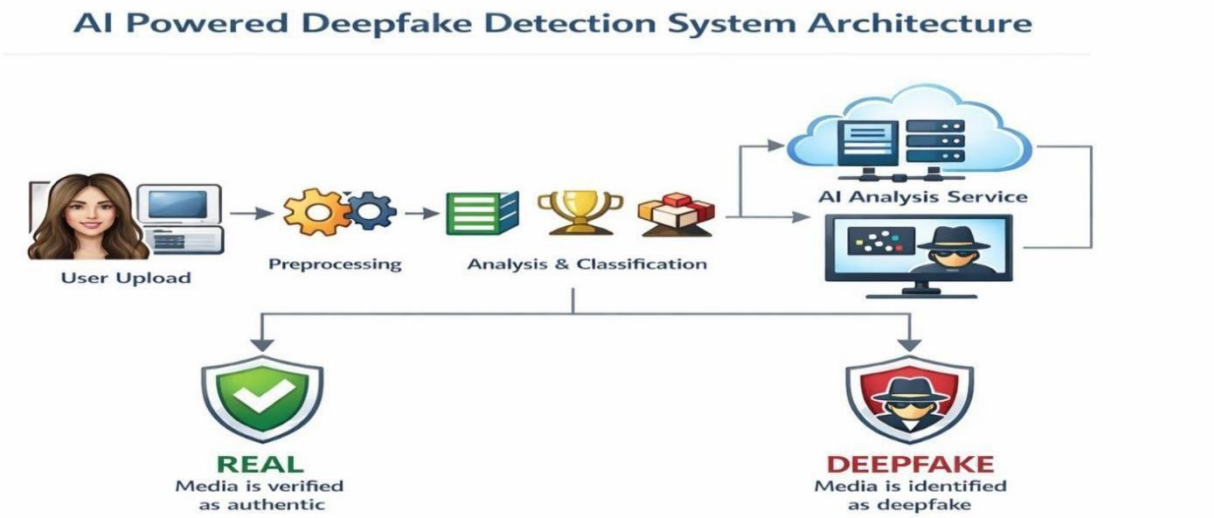


Figure 3.1.1 Deepfake Detection System Architecture

3.2 Modules of the Proposed System

Module	Description
User Interface Module	Provides a web-based interface using React for users to upload images and view analysis results.
Backend Processing Module	Manages API requests, orchestrates communication between frontend and AI service, and handles preprocessing tasks.
AI Analysis Module	Integrates with Google Gemini AI to analyze media for manipulation artifacts and authenticity indicators
Logging & Monitoring Module	Maintains logs of requests and analysis outcomes for monitoring, debugging, and audit purposes.
Security & Access Control Module	Ensures secure handling of user data, validates requests, and protects APIs from misuse

Table 3.2.1 Module

3.3 Key Features

The proposed system offers an intuitive web-based interface for deepfake detection, enabling users to verify the authenticity of images with ease. It integrates AI-powered analysis to provide accurate detection along with confidence scores and explanations. The system is designed to be scalable, secure, and easily extendable for future enhancements.

Features	Benefit
Web-Based User Interface	access the system easily through any browser.
Confidence Score, Explanation	users understand the reliability and reasoning behind results.
Scalable Backend Design	Supports future growth and increased user traffic.
User Friendly Interface	Protects user data and ensures safe processing of uploaded files.

Table 3.3.1 Key Features

3.4 Workflow

The workflow of the proposed system starts with the user uploading an image through the web interface. The frontend securely transmits the media to the backend server, where validation and request handling are performed. The backend

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

forwards the media to the AI analysis service for deepfake detection and forensic evaluation. The analysis results, including the verdict and confidence score, are processed and displayed to the user in a clear and user-friendly format.

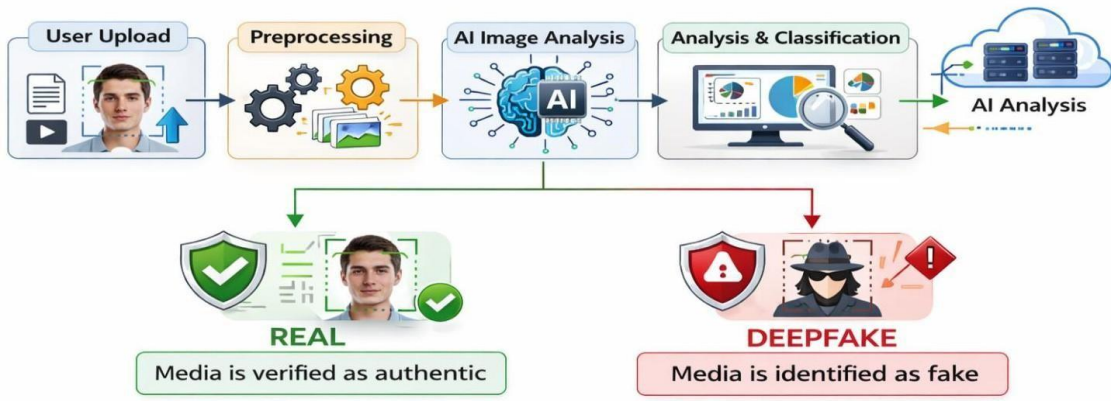


Figure 3.4.1 Workflow

Stage	Activity
User Authentication	User registers and logs into the Veritas
Media Upload	User uploads image through the web interface.
API validation	Frontend sends the media to the backendAPI for validation and handling.
AI Analysis	Backend forwards the media to the AI service for deepfake analysis.
Result evaluate	AI service analyzes the media and returns detection results.
Result process	Backend processes the results and prepares verdict and confidence score.
Dashboard of result	Frontend displays the final results to the user.

Table 3.4.2 Workflow Stages

IV. DATAFLOW DIAGRAM

The data flow in the proposed Deepfake Detection System illustrates how information moves across different system components. The process begins when the user uploads an image through the web interface. The frontend captures the input and forwards it to the backend server for validation and preprocessing. The backend then transmits the media to the AI analysis service for deepfake detection and forensic evaluation. The AI service analyzes the content and generates detection results along with confidence metrics. These results are returned to the backend, processed into a user-friendly format, and finally displayed to the user. This structured data flow ensures secure handling, efficient processing, and clear presentation of analysis outcomes.

Table 4.1 Workflow Stages

Stage	Input	Process	Output
User Input	Image User	User uploads media through the web interface	Media file captured by frontend
Data Preprocessing	Media file	Frontend sends file to backend API	Validated upload request
Model Processing	Media file	Backend Performs validation and preprocessing	Prepared media for analysis
AI Analysis	Prepared media	Media sent to AI analysis service	Verdict, confidence score, explanation
Result Display	AI analysis results	Backend formats response and sends to frontend	Results displayed to the user

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

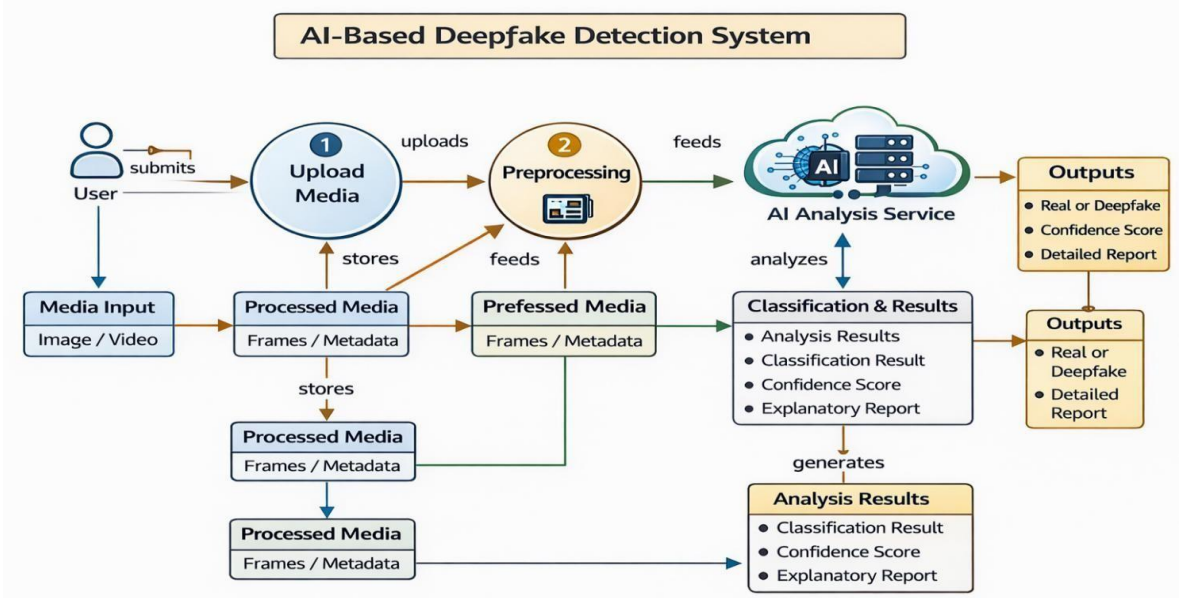


Figure 4.1.1 AI-Based Deepfake Detection System

V. METHODOLOGY

The methodology outlines the systematic approach followed to design, develop, and implement the proposed Deepfake Detection System. It describes the key phases involved, including requirement analysis, system design, development, AI integration, and testing. This structured process ensures reliable implementation, scalability, and effective validation of the system.

5.1 Requirement Analysis and Planning

The initial phase involved identifying the core objectives of the deepfake detection system and understanding user requirements. Functional needs such as media upload, result visualization, and AI-based analysis were clearly defined. Technical requirements including frontend-backend integration, API communication, and secure file handling were documented. This phase helped in establishing a clear development roadmap and selecting appropriate technologies.

5.2 System Design and Architecture

In this phase, the overall system architecture was designed with a separation of concerns between frontend, backend, and AI services. Data flow, module interactions, and communication protocols were planned to ensure scalability and maintainability. Design decisions focused on creating a modular and extensible architecture. This ensured that future enhancements could be integrated with minimal changes to the core system.

5.3 Frontend development

The user interface was developed using React to provide an intuitive and responsive experience. Components were designed for media upload, progress indication, and result visualization. Emphasis was placed on usability, accessibility, and clear presentation of analysis outcomes. This layer enables seamless user interaction with the underlying AI-powered services.

5.4 Backend Development and API Integration

The backend was implemented using Node.js to handle client requests, validate inputs, and orchestrate communication with the AI analysis service. Secure APIs were developed to manage media uploads and responses. Proper error handling, logging, and request management were incorporated to ensure reliability. This layer acts as the central coordinator between the frontend and AI services.

5.5 AI Service Integration and Analysis

The system integrates with an AI service to perform deepfake detection and forensic evaluation of uploaded media. The backend forwards validated inputs to the AI service and receives structured analysis results. These results include authenticity verdicts, confidence scores, and explanatory insights. This approach enables the use of advanced AI capabilities without embedding complex models directly into the application.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

5.6 Testing, Evaluation, and Optimization

Comprehensive testing was conducted to validate system functionality, performance, and reliability. Test cases included different media formats, file sizes, and network conditions. The system was evaluated for accuracy of results, response time, and user experience. Based on observations, performance optimizations and minor refinements were applied to enhance overall system stability.

Category	Recommendation
Data Handling	Enforce strict file validation and size limits to ensure secure uploads.
Security	Implement authentication, rate limiting, and secure API endpoints.
Performance	Use asynchronous processing and efficient API calls to reduce response time.

5.1.1 Methodology of Deepfake detection system

VI. IMPLEMENTATION

The implementation of the proposed Deepfake Detection System was carried out using a modular fullstack approach. The frontend was developed using React to provide a responsive and user-friendly interface for media upload and result visualization. The backend was implemented using Node.js to handle API requests, perform validation, and manage communication with the AI service. Integration with Google Gemini AI enables advanced media analysis and deepfake detection without embedding complex models within the application. Secure handling of user inputs, proper error management, and structured response formatting were incorporated to ensure reliability. The overall implementation supports scalability and future enhancements through a well-defined and maintainable code structure.

VII. BENEFITS AND CHALLENGES

7.1 Benefits

The proposed Deepfake Detection System helps users verify the authenticity of digital media with ease and reliability. It reduces the risk of misinformation, identity misuse, and digital fraud by providing AI-powered analysis. The web-based implementation makes the solution accessible, scalable, and suitable for real-world usage.

7.2 Challenges

Detecting deepfakes remains challenging due to the continuous evolution of manipulation techniques and increasing realism of synthetic media. The accuracy of detection can vary depending on the quality and type of input media. Additionally, reliance on external AI services may introduce limitations related to performance, cost, and availability.

Benefit	Description
Media Authenticity Verification	Helps users identify whether images are real or manipulated.
Misinformation Reduction	Assists in preventing the spread of fake or misleading digital content.
User-Friendly Web Access	Provides easy access through a browser-based interface.
Scalable Architecture	Supports future expansion and integration with other platforms.

7.1.1 Benefits of Proposed System

Challenge	Description
Evolving Deepfake Techniques	New manipulation methods can reduce detection effectiveness over time.
Variability in Media Quality	Low-resolution or compressed media may affect analysis accuracy.
Performance Constraints	Large files and high traffic can increase processing time.
Dependency on AI Services	External AI integration may introduce cost, latency, or availability issues.

7.2.1 Challenges of Proposed System

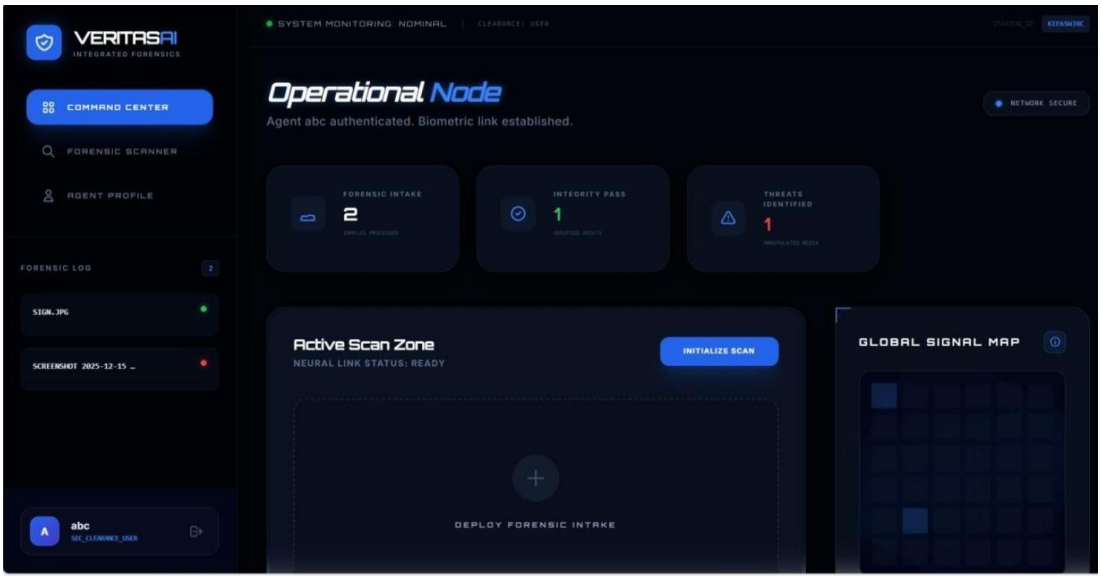
IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

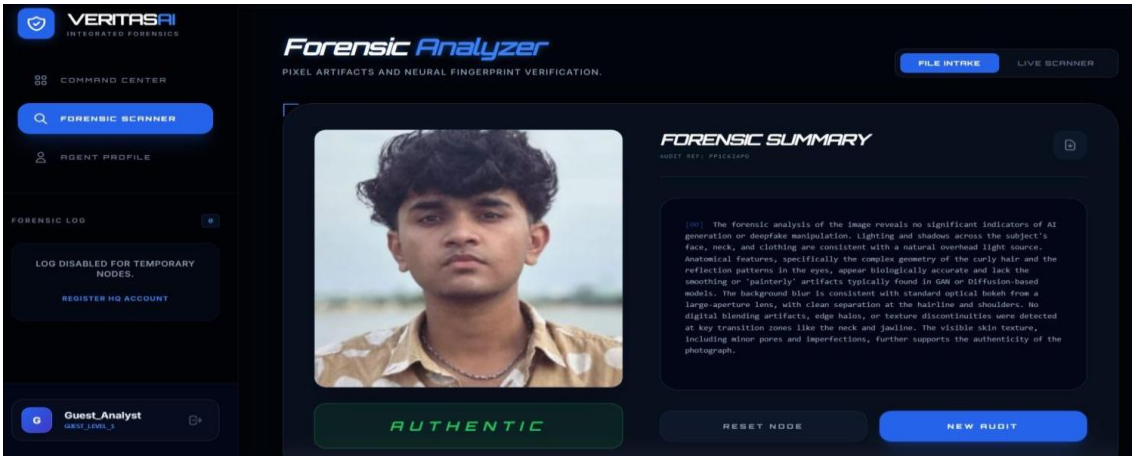
International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

VIII. USER INTERFACE





IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.13302

IX. DISCUSSION AND FUTURE WORK

Discussion:

The proposed Deepfake Detection System demonstrates the effective integration of AI services with modern web technologies to address the challenge of digital media authenticity. The system provides accessible and user-friendly verification of images with meaningful outputs such as verdicts and confidence scores. While the current implementation shows promising results, continuous improvements are required to keep pace with evolving deepfake techniques.

Future Work:

Future enhancements will focus on improving the accuracy and robustness of image-based deepfake detection. Advanced image forensics techniques such as high-frequency analysis, noise pattern detection, and metadata verification can be incorporated to identify subtle manipulation traces. The system can be extended to support high-resolution image analysis and batch image verification for large-scale content screening. Integration with browser extensions and mobile applications can improve accessibility and enable real-time image authenticity checks across digital platforms. Additionally, continuous model training using newly generated deepfake image datasets will help the system adapt to emerging image manipulation techniques and maintain long-term detection reliability.

X. CONCLUSION

The proposed PCOD SmartCare system highlights the potential of Machine Learning in improving women's healthcare through early detection and preventive care. By applying Logistic Regression and Random Forest algorithms, the system analyzes patient symptoms, hormone levels, and medical reports to accurately predict PCOD risk levels. This helps reduce the time and cost involved in traditional diagnosis methods while increasing awareness about the disorder. In addition to prediction, the system provides personalized diet, exercise, and lifestyle recommendations, making it a complete healthcare support platform. Overall, the PCOD SmartCare system offers an intelligent, user-friendly, and cost-effective solution that can assist women in managing PCOD and maintaining a healthier lifestyle.

REFERENCES

[1]. Afchar, D., Nozick, V., Yamagishi, J., and Echizen, I., "MesoNet: A Compact Facial Forgery Detection Network," IEEE International Workshop on Information Forensics and Security (WIFS), 2018, pp. 1–7.
[2]. Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., and Nießner, M., "FaceForensics++: Learning to Detect Manipulated Facial Images," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 1–11.
[3]. Li, Y., and Lyu, S., "Exposing DeepFake by Detecting Face Warping Artifacts," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 46–55.
[4]. Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., and Natarajan, P., "Recurrent Convolutional Strategies for Face Manipulation Detection," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 80–89.
[5]. Verdoliva, L., "Media Forensics and DeepFakes: An Overview," IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 5, 2020, pp. 910–932.
[6]. Mirsky, Y., and Lee, W., "The Creation and Detection of Deepfakes: A Survey," ACM Computing Surveys, vol. 54, no. 1, 2021, pp. 1–41.
[7]. Tolosana, R., Romero-Tapiador, S., Vera-Rodriguez, R., and Fierrez, J., "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," Information Fusion, vol. 64, 2020, pp. 131–148. [8]. Goodfellow, I., Bengio, Y., and Courville, A., *Deep Learning*, MIT Press, 2016.
[8]. Esteva, A., Robicquet, A., Ramsundar, B., et al., "A Guide to Deep Learning in Healthcare," Nature Medicine, vol. 25, no. 1, 2019, pp. 24–29.
[9]. Jiang, F., Jiang, Y., Zhi, H., et al., "Artificial Intelligence in Healthcare: Past, Present and Future," Stroke and Vascular Neurology, vol. 2, no. 4, 2017, pp. 230–243.
[10]. OWASP Foundation, "OWASP Top 10 – Web Application Security Risks," OWASP Publications, 2023.
[11]. Google AI, "Gemini API Documentation," Google Developer Documentation, 2024.

