iJRASET

International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.81720

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

AI Provenance Tracker: A Hybrid Deep Learning and Heuristic Approach for Detecting AI-Generated Images

Jyoti Thakur, Suyash Soni, Kritika Markam, Suresh Hemla, Prof. Vishnu Prasad Verma
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING GOVERNMENT ENGINEERING COLLEGE, BILASPUR

Abstract: Therapidadvancementof generativeartificialintelligencehas significantlyincreasedtheabilityto create highly realistic synthetic images,raisingseriousconcerns regarding digital authenticity, misinformation, and cyber fraud. With the emergence of advanced models such as Generative Adversarial Networks (GANs) and diffusion-based techniques, distinguishing betweenrealandAI- generated images has become increasingly challenging for both humansandtraditionalverification systems.To address this issue,this paper presents an AI Provenance Tracker, an end-to-end hybrid system designed to accurately classify images as real or AI- generated by combining deep learningwithheuristicanalysis. At the core of the proposed system isaResNet50-basedConvolutional Neural Network (CNN) utilizing transferlearning,whichiscapableof extracting complex hierarchical featuressuchasedges,textures,and structuralinconsistenciesfrominput images. The model is trained on a diverse dataset of approximately 40,000 images, consisting of both realandAI-generatedsamples,with appropriate preprocessing techniquesincludingresizing, normalization, and data augmentation.Thenetworkis optimizedusingtheAdamoptimizer and CrossEntropyLoss function, achieving an overall classification accuracy of approximately 90%.

In addition to the deep learning model,aheuristicanalysismoduleis incorporated to enhance detection robustness by examining non- learned features such as noise distribution,compressionartifacts, and metadata irregularities, which are often indicative of synthetic content.Theoutputsfrombothhe CNN model and the heuristic module are combined using an ensemble strategy based on weightedaveraging,improvingthe system's reliability and reducing false predictions.

Thesystemisimplementedusinga scalable architecture, where the backend is developed using FastAPI for efficient request handling and model inference, while the frontend is built using Next.js and Tailwind CSS to provide an interactive and user-friendlyinterface.Thecomplete workflow enables real-time image analysis, making the system suitable for practical deployment in digital content verification scenarios.

Experimental results demonstrate thattheproposedhybridapproach outperforms standalone deep learning models in terms of accuracy, generalization, and robustness. The AI Provenance Tracker offers a practical and scalablesolutiontothegrowing challengeofdetectingAI-generated images, contributing towards enhancing trust and authenticity in digital media ecosystems.

I. INTRODUCTION

Therapidadvancementofartificial intelligence, particularly in generativemodelssuch as GenerativeAdversarialNetworks (GANs) and diffusion-based techniques,hasenabledthecreation of highly realistic synthetic images. While these technologies offer significant benefits in fields like design, entertainment, and virtual content generation, they also raise serious concerns related to misinformation, digital forgery, and cybercrime.AI-generatedimagesare becoming increasingly difficult to distinguish from real ones, making authenticity verification a critical challenge in today's digital ecosystem.

Traditionalimageforensicsmethods relied on detecting visible inconsistencies such as lighting errors, unnatural textures, or facial distortions. However, modern AI models have significantly reduced these artifacts, rendering conventional detection techniques less effective. As a result, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have gained prominence duetotheirabilitytolearncomplex visual patterns and identify subtle differences between real and synthetic images.

Despitetheireffectiveness,deep learning models often face limitations such as poor generalization and sensitivity to unseen data. To address these challenges, this paper proposes an AI Provenance Tracker, a hybrid system that combines a ResNet50-basedCNNwithheuristicanalysis.

Thesystemleveragesbothlearned featuresandrule-basedchecks,such as noise patterns and compression artifacts,toimprovedetection accuracy and robustness.

The proposed solution is designed with a scalable architecture using FastAPIforbackendprocessingand Next.js for frontend interaction, enabling real-time image analysis.

Thisworkaimstoprovidean efficientandpracticalapproachfor detecting AI-generated images, contributing to improved trust and authenticity in digital media.

II. LITERATURE REVIEW

Several researchers have proposed different techniques for detecting deepfakeandAI-generatedcontent. The evolution of these techniques over recent years is summarized as follows:

(2021)

Some studies combined Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks to capturetemporalinformationacross multiple video frames. This approach significantly improved detection accuracy for video-based deepfakesbyanalyzingbothspatial and temporal features.

(2022)

Traditional detection methods proved largely ineffective against highlyrealisticdeepfakesgenerated by modern Generative Adversarial Networks(GANs).Asaresult,deep learning became the primary focus. Detection techniques were broadly categorized into:

- Holisticmethods:analyzingthe entire face
- Feature-basedmethods:focusing on specific regions such as eyes and lips Popular architectures included CNN, VGG-19, and DenseNet.
- Additionally, ensemble and multi-attentionmodelswere introduced to improve generalizationandperformance.

(2023)

Earlierdetectionmethodsreliedon visible inconsistencies such as eye blinking, facial artifacts, and head pose variations. However, advancementsindeepfake generation significantly reduced thesedetectablecues,makingsuch approaches less reliable.

(2023)

Advancedtechniquessuchas transferlearning,ensemblelearning, and multimodal detection systems were proposed to enhance robustnessandgeneralization.These approaches aimed to improve performanceacrossdifferent datasetsandunseendeepfaketypes.

(2024)

Recentresearchprimarilyfocuseson deep learning-based approaches, especially CNN models, for extracting spatial features from imagesandvideoframes.These models analyze complex visual patternstoclassifycontentasrealor manipulated with improved accuracy.

(2025)

Early deepfake detection methods based on physical cues like eye blinking and facial movement have becomeineffectiveagainstadvanced AI-generated content. Current research emphasizes the use of advancedCNNarchitecturessuchas XceptionNet and EfficientNet for spatial feature extraction. For video analysis, RNN and LSTM models are used to detect temporal inconsistencies across frames. Additionally,researchisfocusingon addressing challenges such as adversarialattacksandimproving computationalefficiencyforreal- time detection systems

III. AI PROVENANCE: KEY INNOVATIONS

1) *Innovation1:Hybrid Detection Model*

Unlike existing systems that rely onlyondeeplearning,ourapproach combines ResNet50 CNN + heuristicanalysistoimprove detectionaccuracyandrobustness.

2) *Innovation2:DiverseDataset Training*

Themodelistrainedonalarge dataset (~40,000 images) including multiple sources such as CIFAKE andKaggledatasets,ensuringbetter generalization across different AI- generated images.

3) *Innovation3:Heuristic FeatureAnalysis*

Thesystemanalyzesadditional features like noise patterns, compressionartifacts,andmetadata inconsistencies, which are often ignoredbytraditionalCNNmodels.

4) Innovation4:Ensemble Decision System

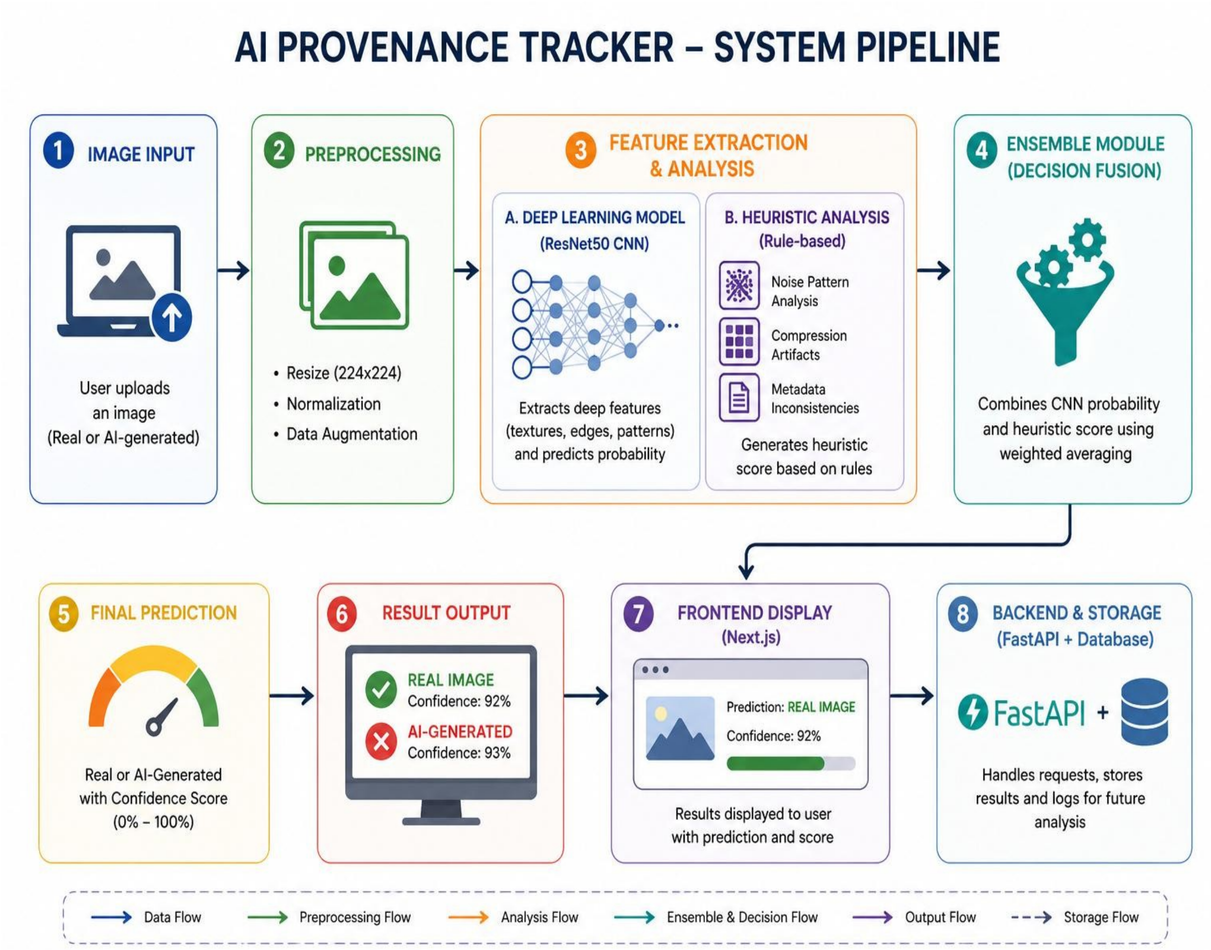
Insteadofrelyingonasinglemodel, outputs from deep learning and Heuristicmodulesarecombined usingaweightedensemblemethod for more reliable predictions.

5) Innovation5:Real-TimeWeb- Based Detection

Thesystemsupportsreal-time detectionthroughaNext.jsfrontend and FastAPI backend, allowing users to upload images and get instant results.

6) Innovation6:Scalableand DeployableArchitecture

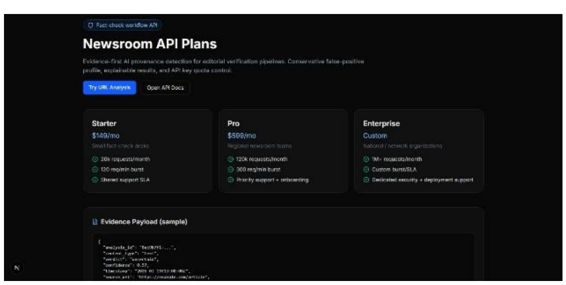
Thearchitectureisdesignedtobe lightweight, scalable, and easily deployable for real-world applicationssuchassocialmedia verificationanddigitalforensics.



Photos:

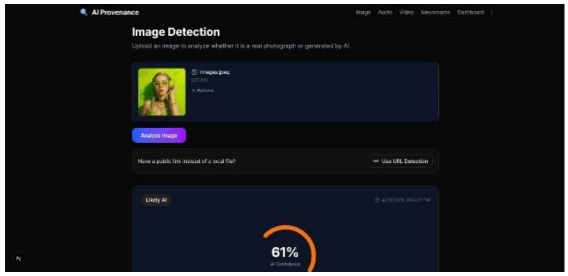
UserInterface:

- NewsRoomAPIPlans

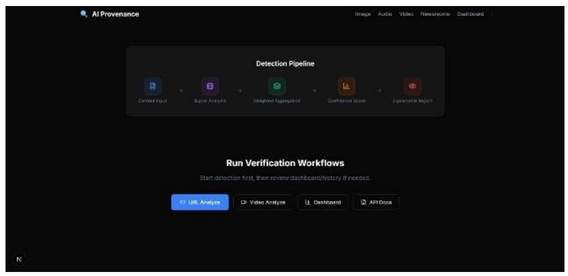


ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

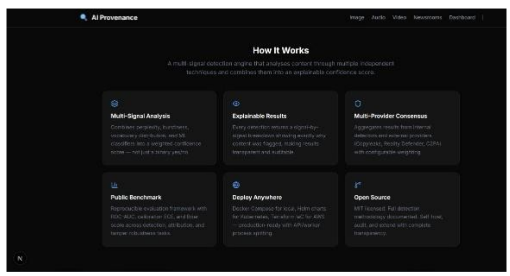
- ImageDetection



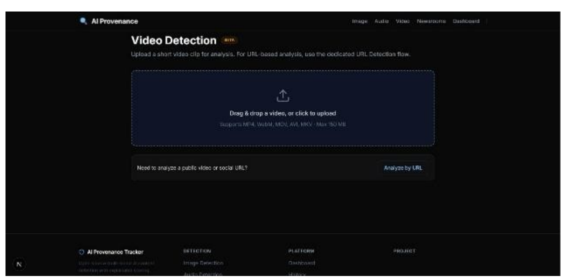
- DetectionPipeline



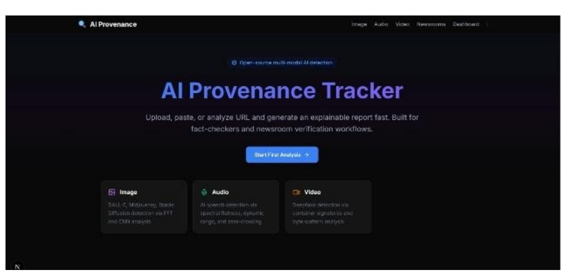
- Working



- FileUploader



- Dashboard



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

- Accuracy

```python
model.eval()
correct = 0
total = 0

with torch.no_grad():
    for images, labels in val_loader:
        images = images.to(device)
        labels = labels.to(device)

        outputs = model(images)
        _, predicted = outputs.max(1)

        total += labels.size(0)
        correct += (predicted == labels).sum().item()

print("Accuracy:", correct / total)
```

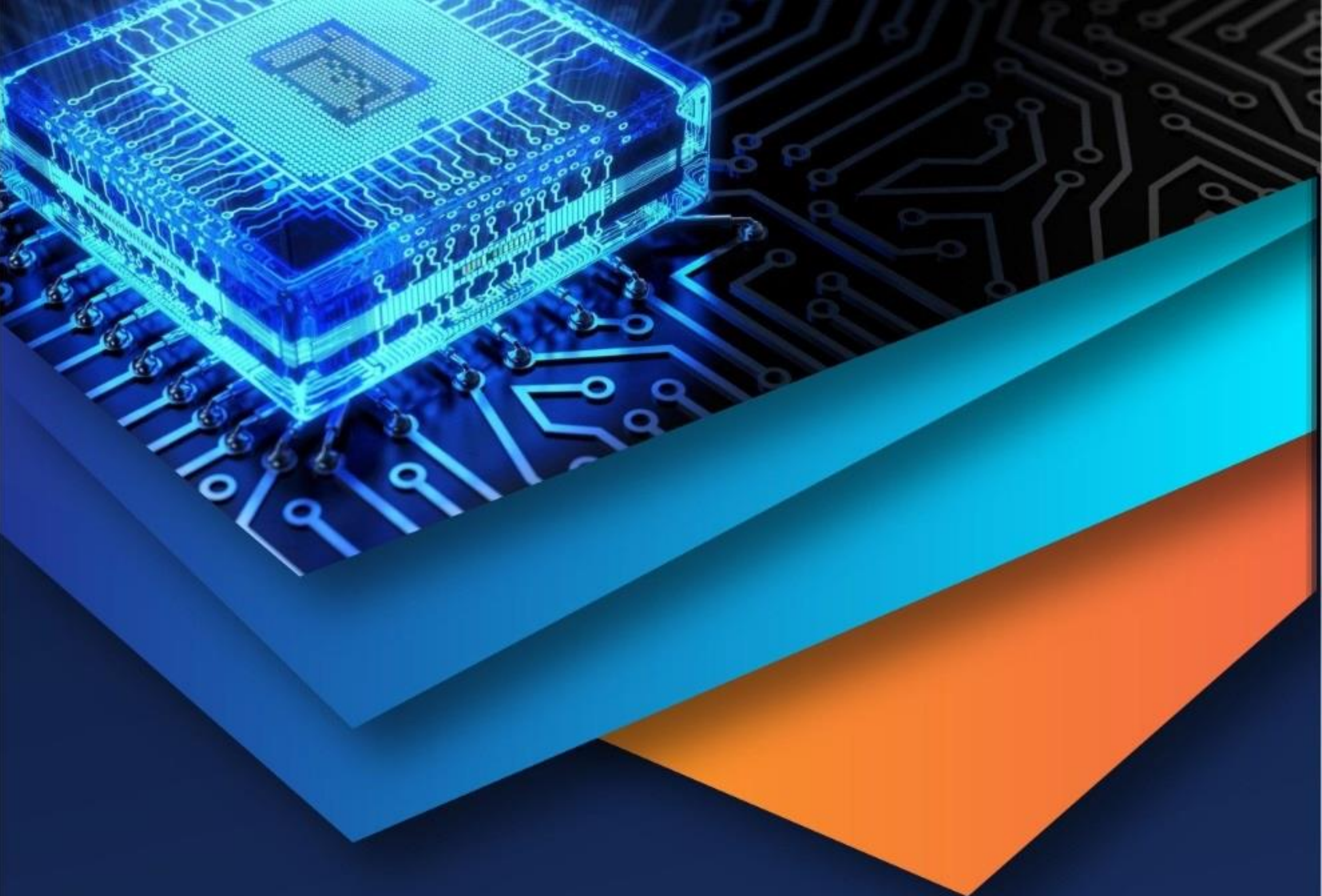
```
Accuracy: 0.9
```

IV. CONCLUSION

This paper presented the AI ProvenanceTracker,ahybridsystem designed to detect whether an image is real or AI-generated. By combining a ResNet50-based deep learning model with heuristic analysis, the proposed approach effectively captures both learned visual patterns and rule-based inconsistencies such as noise and compression artifacts.

The system achieved an accuracy of approximately 90%, demonstrating improved performance compared to standalone methods. The use of an ensemble mechanism further enhanced reliability and reduced false predictions. In addition, the scalable architecture built using FastAPI and Next.js enables real- time detection, making the system suitable for practical applications.

Overall, the proposed solution provides a robust, efficient, and deployable approach for AI- generated image detection. It contributes toward improving trust, authenticity, and verification in digitalmedia,addressingthegrowing challenges posed by advanced generative technologies.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY