International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

Ai Image Fraud Detector

Shreya Shashikant Patil[1], Shital Nivrutti Sutar[2], Prachi Prasad Patil[3], Mrs . Meghana Khare[4]
1-3(CSE (Artificial Intelligence and Data Science), PVPIT College of Budhgaon)

Abstract- Artificial intelligence has made it possible to generate highly realistic images, which can be mis used for misinformation, fraud and identity theft. Detecting such AI- generated images manually is difficult and time consuming. Detecting such AI-generated images has become very important to maintain the authenticity of digital content. This paper presents an AI Image Fraud Detector such that uses deep learning techniques to classify as real or fake. The system integrates YOLO (You Only Look Once) model with a web-based applications developed using Flask and JavaScript. Users can upload images through a user-friendly interface, and the system provides prediction result along with confidence scores. The model processes images in real time and ensures fast detection. Experimental results show that the system performs efficiently with good accuracy depending on the dataset quality. This research contributes to improving digital security by providing an automated solution for detecting AI-generated images. In this research, we developed an AI image fraud detection system using deep learning models such as VGG16, ResNet, and InceptionV3.Thesemodels are trained on a dataset containing both real and AI generated images. The system compares the performance of all three model to find which one give better accuracy. The model is trained on a dataset from Kaggle that contain both real and fake images of Aadhar- id photo and other documents. Image preprocessing techniques are used to improve performance of the model. The result show that deep learning models can effectively detect fake images, with one model performing better based on accuracy and efficiency. The study highlights that using multiple models improve reliability and provides a strong solution for detecting AI-generated images in real world applications. We also tested different settings of the model to understand what works best. Our study shows that it is a strong and reliable method for detecting AI-generated images and can be useful in real-world applications. Model is addressing the increasing challenge of AI-generated image detection, laying a foundation for future research in critical area.

Keywords- paper survey, Explainable artificial intelligence, Feature extraction techniques, input layer, output layer, feature maps, Decision model, Transfer learning, Digital media Detection techniques, Trust in media, Convolutional neural network, GAN based image detection, vision transformer, Image classification, Fake AI-generated image detection.

I. INTRODUCTION

Artificial Intelligence has significantly transformed image generation techniques, particularly with the emergence of Generative Adversarial Networks (GANs). While such advancements offer benefits in fields like media, design, and entertainment, they also pose serious risks including misinformation, identity fraud, and deepfake content.

Manual detection of such images is inefficient and unreliable. Therefore, there is a need for an automated system that can quickly and accurately detects AI- generated images. This research focuses on developing a deep learning based web application that detects image authenticity using the YOLO model. The system provides real-time predictions and enhances trust in digital content.. Traditional approaches based on basic image processing techniques are no longer sufficient to handle the complexity of modern synthetic images. Deep learning models, especially Convolutional Neural Networks (CNNs), have shown significant potential in addressing this problem. In this work, we propose a web-based AI Image Fraud Detection system using three powerful CNN architectures: VGG16, ResNet, and InceptionV3. These models are capable of extracting complex features and identifying subtle differences between real and AI-generated images. The system provides real-time predictions and is designed for practical deployment.

II. PROBLEM STATEMENT:

The rapid advancement of AI has enabled the creation of highly realistic fake images, leading to issues such as misinformation,

© 2026 IJSRET

deepfakes, and digital fraud Detecting these images manually is difficult and unreliable. Existing methods lack accuracy and real-time capability. Therefore, there is a need for an automated, efficient system using deep learning models to accurately identify AI generated images.

III. RELATED WORK

Several researchers have worked on detecting fake or AI-generated images using different techniques. Early methods relied on traditional image processing techniques, where handcrafted features such as texture, color, and noise patterns where used. However, these methods were not accurate and failed to detect complex AI-generated images.

1. CNN-based approach:
With the advancement of deep learning, Convolutional Neural Network (CNNs) became widely used for image classification tasks. These models automatically learn features from images and improve detection accuracy.
However, single CNN models may struggle to generalize across different types of AI-generated images.

2. Advanced Deep learning Models:
Modern approaches use powerful deep learning architectures such as VGG16, ResNet, and InceptionV3. These models are capable of capturing complex patterns and fine details in images, making them more effective in detecting AI-generated content. Transfer learning techniques further enhance performance by using pretrained models.

PROCESS:
Training the model with examples of collecting training data- we start by using the dataset from kaggle to capture a series of images showing real and fake images.

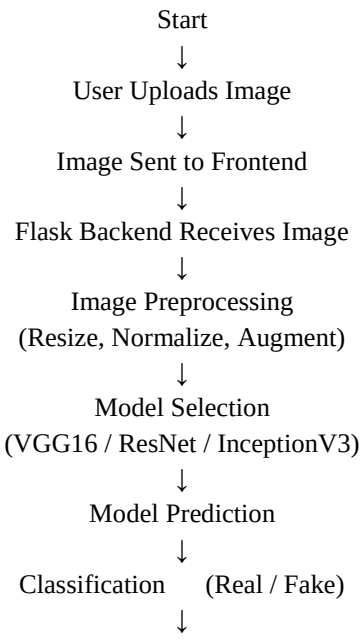
a) **Training Process –** The system is trained using Kaggle dataset that contain images, documents and files. Each image is labelled accordingly. The models learn features such as texture inconsistencies, pixel patterns, and unnatural artifacts

Example: If the model is trained with multiple real and fake face images, it learns to distinguish subtle differences between them.

b) Image Preprocessing –
Before training, images are resized and normalized. Data augmentation techniques such as rotation, flipping, and zooming are applied to improve model performance and accuracy.

c) Real-Time Prediction -
1. User Upload Image
The user uploads an image through the web interface.
2. Model Prediction-
The image is sent to the backend where the trained model analyses it and predicts whether it is real and predicts whether it is real or fake.
3. Continuous Detection-
The system can process multiple images continuously and provide instant results.
4. Display Result -
The final output is displayed with a confidence score indicating the probability of the image being real or AI- generated.
5. Display Result-
The final output is displayed with a confidence score indicating the probability of image being real or AI-generated.

Flowchart of this Process:
The process begins when the user uploads an image through the web interface. The image is sent to the Flask backend, where preprocessing is performed. The processed image is then passed to deep learning models (VGG16, ResNet, or InceptionV3) for prediction. The model classifies the image as real or fake and generates a confidence score. Finally, the result is displayed on the user dashboard.



© 2026 IJSRET

Confidence Score Generated
↓
Result on Dashboard
↓
End

IV. METHODOLOGY

The proposed system is developed using deep learning and web technologies to detect AI- generated images.

a. Dataset Collection
A balanced dataset consisting of real and ai generated images is used for training and evaluation. The dataset includes diverse image categories to improve generalization.

b. Data Preprocessing
1. Image resizing to 224*224 pixels
2. Normalization of pixel values
3. Data augmentation (rotation, flipping, zooming)

c. Mode selection The following deep learning models are used
1. VGG16: Known for its deep architecture and strong feature extraction
2. ResNet: uses residual connections to solve vanishing gradient problems
3. Inception V3: Efficient in capturing multi scale features

d. Training Process
1. Transfer learning is applied using pretrained weights
2. Fine-tuning is performed on the dataset
3. Loss function: Categorical cross entropy
4. Optimizer: Adam

e. System Integration The trained models are integrated into a web application
1. Frontend: HTML, CSS, JavaScript
2. Backend: Flask API
3. Deployment: Local/cloud server
4. Model: YOLO (Ultralytics) 5. Libraries: OpenCV, NumPy

System Workflow
1. User uploads an image through the web interface
2. The image is sent to the flask backend
3. Image preprocessing is performed
4. YOLO model analyzes the image
5. Prediction is generated (Real/Fake) Result is displayed with confidence score

V. SYSTEM REQUIREMENTS

To run the AI-image fraud detector system efficiently, here are some basic system requirements:

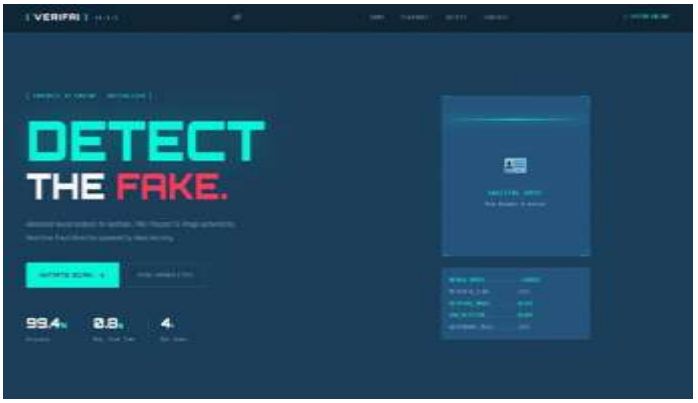
Hardware Requirement –
- Camera – For future real time detection
- Processor – Minimum intel i5 or equivalent (i7or above recommended for faster processing).
- Ram – At least 8 GB ram is required.
- Storage – Minimum 50 GB storage space for storing datasets and model files.
- GPU (Optional): NVIDIA GPU for faster model training and prediction

Software Requirements:
- Operating System – Windows 10 or above version
- Python – Programming language is used for coding the system.
- Framework: Flask (for backend development)
- Libraries –

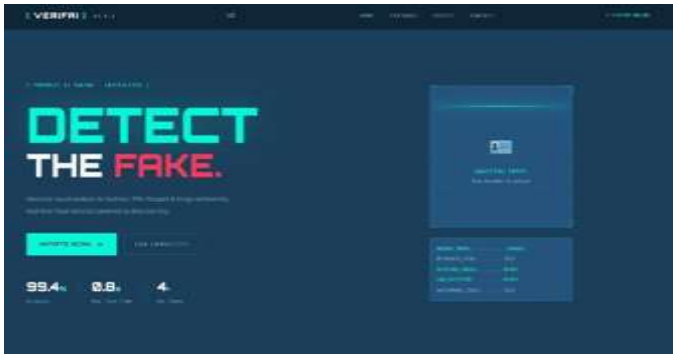
1. OpenCV – for image processing
2. NumPy – for numerical operations
3. TensorFlow / Keras – for deep learning models
4. Matplotlib – for visualization

Fronted Technologies – HTML CSS JavaScript

VI. RESULT



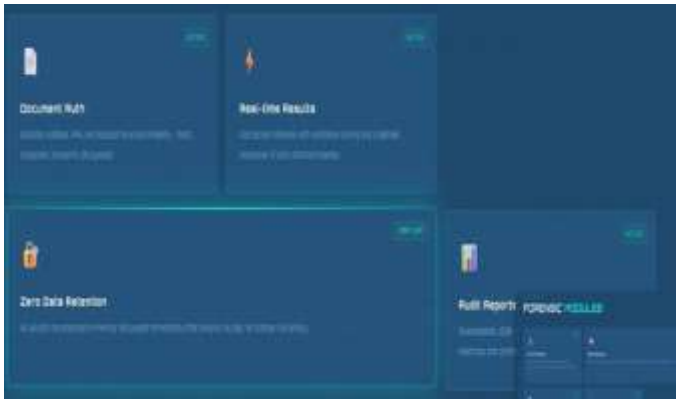
© 2026 IJSRET


DETECT
THE FAKE.


INITIATE FORENSIC SCAN


FORENSIC MODULES


Document Appears Genuine




INITIATE FORENSIC SCAN


UPLOAD & ANALYZE

© 2026 IJSRET

VII. MODELLING AND ANALYSIS

The system architecture consists of three main components:

1. Frontend interface: Allows user to upload images and view results
2. Backend server: Handles API requests and processes images 3. Deep Learning Model: Performs classification

The YOLO model processes the image in a single pass, making it efficient for real-time detection. The performance of the model depends on training data quality and preprocessing techniques.

A. Performance Comparison

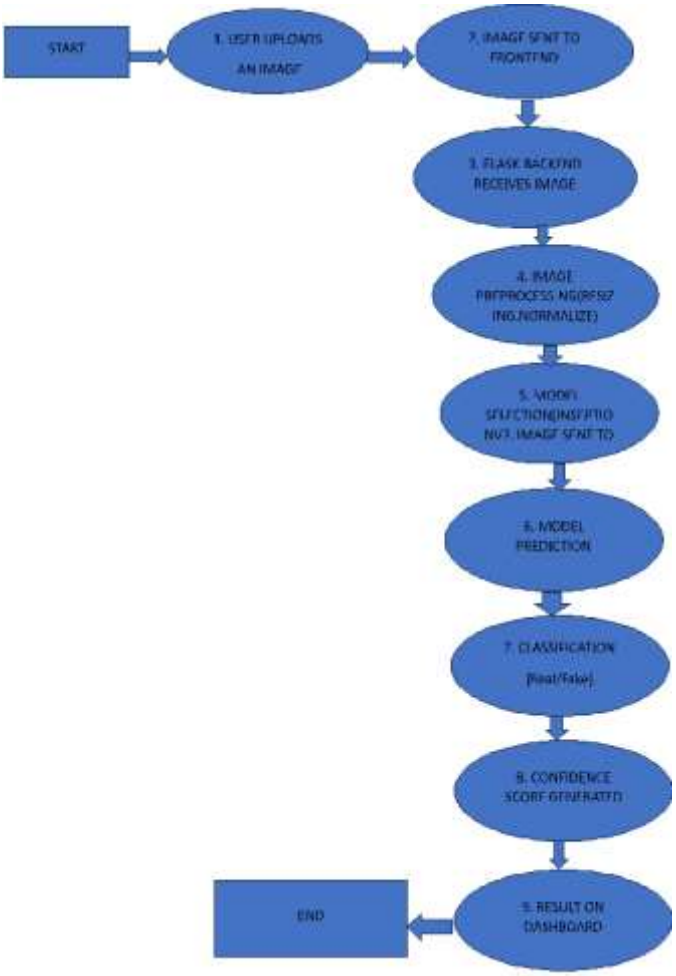
Table7.1 PERFORMANCE COMPARISON

MODEL	ACCURACY	PRECISION	RECALL	SPEED
VGG16	88%	87%	86%	Medium
ResNet	92%	91%	90%	Fast
InceptionV3	94%	93%	92%	Fast

B. Analysis

1. InceptionV3 achieves highest accuracy due to better feature extraction.
2. ResNet provides stable and fast performance.
3. VGG16 performs well but is computationally heavier.

VIII. UML DIAGRAMS



1. Use Case Diagram:

IX. ADVANTAGES

The proposed AI Image Fraud Detection system offers several important advantages:

- High Accuracy: Deep learning models such as VGG16, ResNet, and InceptionV3 can effectively detect subtle differences between real and AI-generated images.
- Real-Time Detection: The system provides fast predictions, making it suitable for real-time applications.

© 2026 IJSRET

- User-Friendly Interface: The web-based platform allows users to easily upload images and view results without technical knowledge.
- Automation: Eliminates the need for manual inspection, reducing time and human effort.
- Scalability: The system can be extended with larger datasets and additional models to improve performance.
- Versatility: Can be applied in various fields such as cybersecurity, social media monitoring, and digital forensics.
- Integration Capability: Can be integrated with other platforms like web applications and security systems for broader usage.

X. CHALLENGES AND LIMITATIONS

Developing an AI-based image fraud detection system involves several challenges. One major challenge is the rapid advancement of AI-generated image techniques, which makes fake images increasingly realistic and difficult to detect. Models must continuously adapt to new types of synthetic data.

Another challenge is the availability of high-quality and diverse datasets. Limited or biased data can reduce the accuracy and generalization ability of the models. Variations in lighting, resolution, compression, and image quality can also affect performance.

The system has certain limitations as well. It may not accurately detect all types of advanced deepfakes or newly generated images that were not part of the training dataset. The performance depends heavily on the quality of training data and model tuning. Additionally, deep learning models require high computational resources, which can affect real-time performance on low-end devices.
Despite these challenges, the system provides a strong foundation for detecting AI-generated images, with scope for further improvements.

FUTURE WORK

The proposed AI Image Fraud Detection system can be further improved in several ways. First, the model can be trained on a larger and more diverse dataset to improve accuracy and generalization across different types of AI-generated images. Advanced deep learning techniques and newer architectures can also be explored to enhance detection performance.

In future, the system can be extended to support real-time video and deepfake detection, which is increasingly important in digital media. A mobile application can be developed to make the system more accessible to users. Additionally, integrating the system with cybersecurity platforms and social media tools can help in automatically detecting and preventing the spread of fake content.

Further improvements can include optimizing the system for faster performance on low-end devices and incorporating explainable AI techniques to provide better understanding of predictions.

XII. CONCLUSION

This paper presents an AI-based Image Fraud Detection system using deep learning models to identify whether an image is real or AI-generated. The system integrates multiple CNN architectures, including VGG16, ResNet, and InceptionV3, to improve detection accuracy and reliability. The results show that deep learning models are effective in capturing complex patterns and distinguishing fake images from real ones.

Among the models used, InceptionV3 achieved the highest accuracy, making it the most suitable for real-time deployment. The integration of the model into a web-based application ensures fast and user-friendly image analysis.

Although the system performs well, its accuracy depends on the quality and diversity of the dataset. Future improvements can focus on using larger datasets, enhancing model performance, and extending the system for real-time video and deepfake detection.

ACKNOWLEDGEMENTS

We sincerely thank our mentors, professors, and institutions for their guidance and support throughout this research. We would like to express our sincere gratitude to our project guide and faculty members for their continuous guidance, valuable suggestions, and support throughout the development of this research work. Their expertise and encouragement helped us successfully complete this project. We are also thankful to our institution, PVPIT College of Budhgaon, for providing the necessary resources and environment to carry out this work. We extend our appreciation to our friends and peers for their support and cooperation during the project. Finally, we would

© 2026 IJSRET

like to thank our family members for their constant motivation and encouragement, which helped us complete this work successfully.

This research is dedicated to improving communication accessibility for individuals with hearing and speech disabilities, and we hope it contributes to a more inclusive society.

REFERENCES

1. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ‘‘You only look once: Unified, real-time object detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788.Ganesh Kumar and P. Vasanth Sena, “Novel Artificial Neural Networks and Logistic Approach for Detecting Credit Card Deceit,” International Journal of Computer Science and Network Security, Vol. 15, issue 9, Sep. 2015, pp. 222-234
2. W. Wang, X. Dong, C. Gan, and S. Kambhampati, ‘‘Image transformers for deepfake detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 1–7.
3. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, ‘‘An image is worth 16×16 words: Transformers for image recognition at scale,’’ 2020, arXiv:2010.11929.
4. S. McCloskey and M. Albright, ‘‘Detecting GAN-generated imagery using color cues,’’ 2018, arXiv:1812.08247
5. J. Zhang, W. Xu, J. Liu, and L. Song, ‘‘Detecting deepfake videos with CNN-based model,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops, Jul. 2019, pp. 1–9
6. H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jegou, ‘‘Training data-efficient image transformers & distillation through attention,’’ in Proc. Int. Conf. Mach. Learn. (ICML)., 2021, pp. 10347– 10357

.

© 2026 IJSRET