Journal of Science, Computing and Engineering Research (JSCER)
Volume-9, Issue- 4, April 2026.
DOI: *https://doi.org/10.46379/jscer.2023.060503*
ISSN: 2708-1079

FAKE IMAGE DETECTION USING DEEP LEARNING

[1]Mohit Sharma,[2]Yash Yadav,[3]Arpit Mukheeja,[4]Vipul Meena,[5]Naresh Kumar

[1]Professor, Department of AI&DS, Modern Institute of Technology and Research Centre, Rajasthan, India.

[2,3,4,5,6]UG Student, Department of AI&DS, Modern Institute of Technology and Research Centre, Rajasthan, India

Article Information

Received : Mar 30 2026

Revised : Mar 30 2026

Accepted : Mar 31 2026

Published : April 02, 2026

Corresponding Author:

Yash Yadav

Abstract— The rapid growth of digital media and social networking platforms has led to an increased spread of manipulated and fake images across the internet. With the advancement of image editing tools and AI-based generation techniques such as deepfakes, it has become extremely difficult to distinguish between real and fake images. These manipulated images can spread misinformation, create security threats, and negatively impact public trust in digital content. Traditional methods for detecting fake images are often manual, time-consuming, and ineffective against advanced manipulation techniques .This project presents a Fake Image Detection System using Deep Learning, designed to automatically identify whether an image is real or fake with high accuracy. The system utilizes deep learning models, particularly Convolutional Neural Networks (CNNs), to analyze complex image features such as textures, pixel distributions, and hidden patterns. These models are capable of learning subtle differences between authentic and manipulated images that are not easily detectable by human observation.The proposed system processes input images through multiple stages including preprocessing, feature extraction, model training, and classification. A labeled dataset containing real and fake images is used to train the model, enabling it to generalize and perform well on unseen data. The system provides quick and reliable predictions, reducing the need for manual verification.Overall, the project offers an efficient, scalable, and practical solution for detecting fake images. It demonstrates the potential of deep learning in digital image forensics and contributes toward improving the authenticity and reliability of visual content in modern digital environments.

Keywords: Fake Image Detection, Deep Learning, CNN, Image Classification, Digital Forensics, Image Manipulation

Citation: Mohit Sharma, Yash Yadav, Arpit Mukheeja, Vipul Meena, Naresh Kumar, "FAKE IMAGE DETECTION USING DEEP LEARNING", Journal of Science, Computing and Engineering Research, 9(04), April 2026.

I. INTRODUCTION

The rapid growth of digital technology and social media platforms has significantly increased the sharing and consumption of images across the internet. Images are widely used in communication, news, entertainment, and social interaction. However, with the advancement of image editing tools and Artificial Intelligence techniques, creating fake or manipulated images has become easier and more sophisticated. Technologies such as deepfakes and Generative Adversarial Networks (GANs) can produce highly realistic images, making it difficult to distinguish between authentic and fake content.

The widespread use of manipulated images poses serious challenges, including the spread of misinformation, fake news, identity misuse, and threats to digital security. These fake images can influence public opinion, damage reputations, and reduce trust in online content. Detecting such images manually is difficult, time-consuming, and often unreliable, especially when the changes are subtle and not easily visible.

Traditional image detection methods rely on basic algorithms or manual verification, which are not effective against modern manipulation techniques. These methods lack accuracy and are not suitable for handling large volumes of data. Therefore, there is a need for an automated and intelligent system that can accurately detect fake images.

With recent advancements in Artificial Intelligence, particularly Deep Learning, more powerful solutions have emerged. Deep learning models, especially Convolutional Neural Networks (CNNs), can analyze complex image features such as textures, edges, and noise patterns to identify hidden inconsistencies. These models are capable of learning from large datasets and can achieve high accuracy in classification tasks.

This project, Fake Image Detection using Deep Learning, aims to develop a system that can automatically classify images as real or fake. The system uses deep learning techniques to analyze image data and detect manipulation effectively. It provides a reliable, efficient,

and scalable solution for improving the authenticity and trustworthiness of digital content.

II. PROBLEM STATEMENT

The rapid growth of digital media and image-sharing platforms has led to a significant increase in the circulation of fake and manipulated images. With the advancement of image editing tools and AI-based technologies such as deepfakes and GANs, creating highly realistic fake images has become easier and more accessible. These images are often used to spread misinformation, manipulate public opinion, create fake news, and pose serious threats to cybersecurity and digital trust.

Detecting such manipulated images manually is extremely difficult, time-consuming, and unreliable. Human observation is not capable of identifying subtle modifications in images, especially when advanced techniques are used. Traditional image verification methods rely on basic algorithms and metadata analysis, which are often ineffective against modern image manipulation techniques.

Existing detection systems also face challenges in maintaining high accuracy across different types of manipulations. Many approaches fail to generalize well on new or unseen data, leading to incorrect predictions. Additionally, handling large volumes of image data in real time remains a major challenge, as manual and semi-automated systems are not scalable.

Another key issue is the lack of robust and intelligent systems that can analyze complex patterns, textures, and inconsistencies within images. Without advanced analysis, it becomes difficult to identify hidden traces of manipulation such as pixel-level changes, noise inconsistencies, and unnatural patterns.

Therefore, there is a strong need for an automated, accurate, and scalable solution that can effectively detect fake images. A deep learning-based approach can address these challenges by learning complex features and patterns directly from data, enabling reliable classification of images as real or fake. Such a system can significantly improve digital security, reduce misinformation, and enhance trust in online content.

III. PROPOSED METHOD

The proposed system is designed to automatically detect fake images using deep learning techniques. It processes input images, extracts meaningful features, and classifies them as real or fake with high accuracy.

A. Image Input and Preprocessing

The process begins with users providing input images. The system reads the images using OpenCV and applies preprocessing techniques such as resizing, normalization, and noise reduction. This ensures consistency in image size and improves model performance.

B. Dataset Collection and Preparation

A dataset containing both real and fake images is collected from reliable sources. The images are labeled accordingly and divided into training, validation, and testing sets. Proper dataset preparation helps the model learn effectively and generalize well on unseen data.

C. Feature Extraction

The system uses Convolutional Neural Networks (CNNs) to extract important features from images. These features include textures, edges, pixel distribution, and noise patterns, which help in identifying hidden inconsistencies in fake images.

D. Model Training

The CNN model is trained on the prepared dataset. During training, the model learns to distinguish between real and fake images by adjusting weights through backpropagation and optimization techniques.

E. Image Classification

After training, the model is used to classify new input images. It predicts whether an image is real or fake based on learned features and provides a confidence score for the prediction.

F. Model Evaluation

The performance of the model is evaluated using metrics such as accuracy, precision, recall, and F1-score. These metrics help in assessing the reliability and effectiveness of the system.

G. System Integration and Implementation

The system integrates all components including preprocessing, feature extraction, model training, and prediction into a single pipeline. It is implemented using Python and deep learning frameworks to ensure efficient and scalable performance.

H. Result Visualization and Output

The final output is displayed through a user interface, showing whether the image is real or fake along with prediction confidence. Visualization tools help users understand the results clearly.

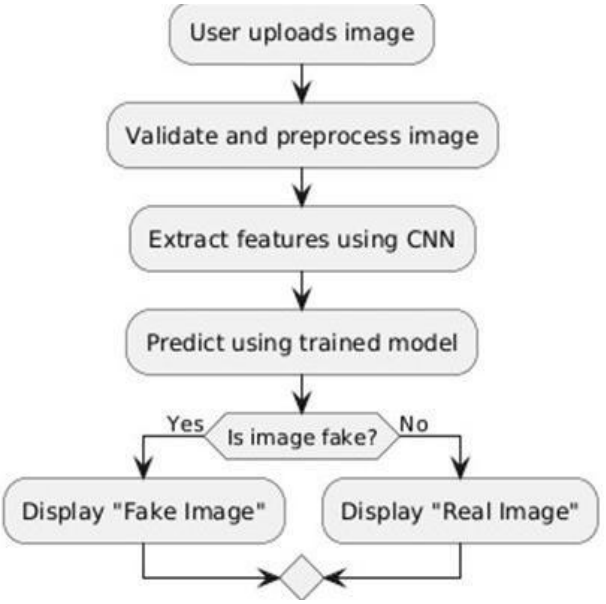

Fig.1.Proposed work Model

IV. TECH STACK

A. Frontend Technologies

The frontend of the system can be developed using Flask, which provides a simple and interactive graphical user interface. It allows users to upload images and instantly view results such as whether the image is real or fake along with prediction confidence. The interface enhances usability and makes the system accessible even for non-technical users.

B. Backend Technologies

The backend is implemented using Python**,** which manages the core logic of the system. It handles image preprocessing, model execution, prediction generation, and communication between different modules. Python's simplicity and strong ecosystem make it ideal for machine learning and image processing tasks.

C. Artificial Intelligence and Deep Learning

The system uses deep learning frameworks such as TensorFlow/Keras or PyTorch to build and train models like Convolutional Neural Networks (CNNs)**.** These models automatically learn complex image features such as textures, edges, and hidden patterns, enabling accurate classification of images as real or fake.

D. Image Processing Libraries

OpenCV is used for image loading, resizing, filtering, and preprocessing operations. NumPy is used for efficient numerical computations and handling image data as arrays. These libraries ensure fast and optimized image processing.

E. Dataset and Training Tools

A labeled dataset containing real and fake images is used for training the model. The dataset is preprocessed, normalized, and split into training and testing sets. Tools like Scikit-learn can be used for data splitting and evaluation support.

F. Model Evaluation and Performance Metrics

The system evaluates performance using metrics such as accuracy, precision, recall, and F1-score. These metrics help in measuring how well the model performs in detecting fake images and ensure reliability of predictions.

G. Visualization and Analysis

Matplotlib is used for visualizing images, training graphs (loss and accuracy), and prediction results. This helps in understanding model performance and improving the system through analysis.

H. Development Tools and Environment

The project is developed using Visual Studio Code for coding and experimentation.The modular design allows easy updates, scalability, and deployment in real-world applications.

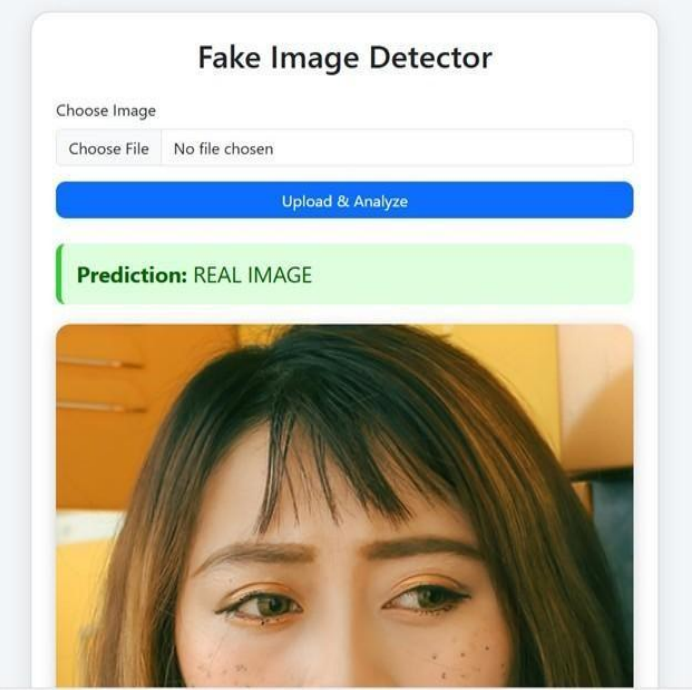

Fig. 2. Home Page

V. RESULTS

The proposed Fake Image Detection System using Deep Learning was successfully implemented and tested on a dataset containing both real and manipulated images. The system was able to effectively learn distinguishing features such as texture inconsistencies, noise patterns, and pixel-level variations between authentic and fake images.

During testing, the model achieved high accuracy in classifying images as real or fake. The performance was evaluated using metrics such as accuracy, precision, recall, and F1-score, which indicated that the system provides reliable and consistent results. The confusion matrix analysis showed that the model correctly identified most of the fake images with minimal misclassification.

The system was also able to process new, unseen images and generate predictions quickly, demonstrating its capability for real-time applications. The output included clear classification results along with confidence scores, making it easy for users to interpret the predictions.

Visualization of training performance, including accuracy and loss graphs, showed that the model learned effectively without significant overfitting. The results confirm that deep learning-based approaches are highly effective for detecting fake images, even when manipulations are subtle and not easily visible to the human eye.

Overall, the system demonstrates strong performance, efficiency, and reliability, making it suitable for applications such as digital forensics, social media monitoring, and content verification.

VI. CONCLUSION

The proposed Fake Image Detection System using Deep Learning demonstrates an effective and reliable approach for identifying manipulated and forged images in the digital world. By leveraging Convolutional Neural Networks (CNNs), the system is capable of analyzing complex image features such as textures, noise patterns, and pixel inconsistencies that are difficult to detect through

traditional methods or human observation.

The system successfully classifies images as real or fake with high accuracy, reducing dependency on manual verification and minimizing errors. The use of deep learning enables the model to continuously improve its performance when trained on diverse datasets, making it adaptable to different types of image manipulations including deepfakes and AI-generated content.

Furthermore, the automated workflow ensures fast processing and efficient handling of large volumes of image data, making it suitable for real-world applications such as social media monitoring, digital forensics, and cybersecurity. The integration of preprocessing, feature extraction, model training, and prediction creates a complete and robust detection pipeline.

Overall, the proposed system provides a scalable, efficient, and practical solution for combating fake images. It highlights the growing importance of deep learning in ensuring digital content authenticity and establishes a strong foundation for future advancements in AI-based image verification systems.

REFERENCES

1. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative Adversarial Nets. Proceedings of the Neural Information Processing Systems (NeurIPS).
2. Rossler, A., Cozzolino, D., Verdoliva, L., et al. (2019). FaceForensics++: Learning to Detect Manipulated Facial Images. IEEE International Conference on Computer Vision (ICCV).
3. Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). MesoNet: A Compact Facial Video Forgery Detection Network. IEEE Workshop on Information Forensics and Security.
4. Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
5. Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition (VGGNet). arXiv.