INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 06, Issue 01, January 2026, pp : 08-10

www.ijprems.com
editor@ijprems.com

e-ISSN : 2583-1062
Impact Factor : 7.001

FAKE IMAGE DETECTION USING DEEP LEARNING

Prof. Nithil Kulkarni[1], Juveriya Bepari[2], Khushi Bagade[3], Namrata Pujari[4], Nuthan KB[5]

[1]Asst. Prof, Department Of Electronics And Communication, K.L.S Vishwanathrao Deshpande Institute Of Technology, Haliyal, Karnataka, India.

[2,3,4,5]Student, Department Of Electronics And Communication, K.L.S Vishwanathrao Deshpande Institute Of Technology, Haliyal, Karnataka, India.

DOI: https://www.doi.org/10.58257/IJPREMS50863

ABSTRACT

The widespread availability of advanced image editing software and AI-based deepfake generation tools has significantly increased the circulation of manipulated images on digital platforms. This makes visual authenticity verification a challenging task. This work presents an automated fake image detection system using deep learning, specifically employing Convolutional Neural Networks (CNNs). The proposed model is trained on a balanced dataset consisting of authentic and manipulated images, enabling it to identify subtle artefacts introduced during image alteration. Data preprocessing and augmentation techniques are applied to enhance generalisation and reduce model bias. The system performance is evaluated using accuracy, precision, recall, and F1-score. Experimental results demonstrate that the proposed approach effectively distinguishes real images from fake ones, highlighting the role of deep learning in combating visual misinformation.

Keywords: Fake Image Detection, Deep Learning, CNN, Image Forensics, Deepfake Analysis, Image Classification, Misinformation Control.

1. INTRODUCTION

Digital images play a crucial role in modern communication through social media, online journalism, and digital documentation. However, the ease of manipulating images using AI-based tools has raised serious concerns regarding content authenticity. Fake images are frequently used to mislead audiences, spread misinformation, or damage reputations. Manual inspection methods are no longer sufficient due to the high realism achieved by modern image synthesis techniques.

Traditional image forensic methods relied on identifying visual inconsistencies such as colour imbalance or lighting anomalies. While effective for basic manipulations, these approaches fail against sophisticated AI-generated forgeries. Consequently, deep learning-based techniques have emerged as reliable solutions for automated image verification.

Convolutional Neural Networks (CNNs) are particularly suitable for image analysis as they automatically extract hierarchical features from visual data. These models can detect a wide range of manipulations, including splicing, copy-move forgery, and deepfake generation. This project focuses on developing a CNN-based framework to improve the accuracy and reliability of fake image detection.

2. METHODOLOGY

The proposed system follows a structured deep learning pipeline to classify images as real or fake. The workflow includes dataset collection, preprocessing, model training, validation, and performance evaluation

2.1 Dataset Collection and Preprocessing

The dataset consists of genuine and manipulated images obtained from verified sources. All images are resized to a uniform resolution to ensure consistency. Noise reduction and image enhancement techniques are applied to improve data quality. Data augmentation methods such as rotation, flipping, and scaling are used to increase dataset diversity and improve model robustness.

2.2 CNN Model Design and Evaluation

A CNN is used as the main model for classification. It has layers that help extract features, layers that reduces the size of data, and layers that make the final classification. The model is trained with images that have labels and uses common loss functions to improve its performance.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 06, Issue 01, January 2026, pp : 08-10
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

3. MODELING AND ANALYSIS

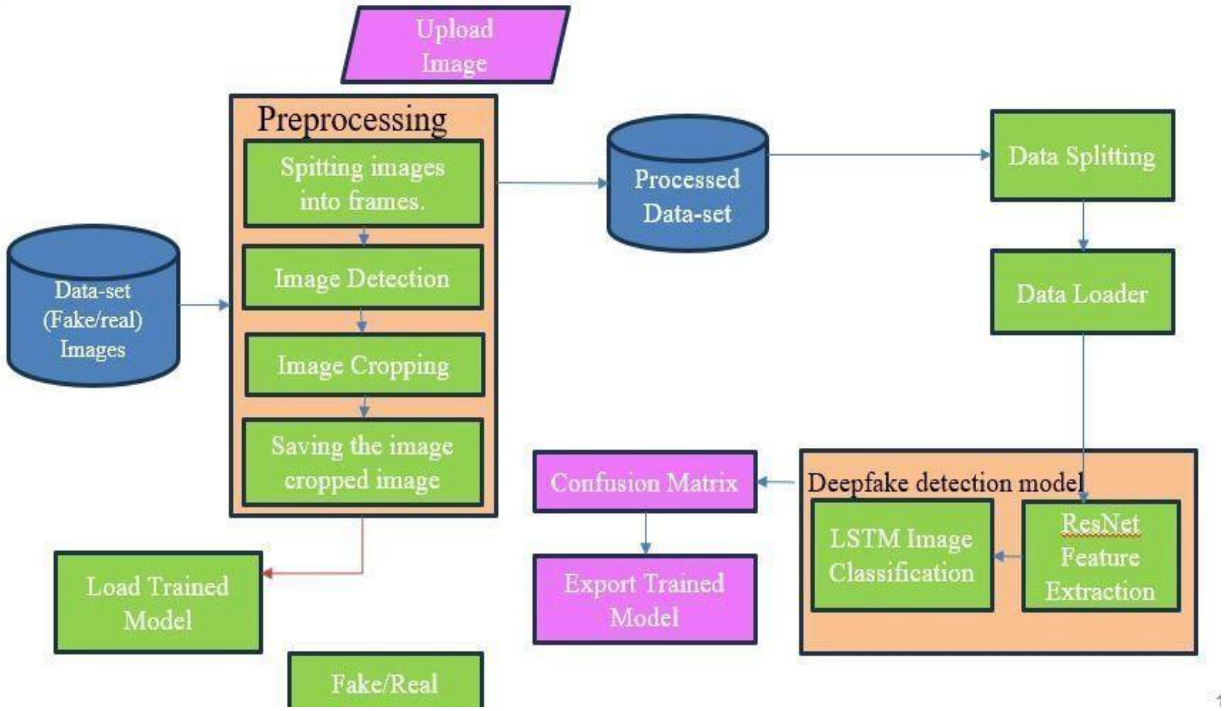

Figure 1: Image detection process.

4. PERFORMANCE EVALUATION

The model is evaluated using accuracy, precision, recall, and F1-score metrics. These metrics provide a comprehensive assessment of the model's classification capability.

5. MODELLING AND ANALYSIS

In this work, a deep learning–based modelling approach is employed to detect deepfake images by learning discriminative visual patterns that differentiate authentic images from manipulated ones. The proposed system is formulated as a supervised binary classification model, where input facial images undergo preprocessing steps such as resizing, normalisation, and data augmentation to ensure consistency and improve generalisation. A convolutional neural network is used as the core feature extraction module due to its ability to capture hierarchical spatial information and subtle artefacts introduced during deepfake generation, including texture inconsistencies, unnatural facial boundaries, and blending irregularities.

The extracted deep features are passed through fully connected layers to perform classification, with a sigmoid activation function producing the probability of an image being fake or real. Model training is carried out using binary cross-entropy loss and optimised through adaptive gradient-based optimisation techniques to ensure stable convergence. Regularisation strategies such as dropout and early stopping are incorporated to minimise overfitting and enhance robustness. The performance of the model is analysed using standard evaluation metrics, including accuracy, precision, recall, and F1-score, which provide a comprehensive assessment of detection reliability. Experimental analysis indicates that the model effectively captures manipulation-specific features and maintains strong performance across different deepfake generation methods, demonstrating the suitability of deep learning–based modelling for reliable and scalable deepfake image detection.

6. RESULTS AND DISCUSSION

Experimental results indicate that the CNN model effectively detects manipulated images by identifying artefacts such as irregular textures, boundary distortions, and lighting inconsistencies. Data augmentation improves generalisation performance. However, detecting highly realistic deepfakes remains a challenge, indicating scope for further improvement.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
www.ijprems.com
editor@ijprems.com
Vol. 06, Issue 01, January 2026, pp : 08-10
e-ISSN : 2583-1062
Impact Factor : 7.001

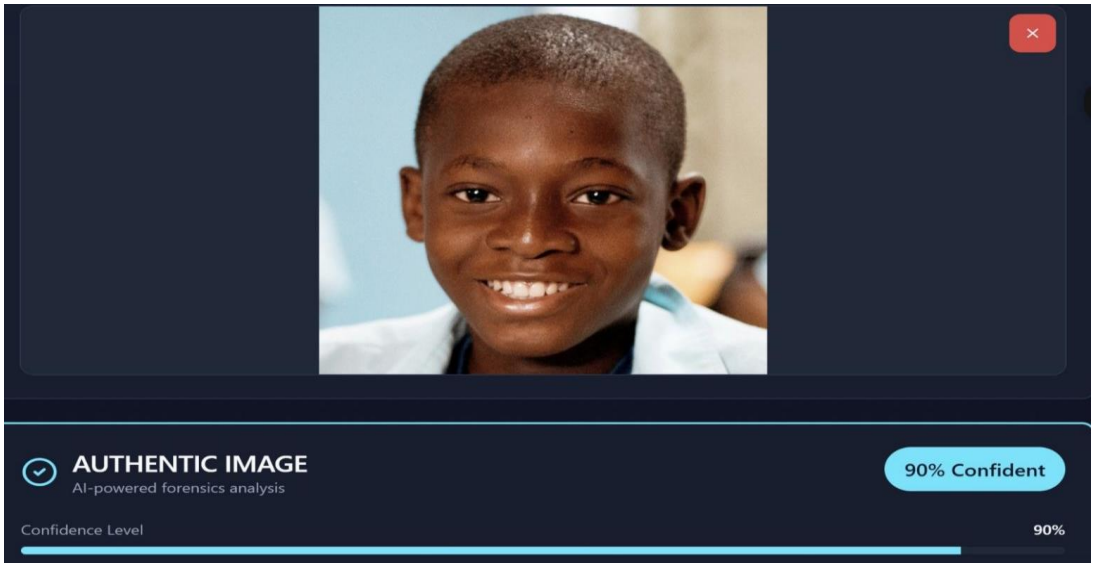

Figure 2: Final result

7. CONCLUSION

This work presents an effective deep learning-based approach for fake image detection. The CNN model demonstrates superior performance compared to traditional forensic techniques. Proper dataset preparation significantly impacts detection accuracy. Although challenges remain with advanced deepfakes, the proposed system serves as a valuable tool for digital forensics and content verification.

ACKNOWLEDGEMENTS

We would like to begin by expressing our heartfelt gratitude to our esteemed Principal, Dr V.A. Kulkarni, for his unwavering support towards our development, also extend our sincere appreciation to our respected Head of Department, Dr Nagaraj Bhat, for his constant inspiration and encouragement throughout this initial phase of the project. Furthermore, we express our sincere thanks to Prof. Nikhil Kulkarni for his exemplary efforts in guiding us, as well as his commitment to our success. His dedication and timely guidance have been invaluable, and we are truly grateful for his patience in addressing our doubts.

Finally, we wish to convey our deepest appreciation to our Major Project Co-Ordinator, Prof. Vijayalaxmi K, Prof. Raghavendra N, and Prof. Ashwini G for their thoughtful advice and invaluable guidance.

8. REFERENCES

[1] Vijayalaxmi C. Handaragall, Jyoti Subhash Metan, “Comprehensive Analysis of Deepfake Detection Techniques for Images and Videos Using Deep Learning,” 2025.

[2] Anjali Singh et al., “Impact of Deep Learning Techniques on Deepfake Image Identification for Digital Investigation,” 2024.

[3] Venkatagopi Narne et al., “Advanced Deep Learning Approaches for Detecting Real and Fake Images,” 2024.

[4] Mohammed Hasan Mutar et al., “Innovative Deep Learning Solutions for Image Forgery Detection,” 2024.