IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, January 2026
Impact Factor: 7.67

Adversarial Media Detection Using Convolutional Neural Networks

Chandana T K[1], D S Chinmayi[2], Kusuma S B[3], Likhitha M P[4], Dr. S. V. Rajashekararadhya[5]
Student, Department of Electronics and Communication Engineering [1-4]
Professor & HOD, Department of Electronics and Communication Engineering[5]
Kalpataru Institute of Technology, Tiptur, India

Abstract: *Often it becomes necessary for an engineer or somebody else to handle drawings or images of objects, which is in existence. One certainly finds oneself ill-equipped when actually facing the task of having to produce drawings and designs with reasonable accuracy in limited time with security. Securing communications is an important aspect in the present era of digital and wireless communication. The objective of communication security is to protect the message from unauthorized users. Hence we describe an effective method for image encryption which employs logical manipulation of pixel value of the imageusing4out of 8 code and performing the reverse process for decrypting the image. Encryption key may be having the length of 'n' alphanumeric characters, where n should be less than or equal to the size of image (no. of pixels in that image). And here we assign4out of8 code to each alphanumeric character and that is used for pixel manipulation for image encryption.*

Keywords: Image, Encryption, Decryption, 4 out of 8 code

I. INTRODUCTION

As modern advancements in artificial intelligence continue to evolve, sophisticated AI techniques are increasingly employed to generate deceptive videos known as adversarial media or deepfakes. These manipulated videos, created by powerful AI algorithms, pose significant threats to society, impacting various social and political dimensions. adversarial media or Deepfake detection using Convolutional Neural Networks (CNNs)represents a critical advancement in the ongoing battle against deceptive media manipulation. With the proliferation of sophisticated AI tools enabling the creation of increasingly convincing deepfakes, there's a pressing need for robust detection mechanisms. CNNs, a class of deep learning models well- suited for image analysis tasks, offer a promising solution. By leveraging their ability to automatically learn hierarchical features from images, CNN-based deepfake detectors can effectively discern subtle visual inconsistencies indicative of tampering. The key strength of CNN-based deepfake detection lies in its capacity to analyze vast amounts of visual data with remarkable speed and accuracy. Through a process called feature extraction, CNNs can identify unique patterns and features within images that may indicate manipulation. By training on large datasets containing both authentic and deepfake content, these networks learn to distinguish between genuine and synthetic media, continually refining their ability to flag suspicious material. Moreover, CNN- based deepfake detection systems can adapt to evolving threats by leveraging techniques such as transfer learning. By finetuning pre-trained CNN models on new data, these systems can quickly adapt to emerging deepfake techniques, enhancing their detection capabilities over time. Additionally, researchers are exploring innovative approaches such as multi modal analysis, combining visual and audio cues to improve detection accuracy further. As deepfake technology continues to evolve, the development of sophisticated CNN-based detection methods remains crucial in preserving trust and integrity in digital media.

II. PROBLEM STATEMENT

Developing a web application utilizing advanced deep learning models for real-time detection of adversarial media images and videos addresses the critical challenge of deceptive media dissemination.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, January 2026
Impact Factor: 7.67

III. METHODOLOGY

The methodology for adversarial media detection consists of several stages:

[1] Data Collection

A dataset containing both real and deepfake images and videos is collected from publicly available sources. The dataset includes variations in lighting, facial expressions, angles, and backgrounds.

[2] Preprocessing

Video files are converted into frames. All images are resized, normalized, and cleaned to remove noise. Data augmentation techniques such as flipping, rotation, and scaling are applied to improve model generalization.

[3] CNN Architecture

A Convolutional Neural Network is designed with convolution layers, pooling layers, activation functions, and fully connected layers. The CNN automatically extracts spatial features from input images and learns patterns associated with deepfake artifacts.

[4] Training and Testing

The dataset is divided into training, validation, and testing sets. The CNN is trained using binary cross-entropy loss and optimized using the Adam optimizer. Performance metrics such as accuracy, precision, recall, and F1-score are evaluated.

[5] Web Application Integration

The trained model is integrated into a Django-based web application. Users can upload media files, which are analyzed by the model, and results are displayed in real time.

IV. SYSTEM ARCHITECTURE

The system architecture consists of the following components:

1. User Interface: Developed using HTML, CSS, and JavaScript.
2. Backend Server: Implemented using Python and Django.
3. CNN Model: Performs feature extraction and classification.
4. Database: MySQL stores user details and detection history.
5. Media Processor: Handles video frame extraction and image preprocessing.

The user uploads an image or video → preprocessing → CNN analysis → classification → result display.

V. RESULTS AND DISCUSSION

The CNN-based adversarial media detection system was tested using multiple real and fake images and videos. The system successfully identified deepfake content with high accuracy. The detection process was fast and efficient, making it suitable for real-time applications.

The model demonstrated strong performance in detecting facial inconsistencies, unnatural lighting, and texture artifacts. False positives and false negatives were minimal. The web interface provided smooth interaction and clear result visualization. Overall, the results confirm that CNNs are effective in detecting adversarial media and can significantly contribute to reducing digital misinformation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, January 2026
Impact Factor: 7.67

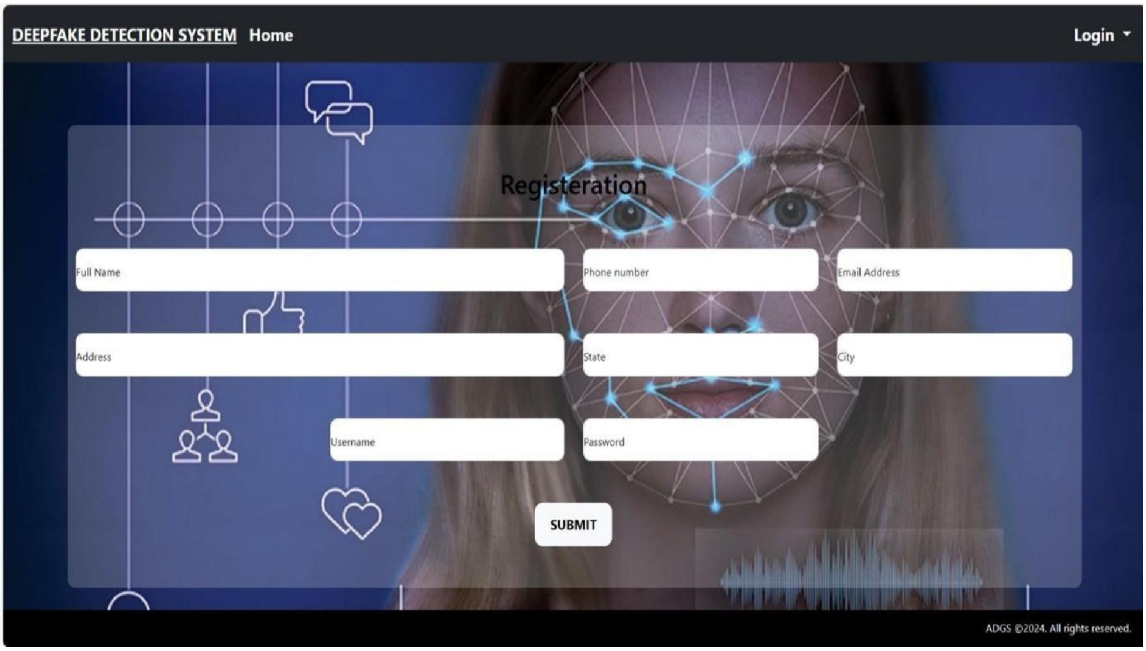

Figure : Registration page

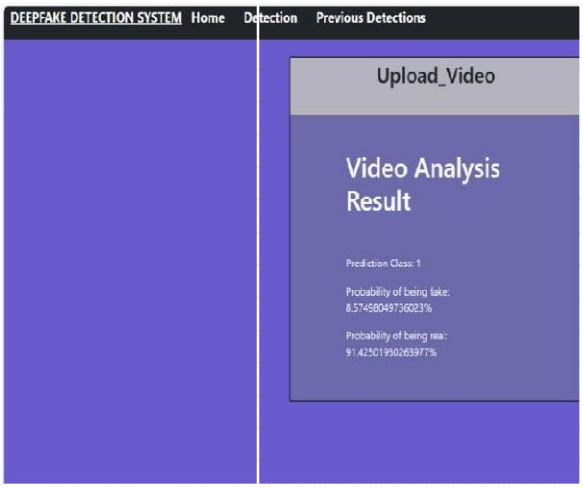

Figure : Detection Page

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, January 2026
Impact Factor: 7.67

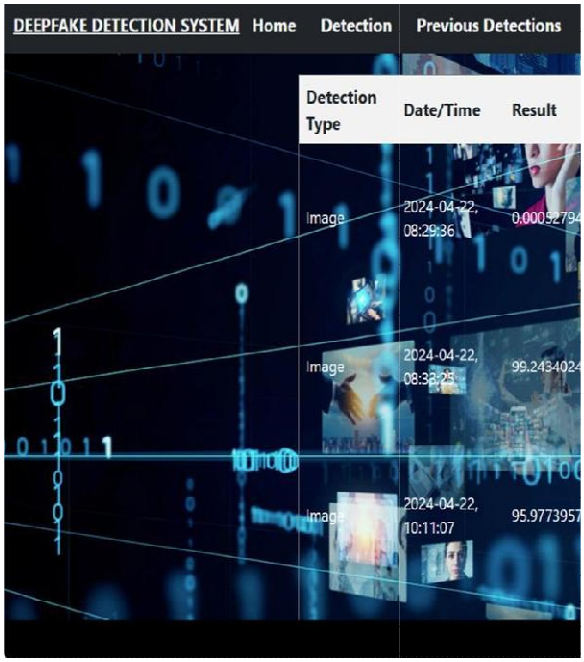

Figure : Previous Detection page

VI. CONCLUSION

The application of Convolutional Neural Networks (CNNs) in adversarial media or deepfake detection represents a significant advancement in combating the proliferation of manipulated media. By leveraging the power of deep learning, CNNs can effectively discern subtle visual cues and patterns indicative of digital tampering, enabling the detection of deepfakes with a high degree of accuracy. This technology holds immense promise in mitigating the potential harms associated with the spread of misinformation and fake content across various online platforms. Furthermore, the robustness of CNN-based deepfake detection systems stems from their ability to adapt and learn from large datasets containing both authentic and manipulated media. Through extensive training on diverse datasets, CNNs can generalize well to new and unseen deepfake variations, enhancing their effectiveness in real-world scenarios. This adaptability is crucial in staying ahead of evolving deepfake generation techniques and maintaining the efficacy of detection systems over time. Moreover, the deployment of CNN-based deepfake detection systems can serve as available tool for both individuals and organizations seeking to verify the authenticity of multimedia content. By integrating these systems into social media platforms, newsagencies, and other online platforms, users can make more informed decisions about the veracity of the media they encounter, thereby mitigating the potential spread of misinformation and preserving trust in digital content. In conclusion, while CNN-based deep fake detection systems represent a significant step forward in combating digital manipulation, ongoing research and development are essential to further improve their accuracy, scalability, and usability. Continued collaboration between researchers, industry stakeholders, and policymakers is vital in advancing the field of deepfake detection and safeguarding the integrity of digital content in the increasingly complex media landscape

VII. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, for providing the necessary facilities and support to carry out this project. We are extremely thankful to our project guide for their valuable guidance, continuous encouragement, and technical support throughout the development of this project.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, January 2026 Impact Factor: 7.67

We also extend our heartfelt thanks to the Head of the Department and all the faculty members of the ECE department for their constant motivation and constructive suggestions. Finally, we would like to thank our family and friends for their support and encouragement, which helped us successfully complete this project.

REFERENCES

[1]. Ankur Nagulwar, Sejal Shingvi, Palak Takhtani. "DEEP FAKE VIDEODETEC-TION USING DEEP LEARNING."

[2]. S Jeevidha, S. Saraswathi, Kaushik J B, Preethi K, NallamVenkataramaya. "DEEPFAKE VIDEO DETECTION USING RES- NEXT CNN ANDLSTM" International Journal of Creative Research Thoughts (IJCRT), 2023.

[3]. Yash Doke, Prajwalita Dongare, Vaibhav Marathe, Mansi Gaikwad, Mayuri Gaik-wad. "DEEP FAKE VIDEO DETECTION USING DEEP LEARNING", Inter-national Journal of Research Publication and Reviews, Vol 3, no 11, pp540-544, November 2022. www.ijrpr.com

[4]. Suganthi, S. T, Mohamed Uvaze Ahamed Ayoobkhan, Nebojsa Bacanin, K. Venkat-achalam, Hubálovský Štěpán, and Trojovský Pavel. "DEEP LEARNINGMODELFOR DEEP FAKE FACE RECOGNITION AND DETECTION." PeerJ Computer Science 8 (2022). e881. DO| 10.7717/peerj-cs.881

[5]. "A NOVEL DEEP LEARNING APPROACH FOR DEEPFAKE IMAGE DETECTION. "Applied Sciences 12,no.19(2022):9820.https://doi.org/10.3390/app121998206

[6]. Artem A. Maksutov , Viacheslav O. Morozov, Aleksander A. Lavrenov, Alexander S. Smirnov."METHODS OF DEEPFAKE DETECTION BASED ON MACHINE LEARNING"