IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141172

AN OVERVIEW OF "AI FAKEBUSTER": A DEEPFAKE DETECTION APP

Prof. Priya Farkade*[1], Uday Lanjewar[2], Ranit Garude[3], Mohan Khobarkhede[4], Rohit Bawanukey[5], Manthan Ukey[6]

Professor Department Of Computer Science And Engineering, Nagarjuna Institute Of Engineering Technology And Management, Nagpur, Maharashtra, India*[1]

UG Student, Department Of Computer Science And Engineering, Nagarjuna Institute Of Engineering Technology & Management, Nagpur, Maharashtra, India[2,3,4,5,6]

Abstract: Deepfake leverage machine learning algorithms, particularly Generative Adversarial Networks (GANs), to create highly realistic images, videos, or audio recordings of individuals. By learning from vast datasets, these models can generate media that mimics real-life behavior, expressions, and voices, making them difficult to identify as fake. The primary objective of this project is to develop a Deepfake Detection System that can effectively identify and classify manipulated media using advanced deep learning and computer vision techniques. This system aims to address the security, ethical, and social implications posed by synthetic media by providing a reliable tool for the detection and analysis of deepfakes in both static images and video sequences.

Keywords: Deepfake, Fake Image Detection, Fake Video Detection, Image Dataset, Video Dataset.

I. INTRODUCTION

In recent years, the emergence of deepfake technology has raised significant concerns regarding its potential to manipulate digital media and deceive audiences worldwide. Deepfakes, which utilize advanced machine learning algorithms to create highly realistic synthetic media, pose a serious threat to the integrity of visual and audio content on the internet. These maliciously altered videos, images, and audio recordings can be used to spread disinformation, impersonate individuals, and undermine trust in online information sources.

In the following sections of this paper, we will delve into the technical details of our proposed AI driven approach to deepfake detection and prevention. We will explore the various machine learning algorithms, computer vision techniques, and audio analysis methods employed in our framework, highlighting their effectiveness in differentiating between authentic and manipulated media. Additionally, we will discuss the importance of proactive measures and collaborative efforts in mitigating the risks associated with deepfake technology and preserving the integrity of digital media in the age of AI.

II. METHODOLOGY

The development of the AI Fakebuster Application follows a systematic methodology to ensure high detection accuracy and model interpretability. The methodology involves multiple phases, including data acquisition, preprocessing, model training, evaluation, and deployment.

1) **Data Collection:**
   Deepfake datasets such as DFDC (Deepfake Detection Challenge) and Celeb-DF were used for training and validation. These datasets contain both authentic and manipulated media to teach the model to distinguish subtle facial distortions and blending artifacts.
2) **Preprocessing:**
   Each frame is extracted from video inputs, and facial regions are detected using MTCNN (Multi-task Cascaded Convolutional Networks). The extracted faces are resized and normalized for training consistency.
3) **Model Training:**
   The model is built on the EfficientNet-B0 architecture, fine-tuned on the Deepfake dataset using TensorFlow. The binary classification layer outputs probabilities for *real* and *fake* predictions.
4) **Grad-CAM Integration:**
   To provide interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) is integrated, allowing visualization of areas contributing to the classification. This improves model transparency and user trust.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141172

5) Application Development:
A Streamlit-based interface was created for easy interaction. The app allows users to upload images/videos or enter URLs. It displays predictions, confidence scores, cropped faces, and Grad-CAM heatmaps.

6) Testing and Validation:
The model is evaluated on unseen test data to measure accuracy, precision, recall, and F1-score. Iterative testing and fine-tuning help reduce false positives and improve generalization.

III. FLOWCHART

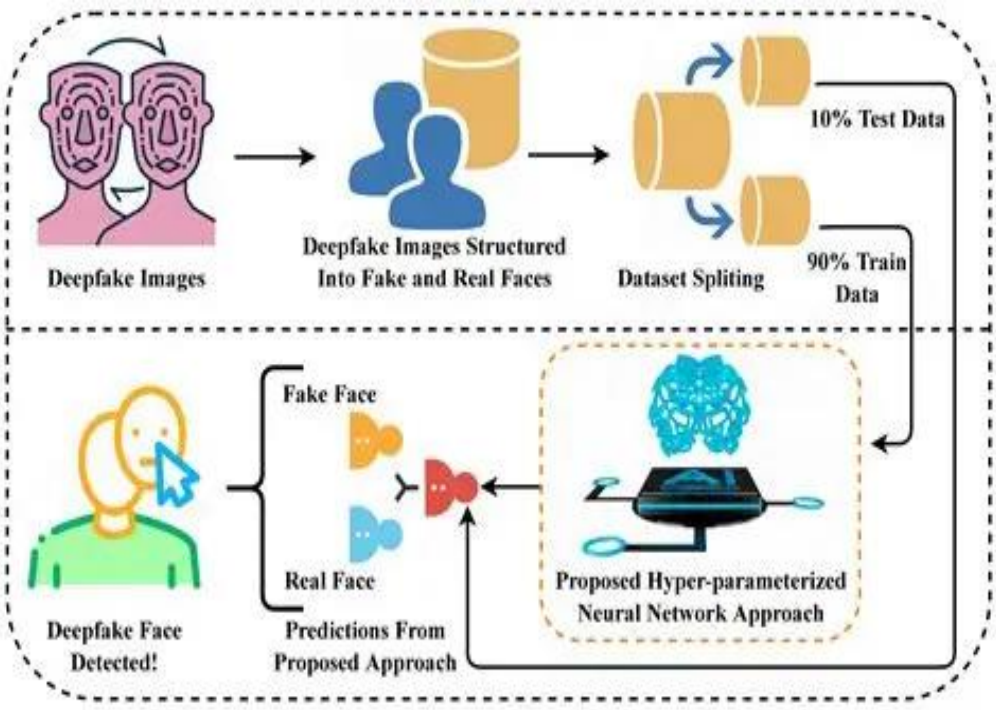

Fig.1: Flow of AI FakeBuster App

IV. RESULTS AND DISCUSSION

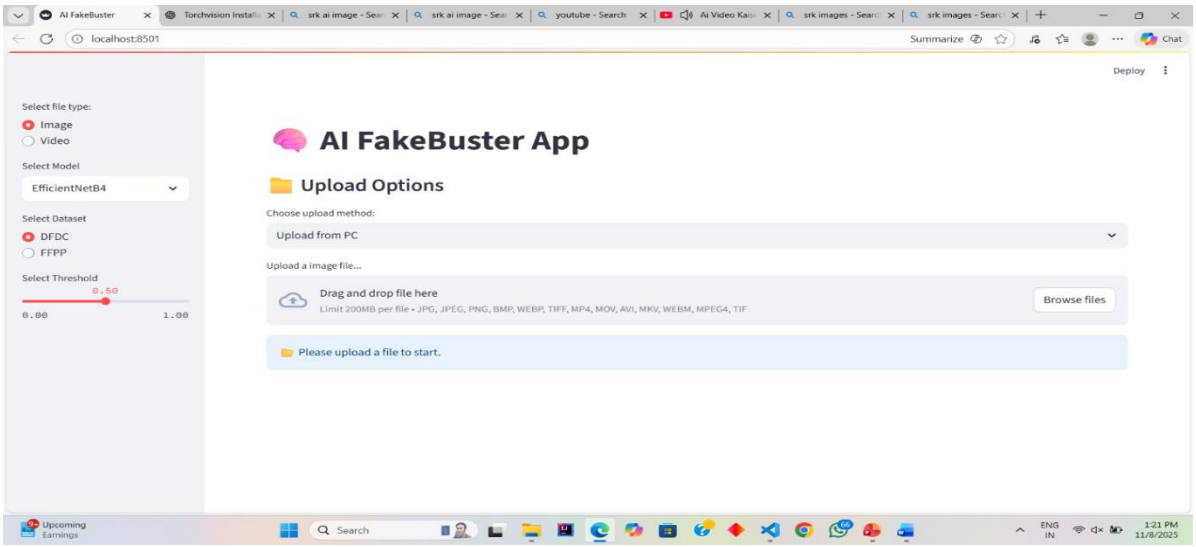

Fig.2: AI FakeBuster App Home Page

Discussion: It is the homepage of the project from where we can navigate to other pages of the website i.e., image, video, module etc.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141172

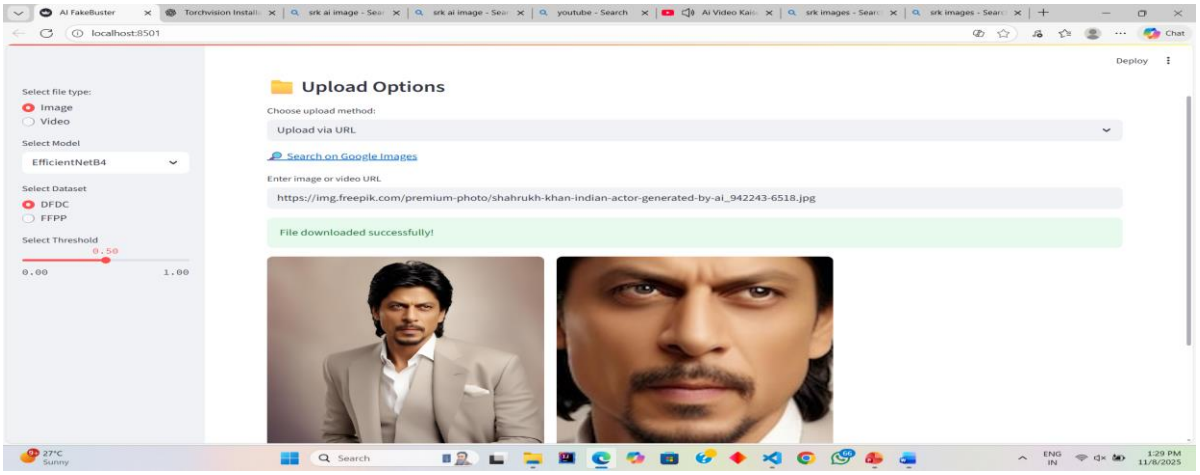

Fig.3: Browse Image/Video

Discussion: Browse Image/Video is the step where can Select Image and Video for Analysis.

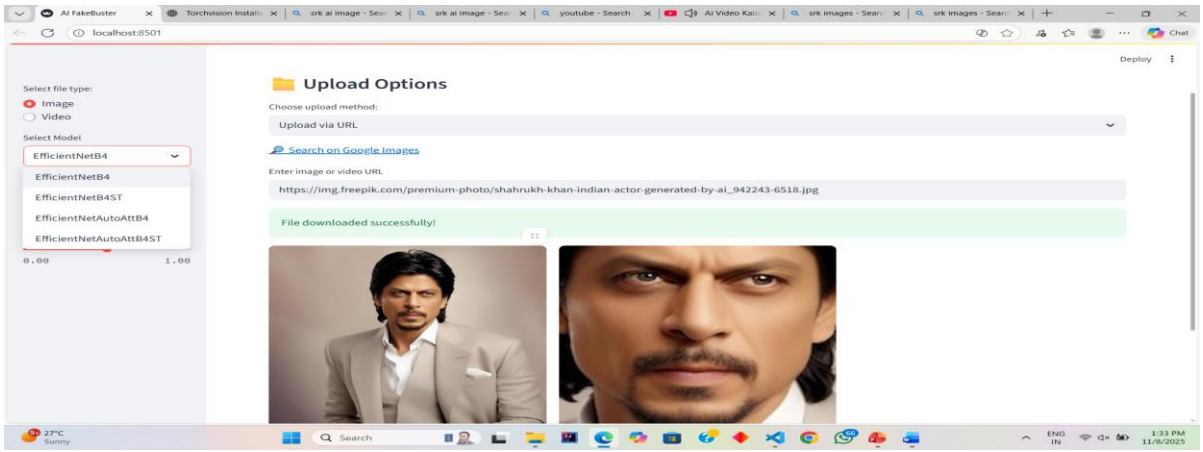

Fig.4: Model Selection

Discussion: Here can select the required Models for the analysis of Image/Video.

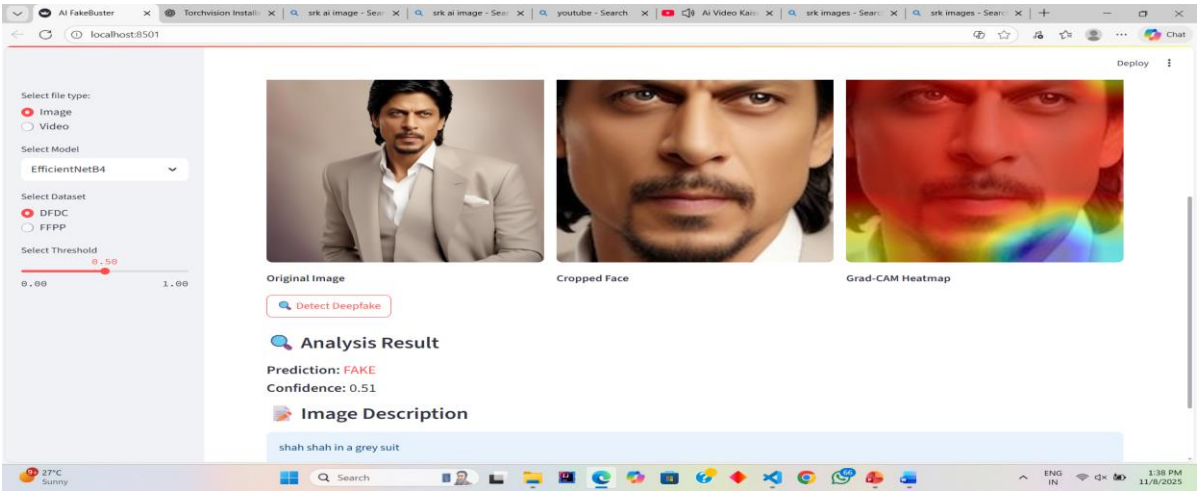

Fig.5: Check for Deepfake

Discussion: Here the Output of Image/Video given as Real or Fake with Probability.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.141172

V. CONCLUSION

In conclusion, the proposed deepfake detection system offers a comprehensive and adaptive approach to combat the malicious use of deepfake technology. By integrating a hybrid of manual inspection, traditional forensic techniques, and state-of-the-art machine learning algorithms, the system achieves high accuracy, scalability, and robustness in identifying manipulated media across various formats.

Throughout this paper, we have highlighted the importance of addressing the multifaceted challenges posed by deepfake technology, including scalability limitations, vulnerability to adversarial attacks, ethical considerations, and regulatory compliance. The proposed system addresses these challenges by leveraging advancements in machine learning, fostering interdisciplinary collaboration, and prioritizing ethical principles and privacy protection.

By adopting a proactive and collaborative approach, we can mitigate the risks associated with deepfake technology and safeguard the integrity of digital media in the age of AI. However, the fight against deepfakes is an ongoing battle, and there are several avenues for future enhancement and research.

REFERENCES

[1]. Dr. M.K. Jayanthi Kannan, Rajat Gore, Apoorva Kharya, Divyansh Sahu, Shivam Kabra (2024). International Journal of Novel Research and Development.
[2]. Ryan Patel, Jennifer Garcia (2023). Deepfake Detection in the Wild: Challenges and Opportunities. ACM Computing Surveys.
[3]. Daniel Rodriguez, Sofia Nguyen, (2022). Deepfake Detection on Social Media: Scale, Efficiency, and Privacy Considerations. Journal of Privacy and Confidentiality.
[4]. Gupta, A., & Jaiswal, S. (2021). DeepFake Detection Techniques: A Survey. International Journal of Advanced Computer Science and Applications.
[5]. Cholleti, S. R., & Reddy, V. U. (2021). DeepFake Detection: A Survey. IEEE Access.
[6]. Zhang, Y., Li, Y., Wang, J., & Li, Y. (2020). Deep Learning for Deepfakes Detection: A Comprehensive Survey. IEEE Transactions on Multimedia.
[7]. Menon, A. K., Balasubramanian, V. N., & Jain, R. (2020). A Survey on Deep Learning Techniques for Video-based Deepfake Detection. Pattern Recognition Letters
[8]. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
[9]. Khan, S. S., & Madden, M. G. (2020). Deepfake videos detection using recurrent neural networks. arXiv preprint arXiv:2001.00157.
[10]. Raghavendra, N., & Venkatesan, S. (2020). Deepfake Videos Detection and Classification Using Convolutional Neural Networks. arXiv preprint arXiv:2010.05540.
[11]. Yuezun Li, Ming-Ching Chang, Hany Farid, Siwei Lyu (2018). In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking.
[12]. Wang, Y., Ma, F., & Luo, C. (2017). Detecting Fake News for Effective News Analysis: A Deep Learning Approach. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (pp. 223-232).