RESEARCH ARTICLE OPEN ACCESS

Deep Fake Image Detection Using Hybrid Of Efficient Net and Xception

C.Ramya[1], R. Vaishnavi[2], K.R. Hemadarshiny[3], V. Subasri[4]

[1] *Professor, Department of Computer Science and Engineering, Parisutham Institute of Technology and Science, Thanjavur, Tamil Nadu – 613006, India*
Emai: v.ramya81@gmail.com

[2]*UG Student, Department of Computer Science and Engineering, Parisutham Institute of Technology and Science, Thanjavur, Tamil Nadu – 613006, India*
Email: : vaishuraman04@gmail.com

[3] *UG Student, Department of Computer Science and Engineering, Parisutham Institute of Technology and Science, Thanjavur, Tamil Nadu 613006, India*
Email darshinyhema519@gmail.com

[4]*UG Student, Department of Computer Science and Engineering, Parisutham Institute of Technology and Science, Thanjavur, Tamil Nadu – 613006, India*
Email: pavisubasri130@gmail.com

Abstract:

The rapid advancement of artificial intelligence and deep learning technologies has led to the widespread creation of deepfake images, posing serious threats to digital security, misinformation control and personal identity protection. Detecting such manipulated content has become a critical challenge in today's digital world. This paper presents a Deepfake Image Detection System that focuses on identifying forged facial images using a hybrid deep learning approach.

The proposed system utilizes a combination of EfficientNet and Xception models to improve detection accuracy. EfficientNet captures global image features with optimized performance, while Xception extracts fine-grained spatial details and manipulation artifacts. By integrating both models, the system enhances the ability to detect subtle inconsistencies present in deepfake images.

The system is implemented with a user-friendly web interface, allowing users to upload images and receive real-time predictions indicating whether the image is real or fake. The backend processes the input image through preprocessing, feature extraction and classification stages to generate accurate results.Experimental results demonstrate that the hybrid model achieves high detection accuracy and robustness compared to individual models. The proposed approach provides an efficient and reliable solution for combating deepfake threats and ensuring digital media authenticity.

Keywords—Deepfake Detection, EfficientNet, Xception, Image Classification, Deep Learning, Digital Security

I. INTRODUCTION

The rapid advancement of artificial intelligence and deep learning technologies has led to the creation of highly realistic deepfake images. These manipulated images can alter facial features and expressions, making it difficult to distinguish between real and fake content. The widespread availability of such technology poses serious threats, including misinformation, identity misuse, digital fraud and loss of trust in visual media. Therefore, detecting deepfake images has become a critical challenge in ensuring digital security and media authenticity.

Traditional image verification methods rely on manual inspection or basic image processing techniques, which are often insufficient to identify subtle manipulations present in deepfake images. As deepfake generation techniques continue to improve, there is a growing need for more advanced and automated detection systems that can accurately analyze and classify images.

To address this challenge, deep learning-based approaches have emerged as effective solutions for image classification and forgery detection. In particular, models like EfficientNet and Xception have shown strong performance in extracting meaningful features from images. EfficientNet focuses on capturing global patterns with optimized scaling, while Xception uses depthwise separable convolutions to detect fine-grained details and manipulation artifacts.

The proposed system utilizes a hybrid approach by combining EfficientNet and Xception to improve deepfake detection accuracy. The system includes an interactive web interface that allows users to upload images and receive real-time predictions. By integrating advanced feature extraction and classification techniques, the system provides a reliable solution for identifying deepfake images while maintaining efficiency and usability.

II. RELATED WORK

Nguyen et al. [1] proposed a deep learning-based approach for detecting deepfake images using convolutional neural networks (CNNs). Their method focuses on identifying visual inconsistencies such as unnatural textures, facial distortions and blending artifacts present in manipulated images. By training CNN models on large datasets, the system learns to differentiate between real and fake images effectively. However, the approach mainly relies on single-model architecture, which may limit detection performance in complex scenarios. The study highlights the importance of feature extraction in deepfake detection tasks.

Rossler et al. [2] introduced the FaceForensics++ dataset and evaluated multiple deep learning models for detecting manipulated facial images. Their work provides a benchmark for analyzing the performance of different detection techniques under various compression levels and manipulation methods. The study demonstrates that deep learning models can effectively capture manipulation artifacts, but performance may degrade with highly compressed images. This work emphasizes the need for robust models capable of handling real-world variations.

Chollet [3] proposed the Xception model, which utilizes depthwise separable convolutions to improve feature extraction efficiency. The architecture is designed to capture fine-grained spatial details in images, making it highly suitable for detecting subtle deepfake artifacts. Its ability to focus on pixel-level inconsistencies makes it a strong candidate for forgery detection tasks. The study highlights improved performance compared to traditional CNN architectures.

Tan and Le [4] introduced EfficientNet, a model that optimizes network scaling to achieve better accuracy with fewer parameters. EfficientNet captures global image features effectively while maintaining computational efficiency. The model's balanced scaling approach allows it to perform well on various image classification tasks. This makes it suitable for identifying overall patterns in deepfake images.

Afchar et al. [5] developed MesoNet, a CNN-based model specifically designed for deepfake detection. Their approach focuses on mesoscopic features, which lie between low-level pixel details and high-level semantic features. The model achieves good performance in detecting common deepfake manipulations but may struggle with highly sophisticated forgeries. The work highlights the importance of combining multiple feature levels for better accuracy.

Dang et al. [6] proposed a multi-task learning framework that jointly detects deepfakes and segments manipulated regions within images. Their approach improves interpretability by highlighting the exact regions where manipulation occurs. The system enhances detection accuracy by learning both classification and localization tasks simultaneously. This work demonstrates the benefit of combining multiple objectives in deepfake detection.

Li et al. [7] explored detection techniques based on biological signals such as eye blinking patterns. Their study identifies unnatural blinking behavior in deepfake videos as a key indicator of manipulation. While effective in certain scenarios, the approach may not generalize well to all types of deepfakes. The work highlights the potential of using domain-specific features for detection.

Tolosana et al. [8] presented a comprehensive survey on deepfake detection methods, covering image, video and audio-based techniques. The study compares various machine learning and deep learning approaches, emphasizing their strengths and limitations. It also identifies challenges such as generalization, dataset bias and robustness against new manipulation techniques. The survey provides valuable insights into current trends and future directions.

Verdoliva [9] reviewed multimedia forensics techniques for detecting manipulated content. The work discusses both traditional and deep learning-based methods for identifying digital forgeries. It highlights the growing need for automated detection systems due to the increasing sophistication of deepfake technologies. The study reinforces the importance of combining multiple detection strategies.

Mirsky and Lee [10] analyzed the evolution of deepfake generation and detection techniques. Their work provides an overview of how generative models have advanced and how detection methods must continuously adapt. The study emphasizes the need for hybrid models that can capture both global and local features to improve detection performance. This supports the effectiveness of combining models like EfficientNet and Xception.

III. PROPOSED METHODOLOGY

The proposed system is designed to detect deepfake images by analyzing facial features using advanced deep learning techniques. The system integrates image acquisition, preprocessing, hybrid feature extraction, classification and result generation within a unified framework. By combining powerful models such as EfficientNet and Xception, the system ensures accurate detection of manipulated images.

EfficientNet captures global image patterns while Xception focuses on fine-grained spatial details and manipulation artifacts. This hybrid approach improves detection accuracy and robustness. The system allows users to upload images and obtain real-time predictions indicating whether the image is real or deepfake. The overall workflow consists of several stages including image acquisition, preprocessing, feature extraction, classification and result generation.

a. System Overview

The system for Deepfake Image Detection using EfficientNet and Xception is designed with multiple modules, including Image Upload, Preprocessing, Feature Extraction, Classification, and Result Display.Initially, the user uploads an image, which is then preprocessed to enhance quality, resize, and normalize it for better analysis. The processed image is passed to a hybrid deep learning model that combines EfficientNet and Xception to extract both global and fine-grained features.These features are used to accurately classify the image as real or deepfake. Finally, the system displays the prediction result to the user. The overall system ensures high accuracy, efficient processing, and reliable performance for real-time deepfake detection

b. User & Image Upload Module

The User & Image Upload Module serves as the entry point of the deepfake detection system, where the user interacts with the application through a simple and user-friendly interface. It allows users to upload images for analysis and initiates the detection process.

The system supports common image formats such as JPG, JPEG, and PNG to ensure compatibility with most user inputs. Once an image is selected, the module performs validation checks, including verifying the file format, checking file size limits, and ensuring that the file is not corrupted or invalid. These checks help maintain system stability and prevent errors during processing.

In addition, the module may provide basic user feedback, such as upload success messages or error notifications if the file does not meet the required criteria. Security measures are also considered to prevent unauthorized or harmful files from entering the system.After successful validation, the uploaded image is securely transferred to the preprocessing module for further analysis. Overall, this module ensures smooth input handling, reliability, and a seamless user experience in the deepfake detection workflow.

c. Image Preprocessing Module

The Image Preprocessing Module plays a crucial role in preparing the uploaded image for accurate deepfake detection. Raw images may vary in size, quality, lighting, and noise levels, which can affect model performance. Therefore, this module applies a series of transformations to standardize and enhance the input data before it is passed to the feature extraction stage.

The preprocessing steps include:

- **Resize:** The image is resized to a fixed dimension required by the EfficientNet and Xception models, ensuring compatibility and consistent input shape.
- **Normalization:** Pixel values are scaled (typically between 0 and 1) to improve computational efficiency and help the model converge faster during prediction.
- **Noise Reduction:** Unwanted distortions, blur, or artifacts are reduced using filtering techniques, improving the clarity of important features.
- **Image Standardization:** The image format, color channels, and intensity distribution are standardized to maintain uniformity across all inputs.

Additionally, this module may include operations such as contrast adjustment and color correction to further enhance image quality. By reducing inconsistencies and improving clarity, preprocessing helps the model focus on meaningful patterns rather than irrelevant variations.

Overall, this module ensures that the input image is clean, consistent, and optimized, significantly improving the accuracy and reliability of the deepfake detection system.

.

d. Database Connection

The Database Connection Module is responsible for managing the storage and retrieval of data within the deepfake detection system. It acts as a bridge between the application and the database, ensuring efficient and secure data handling.

This module stores important information such as uploaded image details, preprocessing results, extracted features (optional), and final classification outputs. It may also maintain user-related data, logs, and system activity for future reference and analysis.

The module establishes a secure connection to the database using appropriate technologies (such as SQL or NoSQL databases) and ensures smooth communication between the backend and storage system. It performs operations like data insertion, retrieval, updating, and deletion as required.

Additionally, it includes validation and error-handling mechanisms to prevent data loss or corruption. Security measures such as authentication and access control can also be implemented to protect sensitive information.

Overall, this module ensures reliable data management, supports system scalability, and helps in maintaining records for monitoring and improving the deepfake detection system.

e. Feature Extraction Module

This module is responsible for extracting meaningful features from the preprocessed image using a hybrid deep learning approach. The system uses both EfficientNet and Xception.

• EfficientNet captures global features such as overall facial structure and patterns
• Xception extracts fine details and manipulation artifacts at pixel level

The outputs from both models are combined to form a strong feature representation. This hybrid approach improves detection capability by analyzing both high-level and low-level features present in the image.

zimage upload, processing and real-time prediction within a unified environment. The system is designed to ensure consistency, efficiency and scalability, making it suitable for practical applications in digital media verification.

Experimental results indicate that the hybrid approach achieves high accuracy and robustness in detecting deepfake images, even in challenging scenarios. By focusing on feature-based detection rather than manual inspection, the system provides a reliable and automated solution for identifying forged content. Overall, this work highlights the effectiveness of deep learning techniques in enhancing digital security and maintaining the authenticity of visual media.

IV CONCLUSION

This work presents a deep learning–based framework for detecting deepfake images using a hybrid model approach. The system integrates advanced feature extraction techniques with an efficient classification mechanism to identify manipulated images while maintaining high accuracy and reliability. By combining EfficientNet and Xception, the model captures both global image patterns and fine-grained manipulation artifacts, enabling effective detection of forged content.

The proposed architecture, implemented using a user-friendly interface, supports seamless image upload, preprocessing, feature extraction and real-time prediction within a unified environment. The detection process is designed to preserve image structure while analyzing inconsistencies, allowing accurate classification without manual intervention. This enables the system to be effectively used in applications such as digital media verification, social platforms and security systems.

Experimental evaluation indicates that the hybrid model achieves high detection accuracy with strong robustness against various types of deepfake manipulations. By emphasizing automated feature-based detection over traditional manual methods, the proposed approach demonstrates a scalable and practical solution for identifying fake images. Overall, this work highlights deep learning as a powerful technique for ensuring digital authenticity and strengthening cybersecurity in modern multimedia systems.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their project guide and faculty members of the Department of Computer Science and Engineering for their continuous support, valuable guidance, and encouragement throughout the course of this research work. Their expertise and direction played a vital role in the successful completion of this project.

The authors also extend their heartfelt thanks to their mentors, peers, and technical contributors who provided valuable suggestions, constructive feedback, and support at various stages of the project. Their insights significantly helped in enhancing the quality, performance, and overall effectiveness of the deepfake image detection system.

Additionally, the authors acknowledge the use of publicly available datasets, deep learning frameworks, and development tools that facilitated the implementation, training, and evaluation of the proposed hybrid model using EfficientNet and Xception. These resources were crucial in achieving accurate, efficient, and reliable results.

Finally, the authors express their gratitude to their institution for providing the necessary infrastructure and environment to carry out this work successfully.

REFERENCES

[1] H. Nguyen, F. Fang, J. Yamagishi, and I. Echizen, “Multi-task Learning for Detecting and Segmenting Manipulated Facial Images,” IEEE International Conference on Biometrics, pp. 1–8, 2019.

[2] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, “FaceForensics++: Learning to Detect Manipulated Facial Images,” IEEE International Conference on Computer Vision (ICCV), pp. 1–11, 2019.

[3] Francois Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258, 2017.

[4] Mingxing Tan and Quoc V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” International Conference on Machine Learning (ICML), pp. 6105–6114, 2019.

[5] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, “MesoNet: A Compact Facial Video Forgery Detection Network,” IEEE International Workshop on Information Forensics and Security (WIFS), pp. 1–7, 2018.

[6] H. Li, B. Li, S. Tan, and J. Huang, "Identification of Deepfake Videos by Detecting Eye Blinking," IEEE International Workshop on Information Forensics and Security (WIFS), pp. 1–7, 2018.
[7] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," Information Fusion, vol. 64, pp. 131–148, 2020.
[8] L. Verdoliva, "Media Forensics and DeepFakes: An Overview," IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 5, pp. 910–932, 2020.
[9] Y. Mirsky and W. Lee, "The Creation and Detection of Deepfakes: A Survey," ACM Computing Surveys, vol. 54, no. 1, pp. 1–41, 2021.
[10] I. Goodfellow et al., "Generative Adversarial Nets," Advances in Neural Information Processing Systems (NeurIPS), pp. 2672–2680, 2014.