iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** X **Month of publication:** October 2024

DOI: https://doi.org/10.22214/ijraset.2024.64450

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue X Oct 2024- Available at www.ijraset.com

Hybrid LSTM and Encoder-Decoder Architecture for Detection of Image and Video Forgeries

Uday Sai Kiran vangaru[1], Thribuvanesh Kurvagadda[2]
Department of Electronics and Communication Engineering, Marri Laxman Reddy Institute of Technology and Management

Abstract: ***This paper introduces a hybrid architecture combining Long Short-Term Memory (LSTM) networks and an encoder-decoder model for the detection and localization of image and video forgeries. The proposed system leverages resampling features and LSTM cells to identify manipulation patterns such as splicing and retouching in multimedia content. By utilizing a combination of spatial and temporal features, the model achieves high precision in detecting forged regions. Extensive testing on diverse datasets demonstrates the robustness of the proposed method.***
Keywords: Image Forgery Detection, Video Forgery Detection, LSTM, Encoder-Decoder, Deep Learning.

I. INTRODUCTION

With the proliferation of digital media and sophisticated editing tools, the alteration of images and videos has become more accessible and widespread. These manipulations can have severe consequences, ranging from misinformation to fraudulent activities. Thus, developing effective techniques for detecting and localizing such forgeries is crucial. This research focuses on a hybrid architecture that integrates LSTM networks with an encoder-decoder model to detect forged regions in both images and videos.

II. LITERATURE REVIEW

Various techniques have been proposed for image and video forgery detection. Traditional methods, such as copy-move detection and digital watermarking, are often limited by their dependency on specific features and lack of adaptability. Recent advancements in deep learning have introduced models like Convolutional Neural Networks (CNNs) for image analysis and Recurrent Neural Networks (RNNs) for temporal data. However, these methods often struggle with localization accuracy. The proposed hybrid model aims to address these limitations by combining spatial and temporal information for precise forgery localization.

III. PROPOSED METHODOLOGY

A. Hybrid Architecture

The hybrid architecture comprises two main components: an LSTM network for temporal feature extraction and an encoder-decoder model for spatial feature analysis. The LSTM network captures temporal dependencies in video sequences, while the encoder-decoder model processes spatial features to identify subtle manipulations in images.

B. Data Preparation

The model is trained on a diverse dataset containing tampered images and videos, labelled with ground-truth masks indicating forged regions. Data augmentation techniques, such as rotation and scaling, are applied to enhance model robustness.

C. Training and Evaluation

The model is trained using a combination of pixel-wise and binary cross-entropy loss functions to optimize both localization and classification performance. The evaluation metrics include precision, recall, F1-score, and Intersection over Union (IoU), which are used to assess the model's accuracy and reliability.

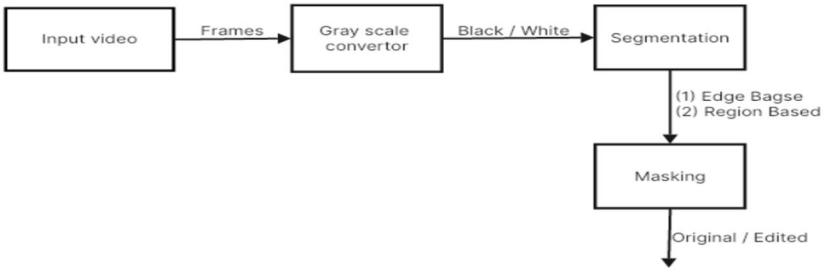


Fig. 1 Encoder

©IJRASET: All Rights are Reserved

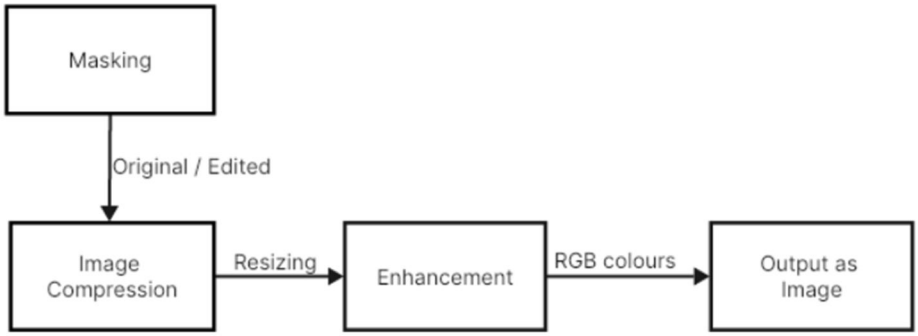


Fig. 2 Decoder

IV. RESULTS

The proposed method was evaluated on multiple datasets, including a large-scale image splicing dataset and a video forgery detection dataset. The results indicate that the hybrid model outperforms traditional CNN and RNN models in terms of precision and localization accuracy. The model achieved an average IoU score of 85% on image datasets and 82% on video datasets, demonstrating its effectiveness in various forgery scenarios.

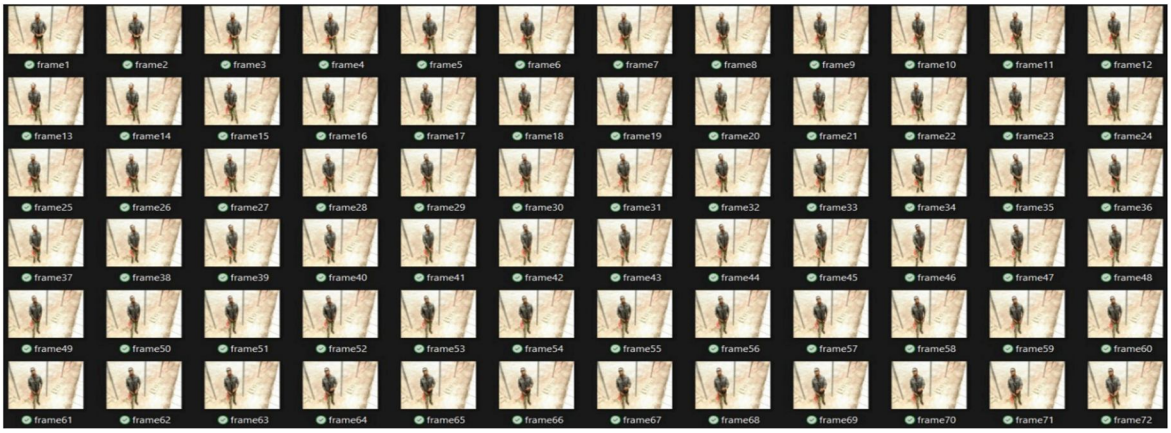


Fig. 3 Video Frames

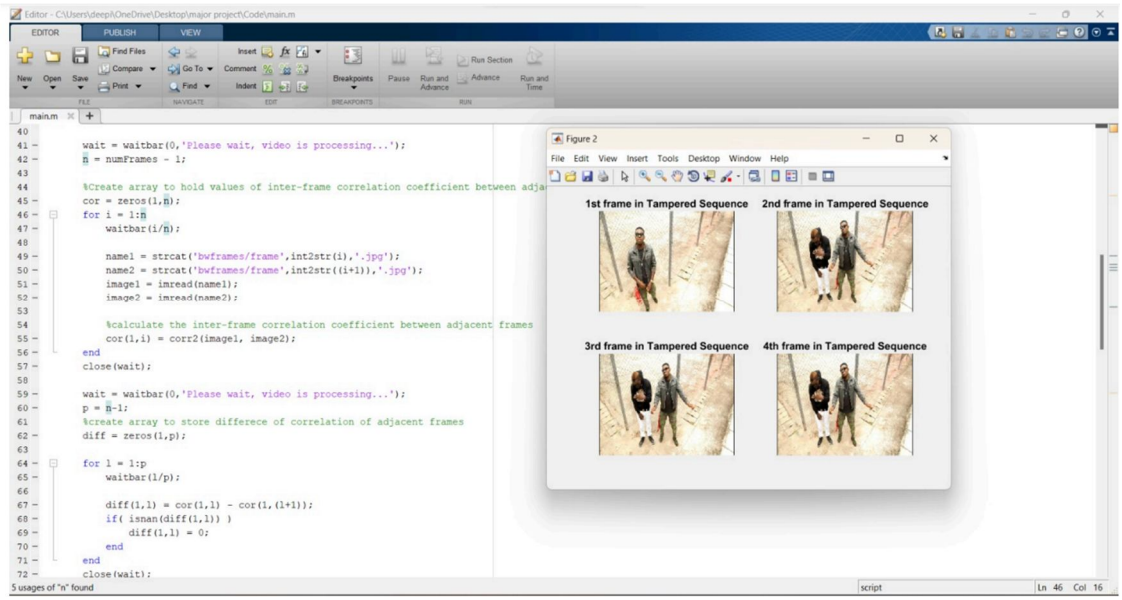


Fig. 4 Tampered video

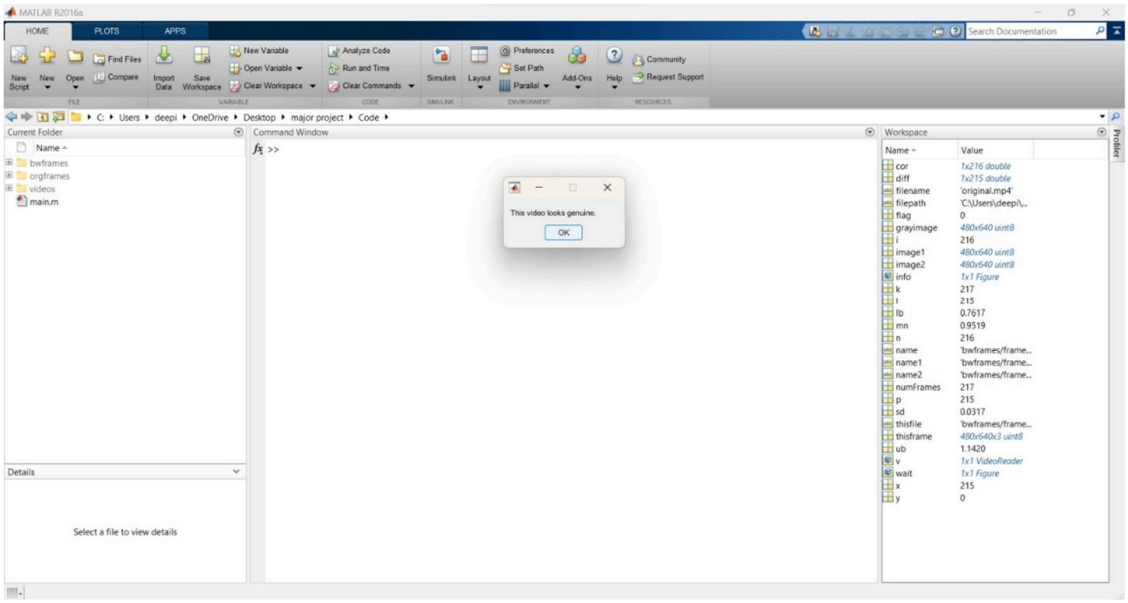


Fig. 5 Final Output

A. Discussion

The integration of LSTM networks with the encoder-decoder model allows for comprehensive analysis of both spatial and temporal features, which is critical for detecting complex forgeries. While the model shows promising results, further improvements can be made by incorporating additional data sources and refining the training process.

V. CONCLUSION

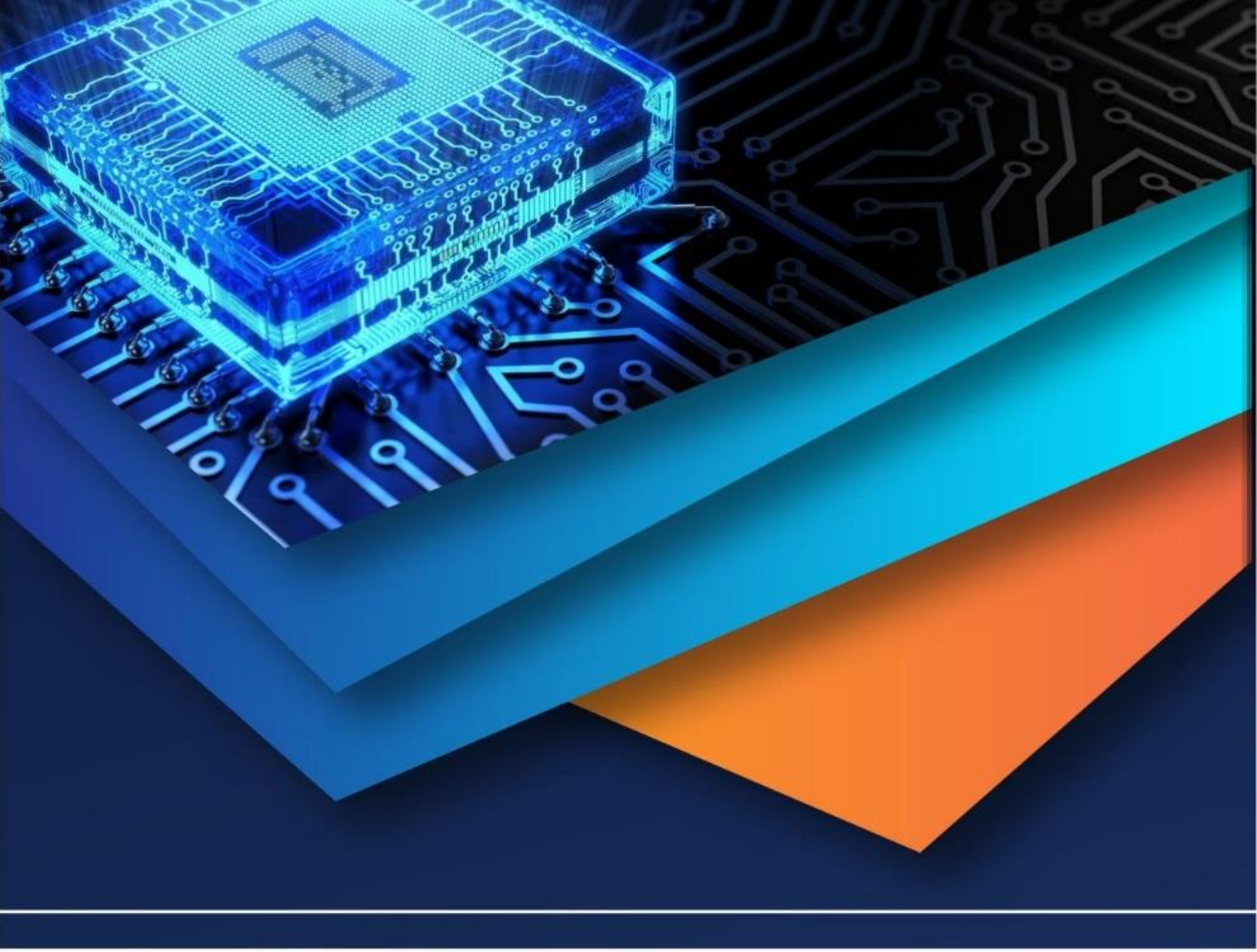
This paper presents a novel hybrid LSTM and encoder-decoder architecture for detecting image and video forgeries. The proposed system effectively localizes forged regions, offering a robust solution for multimedia forensics. Future work will focus on enhancing real-time detection capabilities and expanding the model to handle more diverse types of forgeries.

VI. ACKNOWLEDGMENT

We would like to express our sincere gratitude to all those who contributed to the development of our project on the *"Hybrid LSTM and Encoder-Decoder Architecture for Detection of Image and Video Forgeries"*. We also wish to acknowledge the contributions of our peers at MLRITM institute, whose constructive feedback and collaborative spirit enriched our work. Lastly, we thank our families and friends for their unwavering support and understanding during the course of this project. Their encouragement kept us motivated and focused. This research stands as a testament to the collective efforts of all involved, and we hope it contributes meaningfully to the field of forgery detection.

REFERENCES

[1] A.Gironi, M. Fontani, T. Bianchi, A. Piva, and M. Barni, “A video forensic technique for detection frame deletion and insertion,” IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2014, pp. 6226–6230.
[2] H. Ravi, A.V. Subramanyam, G. Gupta, and B. Avinash Kumar, “Compression noise-based video forgery detection,” IEEE International Conference on Image Processing, 2014, pp. 5352–5356.
[3] M. Chen, J. Fridrich, M. Goljan, and J. Lukás, “Determining image origin and integrity using sensor noise,” IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, pp. 74–90, March 2008.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)