IMAGE AUTHENTICITY VERIFICATION: A HYBRID APPROACH OF USING CAPSULE NETWORKS AND FREQUENCY DOMAIN ANALYSIS

Project Reference No.: *48S_BE_6210*

College : Bangalore Institute Of Technology, Bengaluru
Branch : Department Of Computer Science And Engineering
Guide(S) : Dr. Savitha. S. K
Student(S): Mr. Gurudatta B S
Ms. Deepthi S
Ms. Anagha Shree H J
Mr. Arbaz N

Keywords:

Capsule Networks, Frequency Domain Analysis, Fake Image Detection, Image Manipulation, Deepfake Detection, Support Vector Machines (SVM), Logistic Regression, Keras Neural Networks, Image Segmentation, High-pass Filters, Machine Learning Classification, Cross-validation, Spatial Hierarchy, Bilateral Frequency Filtering, Image Pre-processing, Real-time Analysis, User-friendly Application, Model Accuracy, Confidence Score Comparison, Digital Forensics, Content Integrity, Manipulated Media Detection, Structural Inconsistencies, Hybrid Detection Models, Ethical AI Deployment.

Introduction:

The project involves developing an advanced system to detect fake images, which combines the strengths of Capsule Networks, frequency domain analysis, and machine learning techniques. By leveraging these cutting-edge technologies, the system can effectively identify anomalies and artifacts that indicate manipulation, and classify images as authentic or fake with high accuracy.

The system's architecture is designed to integrate spatial domain analysis through Capsule Networks with frequency domain techniques, allowing for a comprehensive analysis of image features. The extracted features are then fed into machine learning algorithms, including Support Vector Machines, Logistic Regression, and Keras-based

neural networks, which are trained to detect subtle patterns and anomalies that may indicate manipulation.

The resulting system is a powerful tool for detecting fake images, with potential applications in security, journalism, and social media. By providing a robust and accurate means of identifying manipulated media, the system can help to mitigate the risks associated with deepfakes and promote a safer and more trustworthy online environment

Objectives

- To build a robust segmentation model to isolate specific regions of interest,such as faces, for enhanced scrutiny and analysis in manipulated images.
- To design a high-accuracy classification system that can reliablydistinguish between authentic and fake images using enhanced imagequality and segmentation data.
- To implement a scalable solution that integrates image enhancement, segmentation and classification techniques to improve the overall efficiencyof fake image detection.

Methodology:

The methodology for this project involves several systematic steps:

- Pre-processing: Apply image pre-processing techniques such as resizingand normalization to prepare the images for analysis.Feature Extraction: Utilize Capsule Networks to capture spatial hierarchiesin images, enhancing the detection of subtle manipulations.
- ImplementFrequency Analysis using bilateral high-pass filters to identify discrepanciesin image frequency components that may indicate tampering.
- Model Training: Train machine learning models using labelled datasets(real vs. fake) to learn distinguishing features. Employ techniques such ascross-validation to ensure model robustness and prevent over fitting.
- Classification: Develop a classification algorithm that combines outputsfrom both Capsule Networks and frequency analysis to make finalpredictions about

image authenticity. Use metrics such as accuracy,precision, recall, and F1-score to evaluate model performance.

- Deployment: Integrate the detection system into a user-friendly applicationthat can be used on social media platforms for real-time analysis.

Expected Outcome of the project

The anticipated outcomes of this project include:

- High Accuracy Detection: Achieving a detection accuracy exceeding 90%for identifying real vs. fake images through advanced machine learningtechniques.
- User Engagement: Development of an application that allows users toupload images for analysis, receiving immediate feedback on authenticity.
- Educational Resources: Providing resources and information aboutdeepfakes, helping users understand how to identify manipulated contentthemselves.
- Industry Collaboration: Establishing partnerships with social mediaplatforms to implement detection tools that enhance content integrity.

Result and Conclusion:

The project successfully implemented a system that processes images through three models—Capsule Network, Bi-HPF, and Frequency-Domain Analysis, followed by a comparison of confidence scores. The model with the highest confidence score determines the final output, which is sent back to the frontend and displayed on the website. This endto-end system provides a seamless workflow for analyzing uploaded images and generating results in real time**.**

In conclusion, The comparative analysis revealed that each method brought distinct advantages to the table. Capsule Networks were particularly effective in analyzing spatial relationships. Image Authenticity Verification using Deep Neural Networks making them adept at detecting structural inconsistencies. The bilateral high-frequency filter approach was unparalleled in its ability to expose subtle artifacts in the frequency domain, while the Keras prediction model excelled in overall prediction accuracy and computational efficiency

Future Scope:

Future development of this deepfake detection project will focus on several key areas to enhance its impact and address the evolving nature of digital manipulation. This includes integrating video deepfake detection to identify frame-by-frame inconsistencies, creating hybrid models combining spatial and frequency-based approaches for superior accuracy, and ensuring scalability for real-time applications in areas like live streaming. Further priorities include improving dataset diversity to increase generalization over different types and contexts of deepfakes, adapting the framework for cross-domain applications such as audio and text manipulation detection, developing ethical and legal frameworks for responsible AI deployment, and making interfaces more user-friendly to access the framework for a wider popuplation.