Paper Title: Deep Fake Detection: A journey towards Authenticity

Paper link : [Deep Fake Detection: A journey towards Authenticity]

1.Summary

1.1 Motivation

Maintaining confidence in digital material requires taking precautions against the growing threat of deepfakes. Deepfake detection and debunking represents a cutting-edge technological frontier, stretching the limits of artificial intelligence and image processing. Deep fakes are a danger to information authenticity, thus finding a solution to this problem is essential to shielding people and society from false information.It is morally necessary to develop strong deep fake detection techniques in order to promote ethical technology use and stop malicious alteration of visual material.

1.2 Contribution

In order to recognize and lessen the growing threat of deep fake material, "Deep Fake Detection: A Journey towards Authenticity" examines cutting-edge methods and tools. The essay explores how digital authenticity is changing, including everything from forensic examination to sophisticated machine learning algorithms. The objective is to strengthen the barriers against misleading media manipulation by comprehending and tackling the difficulties presented by deep fakes.

1.3 Methodology

Using sophisticated machine learning algorithms to examine minute facial and audio discrepancies, the "Deep Fake Detection: A Journey towards Authenticity" technique accurately identifies deepfake content for improved authenticity verification.

- -> Model Selection
- i) Utilized pre-trained Convolutional Neural Network (CNN) Models
- ii) InceptionResNetV2, DenseNet201, ResNet15V2 and InceptionV3
- -> Model Performance Evaluation
- i)Validation Accuracy and Loss Results
- ii) InceptionResNetV2: Highest Validation Accuracy of 99.87%

1.4 Conclusion

In summary, the thesis introduces a patch-based technique for deepfake localization, contributing to the field of multimedia forensics by enhancing the accuracy and efficiency of deepfake detection through localized analysis and pattern recognition.

2.Limitation

2.1 First Limitation

It could be difficult for the suggested patch-based deepfake localization method to generalize well across a variety of manipulation techniques. Deepfakes that use unique or diversified manipulation techniques that are not sufficiently represented in the training data may be difficult for the model to identify and locate if it is trained largely on certain kinds of visual artifacts or manipulation patterns. The model's resilience in real-world situations where novel manipulation methods could surface could be impacted by this constraint.

2.2 Second Limitation

The caliber and volume of training data utilized have a significant impact on how well the patch-based deepfake localization technique performs. A training dataset that is too small, homogeneous, or does not include a representative sample of deepfake scenarios might make it difficult for the model to accurately depict the complexity and variety seen in real-world altered pictures. Furthermore, the model may become less dependable in real-world applications if it is subjected to changes in illumination, resolutions, or compression artifacts that were not sufficiently covered during training.

3. Synthesis

The suggested technique, called "Patch-Based Deepfake Localization," uses visual artifact analysis to reveal areas that have been altered in photos. The system uses a patch-based methodology to pinpoint and emphasize the precise regions impacted by deepfake manipulation. Through close examination of visual signals and abnormalities, the method reveals the existence of digital modifications, allowing accurate localization of modified material. By identifying areas containing questionable artifacts, this method improves the interpretability of deepfake detection. The synthesis provides an effective way to recognize and comprehend how deepfake modification affects visual material.