MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

A Project Report on

Deep Fake Detection Using Intrinsic Signature

Submitted by,

Sanika Mahendra Thakare (Exam Seat No. 202101070120)

Guided by,

Dr.Vipin Yadav

**A Report submitted to MIT Academy of Engineering, Alandi(D), Pune,
An Autonomous Institute Affiliated to Savitribai Phule Pune University
in partial fulfillment of the requirements of**

THIRD YEAR BACHELOR OF TECHNOLOGY in

Electronics & Telecommunication Engg.

School of Electronics & Telecommunication Engg.

MIT Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

Alandi (D), Pune – 412105

(2023–2024)

MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

CERTIFICATE

It is hereby certified that the work which is being presented in the Third Year Project Design Report entitled **"Deep Fake Detection Using Intrinsic Signature"**, in partial fulfillment of the requirements for the award of the Bachelor of Technology in Electronics & Telecommunication Engg.and submitted to the **School of Electronics & Telecommunication Engg.** of **MIT Academy of Engineering, Alandi(D), Pune, Affiliated to Savitribai Phule Pune University (SPPU), Pune**, is an authentic record of work carried out during Academic Year **2023–2024** Semester **V**, under the supervision of **Dr.Vipin Yadav**, **School of Electronics & Telecommunication Engg.**

Sanika Mahendra Thakare (Exam Seat No. 202101070120)

Dr.Vipin Yadav
Project Advisor

Dr.Gayatri Ambadkar
Project Coordinator

Dr. Dipti Y. Sakhare
Dean

External Examiner

DECLARATION

We the undersigned solemnly declare that the project report is based on our own work carried out during the course of our study under the supervision of **Dr.Vipin Yadav**.

We assert the statements made and conclusions drawn are an outcome of our project work. We further certify that

1. The work contained in the report is original and has been done by us under the general supervision of our supervisor.
2. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this Institute/University or any other Institute/University of India or abroad.
3. We have followed the guidelines provided by the Institute in writing the report.
4. Whenever we have used materials (data, theoretical analysis, and text) from other sources, we have given due credit to them in the text of the report and giving their details in the references.

Sanika Mahendra Thakare (Exam Seat No. 202101070120)

Abstract

The ”Deep Fake Detection using Intrinsic Signatures” is a method use for spotting fake images/videos by analyzing their basic and natural attributes. This method is use to aim classify manipulated content from genuine media without depending on external data. By identifying features within the video/image, this method improves the accuracy and quality of deep fake identification. This approach holds promise in addressing the increasing concern over misleading digital media and contributes to the development of more reliable tools for identifying manipulated content.

Acknowledgment

We want to express our gratitude towards our respected project guide Prof.Vipin Yadav for his constant encouragement and valuable guidance during the completion of this project work. We also want to express our gratitude towards respected School Dean Dr. Dipti Sakhare for her continuous encouragement.We would be failing in our duty if we do not thank all the other staff and faculty members for their experienced advice and evergreen co-operation

202101070206 Shaikh Rumman Rafique

202101060015 Akshata Sharad Thorkar

202101070120 Sanika Mahendra Thakare

202101070202 Atharv Satish Mhatre

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

With the rise of fake videos, it's becoming harder to tell what's real online. Common detection methods using outside clues are not as good anymore.The "Deep Fake Detection using "Intrinsic Signatures" project wants to find a better way. Instead of looking at outside info,it focuses on the stuff inside the video itself. This project aims to create a new and strong method to spot fake videos,even when the usual clues don't work well. It's all about making sure we can trust what we see online and stop fake videos from fooling us.The "Deep Fake Detection using Intrinsic Signatures" project addresses the growing challenge posed by fake videos in the online landscape. As technology advances, distinguishing between authentic and manipulated content has become increasingly difficult. Traditional detection methods that rely on external clues are becoming less reliable. Hence, there is a critical need for innovative approaches that focus on the intrinsic characteristics of the videos themselves.

1.2 Motivation

"Deep Fake Detection using Intrinsic Signatures" aims to face misleading digital content by focusing on intrinsic characteristics within media.This approach provides accurate and reliable detection techniques,maintaining trust in digital media and

combating manipulated content spread.The reason, behind this effort is based on the need to maintain trust in media. As fake content becomes more prevalent its impact goes beyond spreading misinformation. It poses a threat to society by eroding trust in visual information, which can create an atmosphere of doubt and skepticism. This project strives to be a measure that assures individuals they can consume content knowing that advanced detection methods are, in use to protect against deception.

1.3 Project Idea

The concept is to tell if a video is fake by just looking at the video itself, without needing to know where it came from or when it was made. That's the idea behind "Deep Fake Detection using Intrinsic Signatures."It's like finding secret clues inside the video that only real videos have, and fake ones can't copy easily.By studying these hidden clues, we can make a clever way to know if a video is real or fake, which will help us not get tricked by fake videos online.Deep fake is computer program use to create aur manipulate the facial attributes of person making it realistic, where one cannot distinguish between real and fake.The manipulated media is use specially for miscommunication or mallicious activities.

1.4 Proposed Solution

Our method demonstrates simple and interpretable features that can be effectively used for deepfake detection, offering a compelling alternative to complex deep learning models. By utilizing frequency domain characteristics such as wavelet transforms which converts image to 2D spectrum and converting these into 1D spectra through azimuthal averaging, the method achieves high classification accuracy.For classification we have used various machine learning algorithms.By focusing on frequency components, our approach provides a robust way to detect fake images without requiring large amounts of labeled data. The intrinsic signatures in the frequency domain offer a reliable means of identifying fake content, even when the images are highly realistic.

Chapter 2

Literature Review

Table 2.1: Literature review

Author Name	Journal	Title	Method	Observation
Asad Malik [2022]	IEEE Access	Deep Fake detection for human face images and videos: A survey	Deep learning, Deep Fake, CNNs, GANs	Detailed overview of Deep Fake, a GAN-based technology, discussing its benefits and threats models.
Ashwin Swaminathan [2008]	IEEE transactions on information forensics and security	Digital image forensics via intrinsic fingerprints	Image acquisition forensics, intrinsic fingerprints.	Coefficients of linear shift-invariant approximation using blind deconvolution
Praveen Gupta and Rashid Sheikh [2021]	International Research Journal of Engineering and Technology,	Deep fake detection: Survey of facial manipulation detection solutions	Deep Learning, Neural Networks, Deep Fakes.	Analyzed several such states of the art neural networks and compare them against each other, to find an optimal solution.

Chapter 3

Problem Definition and Scope

3.1 Problem statement

Classify real and computer generated image using intrinsic signature and extend the findings for Deep fake detection

3.2 Goals and Objectives

- Study of different intrinsic signature
- Develop ML/DL based model to classify real and computer generated image using intrinsic signature.
- Extend the findings for deep fake detection.

3.3 Scope and Major Constraints

Scope:

1. Classification of real and computer generated images.
2. Identification of deep fake using intrinsic signature

Major Constraints:

1. Limited Dataset Availability: Acquiring a diverse and comprehensive dataset containing a sufficient number of both real and computer-generated images may be challenging.
2. Computational Resources: Developing and training ML/DL models, especially for deep fake detection, can be computationally intensive. Limited computational resources may constrain the model's complexity and training duration.
3. Ethical Considerations: The model's ability to generalize to various types of computer-generated images and adapt to evolving deep fake techniques may pose challenges.
4. Dynamic Nature of Deep Fakes: The landscape of deep fakes is continuously evolving, and the model may need periodic updates to remain effective against new and sophisticated techniques.

3.4 Software Requirements

- Jupyter Notebook
- Python 3.10
- Python Libraries:OpenCV,OS,PyWavelets,Pickle

3.5 Expected Outcomes

1. Classification of real and computer generated image by ML based model using intrinsic signature.
2. Identification of deep fake using intrinsic signature

Chapter 4

System Requirement Specification

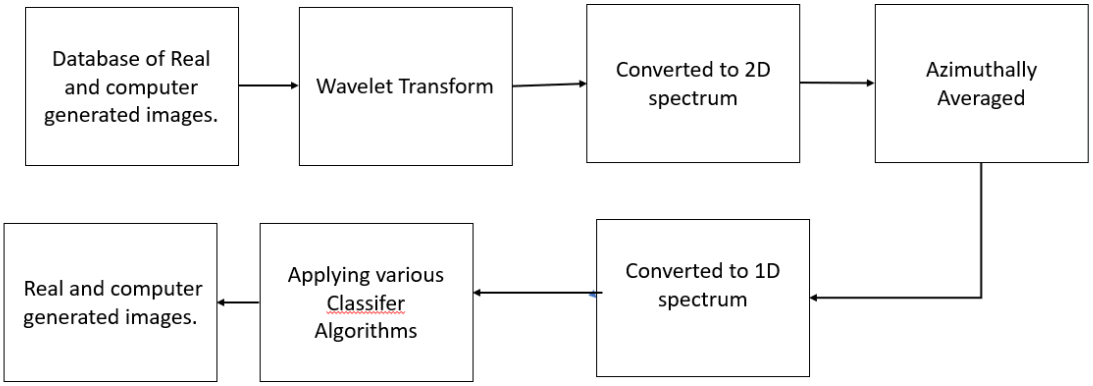

Figure 4.1: Block Diagram

This is the proposed block diagram of the system.This involves collection of database consist of real and computer generated image then convert it into 2D by using wavelet transform again convert this into 1D by doing azimuthal averaging.then for classification of real and computer generated image,we have applied different machine learning algorithms such as ANN,Random forest,SVM,Logistic regression and KNN

4.1 Software Requirements

Software Specification:

1. Jupyter Notebook

2. Python 3.10

3. Python Libraries:OpenCV,OS,PyWavelets,Pickle

4.2 Project Planning

Phase 1: Data Collection and Preparation

1. Collecting Diverse Dataset:

- Gather a diverse dataset comprising both real and computer-generated images to encompass various deep fake scenarios.

2. Data Enhancement:

- Apply preprocessing techniques to standardize and enhance the quality of the collected dataset.

Phase 2: Model Development

3. Feature Extraction and Intrinsic Signature Analysis:

- Implement algorithms to extract intrinsic signatures from images, focusing on features like frequency by using wavelet tranform.

4. Model Selection:

- Choose an appropriate AI/ML model for deep fake detection based on intrinsic signatures.Consider architectures like Artificial Neural Networks (ANNs) or other deep learning models optimized for image analysis.

Phase 3: Training and Testing

5. Training the Model:

-Train the selected model using the prepared dataset, emphasizing the recognition of intrinsic signatures associated with deep fake content.Optimize training parameters to ensure the model's effectiveness.

6. Validation and Testing:

- Validate the model using a dedicated dataset to assess its accuracy and generalization to unseen deep fake variations.

Phase 4: Model Refinement

7. Feedback Integration:

- Collect feedback from initial testing and real-world scenarios to identify false positives/negatives and areas for improvement.Incorporate user insights to enhance the model's performance.

By following this roadmap, the project aims to develop an effective solution for Deep fake detection using intrinsic signature.

Chapter 5

Proposed Methodology

5.1 System Architecture

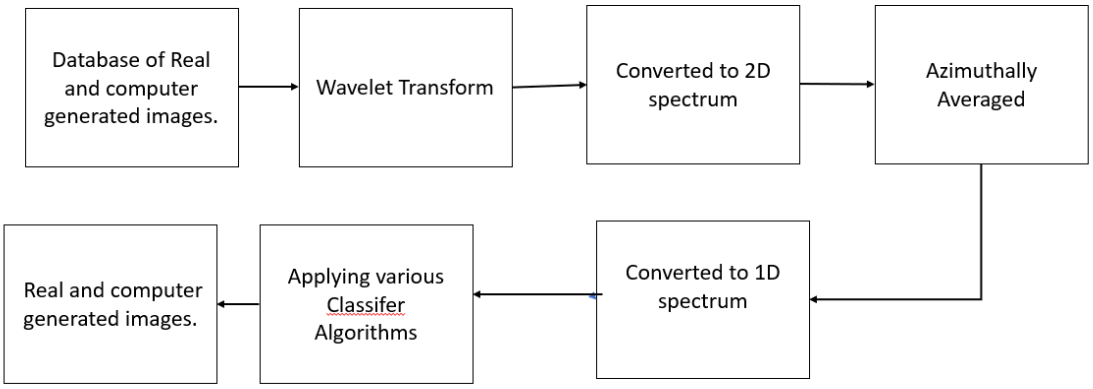

Figure 5.1: System Architecture

This is the proposed system architecture of the system.This involves collection of database consist of real and computer generated image then convert it into 2D by using wavelet transform again convert this into 1D by doing azimuthal averaging.then for classification of real and computer generated image,we have applied different machine learning algorithms such as ANN,Random forest,SVM,Logistic regression and KNN

5.2 Approach

1. Data Collection: Referring the database obtained from kaggle which are of two types real and computer generated.

2. Analyzation of Data: Reading and analyzing the attribute of obtain image.

3. Wavelet Transform: Apply wavelet transform on input images in order to convert them into 2D spectrum.

4. Azimuthal averaging: Converting the obtained 2-dimensional data into 1-dimensional data by this method.

5. machine learning algorithms: we have applied different machine learning algorithms such as ANN,Random forest,SVM,Logistic regression and KNN for the classification of real and computer generated images

6. Model Training: Train the model to recognize real and computer generated image.

7. Post-processing: Refine model output with post-processing.

Chapter 6

Conclusion

6.1 Conclusion

The study successfully demonstrates that simple and interpretable features can be effectively used for deepfake detection, offering a compelling alternative to complex deep learning models. By utilizing basic image statistics and frequency domain characteristics such as wavelet transforms, and converting these into 1D spectra through azimuthal averaging, the method achieves high classification accuracy. Specifically, classifiers like SVM, LR,ANN,KNN,Random Forest to achieve accuracies of 92 and 83,92,93.03,93 respectively, in distinguishing real from computer-generated images. This approach not only proves to be computationally efficient and robust against common post-processing techniques but also maintains high interpretability. Thus, it provides a practical and reliable solution for real-time deepfake detection in resource-constrained environments. do this paragraph shorter than this

6.2 Future Scope

Semester VI: Extend the finding to detect deep fake using intrinsic signature.

References

AlShariah, N. M., Khader, A., & Saudagar, J. (2019). Detecting fake images on social media using machine learning. *International Journal of Advanced Computer Science and Applications*, *10* (12), 170–176.

De Rezende, E. R., Ruppert, G. C., Theophilo, A., Tokuda, E. K., & Carvalho, T. (2018). Exposing computer generated images by using deep convolutional neural networks. *Signal Processing: Image Communication*, *66* , 113–126.

Dirik, A. E., Bayram, S., Sencar, H. T., & Memon, N. (2007). New features to identify computer generated images. In *2007 ieee international conference on image processing* (Vol. 4, pp. IV–433).

Durall, R., Keuper, M., Pfreundt, F.-J., & Keuper, J. (2019). Unmasking deepfakes with simple features. *arXiv preprint arXiv:1911.00686* .

Guarnera, L., Giudice, O., & Battiato, S. (2020). Deepfake detection by analyzing convolutional traces. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition workshops* (pp. 666–667).

Guo, K., & Wang, R. (2011). A method for identifying computer images and real images. In *2011 international conference on electronics, communications and control (icecc)* (pp. 638–641).

Kumar, M., Sharma, H. K., et al. (2023). A gan-based model of deepfake detection in social media. *Procedia Computer Science*, *218* , 2153–2162.

Li, H., Li, B., Tan, S., & Huang, J. (2020). Identification of deep network generated images using disparities in color components. *Signal Processing* , *174* , 107616.

Malik, A., Kuribayashi, M., Abdullahi, S. M., & Khan, A. N. (2022). Deepfake detection for human face images and videos: A survey. *Ieee Access*, *10* , 18757–18775.

Nguyen, H. H., Yamagishi, J., & Echizen, I. (2019). Use of a capsule network to detect fake images and videos. *arXiv preprint arXiv:1910.12467* .

Ni, X., Chen, L., Yuan, L., Wu, G., & Yao, Y. (2019). An evaluation of deep learning-based computer generated image detection approaches. *IEEE Access*, *7* , 130830–130840.

Swaminathan, A., Wu, M., & Liu, K. R. (2008). Digital image forensics via intrinsic fingerprints. *IEEE transactions on information forensics and security*, *3* (1), 101–117.

Sychandran, C., & Shreelekshmi, R. (2023). Performance comparison of deep learning models for computer generated image detection. In *2023 international conference on control, communication and computing (iccc)* (pp. 1–5).

Tariang, D. B., Senguptab, P., Roy, A., Chakraborty, R. S., & Naskar, R. (2019). Classification of computer generated and natural images based on efficient deep convolutional recurrent attention model. In *Cvpr workshops* (pp. 146–152).

Yu, N., Davis, L. S., & Fritz, M. (2019). Attributing fake images to gans: Learning and analyzing gan fingerprints. In *Proceedings of the ieee/cvf international conference on computer vision* (pp. 7556–7566).