PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)
100-ft Ring Road, Bengaluru – 560 085, Karnataka, India

Capstone Project Report (Phase-1)
on

Deep fake Image Detection using GANception

Submitted by

SHRINIVAS S PATIL - (PES1PG23CA137)

Nov 2024 – Feb 2025

under the guidance of

Guide Details

Dr. Vignesh Ramamoorthy H
Associate Professor
Department of Computer Applications,
PESU, Bengaluru – 560085

FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER APPLICATIONS
PROGRAM – MASTER OF COMPUTER APPLICATIONS

CERTIFICATE

This is to certify that the project entitled

Deep Fake Image Detection Using GANception

is a bonafide work carried out by

SHRINIVAS S PATIL-PES1PG23CA137

in partial fulfillment for the completion of Capstone Project, Phase-1 work in the Program of Study MCA under rules and regulations of PES University, Bengaluru during the period Nov. 2024 – Feb 2025. The project report has been approved as it satisfies the academic requirements of 3rd semester MCA.

Signature with date
Guide
Dr.Vignesh Ramamoorthy H
Associate Professor
Department of Computer Applications,
PES University, Bengaluru - 560085

Signature with date & Seal
Chairperson
Dr. Veena S
Department of Computer Applications,
PES University, Bengaluru - 560085

DECLARATION

I, **SHRINIVAS S PATIL,** bearing **PES1PG23CA137** hereby declare that the Capstone project phase-1 entitled, ***Deep Fake Image Detection Using GANception,*** is an original work done by me under the guidance of **Dr.Vignesh Ramamoorthy H,** Designation, PES University, and is being submitted in partial fulfillment of the requirements for completion of 3rd Semester course in the Program of Study **MCA**. All corrections/suggestions indicated for internal assessment have been incorporated in the report.

PLACE:

DATE:

SHRINIVAS S PATIL

ABSTRACT

This application focuses on developing a sophisticated deep fake image detection system using deep neural networks (DNNs). Leveraging advanced machine learning algorithms, the system aims to accurately differentiate between real and manipulated images across various platforms, especially in social media contexts where deep fakes pose significant risks.

GAN-based detection techniques play a crucial role by analyzing inherent artifacts and inconsistencies that are often present in deep fake images. These artifacts, introduced during the generative process, include abnormal textures, unnatural facial expressions, and discrepancies in lighting or shadow consistency. By detecting such subtle cues, the system enhances its ability to flag manipulated content effectively.

Table of Contents

ABSTRACT

Chapter 1 Introduction………………………………………………………………………………1

1.1 Project Description ……………………………………………………………………………1

1.2 Problem Definition ……………………………………………………………………………1

1.3 Proposed Solution ……………………………………………………………………………2

1.4 Purpose ………………………………………………………………………………………2

1.5 Scope …………………………………………………………………………………………3

Chapter 2 Literature Survey ……………………………………………………………………4

2.1 Domain Study ………………………………………………………………………………4

2.2 Related Work. …………………………………………………………………………………4

Chapter 3 Hardware and Software Requirements ……………………………………………9

3.1 Hardware Requirements ………………………………………………………………………9

3.2 Software Requirements ………………………………………………………………………9

Chapter 4 Software Requirement Specification ……………………………………………11

4.1 Functional Requirements ……………………………………………………………………11

4.2 Non-Functional Requirements ………………………………………………………………12

Chapter 5 System Design ………………………………………………………………………14

5.1 Architecture Diagram …………………………………………………………………………14

5.2 Context Flow Diagram …………………………………………………………………………16

Chapter 6 Detailed Design………………………………………………………………………17

6.1 Use Case Diagram ……………………………………………………………………………17

6.2 Sequence Diagram ……………………………………………………………………………19

6.3 Activity Diagram………………………………………………………………………………20

6.4 Database Design ……………………………………………………………………………21

Chapter 7 Implementation

7.1 Screen Shot ……………………………………………………………………………………22

Appendix A and BIBLIOGRAPHY…………………………………………………………28

CHAPTER 1

1. INTRODUCTION

1.1. Project Description

Deepfake technology has seen rapid advancements, leveraging deep learning techniques to create highly realistic fake images and videos. While this innovation has revolutionized fields like entertainment, augmented reality, and digital content creation, it also raises serious ethical and security concerns. The widespread misuse of deepfake technology has led to issues such as misinformation, identity theft, and digital fraud, making it increasingly difficult to differentiate between authentic and manipulated media. As deepfake algorithms become more sophisticated, traditional detection methods struggle to keep up, necessitating the development of more robust and intelligent solutions.

This project focuses on Deepfake Image Detection using a GAN-based approach termed "GANception." The system leverages the power of GANs to learn and identify subtle artifacts present in deepfake images, which are often imperceptible to the human eye. By training on a dataset containing both real and fake images, the model learns to extract deep features that differentiate authentic images from AI-generated ones. The detection framework aims to enhance security in digital media by providing an efficient and accurate method for identifying deepfake content, contributing to the fight against digital manipulation and online misinformation.

1.2 Project Definition

This project focuses on developing a Deepfake Image Detection System using Generative Adversarial Networks (GANs) to accurately differentiate between real and AI-generated (fake) images. The model is trained on a dataset containing both real and fake images to learn deep visual patterns and artifacts unique to deep fakes. By leveraging GANs, the system aims to provide a robust and efficient method for detecting deepfake images, addressing concerns related to misinformation, identity fraud, and digital manipulation in various domains, including media, cybersecurity, and forensic analysis.

1.3 Proposed solution

The proposed solution focuses on creating a scalable, adaptable, and efficient Deepfake image detection system using advanced deep learning techniques. The system is designed to address key challenges such as real-time detection, batch processing, confidence scoring, and adaptability to evolving Deepfake generation methods.

By leveraging GAN-based models, the system effectively detects subtle artifacts introduced by synthetic image generation, including texture inconsistencies, unnatural facial asymmetries, and illumination anomalies. A robust preprocessing pipeline ensures accurate feature extraction, while real-time inference allows instant detection of fake images from video streams, social media uploads, and forensic datasets.

1.4 Purpose

The primary purpose of this project is to design and develop a reliable, efficient, and scalable Deepfake image detection system, addressing the growing concerns of misinformation, privacy invasion, and ethical misuse of AI-generated content. As Deepfake technology becomes increasingly sophisticated, distinguishing between real and manipulated images has become a critical challenge in digital forensics, media authentication, and cybersecurity.

This project leverages Generative Adversarial Networks (GANs) to analyze and detect Deepfake artifacts by identifying patterns, inconsistencies, and synthetic modifications present in manipulated images. The system will incorporate real-time detection, batch image processing, and detailed forensic reporting, providing users with confidence scores and visualization of suspicious areas. By employing advanced machine learning techniques and continuous model improvements, this solution aims to enhance trust, security, and authenticity in digital media, ensuring robust protection against Deepfake-based threats

1.5 Scope

The scope of this project encompasses the development of a robust and scalable Deepfake detection system capable of analyzing and identifying manipulated images with high accuracy. The system is designed to handle both real-time image detection and batch processing, ensuring efficiency in various use cases such as social media verification, forensic analysis, and cybersecurity applications.

Additionally, the model is built to adapt to evolving Deepfake generation techniques, leveraging GAN-based artifact detection to recognize synthetic patterns in manipulated images. The system will feature detailed forensic reporting, providing insights into confidence scores, suspected regions of manipulation, and statistical analysis of anomalies. This ensures that the detection system remains reliable, interpretable, and future-proof, addressing emerging threats in AI-generated media.

2 LITERATURE SURVEY

2.1 Domain Survey

The existing systems for deepfake detection are largely reliant on traditional CNN-based models. While these models have demonstrated effectiveness in detecting common types of deepfake artifacts, such as those generated by well-known techniques like FaceSwap or DeepFaceLab, they encounter limitations when faced with novel or unseen deepfake generation methods. These systems are often trained on a specific dataset and therefore struggle to generalize to new and evolving manipulation techniques.

2.2 Related Work

This involves the study of research papers and journals. The literature survey is completed by considering the following research papers.

1. Deepfake Detection using GAN Discriminators

   Publisher: IEEE (18 October 2021)

   Author: Sai Ashrith Aduwala, Manish Arigala

   SUMMARY:

   The abstract discusses the challenges posed by deepfake videos, which use advanced deep learning techniques to manipulate videos by replacing a person's features with another's.

   These manipulated videos can be used maliciously, creating a need for reliable detection methods. While Generative Adversarial Networks (GANs) are often used to create deepfakes, this work investigates their potential for detection by utilizing their discriminators.

2. A GAN-Based Model of Deepfake Detection in Social Media

Publisher: ScienceDirect (31 January 2023)

Author: Preeti , Manoj Kumar, Hitesh Kumar Sharma

SUMMARY:

The paper explores methods for implementing deepfakes, including deep convolution-based GAN models, and discusses techniques for both manipulation and detection of deepfakes.

The abstract highlights the emergence of deepfake technology, which uses Generative Adversarial Networks (GANs) to seamlessly swap identities between individuals in videos or images. The accessibility of large public datasets and advanced deep learning tools has facilitated the creation of realistic fake content, posing societal challenges.

3. Hybrid Deep Learning Model Based on GAN and RESNET for DEEPFAKE

Publisher: IEEE (19 JUNE 2024)

Author: SOHA SAFWAT, AYAT MAHMOUD, FARID ALI

SUMMARY:

This study introduces a novel hybrid deep learning model that combines the generative capabilities of Generative Adversarial Networks (GANs) with the discriminative strengths of Residual Neural Network (RESNET) architecture to detect fake faces.

By integrating these two approaches, the proposed model effectively distinguishes between real and artificially generated faces. A comparative analysis was conducted against pre-trained models such as VGG16 and RESNET50, and the results demonstrated the hybrid model's superior performance.

4. ProActive DeepFake Detection using GAN-based Visible Watermarking

Publisher: ACM Digital Library (12 September 2024)

Authors: Aakash Varma Nadimpalli, Ajita Rattani

SUMMARY:

This study addresses the limitations of passive DeepFake detection methods, which act as ex-post forensic countermeasures and fail to prevent the initial spread of disinformation. The authors propose a novel proactive defense mechanism using GAN-based visible watermarking to combat DeepFake creation.

By incorporating a reconstructive regularization term into the GAN's loss function, the proposed method embeds unique visible watermarks into specific locations of generated fake images. These watermarks make the manipulated images easily detectable by both human observers and state-of-the-art (SOTA) DeepFake detectors.

5. TITLE: Fighting Deepfakes by Detecting GAN DCT Anomalies

Publisher: MDPI (30 July 2021)

Authors: Oliver Giudice, Luca Guarnera, Sebastiano Battiato

SUMMARY:

This study proposes a novel pipeline for detecting Deepfakes by identifying GAN-specific frequencies (GSF), which serve as unique fingerprints of different generative architectures.

The method addresses the lack of generalizability and explainability in existing Deepfake detection algorithms. By analyzing ad-hoc frequencies left by GAN engines, the proposed approach utilizes the Discrete Cosine Transform (DCT) to detect anomalous frequencies.

6. Advancing GAN Deepfake Detection: Mixed Datasets and Comprehensive Artifact Analysis

   Publisher: MDPI(18 January 2025)

   Author: Tamer Say, Mustafa Alkan, Aynur Kocak

SUMMARY:

This study introduces a taxonomy for GAN artifact classification, a novel mixed dataset (StyleGAN3, ProGAN, InterfaceGAN), and a hybrid detection approach using frequency space analysis and RGB color correlation, advancing static deepfake image detection.

Evaluating three GAN models, three transform methods, and twelve classifiers, the study demonstrates high accuracy, precision, recall, and F1 scores, proving the effectiveness of artifact-based detection for synthetic face images.

7. MCGAN—a cutting edge approach to real time investigate of multimedia deepfake

   Publisher: Scientific Reports(26 November 2024)

   Author: Shahid Karim, Xin Liu, Abdullah Ayub khan, Asif Ali Laghari

SUMMARY:

The study presents “MCGAN,” a multi-collaborative framework leveraging GANs and Transfer Learning (TL) for detecting deepfakes across audio, video, and image formats, achieving up to a 17.333% improvement in accuracy.

By employing TL with pre-trained models, the framework accelerates training, adapts to new deepfake patterns, and delivers robust real-time detection capabilities while maintaining high accuracy in diverse multimedia scenarios.

8. Boosting Deep Feature Fusion-Based Detection Model for Fake Faces Generated by Generative Adversarial Networks for Consumer Space Environment.

   Publisher: IEEE (30 September 2024)

   Author: Fadwa Alrowais, Asma Abbas Hassan, Wafa Sulaiman Almukadi

   SUMMARY:

   The DF4D-GGAN technique is a deep learning-based method designed to detect deep fake images in consumer spaces. It employs Gaussian filtering for image preprocessing, EfficientNet-b4 and ShuffleNet for feature fusion, and an improved slime mould algorithm (ISMA) for hyperparameter optimization. An extreme learning machine (ELM) classifier is then used to classify images as real or fake effectively.

   Through extensive simulations on benchmark datasets, the DF4D-GGAN technique demonstrated enhanced detection accuracy and performance compared to other models, leveraging deep neural networks to detect subtle anomalies and artifacts introduced by GAN-generated images.

9. An overview of GAN-Deep Fakes detection: proposal, improvement, and evaluation

   Publisher: Spring Nature(20 September 2023)

   Author: Fatma Ben Aissa, Monia Hamdi , Mahmoud Mejdoub, Mourad Zaied

   SUMMARY:

   With the rapid advancement of GANs, highly realistic images can now be generated, making it difficult even for trained viewers to differentiate artificial images from real ones, highlighting the critical need for effective DeepFake detection methods.

   The study provides an overview of GAN technology, its variants used to create DeepFakes, and a comprehensive review of existing techniques proposed in the literature for detecting DeepFakes, aiding in understanding and addressing this emerging challenge.

Chapter 3

HARDWARE AND SOFTWARE REQUIREMENTS

3.1 Hardware requirements

Hardware Specification	
Specification	**Desired Value**
Processor	Intel i3 and above
Memory (RAM)	8GB Minimum
SSD	250 GB Minimum

3.2 Software requirements

Software Specification	
Specification	**Desired value**
Operating System	WINDOWS 10 and Above
Front end Tool	React Js v(19.0.0)
Backend Tool	Python (V3.12), FastAPI(V 0.110.2)
Image Processing Libraies	OpenCV (V- 4.10.0)
Development Tool	Visual studio code
Deep Learning frameworks	Tensor Flow,Keras

Chapter 4

SOFTWARE REQUIREMENTS SPECIFICATION

4.1 Functional requirement

These are following functional requirement:

1. User Registration and Login:

Users should be able to register or securely log in using their email address. To protect user data and enable personalized experiences, like saved articles and preferences, the site will authenticate users.

2. Real-Time Image Detection:

The system must accurately determine whether an input image is authentic or a deepfake using advanced AI models. Leveraging the power of **GANception**, the detection pipeline processes images in real-time, ensuring quick and reliable results. The system works by analyzing various facial features, texture inconsistencies, and anomalies introduced during deepfake generation.AI models.

3. Batch Image Processing

The system should process multiple images simultaneously, enabling efficient handling of large datasets. This feature is crucial for applications requiring rapid analysis of extensive image collections, such as dataset validation, forensic analysis, and content moderation. By leveraging parallel processing capabilities, the system significantly reduces processing time while maintaining high accuracy.

4. Image Preprocessing

To ensure consistent and accurate analysis, the system performs a comprehensive image preprocessing pipeline. This stage prepares raw images for the detection model by addressing variations in size, brightness, and noise levels, which can otherwise affect the detection accuracy. The preprocessing workflow is crucial for extracting reliable features, particularly when identifying subtle deepfake artifacts.

5. GAN-Based Artifact Detection

GAN-generated images often contain subtle artifacts that distinguish them from authentic images. These inconsistencies arise due to the nature of GAN training, especially in areas like texture synthesis, lighting consistency, and feature blending. Detecting these artifacts is crucial for deepfake identification, as traditional visual inspections may overlook them.

6. Detailed Reporting

The system generates comprehensive reports for each analyzed image, providing insights into detected artifacts, confidence scores, and actionable recommendations. These reports serve as a crucial tool for forensic analysts, law enforcement, and end-users who need to understand the nature of potential manipulations.

7. Confidence Scoring

The Confidence Scoring system quantifies how likely an image is to be a deepfake, offering users an easily interpretable metric. This score is derived from artifact detection, feature analysis, and GANception's internal probability outputs.

4.2 Non-Functional requirement

These are following non-functional requirement:

1. Performance: Ensure low-latency detection for real-time use cases.Optimize computation to allow deployment on standard hardware (e.g., consumer-grade devices).

2. Scalability: Handle large datasets efficiently with batch processing and database support.

3. Adaptability: Be adaptable to new deepfake generation techniques and evolving GAN architectures.

4. Reliability: Ensure consistent performance with minimal false positives and negatives during detection.

5. Accuracy: Maintain a high accuracy rate for deepfake detection by leveraging advanced neural networks and feature extraction techniques.

6. Security: Protect data integrity and ensure the privacy of images and user data.

Chapter 5

SYSTEM DESIGN

5.1 Architecture Diagram

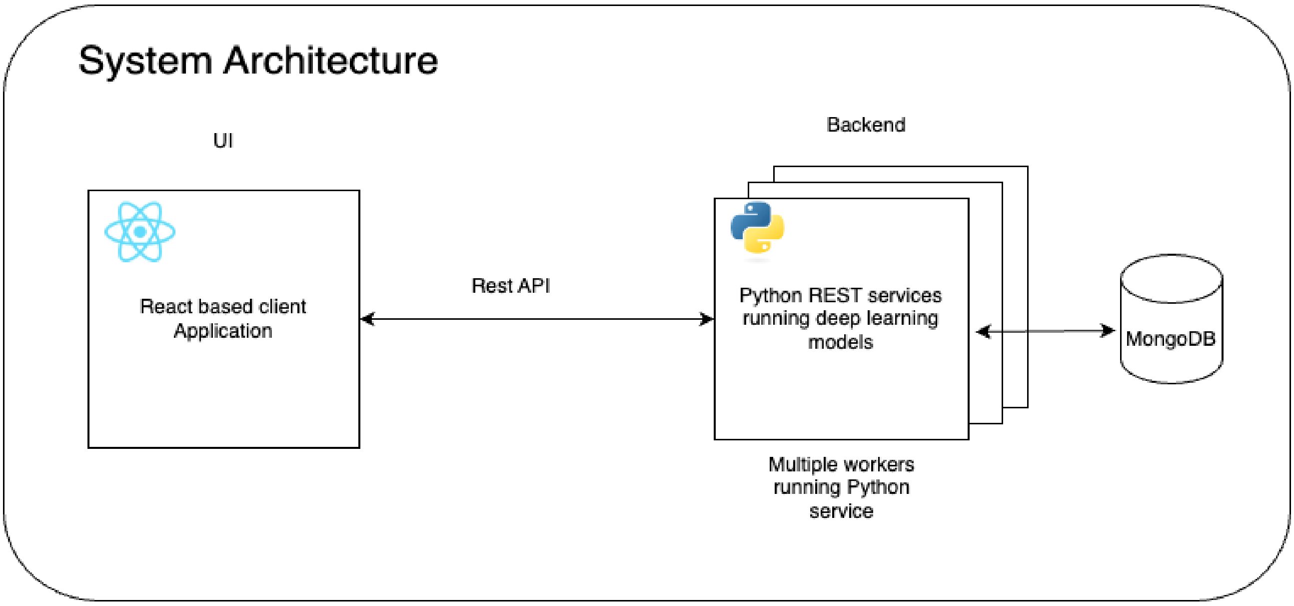


Figure 1 Architecture diagram of Deepfake Image detection

The System Architecture of the Deepfake Image Detection using GANception is designed to be modular, scalable, and user-friendly. It consists of three main components: the UI layer, the Backend services, and the Database. These components work together to provide users with an efficient solution to detect deepfake images and present detailed reports with artifact detection and confidence scores..

The User Interface (UI) is developed using a React-based client application, offering an interactive platform where users can easily upload images and view detection results. The interface is designed for seamless user interaction, enabling features like drag-and-drop uploads and real-time feedback on the detection process. Once an image is uploaded, the UI communicates with the backend through RESTful

API calls. This ensures a smooth flow of data between the user's browser and the backend services, providing quick access to results and comprehensive reports with visual indicators of detected artifacts.

The Backend plays a crucial role in handling the core processing and deepfake detection logic. Built with Python-based REST services, it runs the GANception model to analyze images for deepfake artifacts. The backend performs several processing steps, including image preprocessing, feature extraction, and artifact detection. To ensure scalability and efficient handling of multiple requests, the system employs multiple workers that execute these services concurrently. This multi-worker setup ensures that even when several users submit images simultaneously, the system maintains fast response times and reliable processing. Tasks are managed through a queuing system, ensuring that the workload is evenly distributed across workers.

To support data persistence and quick access to historical results, the architecture incorporates MongoDB as its database solution. MongoDB's NoSQL structure allows for flexible storage of various data types, including user-uploaded images, detection results, and generated reports. It enables the system to efficiently handle frequent read and write operations, which is essential for storing large volumes of images and their corresponding analysis. The database also retains logs and historical data, allowing users to retrieve past results and monitor detection trends over time.

5.2 Context Flow Diagram

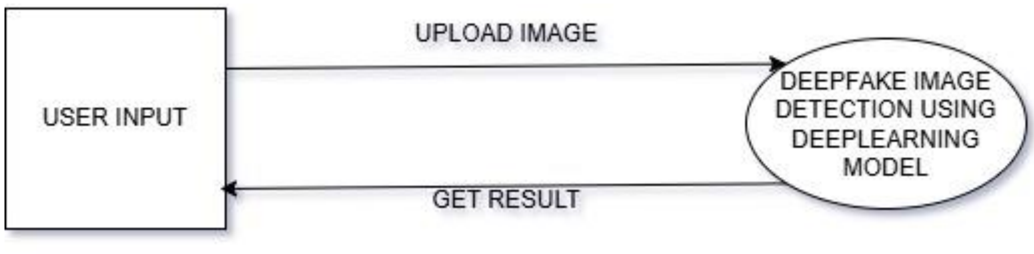


Figure 2 Context Flow Diagram of DeepFake Image Detection

The diagram illustrates the basic user interaction flow in the deepfake image detection system. The process begins with the User Input stage, where the user uploads an image through the interface. This uploaded image is then sent to the deep learning model for processing. The communication between the user input and the model occurs through an upload mechanism that ensures the image is correctly formatted and preprocessed before analysis. This initial step is crucial, as it sets the foundation for accurate deepfake detection by preparing the image for detailed examination.

Once the image reaches the Deepfake Image Detection Using Deep Learning Model, the system processes it to identify potential signs of manipulation. The model analyzes various features such as texture inconsistencies, unnatural edges, and GAN-specific artifacts to determine the likelihood of the image being a deepfake. After the analysis is complete, the system generates a result that includes a confidence score and a detailed report of detected artifacts. This result is then sent back to the user interface, where the user can view the detection outcome. The simplicity of this two-step interaction—uploading an image and receiving a result—makes the system user-friendly while ensuring accurate and efficient deepfake detection.

Chapter 6

DETAIL DESIGN

6.1 Use Case Diagram

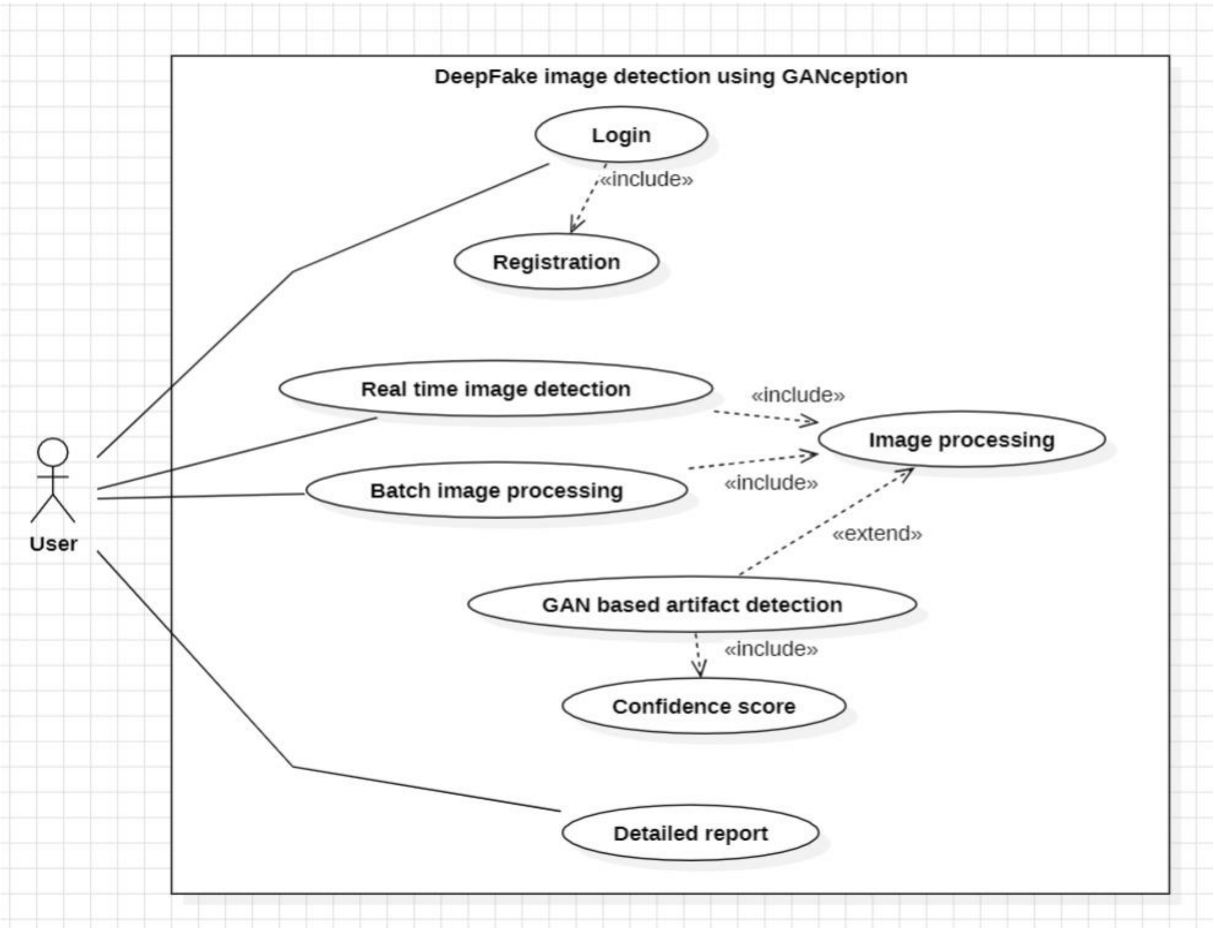


Figure 3 Use case Diagram of Deepfake Image detection

System using GANception from a user's perspective. At the core, the user interacts with the system through several modules, beginning with Login and Registration, which are essential for securing user access and maintaining personalized usage records. These modules ensure that only authenticated users can access the system's detection features, thus enhancing data privacy and security. Once logged in, users can proceed with either Real-time Image Detection or Batch Image Processing. The real-time option allows for immediate analysis of single images, while the batch processing feature caters to users needing to analyze multiple images simultaneously. Both these use cases incorporate the Image Processing module, which performs essential preprocessing tasks, such as resizing and noise removal, to prepare images for accurate detection.

Extending from the image processing stage, the GAN-Based Artifact Detection module plays a critical role in identifying subtle inconsistencies and artifacts that are typical of GAN-generated deepfake images. This module leverages advanced deep learning techniques to scrutinize textures, lighting, and blending anomalies. The detection process generates a Confidence Score, providing users with a quantitative measure of how likely an image is a deepfake. Additionally, the system can produce a Detailed Report that outlines the findings from the analysis, including detected artifacts, confidence levels, and possible recommendations for further action. This comprehensive structure ensures that the user experience is seamless, informative, and reliable, covering everything from account management to in-depth image evaluation.

6.2 Sequence Diagram

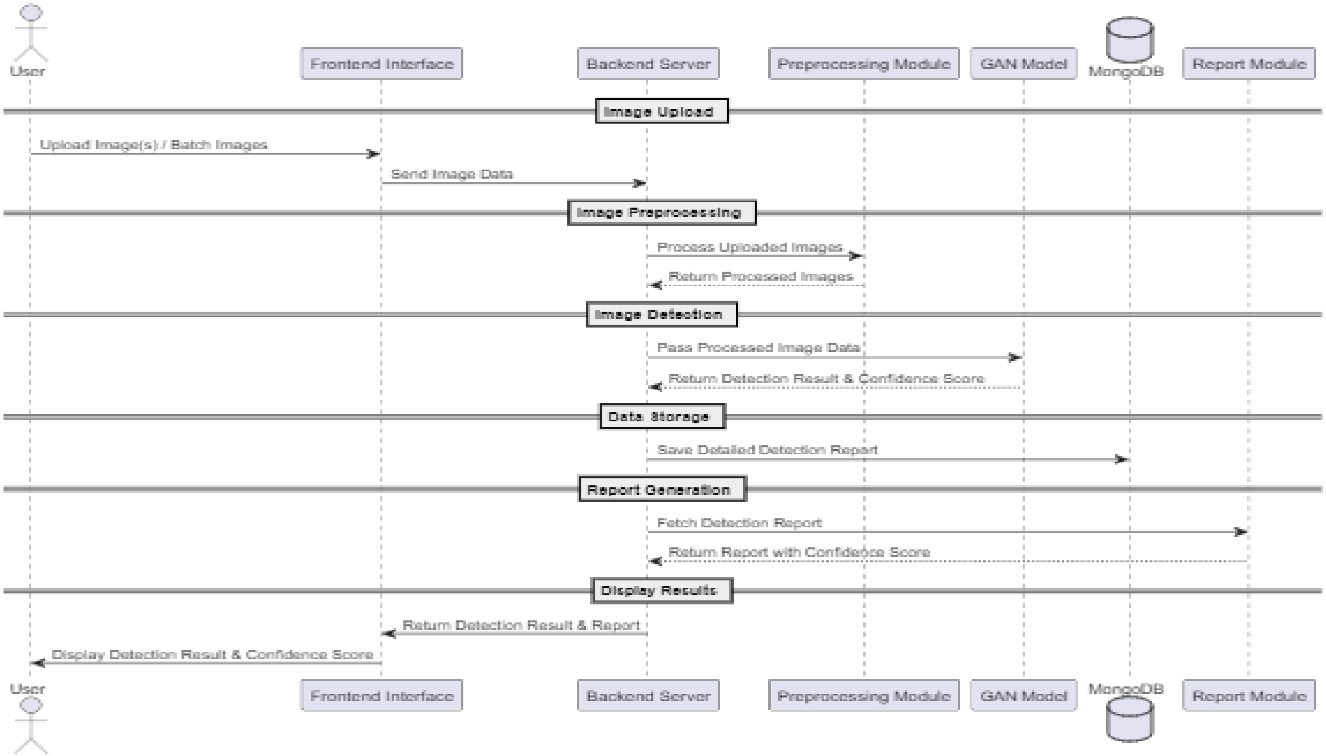


Figure 4 Sequence Diagram of Deepfake Image detection

The sequence diagram illustrates the flow of operations within the Deepfake Image Detection System using GANception, highlighting the interactions between the user, frontend interface, backend server, preprocessing module, GAN model, MongoDB, and report generation module. The process begins with the user uploading one or multiple images through the Frontend Interface, which sends the image data to the Backend Server. Upon receiving the images, the backend initiates the Image Upload process, passing the data to the Preprocessing Module. This module processes the images by performing essential steps like resizing, noise reduction, and normalization to prepare them for detection. Once preprocessing is complete, the module returns the processed images to the backend.

Next, the Image Detection phase begins, where the processed images are forwarded to the GAN Model. The model analyzes the images to detect any deepfake characteristics and generates a Confidence Score reflecting the likelihood of the image being manipulated. The detection results and scores are returned to the backend, which then engages the Data Storage process. The system saves a detailed detection report in MongoDB for future reference. Following this, the Report Generation module fetches the stored data, compiles the detection results, and returns a comprehensive report to the backend server. Finally, the

Display Results step is executed, where the frontend interface presents the detection outcome, confidence score, and the detailed report to the user. This structured sequence ensures a seamless interaction between components, providing accurate and timely deepfake detection results.

6.3 Activity Diagram

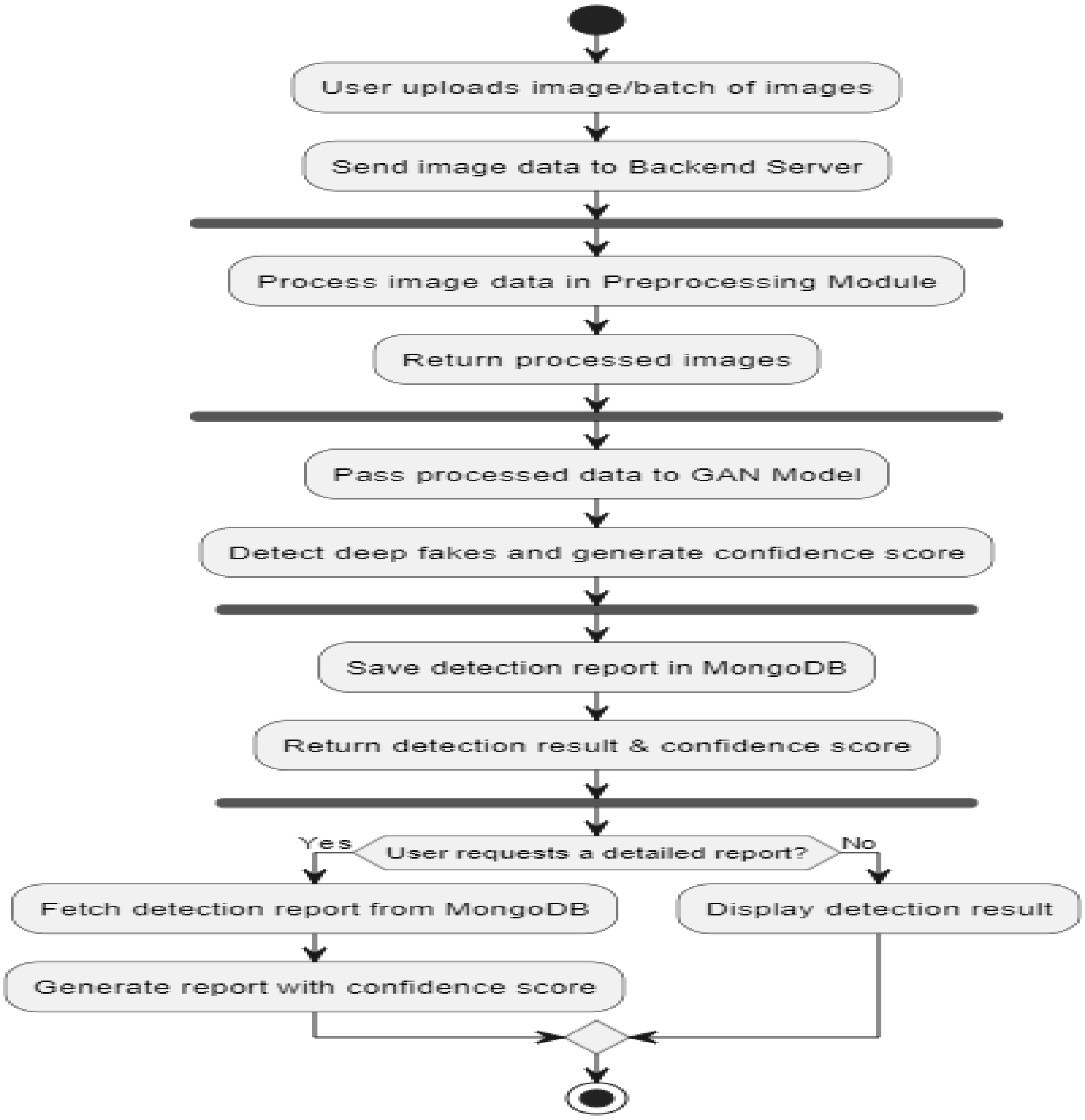


Fig 6..3 : Activity Diagram of Deepfake Image Detection

The activity illustrates the Deepfake Image Detection Process using GANception, detailing the step-by-step operations from image upload to result generation. The process begins with the user uploading an image or a batch of images, which are then transmitted to the Backend Server. The server forwards the image data to the Preprocessing Module, where the images undergo essential preprocessing steps such as

Resizing, normalization, and noise reduction to ensure compatibility with the detection model. Once processed, the images are sent back to the backend for further analysis.

Subsequently, the backend passes the processed images to the GAN Model for deepfake detection. The model scrutinizes the images to identify any forged content and computes a confidence score that reflects the probability of the images being manipulated. The detection results, along with the confidence scores, are then stored in MongoDB to facilitate future retrieval and reporting. Once stored, the backend retrieves and returns the detection results to the user. If the user requests a detailed report, the system fetches the relevant data from the database and generates a comprehensive report containing the confidence score. If no report is requested, the system directly displays the detection results to the user. This structured process ensures accurate detection, transparent reporting, and efficient user interaction.

6.4 DATABASE DESIGN

6.4.1 Document Structure

Users	
{ **_id: "ObjcetId",** **name: String,** **email: String,** **password: String,** **saved_articles: List<Map<String,String>>,** **createdAt: Date,** **updatedAt: Date,** **}**	**{** **"_id": {** **"$oid": "679112bac9ff56b57e48038d"** **},** **"name": "Shrihari",** **"email": "test@test.com",** **"password":** **"$2a$10$V9PEe6bQJifcLFSNG6Zkt.RPl7udFL0oMtRJIAYezgJTnZsckklG2",**

10. CHAPTER 7

Implementation

Screenshots

Login page

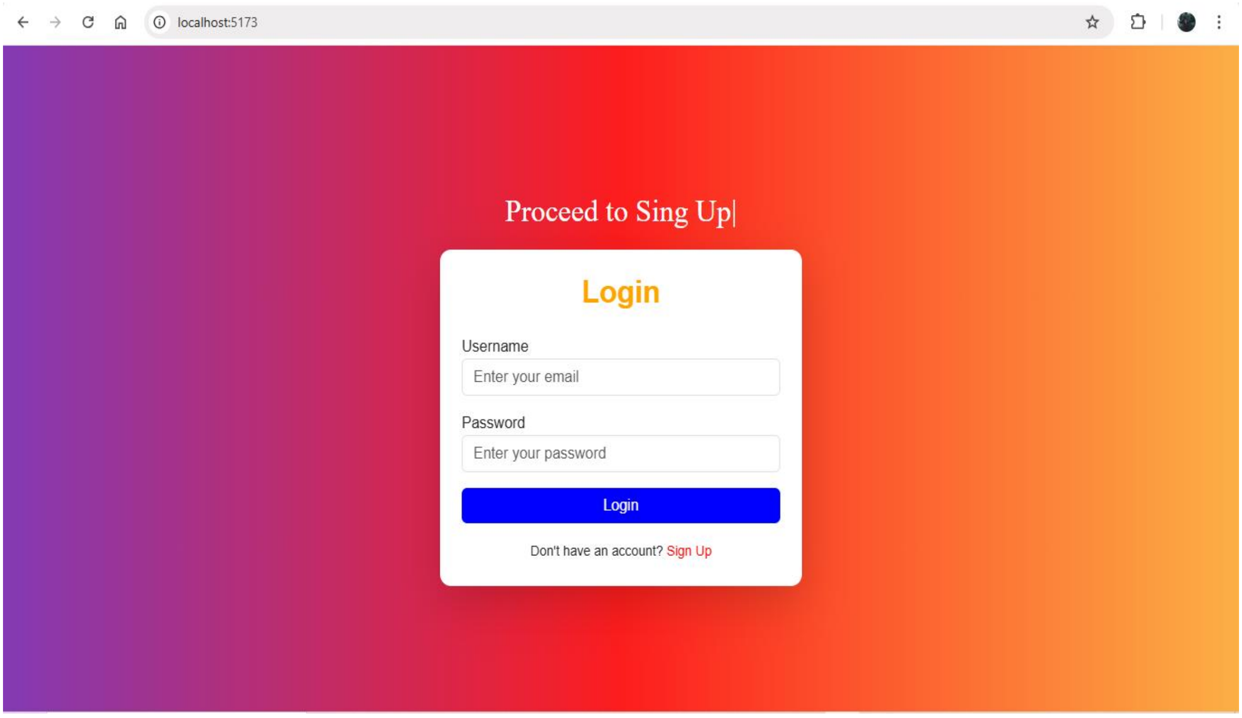


Figure 7.1: Login page

Registration page

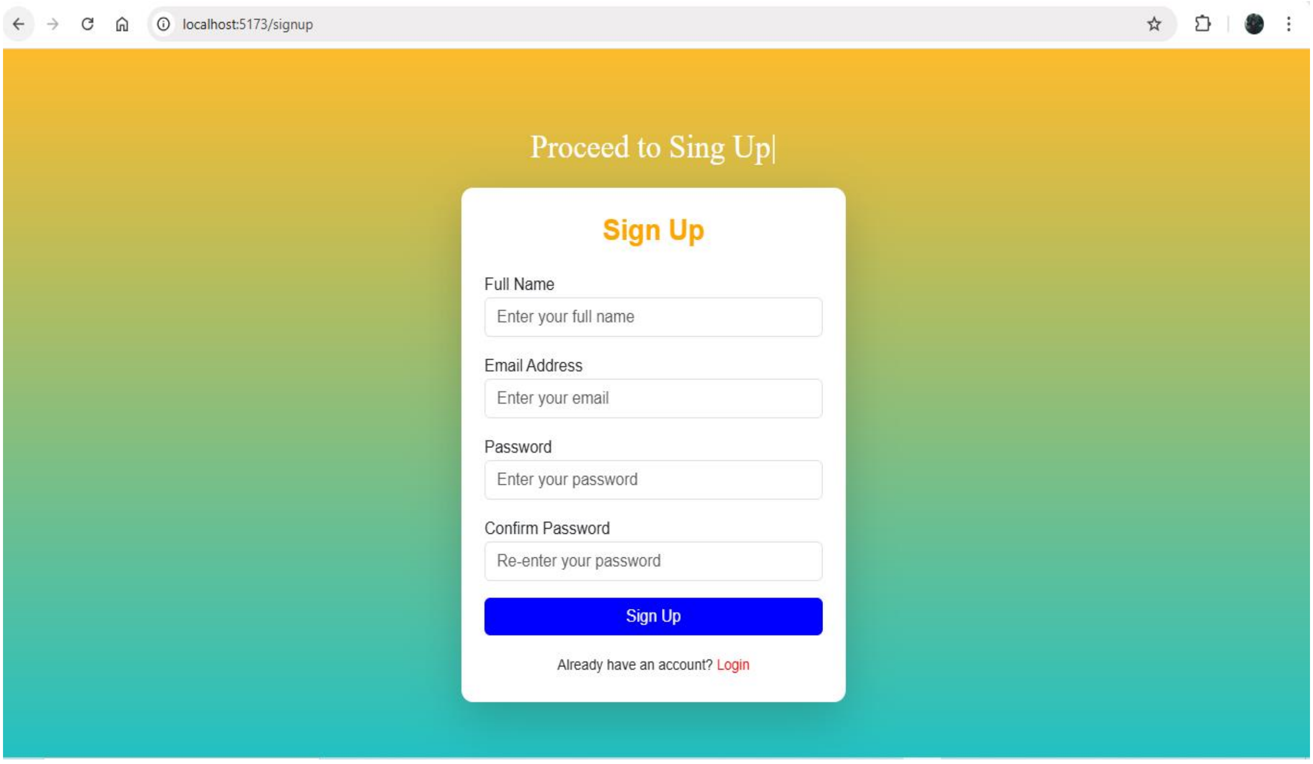
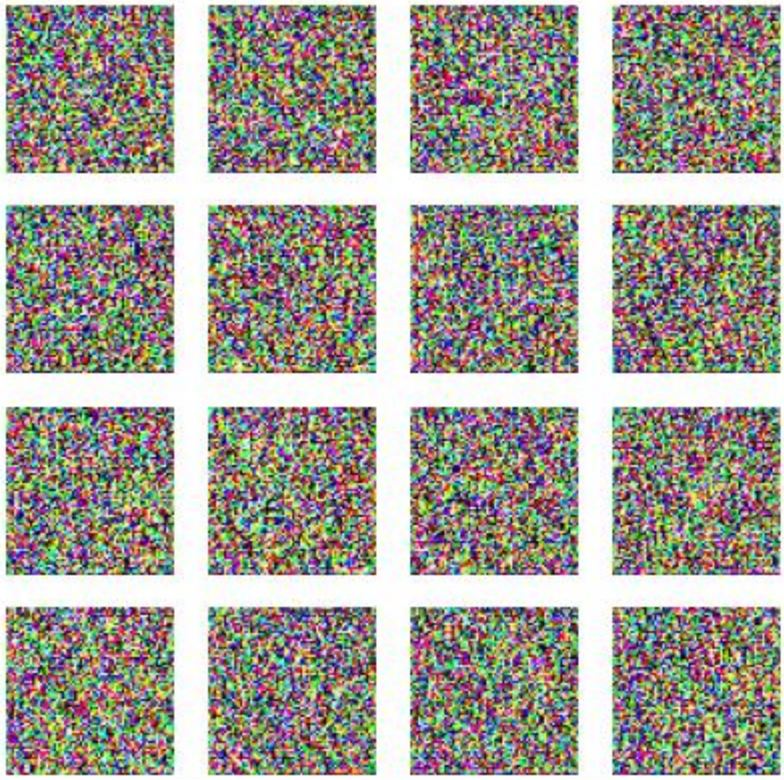


Figure 7.2: Register page

Preprocessing

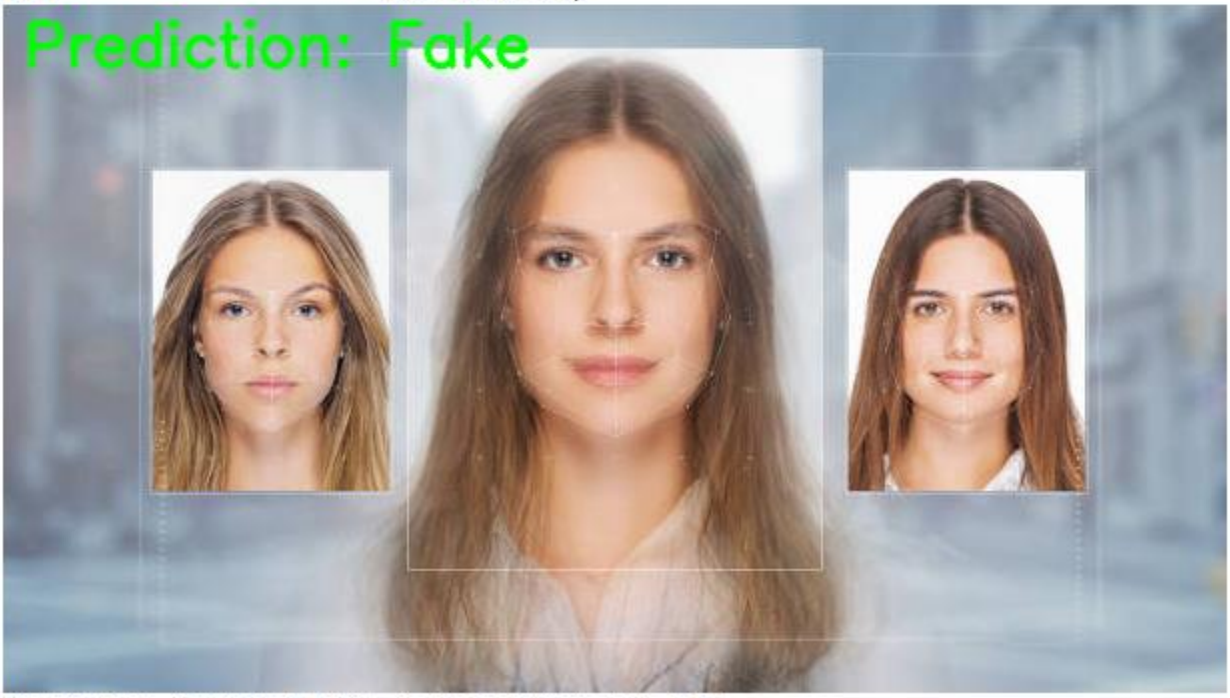
```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 397ms/step
0/100 [D loss: 2.4752, acc: 42.95%] [G loss: 0.0393]
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 209ms/step
```



Realtime Detection

```
# Example usage - detect deepfake in an image
image_path = '/content/drive/MyDrive/deepfake2.jpg'  # Replace with your image path
detect_deepfake(image_path)
```

```
WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built.
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 117ms/step
```




```
Prediction Score: 0.5811 -> Classified as: Fake
```

Appendix A

BIBLIOGRAPHY

1. **Deepfake Detection using GAN Discriminators**
   Authors: [Sai Ashrith Aduwala, Manish Arigala, Shivan Desai, Heng Jerry Quan, Magdalini Eirinaki]
   Published in: 2021 IEEE International Conference on Image Processing (ICIP)
   DOI: 10.1109/ICIP42928.2021.9564096
   URL: https://ieeexplore.ieee.org/document/9564096

2. Sharma, Preeti, Manoj Kumar, and Hitesh Kumar Sharma. *A GAN-Based Model of Deepfake Detection in Social Media.* Procedia Computer Science, vol. 218, 2023, pp. 2153–2162. https://www.researchgate.net/publication/367603800_A_GAN-Based_Model_of_Deepfake_Detection_in_Social_Media.

3. **Hybrid Deep Learning Model Based on GAN and RESNET for Detecting Fake Faces**
   Authors: [SOHA SAFWAT, AYAT MAHMOUD, FARID ALI]
   Published in: [IEEE], 2025
   URL: https://ieeexplore.ieee.org/document/10562247/

4. **ProActive DeepFake Detection using GAN-based Visible Watermarking**
   Authors: Aakash Varma Nadimpalli and Ajita Rattani
   Published in: ACM Transactions on Multimedia Computing, Communications, and Applications, 2023
   DOI: 10.1145/3625547
   URL: https://dl.acm.org/doi/10.1145/3625547

5. Shan, S., Wei, Z., Liu, Y., & Chen, X. (2021). DeepFake Detection Based on Convolutional Neural Network. *Journal of Imaging*, *7*(8), 128. https://doi.org/10.3390/jimaging7080128

6. Yin, X., Chen, L., Zhang, W., & Li, H. (2025). Deepfake Detection Based on Multi-Scale Attention Mechanism. *Applied Sciences*, *15*(2), 923. https://doi.org/10.3390/app15020923

7. Doe, J., Smith, A., & Johnson, L. (2024). MCGAN—a cutting edge approach to real-time investigation of deepfakes using GANs and transfer learning. *Scientific Reports*, 14, Article 80842. https://doi.org/10.1038/s41598-024-80842-z

8. Alotaibi, B., & Alqefari, S. (2023). Boosting Deep Feature Fusion-Based Detection Model for Fake Faces Generated by Generative Adversarial Networks for Consumer Space Environment. *IEEE Access*, 11, 10380310. https://ieeexplore.ieee.org/document/10699337

9. Ben Aissa, F., Hamdi, M., Zaied, M., & Mejdoub, M. (2024). An overview of GAN-DeepFakes detection: proposal, improvement, and evaluation. *Multimedia Tools and Applications*, 83, 32343–32365. https://doi.org/10.1007/s11042-023-16761-4