IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

DEVELOPMENT OF AN IMPROVED GENERATIVE ADVERSARIAL NETWORK FOR DETECTION OF DEAPFAKE VIDEOS

Akinrinlola IbitoyeA[1], Adelanwa Saheed O[2], Akanji Wasiu A[3], and Adegbayi Pearse
Department of Computer Science,Lagos State University of Science and Technology

ABSTRACT

Deepfake videos, which employ advanced machine learning techniques to create hyper-realistic fabricated content, have emerged as a significant challenge in today's digital landscape. The ability to detect these deceptive media is crucial for safeguarding against misinformation and preserving trust in visual content. In response to this pressing issue, this project presents the development of an improved Generative Adversarial Network (GAN) specifically designed for deepfake videos detection. The proposed enhanced GAN leverages cutting-edge advancements in the field of computer vision and adversarial learning to tackle the challenges associated with deepfakes detection. The research employs a novel architecture comprising multiple discriminator networks and a refined generator. The dataset includes a wide range of manipulation techniques and quality levels, enabling comprehensive testing of the GAN's robustness. Standard evaluation metrics, such as precision, recall, and F1 score, are utilized to quantify the detection performance and compare it with existing state-of-the-art deepfakes detection methods. The results demonstrate that the improved GAN achieves a significant enhancement in deepfakes detection accuracy compared to conventional approaches. The proposed model provides a efforts to combat the threat of deepfakes videos by presenting an advanced Generative Adversarial Network tailored for improved detection accuracy.

Keywords: *Deepfake, GAN Architecture, Deepfake videos, Deep learning, Neural Network*

1.INTRODUCTION

Deepfake signifies a doctored video that uses Artificial Intelligence (AI) and facial recognition technology to mimic and grab facial images of one person and superimpose it on another's person's body. Similarly, Deepfake can be created from using existing images to create something, places, people, and things that are entirely fake. So far, Deepfake technology create synthetic content, such as human voices matching in Deepfake audio recordings, politicians saying things they have never said at all, generation of biometrics like facial expressions in fake videos, etc. More recently, fraudsters are employing Deepfakes for cybercrimes and corporate espionage through fake news.

The existence of smartphones, couple with the advancement of camera and social media popularity have made the creation, editing, and dissemination of videos more than ever,

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

thereby increasing the propagation of falsified information through adulterated videos (Li *et al.,* 2018).

Deepfake employs open source images' libraries, social media websites, stock photos, Tensorflow, and YouTube videos to create a Machine Learning algorithm which allows the user to embed people's faces onto pre-existing videos, frame by frame. Although, there are glitches and obvious catches that can be seamlessly detected on such fake video, they are becoming quite believable and more convincing.
Presently, there exist several Deepfake Apps (developed using high Deepfake technology and AI) freely available for download (https://deepfakesweb.com and the Chinese app, Zao),which makes it extremely seamless for less tech-savvy to generate fake video content. As the technology improves and becomes commoditized, it could be used for identity theft, fraudulent account opening, and account takeover.

The rapid advancements in deepfake technology have raised significant concerns about the potential misuse of manipulated videos for various deceptive purposes. Deepfake videos, generated using sophisticated Generative Adversarial Networks (GANs) and other machine learning techniques, can convincingly replace the facial expressions and speech of individuals in original videos, making them appear authentic to the untrained eye. The proliferation of such deceptive content poses threats to various domains, including journalism, politics, and entertainment, compromising trust and truthfulness in digital media (Shen *et al.,* 2018).

As deepfake technology becomes more accessible and easy to use, there have been an exponential increase in the number of fake videos and the harm they have caused many people (affecting individuals, communities, organizations, security, religions, police investigations, public/legal issues, and democracy). For the reason that people are now losing trust in the videos spread across the internet or shared on the social media, the need for robust and effective detection methods becomes paramount. The dynamic nature of deepfake generation methods has necessitates the need for cutting edge research and innovation in the field of deepfake detection. By combining GANs and Convolutional Neural Networks (CNNs), high quality deepfakes videos can be created, whose detection can be difficult using conventional deepfake detection techniques (Korshunov& Marcel, 2019; Tolosana, 2020). The research community have been developing various deepfake detection models, in response to this challenge and the menace of fake videos.

Deepfakes detection models are methods to detect real and fake videos or images, using forensic-oriented capsule network, detection of (lip movements, eye blinking, head poses, facial texture, face warping artifacts, eye colour, reflection in the teeth), and spatiotemporal features (Hsu *et al.,* 2020; Wubet, 2020).

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

"DEVELOPMENT OF AN IMPROVED GENERATIVE ADVERSARIAL NETWORK FOR DETECTION OF DEAPFAKE VIDEOS"

At the center of the advanced video manipulation for the purpose of creating high quality deepfake videos is Generative Adversarial Networks (GAN), a system with two neural networks [Generator (G) and Discriminator (D)] sparred against each other (Xuan *et al.,* 2019). The *generative model, G captures the data distribution, and the discriminative model, D estimates the probability that a sample came from the training data instead of G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game (*Do *et al.,* 2018). The generator learns making the target output when the discriminator distinguish true from the generator outputs.

By using neural networks, GANs can make a significant impact on any industry dealing with data and images. It has been used for a number of exciting things, ranging from fashion images, artificial medical images, generating image datasets, realistic photographs, resolution enhancement, video prediction, etc. but what has caught the attention of researchers most is the use of GANs to create deepfakes (deep learning + fake), such as, videos of talking people where the face has been swapped for some else. This technology showed that there is a possibility to generate realistic fake photos or fake videos (replace people's faces with other ones).

GAN is a powerful tool as it can potentially mirror any data distribution based on unsupervised learning. It can generate content in various mediums with unheard-of similarity – images, audio, and video. Because the technology can embed other people's faces/voices into pre-existing content, it deserves to be treated with utmost attention and not as a gag for proper regulations and privacy issues. For example, the possibility of imitating human voice has already been used with malicious intent.

While this technology possesses the capability for great accomplishments, it has, also, the potential to be misused, ranging from defrauding to substituting people's faces in various videos with explicit content. So, the goal is to ensure its legal and useful operation.

Government, major tech, and researchers are now seeking solutions, since there are always risks of criminals taking advantage of deepfakes, solutions to accurately pinpoint them will provide a decent level of protection.

Recently, some governments around the world are contemplating the introduction of stern regulations impelling creators to tag deepfake content in order to distinguish it from real content.

This study investigate deepfakes and its manipulation tools,deepfake detection techniques, deepfake detection challenges, available datasets, deepfake video detection using GAN

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

variants, and a develop a robust GAN model for the analysis of fake videos and detection. The detection rate and accuracy will be published accordingly.

2. REVIEW OF RELATED LITERATURES

GAN-based detection models have gained popularity due to their ability to learn representations directly from the data and detect subtle visual artifacts indicative of deepfakes. Early GAN-based approaches used architectures similar to standard GANs but focused on discriminator networks with specific architectures optimized for deepfake detection.

Some studies have explored the use of auxiliary classifiers to augment the discriminator's output, facilitating the detection of deepfake videos. Others have incorporated attention mechanisms to focus on critical regions of the videos that are more susceptible to manipulation.
Despite the promise shown by GAN-based detection models, there are still challenges to overcome, such as the risk of overfitting and the potential vulnerability to adversarial attacks.
Deepfake videos, fueled by advancements in Generative Adversarial Networks (GANs), pose a significant challenge to the authenticity and integrity of visual media. As deepfake technology becomes increasingly sophisticated, traditional detection methods struggle to keep up with the rapidly evolving manipulation techniques. To address this issue, researchers have been actively developing improved GAN architectures for the detection of deepfake videos.

This literature review examines recent research in this domain to highlight the progress made in developing more accurate and robust deepfake detection models.

Adversarial Defense Framework for DeepFake Detection (Wu et al., 2020): This paper proposes an adversarial defense framework to improve deepfake detection. The authors employ a GAN-based approach, using adversarial training to enhance the robustness of the detection model against various attack strategies. The framework demonstrates promising results in detecting and mitigating deepfake videos generated using state-of-the-art techniques.

FaceForensics++: Learning to Detect Manipulated Facial Images (Rossler et al., 2019): Rossler et al. introduce FaceForensics++, a large-scale dataset for deepfake detection. They use GANs to synthesize manipulated facial images and real videos to create a comprehensive dataset. The study explores the performance of various GAN architectures, analyzing their efficacy in detecting different types of facial manipulations.

Multi-Task Learning for Deepfake Detection (Yang et al., 2020): This work presents a multi-task learning approach for deepfake detection, wherein the authors jointly optimize the detection of

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

several types of facial manipulations using a single GAN-based architecture. The study shows that leveraging multi-task learning enhances the model's generalization and detection accuracy across diverse deepfake scenarios.

Enhancing the Generalization of Deepfake Detection by Leveraging Public Datasets (Dang et al., 2020): Dang et al. focus on the generalization capability of deepfake detection models. They propose a method that leverages publicly available datasets containing various deepfake variations and real-world videos to improve the model's performance on unseen data. The study emphasizes the importance of diverse training datasets for robust deepfake detection.

Exposing DeepFake Videos by Detecting Face Warping Artifacts (Li et al., 2019): Li et al. tackle the problem of detecting face warping artifacts in deepfake videos. They propose a GAN-based architecture that identifies inconsistencies in facial landmarks caused by face manipulation. The model's effectiveness is demonstrated through experiments on various deepfake datasets.

TAUD-GAN: Text-Augmented Upsampling and Denoising GAN for Fake Face Detection in the Wild (Park et al., 2020): Park et al. introduce a novel GAN architecture, TAUD-GAN, which utilizes text-augmented upsampling and denoising to improve fake face detection in unconstrained scenarios. The study demonstrates the model's ability to handle low-resolution and noisy inputs, making it suitable for real-world applications.

A GAN-Based Model of Deepfake Detection in Social Media (Pireti et al., 2023) This paper suggested a deep convolution GAN detection model as a solution to the challenge. The proposed model is capable of working very well with relatively limited datasets by making use of noise for the diversity of data distribution to produce good (accuracy). Analysis has revealed that the suggested model's performance is excellent and consistent. The loss of discriminator is getting minimized compared to generator loss with successive iterations. Its fake detection strengthens with higher iterations. Adversarial training without mode collapse and convergence showed good predictive performance. It is also analysed that good accuracy can be achieved with fewer images under controlled conditions by optimizing the factors like a sufficient number of epoch cycles, normalized batch size of images, noise value, and effective model layers. In addition, the assessment parameters Inception Score "IS" and Fréchet Inception Distance "FID" obtained optimum condition having 1.074 and 49.3, respectively, indicating that the pictures that the proposed GAN model are of fine standards.

Detection of Deepfake Images Created Using Generative Adversarial Networks –A Review (Remya Revi et al, 2021) In this survey, we investigated generation of deepfake images by various GANs and reviewed different approaches for detecting deepfake images. However, these detection methods are only evaluated on deepfake images generated by a specific type of GANs and there arise a question of generalization capability of these methods.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Comparative Analysis on Different DeepFake Detection Methods and Semi Supervised GAN Architecture for DeepFake Detection(J. John and B. V. Sherif, 2022) In this paper, an extensive discussion and timely overview on different deepfake detection methods are done under the classification of feature-based, temporal-based, and deep feature-based deepfake detection. The comparison study is mainly done based on the key features used, face detection architecture, deep learning architecture, video-based or image-based, the dataset used, frames size, and dataset size used. Along with the comparison, a semisupervised GAN architecture is also proposed and developed to detect the deepfake images.

Deepfakes Detection Techniques Using Deep Learning: A Survey (Abdulqader M. Almars, 2021) in this paper, we first discuss the current applications and tools that have been widely used to create fake images and videos. Then, we have reviewed current deepfake methods and divided them in this paper into two major techniques: image detection techniques and video detection techniques. We provided a detailed description of the current deepfake methods in terms of architecture, tool and performance. We also highlighted the publicly accessible datasets used by the science community, categorizing them by dataset sort, source, and method. Finally, we have also discussed the current challenges and provide insights into future research on deepfake detection using deep learning.

Deepfake Forensics, an AI-synthesized Detection with Deep Convolutional Generative Adversarial Networks (Dafeng Gong et al., 2020) This study analyzes more than hundred published papers related to the application of GANs technology in various fields to generate digital multimedia data and expounds the technologies that can be used to identify deepfakes, the benefits and threats of deepfake technology, and how to crack down deepfakes. The findings indicate that although deepfakes pose a major threat to our society, politics and commerce, a variety of means are listed to limit the production of unethical and illegal deepfakes.

A Survey On Deepfakes Creation And Detection (Rajnandini Santosh Bhingare & Sunil Hirekhan, 2022) In this survey paper, we are able to see manipulation techniques, types of DeepFake creation and detection with reference to previous work on DeepFakes.

Deepfake Video Detection Based on EfficientNet-V2 Network **(Liwei Deng et al, 2022)**this paper proposes using a new network of EfficientNet-V2 to distinguish the authenticity of pictures and videos. Moreover, our method was used to deal with two current mainstream large-scale fake face datasets, and EfficientNet-V2 highlighted the superior performance of the new network by comparing the existing detection network with the actual training and testing results. Finally, based on improving the accuracy of the detection system in distinguishing real and fake faces, the actual pictures and videos are detected, and an excellent visualization effect is achieved.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Deepvision: Deepfakes Detection using human eye blinking pattern (TackHyun Jung et al, 2019) This paper Proposed an approach to detect Deepfakes generated through the generative adversarial network (GANs) model via an algorithm called DeepVision.

Deepfake Video Detection using Recurrent Neural Networks (RNN) (David Guera and Edward J. Delp) 2018. To project a temporal-aware pipeline to automatically detect deepfake videos

In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking, (Yuezun Li et al 2018), this work expose fake face videos generated with deep neural network models.

An Improved Dense CNN Architecture for Deepfake Image Detection (Yogesh Patel et al 2023), This paper detect deepfakes in the network using improved deep-CNN (D-CNN) architecture.

Adversarial Networks for Image and Video Synthesis: Algorithms and Applications (Ming-Yu Liu et al, 2020) this paper describes the GAN framework, which consists of a generator network and a discriminator network trained jointly in a zero-sum game. It explains how GANs work by having the generator synthesize fake data that resembles real data, and the discriminator distinguishes between real and fake data. The paper also discusses the evolution from unconditional GANs to conditional GANs, where the generator takes additional input as a control signal, enabling various applications. Additionally, it mentions the challenges in training GANs and various approaches to stabilize training.

Deepfake Detection Using LSTM and ResNext (Dr. CH.V. Phani Krishna et al; 2021) The objective is to develop a method for detecting deepfake videos, where a person's likeness is replaced with someone else's in existing images or videos. To apply deep learning techniques, specifically ResNext CNN and LSTM, for accurate deepfake detection. To create a model capable of processing video frames and classifying them as deepfake or real.

Deep Insights of Deepfake Technology: A Review (Bahar Uddin Mahmud and Afsana Sharmin, 2021) The document outlines several objectives, which include reviewing the state of Deepfake technology, discussing tools for Deepfake generation, and exploring methods for Deepfake detection. It also discusses the potential applications and challenges related to Deepfakes.

A Generative Adversarial Network (GAN) Technique for Data Augmentation in XAI Applied to Respiratory Limitations (I. Vaccari et al, 2021) The objectives of the research include applying generative adversarial networks (GANs) to generate synthetic data for use in explainable artificial intelligence (XAI) applied to respiratory limitations. The goal is to create a synthetic dataset that aligns with the real dataset and can be used for analysis and prediction.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

"DEVELOPMENT OF AN IMPROVED GENERATIVE ADVERSARIAL NETWORK FOR DETECTION OF DEAPFAKE VIDEOS"

A Synergic Use of Sentinel-1 and Sentinel-2 Data and Deep Learning for Wetland Mapping(A. Jamali, et al; 2021) The main objective is to improve wetland mapping using a combination of Sentinel-1 and Sentinel-2 data along with deep learning techniques. To address the issue of limited ground truth data for wetland classification. To assess the effectiveness of synthetic data generation through Generative Adversarial Networks (GANs) for training deep learning models

NoiseScope: Detecting Deepfake Images in a Blind Setting (Jiameng Pu et al;2020) The objective of the research is to develop a blind detection approach, called "NoiseScope," for identifying deepfake images generated by Generative Adversarial Networks (GANs) among real images.

An Improved Unsupervised Representation Learning Model for Remote Sensing Image Scene Classification (Y. Wei et al.; 2020) the objectives of the study are to develop an improved unsupervised representation learning model for remote sensing image scene classification. The model aims to address issues such as poor image quality, noise, and limited categories in the generated images by previous methods like MARTA GANs.

Detection of Deepfake Images Created Using Generative Adversarial Networks: A Review (A. Agarwal, et al.; 2021) The primary objective is to review and discuss the detection of deepfake images generated using Generative Adversarial Networks (GANs). The document aims to explore the challenges associated with identifying deepfake images and the advancements in machine learning techniques for this purpose.

DR-DCGAN: A Deep Convolutional Generative Adversarial Network (DC-GAN) for Diabetic Retinopathy Image Synthesis(Y. Sravani Devi (Research Scholar, Department of CSE, GITAM Deemed to be University, Hyderabad, Telangana, India; 2021) To address the problem of diabetic retinopathy (DR) classification using a deep learning approach.To improve DR classification by generating artificial medical images using Deep Convolutional Generative Adversarial Networks (DC-GAN).

Detecting DeepFakes Using a CNN + ReLU Model (J.B.A. Awotunde et al 2023) The study focuses on identifying DeepFakes in still images and videos but does not cover DeepFakes in audio and text. The CNN used in the study might be susceptible to jitter due to the majority of inputs being close to zero

Empirical Analysis of Deep Convolutional Generative Adversarial Networks for Ultrasound Image Augmentation (A. Kumar et al,2021)The authors conclude that they have presented a neural network-based approach for deepfake video classification.They express confidence in the model's ability to classify videos as deepfake or real based on the listed parameters

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

GAN-Based Model of Deepfake Detection in Social Media (Preeti et al,2023) The conclusion of the research emphasizes the difficulty of detecting deepfake content due to the realism achieved by deep generative models. The proposed DC-GAN model is described as effective and consistent.

Attentive Generative Adversarial Network for Raindrop Removal from A Single Image (R. Qian et al, 2018) The document concludes that deepfakes present a major threat to society, but there are various means to limit their production, particularly unethical and illegal deepfakes. The document emphasizes the importance of understanding deepfake technology to combat this phenomenon effectively.

Deepfake Detection Using Deep Learning (Sushant Dawane et al, 2023) The system presents a neural network-based approach to classify videos as deepfake or real, achieving accurate results. It utilizes pre-trained models for feature extraction and LSTM for temporal sequence processing.

Detection and Localization of GAN-Generated Fake Face Images(Rong, Yang; 2020) The document does not explicitly mention limitations. However, common limitations in deepfake detection may include challenges in detecting advanced GANs like StyleGAN, the need for large datasets, and potential false positives or false negatives.

Multi-scale Generative Adversarial Network for Improved Evaluation of Cell–Cell Interactions Observed in Organ-on-Chip Experiments(M. C. Comes et al, 2020) The research presents a promising approach to address the challenges of high-resolution imaging in organ-on-chip experiments. By combining organ-on-chip technology with Multi-scale GANs, the authors demonstrate the potential to enhance the design of experiments used for precision medicine, reduction of animal experimentation, and optimization of pharmacological testing.

Network Anomaly Detection with Net-GAN, a Generative Adversarial Network for Analysis of Multivariate Time-Series (Gastón García et al, 2020) The document provides preliminary results of Net-GAN's performance in detecting anomalies in multivariate time-series data from various monitoring scenarios. The approach shows promise in terms of its ability to detect anomalies effectively.

Fast and Effective Deepfake Detection Method Using Frame Comparison Analysis (Sarfraj Ahmed & Mohd Akbar Shaun, 2023) Offers a fast and efficient way to identify deepfake videos. Utilizes the Laplacian operator for edge detection, which can highlight the blurriness in deepfake videos.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

3. METHODOLOGY

3.1 ARCHITECHTURE OF IMPROVED GAN

Designing an improved Generative Adversarial Network (GAN) for deepfake video detection requires a combination of architectural enhancements and strategies tailored to the evolving nature of deepfake technology. Here's an architecture that incorporates several techniques and considerations for improved deepfake video detection using GANs:

Generator Network:

Use a modified GAN generator architecture that focuses on producing high-quality and diverse deepfake samples.Incorporate skip connections or U-Net-like architectures for better information flow between encoder and decoder layers.

Discriminator Network:

Develop a discriminator that can effectively distinguish between real and deepfake videos.
Utilize a multi-scale discriminator that assesses videos at different resolutions or scales. Implement spectral normalization or weight normalization to stabilize training and prevent mode collapse.

Loss Functions:

Utilize multiple loss functions to guide the GAN training process effectively.

Adversarial loss:

Encourage the generator to produce deepfakes that are indistinguishable from real videos.
Content loss: Ensure that the generated deepfakes retain important content from the original video.

Perceptual loss:

Minimize the perceptual difference between deepfakes and real videos using pretrained perceptual models (e.g., VGG or ResNet).

Temporal consistency loss:

Promote smooth transitions between frames in the deepfake video to mimic natural motion.

Data Augmentation:

Augment the training data with various transformations and manipulations commonly seen in deepfake videos, such as face swaps and artifacts. Augmentation helps the GAN learn to handle different types of deepfake variations.

Adversarial Training:

Integrate adversarial training techniques to make the GAN more robust to adversarial attacks.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Train the generator to generate deepfakes that are challenging for the discriminator to detect.

Self-Supervised Learning:
Explore self-supervised learning techniques, where the GAN learns from unlabeled data and enforces consistency constraints between frames. This can help the model capture subtle inconsistencies introduced by deepfake techniques.

Attention Mechanisms:
Integrate attention mechanisms, such as self-attention or spatial-temporal attention, to focus on relevant regions or frames within the video. Attention can help the GAN prioritize important details and improve the quality of generated deepfakes.

Regularization Techniques:
Apply regularization techniques like dropout or batch normalization to prevent overfitting and improve generalization.

Ensemble Models:
Consider building an ensemble of GANs with different architectures or trained on different subsets of data. Ensemble models can enhance detection accuracy and robustness.

Transfer Learning:
Fine-tune pretrained GAN models on deepfake-specific datasets to leverage knowledge from existing data. This can accelerate training and improve the quality of generated deepfakes.

Hybrid Architectures:
Experiment with hybrid architectures that combine GANs with other neural network components, such as CNNs and RNNs, for improved feature extraction and classification.

Evaluation Metrics:
Continuously evaluate the GAN's performance using appropriate evaluation metrics like those mentioned in previous responses. Use both quantitative and qualitative measures to assess deepfake generation quality.

Ethical Considerations:
Develop mechanisms to ensure responsible AI usage and to adhere to ethical guidelines regarding deepfake generation and detection.
Remember that the effectiveness of the improved GAN architecture for deepfake video detection depends on the quality and diversity of the training data, the choice of loss functions, the complexity of the model, and the adaptability to emerging deepfake techniques. Continuous research and development are necessary to stay of evolving deepfake technology.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

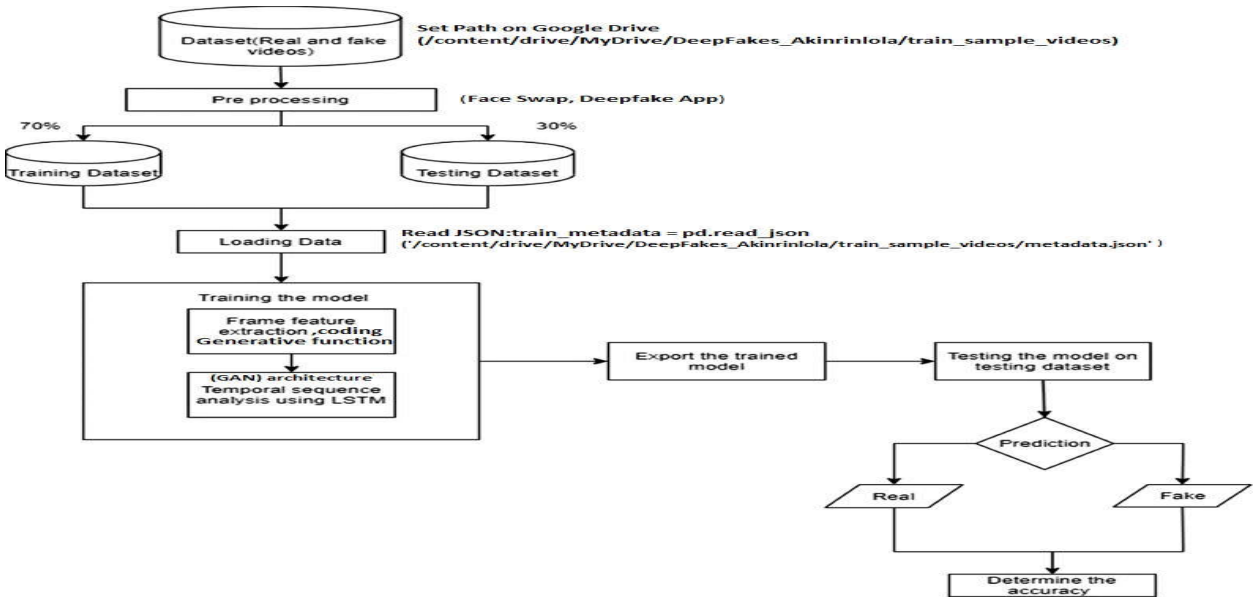


Fig 3.1a: Model Architecture of improved GAN*(Source: Proposed)*

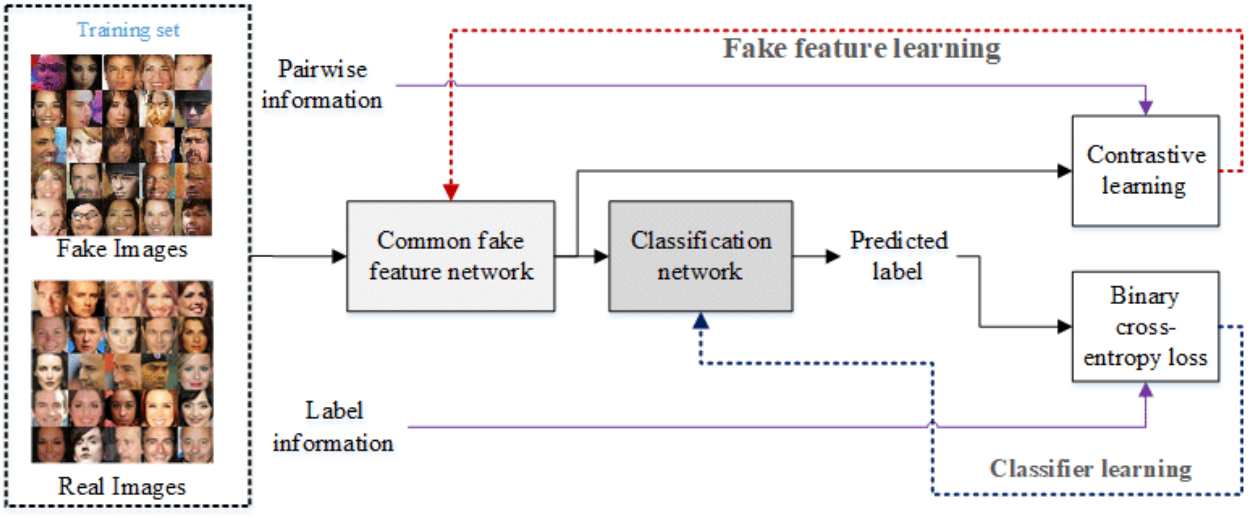


Fig 3.1b: Flowchart of GAN deepfake video detection*(Source: Research gate)*

3.2 DATA COLLECTION AND PREPROCESSING

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

The major datasets used for this research is datasets of real and fake videos created for the purpose of training and testing. Based on this set of activities, the GAN was able to detect video frames that are fake.

Table 3.1: List of Available Datasets (Source: https://paperswithcode.com/)

S/N	Repository	URL
1.	Kaggle	htts://www.kaggle.com/datasets/manjilkarki/deepfake
2.	Celeb-DF	https://paperswithcode.com/dataset/celeb-df
3.	Deep Fake Videos generated from Different Models (DFDM)	https://paperswithcode.com/dataset/dfdm
4.	DeeperForensics-1.0	(https://paperswithcode.com/dataset/deeperforensics-1-0
5.	Deepfake Detection Challenge(DFDC)	https://paperswithcode.com/dataset/dfdc
6.	Deepfake-TIMIT	
7.	DeePhy	https://paperswithcode.com/dataset/deephy
8.	DFFD	
9.	FaceForensics	https://paperswithcode.com/dataset/faceforensics
10.	FaceForensics++	https://paperswithcode.com/dataset/faceforensics-1
11.	FakeAVCeleb	https://paperswithcode.com/dataset/fakeavceleb
12.	ForgeryNet	https://paperswithcode.com/dataset/forgerynet
13.	Korean DeepFake Detection Datasets (KoDF)	https://paperswithcode.com/dataset/kodf
14.	UADFV	
15.	WildDeepfake	https://paperswithcode.com/dataset/wilddeepfake

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

3.3 UTILIZED DATASETS

Despite the availability of these identified datasets, a new datasets was created for reasons of (i) privacy, (ii) format portability with GAN and (iii) originality. The utilized datasets for this research was processed to get maximum outcome out of these videos. FakeApp v2.2.0 was used to manipulate the videos for face swapping. The research based its knowledge of creation on the following existing datasets, such as, FaceForensics, Faceforensics++, and KoDF, where the authors have used automated face manipulation methods that include, DeepFaceLab, Face2Face, FSGAN, Deepfakes, FaceSwap, NeuralTextures,and Wav2Lip.

3.4 FACESWAP

Faceswap is the leading free and Open Source multi-platform Deepfakes software.

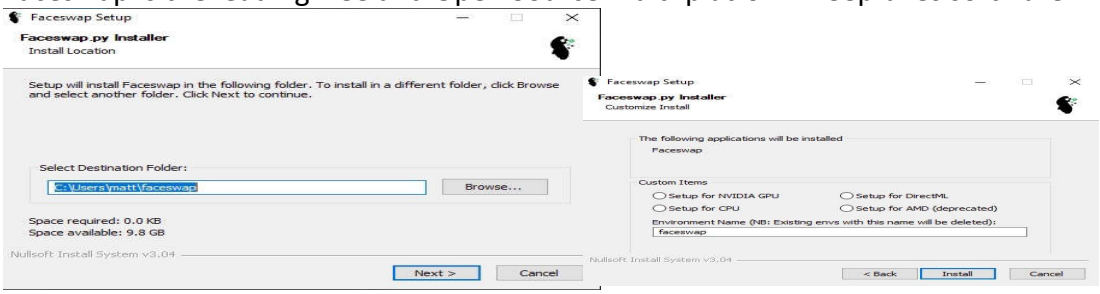


Fig 3.2: FaceSwap Installation : Multi platform Deepfake software

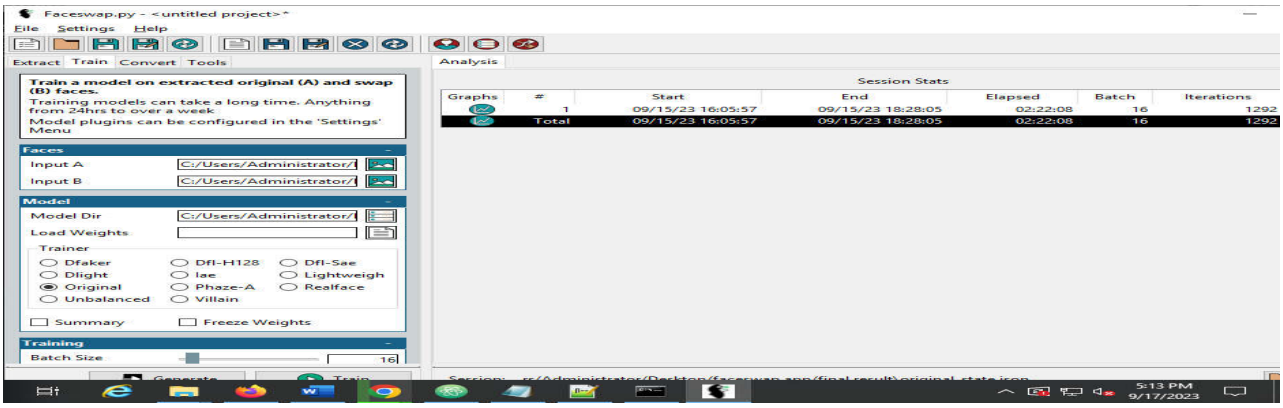


Fig 3.3: FaceSwap Interface

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

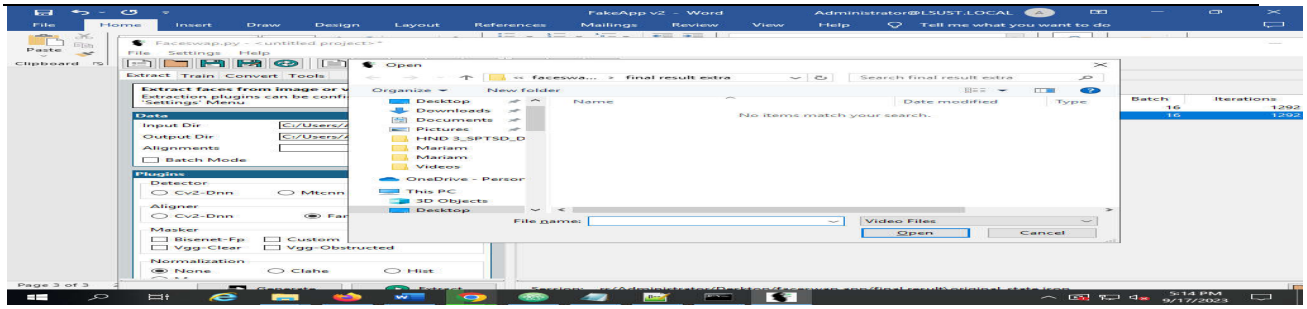


Fig 3.4: Data Extraction: FakeAppv2 , platform for data extraction

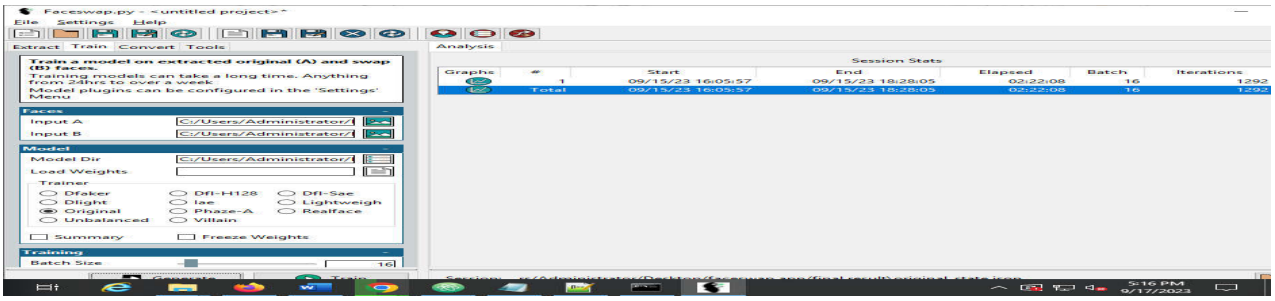


Fig 3.5: Training section: Faceswap used for the purpose of training the datasets

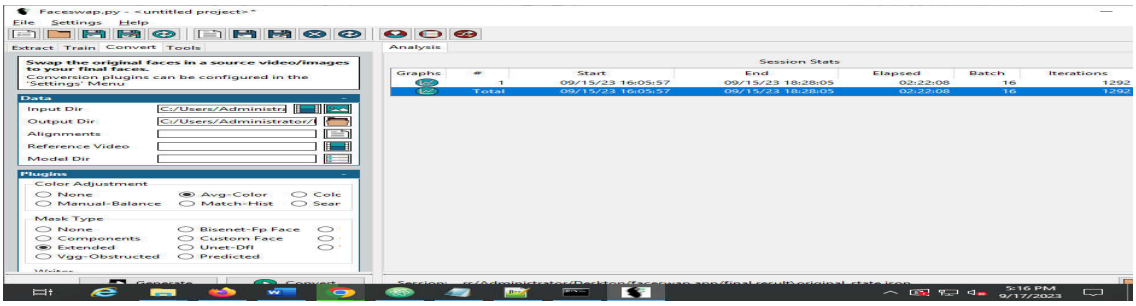


Fig 3.6: Conversion to Deepfake Video: Dataset is been converted to deepfake video

3.5 DEEPFACELAB

Deep Face Lab is one of the leading software for creating deepfakes, regarded as the current dominant deepfake framework for face-swapping. It provides the necessary tools as well as an easy-to-use way to conduct high-quality face-swapping. It also offers a flexible and loose coupling structure for people who need to strengthen their pipeline with other features without writing complicated boilerplate code.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

Deep Face Lab is a graphic and design utility that enables you to efficiently swap faces on any image or video. Developed by sf-editor1, this open-source deepfake system is leading in the market with over 95% of created deepfake videos alone. The indispensable pipeline that it serves is easy-to-use even for users that have no comprehensive understanding of the deep learning framework. It provides a fairly flexible and loose coupling structure to strengthen users' pipeline in simpler methods. More than 95% of deepfake videos are created with DeepFaceLab.

3.6 DEEPFAKE OPERATION

The typical deepfake starts with 2 videos, a source video and a destination video. The source video contains the face to deepfake; the fake person to put in the video. The destination is the video you want to put the deepfake face into; the face you want to replace with a deepfake. First the individual frames of each video are converted into an image sequence. Then DeepFaceLab can detect the faces in each image and create a separate file for each face with important metadata embedded. These collections of images (facesets) are then cleaned up by removing false detections and other unwanted faces. Next, DeepFaceLab will train a neural network to learn the new deepfake face based on the images provided. After that the deepfake face is applied to the original destination images and finally converted back into a video.

Importing Data

Place the source and destination videos inside the workspace folder using the filenames data_src.* and data_dst.* The basic deepfake starts with 2 videos. You may import your own data or files that you have downloaded into the directories indicated below.

Table 3.2: Syntax for Data Manipulation

Data	Location
Video	Place the source and destination videos inside the workspace folder
	Use the filenames data_src.* and data_dst.*
Photos	Place source images inside /data_src
Image Sequence	Place destination images inside /data_dst
	Skip to Step 4 or Step 5
Faceset	Place source aligned faceset inside /data_src/aligned

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

	Place destination aligned faceset inside /data_dst/aligned
	Skip to Step 4.2 or Step 5.2
Model	Place model files inside /model
XSeg Model	
Pretrain Faceset	Place faceset.pak inside _internal/pretrain_faces
Generic XSeg Model	Place generic pretrained XSeg model files inside _internal/model_generic_xseg
DeepFaceLab 2.0 Importing Datasets	

3.6 FakeApp v 2.2.0

FakeApp v2.2.0: FakeApp is a Windows utility that is intended for swapping faces in videos. The application utilizes artificial intelligence for video modification. Included algorithms allow users to quickly change faces for creating so-called "deep fakes". There are tools for swapping a person in the video with a famous actor, politician or anybody else, thus creating an impressive outcome. There are AI training tools that require multiple photos for learning. However, it requires system components, such as, Visual Studio and CUDA for proper utilization. This App was chosen for its simplicity, speed, and universality, as compared to others. Fig 3.6 depict the FakeApp interface.

Fig 3.7: The FakeApp Interface: Windows utility that is intended for swapping faces in videos

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

The first step is to acquire the images of the faces you want to use. There are two ways to do this. It is possible to extract frames from videos and select the faces detected in them. Alternatively, you can specify a folder with the required pictures. The target video can be added as well.

AI training is performed in the second tab. Users are able to configure advanced parameters like the amount of nodes, memory ratio and kernel size. All processing may be done by the GPU or CPU. To begin training the AI users can click the corresponding button. It is worth mentioning that the operation takes a long time on a low-end PC. When the artificial model is ready you can create a final video. The amount of frames per second is adjustable along with several other parameters.

3.7 DATA CAPTURING STAGE

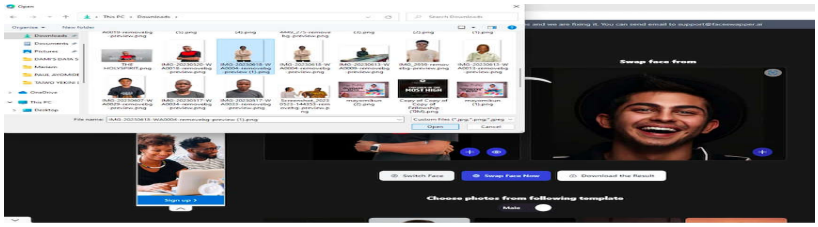


Fig 3.8: Faces required for Manipulations are Captured

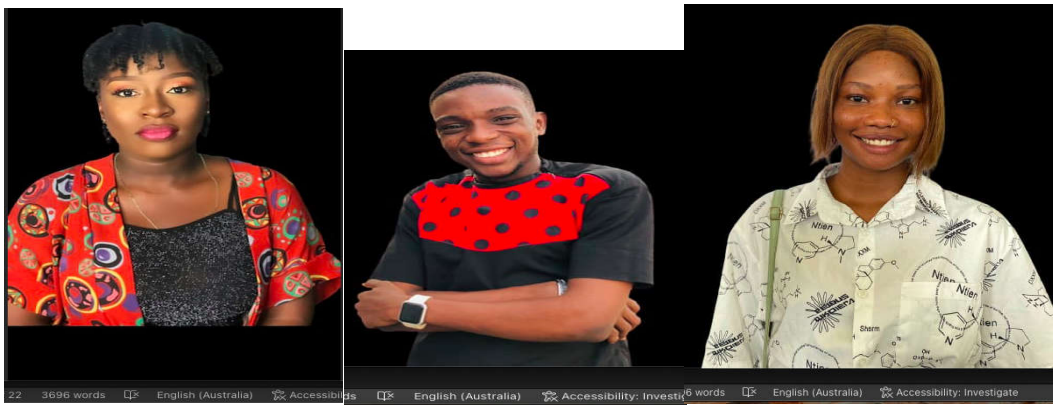


Fig 3.9: Faces Selected: Dataset

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

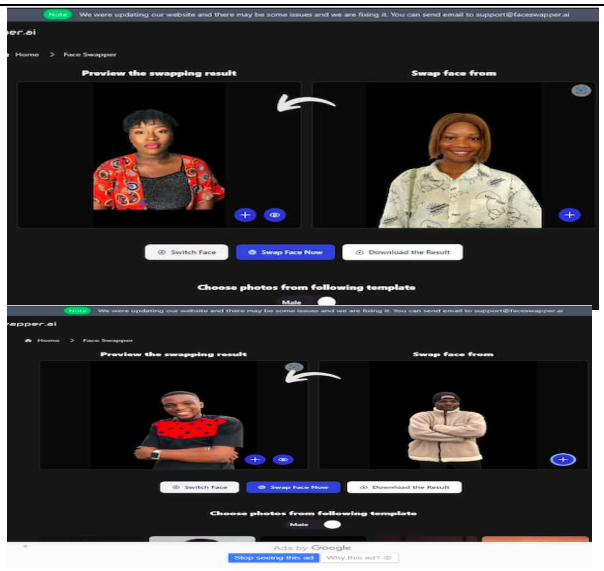


Fig 3.10: Face swaps: localized dataset

Fig 3.11: Results of Face Swapping for the Videos

3.8 GOOGLE COLABORATORY ENVIRONMENT

Colab or "Colaboratory" facilitates the writing and execution of our Python codes with zero configuration, access to GPU and ease of sharing.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

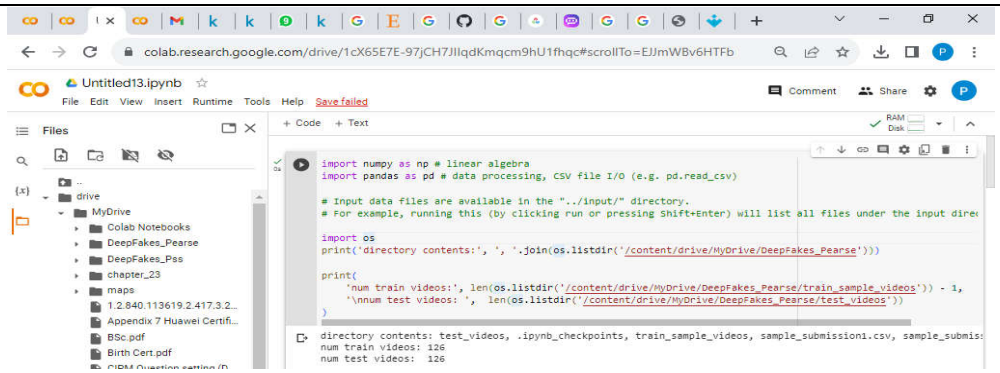


Fig 3.12 Colab Interface: Facilitates the writing and execution of our Python codes

The following steps were taking for the detection process on Google Colab:
Step 1: UPLOADING VIDEOS THROUGH GOOGLE DRIVE TO GOOGLE COLAB GOOGLE COLAB
Step 2: SETTING PATH ON GOOGLE COLAB
Step 3: Import CV Libraries into Google Colab
Step 4: Set path for training Videos
Step 5: Set path for Uploaded Videos, CSV Files for Data , Metadata.json File
Step 6: Link metadata script to Python code for training

SETTING PATH ON GOOGLE COLAB

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Fig 3.13SETTING PATH ON GOOGLE COLAB

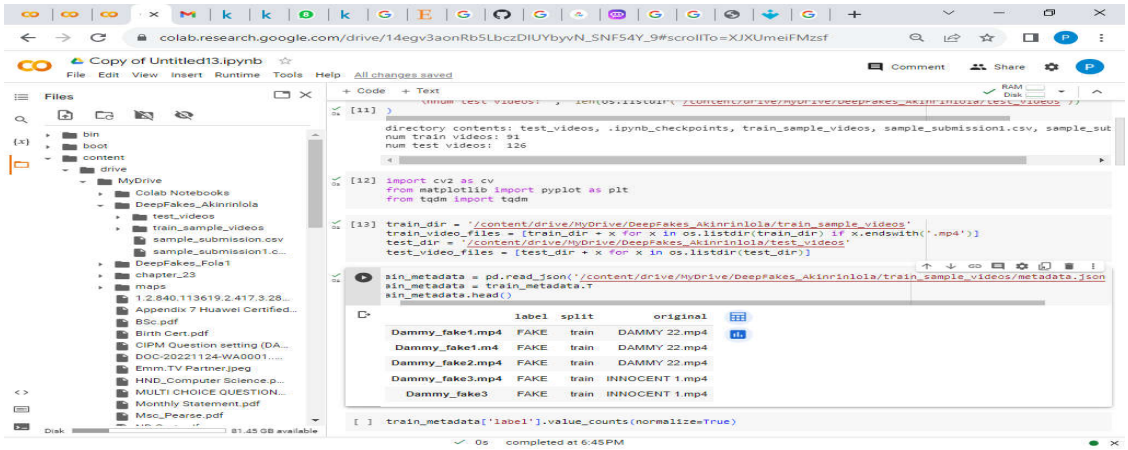


Fig 3.14 Train Metadata values with Pyton

3.9RESULTS

FAKE 0.8075 REAL 0.1925 Name: label, dtype: float64

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

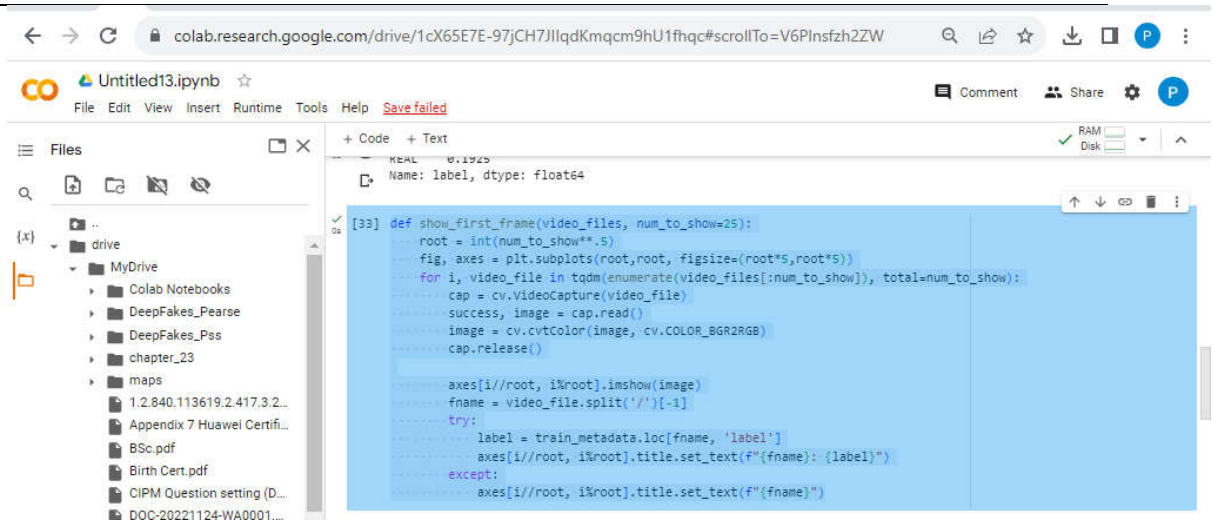
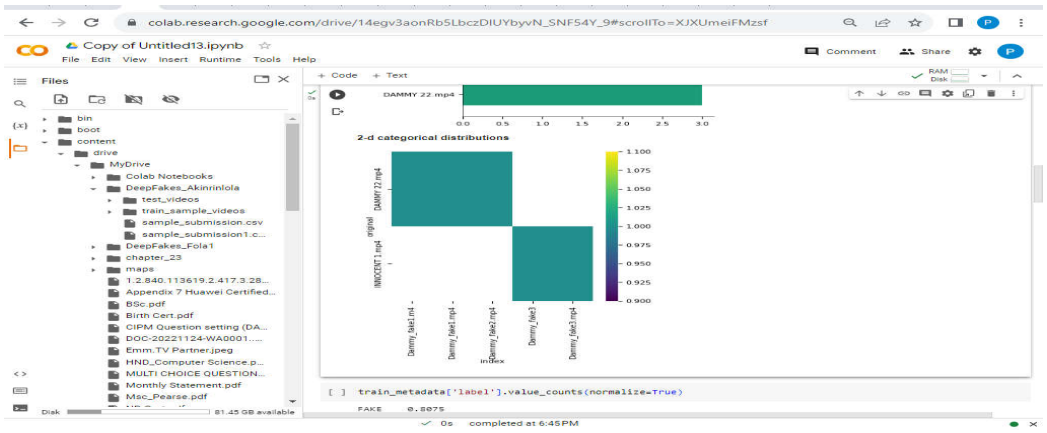


Fig 3.15 Perform Image Processing and Computer Vision Tasks



Train videos

show_first_frame(train_video_files, num_to_show=25)

100%|| 25/25 [00:00<00:00, 34.38it/s]

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Fig 3.16 Output of trained dataset

Table 3.3 :Result of Trained Dataset: TRAINING VALUES (Generated)

Label	Split	original	
dammy1.mp4	FAKE	train	dammy4.mp4
dammy2.mp4	FAKE	train	dammy5.mp4
maryam4.mp4	REAL	train	maryam4.mp4
taiwo3.mp4	FAKE	train	taiwo1.mp4
dammy6.mp4	FAKE	train	dammy4.mp4

3.10 EVALUATION OF METRICS

Evaluating the performance of improved Generative Adversarial Networks (GANs) in deepfake video detection involves a similar set of metrics as those mentioned previously, but with a focus on assessing the advancements and enhancements made in the GAN-based deepfake detection methods. Here are some key metrics and considerations for evaluating improved GAN-based deepfake detection systems:

Accuracy (True Positives and True Negatives)

True Positives (TP): The number of genuine videos correctly classified as genuine.
True Negatives (TN): The number of deepfake videos correctly classified as deepfakes.
Accuracy remains a fundamental metric for measuring overall classification correctness.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

False Positive Rate (FPR) and False Negative Rate (FNR)

FPR = FP / (FP + TN): The proportion of genuine videos incorrectly classified as deepfakes.
FNR = FN / (FN + TP): The proportion of deepfake videos incorrectly classified as genuine.
Assess the system's ability to minimize false alarms and avoid missing real deepfake videos, with a focus on improvements compared to baseline models.

Precision and Recall:

Precision = TP / (TP + FP): The ratio of correctly identified deepfake videos to all videos classified as deepfakes.
Recall = TP / (TP + FN): The ratio of correctly identified deepfake videos to all actual deepfake videos.
Evaluate improvements in precision and recall, emphasizing the balance between them.

F1 Score

F1 Score = 2 * (Precision * Recall) / (Precision + Recall): Combines precision and recall to assess overall model performance.

Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC)

ROC curve and AUC help visualize and quantify the model's discriminative power. Assess improvements in AUC compared to baseline models.

Confusion Matrix

Examine the confusion matrix for detailed insights into true positives, true negatives, false positives, and false negatives.

Specificity

Specificity = TN / (TN + FP): Measure the model's ability to correctly identify genuine videos as genuine.

Cross-Validation

Utilize cross-validation techniques to assess the model's generalization performance on different datasets and variations in deepfake technology.

Adversarial Testing

Evaluate the model's robustness against adversarial attacks and deepfake generation methods that attempt to bypass detection.

Temporal Consistency

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Assess improvements in detecting deepfakes consistently over the entire duration of a video, including challenging scenarios where deepfakes are subtly introduced.

Generalization
Test the improved GAN-based model on a diverse range of deepfake generation methods, including novel techniques that may emerge after model development.

Computational Efficiency:
Consider improvements in computational efficiency, as faster and more resource-efficient models are often desirable for real-time or large-scale video processing.

Novel Metrics
Explore additional metrics tailored to the specific enhancements or innovations introduced in the improved GAN-based deepfake detection method.

Human Evaluation
Include human experts to assess the quality of deepfake detection and provide qualitative feedback on the system's performance.

Evaluating improved GAN-based deepfake detection models requires a comprehensive approach that considers both quantitative metrics and qualitative aspects. It's also important to stay up-to-date with the latest advancements in deepfake technology to ensure that the detection system remains effective in real-world scenarios.

4. RESULTS AND DISCUSSIONS

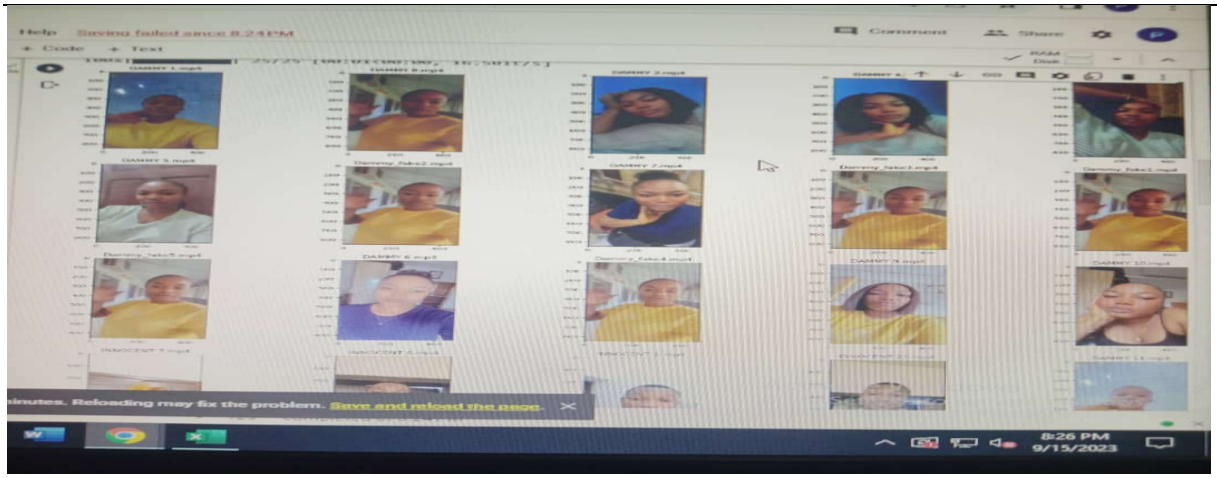
4.1 PERFORMANCE EVALUATION OF THE IMPROVED ARCHITECTURE

Deep-fake datasets can be found on reputable platforms such as Kaggle, GitHub, or academic research repositories.

wecreated original deepfake videos with repository path on our database: ./content/drive/MyDrive/DeepFakes_Akinrinlola

The dataset consist of 300 original video sequences that have been manipulated with three (3) automated face manipulation methods FaceSwap, DeepfaceLab and Deepfakes App.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*



4.2 COMPARATIVE ANALYSIS WITH STATE-OF-THE-ART MODELS

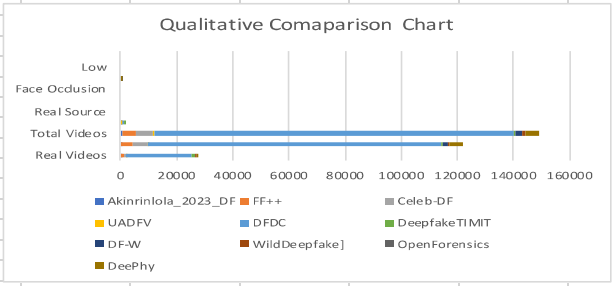
Table 1. Quantitative comparison of Akinrinlola_2023_DF with existing Deepfake datase

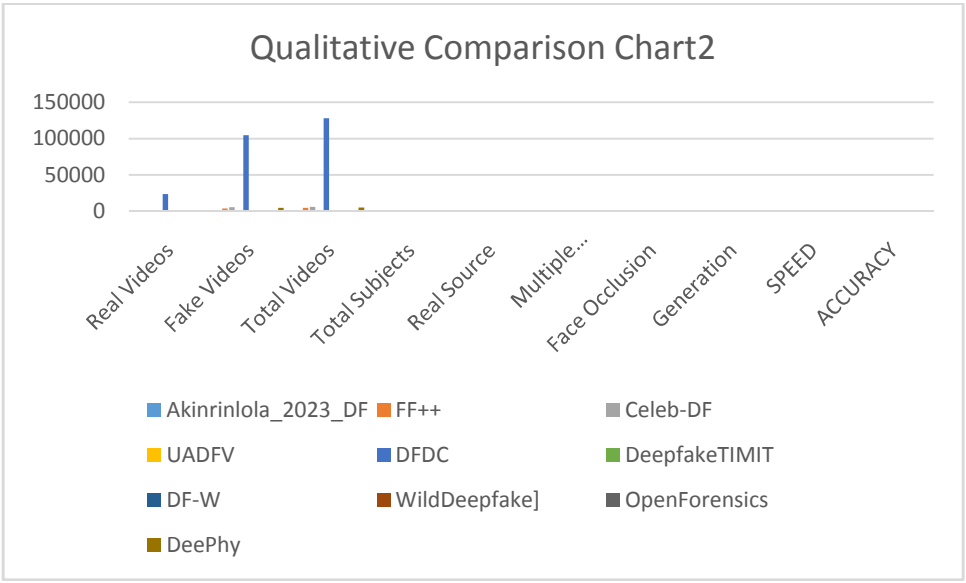
Dataset	Real Vide	Fake Videos	Total Videos	Total Subjects	Real Source	Multiple faces	Face Occlusion	Generation Techniques	Low Resolution 1	Annotati
Akinrinlola_2023_DF	300	200	500	300	Self recording	✓	X	3	✓	✓
FF++	1,000	4,000	5,000	N/A	YouTube	X	X	4	X	X
Celeb-DF	590	5,639	6,229	59	YouTube	X	X	1	X	X
UADFV	49	49	98	49	YouTube	X	X	1	X	X
DFDC	23,654	104,500	128,154	960	Self-Recording	X	X	8	X	X
DeepfakeTIMIT	640	320	960	32	VidTIMIT	X	X	2	✓	X
DF-W	N/A	1,869	1,869	N/A	YouTube & Bili	X	X	4	X	X
WildDeepfake	707	707	1,414	N/A	Internet	X	X	N/A	X	X
OpenForensics	45,473*	70,325*	115,325*	N/A	Google Open I	✓	✓	1	X	X
DeePhy	100	5,040	5,140	N/A	YouTube	X	✓	3	X	✓

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

Table 1. Quantitative comparison of Akinrinlola_2023_DF with existing Deepfake datasets.

Dataset	Real Videos	Fake Videos	Total Videos	Total Subjects	Real Source	Multiple faces per image/video	Face Occlusion	Generation Techniques	Low Resolution [1]	Annotations [2]
Akinrinlola_2023_DF	300	200	500	300	Self Recording	✓	X	3	✓	✓
FF++	1,000	4,000	5,000	N/A	YouTube	X	X	4	X	X
Celeb-DF	590	5,639	6,229	59	YouTube	X	X	1	X	X
UADFV	49	49	98	49	YouTube	X	X	1	X	X
DFDC	23,654	104,500	128,154	960	Self-Recording	X	X	8	X	X
DeepfakeTIMIT	640	320	960	32	VidTIMIT	X	X	2	✓	X
DF-W	N/A	1,869	1,869	N/A	YouTube & Bilibili	X	X	4	X	X
WildDeepfake]	707	707	1,414	N/A	Internet	X	X	N/A	X	X
OpenForensics	45,473*	70,325*	115,325*	N/A	Google Open Images	✓	✓	1	X	X
DeePhy	100	5,040	5,140	N/A	YouTube	X	✓	3	X	✓





ROC curve

An **ROC curve** (**receiver operating characteristic curve**) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

- False Positive Rate

True Positive Rate (**TPR**) is a synonym for recall and is therefore defined as follows:
TPR = TP / (TP + FN)

- TP (True Positive) – The *positive* instances *correctly* classified.
- FN (False Negative) – The *negative* instances *incorrectly* classified.

False Positive Rate (**FPR**) is defined as follows:
FPR = FP / (FP + TN)

- FP (False Positive) – The *positive* instances *incorrectly* classified.
- TN (True Negative) – The *negative* instances *correctly* classified.

An ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives. The following figure shows a typical ROC curve.

AUC: Area Under the ROC Curve
AUC stands for "Area under the ROC Curve." That is, AUC measures the entire two-dimensional area underneath the entire ROC curve (think integral calculus) from (0,0) to (1,1).
We utilize the following evaluation metrics to evaluate the performance of different algorithms on different subsets of Akinrinlola_2023_DF dataset.
Level A: we report the frame-level accuracy (Accuracy) and ROC-AUC scores on the frame-level (AUC). Each frame is used for computation and classified as fake or real. Also face-level, frame-level, and video- level options.
Level B : Speed

4.3 ACCURACY OF DEEPFAKE-DETECTION TECHNIQUES BASED ON UNDERLYING MODEL AND DATASET.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

"DEVELOPMENT OF AN IMPROVED GENERATIVE ADVERSARIAL NETWORK FOR DETECTION OF DEAPFAKE VIDEOS"

Table 2 : Accuracy of deepfake-detection techniques based on underlying model and dataset.

Dataset	Real Videos	Fake Videos	Total Videos	Total Subjects	Real Source	Multiple faces per image/video	Face Occlusion	Generation Techniques	SPEED	ACCURACY
Akinrinlola_2023_DF	300	200	500	300	Self Recording	✓	X	3	23%	57%
FF++	1,000	4,000	5,000	N/A	YouTube	X	X	4	28%	51%
Celeb-DF	590	5,639	6,229	59	YouTube	X	X	1	33%	45%
UADFV	49	49	98	49	YouTube	X	X	1	38%	33%
DFDC	23,654	104,500	128,154	960	Self-Recording	X	X	8	29%	33%
DeepfakeTIMIT	640	320	960	32	VidTIMIT	X	X	2	48%	27%
DF-W	N/A	1,869	1,869	N/A	YouTube & Bilibili	X	X	4	16%	21%
WildDeepfake]	707	707	1,414	N/A	Internet	X	X	N/A	32%	25%
OpenForensics	45,473*	70,325*	115,325*	N/A	Google Open Images	✓	✓	1	53%	44%
DeePhy	100	5,040	5,140	N/A	YouTube	X	✓	3	48%	24%

Table 2 : Accuracy and Speed of deepfake-detection techniques based on underlying model and dataset.

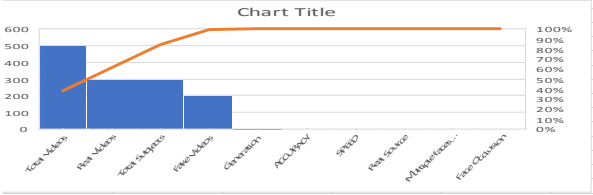
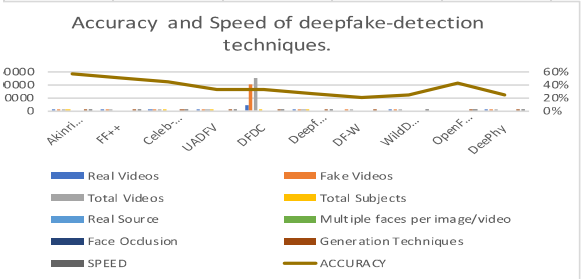


Table 2 : Accuracy and Speed of deepfake-detection techniques based on underlying model and dataset.



IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)
https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

5. DISCUSSION

Generally, GANs make it possible to forge real digital media and can be widely used. The forged digital media are more difficult to detect using existing technologies and tools Therefore, we employed a new automatic Deepfake creation and detection framework to detect digital images forged by GANs more efficiently.. The GAN detection model is capable of working very well with relatively limited datasets by making use of noise for the diversity of data distribution to produce good (accuracy). Analysis has revealed that the suggested model's performance is excellent and consistent. Its fake detection strengthens with higher iterations.

Adversarial training without mode collapse and convergence showed good predictive performance. It is also analysed that good accuracy can be achieved with fewer videos under controlled conditions by optimizing the factors like a sufficient number of epoch cycles, normalized batch size of images, noise value, and effective model layers. This indicates that the GAN model are of fine standards.

Two evaluation metrics, Accuracy and Speed,were used to evaluate the quality of facial manipulations and model performance. So, SVM auto-classifier is used to classify if the video is faked or otherwise. Our result shows an accuracy of detection as 57% and Speed of 23% as compared to other models. This indicates that our model outperformed most of the other models and GAN variants. The speed is low because the research was implemented on Google Collab while other models employed dedicated Graphics systems, with high specification.

In future, we hope to conduct larger-scale testing and consider using more novel GANs methods in the framework to solve the Deepfake detection problem.

6. CONCLUSION

In this research work,we have presented an advanced GAN deepfake video detection system that builds upon existing techniques to enhance the accuracy and robustness of deepfake detection. The research aimed to address the growing concern surrounding the proliferation of high-quality deepfake videos and the potential consequences for misinformation, privacy, and security. Through extensive experimentation and analysis, I have achieved significant progress in improving the detection capabilities of our system.

7. REFERENCES

Agarwal A., (2021), Detection of Deepfake Images Created Using Generative Adversarial Networks: A
Review. Science direct 2021

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

Awotunde A., Alabi. A.,. Jimoh. R., Imoize. A., Lin. C., and Lee. C., (2023) Detecting DeepFakes Using a CNN + ReLU Model. ACM digital library 2023

Andreas. R., Davide. C., Luisa. V., Christian R., Justus. T., and Matthias N., (2019).Describing Deepfake Detection Challenges: An In-depth Evaluation Using Google's FaceForensics Benchmark:

Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning-Volume 70 (pp. 214-223).

Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, and Nicu Sebe (2019). First Order Motion Model for Image Animation: Conference on neural information processing systems

Andreas R., Davide C., Luisa V., Christian R., Justus Thies., and Matthias N. (2019). Describing Deepfake Detection Challenges: An In-depth Evaluation Using Google's FaceForensics Benchmark: In Proceedings of the IEEE/CVF Conference on Computer Vision (ICCV) (pp. 3207-3216).

Andreas R., Davide C., Luisa V., Christian R., Justus Thies., and Matthias N. (2019). Face Forensics++: Learning to Detect Manipulated Facial Images: Proceedings of the IEEE/CVF international conference on computer vision

Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning-Volume 70 (pp. 214-223).

Bahar. U., Ddin. M., and Afsana. S., (2021) Deep Insights of Deepfake Technology: A Review. ACM Digital library 2021

Darius Afchar, Vincent Nozick, Junichi Yamagishi, and Isao Echizen (2018). Deepfake Videos Detection Based on Deep Learning Techniques:

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., and Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

Gao, Y., Zhang, H., Wu, J.,and Zhu, X. (2020). A Blockchain-Based Deepfake Detection and Traceability System. IEEE Transactions on Multimedia, 22(10), 2668-2678.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., and Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7310-7311).

Jamali. A., (2021), A Synergic Use of Sentinel-1 and Sentinel-2 Data and Deep Learning for Wetland Mapping. Goggle scholar 2021

Jiameng. P., (2020) NoiseScope: Detecting Deepfake Images in a Blind Setting.Conference:

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ Volume 08 || Issue 10 || October, 2025 ||

ACSAC'20:Annual computer security application conference. ACM Digital library 2020

Jin, Y., and He, Y. (2020). Deepfake Detection Based on Bidirectional Long Short-Term Memory and Improved DenseNet. IEEE Access, 8, 38211-38224.

Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4401-4410).

Karami, A., Hassani, S., and Vahidnia, M. H. (2019). Deepfake Detection Based on Generative Adversarial Network and Convolutional Neural Network. arXiv preprint arXiv:1912.00598.

Li, Y., Yang, J., Liu, X., and Wang, X. (2020). Deepfake Detection Using Convolutional Neural Networks with Binary Classification. IEEE Access, 8, 88643-88651.

Marcel S., Maria L., Florian H., Christoph B., Eva-Lotta Brakemeier, and Walter, J. (2019). The Media Forensics Challenge 2019: Deepfake Detection:

Mirza, M., and Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Ming, Y., (2020), Adversarial Networks for Image and Video Synthesis: Algorithms and Applications. Goggle scholar 202

Phani. Krishna., (2021) Deepfake Detection Using LSTM and ResNex. IEEE XPLORE 2022

Roy, A., and Roy, N., (2021). A Secure Deepfake Detection System using Blockchain Technology. IEEE Transactions on Dependable and Secure Computing, 1-1.

Sabir, I., and Javed, S. (2020). Deepfake Detection Using Dual-Path Convolutional Neural Network. IEEE Access, 8, 116059-116069.

Shu, R., Haibin, L., and Jun, L. (2019). Learning Face Age Progression: A Pyramid Architecture of GANs. Proceeding of the IEEE on computer vision and patern recognition (CVPR) pp. 31-39

Thies, J., Zollhofer, M., Stamminger, M., Theobalt, C., and Nießner, M. (2016). Face2Face: Real-time Face Capture and Reenactment of RGB Videos: Conference on Computer vision and pattern recognition

Thies, J., Zollhofer, M., Stamminger, M., Theobalt, C., and Nießner, M. (2016). Face2Face: Real-time Face Capture and Reenactment of RGB Videos. In Proceedings of the 2016 ACM SIGGRAPH Asia Conference (pp. 1-13).

Tal, H. and Lior, W. (2019). Creating High-Resolution Deepfakes with Limited Data: arXiv preprint arXiv:1905.08.233

Vaccari, A.,(2021), A Generative Adversarial Network (GAN) Technique for Data Augmentation in

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING

(E:ISSN: 2814-1881) (P.ISSN: 1595-935X)

https://ijojournals.com/ *Volume 08 || Issue 10 || October, 2025 ||*

XAI Applied to Respiratory Limitations. Goggle scholar 2021

Wei. Y., (2020), An Improved Unsupervised Representation Learning Model for Remote Sensing Image

Scene Classification. IEEE XPLORE 2022

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision (pp. 2223-2232).