DeepFake Detection

Archita Todi - 20BCE0529, Parth Maheshwari - 20BCE0220

Mohit Agnihotri - 20BCE2513 , Kajal Patro - 20BCE0736

Abstract

Deepfake detection is formulated as a hypothesis testing problem to classify an image as genuine or GAN-generated. A robust statistics view of GANs is considered to bound the error probability for various GAN implementations in terms of their performance. The bounds are further simplified using a Euclidean approximation for the low error regime. Lastly, relationships between error probability and epidemic thresholds for spreading processes in networks are established.

Better generative models and larger datasets have led to more realistic fake videos that can fool the human eye but produce temporal and spatial artifacts that deep learning approaches can detect. Most current Deepfake detection methods only use individual video frames and therefore fail to learn from temporal information.

Keywords: `Deepfake, GANs, Euclidean approximation, fake videos, deep learning`

1. Introduction

In a narrow definition, deepfakes are synthetic media in which a person in an existing image or video is replaced with someone else's likeness. While the act of faking content is not new, deepfakes leverage powerful techniques from machine learning and artificial intelligence to manipulate or generate visual and audio content with a high potential to deceive.

The deepfake videos can have a heavy social, political and emotional impact on individuals, as well as on the society. The first step to preempt such misleading

deepfake videos from social media is to detect them.

AI based tools can manipulate media in increasingly believable ways, for example by creating a copy of a public person's voice or superimposing one person's face on another person's body.

In this project we intend to identify deepfake videos and prevent them from getting circulated on various social media platforms for misuse. Using hdeeplearning methods such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), and Long short-term memory (LSTM) we can tell whether a video has been tampered with . This system only requires a recording of the video to function and hence can be accessible to the majority of people. Our projet will help in solving of false rumours about people and prevent defaming of famous personalities . The way our project differs from the existing ones available is because it's way more efficient and accurate than the others as we use Our model uses CNN, Convolutional Neural Network for pattern matching which saves computational power..

1.1 **Problem**

Deepfakes are a relatively new off-the-shelf video editing tool that allows anyone to anyone to swap two identities in a single video. Aside from Deepfakes, a number of GAN-based face swapping algorithms have also been published. To counter this looming danger, we have come up with this proposition to help the general mass.

Deepfake's face forgery is frequently used on the internet, causing serious societal implications. How to detect such fraudulent material has recently become a popular research area as well. The current deepfake algorithms present require a lot of computational power and so to provide a solution for it we thought of implementing CNN, Convolutional Neural Network for pattern matching.

1.2 **Motivation**

It is threatening to world security when deepfake methods can be employed to create videos of world leaders with fake speeches for falsification

purposes. Deepfakes can be abused to cause political or religious tensions between countries, fool the public and affect results in election campaigns, or create chaos in financial markets by creating fake news. It can be even used to generate fake satellite images of the Earth to contain objects that do not exist to confuse military analysts, e.g., creating a fake bridge across a river although there is no such a bridge. This can be misleading and alarming to people worldwide

Deepfake is no longer a high-priced technology just for the film industry. People are now better able to take advantage of deepfakes because of developments in synthetic media production. When technology first became ubiquitous in 2017, individuals created amusing celebrity videos. However, as deepfakes got more popular, more people began to use them for nefarious purposes. While celebrities are used to receiving a lot of attention in both positive and negative ways, deepfake technology has started to be utilized in some terrifying ways that may worry ordinary people.

Deepfakes began to be exploited by online predators as well. People began leveraging the ability to substitute anyone's face in an image or video to make pornographic content without their consent. And the deepfake technology allows them to do just that; all they need is a profile image on social media to generate fake material.

1.3 **Objectives**

- To create an easy-to-use interface for detecting whether a video is fake or not.
- To embed features like showing by how much percentage of the threshold is the video fake.
- To implement an ideal face recognition system that has high accuracy and minimal latency.

- To have real-time upgradation of the output threshold values depending on how the face changes.
- To create a fun and simple GUI for the users so that they can easily interact with all the functions of the application.

1.4 Possible Techniques for Proposed System

Deep learning is an effective and useful technique that has been widely applied in a variety of fields, including computer vision, machine vision, and natural language processing. Deepfakes uses deep learning technology to manipulate images and videos of a person that humans cannot differentiate them from the real one.

Several techniques based on deep learning have been proposed including: 1) convolutional neural network (CNN); 2) recurrent neural network (RNN); 3) long short-term memory (LSTM).

Deepfake is a technique that uses the Generative adversarial networks (GANs) methods to generate fictitious photographs and videos. In this section, we fist give an overview of the current applications and tools to create deepfake image and videos.

Deepfake Generation

Generative adversarial networks (GANs) are a form of deep neural network that has been commonly used to generate deep fake. One advantage of GNAs is that it capable to learn from a set of training data set and create a sample of data with the same features and characteristics. For example, GANs can be used to swipe a "real" image or the video of a person with that of a "fake" one.

Deep learning has achieved great success in deepfake detection. Below, we first discuss the Image Detection models using deep learning technologies and then Video Detection models are explained.

Image Detection Models

Different methods have been proposed to detect the GAN generated im-

ages using deep networks. Tariq et al. [24] suggested neural network-based methods for detecting fake GAN videos. This method employs pre-processing techniques to analyses the statistical features of image and enhances the detection of fake face image created by humans

Video Detection Models

For the last years, deep learning methods have been successfully applied for fake image detection. However, the current deep learning methods for image cannot be directly applied for fake videos detection due to the availability of significant loss of frame information after video compression. In the subsection below, we have divided the related work in deepfake video detection into two main categories: biological singles analysis and spatial and temporal features analysis.

1) Biological Singles Analysis

This method uses a convolutional neural network (CNN) with a recursive neural network (RNN) to discover the physiological signals.

2) Spatial and Temporal Features Analysis

Video manipulation can be carried out on multiple frame-level features. Recently, many researches have shown that analyzing the temporal sequence between frames can successfully help to discriminate the real video or the fake one.

2. Literature survey

Sr. No.	Research Paper	Author	Literature Works	Research Gaps
1.	The DeepFake Detection Challenge (DFDC) Dataset	Brian Dolhansky, Joanna Bitton, Ben Pflaum, Jikuo Lu, Russ Howes, Menglin Wang, Cristian Canton Ferrer	Hypothesis is that GAN-like methods work well in limited settings with even lighting, such as news rooms, interviews, or controlled-capture videos.	Deepfake detection is extremely difficult and still an unsolved problem, a Deepfake detection model trained only on the DFDC can generalize to real "in-the-wild" Deepfake videos
2.	Multi-attentional Deepfake Detection	Hanqing Zhao, Wenbo Zhou, Nenghai Yu	Early works detect the forgery through visual biological artifacts, e.g., unnatural eye blinking or inconsistent head pose.	Most existing methods treat the deepfake detection as a universal binary classification problem. They focus on how to construct sophisticated feature extractors and then a dichotomy to distinguish the real and fake faces.
3.	WildDeepfake: A Challenging Real-World Dataset for Deepfake Detection	Bojia Zi, Minghao Chang , Jingjing Chen , Xingjun Ma, Yu-Gang Jiang	While existing works are mostly focused on identifying the flaws in a face image/sequence, most wild deepfake videos are carefully tuned to have no obvious flaws	Since the fake videos are crafted by the researchers rather than real-world deepfakes uploaded to the internet, we denote the fake videos in these datasets as virtual deepfakes. Moreover, most of the source videos are filmed with a few volunteer actors in limited scenes.

4.	DEEPFAKE DETECTION: CURRENT CHALLENGES AND NEXT STEPS	Siwei Lyu	the synthesized videos can be more realistic if they are accompanied with realistic voices, which combines video and audio synthesis together in one tool.	Falsified videos created by AI algorithms, in particular, deep neural networks (DNNs), are a recent twist to the disconcerting problem of online disinformation.
5.	DeepFake Detection by Analyzing Convolutional Traces	Luca Guarnera, Oliver Giudice, Sebastiano Battiato	Many works try to reconstruct the history of an image; others try to identify the anomalies, such as the study on the analysis of interpolation effects through CFA (Color Filtering Array), analyzing compression parameters.	Being able to understand if an image is the result of a generative Neural Network process turns out to be a complicated problem, even for human eyes. They addressed the problem of fake detection as a binary classification problem for each frame of manipulated videos
6.	Generalization Of Audio Deepfake Detection	Tianxiang Chen, Avrosh Kumar, Parav Nagarsheth, Ganesh Sivaraman, Elie Khoury	In this work, we tackle this challenge from a different perspective. Instead of investigating different low level audio features, we try to increase the generalization ability of the model itself. To do so, we use large margin cosine loss function (LMCL) which was	The challenge results show that the current biggest problem in a spoofing detection system is its generalization ability. Traditionally, signal processing researchers tried to overcome this problem by engineering different low-level spectro-temporal features

			initially used for face recognition.	
7.	Joint Audio-Visual Deepfake Detection	Yipin Zhou, Ser-Nam Lim	Video deepfake detection demonstrate that there exists 'fingerprints' for different GAN frameworks that can be used to detected generated images. Many existing spoofed speech detection frameworks rely on extracting acoustic representations like MFCC , STFT and CQCC	Applying late fusion is a straight-forward operation to jointly learn with multiple modalities. We simply extract latest feature representations from the audio and video streams
8.	Learning Self-Consistency for Deepfake Detection	Tianchen Zhao, Xiang Xu, Mingze Xu	There are four common types of deepfake entire image synthesis, modification of facial attribute or expression, and face identity swap. 3D models , AutoEncoders, or Generative Adversarial Networks.	The consistency loss in PCL needs pixel-level annotation about whether a location has been modified. It is generally not available in deepfake detection datasets, on which the re-annotation could be laborious and error-prone.
9.	KoDF: A Large-scale Korean DeepFake Detection Dataset	Gyuhyeon Nam, Sungwoo Park, Gyeongsu Chae	Although face swap is the best-known method for deepfake creation, deepfake technologies are not simple equivalents of face-swap neural networks. There are a number of	The collected real clips are manually inspected for various possible defects: (1) audio-video sync problem, (2) excessive background noise, (3) utterances severely

				utterances severely hindered or stuttered, (4) extreme lighting conditions, and (5) face located far outside of the central region.
10.	M2TR: Multi-modal Multi-scale Transformers for Deepfake Detection	Junke Wang, Zuxuan Wu, Jingjing Chen, Yu-Gang Jiang	Most existing approaches combat Deepfakes with deep neural networks by mapping the input image to a binary prediction without capturing the consistency among different pixels. Decent detection results are achieved by stacked convolutions	It is worth noting using a binary classifier tends to result in overfitted models. We additionally predict the the face region as an auxiliary task to enrich the supervision for training the networks. Perceptual loss is usually used in face inpainting approaches
11.	Deepfake detection: humans vs. machines	Pavel Korshunov, Sebastien Marcel	Autoencoders and generative adversarial networks (GANs) significantly improved the quality and realism of the automated image generation and face swapping, leading to the deepfake phenomena. Many databases with deepfake videos were created to help	Many are starting to believe that the proverb 'seeing is believing' is starting to loose its meaning when it comes to digital video1 .

12.	Understanding the Security of Deepfake Detection	Xiaoyu Cao, Neil Zhenqiang Gong	In this work, we perform systematic measurement studies to understand the security of deepfake detection. The security of a deepfake detection method relies on the security of both the face extractor and the face classifier. Therefore, we perform measurement studies to understand the security of both components.	We find that an attacker can leverage backdoor attacks developed by the adversarial machine learning community to evade a face classifier. Our results highlight that deepfake detection should consider the adversarial nature of the problem.
13.	Adversarial Threats to DeepFake Detection: A Practical Perspective	Paarth Neekhara, Brian Dolhansky, Joanna Bitton	In this work, we study the vulnerabilities of state-of-the-art DeepFake detection methods from a practical stand point. We perform adversarial attacks on DeepFake detectors in a black box setting where the adversary does not have complete knowledge of the classification models.	The best performing methods model the DeepFake detection problem as a per-frame classification problem. While such methods achieve promising results in terms of detection accuracy, they are vulnerable to adversarial examples and can be evaded by adding a carefully crafted perturbation to each frame of a given input video.
14.	Deepfake Detection using	Oscar de Lima, Sean Franklin,	We are comparing the	In this paper, we aim to apply

	Spatiotemporal Convolutional Networks	Shreshtha Basu, Blake Karwoski, Annet George	performance of our video-based methods against a selection of methods that only work on the level of frames and don't learn from temporal information. Better generative models and larger datasets have led to more realistic fake videos that can fool the human eye	techniques used for video classification, that take advantage of 3D input, on the Deepfake classification problem at hand. All the convolutional networks we tested (apart from RCN) were pre-trained on the Kinetics dataset
15.	Limits of Deepfake Detection: A Robust Estimation Viewpoint	Sakshi Agarwal and Lav R. Varshney	This work gives a generalizable statistical framework with guarantees on its reliability. In particular, we build on the information-theo retic study of authentication to cast deepfake detection as a hypothesis testing problem specifically for outputs of GANs, themselves viewed through a generalized robust statistics framework.	Technologies capable of generating hyperrealistic fake images and videos are used for dating scams, "astroturfing," "catfishing," and to gain the victim's trust for blackmail, harassment, or sabotage. With an active community of developers creating free and easy tools to commoditize such technology, there is eroded trust in visual content.
16.	DeepFake Detection Based on Discrepancies Between Faces and their Context	Yuval Nirkin, Lior Wolf, Yosi Keller, and Tal Hassner	We propose a method for detecting face swapping and other identity manipulations in single images. Face swapping methods, such as	We claim that it is no coincidence that all face manipulation methods we know of do not affect the entire head: While human faces

			DeepFake, manipulate the face region, aiming to adjust the face to the appearance of its context, while leaving the context unchanged.	have simple, easily modeled geometries, their context (neck, ears, hair, etc.) are highly irregular and therefore difficult to consistently reconstruct and manipulate.
17.	Towards Solving the DeepFake Problem : An Analysis on Improving DeepFake Detection using Dynamic Face Augmentation	Sowmen Das, Selim Seferbekov, Arup Datta	Most state-of-the-art deepfake generators utilize GANs for face forgery. The GAN is trained with a dataset of face images to learn the identifying characteristics and attributes of a face. During generation, these attributes can be selectively modified to produce a different face.	Overfitting is one of the main challenges in training neural networks. Dropout is one of the most widely used regularization techniques that is used o mitigate this problem. These layers are inserted in-between any standard convolution layers where they randomly drop the activations of a fixed amount of neurons.
18.	FReTAL: Generalizing Deepfake Detection using Knowledge Distillation and Representation Learning	Minha Kim, Shahroz Tariq	This work spans different fields, such as deepfake detection, domain adaptation, knowledge distillation, and representation learning. We use FReTAL to perform domain adaptation tasks on new deepfake datasets, while minimizing the catastrophic forgetting.	As deepfake video generation techniques are continuously evolving, more types of deepfake videos will emerge in the future. Collecting and producing a large number of new deepfake samples for each dataset would be impractical. In order to perform domain adaptation, the

				model is initialized with the pre-trained weights on the source dataset.
19.	DEFAKEHOP: A LIGHT-WEIGHT HIGH-PERFORMANCE DEEPFAKE DETECTOR	Hong-Shuo Chen , Mozhdeh Rouhsedaghat , Hamza Ghani	A light-weight high-performance Deepfake detection method, called DefakeHop, is proposed in this work. Stateof-the-art Deepfake detection methods are built upon deep neural networks. DefakeHop extracts features automatically using the successive subspace learning (SSL) principle from various parts of face images.	Fake videos of the second generation are more realistic, which makes their detection more challenging. In the experiment, we focus on compressed videos since they are challenging for Deepfake detection algorithms.
20.	Not made for each other– Audio-Visual Dissonance-based Deepfake Detection and Localization	Komal Chugh, Parul Gupta, Abhinav Dhall, Ramanathan Subramanian	We propose detection of deepfake videos based on the dissimilarity between the audio and visual modalities, termed as the Modality Dissonance Score (MDS). We hypothesize that manipulation of either modality will lead to dis-harmony between the two modalities, e.g., loss of lip-sync, unnatural facial	Since fake videos are often indistinguishable from genuine counterparts, detection of deepfakes is challenging but topical given their potential for denigration and defamation, especially against women and in propagating misinformation. Part of the challenge in detecting deepfakes via AI approaches is that deepfakes

			and lip movements, etc	are themselves created via AI techniques.

Clearly, better results can be expected by strategies that take explicitly into account the temporal direction. In fact, even if current generation methods are very effective, they perform face manipulation on a frame-by-frame basis and hence may incorrectly follow the face movements. Several methods have been proposed in the literature to exploit this point. A convolutional Long Short Term Memory (LSTM) network is used to exploit such dependencies and improve upon single-frame analysis.

The work shows that CNN-generated images share some common flaws that allow one to trace their origin even on unseen architectures, datasets and training methods. The main idea is to make a very large augmentation in the training step by means of several and different post-processing operations, like blurring and compression and combinations of them, even if they are not performed at test time.

Modern sophisticated manipulations, however, are more and more effective in avoiding such pitfalls and methods which test digital integrity are by far more widespread and represent the current state of the art.

Despite the continuous research efforts and the numerous forensic tools developed in the past, the advent of deep learning, is changing the rules of the game and asking multimedia forensics for new and timely solutions. This phenomenon is also causing a strong acceleration in multimedia forensics research, which often relies itself on deep learning.

Hence, beyond reviewing the conventional media forensics approaches, a special attention will be devoted to deep learning-based approaches and to the strategies designed to fight deepfakes.

3. Proposed work

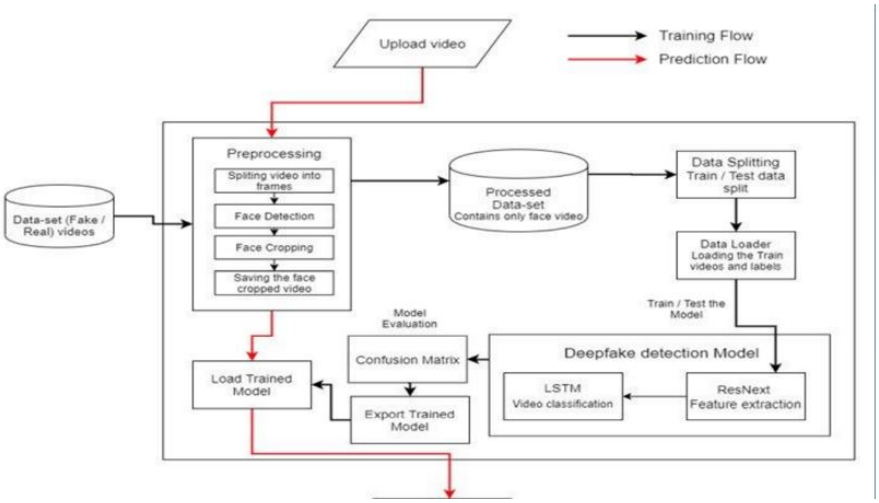
3.1 Proposal formulated from the literature analysis:

We have preferred to use CNN (Convolutional Neural Networks) over GAN (Generative adversarial Network) because CNNs are the most widely used

deep learning architectures in image processing and image recognition. Given their supremacy in the field of vision, it's only natural that implementations on different fields of machine learning would be tried. A CNN uses a system much like a multilayer perceptron that has been designed for reduced processing requirements. The layers of a CNN consist of an input layer, an output layer and a hidden layer that includes multiple convolutional layers, pooling layers, fully connected layers and normalization layers. The removal of limitations and increase in efficiency for image processing results in a system that is far more effective, simpler to trains limited for image processing and natural language processing. Hence all this saves us substantial computational power.

3.2 Architecture Diagram :

The architecture diagram is as follows



There are primarily two types of flows in the architecture diagram - The training and the prediction flow.

The prediction flow follows the uploading of a video and pre-processing of it that is, splitting the video into different frames, cropping the face and

detecting it. And then load the trained model which further predicts if the video is real or fake.

The training flow is related to the training and testing of the model using a processed dataset. The videos are then pickled and loaded into the model from the dataset. It is then passed through CNN and LSTM Video classification. The model is evaluated using the confusion matrix. The trained model is then exported.

3.3 **Datasets used :**

The given dataset is a forensics dataset consisting of 1000 original video sequences that have been manipulated with four automated face manipulation methods: Deepfakes, Face2Face, Face Swap and Neural Textures. The data has been sourced from 977 YouTube videos and all videos contain a trackable mostly frontal face without occlusions which enables automated tampering methods to generate realistic forgeries.

3.4 **Techniques :**

For our project, we have made you use of deep-learning methods such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), and Long short-term memory (LSTM).

A convolutional neural network (CNN) is the most commonly used deep neural network model. CNN, like neural networks, has an input and output layer, as well as one or more hidden layers. In CNN , the hidden layers first read the inputs from the first layer and then apply a convolution mathematical operation on the input values. Here, convolution indicates a matrix multiplication or other dot product. After applying matrix multiplication, CNN uses the nonlinearity activation function such as Rectified Linear Unit (RELU) followed by additional convolutions such as pooling layers. The main goal of pooling layers is to reduce the dimensionality of the data by computing the outputs utilizing functions such as

maximum pooling or average pooling.

LSTM, Long short-term memory is a type of artificial recurrent neural network (RNN) that handles long-term dependencies. LSTM contains feedback connections to learn the entire sequence of data. LSTM has been applied to many fields that based on time series data such as classifying, processing and making predictions. The common architecture of LSTM consists of:

1) input gate;

2) forget gate;

3) and an output gate.

The cell state is long-term memory that remembers values from previous intervals and stores them in the LSTM cell. First, the input gate is responsible of selecting the values that should enter the cell state. The forget gate is reasonable of determining which information is to forget by applying a sigmoid function, which has a range of [0, 1]. The output gate determines which information in the current time should be considered in the next step.

References

[1] B. Dolhansky, R. Howes, B. Pflaum, N. Baram, C. C. Ferrer, The deepfake detection challenge (dfdc) preview dataset, arXiv preprint arXiv:1910.08854.

[2] L. Guarnera, O. Giudice, S. Battiato, Deepfake detection by analyzing convolutional traces, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 666–667.

[3] B. Dolhansky, J. Bitton, B. Pflaum, J. Lu, R. Howes, M. Wang, C. C. Ferrer, The deepfake detection challenge (dfdc) dataset, arXiv preprint arXiv:2006.07397.

[4] H. Zhao, W. Zhou, D. Chen, T. Wei, W. Zhang, N. Yu, Multi-attentional deepfake detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 2185–2194.

[5] B. Zi, M. Chang, J. Chen, X. Ma, Y.-G. Jiang, Wilddeepfake: A challenging real-world dataset for deepfake detection, in: Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 2382–2390.

[6] S. Lyu, Deepfake detection: Current challenges and next steps, in: 2020 IEEE international conference on multimedia & expo workshops (ICMEW), IEEE, 2020, pp. 1–6.

[7] T. Chen, A. Kumar, P. Nagarsheth, G. Sivaraman, E. Khoury, Generalization of audio deepfake detection, in: Proc. Odyssey 2020 The Speaker and Language Recognition Workshop, 2020, pp. 132–137.

[8] Y. Zhou, S.-N. Lim, Joint audio-visual deepfake detection, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 14800–14809.

[9] T. Zhao, X. Xu, M. Xu, H. Ding, Y. Xiong, W. Xia, Learning self-consistency for deepfake detection, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 15023–15033.

[10] P. Kwon, J. You, G. Nam, S. Park, G. Chae, Kodf: A large-scale korean deepfake detection dataset, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10744–10753.

[11] A. A. Maksutov, V. O. Morozov, A. A. Lavrenov, A. S. Smirnov, Methods of deepfake detection based on machine learning, in: 2020 IEEE conference of russian young researchers in electrical and electronic engineering (EIConRus), IEEE, 2020, pp. 408–411.

[12] J. Wang, Z. Wu, J. Chen, Y.-G. Jiang, M2tr: Multi-modal multi-scale transformers for deepfake detection, arXiv preprint arXiv:2104.09770.

[13] P. Korshunov, S. Marcel, Deepfake detection: humans vs. machines, arXiv preprint arXiv:2009.03155.

[14] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, D. Manocha, Emotions don't lie: An audio-visual deepfake detection method using affective cues, in: Proceedings of the 28th ACM international conference on multimedia, 2020, pp. 2823–2832.

[15] X. Cao, N. Z. Gong, Understanding the security of deepfake detection, arXiv preprint arXiv:2107.02045.

[16] P. Neekhara, B. Dolhansky, J. Bitton, C. C. Ferrer, Adversarial threats to deepfake detection: A practical perspective, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 923–932.

[17] O. de Lima, S. Franklin, S. Basu, B. Karwoski, A. George, Deepfake detection using spatiotemporal convolutional networks, arXiv preprint arXiv:2006.14749.

[18] S. Agarwal, L. R. Varshney, Limits of deepfake detection: A robust estimation viewpoint, arXiv preprint arXiv:1905.03493.

[19] Y. Nirkin, L. Wolf, Y. Keller, T. Hassner, Deepfake detection based on discrepancies between faces and their context, IEEE Transactions on Pattern Analysis and Machine Intelligence.

[20] Y. Nirkin, L. Wolf, Y. Keller, T. Hassner, Deepfake detection based on the discrepancy between the face and its context, arXiv preprint arXiv:2008.12262.

[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20].

Contributions of the Team Members:

1) Archita Todi (20BCE0529) - Section 1-Introduction(Discussion about the problem, Motivation, Objectives, Possible technicques for proposed system)

2) Parth Maheshwari(20BCE0220) - Section 2-Literature Review(Making a comparison table between 20 research papers along with a detailed discussion of the literature works, research gaps)

3) Mohit Agnihotri (20BCE2513) - Section 3-Proposed Work(Proposal formulated from the literature analysis, Architecture diagram)

4) Kajal Patro(20BCE0736) - Section 3-Proposed Work(Datasets used, Techniques, References)