**DEEP FAKE ANALYSER**

### A Project Progress Report- 1

### *Submitted in the partial fulfilment for the award of the degree of*

### BACHELOR OF ENGINEERING

### IN

### ARTIFICIAL INTELLIGENCE AND

### MACHINE LEARNING

### Submitted by:

### Arnav R Patnaik

### 20BCS6658

### Himanshu Gidwani

### 20BCS6637

**Aryan Gupta**

**20BCS6656**

**Sahildeep Kaur**

**20BCS6587**

### Under the Supervision of:

**MRS.YASHIKA SHARMA**

### SUPERVISORS NAME

**MR. AMIT GARG**

# DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING APEX INSTITUE OF TECHNOLOGY

### CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

**PUNJAB**

#### MONTH & YEAR

**20 MARCH, 2022**

**DECLARATION**

I, **Arnav R Patnaik**, **Himanshu Gidwani**, **Aryan Gupta** and **Sahildeep Kaur**, students of **‘Bachelor of Engineering in Artificial Intelligence and Machine Learning**, **session: 2020-2024**, Department of Computer Science and Engineering, Apex Institute of Technology, Chandigarh University, Punjab, hereby declare that the work presented in this Project Work entitled ‘**Deep Fake Analyser’** is the outcome of our bona fide work and is correct to the best of our knowledge and this work has been undertaken taking care of Engineering Ethics. It contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

### [20ONAIML2-A]

### (Arnav R Patnaik)

### Candidate UID: 20BCS6658

### (Himanshu Gidwani)

### Candidate UID: 20BCS6637

(**Aryan Gupta**)

Candidate UID: **20BCS6656**

(**Sahildeep Kaur**)

Candidate UID: **20BCS6587**

#### Date:20 March 2022

**Place: Chandigarh University, Mohali, Punjab**

**ii**

**ABSTRACT**

**With the speedy growth in the technology and automation sectors, different techniques have been developed which can easily manipulate multimedia content such as videos and images with the ultimate level of realism. It becomes very hard for a person to find out which of the media content is real and which is fake. With the help of deep learning and GANs (Generative Adversarial Networks), deep fakes can be produced. A Deepfake is a fake video or image where a person in an existing video or image is replaced by someone else by using an algorithm. If we look at the positives, we can see how deep fake technology can be used in advertisements, filmmaking, and video games. But on the contrary, it also results in the spread of fake news and the generation of pornographic content. This seriously impacts society and presents enormous threats. This research paper aims to present a project based on a deepfake analyzer that will be able to identify and recognise fake media with the help of deep neural networks.**

Keywords— deepfake, media content, fake news, manipulate, deep learning, GANs

**ACKNOWLEDGEMENT**

I would like to thank my supervisor, Mr Amit Garg sir, for his guidance and advice through each stage of making this project. I would also like to thank Mrs Yashika Sharma ma’am for giving me and my teammates Mr Himanshu Gidwani, Mr Aryan Gupta and Ms Sahildeep Kaur this opportunity to work on a minor project in which we can show our true potential, creativity, and hard work. I would also like to thank my family and friends, who have been a constant support and have always motivated me to work hard and bring out the best in me. This project’s success and end necessitated a great deal of direction and assistance from many people, and we are extremely fortunate to have received it all as part of the project's completion. We owe everything we've accomplished to their oversight and help, and we'd like to express our gratitude.

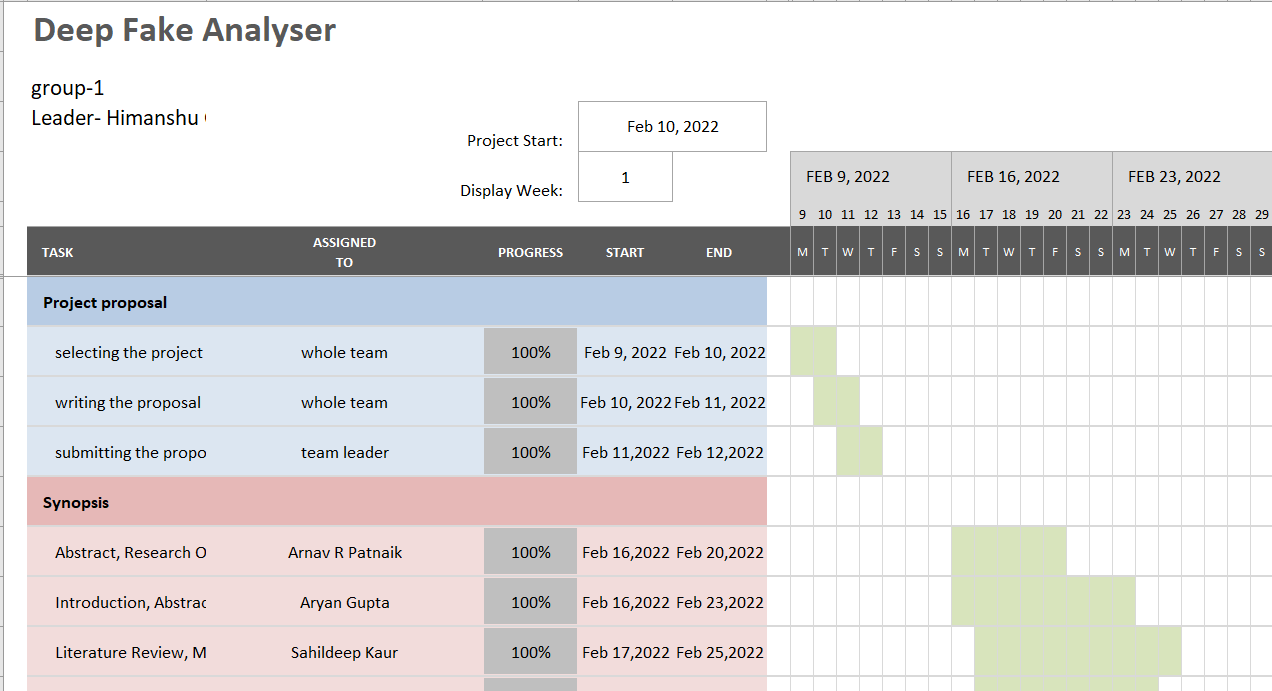
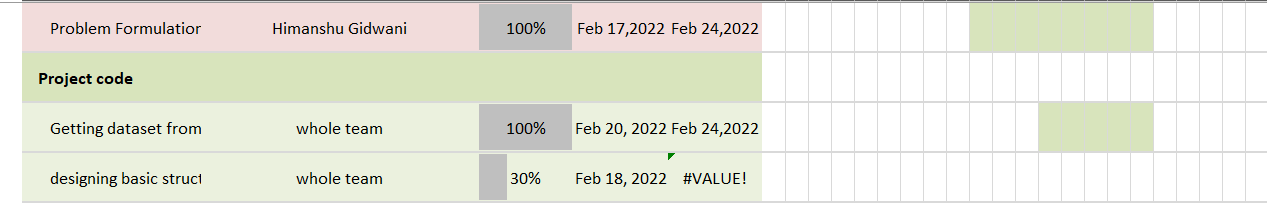
**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| ***Figure*** | ***title*** | ***page*** |
| 1 | Comparison between an original and deep faked image | 6 |
| 2 | Deepfake of politician’s speech going viral | 6 |
| 3 | Flowchart of Deep Learning process | 8 |
| 4 | Deep neural network chain | 9 |
| 5 | System Architecture of the project | 12 |
| 6 | Algorithm chart of DeepFake Analyser | 15 |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| ***Figure*** | ***title*** | ***page*** |
| 2.1 | Literature Review Summary | 13 |

**TIMELINE/GANTT CHART**

# Table of Contents

|  |  |  |
| --- | --- | --- |
| **SR.NO** |  | **PAGE NUMBER** |
| 0 | Title page  Declaration of the Student  Abstract  Acknowledgement  List of Figures  List of Tables  Timeline / Gantt Chart | 1  2  3  4  5  5  5 |
| 1 | **INTRODUCTION\***   * 1. Problem Definition   2. How do Deep Fakes work?   3. Convolution Process   4. Problem Overview   5. Hardware Specification   1.3.1 PC  1.6 Software Specification | 7  7-9  9-10  10  10  10-11 |
| 2 | **LITERATURE SURVEY**   * 1. Existing System Summary | 11-13 |
| 3 | **PROBLEM FORMULATION** | 14 |
| 4 | **RESEARCH OBJECTIVES** | 14 |
| 5 | **METHODOLOGY** | 15 |
| 6 | **TASKS ACCOMPLISHED** | 15-16 |
| 7 | **ONGOING AND UPCOMING TASK** | 16 |
| 8 | **CONCLUSIONS AND DISCUSSION** | 16 |
| 9 | **REFERENCES** | 17 |

# INTRODUCTION

#### PROBLEM DEFINITION

#### A deep fake is a fake, manipulated video or image produced by using deep learning and machine learning techniques that typically features a person’s likeness and voice in a scenario that did not occur to produce false results. This technology can create revenge porn that causes distress to the victim, which includes anger, guilt, paranoia, depression, or even suicide. People use this technology to spread misinformation and false narratives about people.

#### Deepfake technology makes it easy to swap the faces of people in videos or images, so people use it to create videos of trusted people to make other people fall for scams or steal their data or money.

#### 



1. Comparison between an original and deep faked image

#### As shown in the above figure, Facebook's CEO, Mr Mark Zuckerberg, was the subject of a deepfake video in which he appeared to credit the success of the social platform Facebook to a secretive organization. [3]



1. Deepfake of politician’s speech going viral

The image above is a deep fake video of Mr Donald Trump, the former US president, that became popular on the internet. In a roughly 30-second video, Mr Trump claims in a robotic-sounding voice that he will return to social media by joining RuTube, the Russian edition of the video hosting site Youtube. **[4]**

#### HOW DO DEEP FAKES WORK?

Over the years, a large amount of data has been generated which is publicly available on different internet sites, social media, etc. With faster development in machine learning and deep learning techniques like Generative Adversarial Networks, producing realistic fake media has become pretty easy.**[1]**

 Different deepfake techniques can be implemented for manipulating the faces of people. The most commonly used manipulation techniques we see are identity swap, expression swap, and complete face synthesis. We provide the manipulation techniques with larger sets of data that are available publicly.

 Machine learning has a crucial role in deepfakes. Machine learning methods to make deepfakes are built on deep learning techniques and neural networks. The first step of deep faking itself involves training a neural network, which is known as a subset of machine learning, from which we give the model a clear understanding of what a person in an image or video looks like in different conditions like different angles, lighting etc. With the help of machine learning, the model used in deepfake is divided into training and testing. On one hand, model training on a dataset is performed to generate fake videos, while on the other hand, detection is performed to detect the videos.

A very important concept that makes this possible is deep learning. Deep learning is a specialised machine learning technique that instructs and teaches the computer to do what comes to humans naturally. In deep learning, a computer model can classify different objects and tasks just by directly accessing the images, videos (which are broken down into different frames), texts, and sounds. Models learn straight from the data without going through the traditional feature extraction method. Large sets of labelled data and neural network architecture are being utilised to train the models. In contrast, deep learning also requires a good amount of computational power. Here, high-performance GPUs come in handy as they provide a parallel architecture that works efficiently for deep learning. The advantage that deep learning offers over others is that as the amount of data increases, the deep learning models improve as well. This makes Deep Learning achieve greater heights, sometimes outdoing human-level performance.

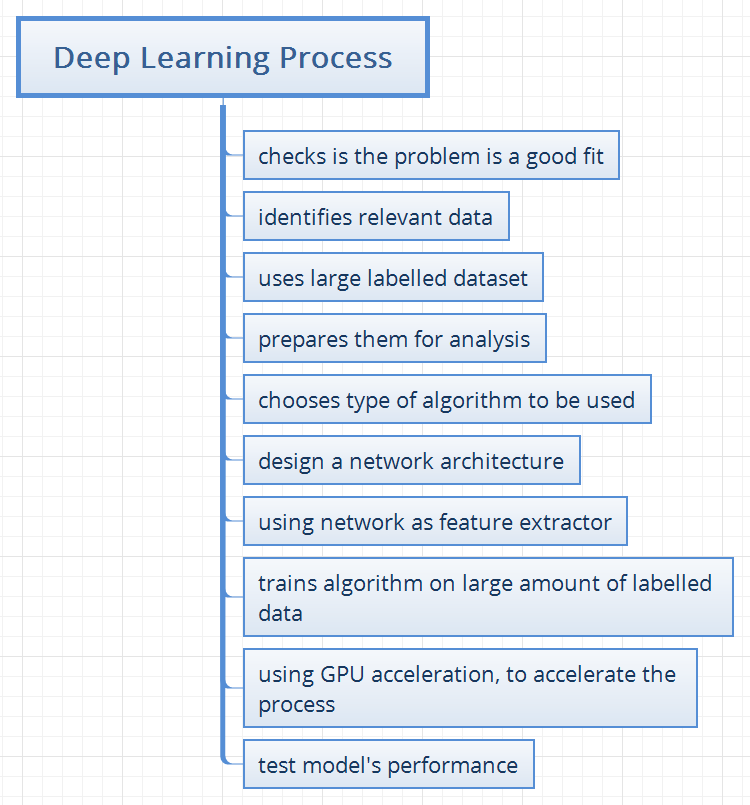


Fig 3: Flowchart of Deep Learning process

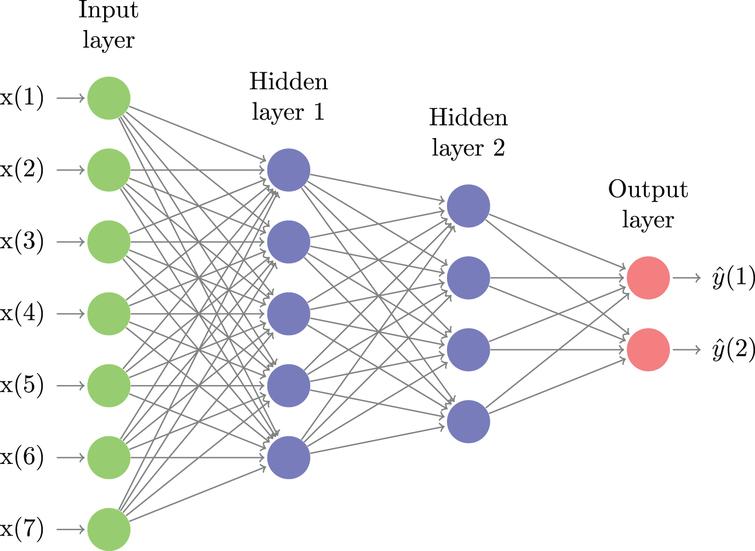


Fig 4:Deep neural network chain **[2]**

A deep neural network is a system that uses many layers of nodes to perform high-level functions from the input that is being provided. A convolutional neural network is used for image recognition purposes and plays a significant part in deep fake detection as well, where we have to detect the image and video frame by frame. Convolutional neural networks are the workhorse behind a lot of what we call artificial intelligence today. There are computational pieces that make them work, and they are called convolution operations. To perform a convolution, we need an image and a kernel image. An image is a two-dimensional array of numbers. The first step that we do is turn the image in the kernel by 180 degrees. This kernel also maps the numbers to colours. The convolution takes the reverse kernel and moves it row by row and column by column, matching it to that patch in the image, multiplying all the elements one at a time by each other and adding up the result. This kernel is the detector of the feature. That is the process of convolution. A convolutional neural network will learn these kernels based on what is more useful and will try to get a better result. A convolutional neural network (CNN) is a type of deep learning algorithm that takes an image as input, assigns learnable weights and biases to objects in an image, and can separate them from one another. These "deep fakes" are hyper-realistic videos and images that apply artificial intelligence to depict someone saying and doing things that have never occurred. Deep fakes are the product of Generative Adversarial Networks (GANs) namely two neural networks working together to generate realistic-looking media. This synthetic content is difficult to detect as it uses real footage and uses authentic-sounding audio.

#### CONVOLUTION PROCESS

What works in deep fake media is the use of generative adversarial networks (GANs). Generative modelling is an unsupervised learning algorithm that involves discovering and learning the patterns in the input data in such a way that the model can come up with new examples that look as if they have been taken from the original dataset. GANs are a brilliant way of training a generative model by framing the problem as a supervised learning problem with two sub-models. One is a generator model, which we train to generate new examples, and the other is a discriminator model that tries to classify examples as real or fake. The two sub-models are trained together until the discriminator model is fooled about half the time that the generator model is actually generating examples from the actual dataset. This helps in generating realistic examples across various problems, which makes it hard for humans to tell what is fake and what is real. Convolution is useful in image processing, and an alternative way to understand it, instead of seeing it as a feature detector, is to see it as a replicator.

Mesonet, a neural network, is used to detect deep fakes. The Meso4 Model is a convolutional model that is being utilised to determine deepfakes in this project. This network starts with four layers of repeated convolutions and pooling and then moves on to a dense network with one hidden layer. The fully-connected layers employ Dropout to regularise and increase their resilience, while the convolutional layers use ReLU activation functions to create non-linearities and Batch Normalization to regularise their output and prevent the vanishing gradient effect. This network has a total of 27,977 trainable parameters.

#### 1.4 PROBLEM OVERVIEW

* Obtain a deep fake dataset from Kaggle and refine it with real-world samples. This aids in the development of a better feature selection system.
* Develop a deep learning and neural network-based algorithm to achieve optimum accuracy.
* We'll add another function to the algorithm that will present the analysis as a bar chart, line chart, or pie chart, giving the user a better understanding of how much the number of deep fakes has grown and how they differ based on their confidence scores.
* Anvil, a free Python-based drag-and-drop web app builder, will be used to create the GUI.

#### 1.5 HARDWARE SPECIFICATION

#### 1.5.1 PC

A PC is a personal computer that can be utilised for multiple purposes depending on its size, capabilities, and price. They are to be operated directly by the ultimate consumer. Personal computers are single-user systems and are portable. Our web application program will be installed on the PC for our clients to use it. This makes it feasible for individual use.

#### SOFTWARE SPECIFICATION

### The software tools that will be utilised in the development of this project are as follows:

|  |  |  |
| --- | --- | --- |
| Software tool used | Description | Logo |
| Jupyter Notebook | Jupyter Notebook is a web-based open-source application that is used for editing, creating, running, and sharing documents that contain live codes, visualisations, text, and equations. Its core supported programming languages are Julia, R, and Python. The Jupyter notebook comes with an IPython kernel that allows the programmer to write programs in Python. There are over 100 kernels other than IPython available for use. |  |
| Atom text editor | Atom is a text and source code editor which works across all operating systems. It speeds up find-and-replace operations by an order of magnitude and improves loading performance for large, single-line files. It’s a desktop application built with HTML, JavaScript, CSS, and Node.js integration. |  |
| Visual Studio Code | Visual studio code is an open-source code editor built for Windows, Mac OS, and Linux which can be used for various programming languages like Java, JavaScript, Python, C, C++, Node.js. It contains support for multiple features like debugging, syntax highlighting, code snippets, code refactoring and git command built-in. |  |
| Anvil | Anvil is a simple way to develop web applications. It consists of a drag and drops design, built-in storage, simple programming in python and easy integration with external services. |  |

# LITERATURE REVIEW

Deep Fake Detection using Inconsistent Head Pose:

**[6]** This project is proposed by Xin Yang, Yuezen Li, and Siwei Lyu. Here synthetic content like deep fake videos and images is detected using inconsistent head poses. The 3D poses correspond to rotation and the translation of worlds coordinates to the corresponding camera coordinates. Here R is a rotation matrix that is of 3x3 order and t is a 3x1 translation vector. In 3D head poses estimation we need to solve the reverse of R and t by using 2D image coordinates. As per the model SVM that is Support Vector Machine is selected to detect synthetic contents. SVM is trained to classify based on differences between head pose estimated using a full set of facial landmarks and those in central face regions to differentiate. Here training and testing data involve a dataset of deep fake and real images and videos. Yang and Li in the year 2018 released an article regarding this topic. The topic released by Yang focuses on analysing synthetic content based on head pose estimation. A technique like this involves a classifier like SVM (Support Vector Machine). The parameters that were used to evaluate this model or design include complexity and accuracy.

Deep Fake Detection:

**[7]** This project is implemented by Piyush Chandra during the deep fake detection Challenge that is also known as DFDC. This project can be found on Kaggle This model is designed to detect deep fake videos. To achieve this technology face recognition is used that will capture the video in particular frames thus giving time to the model to detect whether it is real or fake. For a baseline, the submission facenet is used. Facenet is a face recognition pipeline that learns a mapping from faces to positions in multidimensional space where the distance between the points directly corresponds to a measure of face similarity. The final output that is whether the video is real or fake has been provided by plotting a histogram. The plotted histogram provides us with the confidence level that whether the model is confident about the prediction that has been made. The parameters that were used to evaluate this model or design include complexity and accuracy. But this model lacks a graphical user interface, which just makes it a simple text-based program. Besides this, the final output was not clear as it was provided in the form of a graph.

Deep Learning Deep Fake Detection (ResNext and LSTM):

**[5]** This project was designed by a college student team. The technique that they adopted is called LSTM, a neural network. Aside from neural networks, web applications were also created using the Django application. The parameters that were used to evaluate this model or design include complexity and accuracy. This project is designed to detect videos that fall under the category of "deep fakes." Here, deep learning techniques like LSTM and ResNext CNN are used to obtain a feature vector. A further LSTM layer is trained using these features. The system architecture of this model includes the upload action of videos. Then comes the pre-processing part, where the splitting of data takes place. This is followed by a deep fake detection model that includes ResNext Feature extraction and LSTM video classification. After deepfake detection comes the model evaluation, where a confusion matrix is used that gives us our final and required output in the form of a label that says whether the video is authentic (which is original), or synthetic (which comes under the tag fake).

A website has been formed by making use of the Django application to create a real product. Django is a high-level Python web framework that encourages rapid development and clean code. Django is a collection of Python libraries allowing you to efficiently create a quality web application and is suitable for both the backend and the frontend.

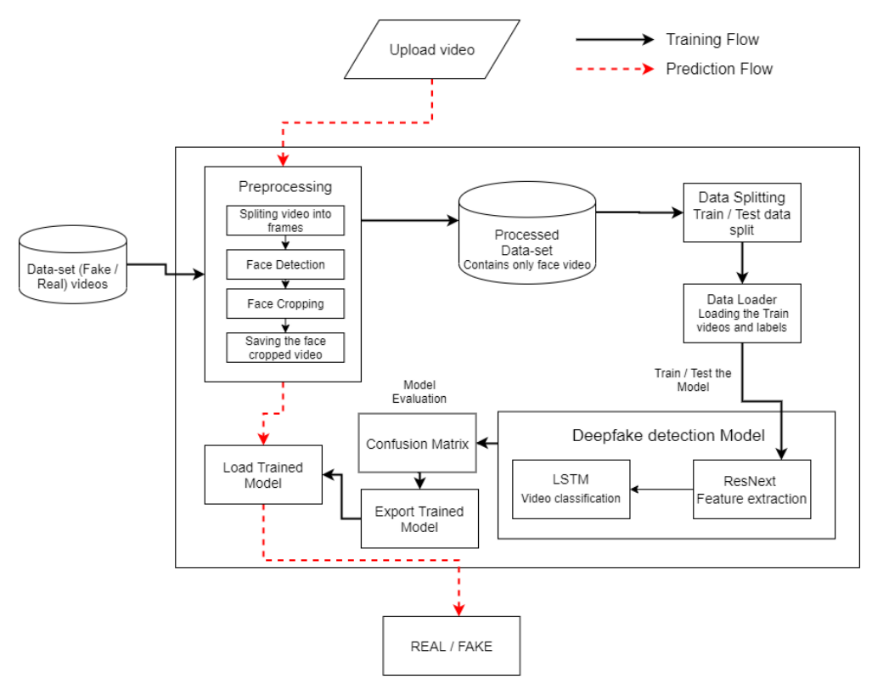


Fig 5: System Architecture of the project

**2.1 Literature Review Summary**

Table 2.1: Literature review summary

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year and citation** | **Article Title** | **Purpose of study** | **Tools / Software Used** | **Comparison of technique** | **Source (Journal / Conference)** | **Findings** | **Data set (if used)** | **Evaluation parameters** |
| Yang, X., Li, Y. and Lyu, S. (2018). Exposing Deep Fakes Using Inconsistent Head Poses. *arXiv:1811.00661 [cs]*. [online] Available at: https://arxiv.org/abs/1811.00661 | Exposing deep fakes using inconsistent head pose | It focuses on analyzing deep fake contents using inconsistent Head poses estimation. In this kind of technique classifier like SVM is used. | Inconsistent head poses estimation | * Deep Learning * Head pose estimation | https://arxiv.org/abs/1811.00661 | Fake image and videos detection | paperswithcode.com/dataset/faceforensics-1 | * Complexity * accuracy |
| kaggle.com. (n.d.). *Deepfake Detection*. [online] Available at: https://www.kaggle.com/piyush357/deepfake-detection. | Deepfake Detection | Used to detect synthetic video by making use of face recognition. | * Face recognition * Jupyter Notebook | Deep Learning | https://www.kaggle.com/piyush357/deepfake-detection | Detection of fake images and videos | piyush357/deepfake-detection/data | * Complexity * accuracy |
| Jadhav, A.H. (2022). *Deepfake detection using Deep Learning (ResNext and LSTM)*. [online] GitHub. Available at: https://github.com/abhijitjadhav1998/Deepfake\_detection\_using\_deep\_learning [Accessed 26 Feb. 2022] | Deepfake detection using Deep learning (ResNext and LSTM) | Detects deep fake video by making use of a neural network like LSTM. Also involves Django application to create a web application. | * Django Application * Jupyter Notebook | * Deep learning * LSTM | https://github.com/abhijitjadhav1998/Deepfake\_detection\_using\_deep\_learning | Fake images and videos are detected. | piyush357/deepfake-detection/data | * Complexity * accuracy |

# PROBLEM FORMULATION

### The major problem arising in the project work is that we do not have enough datasets for real and deepfake prediction. We should have more datasets to improve the model's prediction accuracy. It's critical to overcome this problem since it won't be able to anticipate real-world deepfake pictures because the vast majority of the deepfake dataset comes from pornographic sites or a specific portion thereof. When a person submits an image to see if it is real or a deep fake, the problem becomes even worse. They could get erroneous findings.

### There is a need to make a bigger and better dataset with varied types of images and people of different ethnicities. We can come up with a few methods to tackle this problem. We can either create or arrange a platform for users to upload their images, or we can ask a few large image platforms, such as Meta, which deals in a variety of different types of user images, to provide us with datasets for research purposes, or we can merge the publicly or freely available datasets of deepfake and real photos into one.

### Second, there is another issue: deep fake analysis is not the ultimate solution to reducing the amount of synthetic information available on the internet. The final answer, we discovered, is verification of the material accessible on various websites on the internet. In the real world, this may be accomplished by large IT businesses banding together and including some type of identification into their platform while publishing material.

# RESEARCH OBJECTIVES

The proposed study is to conduct research that will lead to the creation of a method for identifying deepfakes. The Deepfake Analyser, which is the proposed project, will be accomplished by separating it into the following goals:

1. Obtain a deep fake dataset from Kaggle and refine it with real-world samples. This aids in the development of a better feature selection system.
2. Develop a deep learning and neural network-based algorithm to achieve optimum accuracy.
3. We'll add another function to the algorithm that will present the analysis as a bar chart, line chart, or pie chart, giving the user a better understanding of how much the number of deep fakes has grown and how they differ based on their confidence scores.
4. Anvil, a free Python-based drag-and-drop web app builder, will be used to create the GUI.

# METHODOLOGY

The following methodology will be followed to achieve the objectives defined for the proposed research work:



Fig. 6: Algorithm chart of DeepFake Analyser

1. **TASK ACCOMPLISHED**

The project's goal is to identify hyper-realistic content, such as videos and images. This sort of detection system has the potential to be extremely useful in both the public and private sectors. The amount of research done on this subject, namely the identification of synthetic material, sometimes known as "deep fakes," is limited. Our group's goal is to develop a model that can distinguish between real and fake objects quickly.

The first goal was to learn more about what kind of data set we could work with, which model would provide a good response to our inputs, how accurate these models could be, and so on. Following the collecting of the necessary information, we moved on to the implementation phase.

Importing the appropriate libraries or modules is the first step in the implementation process. Simple and helpful modules like NumPy and matplotlib were imported first, then as we progressed, we moved on to TensorFlow, Keras, Adam, and other modules and classes.

The model's structure, such as how many layers the model will have, whether padding will be utilised or not, and what activation unit will be used, all go under the Meso4 class, which takes the classifier class as a parameter. Meso4 is a Convolutional Neural Network model with four layers in its fundamental layout. MaxPooling and batch normalisation are two of the technologies covered. Max Pooling is used to minimise the dimensionality of picture output by reducing the amount of pixels in the output from the preceding convolutional layer, which is a necessary step. This gave us a hard time because we had no idea how this would function or what the outcome would be produced.

When it comes to padding, zero padding is employed when the original size of a picture must be preserved. If padding is true, it indicates there isn't any padding. Padding is done to make the output size the same as the input size if it is specified as the same.

Now we loaded the weights for our input transformation using hidden layers. Weight is the parameter within a neural network that transforms input data within the network's hidden layers. Now we rescale our original images as rescale is a value by which we will multiply the data before any other processing. Our original images consist in RGB coefficients in the 0-255, but such values would be too high for our model to process (given a typical learning rate), so we target values between 0 and 1 instead by scaling with a 1/255. factor. Next, we start indexing our dataset in 0 and 1. 0 we set for deepfake and 1 for real images. Now we pick random images one by one and start passing through our prediction model. We passed all images from our dataset through our model and we will categorize them in 8 different categories correct real, correct\_real\_pred,correct\_deepfake,correct\_deepfake\_pred,misclassified\_real,misclassified\_real\_pred,misclassified\_deepfake ,misclassified deepfake pred then we plotted these all categorized images with their confidence score.

**All of these activities were finished by the 15th of March.**

1. **ONGOING AND UPCOMING TASKS**

**Ongoing Tasks:**

* Integrating the backend with the frontend
* Testing GUI with tkinter

**Future Tasks:**

* Finding the best optimal way for the GUI
* Flutter is taken in consideration for the GUI

1. **RESULTS AND DISCUSSION**

Deep learning is now making waves in the field of artificial intelligence, and image-based challenges have aided in this achievement. Because of recent developments in convolutional neural networks, we believe Ai will surpass humans in most traditional vision-based activities. There are true and significant ramifications for society, information, and the media that will have a big influence for years to come now that AI is capable of making convincing fake photographs and videos.

1. **REFERENCES**
2. ‌ Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A. and Ortega-Garcia, J. (2020). Deepfakes and beyond: A Survey of face manipulation and fake detection. *Information Fusion*, [online] 64, pp.131–148. Available at: <https://arxiv.org/pdf/2001.00179.pdf>.
3. Dixon, M., Klabjan, D. and Bang, J.H. (2017). Classification-based financial markets prediction using deep neural networks. *Algorithmic Finance*, 6(3-4), pp.67–77.
4. Dailymail.co.uk. (2022). [online] Available at: https://i.dailymail.co.uk/1s/2019/06/12/00/14669252-0-image-a-2\_1560294887875.jpg [Accessed 5 Apr. 2022].
5. ‌ Palmer, E. (2022). *Donald Trump’ deepfake shows ex-president joining Russia’s version of YouTube*. [online] Newsweek. Available at: https://www.newsweek.com/trump-deepfake-russia-rutube-1687647 [Accessed 5 Apr. 2022].
6. Jadhav, A.H. (2022). *Deepfake detection using Deep Learning (ResNext and LSTM)*. [online] GitHub. Available at: https://github.com/abhijitjadhav1998/Deepfake\_detection\_using\_deep\_learning.
7. Yang, X., Li, Y. and Lyu, S. (n.d.). *EXPOSING DEEP FAKES USING INCONSISTENT HEAD POSES*. [online] Available at: https://arxiv.org/pdf/1811.00661v2.pdf [Accessed 7 Apr. 2022].
8. kaggle.com. (n.d.). *Deepfake Detection*. [online] Available at: https://www.kaggle.com/code/piyush357/deepfake-detection/notebook [Accessed 7 Apr. 2022].