IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, November 2026
Impact Factor: 8.2

Detection of Deepfake Videos Using Long-Distance Attention

Pruthviraj T J
PES Institute of Technology and Management (PESITM), Shivamogga, Karnataka, India.

Abstract : *The "Detection of Deepfake Videos Using Long-Distance Attention" project uses deep learning algorithms to determine if a video is authentic or fraudulent. These days, artificial intelligence technology is developing quickly, and by altering a person's face, attitudes, and movements, it is feasible to produce phony videos. We refer to these videos as deepfake videos. They can be used to misuse someone's identity, disseminate fake information, and produce deceptive material. As a result, identifying phony videos has grown to be a significant difficulty.This project's primary goal is to create a deep learning-based online service that enables users to upload videos and determine if they are authentic or fraudulent. Convolutional Neural Network (CNN) and Long-Distance Multi-Head Self-Attention are combined by the system to analyze video frames and detect significant temporal and visual characteristics.Real and false video datasets are gathered for this project and kept in different files for model testing and training. The system preprocesses each video by extracting specific frames and resizing them to a predetermined image size.*

Important features are extracted from the video frames using the CNN model. The Long-Distance Attention mechanism examines the connections between several frames after feature extraction. Lastly, a classification layer determines if the video falls into the REAL or FAKE category.Python, PyTorch, OpenCV, Flask, HTML, CSS, and JavaScript are used in the project's implementation. OpenCV is used for video processing, Flask is used to create the web application, PyTorch is used to create and train the deep learning model, and Python is used for model development and implementation.The application has a result page, video upload module, user dashboard, admin page, and login page.The deepfake detection model can be trained by the administrator after the dataset has been prepared. The program allows users to check the prediction results, add video files, and log in. The system shows the anticipated category as REAL or FAKE along with confidence scores after processing the uploaded video.The creation of a lightweight model that can function with constrained computing resources is another goal of the research. Effective model training and prediction are supported by methods like Group Normalization, decreased frame processing, and smaller image dimensions.

Keywords: Detection of Deepfake Videos Using Long-Distance Attention

I. INTRODUCTION

Deep learning and artificial intelligence (AI) technologies have advanced quickly in recent years. Numerous industries, including education, healthcare, entertainment, communication, and digital media, use these technology. Deepfake technology is one of the significant advancements in this field. Videos can be made or altered using deepfake technology by altering a person's voice, body language, facial expressions, or face. It can make a phony video appear similar to a real one, making it challenging for viewers to distinguish between the two.These days, social media, websites, and other internet apps are used extensively to exchange videos. Deepfake technology is sometimes abused to produce deceptive videos, disseminate false information, harm someone's reputation, and confuse the public.The goal of the project "Detection of Deepfake Videos Using Long-Distance Attention" is to determine if a video that has been uploaded is authentic or fraudulent. This project's primary goal is to create a deep learning-based system that can examine video frames and identify potential modifications. In order to comprehend crucial information from several

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, November 2026
Impact Factor: 8.2

video frames and enhance the detection process, the system employs a Long-Distance Attention technique.Real and fraudulent videos are gathered and arranged into distinct datasets for testing and training in this project.

The algorithm is used to forecast the outcome of newly uploaded movies by users after it has been trained.The Flask framework is also used to create an online application. Users can log in, access the dashboard, and upload video files for detection using its straightforward interface. Once a video has been uploaded, the algorithm analyzes it and shows the final forecast along with confidence percentages, indicating whether it is REAL or FAKE.The Python programming language and many technologies, including PyTorch, OpenCV, Flask, HTML, CSS, and JavaScript, are used to implement the project. The deep learning model is created and trained using PyTorch, the video is processed using OpenCV, and the trained model is connected to the web application using Flask. In summary, the goal of this research is to apply Long-Distance Attention to create a deep learning-based deepfake video detection system. It offers an automated way to identify uploaded movies as REAL or FAKE by combining web application technologies, deep learning, and video processing.

1.1. PROBLEM STATEMENT

Artificial intelligence and deep learning technologies are being used more and more these days. By altering a person's face, attitudes, and movements, these technologies enable the creation of bogus videos. We refer to these videos as deepfake videos. Because these altered videos resemble the originals so closely, it is challenging for regular users to determine whether the content is authentic or not.The primary issue is the ease with which phony videos can be disseminated via websites, social media platforms, and other online resources. These movies could be used to mislead the public, propagate incorrect information, cause misconceptions, and harm someone's reputation. Because the alterations performed to the movies might not be apparent to the human eye, it is challenging to manually identify deepfake videos.

Advanced facial modifications may not always be correctly detected by conventional video verification techniques. As a result, an automated system that can examine video frames and spot potential distinctions between authentic and fraudulent videos is required.By creating a deep learning-based detection system, the proposed project "Detection of Deepfake Videos Using Long-Distance Attention" seeks to address this issue. Using a trained model, the system examines the uploaded video frames to determine whether the video is authentic or fraudulent. Additionally, it shows confidence percentages to aid users in comprehending the prediction outcome.Through a straightforward interface, users can upload video files to the system's web-based application and get the detection result. The primary objective is to increase user accessibility, speed, and ease of deepfake video detection.

1.2 PROJECT DESCRIPTION

It was created to use deep learning technology to determine whether a video is REAL or FAKE. Using an attention-based model, the system examines video frames to identify potential facial modifications.A online application allows users to submit video files. The system uses OpenCV to process the submitted video and a trained PyTorch model to forecast the outcome. Along with confidence percentages, the outcome is shown as REAL or FAKE.Python, PyTorch, OpenCV, Flask, HTML, CSS, and JavaScript are used in the development of the application. It offers a straightforward and easy-to-use interface for identifying deepfake videos.

1.3 OBJECTIVES

This project's primary goals are:

- To create a deep learning system that can distinguish between authentic and fraudulent videos.
- To use Long-Distance Attention techniques to analyze video frames.
- To spot potential facial modification in videos that have been uploaded.
- To use both genuine and fake video datasets to train the model.
- To offer a straightforward web application for testing and uploading movies.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, November 2026
Impact Factor: 8.2

- To show the forecast outcome with confidence percentages as either REAL or FAKE.
- To lessen the dissemination of false information and raise awareness of deepfake videos.

1.4 SCOPE

The goal of this research is to create an automated deep learning-based system that can distinguish between authentic and fraudulent videos. The system is built to use Long-Distance Attention techniques to analyze submitted video files and detect potential facial changes.Via the application's web-based interface, users can upload videos. Following upload, the system uses OpenCV to extract specific video frames, which it then processes using a deep learning model that has been trained. The technology determines whether the video is REAL or FAKE based on the analysis and shows the outcome with confidence percentages.A user-friendly dashboard with login, video upload, prediction, and logout features is also offered by the project. It raises awareness of the improper use of deepfake technology and assists users in determining the veracity of digital video material. Students, scholars, social media users, and anyone looking to examine potentially edited films can all benefit from this technique. The project can be expanded in the future by utilizing bigger datasets, enhancing model performance, supporting more video formats, and creating more sophisticated deepfake detection techniques.

1.5 PURPOSE

The goal of this project is to use deep learning technology to create an automated system for identifying deepfake videos. Fake videos can now be produced by altering a person's visage, facial expressions, and body language to make them appear authentic. This project uses a trained deep learning model to analyze the frames of an uploaded video to assist users determine if it is REAL or FAKE.Users can upload films and examine the prediction results with confidence percentages using the system's straightforward online application. It seeks to lessen the dissemination of inaccurate and deceptive video content, increase knowledge of deepfake technology, and facilitate video verification.

II. LITERATURE SURVEY

The literature review offers details on earlier studies and various methods for identifying deepfake videos. Numerous researchers have created techniques to detect altered photos and videos using artificial intelligence, machine learning, and deep learning. These studies aid in comprehending current techniques, their benefits, and their significance in creating a better deepfake detection system.

[1] Celeb-DF: An Extensive and Difficult Dataset for DeepFake Forensics (2020):

In 2020, Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu presented the Celeb-DF dataset. Both actual and deepfake videos made with sophisticated face-swapping algorithms are included in this dataset. This study's primary goal was to offer a difficult dataset for assessing deepfake detection techniques. High-quality altered videos that resemble original footage are included. This study aids engineers in comprehending the challenges associated with recognizing realistic-looking phony films and enhancing detection precision.

[2] Convolutional Vision Transformer for Deepfake Video Detection (2021):

A deepfake video detection technique utilizing Convolutional Vision Transformer technology was proposed by D. Wodajo and S. Atnafu. This method extracts key features from video frames by combining transformer-based methods with convolutional neural networks. While the transformer mechanism aids in understanding the links between various image components, CNN assists in identifying local visual patterns. This study offers a practical method for using deep learning to identify altered facial content.

[3] Deepfake Detection Scheme Based on Vision Transformer and Distillation (2021)

Y.-J. Heo and colleagues presented a deepfake detection technique based on knowledge distillation and Vision Transformer. To find crucial visual information, Vision Transformer splits images into smaller areas and examines the connections between them. Transferring valuable information from a larger model to a smaller model is facilitated by

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 6, November 2026 Impact Factor: 8.2

knowledge distillation. This method is helpful in lowering the computing demands of deep learning models and enhancing deepfake detection.

[4] FaceForensics++: Identifying Photoshopped Facial Images (2019)

In 2019, A. Rössler and colleagues presented FaceForensics++. It is a popular dataset for identifying photos and videos of faces that have been altered. The collection include both authentic footage and other kinds of altered videos made with face-manipulation methods. It supports the training and assessment of deepfake detection algorithms by researchers. This study offers a crucial basis for creating systems that can recognize altered facial features and identify modified video information.

[5] FakeCatcher: Biological Signals-Based Identification of Synthetic Portrait Videos (2020)

FakeCatcher, a method for identifying fake portrait videos using biological signals, was presented by U. A. Çiftçi, İ. Demir, and L. Yin. This technique examines minute variations in facial skin tone brought on by blood flow. Although these changes are hard to see with the naked eye, video frames can be used to measure them. These biological cues are used by the algorithm to differentiate between actual and synthetic films. By concentrating on physiological data rather than just visual facial traits, this study presented an alternative method for deepfake identification.

2.1 EXISTING SYSTEM

Convolutional Neural Networks (CNN), machine learning, and image processing techniques are some of the approaches used by the current deepfake video detection system. In order to identify modified content, these systems examine facial features, expressions, and visual changes in video frames.Some current techniques train and evaluate deepfake detection models using datasets like FaceForensics++ and Celeb-DF. These techniques aid in the detection of phony movies, but they may encounter challenges in identifying sophisticated facial modifications, high-quality deepfakes, and various video attributes.The majority of current systems need computing power and technical expertise for testing and training. Additionally, some approaches may not completely analyze the links between several video frames since they concentrate primarily on individual frames.As a result, current techniques are limited in their ability to reliably identify sophisticated deepfake films and offer a user-friendly detection tool.

2.1.1 DISADVANTAGES

- High-quality deepfake films might get undetected by current systems.
- When video quality is poor, detection accuracy could drop.
- . Large datasets are necessary for training some techniques.
- A lot of processing power is needed to train deep learning models.
- Some systems may overlook changes between video frames because they simply examine individual frames.
- A.It is challenging to recognize advanced facial modification techniques.
- Some of the current systems are complicated and need technical expertise to function.
- When testing films differ from the training dataset, detection performance may suffer.

2.2 PROPOSED SYSTEM

The "Detection of Deepfake Videos Using Long-Distance Attention" system is designed to use deep learning techniques to determine whether an uploaded video is REAL or FAKE. In order to identify potential face modifications, the algorithm examines specific video frames and extracts significant visual characteristics.The suggested system processes video frames and distinguishes between authentic and fraudulent videos using a Frame Encoder and an Attention-based deep learning model. The detection model is trained and executed using PyTorch, while frames are extracted and processed using OpenCV JavaScript, HTML, CSS, and Flask are used in the development of a web application. After logging in, users can check the prediction results, upload video files, and access the dashboard. Along with confidence percentages, the system indicates whether the result is REAL or

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 6, November 2026 Impact Factor: 8.2

FAKE.The suggested solution offers an easy-to-use platform for identifying deepfake videos. It seeks to facilitate video analysis and assist consumers in spotting potentially altered digital information.

2.2.1 ADVANTAGES

- automatically determines if a video uploaded is authentic or fraudulent.
- Analyzes video frames using deep learning algorithms.
- Uses Long-Distance Attention to comprehend data from various frames.
- Offers forecast outcomes along with percentages of confidence.
- Provides an easy-to-use web application.
- Makes manual video verification less necessary.
- Allows for automatic analysis and video uploading.a
- Can be enhanced with more sophisticated detection methods and bigger datasets.
- Offers a handy dashboard with the ability to log in and exit.

III. SOFTWARE REQUIREMENT SPECIFICATION

3.1 REQUIREMENT ANALYSIS

An essential stage of software development is requirement analysis. It specifies the software and hardware needed to create and operate the suggested deepfake video detection system. To evaluate films and determine if they are REAL or FAKE, the system needs the right hardware, software tools, computer languages, and datasets.The suggested project makes use of Flask for web application development, OpenCV for video processing, PyTorch for deep learning model construction, and Python for implementation. The user interface is made with HTML, CSS, and JavaScript.Video uploading, frame extraction, deep learning model training, video prediction, and showing the detection results with confidence percentages are the primary requirements of this project. These specifications enable the system to function correctly and give users a straightforward experience.

3.1.1 FUNCTIONAL REQUIREMENTS

The primary tasks and functions carried out by the suggested system are outlined in the functional requirements. They describe how users engage with the system and how the program functions. Below is an explanation of the Deepfake Video Detection System's primary functional needs.

1. User Login

Authorized users can access the web application by using their login credentials thanks to the system's login feature. It aids in managing application access.

2. Dashboard

The user can access the dashboard following a successful login. It shows the model status, video upload option, application details, and other significant elements.

3. Video Upload

Users can choose and upload video files from their PC using the system. The deepfake detection technique uses the uploaded video as input.

4. Video Processing

The system reads the movie and extracts specific frames using OpenCV once the video has been uploaded. For model processing, the extracted frames are scaled and transformed into the necessary format.

5. Feature Extraction

From the chosen video frames, the Frame Encoder retrieves crucial visual information. The deep learning model analyzes the video footage using these features.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 6, November 2026 Impact Factor: 8.2

6. Deepfake Detection

In order to determine whether the uploaded video is REAL or FAKE, the trained deep learning model analyzes the extracted characteristics using Long-Distance Attention techniques.

7. Result Display

The system shows the prediction result on the dashboard once the detecting procedure is finished. To aid users in understanding the outcome, it also displays REAL and FAKE confidence percentages.

8. Logout

After the video detection process is finished, users can safely exit their session by using the system's logout option.

3.1.2 NON-FUNCTIONAL REQUIREMENTS

The quality, performance, security, and dependability of the suggested system are described by non-functional criteria. They specify how the deepfake video detection program should function and how well it should carry out its tasks. These specifications aid in ensuring that the program is dependable, user-friendly, and appropriate for video detection. The following explains the project's primary non-functional requirements:

1. Performance

In a fair length of time, the system should process submitted video files and deliver prediction results. Reading the video, extracting certain frames, and analyzing them with the trained deep learning model are all part of video processing. The size, quality, and hardware resources of the video may affect its performance.

2. Accuracy

By analyzing significant visual characteristics from video frames, the system ought to produce accurate forecasts. Using the supplied training dataset, the trained model should be able to distinguish between actual and altered videos. The quality and diversity of the dataset and model training determine the detection performance.

3. Usability

The application ought to have an easy-to-use interface. Without the need for sophisticated technological knowledge, users should be able to log in, access the dashboard, upload video files, and examine the prediction results. Information should be displayed clearly on the interface.

4. Security To manage access to the online application, the system ought to include login and logout capabilities. To stop unwanted access, user sessions should be appropriately managed. Maintaining application security requires careful handling of uploaded video data.

5. Reliability

The system should successfully upload, process, and anticipate videos. It should appropriately handle processing errors and incorrect files. When an uploaded video cannot be processed, the program ought to display suitable messages.

6. Compatibility

The online application should function with popular web browsers like Microsoft Edge and Google Chrome. It should process video files in accordance with the necessary model parameters and support the video file formats that the program accepts.

7. Maintainability

Future modifications and enhancements should be made easier by the system's design. The deep learning model should be updated and modified by developers.

3.2 SYSTEM DESIGN

System design is a crucial stage of software development that describes the general layout and operation of the suggested system. The suggested Deepfake Video Detection System uses deep learning algorithms to determine if an uploaded video is authentic or fraudulent. When a user uploads a video file and logs into the online application, the system is activated. OpenCV is used to process the uploaded video and extract certain frames.The trained deep learning model receives the resized extracted frames. While the Long-Distance Attention mechanism aids in the analysis of

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 6, November 2026 Impact Factor: 8.2

relationships between several frames, the Frame Encoder extracts significant visual characteristics. The model determines if the video is REAL or FAKE based on the extracted data.

Lastly, confidence percentages and the forecast outcome are shown on the online application dashboard. Flask controls the application's operations and links the deep learning model to the user interface.User login, dashboard, video upload, video processing, feature extraction, deep learning prediction, and result presentation are the system's primary constituents. Together, these elements create an automatic and intuitive deepfake video identification method.

3.2.1 SYSTEM ARCHITECTURE

The general layout and operation of the suggested Deepfake Video Detection System are described by the system architecture. It describes how various modules work together to determine if a submitted video is authentic or fraudulent.From user login and video upload to video processing and final prediction, the system operates in a sequential manner. The user first opens the dashboard after gaining access to the web application via the login page. After that, the user submits a video file for identification.OpenCV examines the video after it has been uploaded and extracts specific frames. The extracted frames are transformed into the necessary format and resized. The deep learning model receives these frames and uses them to extract and analyze features.

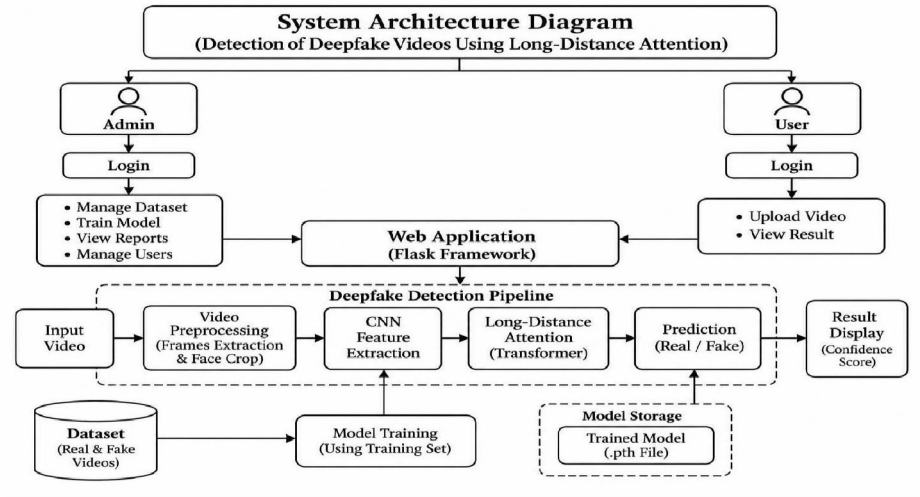

Fig 3.2.1 System Architecture

Important visual elements are extracted from the video frames by the Frame Encoder. The relationships between various frames are examined by the Long-Distance Attention mechanism. The video is then categorized as either REAL or FAKE by the trained model.The prediction result is then delivered to the Flask web application, where it is shown with confidence percentages on the dashboard. The deep learning model, result display module, video processing module, and user interface are all connected by the architecture.