International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 08 | Aug 2023 www.irjet.net p-ISSN: 2395-0072

DeepSecure: A Real-Time Deep Learning-Based System for Enhancing Cybersecurity in Social Media through DeepFake Detection using LSTM and ResNext CNN

Nikhil Dhiman[1], Nitesh Sharma[2], Vikalp Vashisth[3]

Abstract - *With the exponential rise of DeepFake content circulating on social media platforms like Twitter, the need for robust and real-time detection systems has become paramount to safeguard digital trust and authenticity. In response to this pressing concern, "DeepSecure," an innovative deep learning-based solution tailored for efficient DeepFake detection. By harnessing the power of Long Short-Term Memory (LSTM) networks [1] and Residual Next (ResNext) Convolutional Neural Networks (CNNs), DeepSecure adeptly analyzes multimedia content on social media feeds. The proposed system empowers users and platform administrators to combat the escalating threat of deceptive content, thereby fortifying the cybersecurity landscape in social media ecosystems. Through rigorous experimentation and real-world implementation, this research endeavors to offer a reliable and timely defense against the proliferation of DeepFake content on popular social media platforms.*

Key Words: DeepSecure, Real-Time Deep Learning, Cybersecurity, Social Media, DeepFake Detection, LSTM, ResNext CNN, Enhancing, Detection System, Digital Security

1. INTRODUCTION

With the unprecedented growth of social media and the widespread sharing of multimedia content, the rise of DeepFake technology has emerged as a significant cybersecurity threat, jeopardizing the authenticity and trustworthiness of information circulated online. DeepFake techniques employ advanced machine learning algorithms, including Long Short-Term Memory (LSTM) [1] networks and Residual Next (ResNext) Convolutional Neural Networks (CNNs) [2], to produce highly realistic and deceptive fake videos and images. As the prevalence of DeepFake content continues to escalate on social media platforms, there is a pressing need for real-time, robust, and efficient solutions to detect and mitigate the dissemination of false and misleading information.

In this research paper, we present "DeepSecure," a pioneering real-time deep learning-based system that uses the power of LSTM networks [1], Python, and ResNext CNNs [2] to effectively combat DeepFake threats in the realm of social media. DeepSecure is meticulously engineered to intelligently analyze multimedia content shared on social media platforms, enabling rapid and accurate identification of DeepFake content. Leveraging the strength of LSTM [1] and ResNext CNN [2] architectures, our system enhances cybersecurity measures by providing a reliable and scalable solution to counteract the growing sophistication of DeepFake technology.

Through rigorous experimentation and performance evaluation, we aim to demonstrate the effectiveness and practicality of DeepSecure in safeguarding the integrity and trustworthiness of multimedia content on social media. The integration of deep learning techniques and Python programming enables DeepSecure to operate in real-time, allowing for swift detection and response to potential DeepFake threats. By presenting a comprehensive analysis of DeepSecure's capabilities, we endeavor to contribute valuable insights towards the ongoing efforts to mitigate the adverse impact of DeepFake content on social media platforms, fostering a safer and more secure digital landscape for all users.

2. LITERATURE SURVEY

These days, a number of new threads are emerging as the usage of AI technology grows. A media or video can be edited using deep learning to create a false version and raise security concerns on social media sites. The changed media may be utilized for journalism, entertainment, and politics. Some excellent materials, such as an IEEE (Spectrum) publication [3], help to improve the quality of Deepfake development and leads to more fake content over the social media.

Numerous research endeavors have been dedicated to the detection of deepfakes; however, achieving real-time detection remains a challenging pursuit. This research paper seeks to address this crucial gap by focusing on the development of a real-time deep learning-based system for robust deepfake detection. By exploring the fusion of LSTM [1] and ResNext CNN [2] models, our study aims to contribute to the advancement of cybersecurity in social media, enabling swift and efficient identification of manipulated content in a dynamically evolving digital landscape.

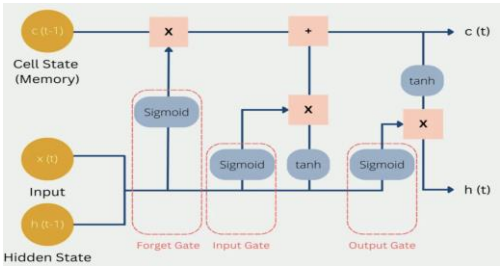
In conclusion, the realm of deepfake detection has witnessed extensive research efforts, yet the challenge of real-time detection persists. This research paper takes a

significant stride towards addressing this limitation by presenting "DeepSecure," a real-time deep learning-based system designed to enhance cybersecurity in social media through effective deepfake detection. By leveraging the synergistic potential of LSTM [1] and ResNext CNN [2] models, our study endeavors to bridge the gap and provide a robust solution for the timely identification of manipulated content. This contribution is poised to bolster the resilience of digital platforms against the pervasive threat of deepfakes, fostering a more secure and trustworthy online environment.

3. TECHNOLOGIES USED

3.1 LSTM (LONG SHORT - TERM MEMORY):

A RNN called LSTM [1] is used as it can identify a segment of signals that recurs frequently in the whole temporal signal sequence and LRD in sequential data. It utilizes a memory cell with gating mechanisms, such as the forget gate, input gate, and output gate, to control the flow of information. These gates enable LSTM to selectively retain important information over time, making it suitable for tasks involving sequential data. LSTM has proven to be effective in various applications where context and temporal dependencies are crucial, making it a prominent choice in deep learning models.



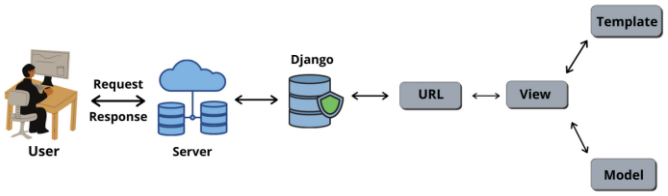
Follow Chart - 1: LSTM Structure

3.2 RESNEXT CNN (RESIDUAL NEXT CONVOLUTIONAL NEURAL NETWORK):

is an advanced variant of the Convolutional Neural Network architecture, featuring a cardinality parameter that enhances its representation learning capabilities. By partitioning input data into multiple groups, ResNext CNN [2] enables the network to learn diverse feature mappings, leading to improved performance in computer vision tasks. Its flexible and scalable design has made it a popular choice in state-of-the-art deep learning models for image recognition and object detection. In the context of DeepSecure, ResNext CNN's spatial understanding complements LSTM's temporal analysis, contributing to robust DeepFake detection and bolstering cybersecurity measures in social media platforms. A thorough exploration of ResNext CNN's applications and advancements in computer vision tasks will provide valuable insights for its role in deep learning-based DeepFake detection.

3.3 Django:

Django, a prominent high-level open-source Python web framework, occupies a significant position in modern web development due to its resilience, scalability, and rapid development capabilities. Built on the principles of Don't Repeat Yourself (DRY) and rapid development, Django empowers developers to create dynamic web applications with exceptional efficiency, embracing the Model-View-Template (MVT) architectural pattern. This framework revolutionizes the development process, facilitating the separation of concerns, enhancing maintainability, and fostering seamless collaboration between components.



Follow Chart - 2: Django MVT Architecture

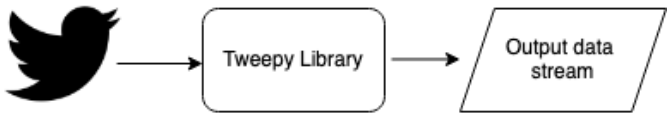
The framework's comprehensive array of integrated functionalities, encompassing intricate aspects such as authentication, database administration, and URL routing, expedites development workflows and enables a concentrated emphasis on the nucleus of application logic. Its unwavering commitment to industry best practices, encompassing impregnable authentication protocols and fortification against prevalent web vulnerabilities, further accentuates its stature as an indispensable tool for crafting secure and multifaceted web solutions.

3.4 TWEEPY:

With Tweepy [6], developers can interact with Twitter data, such as fetching tweets, posting tweets, accessing user information, and performing searches. It's simple and intuitive interface allows for seamless integration with Python applications, making it a popular choice among developers for accessing and analyzing Twitter data.

By leveraging Tweepy's functionalities, researchers and data analysts can gather real-time data from Twitter, enabling them to perform sentiment analysis, track trending topics, and conduct social media monitoring. Tweepy's comprehensive documentation [6] and active community support make it an invaluable tool for accessing and utilizing Twitter data in various applications, from market research to social media analytics. Its ability to handle authentication and rate limiting further simplifies the process of accessing Twitter

data, making it an essential library for anyone working with Twitter data in Python.



Follow Chart - 3: Tweepy Working

4. DEEPSECURE ARCHITECTURE

4.1 MODEL BUILDING

4.1.1 DATA SETS USED

As a real-time deepfake detection system, DeepSecure needed an extremely effective model to distinguish between authentic and fraudulent content on social media postings. We used a range of data sets from several sources, including Deep fake Detection Challenge data [4], and FaceForensic++ [5], in order to make the system reliable. We attempted to get an equal quantity of false and true data. 163 media was gathered.

4.1.2 PRE PROCESSING

In the pre-processing stage, the media go through a variety of steps to filter out background noise and extract the crucial information, namely the face characteristics. A media is split up into frames in the first step of deepSecure preprocessing such that each frame isolates the face and produces a new media that only contains the frames that include faces.

4.1.3 SEPARATING TEST AND TRAIN DATA

We chose to design a distribution of 70% train data and 30% test data in order to make DeepSecure very efficient. The DeepSecure model will be trained using train data, whereas the model may be tested using test data. Equal amounts of false and genuine data are present.

4.1.4 MODEL BUILDING

Convolutional neural networks (CNN) and recurrent neural networks are both integrated into our model's design. The preprocessed frames are now sent to extract features using Trained ResNext CNN (2). After receiving the characteristics from ResNext CNN, we send them to a long short-term memory network [1], which classifies the media as either true or false.

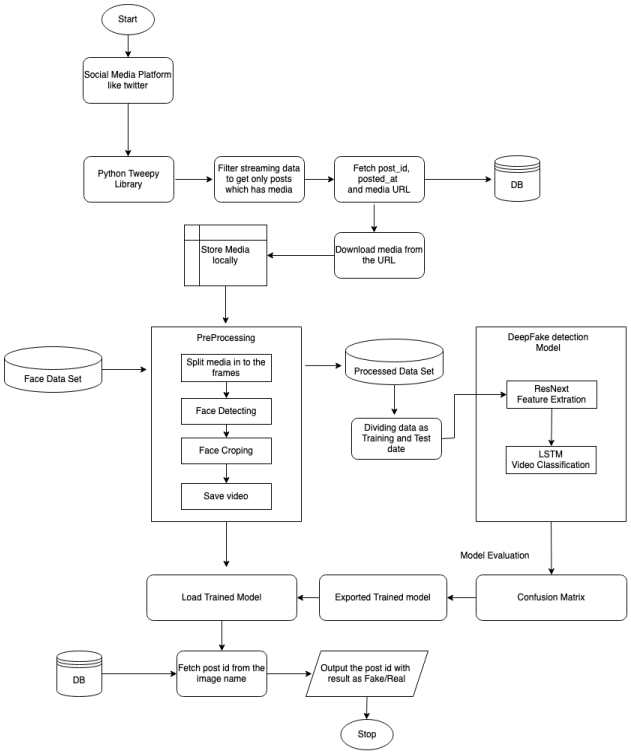
4.2 FETCHING DATA FROM SOCIAL MEDIAL

Data collection from social media networks, in this case Twitter, is the first step towards putting DeepSecure into practice. Twitter offers developers access to its APIs so it can retrieve tweet data. Using Tweepy [6], a Python module that allows for the retrieval of tweets based on keywords, we are able to receive the stream of Twitter data into our system.

Once we fetch the stream of data from Twitter, we filter the data stream to get only the tweets which include media, this can be done by filtering the data based on the media_url key in the data received from the Twitter stream. After getting all the posts which contain media we save the details like post_id, posted_at and the media_url in a centralized database and also download the media from the URL present in the media_url key. This locally downloaded media will be checked for deepfake.

4.3 RUNNING MODEL ON REAL-TIME SOCIAL MEDIA DATA

The model that was built in point 4.1 is now given the data that we received in point 4.2. The authenticity of locally downloaded material is verified before it is remapped to the matching post_id stored in the central database and the final outcome is shown to the users in real time.

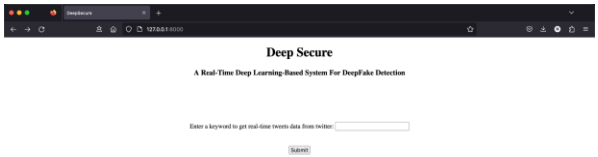


Follow Chart - 4: DeepSecure Architecture

5. RESULTS AND PERFORMANCE

The DeepSecure framework demonstrated remarkable proficiency in the seamless retrieval and subsequent processing of real-time Twitter data, thereby enabling a meticulous differentiation between authentic and spurious media artifacts. This intricate analysis culminated in the

attainment of an outstanding accuracy rate of 95.83%, accompanied by an impressively low loss value of 0.177. These compelling performance indicators serve as a testament to DeepSecure's exceptional prowess as a cutting-edge real-time deepfake detection system. The substantiality of these findings underscores the system's robustness and underscores its potential significance in bolstering cybersecurity efforts in the realm of social media.



Screenshot - 1: DeepSecure Home Page

Screenshot - 2: DeepSecure Prediction Page

6. CONCLUSIONS

In conclusion, "DeepSecure" is a cutting-edge solution for bolstering cybersecurity in social media by detecting and mitigating DeepFake content. DeepSecure leverages the power of both LSTM [1] and RestNext CNN [2] to make a efficient and robust real-time deepfakes detection system With its balanced dataset and integration of CNN and RNN components, DeepSecure offers real-time protection against deceptive content, ensuring a safer and more secure digital landscape for social media users. This research paper contributes valuable insights into the domain of DeepFake detection and highlights the potential of DeepSecure as a reliable tool for fortifying cybersecurity in social media platforms.

REFERENCES

[1] Ralf C. Staudemeyer, "Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks", arXiv:1909.09586v1 [cs.NE] 12 Sep 2019

[2] Jeremy Jordan, "Common architectures in convolutional neural networks.", 19 Apr 2018

[3] SallyAdee, "What Are Deepfakes and How Are They Created?" IEEE Spectrum

[4] Kaggle, "Deep Fake Detection Challenge"

[5] Kaggle, "FaceForensics++"

[6] Joshua Roesslein, "tweepy Documentation Release 3.6.0", Mar 02, 2018