Applied GIS

ISSN: 1832-5505

Vol-10, Issue-01, 2022

DEEP FAKE VIDEO DETECTION USING DEEP LEARNING

Ms. S. Chandra Priyadharshini, Ms. R. Roopa, Mr. S. Satheesh, Mr. S. Sugavanam

Assistant Professor [2,3], Associate Professor [1,4]

chandrapriyadharshini.s@actechnology.in, roopa.r@actechnology.in,

ssatheesh@actechnology.in, sugavanam.s@actechnology.in

Department of CSE, Arjun College of Technology, Thamaraikulam, Coimbatore-Pollachi Highway, Coimbatore, Tamilnadu-642 120

ABSTRACT

Deep fakes have recently exploded in popularity; they are films and photographs that have been digitally changed to seem quite genuine. Researchers are looking at a lot of amazing applications for this technology. There has been a recent uptick in the fraudulent usage of films online, including those promoting false news, pornographic movies featuring celebrities, and financial schemes. Therefore, the Deep fake detection problem is especially harmful to politicians, celebrities, and other prominent people. Many algorithms based on deep learning have been proposed to identify deep fake films or photographs, and a great deal of study has been conducted in the last few years to comprehend the inner workings of deep fakes. Research in this area has focused on deep fake creation and detection systems that use various deep learning techniques. Also covered will be the limitations of existing methods as well as the societal availability of databases. The development of an automated, deep fake detection system. The world faces a major challenge due to the absence of an efficient deep fake detection system, since deep fake films and pictures may be easily made and circulated. Still, many have tried to solve this problem, and the ones that include deep learning have shown to be more effective than the more conventional methods. With these features, we can train ResNext to detect when a video has been manipulated and when it hasn't, as well as to spot the time discrepancies between frames shown by DF introduction tools.

Keywords: Deep Fakes, Deep Learning, Fake Generation, Fake Detection, Machine Learning.

I.INTRODUCTION

ISSN: 1832-5505
Vol-10, Issue-01, 2022

Motivation

This study provides a comprehensive evaluation of deep fake production and detection systems that use several deep learning algorithms. In addition, we will look at how accessible databases are to different groups of people and how current approaches are lacking in some areas. A reliable automated method for detecting deepfakes. The ease with which deepfake videos and photographs may be made and circulated makes the lack of an effective deep fake detection system a serious danger to the world community. But there have been a lot of attempts to fix this, and solutions connected to deep learning outperform the old ways.

Problem definition

Due to the huge loss of frame content during video compression, existing deep learning algorithms for image identification cannot effectively detect bogus videos. The severe deterioration of the frame data following video compression prevents the majority of image recognition techniques from being employed for videos. Additionally, videos provide a problem for techniques intended to identify only still fake images since their temporal features vary across sets of frames.

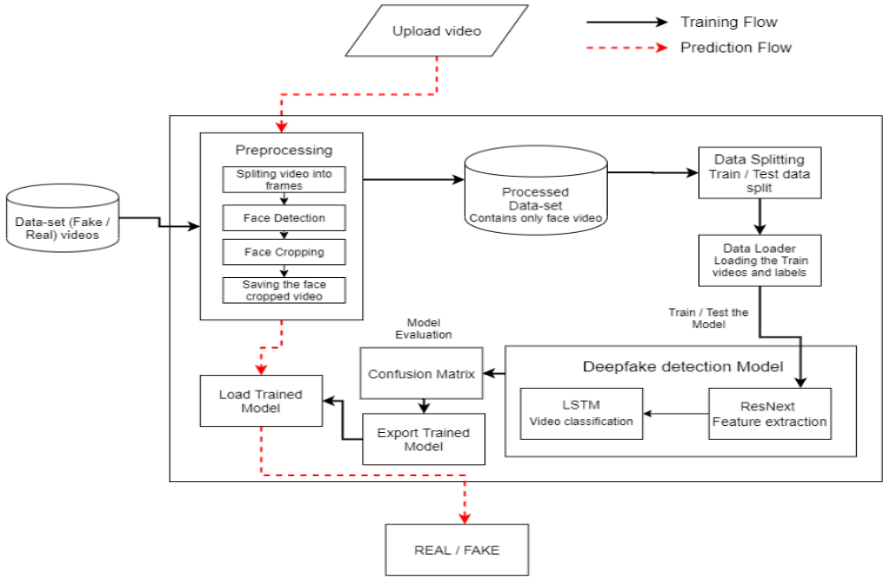
Objective of project

A structure that will allow low-level face manipulation flaws to manifest as interframe inconsistencies and temporal distortions. On the other hand, most of the face shots used by deep learning algorithms have wide eyes, and you won't find many pictures of individuals with closed ones online. Consequently, without images of real individuals blinking, deep fake algorithms can't make up faces that blink often. To rephrase, the blink rates of deep fakes are far lower than those of ordinary videos.

Scope of project

Research and development in the domain of deep learning for the detection of deep fake pictures and videos is vital and constantly growing. Technological developments and the social effects of deep fake technologies are both covered by this expansive area of study. When thinking about deep fake detection utilising deep learning algorithms, there are a few important things to keep in mind:

ISSN: 1832-5505
Vol-10, Issue-01, 2022



II.EXISTING SYSTEM

The self-consistency of local source features—details of images that are spatially-local but independent of content—was recently presented by Zhao et al. as a technique for deep fake detection. Pairwise self-consistency learning is the special representation learning technique that a convolutional neural network (CNN) model uses to retrieve the source features from down-sampled feature maps. By doing so, we want to discourage matching feature vectors with locations in the same image that have low cosine similarity scores. The use of technologies that immediately output the whole image and maintain consistent source characteristics across each point within each image might provide a disadvantage when dealing with misleading photographs.

What are known as "DeepFake" (DF) films are recent creations that use free software techniques based on deep learning to create convincing face swaps in videos with little signs of manipulation.

While it's true that digital video manipulation using visual effects has been around for a while, the advent of deep learning has made it more easier and more realistic to generate such material.

Disadvantages of existing system:

- Since, fake image-based methods use error functions for real or fake image detection. For video, it needs lots of computational power and

ISSN: 1832-5505
Vol-10, Issue-01, 2022

is hence time-consuming by using such methods.

- Some poorly created deep fake videos keep some visual artifacts behind, which can be used for deepfake detection. Thus we can group methods used for classification based on classifiers used i.e either deep or shallow.

III.PROPOSED SYSTEM

When it comes to DeepFakes, there are plenty of tools for making them, but very few for detecting them. In order to prevent the DF from spreading throughout the internet, our method for recognising it will be an invaluable asset. To help users determine whether a video is authentic or not, we are developing a web-based platform. Building a web-based platform is often the first step in this project's escalation to creating a browser plugin for automated DF detections. If major apps like WhatsApp and Facebook want to make it simple to identify DF before transmitting it to another user, they may include this project into their own programme. Assessing its efficacy and acceptance in regard to safety, ease of use, precision, and dependability is a key aim. Our approach is designed to identify several forms of DF, including interpersonal, retrenchment, and replacement DF.

.

Advantages of proposed system

- Deep learning has shown considerable achievement in the identification of deep fakes.
- In order to recognize fake videos & photos properly must be enhanced current deep learning approaches.
- It primarily covers classic detection methods as well as deep Learning based methods such as CNN, RNN, and LSTM.

IV.MODULES:

Dataset: It is essential to have access to real-world data in order to construct any deep learning or machine learning model. For starters, we gathered information from many sources, such as Celeb-DF[8], FaceForensic, and Kaggle's Deepfake Detection competition. There are three thousand films in the DeepFake detection challenge on Kaggle, with half of the data being authentic and half being altered. Celeb-DF features videos of well-known celebrities; out of a total of

ISSN: 1832-5505
Vol-10, Issue-01, 2022

1000 videos, 500 are authentic and 500 have been digitally altered.The FaceForensic ++ dataset has two thousand films, with one thousand being authentic and the other thousand being altered. The next step is to combine the three datasets and send them on to data preparation.

Data Preprocessing: One crucial step is data preparation, which involves attempting to extract useful information from raw data. From the original data, we remove any extraneous information. One aspect of preparing the dataset is dividing the video into frames. After that, we crop the frame to just include the one that has a face in it. Finding the average of the video dataset allows us to maintain frame consistency, therefore we build a new processed face cropped dataset with the same amount of frames as the mean. Preprocessing does not take into account frames that are faceless. Lots of processing power is required for a 10-second video at 30 frames per second, which is 300 frames in total. To keep things simple and allow for more experimentation, we suggest training the model with only the first 100 frames.

Model: The model is made up of resnext50 32x4d and one LSTM layer. The Data Loader loads the preprocessed face cropped films and divides them into two groups: train and test. In addition, the frames from the processed videos are supplied to the model in tiny batches for training and testing.

Res Next CNN for Feature Extraction: Rather than rebuilding the classifier, we suggest utilising the ResNext CNN classifier to consistently recognise frame-level properties and extract features. The next step is to adjust the network's parameters, such as the number of layers and the learning rate, to make sure the model's gradient descent is convergent. LSTM for Processing Sequences: Picture a two-node neural network that takes as input the likelihood that the given sequence is either a deep fake video or an unaltered video and outputs a series of ResNext CNN feature vectors derived from the input frames. Designing a model that can meaningfully recursively process a sequence is the primary difficulty we need to tackle. We suggest a 2048 LSTM unit with a 0.4 dropout probability for our assignment, which is sufficient to accomplish our objective. By comparing the frame at 't' second with the frame at 't' second, the LSTM is used to perform a temporal analysis of the video via sequential examination of the frames.

ISSN: 1832-5505
Vol-10, Issue-01, 2022

Predict: The trained model is given a new video to forecast. A fresh video is also preprocessed to incorporate the trained model's format. The video is divided into frames, then face cropped, and instead of keeping the video locally, the cropped frames are sent immediately to the trained model for identification.

V.ALGORITHMS:

Long short-term memory (LSTM):
One design used in deep learning is long short-term memory, which is an artificial recurrent neural network (RNN). In contrast to traditional feedforward neural networks, LSTM incorporates feedback connections. It is capable of processing both discrete data points (like pictures) and continuous data sequences (like audio or video). Some examples of jobs that LSTM can handle are unsegmented, linked handwriting recognition, voice recognition, and intrusion detection systems (IDSs) or network traffic anomaly detection.

A typical long short-term memory (LSTM) unit has a cell, an input gate, an output gate, and a forget gate. With the help of the three gates, the cell is able to control the influx and outflow of information and remember values for indeterminate durations.

Since there may be delays of undetermined length between significant events in a time series, LSTM networks excel at processing, classifying, and generating predictions using this kind of data. A common issue while training classic RNNs is the vanishing gradient, which led to the development of LSTMs as a solution. In contrast to RNNs, hidden Markov models, and other sequence learning techniques, LSTM is relatively unaffected by gap length, which is a huge plus in a wide variety of contexts.

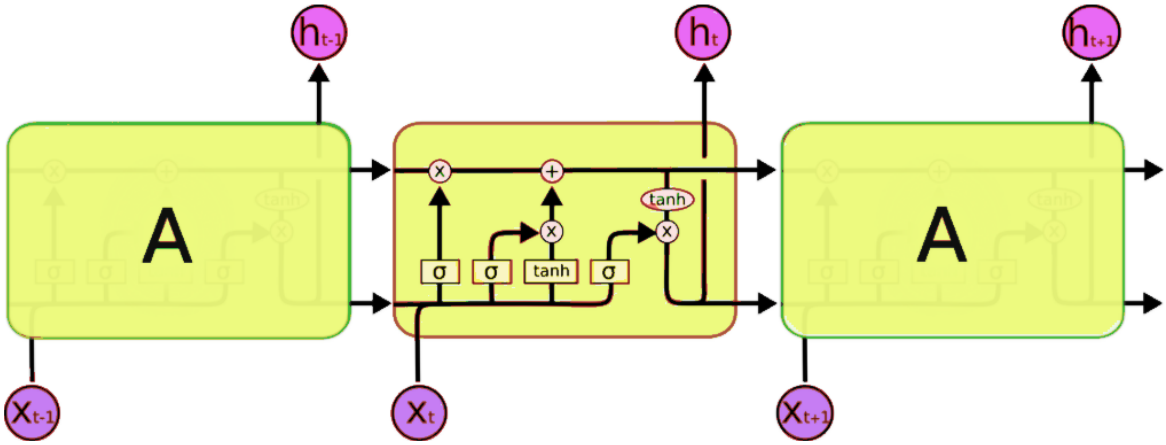
Training:

An RNN using LSTM units can be trained in a supervised fashion, on a set of training sequences, using an optimization algorithm, like gradient descent, combined with backpropagation through time to compute the gradients needed during the optimization process, in order to change each weight of the LSTM network in proportion to the derivative of the error (at the output layer of the LSTM network) with respect to corresponding weight.

A problem with using gradient descent for standard RNNs is that error gradients vanish exponentially quickly with the size of the time lag between important events. However, with LSTM

ISSN: 1832-5505
Vol-10, Issue-01, 2022

units, when error values are back-propagated from the output layer, the error remains in the LSTM unit's cell. This "error carousel" continuously feeds error back to each of the LSTM unit's gates, until they learn to cut off the value.

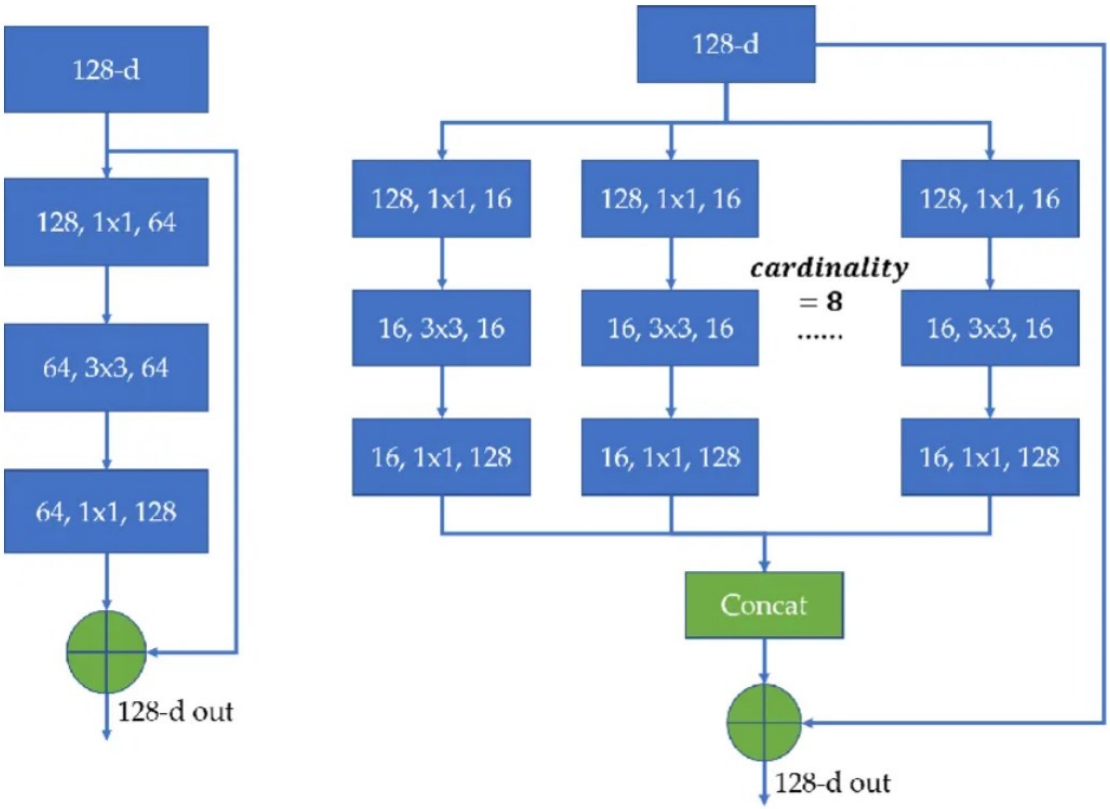


ResNeXt:

ResNeXt is a Convolutional Neural Network (CNN) architecture, which is a deep learning model. ResNeXt was developed by Microsoft Research and introduced in 2017 in a paper titled "Aggregated Residual Transformations for Deep Neural Networks."

ResNeXt uses the basic ideas of the ResNet (Residual Network) model, but unlike ResNet, it uses "groups" instead of many smaller paths. These groups contain multiple parallel paths, and each path is used to learn different features. This allows the network to learn more features more effectively, increasing its representational power.

ISSN: 1832-5505
Vol-10, Issue-01, 2022



The main features and advantages of ResNeXt are:

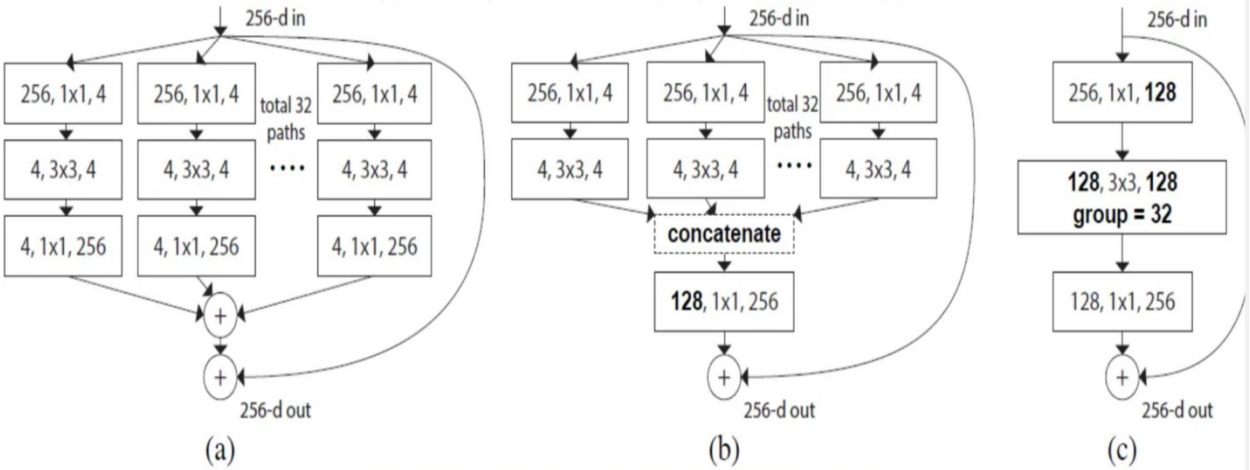
Parallel Paths: ResNeXt is based on the use of multiple parallel paths (or groups) in the same layer. This allows the network to learn a broader and more diverse set of features.

Depth and Width: ResNeXt combines two basic methods, both increasing the depth of the network and increasing the width of the network by increasing the number of groups in each layer. This allows using more parameters to achieve better performance.

State-of-the-Art Performance: ResNeXt has demonstrated state-of-the-art performance on a variety of tasks. It has achieved successful results especially in image classification, object recognition and other visual processing tasks.

Transfer Learning: ResNeXt can be effectively used to adapt pre-trained models to other tasks. This is important for transfer learning applications.

ISSN: 1832-5505
Vol-10, Issue-01, 2022



ResNeXt is used in many application areas, particularly in deep learning problems working with visual and text data, such as image classification, object detection, face recognition, natural language processing (NLP) and medical image analysis. This model performs particularly well on large data sets and is also a suitable option for transfer learning applications.

VI.CONCLUSION:

Many different deep-learning methods for deep-fake photos and movies have been developed by different researchers. Deep fakes gained traction as a result of the widespread availability of visual content in social media content. Especially on social media platforms where users may easily disseminate such false information, this is of the utmost importance. In response to this issue, a plethora of deep learning-based methods have been published that successfully detect and remove bogus media. Fake images and movies may be easily created using the programmes and technology covered in the first section. Next, we'll go over the various methods utilised to create deep fake films and photos in Section 2. Give specifics about the assessment measures and datasets utilised for deep fake detection as well. The quality of deep fakes has been rising, even if deep learning has been successful in detecting them. Present deep learning methods need improvement to accurately identify phoney films and images.

Along with the confidence of the suggested model, we offered a neural network-based approach to categorise

ISSN: 1832-5505
Vol-10, Issue-01, 2022

the video as deep fake or real. Frame stage detection is handled by our method using ResNext CNN, while video class is handled by LSTM. Based on the factors stated in the research, the suggested technique successfully detects whether the movie is a deep fake or real. When applied to real-time data, we anticipate that it will provide very precise results.

VII.REFERENCES

[1] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.

[2] Y. Bengio, P. Simard, and P. Frasconi, "Long short-term memory," IEEE Trans. Neural Netw, vol. 5, pp. 157–166, 1994.

[3] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016. [4] S. Hochreiter, "Ja1 4 rgen schmidhuber (1997)."long short-term memory"," Neural Computation, vol. 9, no. 8.

[5] M. Schuster and K. Paliwal, "Networks bidirectional reccurent neural," IEEE Trans Signal Proces, vol. 45, pp. 2673–2681, 1997.

[6] J. Hopfield et al., "Rigorous bounds on the storage capacity of the dilute hopfield model," Proceedings of the National Academy of Sciences, vol. 79, pp. 2554–2558, 1982.

[7] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, et al., "Google's neural machine translation system: Bridging the gap between human and machine translation," arXiv preprint arXiv:1609.08144, 2016.