International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Fake Video Detection: A Deep Learning Based Approach

Sumit Tokhare[1], Prof. Priyanka Choudhary[2]

Abstract: With increasing use of smart phones and social media applications, the sharing of videos has become extremely common. This has also led to the sharing of fake videos. Due to the size and complexity of the data being shared, it is almost infeasible for manual detection of video forgery. Since the data size to be analysed by time critical applications is enormous indeed, therefore the conventional techniques prove to be infeasible to detect fake videos with high level of accuracy, which primarily leads the focus to artificial intelligence and machine learning tools for the same. The proposed approach presents a frame decomposition, statistical feature extraction and machine learning based approach for fake video detection. The convolutional neural network has been used in this approach for detection of fake videos. The system achieves an accuracy of 98.5% with convergence at 1170 iterations. The results indicate that the proposed approach attains higher classification accuracy compared to existing techniques.

Keywords: Video Forgery, Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Networks (CNN), Classification Accuracy

I. INTRODUCTION

Recent technological developments have exponentially increased the amount of visual data (billions of images and videos) generated every day on the web and by social networks. Facebook, Twitter, YouTube and Instagram are the most popular online websites enabling people to upload and share billions of pictures. Nowadays, social media websites are playing a more important role in our daily life. They help users to express themselves, make new friendships and share their interests and ideas with others. In the digital multimedia era, digital forensics is becoming an emerging area of research thanks to the large amount of image and video files generated. Ensuring the integrity of such media is of great importance in many situations. This task has become more complex, especially with the progress of symmetrical and asymmetrical network structures which make their authenticity difficult. Consequently, it is absolutely imperative to discover all possible modes of manipulation through the development of new forensics detector tools. Although many solutions have been developed, tamper-detection performance is far from reliable and it leaves this problem widely open for further investigation

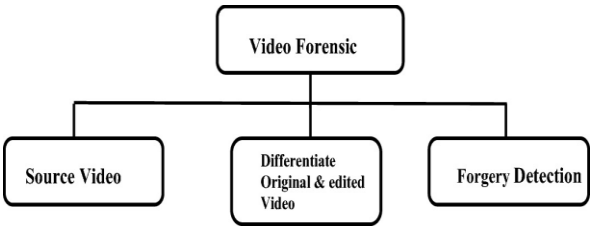

Fig. 1 The Video Forgery Model

The term video attack has gained attention under the name video forgery. The simplest type of video forgery is copy-move tampering which can be detected by human eyes. The complex type of video forgery is video falsifying which is more professional than copy-move as highly improved techniques are needed to detect a falsified video. The difficulty of detecting video falsifying attack because of changing the semantic meaning of the original videos by creating fake videos can be conducted by editing, combining or generating a new video content.

Multimedia Tampering: Image (or video) tampering can be defined as the action of "adding or removing

important features from an image (or video) without leaving any obvious traces of tampering" . Generally, the most common applied tampering operations are: (i) deleting (or hiding) a region in the image, (ii) adding a new object into the image, and (iii) misrepresenting the image information (e.g., resizing an object within the image). Despite this problem of digital forensics that has attracted much attention, however, most research in this area still lacks rigorous and solid results and discussions. In addition, several methods have apparent limitations and are difficult to implement.

II. DEEP FAKES

A growing unease has settled around evolving deepfake technologies that make it possible to create evidence of scenes that never happened .Deepfakes are fake videos or audio recordings that look and sound just like the real thing. Concerns about deepfakes have led to a proliferation of countermeasures. New laws aim to stop people from making and distributing them. Earlier this year, social media platforms including Facebook and Twitter banned deepfakes from their networks. And computer vision and graphics conferences teem with presentations describing methods to defend against them. Deepfakes technology can seamlessly stitch anyone in the world into a video or photo they never actually participated in. Such capabilities have existed for decades. Deepfake (stemming from "deep learning" and "fake") is a technique that can superimpose face images of a target person to a video of a source person to create a video of the target person doing or saying things the source person does. The underlying mechanism for deepfake creation is deep learning models such as autoencoders and generative adversarial networks, which have been applied widely in the computer vision domain

The deepfkae videos are depicted in figure 2.

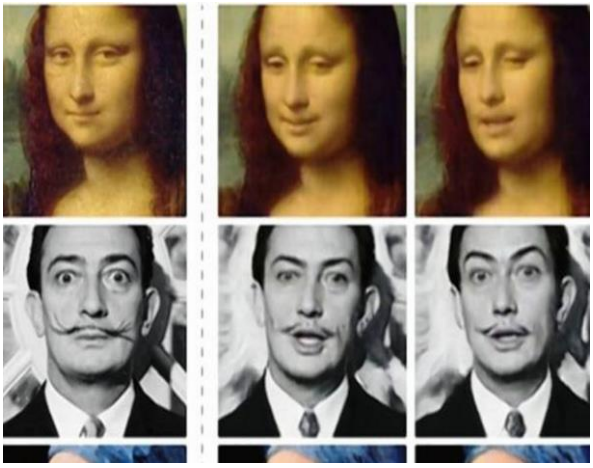

Fig.2 Example of Using DeepFake to alter Facial Expressions

The figure above illustrates the fact that the images which are under consideration are generally edited using sophisticated tools which makes it extremely difficult to recognize the forgeries with naked eyes. Moreover, for real time applications, automated systems need to be designed which can work in the bound time limits to meet the requirements of real time critical applications.

III. PROPOSED METHODOLOGY

As mentioned earlier, detection of deep fakes of morphed videos is a challenging task typically due to the following reasons:

1) Copious amounts of data to be analysed since video files are typically contain the maximum amount of data compared to all data formats.
2) Video editing tools have become extremely sophisticated in the recent years.
3) The amount of time needed to analyse the video data is non-trivial.

Hence artificial intelligence and machine learning based approaches are needed to detect fake videos and make the detection mechanism relevant for time critical applications. Today, the danger of fake news is widely acknowledged and in a context where more than 100 million hours of video content are watched daily on social networks, the spread of falsified video raises more and more concerns. While significant improvements have been made for image forgery detection, digital video falsification detection still remains a difficult task. Indeed, most methods used with images cannot be directly extended to videos, which is mainly due to the strong degradation of the frames after video compression. . Current video forensic studies mainly focus on the video re-encoding and video

recapture however video edition is still challenging to detect. For the last years, deep learning methods have been successfully employed for digital image forensics
The various machine learning algorithms are typically categorized as:

1) Unsupervised Learning: In this approach, the data set is not labelled or categorized prior to training a model. This typically is the most-crude form of training wherein the least amount of apriori information is available regarding the data sets.

2) Supervised Learning: In this approach, the data is labelled or categorized or clustered prior to the training process. This is typically possible in case the apriori information us available regarding the data set under consideration.

3) Semi-Supervised Learning: This approach is a combination of the above mentioned supervised and unsupervised approaches. The data is demarcated in two categories. In one category, some amount of the data is labelled or categorized. This is generally not the larger chunk of the data. In the other category, a larger chunk of data is unlabeled and hence the data is a mixture of both labelled and unlabeled data groups. The neural networks often tend to analyse extremely complex data patterns. For neural architectures which need to analyse complex image based data, the deep neural networks (DNN) are used which happen to have multiple hidden layers. It is shown in the figure 3.

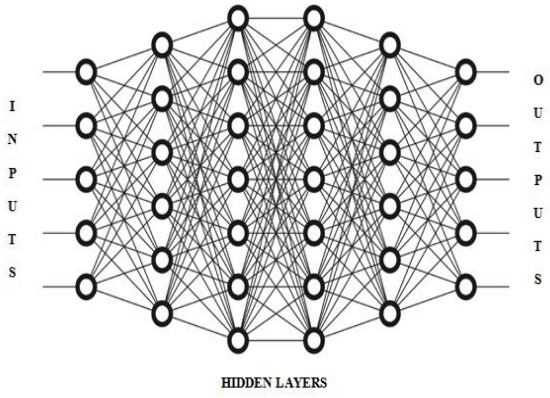

Fig.3 Internal structure of Deep Neural Network

Generally, it is used for extremely large and complex datasets with highly uncorrelated patterns in the data. In general, any type of video forgery is done by using encoding techniques on the frames. A generic model for generating forged video frames in depicted in fig. 4

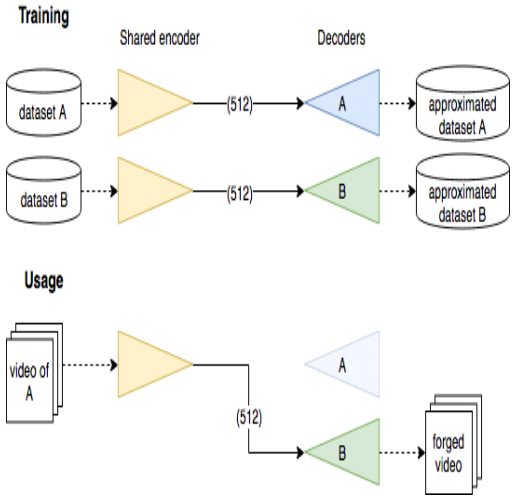

Fig.4 The process of Forged Video

It can be observed that shared encoders process the frames to generate composite fake videos. Thus it becomes customary to split the images into frames for the detection of fake videos. Statistical features of the frames can also augment the classifier's accuracy of classification. The CNN is an extremely effective deep learning based classifier which performs pattern recognition in each of its layers based on stochastic computing. The fundamental operation in the CNN hidden layers is the convolution operation mathematically given by:

$$x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$

Here,
x(t) is the input
h(t) is the system
y is the output
*is the convolution operation in continuous domain
For a discrete or digital counterpart of the data sequence, the convolution is computed as:

$$y(n) = \sum_{-\infty}^{\infty} x(k)h(n-k)$$

Here,
x(n) is the input
h(n) is the system
y is the output
*is the convolution operation in discrete domain

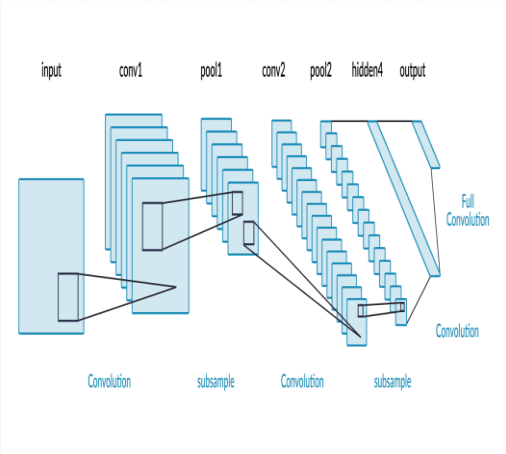

Fig.5 The structure of the CNN

A CNN has the following salient features:

1) **Strided convolution**: While conventional convolution is an overlap between the system and the data, strided convolutions help in covering all the data samples rather than just the internal samples of the data matrix. The stride over is just a hop in convolution. Mathematically, for an (n x n) and (f x f) convolution, if 'p' is the number of strides, then the number of samples in the output are:

$$Y_{Samples} = \left(\frac{n+2p-f}{s}+1\right)\left(\frac{n+2p-f}{s}-1\right)$$

Here,
n is the input sample matrix dimension
f is the system sample matrix dimension
p is the stride length

2) **Pooling and Max Pooling:** The pooling is an operation to make the features more robust and reduce the dimensionality. Typically, max-pooling is employed.
3) **Employing Weighted Gradient Descent:** The gradient descent is used as the most common and effective cost function optimization based CNN training algorithm. It is given by:

$$w_{k+1} = w_k - \alpha\frac{\partial e}{\partial w}$$

Here,
w_{k+1} is the weight of the next iteration
w_k is the weight of the present iteration
e is the error
α is the learning rate

IV. SIMULATION RESULTS

This section illustrates and explains the results obtained. The steps depict the frame separation followed by the CNN based classification.

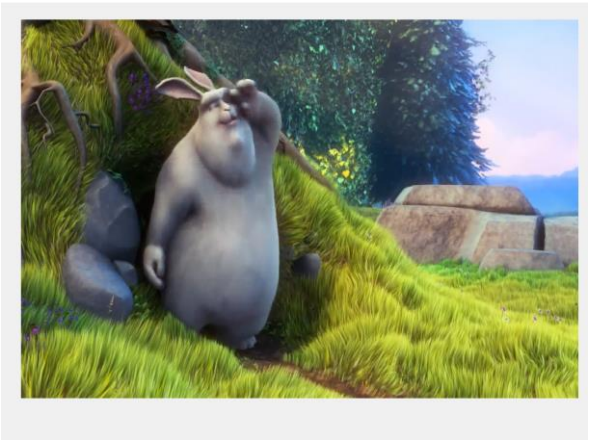

Fig. 6 Loaded Video Frame

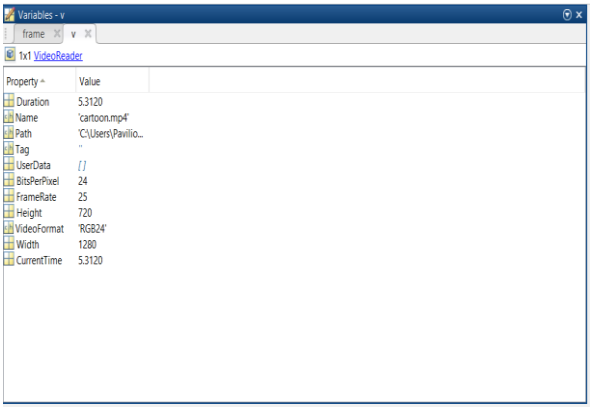

Fig. 7 Frame Properties

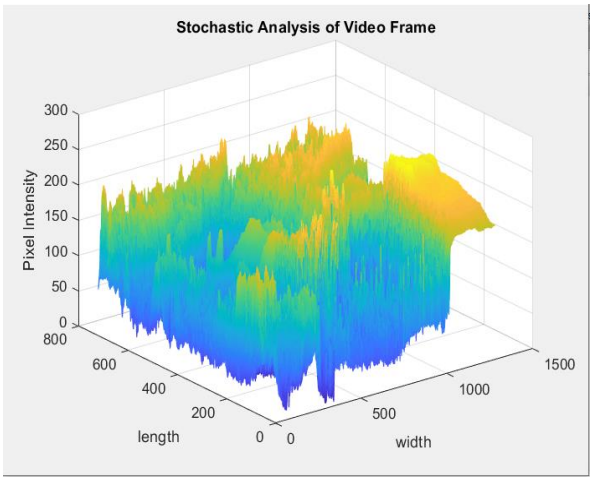

Fig. 8 Stochastic Analysis of Video

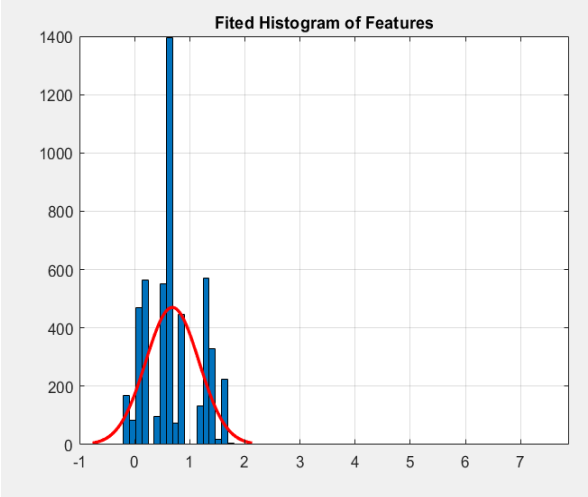

Fig. 9 Histogram of Features

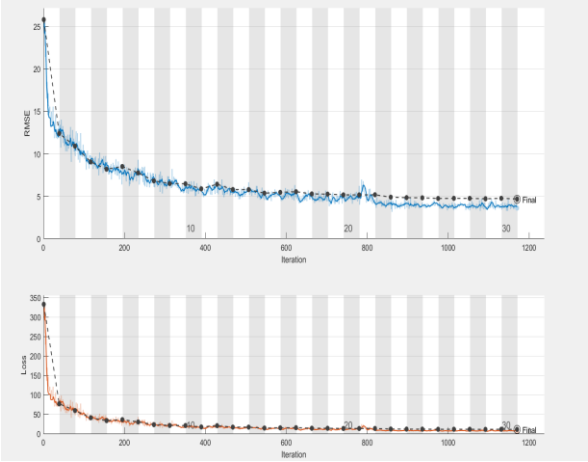

Fig.10 CNN Training

The figure depicts the variation of the root mean square error (rmse) of the system as a function of epochs. It can be seen that the training converges with the rmse stabilizing around 1200 iterations.

Table.1 Summary of Results

S.No.	Parameter	Value
1.	Deep Leaning Model	CNN
2.	Time	46s
3.	Iterations	1170
4.	Learning Rate	0.0001
5.	Hardware Resource	Single GPU
6.	Accuracy	98.5%

Conclusion: It can be concluded from pervious discussions that there has been a rapid increase in sharing of fake videos with emergence of social media applications. Due to the size and complexity of the data being shared, it is almost infeasible for manual detection of video forgery. Since the data size to be analysed by time critical applications is enormous indeed, therefore the conventional techniques prove to be infeasible to detect fake videos with high level of accuracy, which primarily leads the focus to artificial intelligence and machine learning tools for the same. The proposed approach presents a frame decomposition, statistical feature extraction and machine learning based approach for fake video detection. The convolutional neural network has been used in this approach for detection of fake videos. The system achieves an accuracy of 98.5% with convergence at 1170 iterations. The results indicate that the proposed approach attains higher classification accuracy compared to existing techniques.

References

[1] S Suratkar, F Kazi, “Deep fake video detection using transfer learning approach”, Arabian Journal for Science and Engineering, Springer, 2023, vol.48, pp. 9727–9737.
[2] A. Hamza et al., "Deepfake Audio Detection via MFCC Features Using Machine Learning," in IEEE Access, vol. 10, pp. 134018-134028, 2022.
[3] A Mitra, SP Mohanty, P Corcoran, E Kougianos, “A machine learning based approach for deepfake detection in social media through key video frame extraction”, SN Computer Science, Springer 2021, vol.2.no.98.
[4] J. K. Lewis et al., "Deepfake Video Detection Based on Spatial, Spectral, and Temporal Inconsistencies Using Multimodal Deep Learning," 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington DC, DC, USA, 2020, pp. 1-9.
[5] Darius Afchar, Vincent Nozick, Junichi Yamagishi, Isao Echizen “MesoNet: a Compact Facial Video Forgery Detection Network”, IEEE 2019
[6] Pavel Korshunov; Sébastien Marcel et al., “Vulnerability assessment and detection of Deepfake videos”, IEEE 2019
[7] Faten F. Kharbat; Tarik Elamsy; Ahmed Mahmoud; Rami Abdullah., “Image Feature Detectors for Deepfake Video Detection”, IEEE 2019

[8] Miljan Đorđević; Milan Milivojević; Ana Gavrovska., "DeepFake Video Analysis using SIFT Features", IEEE 2019
[9] David Güera; Edward J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks", IEEE 2018
[10] Luciano Floridi.," Artificial Intelligence, Deepfakes and a Future of Ectypes", Springer 2018.
[11] Ali Khodabakhsh; Raghavendra Ramachandra; Kiran Raja; Pankaj Wasnik; Christoph Busch, "Fake Face Detection Methods: Can They Be Generalized?", IEEE 2018.
[12] Lichao Su; Cuihua Li; Yuecong Lai; Jianmei Yang, "A Fast Forgery Detection Algorithm Based on Exponential-Fourier Moments for Video Region Duplication",IEEE 2017.
[13] Omar Ismael Al-Sanjary , Nurulhuda Ghazali, Ahmed Abdullah Ahmed, Ghazali Sulong, "Semi-automatic Methods in Video Forgery Detection Based on Multi-view Dimension", Springer 2017
[14] Sami Bourouis; F. R. Al-Osaimi; Nizar Bouguila; Hassen Sallay; Fahd Aldosari; Mohamed Al Mashr "Video Forgery Detection Using a Bayesian RJMCMC-Based Approach", IEEE 2017
[15] Chee Cheun Huang; Ying Zhang; Vrizlynn L. L. Thing, "Chee Cheun Huang; Ying Zhang; Vrizlynn L. L. Thing", IEEE 2017
[16] G Lynch, FY Shih, HYM Liao, "An efficient expanding block algorithm for image copy-move forgery detection", Elsevier 2013
[17] M Hussain, G Muhammad, SQ Saleh, AM Mirza, "Image forgery detection using multi-resolution Weber local descriptors", IEEE 2013
[18] MF Hashmi, AR Hambarde, "Copy move forgery detection using DWT and SIFT features", IEEE 2013
[19] G Muhammad, M Hussain, G Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform", Elsevier 2012
[20] W Fan, K Wang, F Cayre, Z Xiong, "3D lighting-based image forgery detection using shape-from-shading", IEEE 2012
[21] M Hussain, G Muhammad, SQ Saleh, "Copy-move image forgery detection using multi-resolution weber descriptors", IEEE 2012
[22] H Yao, S Wang, Y Zhao, X Zhang, "Detecting image forgery using perspective constraints", IEEE 2011
[23] G Muhammad, M Hussain, K Khawaji, "Blind copy move image forgery detection using dyadic undecimated wavelet transform", IEEE 2011
[24] H Yao, S Wang, Y Zhao, X Zhang, "Detecting image forgery using perspective constraints", IEEE 2011