International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Fake Detection using LSTM and Survey of Deep Fake Creation Technologies

Shilpa Pant[1], Chhaya Gosavi*[2], Sheetal Barekar[3]

Submitted: 28/09/2023 **Revised:** 14/11/2023 **Accepted:** 26/11/2023

Abstract*:* Artificial intelligence known as Deep Fake is one of many techniques that have been successfully developed in recent years for altering faces in images and videos. It can produce convincingly faked images, audio, and video. Deep Fake can create problems, especially when there is a media component involved. Even if it is helpful, when it is used maliciously, such as for disseminating fake news or cyberbullying, it can pose a threat to society. It is necessary to develop a complete fake detection method to handle such issues. Too far, numerous methods have been developed to distinguish between authentic and fraudulent videos. The objective of this work is to give a summary of different approaches for Deep Fake creation and to provide an overview of LSTM algorithms for deep fake video detection.

Keywords: Convolutional Neural Networks; Deep Fake; Deep Learning; GAN; LSTM

1. Introduction

This template, modified in MS Word 2007 and saved as a "Word 97-2003 Document" for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: ease of use when formatting individual papers, automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and conformity of style throughout a manuscripts. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

Huge advancements have been made in the field of image and video editing over the last several years. In court and during police investigations, videos, photos, and audio snippets are frequently utilized as evidence. These bits of information were ambiguous and inaccurate due to the sophisticated technologies used in deep fake creation. Because of this, before being offered in court, this kind of evidence needs to be scrutinized. In recent times, Deep Fake has been the most popular tampering method. Deep Learning technology is where the term "Deep Fake" originated. "Deep fakes" are produced by deep learning algorithms that are trained to swap out faces [1].

Deep fake can be used to switch the face of one person for another. In many instances, this method can be used to fabricate movies of a celebrity, alter the face expressions, or even in politics. This enables one to fully modify the speaker's identity [2].

People can easily use a variety of free tools, such as FaceApp [3], Wombo [4], Deepfakesweb, etc., and use them to produce media with the desired results [5]. Many pictures and video data are often required to train the models for deep fake techniques.

Celebrities and politicians can be found online in a huge number of videos and photos due to their status as public figures. The deep fake is made in this initial step. Deep Fake can be used to produce hate speeches. It can lead to conflict between countries, political groups, and even religions. It can fabricate financial news to deceive the public. It can fabricate financial news to deceive the public. Using such fake videos and digital material can also cause disruption in the armed forces, the defense, and other sectors [6].

Deep fakes can be used effectively to produce visual effects in movies. Examples of this technique include Snap chat filters, giving individuals their lost voices, and updating films without reshooting them. Face App is a well-known selfie editor that offers several AI filters, background changes, and other effects [7].

Fig. 1. Real Image (left) and its corresponding deep fake (right) [8]

Deep fakes are frequently employed for malicious purposes as well; this negates their advantageous uses. Deep fakes were most recently employed during the Ukraine-Russian war to disseminate false information and propaganda. Videos of several Ukrainians who supported Russia went popular on social media. They were eventually discovered to be profound fakes [8].

Researchers and forensic experts can use deep fake detection tools to confirm the authenticity of viral videos. Profound fakes could consequently represent a danger to commoners as well as open

[1] *Comp Dept. Cummins College of Engineering for Women, Pune, India*
ORCID ID : 0000-0002-6709-5548
[2] *Comp Dept. Cummins College of Engineering for Women, Pune, India*
ORCID ID : 0000-0002-2348-0698
[3] *Comp Dept. Cummins College of Engineering for Women, Pune, India*
ORCID ID : 0000-0001-5532-3884
* *Corresponding Author Email: chhaya.gosavi@cumminscollege.in*

pioneers. For example, a President was swindled of $243,000 utilizing a voice deep fake [9]. A Chinese app called Zao has gone viral, which enables any users to Place them in well-known films or TV clips by putting their faces on the bodies of movie stars [10].

Accordingly finding reality in the digital domain turns out to be seriously difficult. Deep fakes can be identified by a variety of methods. Most of them make use of deep learning. The development and detection of deep fakes are discussed in detail in this paper. It examines various methods for identifying deep fakes. Additionally, the paper also covers the implementation, results, and application of the LSTM method in deep fake detection.

2. Deep Fake Creation

The videos created by utilizing Deep Fake methods have become extremely famous these days. The nature of such videos is awesome, and they show up extremely practical. These videos can be easily created by anyone with a basic understanding of computers. Many of the applications for this make use of deep learning. Deep learning can be used to represent complex data and is best suited to the production of high-quality videos [11].

To create deep fakes of this kind, numerous online tools are available. The resultant videos, movies made by using such tools can range from simple or basic ones that are easily faked at a look to heavily edited ones that not even a skilled viewer can spot. A type of convolutional neural network is used for this, called an Auto-encoder. An Auto-encoder learns to efficiently compress and encode input images and then reconstruct the information from the compressed, encoded representation as close as possible to the input representation [12].

Figure 2 is a deep fake example where decoder B is linked to the feature set of face A to reconstruct face B from the original face A.

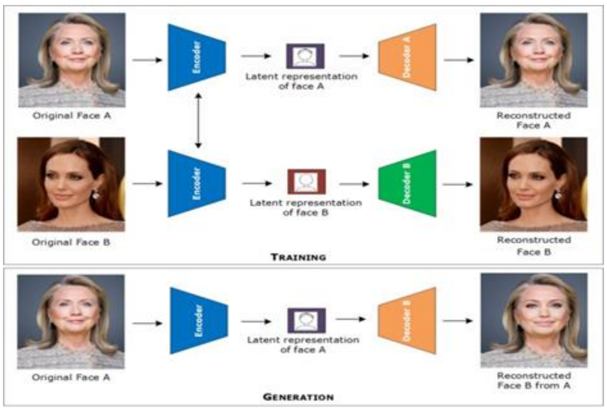

Fig. 2. Creation of a Deep fake using an Auto encoder and Decoder [11]

which can accurately create facial features and then apply them to the videos. A Reddit user created the fake app utilizing the auto encoder-decoder pairing structure. This strategy is used in several works, such as Tensor flow-based deep fakes Deep-Face Lab, DFaker, and Deep Fake it [14].

Table 1 gives a Summary of different Deep Fake Tools [15].

Generative Adversarial Network (GAN) is an improvement to this approach. GAN raises the level of Deep Fake creation quality. Figure 3 displays a general block design of GAN used to create fake images.

Deep learning is used by the program Face Swap to identify, and swap faces in images and movies. Few-Shot Face Translation is another tool which can produce faces that have its gaze direction; glasses being consistent with the given source face [16]

Table 1 Deep Fake Tools

Tools	Key Elements
Faceswap	Free and Open-Source multi-platform. Two encoder-decoder pairs are used.
Faceswap-GAN	An auto encoder architecture is enhanced using adversarial loss and perceptual loss (VGGface).
Few-Shot Face Translation	Face swapping video from a single image without training for any face.
DeepFaceLab	An open source project. Adding new models to the Faceswap technique
DFaker	Built with the Keras library. To reconstruct the face, one uses the DSSIM loss function.
AvatarMe	The first technique is that can create photorealistic 3D faces from a single "in-the-wild" photograph. Can create accurate 3D faces with a resolution of 4K by 6K from a single low-quality photograph.
MarioNETte	A framework that conceals the target's identity for a few-shot face recreation. Identity adaptation does not require a further fine-tuning phase.
StyleRig	Creates photorealistic portraits of faces with context, eyes, teeth, and hair. Self-supervised without manual annotations
“Do as I Do” Motion Transfer	Synthesizes new videos of people in a different context. Can produce a dancing video with many characters that moves in time

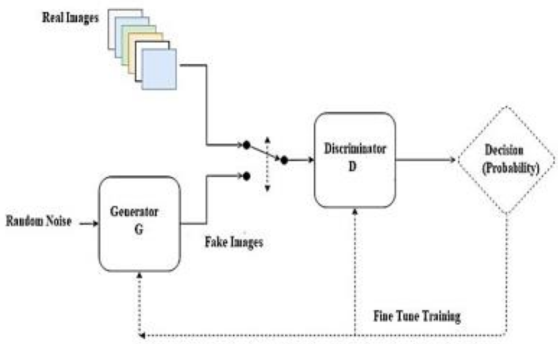

Fig. 3. Generative Adversarial Network general block diagram [13]

3. Deep Fake Detection

Deep Fakes pose a growing threat to democracy, public safety, and individual privacy. The detection of deep fakes has been made much easier by deep learning. Deep Fakes stands out from other methods of video manipulation because it can produce results that look like photos. The final videos can be very convincing if there are enough photos of both performers and enough time spent training on computers. To create such videos does not require any expertise [17].

3.1. Background Comparison

The online GIF archive website "Gyfcat" has demonstrated scalable AI-based counter Deep Fake initiatives. The Background Comparison method is one of the very basic approaches to detect Deep Fake films. The service uses facial recognition models to identify issues with how a video's facial region is rendered.

This approach has several drawbacks. Backgrounds could be wholly fabricated and even composites made with entirely fresh video wouldn't be able to be recognized [18].

3.2. Temporal Pattern Analysis

Sequence-based information has been examined as a validation strategy since a person's activity is best described as a period arrangement of motions. Experts integrated a long short-term memory (LSTM) and a convolutional neural network (CNN) to comprehend temporal successions (LSTM). The network has the possibility to learn explicit development-based habits of its subjects by skipping each frame of a movie into a CNN and producing a series of highlight maps for the LSTM [19].

It has been suggested to use individual distinct combinations of (CNNs) to recognize video action. Most existing models were designed for cropped videos, which can only analyze brief segments and capture segment-level motion information; sequential learning and logic are inappropriate. Recurrent and convolutional tasks that are handled in a small area are probably not suitable for determining the long-extend temporal structure. Furthermore, because of noise gradients and fading, recurrent networks have had significant difficulty learning lengthy sequences of video scenes. An excellent, novel network to approach this problem is the LSTM class of recurrent neural networks (RNNs) [20].

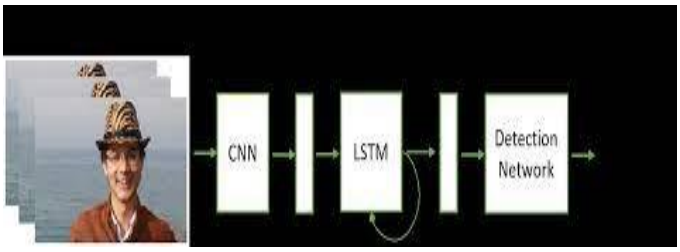

Fig. 4. A deep fake detection technique using extended short-term memory and convolutional neural networks (CNN)

3.3. Facial Artifacts

Recently, experts from UC Berkeley and Adobe worked together to create a technique for identifying intentionally altered photographs by identifying mild facial distortion. An approval accuracy of 99% might be achieved as a human eyewitness by prepping a CNN on examples of images managed using the well-known Face-Aware liquefied feature (53%). While these results are encouraging, the lack of GAN-produced models typically indicates that such a network is only able to recognize manually altered photos, making it worthless for detecting Deep Fakes [21].

Their strategy was dependent on Deep fake algorithms' propensity to produce fixed sizes with lesser resolutions for computational efficiency due to resource and processing time constraints. The Deep fake algorithm can only create facial photos of a specific size. Following that, these images would go through scaling and affine transformation. Based on that, it ought to go through an affine warping to match the source's face arrangement [22].

Preparing and training the CNN model while expediting the process by accurately recreating the resolution discrepancy in affine face warping is very tedious. These steps include extracting the transform matrices to change the default configuration to the faces, first separate the faces and landmarks. Using the opposite of the assessed change grid, the original picture is returned from the affine warped face after adjusting the face with Gaussian obscuring [23].

3.4. Mesoscopic Analysis

Experts from Tokyo's National Institute of Informatics have demonstrated how sophisticated neural networks constructed with an artificially low number of layers are suitable for identifying minute differences detected in Deep fakes at high computational effectiveness and with this 90% and more accuracy is achieved. The presentation of simple networks hinges on how a limited number of layers and an unexpectedly small number of parameters encourage the identification of smaller, more basic examples with the decreasing number of convolution layers of the image during a forward pass [24].

The Mesoscopic level of research is used in this method to identify altered faces in videos. In a compressed video situation where the picture noise is severely deteriorated, microscopic analysis exams based on the noise of the image cannot be used. Likewise, the human eye struggles to detect manipulated images at a higher level of the semantic. The reason is that this strategy suggests using a powerful neural network with fewer layers to take a middle-of-the-road approach. The following two models, with a low level of representation and a high degree of complexity, have achieved the top classification scores across all tests on a remarkably small number of factors [25].

3.5. Head Pose Estimation

Researchers at the University of Albany have demonstrated that using landmark focuses, it is possible to discern systematically between straightforward Deep Fake outputs and the objective facial posture. With blurrier photos, the model's presentation was deemed bad since landmark assignment would also suffer. The method relies on Deep Fakes, which are produced by splicing a synthetic human face area into the original image observations. Common errors can be found when three-dimensional head positions are assessed using photos of human faces. A support vector machine (SVM) classifier is evaluated using several Deep fakes and real photos using highlights based on this question [27].

To distinguish Deep fakes from real videos or images, additional trained SVM classifiers are dependent on the variances between head postures that are evaluated using the whole set of authority landmarks and those in the focal face districts [28].

The following techniques are used to extract the features.

- Programming bundle DLib is utilized to execute a face detector that concentrates 68 facial landmarks for each video frame or image.
- Following that, the entire face is analyzed individually, and the head poses from the focused face region using the equivalent of 68 points from OpenFace2 and the common 3D facial landmark model.
- Detaching by its standard deviation.

3.6. Eye Blinking

This method locates Deep fake films by spotting the lack of eye blinking in fake faces. The brief opening and closing of the eyelid are referred to as eye blinking. The average human blink lasts between 0.1 and 0.4 seconds. Humans blink at intervals of 2 to 10 seconds. If the typical video scene length is 1/30 second, then there is a 7.5% chance of getting a picture of someone blinking. Since most internet-published photographs do not feature people with their eyes closed, a Deep Fake movie can be identified by the absence of blinking eyelids. This technique makes use of long-term recurrent CNN (LRCN), an RNN and CNN hybrid.

The difference between consecutive frames is how Torricelli et al. analyzed the condition of the eyes. Deep Neural Network techniques for the detection of eye blinking are not yet known. This approach utilizes an LRCN-based classifier that is CNN-based. To spot synthetic faces in Deep fake movies, the strategy of watching for eye blinking is rather simple [29].

Table 2 summarizes Common Deep fake Detection Techniques.

4. Implementation

As part of the proposed architecture, Efficient Net B0 serves as a feature extractor. The vision transformer examines the features that Efficient Net B0 extracted. The vision transformer eventually produces the likelihood of face modification [30]. The proposed model used an LSTM network to analyze temporal discrepancies in the sequences after using the Inception V3 model as a feature extractor. 300 deep fake videos from various websites and 300 real videos from the HOHA dataset were used to train the model. The CNN model is used for feature extraction and LSTM for feature analysis [31].

Due to the way it manages long-term dependencies and resolves the vanishing gradient issue that other RNNs have, the LSTM is a special sort of recurrent neural network. Since a movie is made up of a series of frames, LSTMs can effectively detect deep fake videos. A 1792-wide LSTM unit is utilized for sequence processing after Efficient Net B4 is employed for feature extraction [32].

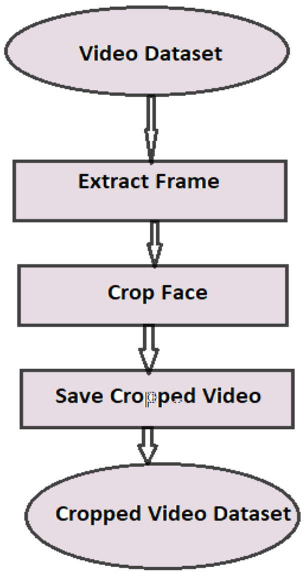

Fig. 5. Pre-processing Sequence

A Fully connected layer is added after the LSTM unit to determine whether the videos are authentic or not. Fig.5 shows pre-processing steps and Fig.6 shows LSTM architecture.

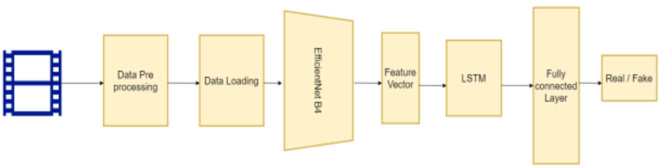

Fig. 6. LSTM System Architecture

Fig.7 shows training and testing flow. Firstly, checked if the cropped video dataset has corrupted videos and remove them. Then, the cropped video dataset was split into three sets – training set, validation set, and testing set in the 80:10:10 ratio. Three corresponding instances of data loader are created for training, testing and validation. The data loader extracts frames from the cropped videos and stacks them together to create the corresponding tensor of the video. The stacked tensor and label of the video is loaded in the model.

Before loading the data in the model, the frames are cropped again to suit the input requirements of our models:

- LSTM – 112x112
- EfficientNet B1 – 240x240
- EfficientNet B2 – 260x260
- EfficientNet B3 – 300x300
- EfficientNet B4 – 380x380

Table 2. Summary of Deep Fake Detection Techniques

Methods	Classifiers/ Techniques	Key Features	Dataset	Media
Xception Net CNN [24]	Convolutional Recurrent Neural Network	Robustness is tested for both compression and out-of-sample inference	FaceForensics++, Celeb-DF video datasets and ASVSpoof 2019 Logical Access audio dataset.	Audio / Video
MesoNet [25]	CNN	On the deep fake and FaceForensics datasets, accuracy was achieved at a rate of 98% and 95%, respectively.	Two datasets: the FaceForensics one, produced using the Face2Face method, and the DeepFake one, constructed from online videos	Videos
Head poses [27]	SVM	68 landmarks from the dataset are used to extract features face area.	- The 49 deep fake videos that make up UADFV are accompanied by corresponding actual videos.	Videos / Images
Eye, teach and facial texture [31]	Logistic regression and neural network	Use the variations in facial texture and the lack of reflections and details in the areas around the eyes and teeth in deep fakes.	A video dataset downloaded from YouTube.	Videos
Eye blinking [29]	Learn the temporal patterns of the eyes using LRCN. Based on the observation that Deepfakes blink far less frequently than is typical.	Highlighted how the human blinking rate in Deepfake films differs from that in real videos.	Real: 49 real interview and presentation videos Fake: generated from above-mentioned videos	Videos

For both EfficientNet and LSTM, trained model on 20 epochs with a learning rate of 0.00001. Adam optimizer was used for reducing the loss and modifying the weights during the learning process. To avoid over fitting, a weight decay of 0.00001 was used. Cross Entropy Loss was used to measure the loss after every epoch. For training, a mini-batch size of 4 was used. After training and validation was concluded, and saved the model and tested it on testing dataset.

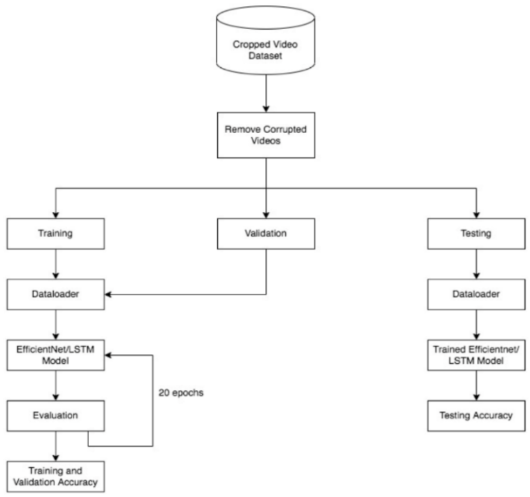

Fig. 7. Training and Testing flow

4.1. Algorithm Details

Step 1: FRAME EXTRACTION
1. Give I/P path of the video
2. Capture Video.
3. Store the path of the folder where all the videos are stored.
4. Convert data frame to csv.

Step 2: CROP VIDEOS
Step 3: TRAINING
Input to the training function – number of epochs, data loader, model, criterion
Step 4: TESTING
Step 5: FORWARD LOOP OF LSTM
Input to the forward loop – stack of tensor with dimensions
(Batch size, sequence length, channels, height, width)
1. Reshape tensor. New dimensions – (batch size ∗ sequence length, channels, height, width)
2. Pass tensor to EfficientNet B4 for feature extraction. EfficientNet B4 returns a feature.
vector of 1792 features.
3. Perform average pooling of feature map
4. Reshape tensor to dimensions (batch size, sequence length, 1792)
5. Pass the reshaped tensor to LSTM model
6. Pass output of LSTM to linear layer
7. Return output of linear layer and feature map

5. Results and Discussions

Results obtained for the LSTM method are as shown in Table 3. The accuracy for the sequence length of 30 is 73.62 % and for the sequence of length 50 it is more than 80 %. So, the highest accuracy obtained with the use of the LSTM model is for the sequence length of 50.

Table 3. Deep Fake Detection using LSTM.

Sequence Length	Training Accuracy	Testing Accuracy
30	92.57%	73.62%
40	93.46%	76.66%
50	97.12%	80.37%

There is still scope to add more videos to the dataset, allowing us to get a well-balanced dataset with lots of footage.

Table 4 shows the results obtained for EfficientNet experiments. Highest accuracy was obtained for EfficientNet B2 with two fully connected layers. This contrasts with results obtained for the ImageNet dataset. For the ImageNet dataset, EfficientNet B4 had the highest accuracy among EfficientNet B1, B2, B3 and B4. This shows that there can be differences in the results for transfer learning and results for pre trained models. It is also observed that each EfficientNet model with two fully connected layers gives a higher accuracy than the corresponding model with three fully connected layers.

Table 4. EfficientNet results

Model	No of Fully Connected Layers	Training Loss	Validation Loss	Testing Accuracy
B1	3	0.000288	0.002155	67.77%
	2	0.000272	0.002164	70.00%
B2	3	0.000297	0.002052	74.43%
	2	0.000274	0.002056	75.55%
B3	3	0.000295	0.002371	67.03%
	2	0.000268	0.002003	71.48%
B4	3	0.000288	0.002064	70.37%
	2	0.000271	0.002282	72.96%

The deep fake detection accuracies of LSTM are quite good. If we compare RNNs model for the same, it is observed that RNNs are better suited for analyzing sequences [33]. That is why LSTM models can predict more accurately than Efficient Net. Fig. 8 shows a confusion matrix for the implemented model.

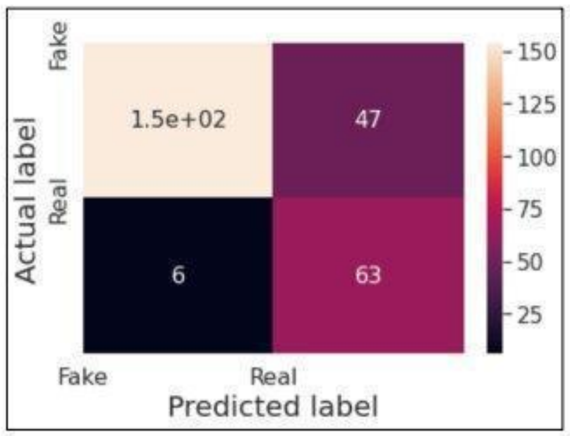

Fig. 8. Confusion Matrix for result using LSTM

6. Conclusion and Future Work

It will probably soon be impossible to identify fake videos because Deep fake technology is moving closer to producing fake content of noticeably higher quality. Therefore, it's crucial to act quickly and carefully in response to the threat that deep fakes offer.
One or two effective tools are insufficient to tackle deep fake on their own. Integrating detection is a different area of study technologies into distribution channels like social media to boost its efficiency in addressing the deep fakes' extensive effects. While it's vital to use detection techniques to identify Deepfakes, it's even more crucial to comprehend the motivations behind those who post Deepfakes.

Digital data creation and manipulation are now simpler than in the past because of technological advancements. Technology improvements have made it easier than ever before to create and manipulate digital data. LSTM with a sequence length of 50 was tested. The precision of the LSTM model still has room for improvement. If the model is

trained on all the movies from these three datasets, it will grow more accurate and reliable. Performing hyper parameter tuning for LSTM by extending the sequence length to values greater than 100 is another intriguing area of research.
Thus, combining CNN with LSTM can give better results.

Acknowledgements

We would like to express our gratitude to the Cummins College of Engineering's Computer Engineering Department for its assistance and encouragement.

Author contributions

Shilpa Pant[1]: Conceptualization, Background study, Methodology. Chhaya Gosavi[2]: Data collection, Implementation, Writing-Original draft preparation, Result Validation. Shital Barekar[3]: Figures and Tables, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] "Deep Fakes Web," [Online]. Available:https://deepfakesweb.com/. [Accessed 15 May 2022].
[2] Y. Li, X. Yang, P. Sun, H. Qi and S. Lyu, "Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics," in IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.
[3] "Face App," [Online]. Available:https://www.faceapp.com/. [Accessed 15 May 2022].
[4] "Wombo," [Online]. Available:https://www.wombo.ai/. [Accessed 15 May 2022].
[5] Y. Pashaeva, "Scammers Are Using Deepfake Videos Now," Slate, 13 September 2021. [Online]. Available: https://slate.com/technology/2021/09/deepfake-video-scams.html. [Accessed 15 May 2022].
[6] Chesney, Robert and Citron, Danielle Keats, DeepFakes: A Looming Challenge for Privacy, Democracy,and National Security (July 14, 2018). 107 California Law Review (2019, Forthcoming); U of Texas Law,Public Law Research Paper No. 692; U of Maryland Legal Studies Research Paper No. 2018-21.
[7] Jessica and Silbey Woodrow Hartzog , "The Upside of Deep Fakes", Maryland Law Review, Volume 78 issue 4,2019
[8] Schwartz, Oscar (12 November 2018). "You thought fake news was bad? Deep fakes are where the truth goes to die". The Guardian.Sik-Ho Tsang, "Review: Inception-v3 — 1st Runner Up (Image Classification) in ILSVRC 2015" ,https://medium.com/@sh.tsang/review-inception-v3-1st-runner-up-image-classification-in-ilsvrc-2015-17915421f77c.
[9] A Voice Deepfake Was Used To Scam A CEO Out Of $243,000 https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/#31bb21a32241, September 3, 2019
[10] Josh Horwitz, https://www.reuters.com/article/us-china-tech-zao-idUSKCN1VN0G9.
[11] Momina Masood1 & Mariam Nawaz & Khalid Mahmood Malik & Ali Javed & Aun Irtaza & Hafiz Malik."Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward." 11 May 2022
[12] Thanh Thi Nguyen, Quoc Viet Hung Nguyen, Cuong M. Nguyen, Dung Nguyen, Duc Thanh Nguyen, Saeid Nahavandi ,"Deep Learning for Deep Fakes Creation and Detection: A Survey",26 April 2021.
[13] Remya Revi K.1 , Vidya K. R.1 and M. Wilscy," Detection of Deepfake Images Created Using Generative Adversarial Networks – A Review ", February 2021
[14] PavelKorshunov, Sebastien Marcel, "DeepFakes: a New Threat to Face Recognition? Assessment and Detection", citing arXiv: 1812.08685[cs.CV], 20 Dec 2018.
[15] M. Grothaus, "Deepfake videos are the latest cruel form of school bullying – parents and teachers must watch out," iNews UK, 9 November 2021.[Online]. Available:https://inews.co.uk/news/technology/deepfake-videos-school-bullying-cyberbullying-ai-apps-parents-teachers-children-1290664. [Accessed 15 May 2022].
[16] Crescendo, "YouTube," 16 April 2020. [Online].Available:https://www.youtube.com/watch?v=xkqflKC64IM [Accessed 02 April 2022].
[17] A. Jadhav, A. Patange, J. Patel, H. Patil and M.Mahajan, "Deepfake Video Detection Using Neural Networks," International Journal for Scientific Research and Development, vol. 8, no.1, 2020.
[18] Abbreviated View of Deepfake Videos Detection Techniques, Mohammed A. Younus ,Taha M. Hasan.6 th international engineering conference" Sustainable Technology and Development" , (IEC-2020), Erbil , Iraq
[19] N. Bonettini, E. D. Cannas, S. Mandelli, L. Bondi,P. Bestagini and S. Tubaro, "Video Face Manipulation Detection Through Ensemble of CNNs," in 2020 25th International Conference on Pattern Recognition (ICPR), 2021.
[20] D. Güera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018.
[21] Yuezun Li, and SiweiLyu, "Exposing DeepFake Videos By Detecting Face Warping Artifacts",arXiv:1811.00656v3 [cs.CV] 22 May 2019 https://doi.org/10.1016/S0969-4765(19)30137-7.
[22] FalkoMatern , Christian Riess and Marc Stamminger,"Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations ", 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW) https://doi.org/10.1109/WACVW.2019.00020
[23] DariusAfchar, Vincent Nozick, Junichi Yamagishi and Isao Echizen, "MesoNet: a Compact Facial Video Forgery Detection Network", arXiv:1809.00888v1[cs.CV] 4 Sep 2018 https://doi.org/10.1109/WIFS.2018.8630761
[24] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess,J. Thies and M. Nießner, "FaceForensics++:Learning to Detect Manipulated Facial Images," in IEEE/CVF International Conference on Computer Vision, 2019.
[25] A. Karandikar, V. Deshpande, S. Singh, S.Nagbhidkar and S. Agarwal, "Deepfake VideoDetection Using Convolutional Neural Network,"International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, 2020.
[26] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in arXiv 10.48550/ARXIV.1409.1556, 2014.
[27] Xin Yang ,Yuezun Li and SiweiLyu, "Exposing Deep Fakes Using Inconsistent Head Poses" , ICASSP 2019 -2019 IEEE ICASSP,17 May 2019
[28] A. Jadhav, A. Patange, J. Patel, H. Patil and M.Mahajan, "Deepfake Video Detection Using Neural Networks," International Journal for Scientific Research and Development, vol. 8, no.1, 2020.
[29] Yuezun Li, Ming-Ching Chang and SiweiLyu,In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking, arXiv:1806.02877v2 [cs.CV] 11Jun 2018.
[30] A. Chintha et al., "Recurrent Convolutional Structures for Audio Spoof and Video Deepfake Detection," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 5, pp. 1024-1037, Aug. 2020, doi: 10.1109/JSTSP.2020.2999185.
[31] A. A. Pokroy and A. D. Egorov, "EfficientNets for DeepFake Detection: Comparison of Pretrained Models," in IEEE Conference of Russian Young. Researchers in Electrical and Electronic Engineering, 2021 .
[32] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in International Conference on Machine Learning, 2019.
[33] Nguyen, H.H., Yamagishi, J., Echizen, I. "Capsule-Forensics Networks for Deepfake Detection", Handbook of Digital Face Manipulation and Detection. Advances in Computer Vision and Pattern Recognition. Springer, Cham, 2022