Earthline Journal of Mathematical Sciences
E-ISSN: 2581-8147
Volume 14, Number 1, 2024, Pages 119-132
https://doi.org/10.34198/ejms.14124.119132

A Lightweight Algorithm for Detecting Fake Multimedia Contents on Social Media

Arnold Mashud Abukari[1,*], Jhansi Bharathi Madavarapu[2] and Edem Kwedzo Bankas[3]

[1] Department of Computer Science, Tamale Technical University, Ghana
e-mail: amashud@tatu.edu.gh

[2] Department of Information Technology, University of the Cumberlands, USA

[3] Department of Business Computing, C. K. Tedam University of Technology and Applied Sciences, Ghana

Abstract

The significant growth of the fourth industrial revolution (Industry 4.0) coupled with the widespread adoption of social media across the world has initiated new challenges that deserve the attention of researchers and industry leaders especially in detecting and preventing fake multimedia contents on social media. The forging of multimedia contents like videos and images for malicious activities is gradually becoming very rampant and this has serious psychological, health, political and economic consequences on the targeted individuals or close associates of the victims. The application of deepfake algorithms to make manipulated videos and images has contributed in making it very difficult to identify fake videos and images from the real multimedia contents. The availability of the internet and social media has made the spread of deepfake videos and images very fast and at an alarming rate. This research work understanding the dire need to detect deepfake videos and images (multimedia contents) proposes a lightweight algorithm to detect deepfake videos and images on social media platforms. The need for a lightweight algorithm is essential to enable low computational devices to be able to apply the algorithm without computational challenges and overheads. The proposed model has demonstrated a significant reduction in the computational and time $O(n)$ complexities. The research work also presented a comparative analysis of some selected deep learning models with emphasis on the datasets used, their features and challenges identified.

Received: September 23, 2023; Accepted: October 31, 2023; Published: November 16, 2023

2020 Mathematics Subject Classification: 68-XX.

Keywords and phrases: deepfake, algorithm, deep learning (DL), machine learning (ML), artificial intelligence (AI), CNN, social media.

*Corresponding author

Introduction

Social Media in recent years have grown to be one of the most widely used technology platform for sharing information, news, ideologies, advertisements and multi-media contents like videos, images and audio files. Consumers of the social media downloads information in the form of files, videos, images and audios in multiple times without detecting the same content has been downloaded using the same device or identifying that some of the videos and images are manipulated. These downloads of files in multiple forms have serious security, trust and cost implications on the user and the computing device used. Duplicate multi-media and fake contents have the potential to negatively influence the performance of the computing device being it smartphone or a personal computer (PC) due to resource allocation and utilization. This phenomenon has caught the attention of researcher across the globe and techniques such as Content Hashing, finger printing and deep learning models and techniques have been adopted and implemented to prevent the downloads of duplicate multi-media contents on social media. The downloading of duplicate multi-media contents has serious financial cost on the user based on the frequency of downloading duplicate multi-media and manipulated content, the size of the multi-media content, the storage requirement of the multi-media content as well as the bandwidth required to download the multi-media and fake content. Resources such as storage space, bandwidth and processors are significantly wasted due to duplicate multi-media content downloads which eventually lead to increasing the cost of Infrastructure. In the digital space in recent years, forgery of images and deepfake videos are gradually gaining prominence as an emerging form of cyber attacks and this has serious security implications especially users of social media (Karras et al. [3]). The term "Deepfake" is coined from two terms namely "Deep Learning" and "Fake". The Deep Neural Networks (DNN) has made it possible to create and manipulate videos and images faster and easier. According to Mitra et al. [2], replacing a person's face, voice, emotions, speech with another person using deep learning technology is known as deepfake. The manipulation is done in such a way that makes traces of the manipulation very difficult to detect. These deepfake videos come with social, political and economic orientations to target individuals or group of people. The most common platform for defaming people and spreading the deepfake videos through vulnerable people is social media (Mitra et al. [2]). The social media platforms and their users are highly susceptible to digital or cyber attacks in the form of blackmailing and character assassination. Malicious applications developed by cyber criminals are attached to these manipulated videos and eventually compromises unsuspecting social media users. There is a growing interest from researchers in detecting deepfake contents in the digital space especially on social media platforms. The most important step to combat the threats posed by deepfake videos and images is to be able to detect them. Manipulated videos and images can serious implications on a person's mental health and social status (Chou and Edge [4]). The challenge identified by Mitra et al. [2] is the difficulty to work with compressed multimedia contents using uncompressed techniques since uploading multimedia contents on social media comes with compressing and resizing of the contents. This research paper focuses on the handling of compressed multimedia deepfake contents.

Social Media and Deepfake Multimedia contents

Technological advancements in the area of video and image forgery are getting alarming. The ability to create fake videos and images of individuals, political figures, religious leaders, organizations and government agencies are increasingly becoming very difficult to distinguish between real multimedia contents from the fake contents. Artificial Intelligence (AI) has succeeded in making images and videos generated by AI very accessible with the help of social media platforms like Facebook, X (formerly Twitter), Instagram and Tiktok among others. Schoenherr [33] reports that social media has been used as a medium to propagate fake videos and images of “Stylish Pope Francis”, “Elon Musk protesting in New York”, “Donald Trump resisting arrest”, Ukranian President in the midst of the Ukraine-Russia war and many others.

Figure 1: AI generated image of Pope Francis (Schoenherr [33]).

Figure 2: Deepfake video of Ukrainian President (Miller [34]).

Schoenherr [33] alleges that social media influencers are being paid to spread fake multimedia contents and fake news. Vota [36] after analyzing a research work conducted by Farhana et al. [35] concluded that fake videos are perceived to be more harmful compared to other threat models of misinformation shared on social media. There is no solution that addresses all concerns of deepfake multimedia contents and there are varying perceptions from social media users on what constitute fake videos or images. Farhana et al. [35] identified lack of skills and willingness to recognize fake videos as the most harmful and risky aspect of the spread of fake videos and images on social media. Satariano and Mozur [37] argues that a news anchor who reported on "the shameful lack of action against gun violence" in the United State of America (USA) and a female news anchor who "heralded China's role in geopolitical relations at an international summit meeting" were all fake. Satariano and Mozur [37] reveal that their voices were stilted and failed to be in sync with the movement of their lips, with pixelated face and hair appeared unreal. Satariano and Mozur [37] concluded that the two news anchors were not real people but Artificial Intelligence generated software.

Deep Learning

The Machine Learning (ML) concept has gained significant attention among researchers and industry practitioners globally. The ability to make machines learn and perform as required of them based on the algorithms is fascinating and this has increased productivity among the various industries. Machine Learning (ML) has direct or indirect connection to several industries and applications (Tayyab et al. [22]). Machine Learning (ML) is an open access concept model that is applied to learn and train complex models (Vedaldi and Lenc [39]). According to Zuo et al. [23], Deep Learning is a subset of Machine Learning (ML). Deep Learning is the application of well structured and specific sets of algorithms that helps machines in pattern recognition based on historic data set. The pattern recognition is to help machines to make predictions and/or make very intelligent decisions based on the patterns recognized emanating from the historic data. One of the commonest and key algorithms used in deep learning is the Artificial Neural Network (ANN) which is used to exploit a data input in order to predict an unknown output. The Artificial Neural Networks (ANN) is made up of interconnected network of computational units called Artificial neurons (Zuo et al. [23]). The deep learning concept of the Artificial Neural Network (ANN) was first proposed by Rosenblatt [24] in an article titled "The perceptron: A probabilistic model for information storage and organization in the brain" which succeeded in allowing neurons to learn. In 1989, LeCun et al. [25] Convolutional Neural Network (CNN) which established a solid foundation and pave the way for deep learning in computer vision. Deep Learning and Computer Vision models like Long short Term Memory (LSTM) and Recurrent Neural Networks (RNN) were developed after the foundation established by LeCun et al. [25]. The deep learning concept gained more prominence when Hinton et al. [26] proposed and renamed a multilayer neural network-related learning algorithm to Deep Learning which led to an improved computer hardware performance and the development of the Graphics Processing Unit (GPU). Krizhevsky et al. [27] proposed the Deep CNN Architecture which became a very dominant framework for deep learning. Graph Neural Network (GNN) in recent years is also gaining serious attention after Xu et al. [28] and Zhou et al. [29] made serious justifications for an effective graphic learning framework using non-Euclidean graph structured data. Yang et al. [30] has been able to establish a deterministic conversion rule that will convert Convolutional Neural Network (CNN) to Spiking Convolutional Neural Network (SCNN). There have been several deep learning algorithms proposed in recent years especially in the wake of the significant growth of Artificial Intelligence (AI). Deep learning algorithms are a set of pre-trained decision making networks geared towards performing a task most especially mimicking humans. The deep learning technology is a force to reckon with and now considered by some researchers to be a core and essential technology for the fourth Industrial revolution (Industry 4.0) (Sarker [31]).

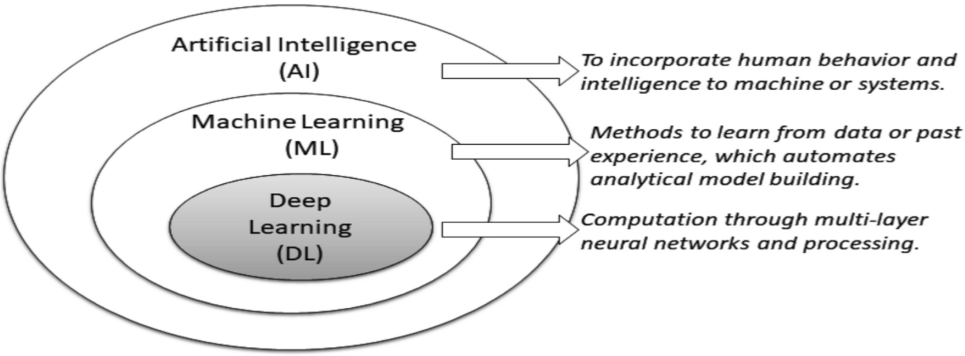

Figure 3: Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) Relationship (Sarker [31]).

Artificial Intelligence (AI) Generated Multimedia Contents

Artificial Intelligence (AI) has contributed greatly to how multimedia contents like videos and images are manipulated and shared on social media. Artificial Intelligence (AI) systems have succeeded in creating "realistic" multimedia contents using a robust three-dimensional (3D) computer graphic in-built model. There has been recent development on the application of deep learning algorithms based on the Generative Adversary Networks (GANs) to generate fake videos. The researchers in Goodfellow et al. [6] first proposed the Generative Adversary Networks by using two networks called generator and discriminator. The generator network according to Goodfellow et al. [6] was used to generate a sample output that were not different from data distribution trained and the discriminator was used to assess the sample generated by the generator network. Radford et al. [8] demonstrated the potential and ability of unsupervised learning by proposing Deep Convolutional Generative Adversary Networks (DCGANs) whiles Isola et al. [9] learn the mapping from input image to output image by applying Conditional Adversarial Networks (CANs) coupled with a loss function to train the mappings. Despite the progress made by Isola et al. [9], Taigman et al. [10] presented Domain Transfer Network (DTN) which was capable of mapping image samples from one domain to another domain. The proposed Domain Transfer Network (DTN) approach by Taigman et al. [10] achieved favorable results and performance when applied to small resolution faces and digital images. Shrisvastava et al. [11] worked and found a way to minimize the gap between real images and synthetic images distribution by applying self-regularization loss and adversarial loss whiles Liu et al. [12] made a case to propose unsupervised image to image translation frame based on the Generative Adversary Networks (GANs) with distribution of images in different domains. Li et al. [5] exposed Artificial Intelligence (AI) generated fake faces using an algorithm that detects eye blinking as a differentiating factor applied in neural networks. Li et al. [5] argues that the blinking of eyes is not well presented in synthesized fake videos and the results of their research shown promising performance in detecting fake multimedia contents like synthesized videos.

Deepfake Detection Algorithms

The human face is one of the key features that identifies an individual since people are recognized largely based on their faces. To attack an individual using fake multimedia contents or deepfake multimedia contents, the attackers or malicious individuals targets the faces of people they intend to attack. Face manipulation and forgery of individuals has increased significantly over the years and gained a lot of attention in recent times. Garrido et al. [13] succeeded and replaced the face of an individual without changing the expressions. In image and video forgery, the expression through the lips is also key in order to get people believe in the fake multimedia content. Suwajanakorn et al. [14] succeeded in syncing the lips of fake or forged video. In 2017, the concept of deepfake was born when deep learning networks was adopted to create some fake videos. The combination of convolutional encoders and the Generative Adversarial Networks (GANs) made the fake videos very real and difficult to differentiate it from the original video. In this digital era, smartphones now have image forgery applications like FaceApp, Reflect face swap, Face swap live, AgingBooth, Meitu and MSQRD geared towards manipulating images and videos. The emergence of the deepfake technology is increasingly making the level of distortion and manipulation of images and videos highly unacceptable. The deepfake technology is highly considered as a disruptive technology that has the ability to change the truth and mislead people at a given time especially on social media platforms. According to Chesney and Citron [15] the deepfake technology is a serious security threat to countries and individuals. Deepfake technology has serious security and ethical issues since it can defame an individual as well as invade their privacy using the technology. News created using images and videos generated from deepfake technology has succeeded in making people losing trust and faith in the images and videos presented during news cast. Deepfake technology has the potential to create serious tension that can destabilize a country and its citizens. The dreams and careers of individuals could be shattered through the deepfake technology. Despite the euphoria on the negative implications of the use of the deepfake technology, the deepfake technology can be used for positive situations like assisting a hearing impaired person to generate lips movement based on an audio. The movie industry can also adopt the deepfake technology in the making of realistic movies with many languages (multilingual movies). Realizing the need to prevent the spread of deepfake multimedia contents in the digital space, technology giant companies namely Microsoft, AWS and Facebook collaborated with partnership on Artificial Intelligence (AI) and Academic Institutions. The main purpose for the collaboration was to build a Deepfake Detection Challenge (DFDC) datasets for researchers to use across the globe. Another technology giant Google also partnered with Jigsaw and presented another dataset called FaceForensics ++ (FF++) for the detection of deepfake videos.

Challenges in Detecting Deepfake Contents on Social Media

Multimedia contents uploaded on social media are compressed and social media users uses different devices including smartphones, tablets and computers among others to access their social media accounts. The existing techniques in detecting deepfake images and videos are largely for uncompressed data and hence do not work well when applied to the social media environment. This paper argues that the detection of deepfake images and videos in the social media environment needs to have a comprehensive model since devices with low computational resources like smartphones are largely used. To detect and extract deepfake images and videos, the model applied needs to be lighter in order to enable low computational device work effectively. This research paper is seeking to develop a model that will detect deepfake multimedia contents like images and videos in their compressed form or lighter versions whiles ensuring a low computational overheads on resources. The research paper presents a comprehensive comparative analysis of some selected deepfake models based on the datasets used, their features and challenges identified to be associated with the selected models.

Table 1: Comparative Analysis of Deepfake models.

Researcher(s)	**Data Set Used**	**Features**	**Challenge (s) / Remarks**
Guera and Delp [17]	Hollywood-2 to Human Actions (HOHA)	Inception-V3 + Long Short Term Memory (LSTM)	Applicable to uncompressed videos
Li et al. [5]	Closed Eyes in the Wild (CEW)	Measured eye blinking rate and used Long Term Convolutional Networks	Applicable to uncompressed videos
Afchur et al. [18]	Internet processed dataset	Meso-4 and MesoInception-4 applied. 2-inception modules, 2-convolution layers and 2-FC layers	Applicable to compressed videos but accuracy is very low.
Nguyen et al. [19]	Four datasets	VGG-19 + Capsule Network	Applicable to compressed videos

			but accuracy is very low for highly compressed videos
Hashmi et al. [20]	Deepfake Detection Challenge (DFDC)	CNN + LSTM. Used facial landmarks	High computational Complexity and works well with long videos (Minimum length of 10 seconds)
Kumar et al. [21]	FaceForensics ++, Celeb-DF	Triplet architecture and metric learning approach	Applicable for highly compressed videos
Mitra et al. [16]	FaceForensics ++	Face artifact analysis, XceptionNet + Classifiers	Applicable to compressed videos with high accuracy

Proposed Deepfake Detection Algorithms

The proposed algorithm for the detection of deepfake multimedia contents is built on the research work presented by Mitra et al. [16] to detect each frame coupled with a proposed deepfake extraction algorithm to help extract key deepfake video frames. The proposed model presented in this paper limits the data based on the frames in the video by concentrating on the extraction of key video frames from each video under consideration. The technique adopted is to reduce the number frames to be checked on each video without compromising on accuracy and this technique reduces the computational overheads on the devices used and thereby making it suitable for low computational devices like smartphones to use. The proposed model presented in this research paper checks only the key video frames for authenticity and this reduces the work load on the computing device used to access the social media platforms. The research work aimed at developing a model to detect deepfake images and videos and this was achieved by using computer vision technique applied in a network of classifiers. Considering I being an input image and F being a feature detector of Filter, the Feature Map is represented as follows:

$$(I * F)(x, y) = \int_0^x \int_0^y I(x-i, y-j)F(i,j)d_i d_j \, , \tag{1}$$

where x and y are the filter value and pixel value of the input image respectively. The equation one (1) is applied to determine the features of the videos or images during the feature vector extraction process. The research work adopted the Least square loss and Least Square Error of the neural network presented by Srivastava et al. [32] as indicated in equation 2 and 3 respectively.

$$L_{SL} = \frac{1}{2}\left(t - \sum_{i=1}^{n} w_i' I_i\right)^2 \tag{2}$$

$$L_{SE} = \frac{1}{2}\left(t - \sum_{i=1}^{n} \delta_i w_i \ I_i\right)^2 \tag{3}$$

Neural networks with dropout layers helps to stop overfitting of the neural networks and the least square loss and least square error are used with activation function $f(x) = x$ to prevent overfitting in the proposed model where δ_i is a Bernoulli random variable. The prediction of the deepfake video or image based on the classification of either being real or fake is determined using equation 4. The proposed model takes in a dimensional vector of the video or image (multimedia content) and creates another vector with values between 0 and 1 with same size as the dimensional vector. Considering the probability of the prediction to be $P(D_n)$ for video or image input data D_n, the predicted output D' is calculated as follows by taking the arguments of the maxima:

$$D' = argmax_{n=2} P(D_n) \tag{4}$$

Algorithm: Proposed Steps to detect fake multimedia content

---

1 **Input**: video or image (v), Proposed Model

2 **Output**: *PredictedTag*

3 Declare and initialize frames(F), specific key frame(K_f), cropped face (C_f) and resized face (R_f) to 0

4 Assign the respective values to F, K_f, C_f and R_f

5 Declare and initialize prediction tags for real video (R_v) and Fake video (F_v) to 0

6 Assign real video probability and fake video probability to R_v and F_v respectively

7 Set *PredictedTag* = False

8 Extract key video frame K_f from v

9 Save the extracted frame to F

10 For $K_f \in F$ *do*

Detect C_f for K_f

Crop the face and save it into C_f

Resize the image and save it into R_f

Load the proposed model

Predict R_f

Set R_v and F_v

If $R_v > F_v$

Continue

Else

Set PredictedTag = True

Consider v as fake

Break

Proposed Model Framework

In this proposed model, a video or image serves as an input. The input video or image undergoes Key Frame Extraction process to identify and extract a key frame to be used. The model performs face detection and crop as well as resize the detected face for further processing. The resized face is submitted to the Convolutional Neural Network (CNN) as presented in Figure 4.

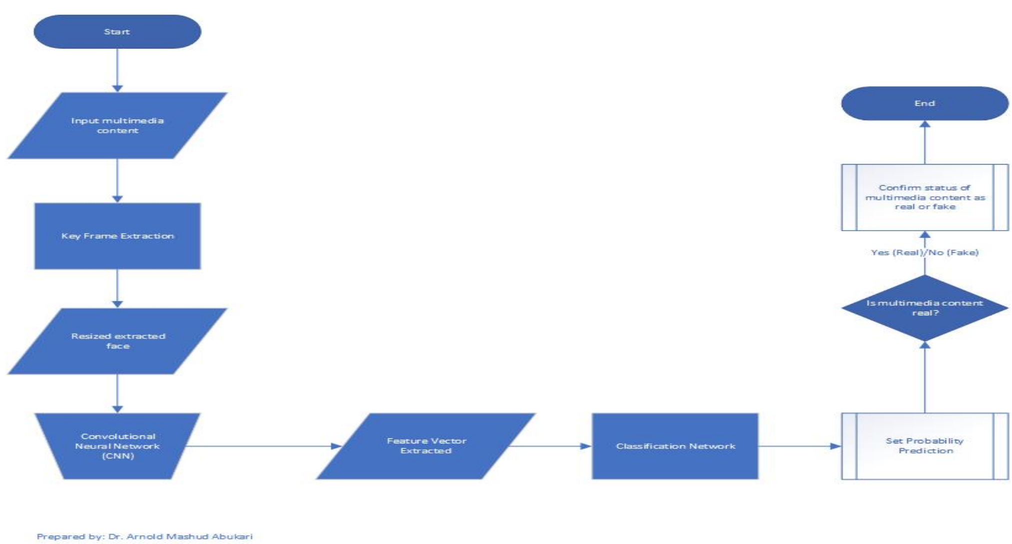

Figure 4: Proposed Model Framework.

The Convolutional Neural Network (CNN) in the proposed model is responsible for Feature Vector Extraction of the video or image. The feature vector extracted is sent to the Classification Network to enable the proposed model classify the extracted feature as real or fake video or image (multimedia) content. The real Probability of prediction and Fake Probability of predictions are set and applied. A comparison of the extracted key frame of the video or image is compared to a dataset and the Real probability of prediction and the fake probability of prediction are used to determine the status of a video or image as either real or fake.

Conclusion

The proposed model presented in this research paper reduces computational resources and complexities in determining and analyzing deepfake multimedia contents like videos and images on social media. Considering a time complexity of $O(n)$ where n is the number of frames, the proposed model takes lesser time since only the extracted key frame is compared instead of comparing to every single frame of the video. The reduction in the computational and time complexities signifies that the proposed model is ideal for low computational devices like smartphones with social media applications. Deepfake multimedia contents like videos and images on social media are a growing concern and more research work is required to reduce or eradicate deepfake multimedia contents. The ability of the proposed model to function and perform on a compressed video makes it very suitable for social media platforms since multimedia contents like videos and images uploaded on the social media platforms like Facebook, X (Twitter), Instagram, TikTok and many others are compressed videos or images. There may be several other approaches to address the concerns brought about by deepfake multimedia contents like videos and images. Eye blinking frequency, inconsistencies in synching audio with the mouth, unnatural key features like hairstyle, extremely smooth skin and many others could be investigated further. This research paper recommends further research work on the implementation of the proposed model presented and other innovative approach in combating deepfake multimedia contents like videos and images on social media.

References

[1] Rafique, R., Gantassi, R., Amin, R., et al. (2023). Deep fake detection and classification using error-level analysis and deep learning. *Scientific Reports*, *13*, 7422. https://doi.org/10.1038/s41598-023-34629-3

[2] Mitra, A., Mohanty, S. P., Corcoran, P., et al. (2021). A machine learning based approach for deepfake detection in social media through key video frame extraction. *SN Computer Science*, *2*(98). https://doi.org/10.1007/s42979-021-00495-x

[3] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*) (pp. 8107-8116).

[4] Chou, H. T. G., & Edge, N. (2012). They are happier and having better lives than I am: The impact of using Facebook on perceptions of others' lives. *Cyberpsychology, Behavior, and Social Networking*, *15*(2), 117-121.

[5] Li, Y., Chang, M., & Lyu, S. (2018). In ictu oculi: Exposing AI generated fake face videos by detecting eye blinking. arXiv preprint: arXiv:1806.02877.

[6] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672-2680).

[7] Denton, E. L., Chintala, S., Fergus, R., et al. (2015). Deep generative image models using a Laplacian pyramid of adversarial networks. In *Advances in Neural Information Processing Systems* (pp. 1486-1494).

[8] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint: arXiv:1511.06434.

[9] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. arXiv preprint: arXiv:1703.06352.

[10] Taigman, Y., Polyak, A., & Wolf, L. (2016). Unsupervised cross-domain image generation. arXiv preprint: arXiv:1611.02200.

[11] Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., & Webb, R. (2017). Learning from simulated and unsupervised images through adversarial training. In *CVPR* (Vol. 3, No. 4, p. 6).

[12] Liu, M.-Y., Breuel, T., & Kautz, J. (2017). Unsupervised image-to-image translation networks. In *NIPS* (pp. 700-708).

[13] Garrido, P., Valgaerts, L., Rehmsen, O., Thormaehlen, T., Perez, P., & Theobalt, C. (2014). Automatic face reenactment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4217-4224).

[14] Suwajanakorn, S., Seitz, S. M., & Kemelmacher-Shlizerman, I. (2017). Synthesizing Obama: Learning lip sync from audio. *ACM Transactions on Graphics* (*TOG*), *36*(4).

[15] Chesney, R., & Citron, D. (2018). Deepfakes: A looming crisis for national security, democracy, and privacy? [Blog post]. Lawfare Blog. Retrieved from https://www.lawfareblog.com/deepfakes-looming-crisis-national-security-democracy-and-privacy

[16] Mitra, A., Mohanty, S. P., Corcoran, P., & Kougianos, E. (2020). A novel machine learning-based method for deepfake video detection in social media. In *Proceedings of the 6th IEEE International Symposium on Smart Electronic Systems* (*iSES*) (In Press).

[17] Güera, D., & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. In *Proceedings of the 15th IEEE International Conference on Advanced Video and Signal Based Surveillance* (pp. 1-6).

[18] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). MesoNet: A compact facial video forgery detection network. In *Proceedings of IEEE International Workshop on Information Forensics and Security* (*WIFS*) (pp. 1-7).

[19] Nguyen, H. H., Yamagishi, J., & Echizen, I. (2019). Capsule-Forensics: Using capsule networks to detect forged images and videos. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*) (pp. 2307-2311).

[20] Hashmi, M. F., Ashish, B. K. K., Keskar, A. G., Bokde, N. D., Yoon, J. H., & Geem, Z. W. (2020). An exploratory analysis on visual counterfeits using conv-lstm hybrid architecture. *IEEE Access*, *8*, 101293-101308.

[21] Kumar, A., Bhavsar, A., & Verma, R. (2020). Detecting deepfakes with metric learning. In *2020 8th International Workshop on Biometrics and Forensics* (*IWBF*) (pp. 1-6). IEEE.

[22] Tayyab, M., Marjani, M., Jhanjhi, N. Z., Hashem, I. A. T., Usmani, R. S. A., & Qamar, F. (2023). A comprehensive review on deep learning algorithms: Security and privacy issues. *Computers & Security*, *131*, 103297.

[23] Zuo, C., Qian, J., Feng, S., et al. (2022). Deep learning in optical metrology: a review. *Light: Science & Applications*, *11*(1), 39. https://doi.org/10.1038/s41377-022-00714-x

[24] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386-408.

[25] LeCun, Y., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*(4), 541-551.

[26] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527-1554.

[27] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (pp. 1097-1105).

[28] Xu, K., et al. (2018). How powerful are graph neural networks? In *Proceedings of the 7th International Conference on Learning Representations* (ICLR).

[29] Zhou, J., et al. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57-81.

[30] Yang, X., Zhang, Z., Zhu, W., et al. (2020). Deterministic conversion rule for CNNs to efficient spiking convolutional neural networks. *Science China Information Sciences*, *63*(12), 122402. https://doi.org/10.1007/s11432-019-1468-0

[31] Sarker, I. H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications, and research directions. *SN Computer Science*, *2*, 420. https://doi.org/10.1007/s42979-021-00815-1

[32] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(1), 1929-1958

[33] Schoenherr, J. R. (2023, April 12). The disturbing trend of state media use of deepfakes. *The Conversation*. https://theconversation.com/the-disturbing-trend-of-state-media-use-of-deepfakes-203368

[34] Miller, J. R. (2022, March 17). *Deepfake video of Zelensky telling Ukrainians to surrender removed from social platforms*. New York Post. https://nypost.com/2022/03/17/deepfake-video-shows-volodymyr-zelensky-telling-ukrainians-to-surrender/

[35] Shahid, F., Kamath, S., Sidotam, A., Jiang, V., Batino, A., & Vashistha, A. (2022). "It Matches My Worldview": Examining Perceptions and Attitudes Around Fake Videos. In *CHI Conference on Human Factors in Computing Systems* (CHI '22) (pp. 1-15). ACM. https://doi.org/10.1145/3491102.3517646

[36] Vota, W. (2022, August 25). *Deepfake videos deeply fool Indian social media users*. ICTworks. https://www.ictworks.org/deepfake-videos-deeply-fool-indian-social-media-users/

[37] Satariano, A., & Mozur, P. (2023, February 7). *The people onscreen are fake. The disinformation is real*. The New York Times. https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html

[38] Hasan, H. R., & Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. *IEEE Access*, *7*, 41596-41606.

[39] Vedaldi, A., & Lenc, K. (2015). MatConvNet - Convolutional Neural Networks for MATLAB. arXiv preprint: arXiv:1412.4564. https://doi.org/10.48550/arXiv.1412.4564