The Deepfake Challenges and Deepfake Video Detection

Worku Muluye Wubet

Abstract: Deepfake is a combination of fake and deep-learning technology. Deep learning is the function of artificial intelligence that can be used to create and detect deepfakes. Deepfakes are created using generative adversarial networks, in which two machine learning models exit. One model trains on a dataset and then creates the deepfakes, and the other model tries to detect the deepfakes. The forger creates deepfakes until the other model can't detect the deepfakes. Deepfakes creating fake videos, images, news, and terrorist events. When deepfake videos, and images increase on social media people will ignore to trust the truth. Deepfakes are increasingly affecting individuals, communities, organizations, security, religions, and democracy. This paper aims to investigate deepfake challenges, and to detect deepfake videos by using eye blinking. Deepfake detections are methods to detect real or deepfake images and videos on social media. Deepfake detection techniques are needed original and fake images or video datasets to train the detection models. In this study, first discussed deepfake technology and its challenges, then identified available video datasets. Following, convolutional neural networks to classify the eye states and long short term memory for sequence learning has been used. Furthermore, the eye aspect ratio was used to calculate the height and width of open and closed eyes and to detect the blinking intervals. The model trained on UADFV dataset to detect fake and real video by using eye blinking and detects 18.4 eye blinks per minute on the real videos and 4.28 eye blinks per minute on fake videos. The overall detection accuracy on real and fake videos was 93.23% and 98.30% respectively. In the future research and development needs more scalable, accurate, reliable and cross-platform deepfake detection techniques.

Keywords: Deepfake, deepfake detection, deep learning, detection techniques, eye blinking

I. INTRODUCTION

Photos and videos are frequently used as evidence in police investigations to resolve legal cases since they are considered to be reliable sources. However, sophisticated technology increases the development of fake videos, and photos that have potentially made these pieces of evidence unreliable [1]. Fake videos and images created by deepfake techniques have been become a great public issue recently [2]. The authors in [3] define the term deepfake as it is a deep learning-based method to create deepfake images or videos by altering the face or full-body of a person in an image or video by the face or full-body of another person. Deep learning is the arrangement of algorithms that can learn the dataset and make intelligent decisions on their own [4].

Generative Adversarial Networks (GANs) [5] is the recent advanced image and video manipulating tool to create high quality manipulated deepfake videos and images, and the media increases the fast distribution of these fake images and videos [6]. The GAN models were trained using a large number of images or videos, it can generate realistic faces or full-body that can be seamlessly spliced into the original video, and the generated video can lead to forgery of the subject's identity in the video [7]. Deepfake manipulation allows a user to replace the face or the full-body of a person in a video with the face or the full-body of another person, provided that enough images may be a large number of images are available of both persons; these videos are called deepfake videos [8]. The authors in [3] state that by using the combination of GANs and Convolutional Neural Networks (CNNs) can create quality deepfakes that the deepfake detection techniques can't detect them.

The existence of, open software mobile applications increasing to everyone to generate fake videos and images [3]. The smartphone availability, advancement of cameras, and social media popularity have made the editing, creation, and dissemination of images and videos more than ever. This increases the tampering of videos and makes effective to propagate falsified information [7]. To detect deepfakes, various detection methods have been proposed after deepfakes were introduced. Deepfake detections are methods to detect real and fake images or videos. The detection methods detect the deepfakes by eye blinking, eye teach and facial texture, head poses, face warping artifacts, eye color, lip movements, audio speakers, reflections in the teeth, spatiotemporal features and capsule forensics [9]-[13]. In this study, investigate deepfakes, deepfake manipulation tools, available datasets, deepfake challenges, deepfake detection challenges, and deepfake detection techniques, deepfake detection by using eye blinking. Finally, this study presents eye blinking detection accuracy and overall detection accuracy results.

II. RELATED WORKS

The authors in [14] present a preview of deepfake detection challenges public a dataset and feature two facial modification algorithms. In [15] states the challenges and opportunities of fake news and detecting fake news by proposing algorithms which detect fake news form the web services. This survey [16], provides a comprehensive review of the recent developments on deep face recognition, that covers databases and protocols, algorithm designs, and application scenes. This paper [3] presents a review of face image manipulation techniques, deepfake methods, and methods to detect manipulations. Deepfake, a deep learning-based technology,

to alter images and videos. In most cases images and videos are used as evidence in investigations and court; but, deepfake, have potentially made these pieces of evidence unreliable.

This research examines its origin and history while assessing how deepfake photos videos are created and present deepfake impacts on society. Different deepfakes detection methods have been developed such as face detection, watermarking, multimedia forensics, and convolutional neural networks (CNNs). Each method uses machine learning and deep learning, to detect any kind of manipulation in photos and videos [1], [6], [9].

In [1] indicated a deepfake detection technique by using multimedia forensics to detects the detailed history of images and determines whether images have been altered. As [36], [38] SVM, RF, and MLP are extracting differentiated features from image datasets and forward these features into SVM, RF, and MLP detection techniques for binary classification. SVM extracts head poses and classifies the extracted features of videos/images. The authors in [9] have used CNN with CFFN to detect fake videos. CFFN performs two-phase feature extraction in an image dataset via Siamese network architecture. CNN [10], [19] performs classification for face tampering in videos generated through deepfake and Face2Face apps [10]. It used to detect any minor modifications that have been made to an image or video. CNN models [37] are learned on a dataset to detect resolution inconsistency between the changed and the nearby face area of video datasets. The authors in [37] have used the CNN detection technique to detect face warping artifacts from deepfake videos. As the experimental result shows ResNet50 has the best performance than other models [8].

III. DEEPFAKES

Deepfakes are created using Generative Adversarial Networks (GANs), in which two machine learning models exit. One model train on a dataset and then creates video forgeries, while the other model attempts to detect the forgeries. The forger creates fakes until the other model can't detect the forgery. The larger the training dataset, the easier it is for the forger to create a believable deepfakes [13], [17].

The first deepfake was created by using an autoencoder-decoder structure. The autoencoder extracts hidden features of face images and the decoder reconstruct the face images. To manipulate faces images, two encoder-decoder pairs train on a dataset [3], [13]. This method allows the encoder to encode and learn the image datasets. In the deepfake creation process, the person one face is linked with the decoder person two faces to reconstruct person two faces from the original person one face. This method is used in different works such as DFaker, DeepFaceLab, and TensorFlow-based deepfakes [13].

According to [13],[17]-[20] FaceSwap, DeepFaceLab, Faceswap-GAN, DFaker, DeepFake-tf, Face2Face, and NeuralTextures are deepfake creation tools. FaceSwap is a graphics-based method to alter the face image from a source video to a target video by using two encoder-decoder pairs. DeepFaceLab is expanding from the Faceswap model with new models and Support multiple face extraction modes. DFaker reconstructs the face and implemented based on the Keras library. Also, the DeepFake-tf is the same as with DFaker but the DeepFake-tf applied based on the TensorFlow library. The NeuralTextures with rendering network uses the video dataset to learn the texture of the target person [13], [14].
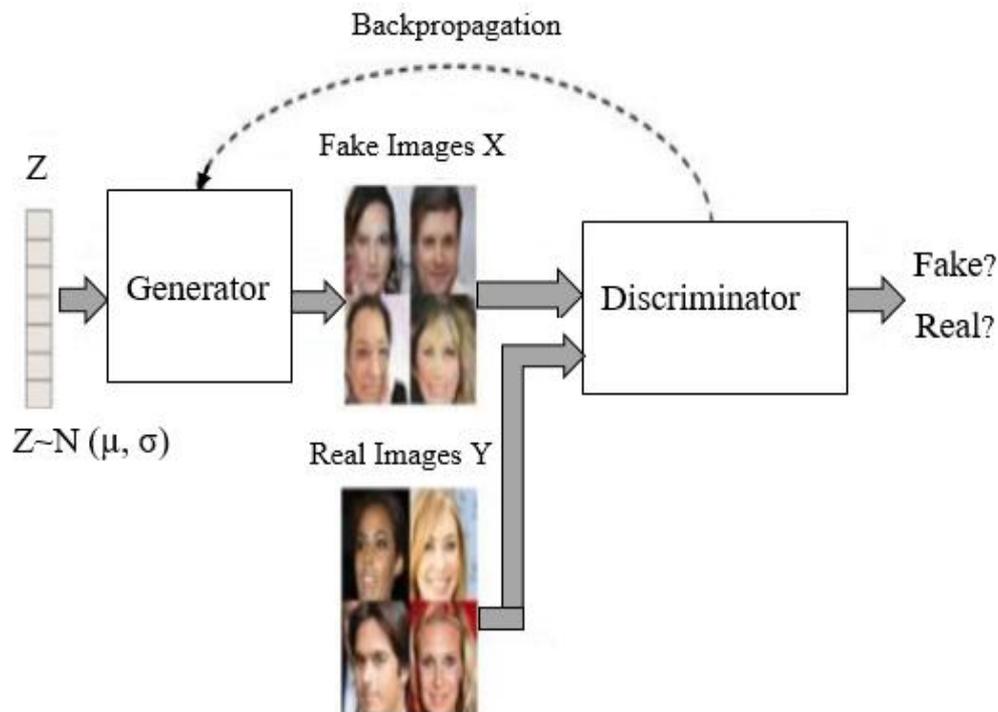

Fig. 1. Generative adversarial network architecture [12].

A. Deepfake Datasets

A dataset plays a crucial role to train the deepfake detection models. Regarding this, different detection techniques are needed well-organized in deepfake image and video contents. Currently, deepfake detection techniques require a huge amount of training dataset and need to be evaluated. As such, there is an increasing need for a huge amount of deepfake image and video datasets to evaluate the performance of detection techniques [21].

Table-I: List of deepfake image and video datasets [6], [13], [14], [18], [21], [22].

Datasets	Total videos	Description
UADFV [13], [21]	98	This dataset has 49 fake and 49 real videos and these videos contain 32,752 frames. These videos are created using the DNN model with FakeAPP.
DeepfakeTIMIT [14], [21], [23]	320	It contains low and high-quality videos with a total of 10537 real and 34,023 fake images extracted from 320 total videos. It is created by faceswap-GAN.
VTD [18]	1000	Video Tampering Dataset is a large, manipulated video dataset.
FaceForensics	2008	It is created from 1004 video datasets collected from YouTube. The video was altered by using the Face2Face approach.
DFD [21]	3068	The Google or Jigsaw deepfake detection dataset deepfake videos generated based on 3068 real videos and 28 consented individuals of different ages, genders, and ethnic groups.
DeepFakeDetection [23]	3363	It contains 363 original sequences and more than 3000 altered videos by using deepfakes and their corresponding binary masks.
FaceForensics++ [22]	5000	It is an extended dataset from FaceForensics which contains real and fake videos generated using faceswap.
DFDC [14]	5214	The Facebook deepfake detection challenge has 4113 deepfake videos created based on 1131 real videos of 66 consented individuals of different ages, genders, and ethnic groups.
Celeb-DF [21]	6229	It contained 590 real videos and 5639 deepfake videos that corresponding to above 2 million video frames.

B. Deepfake Challenges

Deepfakes are affecting the world since people around the world are using deepfakes for multiple reasons such as face-swapping, recreating pornographic videos with someone's face or body, and to create and disseminate fake news [1]. Deepfakes are more and more affecting democracy, privacy, security, religion and cultures of the people. Deepfakes are increasing from time to time, but there is no standard to evaluate deepfake detection techniques. The number of deepfake videos and images found online has nearly doubled since 2018. Massachusetts Institute of Technology (MIT) analyzed 126,000 news disseminated by 3,000,000 users for more than 10 years. Finally, they concluded that fake news spreads 1,500 people 6 times more rapidly than true news. Deepfakes creating fake news, images, videos, and terrorist events. Deepfake erodes people's trust in media and causes to social and financial fraud. Deepfake affects religions, organizations, politicians, artists, and voters. When deepfake videos and images increase in social media people will ignore to trust the truth [4], [24]-[27].

The authors in [1], [24] analyze deepfakes that have the potential to harm individuals and societies. Using deepfakes to harm other people are yet to be largely seen including joke to embarrass a coworker, identity theft or even to spur violence, a porn video for someone's gratification and so on. Also, deepfakes are used to fake terrorism events, blackmail, defame individuals, and to create political distress. Although nobody is safe from deepfakes, some people are more vulnerable than others [13], [24]. With minimum data and computing power, somebody can create a video the country leader saying something leading to civil conflict [26]-[28].

Deepfakes negatively affects targeted person, increase fake news and hate speech, create political tension, distress the public or create war. For example, a person can modify the contents of the video and people in a video to spread fake news, which may lead to war between nations; especially a country that contains diverse nations and nationalities. Deepfakes creation is increasing and the social media disseminate those deepfake images and videos quickly [2], [13], [17], [22],[31]. s

IV. DEEPFAKE DETECTION

To detect deepfakes, various methods have been proposed after this threat was introduced. In a binary classification deepfake detection technique, the classifiers, classify the manipulated and real videos. This type of detection technique needs a huge dataset of real and fake videos to train the machine. A lot of deepfake videos are available online on the Internet, but it is still limited to set a standard to evaluate different deepfake detection techniques. So, the authors in [6] produced a deepfake dataset that contains 620 videos created by using Faceswap-GAN [13], [29].

VidTIMIT publicly available database contains both low and high-quality deepfake videos, which can accurately impersonate the lip movements, facial fallings, and eye blinking of the person. These dataset videos were then used to test various deepfake detection methods. The result shows that the Facenet and VGG face recognition systems are unable to detect deepfakes effectively. The image quality metrics and the lip-syncing method with Support Vector Machine (SVM) show an error when trying to detect deepfake videos. This increases the critical need for more effective and efficient deepfake detection algorithms [13].

The reviewed articles suggested that there are ways to combat deepfakes such as using existing laws, legislation and regulation, additional action from social media companies, corporate policies and voluntary action, education, and training, deploying digital literacy curriculum in schools, improve media literacy, and anti-deepfake technology, content authentication, a deepfake prevention and develop deepfake detection methods [25], [27], [30].

The authors in [18] present a forgery detection approaches the first one human observer and the second approach is automatic forgery detection such as based on stage analysis features and learned features.

A. Deepfake Video Detection

The authors in [18] aimed to detect fakes of facial images by using different face tracking methods. The DARPA accelerating the development of fake digital visual media detection methods [13].

The swapped face images made by CNN and GAN deep learning are more challenging for forensics models to detect the forgeries [13]. The authors in [11] proposed a multitask learning method to perform classification and segmentation of manipulated facial images simultaneously.

In the detection process, the temporal consistency of the video is not imposed efficiently and enforced to use the Spatiotemporal contents of the video to detect deepfakes [13]. The integration of conventional networks and recurrent unit manipulate the temporal inconsistencies of the frames [32]. Using both Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) able to extract temporal features of the video, that is denoted through the sequence of frames. The deepfake detection network takes the sequence frames and estimates the possibilities of the frame sequence that can be either original or deepfake. CNN is used to extract features of the frame and then fed them into the LSTM to generate temporal sequence frames [3], [13], [33], [34].

Using eye blinking, we can detect deepfakes in which a person in deepfakes has fewer or higher eye blinking rate than in real videos. A normal person's eye will blink between 2 to 10 seconds, and each eye blink will take between 0.1 and 0.4 seconds. In most cases, deepfake creation tools can't create eyes that can blink like a normal person. Most of the time eye blinking speed is manipulated videos are slower than in real videos. The color of each eye is extracted using computer vision and the difference in eye color is also used to detect deepfakes [1]. The deepfake detection technique uses the audio-visual dataset to detect inconsistencies between lip-movements, speech, and numerous differences of slow images. The deepfake detection technique through lip-synching differentiates the real videos from altered videos, where the lip-movement and the speech are matched [35], [36].

B. Deepfake Detection Techniques

Deepfake detection techniques are methods to detect fake videos and images. Deepfake detection techniques are grouped into fake image and video detection techniques. Because if we are unable to detect fake images and videos, we may soon be forced to distrust everything we hear and see. The common deepfake detection methods such as Convolutional Neural Network (CNN) to extract frame feature, LRCN to capture the eye blinking temporal patterns, Recurrent Neural Network(RNN) to discover temporal discrepancies across frames and Long Short Term Memory (LSTM) for temporal sequence analysis [1].

Long-Term Recurrent Convolutional Networks (LRCN) [7] dynamically predicts eye area sequences. LRCN captures temporal patterns of eye blinking in the videos since eye blinking frequency of deepfakes is slower or extremely faster than normal videos. Also, it contains feature extractor that extracts depend on CNN, a sequence learning through LSTM, and to predict the possibility of open and closed eye state. The eye blinking has temporal patterns and the LSTM can capture these temporal patterns effectively. The eye blinking rate is estimated as a peak above the threshold of 0.5 with a duration of fewer than 7 frames. The extremely frequent eye blinking may also be the criteria of altered videos [13]. CNN and LSTM [33] CNN detects frame features then disseminates to LSTM and LSTM detect temporal inconsistencies in videos to set frame sequences for classification.

Recurrent Convolutional Neural Networks (RCN) [39] discovered temporal inconsistencies across the frame. It uses Spatiotemporal features of videos. Deep recurrent network models mostly use temporal patterns across video frames to differentiate deepfake videos. The other approach that fragment videos into frames and identify visual objects inside single frames to get discriminant features. These features are then disseminated into a low or high classifier to detect fake and real videos. When deepfake videos are generated with low resolutions are needs face rotation, zooming, and cutting to match with real ones. The resolution discrepancy between the nearby and the changing face area can be detected by CNN models [3], [13].

The capsule network is a deepfake detection technique. The capsule network classifies the hidden features and dynamic routes the 3 capsules output to 2 capsules output, one output for fake and the other output for real images/videos. It deals with videos/images dataset. Logistic regression [8] and neural networks are discovered facial texture and missing reflections of eye and teeth areas and classifying videos.

The deepfake detection methods detect the deepfake videos and images by eye teach and facial texture, head poses, eye blinking, lip-synching, face warping artifacts, physiological signal, reflections in the teeth, spatiotemporal features and capsule forensics [1], [8], [13].

V. METHODS

In this study, deepfake videos have been detected based on eye blinking by using UADFV datasets. First, the publically available datasets that are appropriate for detection techniques have been compared. Second, deepfake detection techniques have been compared and CNN and LSTM selected that are suitable to detect deepfakes using eye blinking speed.In the method section, describe the way to detect fake videos via eye blinking. Eye blinking is the quick opening and closing movement of the eyelid. First focuses on the face, next detects the eyes, and then determines the eye blinking rate from the video. Finally, present the eye blinking and overall detection accuracy results. In order to detect the eye area in each video frame, the eye aspect ratio between height and width of the eye is calculated.
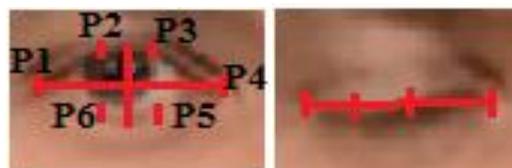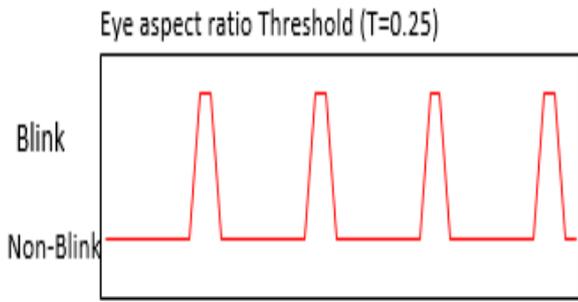
Fig.2: Eye aspect ratio and threshold values.

$$\text{Aspect Ratio} = \frac{||p2-p6||+||p3-p5||}{2||p1-p4||} \qquad (1)$$

where p1, . . ., p6 are the two-dimensional eye area and p2, p3, p5, and p6 measures the height while p1 and p4 measure the eyes width [40]. It decides whether the eyes are in a closed or open state. The distance between the height and width of the eye is calculated. In the experiment 0.25 is a threshold and when the eye aspect ratio smaller than 0.25 or close to zero considered as a closed eye and higher than 0.25 is an open eye. In this study, the normal person eye blinking rate is used as a threshold to detect and count the eye blink and blink intervals. Since a normal person's eye will blink between 2 to 10 seconds, and each eye blink will take between 0.1 and 0.4 seconds [1]. In most cases, the eye blinking rate in fake videos slower or extremely faster than the normal person blinking rate. This enables to detect fake video from real videos by using eye blinking rate.
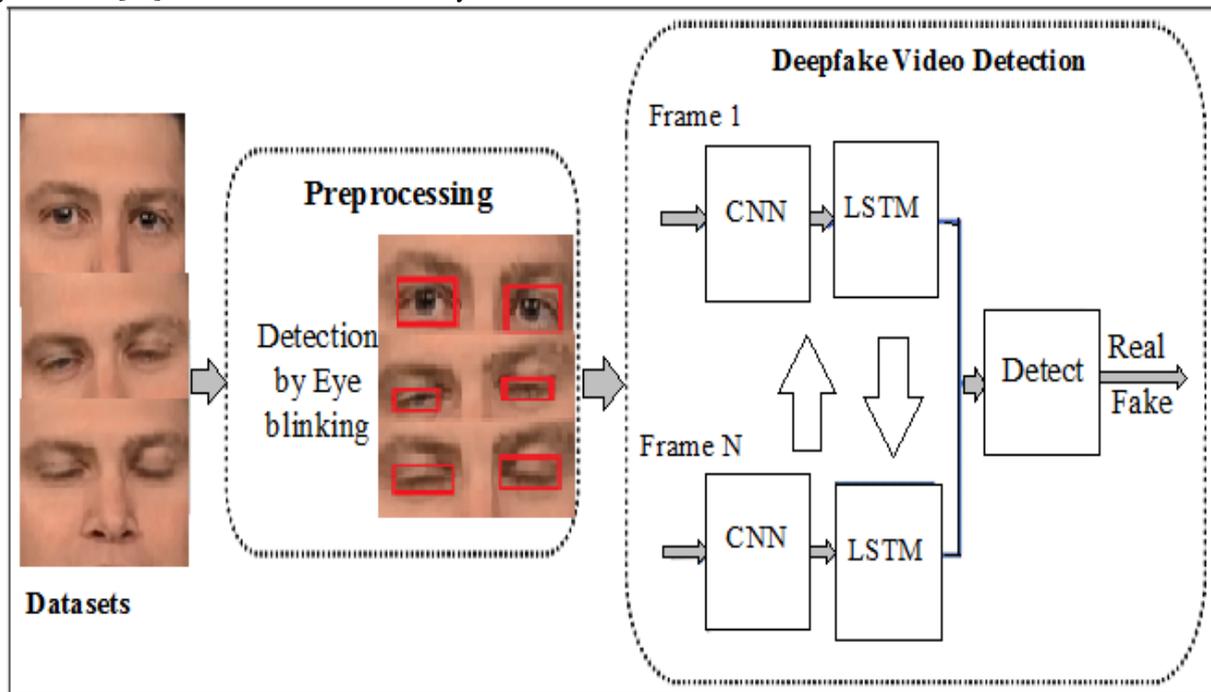

Fig. 3. Detects fake videos based on eye blinking with CNN to classify the eye states and LSTM for sequence learning.

The CNN classifies the eye states and the eye aspect ratio classifies the eye blinking rate. Then CNN and eye aspect ratio detect the eye blinking intervals.

A. Model Training

The VGG16 and ResNet-50 based CNN model was trained on a training dataset that contains open and closed eyes regions. The VGG16 and ResNet-50 based CNN model was trained on a training dataset that contains open and closed eyes regions. The MRL [41] is an eye image dataset that contains approximately 15000 closed and open eye images to train the models. In the training 100 closed and open eye images were used to train the models and 98 fake and real videos as the training set for the model.

B. Results and Discussion

The deepfake video detection using CNN and LSTM based on eye blinking rate was tested on UADFV publicly available dataset.

UADFV has 98 videos of which 49 fake and 49 real videos and these videos contain 32,752 frame sequences. In the dataset, each videos length ranges between 2 to 44 seconds and the average length is 11.26 seconds. As illustrated in the table below Table-II, in the experiment UADFV publicly available 49 real video datasets contains 16,376 frame sequences.

The trained model detects 18.4 eye blinks per minute from the real 49 videos. The normal person eye blinking rate was used as a threshold so that the proposed method detects the video as a real video that exists within the threshold.

The trained model detects 4.28 eye blinks per minute on 49 fake videos which are slower than the threshold and the trained model detects the fake videos that exists above the threshold.

Retrieval Number: E2779039520/2020©BEIESP
DOI: 10.35940/ijitee.E2779.049620
Journal Website: www.ijitee.org

793

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table-II: Detection performance on UADFV video dataset.

Video Dataset	Number of Videos	Average video length	Frames Per Second(FPS)	Eye blinking rate
Real Videos	49	11.26 seconds	28 FPS	14.42/minute
Fake Videos	49	11.26 seconds	28 FPS	4.28/minute

The eye blink detection accuracy and the overall fake video detection accuracy is computed by using (2) and (3) equations, respectively.

$$\text{Blinking Detection Accuracy} = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$\text{Total Detection Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \quad (3)$$

Where TP (True Positive), FN (False Negative), FP (False Positive), and TN (True Negative) are the total number of frames that are detected as eye blinks, shows eye blinks but not detected; detected as eye blinks but they are not blinks and non-blinks detected as no blinks respectively [42].

Table-III: Eye blinking detection accuracy results on real and fake videos.

Criteria	Blink [True]	Not Blink [False]
Number of frames detected as blink on real videos [Positive]	2374 [TP]	218 [FP]
Number of frames detected as not blink on real videos [Negative]	12894 [TN]	890 [FN]
Number of frames detected as blink on fake videos [Positive]	946 [TP]	102 [FP]
Number of frames detected as not blink on fake videos [Negative]	15152 [TN]	176 [FN]

In the experiment, the eye blinking detection accuracy result on real videos is 91.59% and eye blinking detection accuracy on fake videos 90.27%. Furthermore, the overall detection accuracy on real videos is 93.23% and the overall detection accuracy of fake videos is 98.30%. In the eye blinking detection, when the person moves his/her head quickly and when the eye focus on the area below them the eyelids cover the eye and the eye detected as blink or closed this affects the accuracy of the model.

C. Deepfake Detection Challenges

Now a day different deepfake detection techniques have been introduced and evaluated however using unorganized datasets. Improving the performance of detection techniques, growing standardized datasets and evaluating the deepfake detection techniques will simplify the development of accurate and effective deep learning-based detection techniques, which requires a standard dataset [1], [3], [14].

The deepfakes' quality has been improving and the performance of detection algorithms requires to be enhanced consequently. Developing better detection methods to combat fakes generated by deepfakes because artificial intelligence that power deepfakes are rapidly developing, making deepfakes so hard to detect. Also, currently, detection techniques typically emphasize the weaknesses of the deepfake creation tools. This approach is not always functional to detect deepfakes created by all deepfake creation tools. This is a challenge for deepfake detection techniques development and in future research requires an emphasis on proposing scalable, accurate, reliable and cross-platform detection techniques [14].

Facebook has prepared a deepfake detection competition with the partnership on Artificial Intelligence, Amazon, Microsoft, and with some Universities. This competition aims to invite researchers around the world to develop better detection techniques to detect deepfakes. This helps to build better detection techniques to combat fakes generated by deepfakes because deep learning that power deepfakes are rapidly developing and making so hard to detect [14].

VI. CONCLUSION

Deep learning can be used as a deepfake creation, and detection methods. Deepfake creates forged images or videos that persons cannot differentiate from real images or videos. Deepfakes are created using generative adversarial networks, in which two machine learning models exit. One model trains on a dataset and the other model tries to detect the deepfakes. The forger creates fakes until the other model can't detect the forgery. Deepfakes creating fake news, videos, images, and terrorism events that can cause social, and financial fraud. It is increasing affects religions, organizations, individuals and communities', culture, security, and democracy. When deepfake videos and images increase on social media people will ignore to trust the truth. In this study, the available datasets, deepfake creation tools, deepfake challenges, fake video detection techniques and detect fake video by using eye blinking were discussed. Also, the detection models trained on the datasets and the total and the eye-blink detection accuracy results were computed. Deepfake detection is a method to detect real and fake images or videos. In this study, the CNN to extract frame feature and to classify the eye states, and LSTM for temporal sequence analysis have been used. Also, the eye aspect ratio, used for eye blinking rate classification and the CNN and eye aspect ratio detect the eye blinking intervals. The detection models have been trained on UADFV publically available real and fake videos. The deepfake detection methods detect the deepfakes by eye blinking. In the experiment, the eye blinking detection accuracy result on real videos is 91.59% and eye blinking detection accuracy on fake videos 90.27%. Furthermore, the overall detection accuracy results on real videos is 93.23% and the overall detection accuracy on fake videos is 98.30%. In the eye blinking detection, when the person moves his/her head quickly and when the eye focus on the area below them the eyelids cover the eye and the eye detected as blink or closed this affects the accuracy of the model.

Now a day deepfake creation tools can create fake videos by mimic facial expressions of the person exactly so that it is become difficult to detect deepfakes by using facial expressions like eye blinking, and lip-movement. Therefore, both image and video deepfake detection techniques are needed performance improvement, evaluation standards, and parameters.

Future work will focus on evaluating different detection methods by using real and manipulated datasets. Due to the advancement of technology full-body deepfakes are released. The continuous advancements of the face and full-body deepfakes development will be difficult to detect by the existing detection techniques. So, deepfake datasets and cross-platform detection techniques need to be developed in the future. Furthermore, due to the high computational cost, most detection techniques are unfit for mobile applications. This needs efficient, reliable and robust mobile detectors to detect deepfakes in widely used mobile devices. Moreover, will improve deepfake detection by integrating deepfake detection and object detection algorithms.

REFERENCES

1   J. A. Marwan Albahar, Deepfakes Threats and Countermeasures Systematic Review, JTAIT, vol. 97, no. 22, pp. 3242-3250, 30 November 2019.

2   D. D. Zahid Akhtar, A Comparative Evaluation of Local Feature Descriptors for DeepFakes Detection, 07 November 2019.

3   R. V.-R. J. F. A. M. J. O.-G. Ruben Tolosana, DeepFakes and Beyond A Survey of Face Manipulation and Fake Detection, pp. 1-15, 2020.

4   G. Shao, What 'deepfakes' are and how they may be dangerous, 13 October 2019. [Online]. Available: https://deepfakedetectionchallenge.ai/.

5   B. P. W. W. J. D. Xinsheng Xuan, On the Generalization of Generative Adversarial Networks Image Forensics, pp. 1-8, 2019.

6   S. M. Pavel Korshunov, Vulnerability assessment and detection of Deepfake videos, pp. 1-6, 2019.

7   M.-C. C. a. S. L. Yuezun Li, In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking, pp. 1-7, 2018.

8   A. M. R. Z. G. Marissa Koopman, Detection of Deepfake Video Manipulation, in Proceedings of the 20th Irish Machine Vision and Image Processing Conference, Belfast, 2018.

9   Y.-X. Z. a. C.-Y. L. Chih-Chung Hsu, Deep Fake Image Detection Based on Pairwise Learning, Applied science, pp. 1-14, 2020.

10   V. N. J. Y. I. E. Darius Afchar, MesoNet a Compact Facial Video Forgery Detection Network, pp. 1-7, 2018.

11   F. F. J. Y. I. E. Huy H. Nguyen, Multi-task Learning For Detecting and Segmenting Manipulated Facial Images and Videos, pp. 1-8, 2019.

12   I.-S. N. S.-H. K. Nhu-Tai Do, Forensics Face Detection From GANs Using Convolutional Neural Network, pp. 1-5, 2019.

13   C. M. N. D. T. N. D. T. N. S. N. Thanh Thi Nguyen, Deep Learning for Deepfakes Creation and Detection, pp. 1-16, 2019.

14   R. H. B. P. N. B. C. C. F. Brian Dolhansky, The Deepfake Detection Challenge (DFDC) Preview Dataset, Deepfake Detection Challenge, pp. 1-4, 2019.

15   L. O. Alvaro Figueira, The current state of fake news: challenges and opportunities, in ICHSCIST, Barcelona, 2017.

16   W. D. Mei Wang, Deep Face Recognition: A Survey, pp. 1-26, 2019.

17   J. D. Britt Paris, Deepfakes and Cheap Fakes the Manipulation of Audio and Visual Evidence, Data and Society's Media Manipulation, Los Angeles, 2018.

18   A. R. D. C. L. V. C. R. J. T. M. Nießner, FaceForensics++: Learning to Detect Manipulated Facial Images, pp. 1-14, 2019.

19   M. V. R. S. Prabhat Kumar, Detecting Face2Face Facial Reenactment in Videos, pp. 1-9, 2020.

20   D. D. B. B. Zahid Akhtara, Face Authenticity An Overview of Face Manipulation Generation, Detection and Recognition, in International Conference on Communication and Information Processing (ICCIP-2019), 2019.

21   X. Y. P. S. H. Q. S. L. Yuezun Li, Celeb-DF (v2): A New Dataset for DeepFake Forensics, pp. 1-6, 2019.

22   D. C. L. V. C. R. J. T. M. N. Andreas Rössler, FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces, Computer Vision and Pattern Recognition, pp. 1-21, 2018.

23   M. K. F.-J. P. J. K. Ricard Durall, Unmasking DeepFakes with simple Features, pp. 1-8, 2019.

24   T. Farago, Deepfakes an Emerging Risk to Individuals and Societies Alike, 2019.

25   H. K. Hall, Deepfake Videos When Seeing Isn't Believing, Catholic University Journal of Law and Technology, vol. 27, no. 1, pp. 51-76, 2018.

26   H. F. Shruti Agarwal, ProtectingWorld Leaders Against DeepFakes, Computer Vision Foundation, pp. 38-45, 2019.

27   D. C. Bobby Chesney, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, California: California Law Review, 2019.

28   R. Chawla, Deepfakes: How a pervert shook the world, International Journal of Advanced Research and Development, vol. 4, no. 6, pp. 4-8, 2019.

29   D. L. Y. Umur Aybars Ciftci, FakeCatcher: Detection of Synthetic Portrait Videos using Biological Signals, pp. 1-14, 2019.

30   Mika Westerlund, The Emergence of Deepfake Technology: A Review, vol. 9, no. 11, pp. 39-52, 2019.

31   Z. R. E. C. L. E. V. O. P. K. M. H. Xinyi Ding, Swapped Face Detection using Deep Learning and Subjective Assessment, pp. 1-8, 2019.

32   J. C. A. J. W. A. I. M. P. N. Ekraam Sabir, Recurrent Convolutional Strategies for Face Manipulation Detection in Videos, pp. 80-87, 2019.

33 E. J. D. David Guera, Deepfake video detection using recurrent neural networks, AVSS, pp. 1-6, 2018.

34 J. Y. I. E. Huy H. Nguyen, Capsule-Forensics Using Capsule Networks to Detect Forged Images snd Videos, pp. 1-5, 2018.

35 S. M. Pavel Korshunov, DeepFakes A New Threat to Face Recognition, Assessment and Detection, pp. 1-5, 2018.

3] L. Z. V. L. L. T. Ying Zhang, Automated face swapping and its detection, International Conference on Signal and Image Processing, pp. 15-19, 2017.

37 S. L. Yuezun Li, Exposing DeepFake Videos By Detecting Face Warping Artifacts, New York, 2019.

38 Y. L. S. L. Xin Yang, Exposing Deepfakes Using Inconsistent Head Poses, ICASSP, pp. 8261- 8265, 2019.

39 E. C. J. J. A. A. W. M. I. N. P. Sabir, Recurrent convolutional strategies for face manipulation detection in videos, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 80-87, 2019.

40 T. S. a. J. Cech, Real-Time Eye Blink Detection using Facial Landmarks, Computer Vision Winter Workshop, pp. 1-8, 2016.

41 R. Fusek, "Pupil localization using geodesic distance," Lecture Notes in Artifical Intelligence, vol. 11241, pp. 433-444, 2018.

42 T. S. Chinnawat Devahasdin Na Ayudhya, A Method for Real-Time Eye Blink Detection and Its Application, pp. 1-6, 2009.

AUTHOR PROFILE

Worku Muluye received his first degree in Information Systems from Wollo University at Kombolcha Institute of Technology, Ethiopia in 2013 and his Master's Degree from the Adama Science and Technology University, Ethiopia in 2017. From August 2013 to October 2014 he was working as a graduate assistant and assistant lecturer at Wolkite University, Wolkite Ethiopia, to which he returned in 2017 as a full lecturer and researcher. He has two publications in reputable journals. His fields of interest cover Software-Defined Networking, Li-Fi, Artificial Intelligence, Natural Language Processing, and Image Processing.