# Deepfake: A Survey on Facial Forgery Technique Using Generative Adversarial Network

Digvijay Yadav
Dept of Master of Computer Applications
Sardar Patel Institute of Technology
Mumbai,India
diggydigvijay8723@gmail.com

Sakina Salmani
Dept of Master of Computer Applications
Sardar Patel Institute of Technology
Mumbai,India
sakina.shaikh30@gmail.com

*Abstract*— **"Deepfake" it is an incipiently emerging face video forgery technique predicated on AI technology which is used for creating the fake video. It takes images and video as source and it coalesces these to make a new video using the generative adversarial network and the output is very convincing. This technique is utilized for generating the unauthentic spurious video and it is capable of making it possible to generate an unauthentic spurious video of authentic people verbally expressing and doing things that they never did by swapping the face of the person in the video. Deepfake can create disputes in countries by influencing their election process by defaming the character of the politician. This technique is now being used for character defamation of celebrities and high-profile politician just by swapping the face with someone else. If it is utilized in unethical ways, this could lead to a serious problem. Someone can use this technique for taking revenge from the person by swapping face in video and then posting it to a social media platform. In this paper, working of Deepfake technique along with how it can swap faces with maximum precision in the video has been presented. Further explained are the different ways through which we can identify if the video is generated by Deepfake and its advantages and drawback have been listed.**

*Keywords—Deepfake, Machine Learning, Deep Learning, Generative Adversarial Network, Neural Network.*

## I. INTRODUCTION

Over the past few years, Technology is advancing rapidly and it is influencing the Cyber world availability and because of this access and availability of social media platform is additionally growing in numbers. Now social media platforms are used for expressing individual view and opinion. Now information about anything is available on the internet in vast majorities and the internet era has provided an opportunity to share information on the web with each other. Now internet can be easily accessed by anyone because of this people are started spreading fake news and rumours by misusing it. Deepfake is one of them and it is the concept of swapping the face of one person to another using a generative adversarial network algorithm. The output of this algorithm is very realistic and authentic to human eyes [1]. Deepfake is the technologies which will make people do things which they never did in the past. Now people are using this technology for spreading the fake news against people by swapping their face in the video. It is also used for character assassination and taking revenge from people. This technique is very potent because its output is very authentic as the algorithm work on the tribulation and error method. The algorithm generates a fake image and then it compares it with the real image just to make sure the output looks much more realistic and it carries out this task until it gets the perfect fake image which looks authentic to others [1,2]. A few years back it was difficult to generate Deepfake but now the developer is making an app which needs only video files and it will create small data set of images and then it will train the algorithm for generating the images. Deepfake is still a new technology and it still has room for more development and this technology is still not known to everyone. Through this paper, we will understand Deepfake in detail and see how harmful it would be if it is used in unethical ways. First generative adversarial network is explained as how it used for swapping face in Deepfake videos. Second, advantages and drawbacks of deep fake are listed in ethical and unethical ways, then different algorithms have been listed and ways through which we can identify Deepfake videos and then limitations of Deepfake has been listed.

## II. METHODOLOGY

### A. Generative Adversarial Networks.

The Generative Adversarial Networks (GAN) are the best algorithm for Deepfake and it combines two neural networks that are capable of creating realistic images. This machine learning technique is capable for learning from the collection of images and then it creates an image which looks authentic to human eyes just by combining these images, for example, with help of GAN we can create realistic images, of animals, clothes design and anything else for which GAN is trained for [3]. The generative adversarial networks combine two neural networks, one of them is called generator and the other one is the discriminator. Generator is the neural network which tries to generate the fake images from the images data set which is provided to it and it tries to generate a realistic images and the discriminator is the other neural network which evaluates the images produced by the generator for authenticity of the images, it helps in generating more realistic images which look real to human eyes. The generator and discriminator are being trained in min-max method [2]. The min represents 0 and max represents 1. The 0 represents fake output and 1 represent the authentic output. Discriminator tries to get closer to 1 in order to create a realistic Deepfake. The generative adversarial network needs a longer time to reach max value for generating better output. If we train the generator and discriminator for a longer duration, then it will help in generating more realistic and authentic Deepfake images and then it will be used for swapping the face of person A with person B in the video.

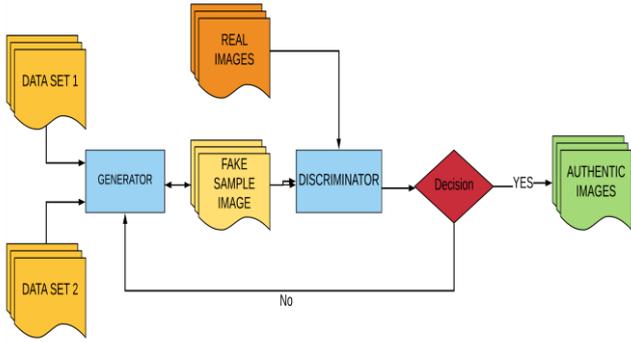## B. How Generative Adversarial Networks works.



**Fig. 1** Face swapping Generative Adversarial Networks architecture for generating Deepfake video.

Generative Adversarial Networks is the most important neural networks and it is part of machine learning. It is a fusion of two neural network generator and discriminator. The generator takes two data sets 1 and 2. These data sets consist of images of two different people. The generator tries to combines these two images to create a fake image and then the image goes to the discriminator. The discriminator is the second neural network which will compare the fake image with the real images for the authenticity of the fake image. If the fake image doesn't look authentic to the discriminator, then it will reject it and then the generator will create another fake image and again it is validated for the authenticity. The discriminator tries to produce accurate output by comparing the fake image with the original images. The algorithm will go in a loop until it finds an authentic fake image for Deepfake. GNA uses Mathematical value function *V(G,D)* **G** represents generator and **D** represent discriminator [2]. The value function is given by equation 1:

$$min_G \ max_D \ V(D,G) = Ex_{\sim\rho(x)}[log \ D(x)] + Ez_{\sim\rho(z)}[log(1-D(G(z)))] \ \dots\dots (1)$$

The generative adversarial network works on min-max method. The value function *V(G,D)* is used for creating authentic images. The discriminator **D** and it remember data from *p(x)*, hence *Ex~ρ(x)[log D(x)]*. Where E represents expectation and the maximization of this equation allow **D** to predict precisely *D(x)=1 when x~ρ(x)*. The next equation is for Generator **G** which tries to convince discriminator. *Ez~ρ(z)[log(1-D(G(z)))]* and if we get *D(G(z))~0* then it means **G** is not able to trick **D**. Training the generator and discriminator is very time consuming and the longer we train the neural networks it will help in generating the authentic output. Once the discriminator gets the correct values after training several days the generator and discriminator will reach a point where both of them cannot improve further. At this point, the generator will start generating realistic Deepfake images and discriminator will not be able to differentiate between the generated Deepfakes [4]. Once the discriminator gets the accurate image it can now be used for swapping the face in the video to create a Deepfake.

## C. Result of Deepfake using generative adversarial network

To generate a perfect and authentic Deepfake we need to train the algorithm for more time and it required good processing power. The better graphic card you use will result in better Deepfake images. We need a high end graphic card

which has lots of CUDA cores for processing image data sets which will help in generating better Deepfake and we need data sets of a high resolution images because these images will be cropped and re-sized according to the face orientation of the target person. This will result in authentic Deepfake for swapping person face in the video and it will look real to human eyes. For generating the result, generative adversarial network is used which will help in creating Deepfake images of the target person and then we will use those Deepfake images to swap the face of the person in the target video. We have collected some image data sets of Henry Cavill and Christopher Reeve, and then we have cropped the face of both person to 256 pixels and removed all the blurred images from the data sets. Then we started the training of neural networks for generation of faces and these faces are getting evaluated by the discriminators.



**Fig. 2** The images are taken while training the Neural Networks and images of Henry Cavill and Christopher Reeve are used as data sets which will be used for generating Deepfake.

The Fig. 2 is taken when the Neural Networks is training for generating Deepfake images. These images are start and end screen of training. In the start, there were only 2 columns and at the end, all columns are filled with Deepfake images which are generated by generator neural network. The first column consists of Christopher Reeve images and the third column consist for Henry Cavill and the other remaining column is for neural networks to train by combining the faces of both actors to generate the realistic Deepfake images and later we will use these images to swap Henry Cavill in Justice League Movie. The graph on the top represents the Learning curves of the generator and the discriminator neural network. To generate realistic Deepfake we have used images with closed and open eyes for creating eye blinking effect and we have used images with faces in different angles.



**Fig. 3** The above images are taken from movie the justice league in the original image the superman is played by Henry Cavill and the second image is a Deepfake image in which the face is swapped with Christopher Reeve who played the role of Superman in late '70s.

The Fig. 3 is the screenshot which is taken after generating Deepfake video of Christopher Reeve. We have replaced Henry Cavill face with Christopher Reeve in the

video which was taken from the Justice League movie which was released in 2017 using the Generative Adversarial Networks. Henry Cavill played the role of Superman in 2017 and Christopher Reeve played the role of Superman in the late '70s. The video looks very authentic and the human eye will not be able to differentiate between the original and the fake video. This output image shows that if we trained the generative model with good image data sets then it will generate authentic images which can be used for swapping the face in the video, it will become very difficult for human eyes to suspect any changes made in the video. The Generative Adversarial Networks is an optimal algorithm for generating authentic Deepfake. The video is nothing but a collection of images which are played at 30 frames per second in order to make a video [5]. So, with the help of Generative Adversarial Networks, we are editing and modifying on each frame, with this we are able to edit the video in such a way so that it will look realistic even after modification. This can be only possible with good data sets and a high-end computer with good processing power [6]. The data sets should be of high-resolution images with a similar facial feature, skin tones and the images should be taken under the same lighting condition. With this, we have created the Deepfake image and swapped actor face in Fig 3.



**Fig. 4** The above images are taken from movie Goldfinger in the original image the role of James Bond is played by Sean Connery and the second image is a Deepfake image in which the face is swapped with Elon Musk.

In Fig. 4, we have swapped Sean Connery face with Elon Musk face. Fig. 4 is a screenshot of a small video of the movie Goldfinger in which the actor Sean Connery played the role of James Bond. The Generative Adversarial Networks is a very powerful algorithm which helps in swapping of the face in the video. Fig. 4 is another example of Deepfake video which we have generated using Generative Adversarial Networks. We followed the same procedure as we did for Fig. 3 in which we swapped Henry Cavill face with Christopher Reeve face. This shows that if we have good data sets for training Neural networks then it will help in generating better output which looks authentic to human eyes. The changes in the above images are hardly recognizable due to the similarities between the images which have similar facial features, same skin tones and images shot under similar lightning conditions. The Deepfake technology requires lot of training and good processing power for generating Deepfakes. Even though it is optimal algorithm sometimes neural networks are trained properly but they are not able to swap the face in the video if the data set contains blurred images and sometimes the images are not taken under same lightning condition. Even though the Deepfake is optimal technique it has some advantages, drawbacks and limitations.

## III. ADVANTAGES OF DEEPFAKE

There are some advantages of using Deepfake technology if it is the used in correct and proper way. The most important use of deep-fakes comes under education and art [7].

### A. *Deepfake in field of education.*

The Deepfake is capable of generating fake video by swapping the face of the person. This quality of Deepfake can help a lot in the field of education [7]. We can create a Deepfake of freedom fighters, scientist, doctors and any other person who invented new thing and now they are dead. Deepfake can create a fake video of these people which will help the student in their studies. It will be an innovative way of teaching student which will increase the interest of student towards the study. Deepfake will help in creating videos of the historical figure which will communicate directly to students [7].

### B. *Deepfake in film and art creation.*

Deepfake is capable of swapping face in videos and images with a realistic look and feel. There are some scenarios in the film where the film creator wants to change the face of actors due to some reasons. In fast and furious 7 when the actor Paul Walker died in a car accident then his brother Cody Walker who looks similar to Paul finished the last scene in the Movie. Think if he has no brother then in that case the film producer has to use lots of money in CGI and video editing for swapping the face in the movie or they have to start shooting of movie from the scratch. For swapping the face in movies with CGI and video editing, the editor needs to be a skilled person then only it will work and it required lots of money. But if we use Deepfake for swapping the face of actors in movies then it will only cost for high-end computer and it required a few months to swap the face in the movies. This technique is very effective and costs saving. In Star Wars Rogue One CGI was used to show a younger version of Princess Leia which is portrayed by Carrie Fisher, they used lots of money and skilled person for that, it would be done using Deepfake which would have saved lots of money and time.

### C. *Deepfake for pattern and design creation.*

Deepfake is one of the best techniques which learn and generate output from multiple inputs. It is used for face swapping but it can also be used for generating images. In fashion industry design and patterns play a very important role and as the time passes designer come up with new design and pattern for clothes. This technique can be used in creating pattern and design from the combination of old design and patterns which help the fashion designer in creating a new design for cloth foot-wears, bag, and wallets with the help of generative adversarial network, the designer will create a new design by providing images of cloth and bag they can generate new shoes or by providing images of shoes and wallet it will generate a design for cloth. If Deepfake is used in the fashion industry, it will save lots of time and effort of designer for creating new design and patterns. Now the developer has started creating an app which can be used by novice person and you have to

provide images and it will train algorithm and generate a new design from it.

## IV. DISADVANTAGES OF DEEPFAKE

This technology was created for entertainment purpose but now as the Deepfake gaining popularity its miss use has been increased. This technique is now being used for character defamation of celebrities and high-profile politician just by swapping the face with someone else. This technique could be problematic if it is utilized in unethical ways. The most important disadvantages of Deepfake are as follows:

### A. Individual character defamation and assassination

Deepfake is capable of making a dead person alive just by swapping the face in the video. It can be used against any individual for defaming the character of an individual person. it will become harmful for a person because it will show the thing which he didn't do. The impact of the video will be much higher because people will believe what they saw in the fake video and it will become difficult for the individual to prove he is innocent and did nothing. It will lead to a person character assassination just by using fake video to defame his character.

### B. Spreading fake news with Deepfake.

The Deepfake is now improving day by day and its flaws are getting fixed as the technology getting popular. The output of the Deepfake is getting much better day by day and it looks authentic to human eyes and it getting very difficult to distinguish between real and fake video. People start spreading fake news with evidence which looks authentic to user and this is not new, previously people use to do it by spreading fake message on social media and now they are doing it with fake video which works as icing on the cake because it's impact will be much more high.

### C. Using Deepfake against celebrities.

Deepfake was created for entertainment purpose and now it is being used as weapons against celebrities for character assassination. In 2017 people started targeting celebrity by a making fake adult video of them by swapping their face with an adult actress in the video. In old days' people used image editing software and video editing software for swapping the face of actresses and it was easy to detect such changes in images but the result of Deepfake look very authentic and convincing to human eyes. This one is the major disadvantage of Deepfake because people started using this for making a fake adult video of celebrities. Deepfake made celebrities to suffer a lot because it is a threat to their public image and their fame and it is affecting very badly because people easily believe in such rumors.

### D. Deepfake is a threat to democracy.

Deepfake is now creating the boom of fake new and it is spreading like fire and people who are not aware of this technology are becoming the victim because they think it is an authentic content as Deepfake is capable of swapping face in the video. If this technology is used against politician then it will damage their reputation and character. The attacker will make them appear in the place where politicians never visited and making them behave badly with people which they never did. The Adobe Company is working on an application which will work as Photoshop of audio which was demonstrated by Adobe developer Zeyu Jin on Adobe MAX conference 2016 in San Diego, California. It is like the text to speech Software which is capable of learning human voice and then giving output in that particular person voice and it was demonstrated by Adobe developer Zeyu Jin in the conference [8]. So if people used these two technologies together then it will become harmful and it will be used in altering the election just by defaming the politician in a situation which they never did. It is a big threat to democracy which will change the outcome of the election.

### E. Deepfake is a threat to law enforcement agencies.

In many court cases if we want to prove if someone is guilty or not, we use evidence like voice, CCTV footage, witness and images because these are the strong pieces of evidence against any criminal and it will help in taking the decision against them. This evidence provides valuable information against the criminal and it can be used against anyone just by swapping the face of criminal with an innocent person to create a Deepfake and it will portray an innocent person doing the criminal activity which he never did and it can be used as evidence against the person. As the technology improving the output is also getting better and it is becoming hard to detect Deepfake. NVidia has created Style-Based Generator for Generative Adversarial Networks which is capable of generating face which look very accurate and it is hard to detect whether it is fake or real [9]. If these things are used as evidence, then it will be a disaster for law enforcement agencies.

## V. DEEPFAKE DETECTION TECHNIQUES

### A. Detecting 'Deepfake' videos by the blinking of an eye.

Deepfake is capable of making the thing real which never happened but for that, we need lots of images for making the fake video which looks realistic to human eyes and the algorithm take lots of time just to create the perfect combination of images. As the algorithm need lots of processing power and not everyone owns a high-end system for making the Deepfake and because of this sometimes the algorithm generates inappropriate video which can be easily detected by human eyes. As the misuse of Deepfake is increasing day by day so it becomes necessary to spread awareness about this technique and develop some way in which it can be easily detected. Sometime Deepfake can be easily detected by human eyes just by looking to face because the images generated by this algorithm sometime were not able to mimic blinking of an eye [10]. This method becomes very useful in detecting Deepfake because there are 3 types of blinking of the eye, such as spontaneous blinking, reflex blinking and voluntary blinking [11]. A normal person blinks in between 1 to 10 sec and this cannot be possible with Deepfake unless and until you provide images of a person with a closed and open eye. So, this can

be a very important technique through which Deepfake can be detected.

### B. Detecting Deepfake with MesoNet.

MesoNet is capable of automatically detecting facing tampering in the Deepfake video using deep learning [12]. This method tells the difference between computer-generated images and real images in the Deepfake video by using two network architecture meso-4 and mesoinception-4. The main purpose of these two architectures is to detect facial video forgery with high accuracy and it can differentiate between image property such as image noise, image accuracy, classification and aggregation of the image. After analyzing these features of the image, the both meso-4 and mesoinception-4 are capable of detecting the Deepfake video with the accuracy of 95% to 98%. The Deepfake is getting popular day by day and people are coming up with a new way of enhancing Deepfake which look more authentic to human eyes and people using this video for spreading fake news and videos. The mesoNet technique can provide an aid to detecting such fake videos and help in preventing fake news and video which are created through Deepfake.

### C. Detection Of Deepfake Using Recurrent Neural Network

Previously only a few people have the knowledge about the Deepfake algorithm and now it is becoming more popular and the developer is making an app which made Deepfake easier and now laymen can create Deepfake with a much easier process using the fake app. Recurrent Neural Networks is another way of detecting the Deepfake video [13]. Deepfake Video Detection Using Recurrent Neural Networks. The concept behind this to train the convolutional neural network for extracting the feature of images and then training the Recurrent Neural Networks for detecting if any changes are made in the video. RNN long-short-term-memory (LSTM) is capable of remembering the previous output which is used as input in the next step. CNN is the best known for its visual recognition and if it is combined with LSTM it can easily detect changes in the frames and then this information is used for detecting the Deepfake which are created using the fake app. fake app is not full-fledged software and it generates video with inconsistency and it can be easily detected by recurrent neural network.

### VI. LIMITATIONS OF DEEPFAKE

### A. Need lots of image data set.

No doubt Deepfake is very powerful technology. As said with great power come great responsibility. In order to make realistic Deepfake video we need lots of images of person x and person y whose face we want to swap. At least we need thousands of images in order to train the algorithm to learn and generate the authentic and realistic Deepfake and it's not possible to get 1000 of images of person until and unless he is celebrity of V.I.P. Even though you get 1000 of images of them it's not possible to generate a Deepfake which can mimic blinking of eyes because nobody posts images with closed eyes. This is the major problem with the Deepfake. So we need to extract images from the videos to create the

data sets and it makes more time consuming some time we do not get desired images which we required for data sets.

### B. Training and swapping is time-consuming.

To train the generative adversarial network for Deepfake it required lots of time because it very slow process it might take weeks for generating a realistic output. As the algorithm works with two neural network generator and discriminator, the generator generates the output and it is evaluated by the discriminator it verifies whether the image will look authentic to human eyes and if the image is fine then it will start the swapping process or else it tells generator to generate another image and the process goes. The algorithm needs a minimum of 72 hours' hour or more for training and then it will take time for swapping which become more time-consuming. Then we get a realistic Deepfake.

### C. Deepfake process is very costlier.

Training and swapping are heavy process and it requires a lot of graphics processing because CPU will take time in processing those images. So, it is not possible to generate a Deepfake without a graphics card and that is why not everyone can do it. For this, we need a very high-end computer which can handle both training and swapping process. That's why only a few people with good graphics processing power are able to generate it. The cost of graphic card is very high not everyone can afford it. This makes the Deepfake technique vey costlier.

### D. Similar faces and skin tones of people.

In order to make a realistic Deepfake we need people with similar faces and skin tones because if the feature of both person x and person y is not similar then it will result in terrible Deepfake which can be easily detected. So, the feature of the face needs to be same if person x has the beard then person y should also have the beard otherwise it will not work and the skin tone of person x need to be similar with person y because it will not look authentic to human eyes.

### E. Sometime output is not optimal.

Sometimes the output of the neural networks are good but while swapping the orientation of face is not matched with the target video then the output of Deepfake video is not optimal. Deepfake which is not optimal can be easily identified by human eyes. When the swapping of the face is not done properly then, if the person moves his head in different directions then we can easily identify if there is something wrong in the video and anyone can detect and identify if the video is fake or not by looking at the face. This kind of mistake happens in Deepfake when we don't build the image data sets with faces in a different direction and similar skin tone of both people. This happens a lot in Deepfake because it's difficult to get images of faces in different directions and similar skin tone.

## VII. CONCLUSION

This survey paper draws the conclusion that using Deepfake technique with the generative adversarial network can generate an output, which looks realistic to human eyes. Generative Adversarial Network creates a fake image by combining two different images but the images of person A and person B need to similar in terms of facial feature and skin tone and the images must be taken under the same lighting condition. Deepfake can be used in two ways either in good ways by implementing Deepfake in education domain for changing faces of historical people and using them as study material, and in arts it can be used for swapping face of actors in movies which will save lots of money from using CGI and VFX, and Deepfake can help in generating design and pattern for the fashion industry. Deepfake technique can be misused. For spreading fake new by swapping the face of people in the video and it will make them do the thing which they never did. Deepfake can be used for defaming political leaders which will influence the result of an election. Most dangerous use of Deepfake is to taking revenge from another person by swapping his face in adult movies and the female celebrities are becoming the victim of Deepfake. There are different ways through which we can detect whether the video is Deepfake by using different algorithm Even though Deepfake is an optimal technique for generating fake video. This technique is more optimal because its output looks very realistic and authentic to human eyes, which make it more optimal and harder to detect.

## REFERENCES

[1] Radhakrishnan, Sreedhar & Jay Kuo, C.-C. (2018). Synthetic to Real World Image Translation Using Generative Adversarial Networks. 1-6. 10.1109/ICCCNT.2018.8493745.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672– 2680.

[3] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017.

[4] Jonathan Hui: GAN-Why it is so hard to train Generative Adversarial Networks!. 2018. https://medium.com/@jonathan_hui/gan-why-it-is-so-hard -to-train-generative-advisory-networks-819a86b3750b

[5] Wikipedia: Video. https://en.wikipedia.org/wiki/Video

[6] KDnuggets Exploring DeepFakes. 2018. https://www.kdnuggets.com/

[7] Chesney, Robert and Citron, Danielle Keats, Deep Fakes: 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. Available at SSRN: https://ssrn.com/abstract=3213954

[8] Zeyu Jin, Adam Finkelstein, Gautham J. Mysore, and Jingwan Lu. FFTNet: a Real-Time Speaker-Dependent Neural Vocoder. The 43rd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2018.

[9] Tero Karras, Samuli Laine, Timo Aila: A Style-Based Generator Architecture for Generative Adversarial Networks. 12 Dec 2018.

[10] Yuezun Li, Ming-Ching Chang and Siwei Lyu: In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking, in IEEE International Workshop on Information Forensics and Security (WIFS), 2018.

[11] Matteo Bologna and Rocco Agostino and Bruno De Gregori and Daniele Belvisi and Donatella Ottaviani and C. Colosimo and Giovanni Fabbrini and Alfredo Berardelli: Voluntary, spontaneous and reflex blinking in patients with clinically probable progressive supranuclear palsy, in Brain: a journal of neurology, 2008.

[12] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018, September). MesoNet: A Compact Facial Video Forgery Detection Network. In IEEE Workshop on Information Forensics and Security, WIFS 2018.

[13] Guera, David & Delp, Edward: Deepfake Video Detection Using Recurrent Neural Networks, in IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018.