IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 14, 2020, accepted April 8, 2020, date of publication April 20, 2020, date of current version May 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2988660

DeepVision: Deepfakes Detection Using Human Eye Blinking Pattern

TACKHYUN JUNG[1], SANGWON KIM[2], AND KEECHEON KIM[3]
[1]Department of IT Convergence Information Security, Konkuk University, Seoul 05029, South Korea
[2]Department of Computer, Information, and Communications Engineering, Konkuk University, Seoul 05029, South Korea
[3]Department of Computer Science and Engineering, Konkuk University, Seoul 05029, South Korea

Corresponding author: Keecheon Kim (kckim@konkuk.ac.kr)

This work was supported by the Konkuk University, in 2019.

ABSTRACT In this paper, we propose a new approach to detect Deepfakes generated through the generative adversarial network (GANs) model via an algorithm called DeepVision to analyze a significant change in the pattern of blinking, which is a voluntary and spontaneous action that does not require conscious effort. Human eye blinking pattern has been known to significantly change according to the person's overall physical conditions, cognitive activities, biological factors, and information processing level. For example, an individual's gender or age, the time of day, or the person's emotional state or degree of alertness can all influence the pattern. As a result, Deepfakes can be determined through integrity verification by tracking significant changes in the eye blinking patterns in deepfakes by means of a heuristic method based on the results of medicine, biology, and brain engineering research, as well as machine learning and various algorithms based on engineering and statistical knowledge. This means we can perform integrity verification through tracking significant changes in the eye blinking pattern of a subject in a video. The proposed method called DeepVision is implemented as a measure to verify an anomaly based on the period, repeated number, and elapsed eye blink time when eye blinks were continuously repeated within a very short period of time. DeepVision accurately detected Deepfakes in seven out of eight types of videos (87.5% accuracy rate), suggesting we can overcome the limitations of integrity verification algorithms performed only on the basis of pixels.

INDEX TERMS Cyber security, deep-fake, GANs, deep learning.

I. INTRODUCTION

In recent years, various social issues have arisen because of fake videos called Deepfakes. Generated using the generative adversarial network (GANs) model, Deepfakes are created by iterating an actual data-based generation and verification task through two opposite deep learning models [1].

This principle means that faces or specific body portions in videos or photographs can be synthesized to artificially obtain the information of other people. At its early stages, Deepfakes videos could be detected through the naked eyes because of the pixel's collapse phenomena that generate unnatural visual artifacts in the skin tone or face contour of images or frequent visual artifacts. However, with the technology's advancement', Deepfakes have evolved to be highly indistinguishable from natural images [2], [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Mamoun Alazab.

As a result of the technical advancement, there has been an increase in the frequency of its improper using Deepfake, a large number of pornographic photographs of celebrities and politicians have been produced for the purposes of spreading propaganda and fake news, causing a wide variety of social problems [4].

According to the Washington Post [5], the crime victims of these Deepfakes photographs have been expanded to the general public, and face photos and pornographic photographs are now skillfully synthesized and spread through social networking services without the consent and permission of the related parties. Some companies even specialize in providing such Deepfakes services.

Thus, considerable research attention has been paid to the development of a method that will verify the integrity of Deepfakes. As one of the most actively studied integrity verification methods, a method that detects the collapse of pixels and visual artifacts in Deepfakes has been proposed.

However, as the generator and discriminator [1] in the GANs model have advanced to bypass such verification [2], [3], this measure has faced a problem in its utilization. Thus, eye blinking, which is a unique action that is iterative and occurs unconsciously, provides an alternative solution to find the integrity verification indices that are difficult to verify using the discriminator in the GANs model.

If the eye blinking pattern that occurs irregularly can be formulated and analyzed through a number of algorithms, it may not only be difficult to verify using the discriminator but also be highly useful in terms of integrity verification. Thus, in this study, we aimed to implement an algorithm that observes and analyzes various cognitive and behavioral indicators that affect eye blinking, thereby discussing the possibility of identifying Deepfakes based on the information pattern of eye blinks, which is a voluntary and unconscious behavior.

II. RELATED WORKS

A. RESEARCH TRENDS OF DEEPFAKES

The first proposed GANs model [1] has great significance in that it invented a new way of learning by producing data with the Generator and validating it with the Discriminator. However, it has faults such as the minimax problem [2] or the saddle problem [2], resulting in unnatural spectra in the outline and shade of generated pictures.

In 2016, DCGAN(Deep Convolutional Generative Adversarial Networks) proposed by Alec Radford *et al.* made possible arithmetic operations with filters between images using latent vector by applying CNN (Convolutional Neural Network) [2] models to GANs, emerging more clever forgeries.

This development was further built upon in 2017 by a research team from the University of Washington that produced sophisticated fake videos that matched a speaker's voice and mouth in a video and produced the shape of his mouth for every moment. Through this, the previous limits of pixel crush, jaw form, wrinkles, etc. were greatly improved upon by applying methods such as jaw correction [3].

The continuous development of the GANs model makes it more difficult to verify the integrity of Deepfakes. Previous integrity verification methods, which detected the crushing of pixels or inconsistencies in the outline, have significantly lost their effectiveness, and this methodology is expected to be significantly underutilized as these parts continue to be improved in the GANs model. Therefore, this study is mainly focused on finding elements of new integrity verification.

B. DETECTION OF DEEPFAKES

We considered various methods used in the forensic community for generic fake video detection [6], [7], face matching detection [8], [9] and eye blinking detection [10]. Of all, the most widely used detection methods of Deepfakes is training a dataset of facial forgeries in deep neural networks or detecting a pixel's anomaly [23], [24].

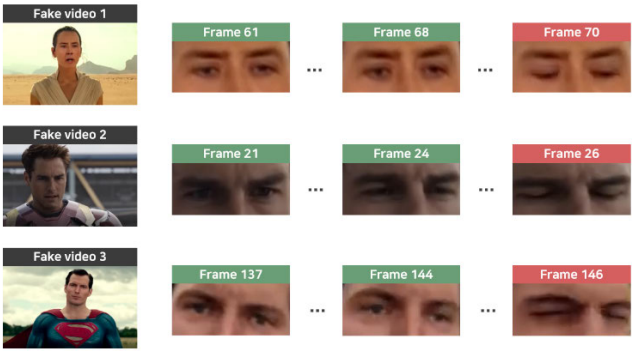

FIGURE 1. Deepfakes have become more elaborate with eye blinks.

FaceForensics++ [8] is an effective dataset of facial forgeries that enables to train deep learning-based approaches. And research [9] is approach of detection fake video through trained CNN (Convolution Neural Network) [11]. These methods promise trustworthy results but require a lot of data and need to be improved periodically. Thus, we focused on research [10], which does not require as much data, but is likely to be used more widely.

Research [10] found that many faces generated in Deepfakes do not eye blink. However, as seen in Fig. 1, a number of new cases that have adjusted the discriminator to verify blinking have recently emerged to circumvent these detection techniques. [28].

Such technological improvements raise the need for more advanced integrity verification technologies, such as DeepVision [10]. DeepVision performs integrity verification by tracking significant changes in the eye blinks in deepfakes by means of a heuristic method based on the results of medicine, biology, and brain engineering research as well as machine learning and various algorithms based on engineering and statistical knowledge. This comprehensive method will aid us overcome the limitations of integrity verification performed only on the basis of pixels.

C. HUMAN'S EYE BLINK

Humans blink iteratively and unconsciously daily, to maintain a certain thickness of the tear film on the cornea [12]. However, eye blinks serve more purposes than maintaining the cornea [14], as suggested by how there is a difference in the frequency of eye blinks between adults and infants [13].

Notably, blinking frequency fluctuates based on a person's activity. When reading out loud specific sentences or performing rehearsals of presented visual information, the number of eye blinks increases, whereas it decreases when a person concentrates on visual information or silently reads sentences [15].

In addition, a study by Ponder in 1928 reported that when people conversed with one another, their eye blinks increased [16]. This implies that eye blinks are also affected by cognitive activities and behavioral factors [17]. Moreover, the number of eye blinks varies throughout the day depending

on time: the highest number of eye blinks is usually observed at nighttime around 8 pm [18].

The fact that blinking frequency is affected by a variety of factors, such as an individual's physical condition, cognitive activities, physiological factors and information processing level [14], [15], means that, by collecting and statistically analyzing this information, the number and range of eye blinks can be predicted to some extent. In fact, the three Deepfakes measured in Fig. 1, which all showed an unnatural visual effect, also had less than five blinks per minute, significantly less than the average number of eye blinks [22].

Therefore, we conducted an experiment to verify the integrity of Deepfakes by implementing a method that predicts the number of eye blinks that people of a given age and gender will perform under certain conditions.

III. DEEP VISION

A. ARCHITECTURE OF DeepVision

Here, we present an architecture for Deepfakes detection using analysis of eye blinking. The proposed method called DeepVision has a simple process structure, as seen in Fig. 2.

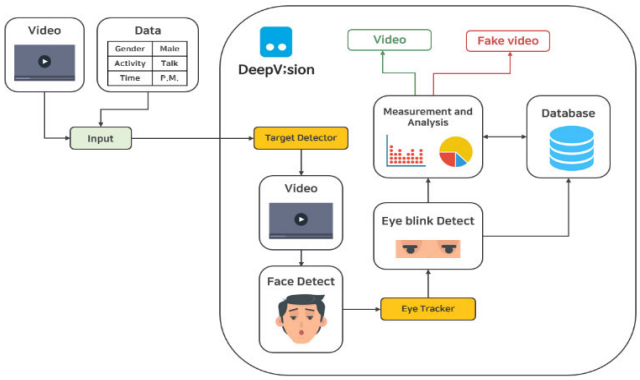

FIGURE 2. This architecture is the DeepVision that we propose. It is able to detect a face area, locate face and eyes landmarks, track the human's eye blink, and classify a given video as Deepfakes or generic video.

First, DeepVision's architecture has a pre-process, which receives input of information. Through this process, data such as gender, age, activity, and time are inputted as important parameters that can verify changes in the human's eye blinks. After this pre-process, DeepVision conducts measurements through the Target Detector, which detects objects in the video, and the Eye Tracker, which tracks the blinking. These processes are performed in frame unit, and the measured data is compared with DeepVision's database of natural movements to verify that a human's eye blink is either natural or fake.

B. INPUT DATA IN PRE-PROCESS

DeepVision aims to track blinking patterns that significantly fluctuates with regards to gender, age, activity, and time factors. Thus, in the pre-process, we watch the sample video and extract and input these variables into the DeepVision through parameters. At this time, the type of data input is defined as Table 1.

TABLE 1. Type of input data in pre-process.

Category	Input
Gender	Male, Female
Age	<20, 20-30, 30-40, 40-50, 50-60, 65+
Activity	Dynamic / Static
Time	A.M / P.M

There is a difference in the average number of eye blinks between males and females [15]. Thus, gender data is inputted to track these differences and changes. In addition to gender, age is also directly related to the number and period of eye blinks [18]. Thus, age data is also inputted for consistency, with the data subcategorized into one of six groups, ranging from less than 20 years old (<20) to over 65 years old (65+).

Blinking frequency also fluctuated based on the type of activity a person is engaged in and on external recognition [15]. Thus, activity data was inputted for measure. For example, the number of blinking decreased and was less than average while performing a static activity such as "reading a book" that focuses on visual information [15], [19]. On the other hand, the number of blinking increased and was above average while performing a dynamic activity such as "talking" or conducting "physical movement", or during a "moment of recall of a particular sentence" [15].

In addition, blinking significantly changes over time [18]. Therefore, time data was inputted, categorized as A.M. or P.M. Each data entered in this process was then transferred to the next step, Target Detector, with the analysis target (video).

C. THE TARGET DETECTOR

As shown in Fig. 3, DeepVision fuses the Fast-HyperFace (face detect) [20] and EAR algorithm (eye detect) [21] to track the blinking, utilizing the synergy of their performance advantages.

Fast-HyperFace was invented by Rajeev Ranjan *et al.* [20] It is an algorithm for face detection, landmarks localization, pose estimation, and gender recognition. DeepVision's Target Detector was implemented by utilizing this algorithm:

Algorithm 1 The Target Detector

Input: Trained Fast-HyperFace(model) *hf*

Loop(frames)**:**
frame ← **Pre_Process**(frame) # astype, resize, normalization
landmarks, detections, poses, genders, rects ← *hf* (frame)
IF detections > 0.7:
crop_area ←**Make_Outer_line**(landmarks, frame, rects)
Forward_to_EyeTracker(crop_area, landmarks, rects)
ELSE: PASS

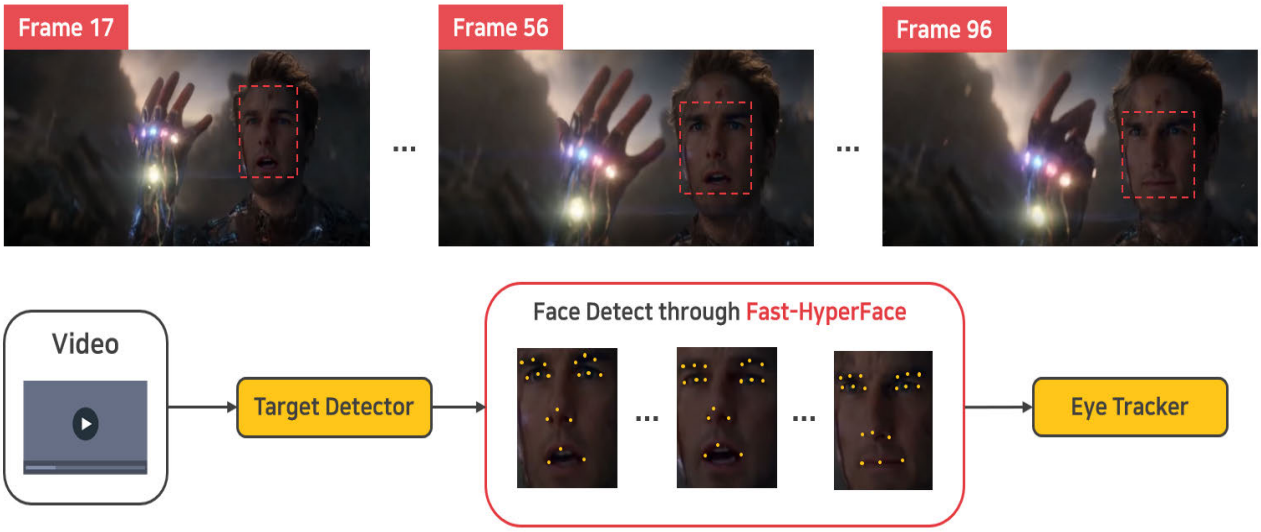

FIGURE 3. This is a visualization of DeepVision's target detector stage. It is able to slice a given video into frame units, detect a face through Fast-HyperFace in frames, and forward the detection results to the Eye Tracker.

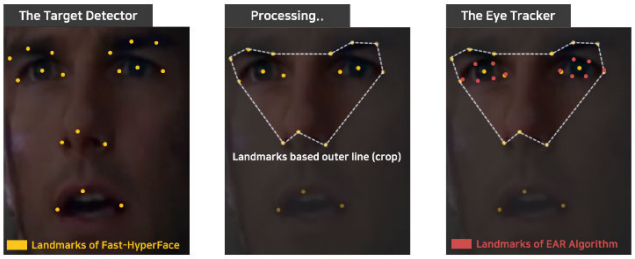

FIGURE 4. This is a visualization of the process that fuses the target detector and eye tracker. The EAR algorithm is applied after the detection area is limited through outer lines with Fast-Hyperface's high detection performance. This method performs better than using EAR alone.

Algorithm 1. shows the sequence and principle of performing face detection through the Target Detector and links the results to the next step (Eye Tracker). In the Pre-Process, it is performed that data type conversion, image resizing and normalization. Then, the face detection is performed through a trained Fast-HyperFace model.

We define the detection rate of each frame above 70 percent as accurate detection. If the conditions apply, it makes an outer line based on Fast-HyperFace's Landmarks.

As seen in Fig. 4, this improves detection performance by effectively limiting the detection coverage of EAR [21] through utilizing Fast-HyperFace's Landmarks based outer line.

Fast-HyperFace has high detection performance, although it is not able to detect eye blinking alone. In contrast, the EAR algorithm is able to detect eye blinking on its own, but the detection performance is poor. Thus, the proposed architecture in this study utilizes the performance of both of these algorithms, taking advantage of the synergy.

D. THE EYE TRACKER

The Eye Tracker was implemented based on EAR (Eye-Aspect-Ratio) [21]. Invented by Tereza Soukupova and Jan Cech in 2016, EAR takes six points(p_i) around the eyes and calculates the absolute area of the horizontal axis and vertical axis.

$$\mathrm{EAR} = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \tag{1}$$

Eq. (1) is the EAR formula used to detect eye blinks, as defined in research [21]. Points p_1 and p_4 refer to the horizontal axis in the eye area, as shown in Fig. 6, and the other points refer to the vertical axis. Thus, EAR is an absolute value of the size calculated through the area of the horizontal and vertical axes.

In general, eye blinks occur simultaneously in both eyes. Thus, we used Eq. (2), which sums and divides in half the eye's ratio (EAR_i) through using the value of the left eye (EAR_l) and right eye (EAR_r).

$$EAR_i = \frac{(EAR_l + EAR_r)}{2} \tag{2}$$

The value of EAR_i is able to detect an eye blink that is smaller than a threshold. The thresholds used are defined in Fig. 7.

$$\sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}} \tag{3}$$

Eye blinking is performed in two ways: one is a process of closing one's eyelids and the other is a process of lifting closed eyelids. Thus, we implemented a method for detecting eye blinks by utilizing EAR_i, as seen in Fig. 8.

When eyes are closed, the EAR_i is reduced and drops below the threshold in the consecutive video frames. Then, when eyes are open, the EAR_i is restored as before. In this

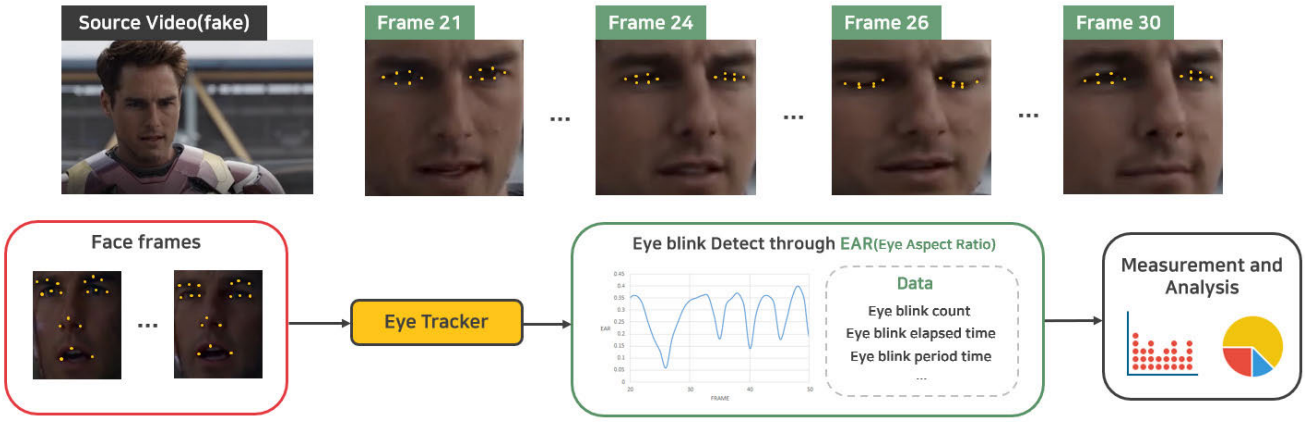

FIGURE 5. This is a visualization of DeepVision's eye tracker. It is able to measurement the blinking count, period, and more through the EAR (Eye-Aspect-Ratio) algorithm.

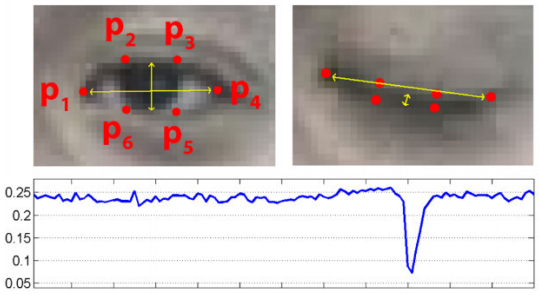

FIGURE 6. This equation shows the calculations of EAR in the frame unit. The vertical axis represents the value of EAR, and the horizontal axis represent the time [21].

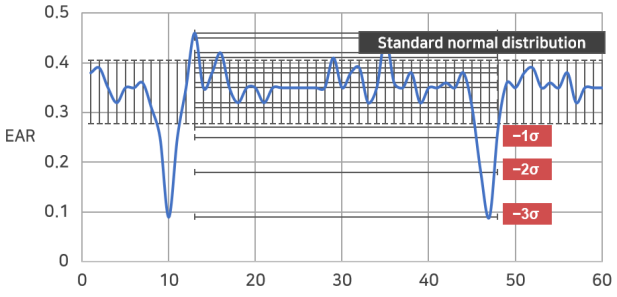

FIGURE 7. This figure shows the proposed method of finding the appropriate threshold. In this study, the threshold value was defined by the minimum value that was outside the range of -2σ in the standard deviation. Eq. (3) was used to calculate the standard deviation, where x means the average of the sample and n means the size of the sample.

process, the time required to blink, and blinking frequency is obtained through measuring the time and changes of the EAR. Fig. 5 visualizes these measurement methods.

Fig. 9 shows how to measure the period of an eye blink through EAR_i, When an eye blink occurs, the period is calculated from the end time point to the next eye blink start time point. The calculated period can be used to detect various abnormal patterns that result from randomly generated eye blinks through a loop or a specific algorithm.

Algorithm 2 The Eye Tracker

Input: threshold **t,** left eye**l,** right eye**r**

output: log(blink count, time, period, elapsed time and etc)

```
def Track_EAR(eye):
    h_axis ← dist.euclidean(p [1], p [4])
    v_axis1 ← dist.euclidean(p [2], p [6])
    v_axis2 ← dist.euclidean(p [3], p [5])
return (v_axis1 + v_axis2) / (2 ∗ h_axis)

Main:
    Loop(frame in frames):
        LeftEAR ← Track_EAR(l)
        RightEAR ← Track_EAR(r)
        EAR ← (LeftEAR + RightEAR) / 2
        logs ← logging(frame, EAR, time_capture())
    Loop(log in logs):
        IF EAR < t:
         blink_count + = 1
         blink_time, elapsed_time AND etc ←
         time_analysis()
         blink_period AND etc ← period_analysis()
```

Fig. 10 shows a graph of consecutive eye blinks, which occurs within a very short time period for many humans. DeepVision was implemented as a measure to verify an anomaly based on the period, repeated number, and elapsed eye blink time when eye blinks were continuously repeated within a very short period of time. If eye blinks were generated arbitrarily in Deepfakes, this detection method could be an important element of integrity verification.

Here, [Algorithm 2] is the specification of a series of operating procedures for the Eye Tracker described in this section. The second paragraph of this specification is the process of calculating the aspect ratio of the eye through

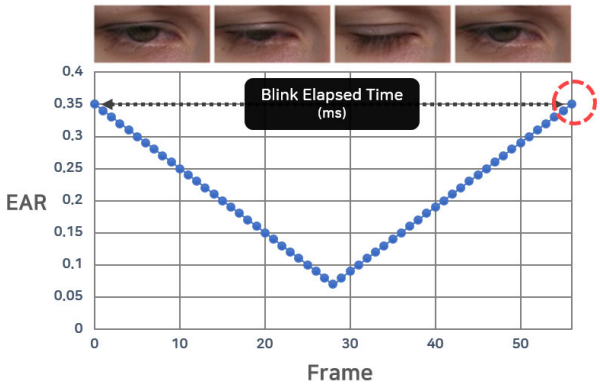

FIGURE 8. This figure shows the method used to measure the elapsed time it takes to blink. The vertical axis represents the value of EAR(i), and the horizontal axis represent the frames.

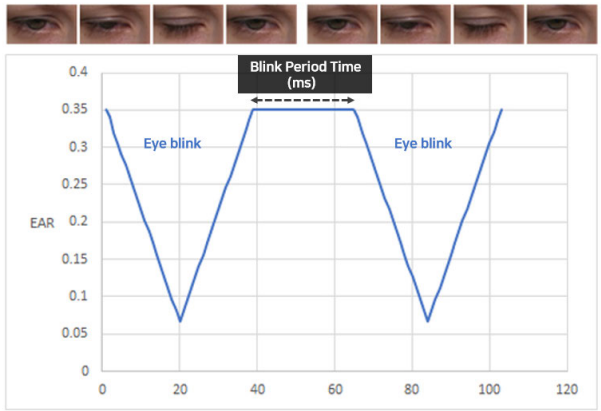

FIGURE 9. This figure shows the method used to measure the eye blink period. The vertical axis represents the value of EAR(i), and the horizontal axis represent the frames.

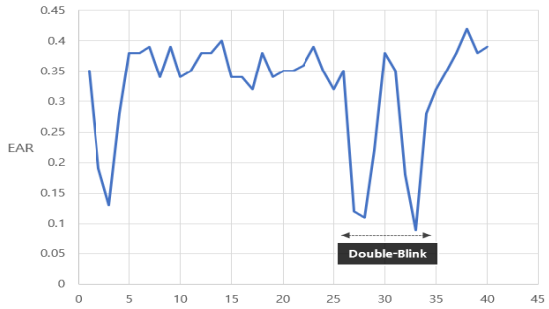

FIGURE 10. This figure shows the method used to measure consecutive eye blinking. The vertical axis represents the value of EAR(i), and the horizontal axis represent the frames.

the EAR algorithm, which is calculated from the left and right eyes respectively, in the third paragraph. In this process, the average of both eyes is obtained, and the value of EAR, the time, and the frame number are continuously recorded.

In the last paragraph of this specification, blinking count and period are calculated based on the recorded logs. This algorithm is a PoC(Proof of Concept) for better understanding, and it is possible to implement a real-time method based on the units of frames depending on the location of the function calls.

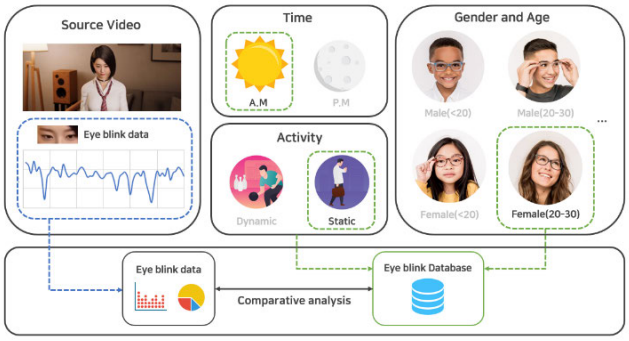

FIGURE 11. This is a visualization of the DeepVision integrity verification. It is performed by finding and contrasting information in the DB that matches both the information of eye blinks measured and the corresponding gender, age, time, and activity.

E. INTEGRITY VERIFICATION IN DeepVision

Fig. 11 shows the method for verifying integrity by utilizing information from the blink of the eye measured in the video. This is done by finding and comparing pattern information matching the corresponding gender, age, activity, and time in the pre-configured database with speed and frequency of blinking measured from the target in the video. In this case, the information entered in the Pre-Process for gender/ age/behavior/time is used as search criteria for the DB. Details of DB search method are shown in Fig. 12.

DeepVision performs integrity verification by tracking the fluctuation of eye blinks based on the four factors above. In comparing the information from blinking stored in DB with the information from actual measured blinking, DeepVision requires appropriate methods and criteria for determining the level of acceptance.

This is because people with the same factors may have similar, but not entirely consistent, blinking counts. Thus, this study established an initial DB of blinking count based on the Eye Blinking Prediction dataset (Kaggle) [26] and conducted a study on the acceptable range.

The dataset (eye blinking prediction) was invented by Rösler and Suendermann [27] It is a reliable and effective data of eye blinking created by an Emotive headset device with 14 sensors. The sensors continuously record brain signals and eye states, after which observed eye states were manually added.

The dataset [26] is very valuable on its own, as measurements have been carried out by professional experiments and outstanding researchers. However, they did not consider gender, age, and changes of various states in the measurement.

Therefore, we constructed the initial DB based on this dataset and related research of eye blinks [13]–[18], then obtained data for further experiment by calculating the increased or decreased value according to conditions such as gender and age through several experiments. We are going to

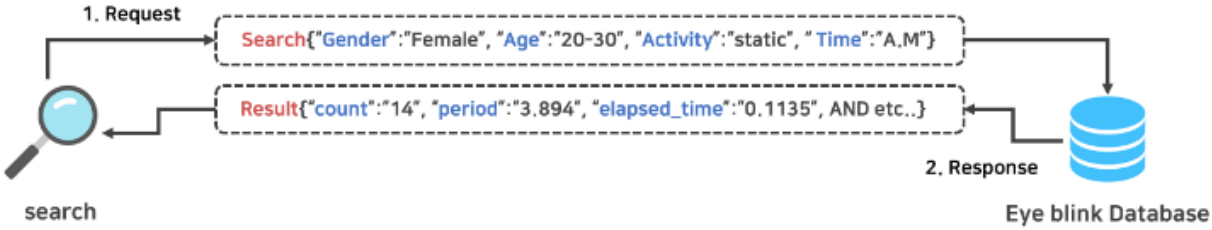

FIGURE 12. This figure shows the search and database connection and process. Using a JSON format, the function can communicate the eye blink count, period, elapsed time, etc.

improve this in the direction of collecting statistic information through more experiments in subsequent studies and through the open public in Kaggle or Github.

Algorithm 3 A Method of Comparative Analysis

Input: Eye_blink $\boldsymbol{E}$ [count, period **AND** etc],
DB_Data $\boldsymbol{D}$ [count, period **AND** etc]
Output: Fake **OR** Not

Loop[i]:
IF($\boldsymbol{E}$[i] < $\boldsymbol{D}$[i]){**IF**($\boldsymbol{D}$[i]− $\boldsymbol{E}$[i]) >= allowable range{**return Fake**}}
ELSE{**IF**($\boldsymbol{E}$[i]− $\boldsymbol{D}$[i]) >= allowable range{**return Fake**}}
i++

[Algorithm 3] shows a series of processes that perform integrity verification through the process of comparing and analyzing the information of blinking eyes measured in the video with DB. In the algorithm, $\boldsymbol{E}$ means the blinking information measured in the video, and $\boldsymbol{D}$ means the blinking information stored in the DB.

It is key to compare and analyze the number of blinks, cycles, average cycles, and duration in order. After subtracting the larger value from the other, the video can be considered as fake if the resulting value is larger than the allowable range.

As seen above, because DeepVision's integrity verification routine utilizes a combination of various factors to perform validation, it is very difficult to bypass it or falsify data by forging cycles, times, counts, etc.

IV. EXPERIMENTAL OF DeepVision

A. ACCURACY OF FAKE DETECTION

We evaluated Deepfakes Integrity verification following the frame analysis method conducted by Li *et al.* [10].

Case 1: Using DeepVision, we performed an experiment on Deepfakes within the timeframe of a minute. The results showed that during the minute, the value of EAR was almost constant, as shown in Fig. 13, indicating that there was no blinking during the experiment. Such lack of blinking can't occur in real humans, so the video was determined as "Fake".

Case 2: Using DeepVision, we conducted an additional experiment on another Deepfakes video under the same conditions. Around nine seconds after the measurement began,

TABLE 2. Accuracy of fake detection through DeepVision.

Case	Data of Measured Eye Blinks			Prediction
	Repeated Number	Period	Elapsed Time	
1	-/min	-/sec	-/sec	Fake
2	1/min	-/sec	0.1135/sec	Fake
3	6/min	6.3333/sec (average)	0.1321/sec (average)	Fake

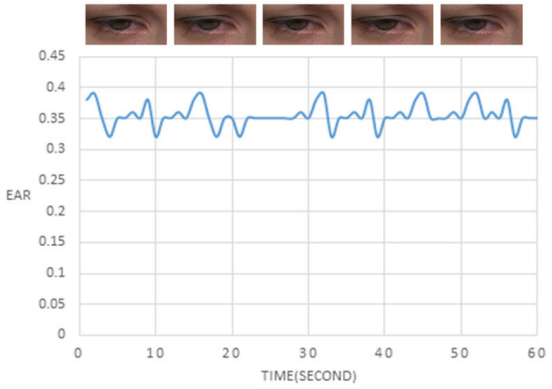

FIGURE 13. This figure shows the result of the Case 1 experiment. The vertical axis represents the value of EAR, and the horizontal axis represents the time. The EAR remained almost unchanged.

a blink was detected once, and the EAR level remained constant afterwards Fig. 14. The number of blinks (/min) and the period (/sec) measured in the experiment were significantly lower than the average (number, cycle, etc) of an actual human blink corresponding to the same condition (female+40-50+ Static+AM). Thus, the video was determined as "Fake".

Case 3: Using DeepVision, we conducted one more experiment on another Deepfakes video. The blink duration was similar to the natural human pattern corresponding to the same conditions (female+30-40+ Static +A.M). However, the measured eye blink number was as small as six times per minute Fig. 15, whereas the actual number of eye blinks corresponding to the condition is about 17 to 22 times per minute [22], [26]. In addition to the infrequent blinking,

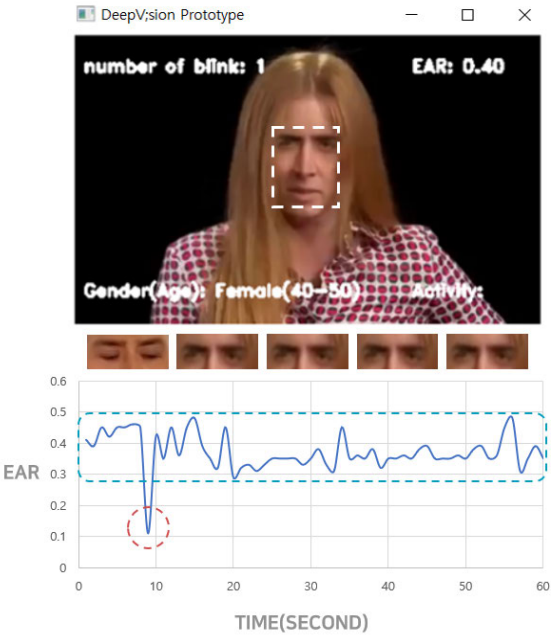

FIGURE 14. This figure shows the result of the Case 2 experiment. The vertical axis represents the value of EAR, and the horizontal axis represents the time. The value of EAR was lower than the threshold only once.

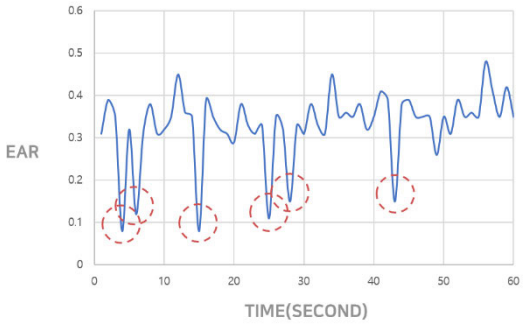

FIGURE 15. This figure shows the result of the Case 3 experiment. The vertical axis represents the value of EAR, and the horizontal axis represents the time. The EAR remained almost unchanged.

the periodic inspection was also determined to be abnormal. Therefore, the video was determined as ''Fake''.

These three Deepfakes video measured in this experiment which all showed an unnatural visual effect with eye blink, also had less than six blinks per minute, significantly less than the average number of eye blinks. This means that most deepfakes videos cannot fully forgery to accurate eye blinks.

B. BENCHMARKS COMPARISON

We tried to compare benchmarks in Deepfakes, following the method by Rössler *et al.* [8]. However, we realized that the benchmarks of previous related research compiled in Table 3. are different from our research [35].

Previous Deepfakes detection was performed based on two-dimensional pixels, as seen in Fig. 16. However, DeepVision is a new integrity verification method performing on

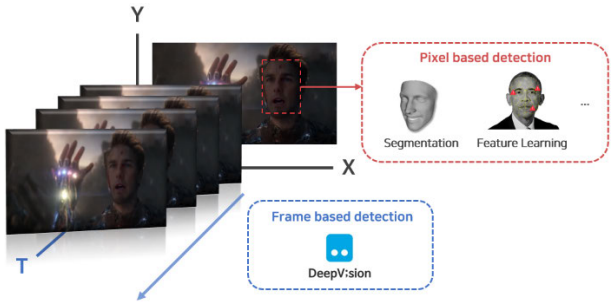

FIGURE 16. This figure distinguishes the various research areas for detecting Deepfakes. DeepVision is based on the frame of T axis, and in this regard, it is different from the previous related studies [6]–[9], which were performed on the pixel basis of two-dimensional images.

FIGURE 17. This dataset was previously used by related research to measure benchmarks [35]. It is a thousand static Deepfakes (images). This cannot be used in frame-based research because the images are all different and can't be repeated.

a frame basis. Thus, it was not possible to measure Benchmark using the same dataset [35] used in previous related researches.

Therefore, we measured Benchmark using our dataset composed of various Deepfakes videos. The result showed that Deepfakes were accurately identified in seven out of eight scenarios, illustrating a high accuracy rate of 87.5%.

In this Benchmark, Most of Deepfake videos had periodical, unnatural eye blink patterns, and only videos that forged mouth and nose except eyes could bypass DeepVision. Additional details of the benchmark are available in the reference section. Our dataset is available online at (https://github.com/takhyun12/Dataset-of-Deepfakes).

C. FEATURES OF DeepVision

We conducted an additional experiment to test the possibility of creating a Deepfake that can mimic the natural blinking pattern and bypass the DeepVision algorithm.

First, we schematized the measured eye blinks data from dataset [26]. Fig. 18. In the graph, the horizontal axis represents the index of the frame and the vertical axis represents the state of the eye, with 1 meaning the eye is closed and 0 meaning the eye is open. The two graphs illustrate two separate measurements of the same subject under the same conditions.

TABLE 3. Benchmark of Deepfakes detection with various methods [8].

Based on	Method	Accuracy
Pixels (Image)	FaceForensics++ [8]	0.964
	(Andreas Rössler et al.) FaceForensics++: Learning to Detect Manipulated Facial Images. ICCV 2019	
	MesoNet [32]	0.873
	(Darius Afchar et al.) Mesonet: a compact facial video forgery detection network. arXiv	
	XceptionNet Full Image [8]	0.745
	(Andreas Rössler et al.) FaceForensics++: Learning to Detect Manipulated Facial Images. ICCV 2019	
	Bayar and Stamm [6]	0.845
	(Belhassen Bayar et al.) A deep learning approach to universal image manipulation detection using a new convolutional layer. ACM Workshop on Information Hiding and Multimedia Security	
	Rahmouni [33]	0.855
	(Nicolas Rahmouni et al.) Distinguishing computer graphics from natural images using convolution neural networks. IEEE Workshop on Information Forensics and Security,	
	Recasting [7]	0.855
	(Davide Cozzolino et al.) Recasting residual-based local descriptors as convolutional neural networks: an application to image forgery detection. ACM Workshop on Information Hiding and Multimedia Security	
	Steganalysis Features [34]	0.736
	(Jessica Fridrich et al.) Rich Models for Steganalysis of Digital Images. IEEE Transactions on Information Forensics and Security	
Frames (Video)	**DeepVision [This Paper]**	**0.875**

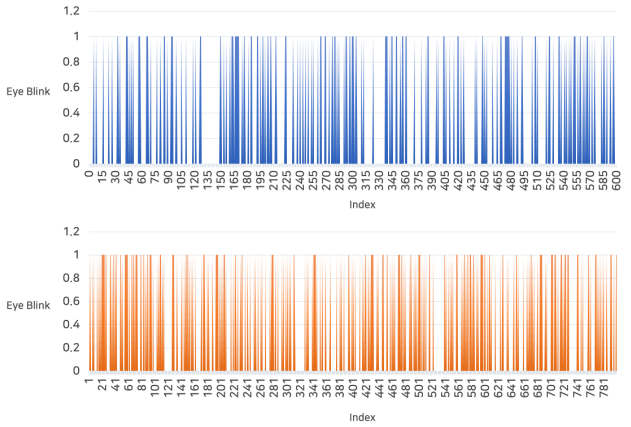

FIGURE 18. This figure illustrates the blinking distribution in dataset [26]. The x axis represents the index, and the y axis represents the state of the eye, with 1 indicating a closed eye and 0 indicating an open eye. Both plots show two different measurements of the same person.

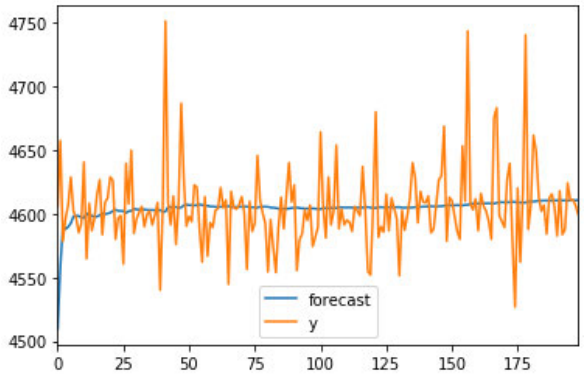

FIGURE 19. This figure shows the results of the ARIMA algorithm that utilized a time series model using the values of dataset [26] (represented as y). The t-test value ($P > |z|$) for the constant of the trained model is 0.120, and we can see that the value predicted by the real model (represented as a forecast) is very different from the distribution of an actual human blink.

We then analyzed the blinking distribution in both graphs. The results showed that the number of blinks was similar, but that the period pattern and timing were dissimilar. This signifies that it is difficult for algorithms or Deepfakes producers to easily predict the pattern of eye blinks.

In another experiment, we tried to determine whether it is possible to predict eye blinks through time series analyses such as the Prophet algorithm [30] and the ARIMA algorithm [31]. Using values of dataset [26] and the ARIMA [30] time series prediction model, it was determined that the t-test value was 0.120, which was not valid at the p-value of 0.05. By comparing the predicted value through the trained model and the actual value of the graph, through it is evident that it is impossible to accurately predict a human's blinking pattern, as shown in Fig. 19.

This means that because human eye blinking occurs unconsciously and spontaneously, [13], [17] the algorithm is unpredictable, bolstering its security. An attacker will find it difficult to use the GANs model to disable DeepVision.

However, we found some limitations in the experiment. The number of eye blinks was correlated with a mental illness closely connected with dopamine activity [15]. The study results revealed that the number (27 times/min) of blinking in patients with schizophrenia was considerably higher than

that of normal people (17 times/min), and that their blinking count was uninfluenced by the medicine administered [25].

In addition, Parkinson's disease, spasmodic torticollis, Tourette syndrome, and attention deficit hyperactivity disorder (ADHD) were found to be closely correlated with the number of eye blinks [15], [25]. Thus, the integrity verification of DeepVision may be limited by mental illnesses or problems in nerve conduction pathways.

V. CONCLUSION

In this study, we proposed and developed a method to analyze significant changes in eye blinking, which is a spontaneous and unconscious human function, as an approach to detect the Deepfakes generated using the GANs model.

Blinking patterns vary according to an individual's gender, age, and cognitive behavior, and fluctuates based on the time of day [34]. Thus, the proposed algorithm (DeepVision) observed these changes using machine learning, several algorithms, and a heuristic method to verify the integrity of Deepfakes. The proposed algorithm implemented using the results of various previous studies consistently showed a significant possibility of verifying the integrity of Deepfakes and normal videos, accurately detecting Deepfakes in seven out of eight videos (87.5%). However, a limitation of the study is that blinking is also correlated with mental illness and dopamine activity. The integrity verification may not be applicable to people with mental illnesses or problems in nerve conduction pathways.

However, this can be improved through a number of measures because cyber-security attack and defense evolve continuously. The proposed algorithm suggests a new direction that can overcome the limitations of integrity verification algorithms performed only on the basis of pixels.

REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[2] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*. [Online]. Available: http://arxiv.org/abs/1511.06434

[3] S. Suwajanakorn, S. M. Seitz, and I. Kemelmacher-Shlizerman, "Synthesizing Obama: Learning lip sync from audio," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–13, Jul. 2017.

[4] C.-C. Hsu, C.-Y. Lee, and Y.-X. Zhuang, "Learning to detect fake face images in the wild," in *Proc. Int. Symp. Comput., Consum. Control (IS3C)*, Taichung, Taiwan, Dec. 2018, pp. 388–391.

[5] S. Edwards and S. Livingston, "Fake news is about to get a lot worse. That will make it easier to violate human rights-and get away with it," *Washington Post*, Apr. 2018, [Online]. Available: https://www.washingtonpost.com/news/monkey-cage/wp/2018/04/ 03/fake-news-is-about-to-get-a-lot-worse-that-will-make-it-easier-to-violate-human-rights-and-get-away-with-it/?utm_term=.23f7e3a1be9b

[6] B. Bayar and M. C. Stamm, "A deep learning approach to universal image manipulation detection using a new convolutional layer," in *Proc. 4th ACM Workshop Inf. Hiding Multimedia Secur. (IH&MMSec)*, 2016, pp. 5–10.

[7] D. Cozzolino, G. Poggi, and L. Verdoliva, "Recasting residual-based local descriptors as convolutional neural networks: An application to image forgery detection," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur. (IHMMSec)*, 2017, pp. 159–164.

[8] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," 2019, *arXiv:1901.08971*. [Online]. Available: http://arxiv.org/abs/1901.08971

[9] D. H. Kim, S. W. Choi, and S. Y. Kwak, "Deep learning based fake face detection," *J. Korea Ind. Inf. Syst. Res.*, vol. 23, no. 5, pp. 9–17, Oct. 2018.

[10] Y. Li, M.-C. Chang, and S. Lyu, "In ictu oculi: Exposing AI generated fake face videos by detecting eye blinking," 2018, *arXiv:1806.02877*. [Online]. Available: http://arxiv.org/abs/1806.02877

[11] S. Lawrence, C. L. Giles, A. Chung Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.

[12] J. G. Lawrenson, R. Birhah, and P. J. Murphy, "Tear-film lipid layer morphology and corneal sensation in the development of blinking in neonates and infants," *J. Anatomy*, vol. 206, no. 3, pp. 265–270, Mar. 2005.

[13] A. J. Zametkin, J. R. Stevens, and R. Pittman, "Ontogeny of spontaneous blinking and of habituation of the blink reflex," *Ann. Neurol.*, vol. 5, no. 5, pp. 453–457, May 1979.

[14] P. J. De Jong and H. Merckelbach, "Eyeblink frequency, rehearsal activity, and sympathetic arousal," *Int. J. Neurosci.*, vol. 51, nos. 1–2, pp. 89–94, Jan. 1990.

[15] J. Oh and J. Jeong, "Potential significance of eyeblinks as a behavior marker of neuropsychiatric disorders," *Korean J. Biol. Psychiatry*, vol. 19, no. 1, pp. 9–20, 2012.

[16] E. Ponder and W. P. Kennedy, "On the act of blinking," *Quart. J. Exp. Physiol.*, vol. 18, no. 2, pp. 89–110, Jul. 1927.

[17] L. C. Dang, G. R. Samanez-Larkin, J. J. Castrellon, S. F. Perkins, R. L. Cowan, P. A. Newhouse, and D. H. Zald, "Spontaneous eye blink rate (EBR) is uncorrelated with dopamine D2 receptor availability and unmodulated by dopamine agonism in healthy adults," *eNeuro*, vol. 4, no. 5, pp. 1–11, Sep. 2017.

[18] G. Barbato, G. Ficca, G. Muscettola, M. Fichele, M. Beatrice, and F. Rinaldi, "Diurnal variation in spontaneous eye-blink rate," *Psychiatry Res.*, vol. 93, no. 2, pp. 145–151, Mar. 2000.

[19] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese, "Analysis of blink rate patterns in normal subjects," *Movement Disorders*, vol. 12, no. 6, pp. 1028–1034, Nov. 1997.

[20] R. Ranjan, V. M. Patel, and R. Chellappa, "HyperFace: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," 2016, *arXiv:1603.01249*. [Online]. Available: http://arxiv.org/abs/1603.01249

[21] T. Soukupová and J. Cech, "Real-time eye blink detection using facial landmarks," in *Proc. Comput. Vis. Winter Workshop (CVWW)*, 2016, pp. 42–50.

[22] D. von Cramon and U. Schuri, "Blink frequency and speech motor activity," *Neuropsychologia*, vol. 18, nos. 4–5, pp. 603–606, Jan. 1980.

[23] D. Guera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Nov. 2018, pp. 1–6.

[24] M. Koopman, A. M. Rodriguez, and Z. Geradts, "Detection of deepfake video manipulation," in *Proc. IMVIP*, Belfast, U.K., 2018, pp. 133–136.

[25] W. J. Freed, J. E. Kleinman, C. N. Karson, S. G. Potkin, D. L. Murphy, and R. J. Wyatt, "Eye-blink rates and platelet monoamine oxidase activity in chronic schizophrenic patients," *Biol. psychiatry.*, vol. 15, no. 2, pp. 329–332, 1980.

[26] Kaggle, "Eye blinking prediction," in *Proc. CompOmics Summer Competition*, 2018. [Online]. Available: https://www.kaggle.com/c/compomicssummer2018/data

[27] O. Rösler and D. Suendermann, "A first step towards eye state prediction using EEG," in *Proc. AIHLS*, Istanbul, Turkey, 2013.

[28] Robot. *Deepfake Sample*. Accessed: Dec. 4, 2019. [Online]. Available: https://www.youtube.com/watch?v=4rs0SU-BLMo&feature=youtu.be

[29] Kendrae. (2018). Google AI: Predicting heart disease in the blink of an eye. Assignment: RC TOM Challenge. Accessed: Nov. 12, 2018. [Online]. Available: https://digital.hbs.edu/platform-rctom/submission/google-ai-predicting-heart-disease-in-the-blink-of-an-eye/

[30] B. L. Sean and J. Taylor, "Forecasting at scale," *Amer. Statistician*, vol. 72, no. 1, pp. 37–45, 2017.

[31] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2016.

[32] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "MesoNet: A compact facial video forgery detection network," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2018, pp. 1–7.

[33] N. Rahmouni, V. Nozick, J. Yamagishi, and I. Echizen, "Distinguishing computer graphics from natural images using convolution neural networks," in *Proc. IEEE Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2017, pp. 1–6.

[34] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, Jun. 2012.
[35] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics: A large-scale video dataset for forgery detection in human faces," 2018, *arXiv:1803.09179*. [Online]. Available: http://arxiv.org/abs/1803.09179

TACKHYUN JUNG is currently pursuing the master's degree in IT convergence information security with the IN&S Lab, Konkuk University. Before entering Konkuk University, he was a Master Sergeant with the Republic of Korea (ROK) Air Force, and he worked on national information security for four and a half years. His research interests are cybersecurity, AI, and computer vision.

SANGWON KIM received the B.S. degree in computer science and engineering from Konkuk University, Seoul, South Korea, in 2019, where he is currently pursuing the M.S. degree in computer and information communication engineering. In 2018, he was selected for the Software Maestro Course hosted by the Ministry of Science and ICT, South Korea, and the Institute for Information and Communication Technology Planning and Evaluation. His research interest includes the convergence of network and artificial intelligence techniques in terms of resource optimization, deterministic networking, and smart security.

KEECHEON KIM received the Ph.D. degree in computer science from Northwestern University, IL, USA, in 1992. He is the Dean of the Graduate School of Information and Telecommunications. He had worked with Sprint Nextel, Inc., USA, and had worked with Shinsegei Telecom (SK Telecom) and Korea Telecom, South Korea, before he joined Konkuk University. He has been actively working in the fields of mobile data communication and network security. His current research interests are AI convergence networking, network security, the IoT, SDN, AI cryptography, and C-ITS.