symmetry

MDPI

Article

A Face Image Virtualization Mechanism for Privacy Intrusion Prevention in Healthcare Video Surveillance Systems

Jinsu Kim [1] and Namje Park [1,2,*]

[1] Department of Convergence Information Security, Graduate School, Jeju National University, Jeju City 63243, Korea; kimjinsu@jejunu.ac.kr

[2] Department of Computer Education, Teachers College, Jeju National University, Jeju City 63243, Korea

[*] Correspondence: namjepark@jejunu.ac.kr

Received: 31 March 2020; Accepted: 27 May 2020; Published: 1 June 2020

Abstract: Closed-circuit television (CCTV) and video surveillance systems (VSSs) are becoming increasingly more common each year to help prevent incidents/accidents and ensure the security of public places and facilities. The increased presence of VSS is also increasing the number of per capita exposures to CCTV cameras. To help protect the privacy of the exposed objects, attention is being drawn to technologies that utilize intelligent video surveillance systems (IVSSs). IVSSs execute a wide range of surveillance duties—from simple identification of objects in the recorded video data, to understanding and identifying the behavioral patterns of objects and the situations at the incident/accident scenes, as well as the processing of video information to protect the privacy of the recorded objects against leakage. Besides, the recorded privacy information is encrypted and recorded using blockchain technology to prevent forgery of the image. The technology herein proposed (the "proposed mechanism") is implemented to a VSS, where the mechanism converts the original visual information recorded on a VSS into a similarly constructed image information, so that the original information can be protected against leakage. The face area extracted from the image information is recorded in a separate database, allowing the creation of a restored image that is in perfect symmetry with the original image for images with virtualized face areas. Specifically, the main section of this study proposes an image modification mechanism that inserts a virtual face image that closely matches a predetermined similarity and uses a blockchain as the storage area.

Keywords: closed-circuit television; CCTV; intelligent video surveillance system; privacy; virtualization; virtual face image

1. Introduction

Closed-circuit television (CCTV) and video surveillance systems (VSSs), installed for incident/accident prevention in public places or investigation, are becoming increasingly more common each year. With their numbers on the rise, controversy over the privacy violation of objects being recorded is also rising [1–3]. Accordingly, the public's interest in intelligent visual surveillance systems (IVSSs) is growing, as the system combines technologies that allow the understanding of the scenes of incidents/accidents and the behavioral patterns of the objects, to predict incidents/accidents and warn the VSS users, and prevent the leakage of personal information [4–7]. An IVSS thus records the persons in areas targeted for recording to help understand the locations and detect the faces of the persons so that they can be used as the basis for assessment and decision-making [8–11].

Detecting a person's face and using the image, later on, can compromise the personal information of the individual that was recorded in the original video data. Hence, an IVSS implements de-identification to protect the privacy of the objects recorded on the system [12–14]. For de-identification, a certain portion(s) of the visual information is removed to prevent the guessing of the identity of a particular individual(s) or is replaced with other information to protect the personal information of the object(s). The most common cause of the implementation is masking, a type of de-identification, which inserts non-meaningful strings into the personal information-carrying portions/frames of the video and changes the identification information in the visual data [15–17]. The identification information contained in the altered visual data is safeguarded against privacy infiltration even if it is leaked to outside entities [18–20]. Security against personal information breaches in such images can be applied in a variety of fields, particularly in the medical environment, where they are often located and are continuously photographing in the same location; thus, the identification of patients is a valid technique to protect the privacy of patients [21–24]. However, statistical processing using the recorded visual data or legitimate requests for using the data does not allow the use of masked visual information. In the foregoing, an alternative is to use the original video data as the basis of generating virtual data, with the new information helping to prevent the leakage of the original visual data and allows for statistical processing using the visual information to take place [25–27]. The main part of this study follows a predetermined similarity to generate a virtual face image and inserts the image into recorded visual information so that it can identify public information with a virtualized face image. When reconstruction of the image is required, the original image and the reconstructed image can be perfectly symmetrical using the original face image recorded in the database. This paper is structured as follows: Section 2 examines the current status of techniques and studies related to the de-identification of visual data; Section 3 introduces the face image modification mechanism; and Section 4 presents the differences between the commercialized or previously introduced techniques and the proposed system.

2. Examination of the Existing Data Privacy Preservation Techniques

Intelligent video surveillance systems (IVSSs) implement various techniques to preserve the privacy of users [28–31]. The information presented in the following describe privacy protection methods that are typically used in IVSSs [32–35].

2.1. Blurring

Using weighted masks in the visual data, blurring multiplies the pixel values in the data by the weighted means/averages to obfuscate a certain portion(s) of the visual data. Typically, blurring is achieved by using the gaussian function to set the sigma parameter. Once the visual data are subject to blurring, they cannot be recovered to their original state [36–38]. This blurring technique blurs the edges, creating a clump of photos, and generally is applied to personal information about the subjects taken from the images to prevent the identification of targets by arbitrarily crushing them [39–42]. This effect becomes more blurred relative to the size of the mask applied, but as the size of the mask increases, the complexity of the calculation increases, requiring greater resource consumption. Figure 1 is a brief representation of the general logic of the blurring technique. Blurring makes a different value by applying a convention operation to a mask weighted against a given unit of zone.

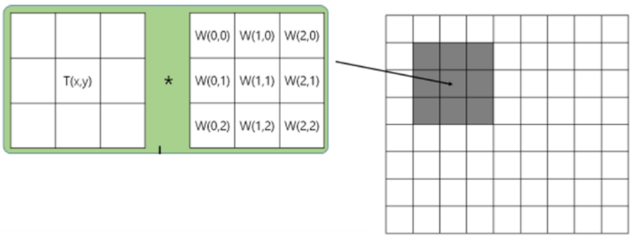


Figure 1. The blurring technique.

2.2. Mosaic Masking

With mosaic masking, the pixel values of a certain region(s) in the video data are summed and then averaged. This average is then entered into the same region(s) unilaterally so that the original visual information in the region(s) cannot be recognized by third parties. Image mosaic applies the average parameter to the entirety of the region(s), making it difficult to recover the original images [43–46]. Such mosaic techniques can be used as a means to mask the personal information existing in images; in turn, the restoration of these techniques is currently being developed by the development of artificial intelligence [47–50]. One notable example is Google's artificial intelligence-applied mosaic-replicating technique, in which photos with mosaics are restored to the point where it is difficult to identify with a third party. This methodology suggests that while accurate restoration may be difficult, it may be possible to leak personal information in that it can be added. Figure 2 is a brief description of the mosaic technique and shows that a certain area is synthesized by averaging the average value as one area.

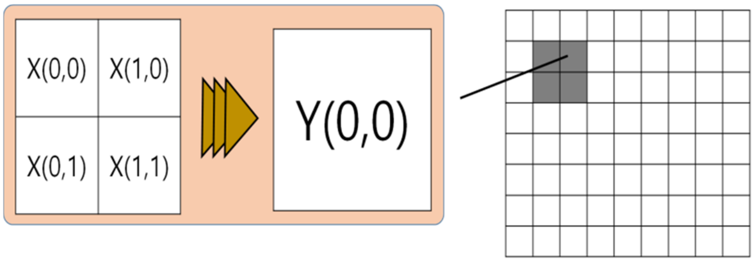


Figure 2. The image mosaic technique.

2.3. Removal and Transformation

Removal and transformation either completely remove or transform the pixels containing the identification information of the object(s) recorded in the visual data. The technique intentionally compromises the original data. Such arbitrary removal or modification of the part where the personal information of the image exists will make it impossible to restore it to the original without additional measures to restore it [51–55]. To restore the image removed from within the image, the removed image must be kept separately and the images stored separately must be identified as those of the original image [56–59]. However, the harder the video restoration process becomes, the more computing resources are required and the more maintenance costs are required. Figure 3 illustrates the application principle of the Removal and Transformation technique, which shows the process of removing or synthesizing a certain pixel region in the image, removing the existing value and changing it to a completely new value.

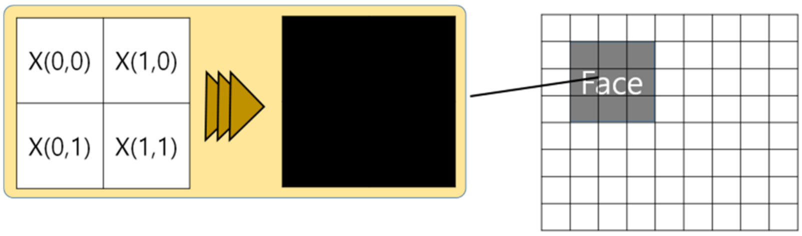


Figure 3. The removal and transformation technique.

2.4. Encryption

Encryption implements encryption keys to visual data to convert the visual information in the recorded images into cryptograms. Any authorized users with decryption keys to the coded texts can readily obtain the original visual information. Hence, additional safeguarding technologies are required to prevent decryption by unauthorized third parties [60–62]. Figure 4 shows the process of creating a cryptographic statement that can be restored but the original cannot be recognized by applying cryptographic keys to pixel values present in the image.

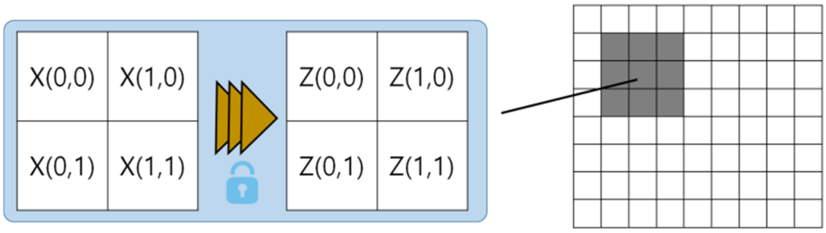


Figure 4. The encryption technique.

2.5. Related Research

Through the de-identification of images, this paper prevented leakage by third parties and further selected and synthesized virtualized face images in the images, making it difficult to identify the objects of images, while enabling the acquisition of general data, enabling digitized statistics. Thus, the main objective of the proposed mechanism was summarized and summarized for other relevant studies applying non-identification.

Uddin pointed out that recent video surveillance systems have increased significantly, generating large amounts of video data, and at the same time increasing the need for distributed processing of video data. Most of the existing solutions do not support more complex application scenarios while using the existing Client/Server framework to perform face and object recognition. It also pointed out that these frameworks were rarely processed in an extensible manner using distributed computing and that existing works do not support low-level distributed video processing APIs (Application Programming Interfaces). To address this problem, a distributed video analysis framework for intelligent video surveillance, called SIAT, was proposed, and SIAT introduced state-of-the-art distributed computing technologies to handle real-time video streams and batch video analytics to ensure scalability, efficiency and fault tolerance [63].

Brkić proposed a computer vision-based non-identification pipeline to help protect human privacy in video sequins by maintaining the naturalness and usefulness of the unidentified data and blurring faces. The proposed pipeline can be recognized when applying simple techniques, such as fade, by de-identifying features such as clothing, hair and skin color. Assuming a surveillance scenario, we combined background subtraction based on Gaussian blend with an improved version of the GrabCut algorithm to find and classify pedestrians and to change the appearance of segment pedestrians through neural technology algorithms that render pedestrian images in different styles using responses from deep neural networks [64].

Gros pointed out that the development of camera and computing equipment hardware has made extensive video data capture and storage simpler, and provides ample opportunity to share video sequins, pointing out the need for automated ways of de-identifying images, especially facial areas, to protect the privacy of subjects shown in the video. To solve this problem, a wide range of experiments have demonstrated that pixelation and blurring are very poor in privacy protection, while greatly distorting data; the Google Algorithm has proposed a new technique by combining model-based face image parameterization with official privacy models [65].

3. Proposed Mechanism for Changing Face Information to Prevent Privacy Infiltration

The mechanism proposed in this section of the study (hereinafter, the “proposed mechanism”) substitutes a face image(s) contained in the recorded visual data with a virtual face image(s) stored in the existing database (DB). Hence, the foregoing is named “virtualization”, and the face image information stored in the DB is abbreviated as “virtual face information” for this study. The proposed mechanism 1) detects face information contained in the recorded visual data; 2) generates virtual face information by using virtualization random numbers; and 3) then replaces the face information in the visual data with the virtual face information with comparable similarities to only allow the identification of sex, age, race and such other public information. Such published information may then be used in statistics of the population that has moved to that location. Virtual face image data can later be recovered using the original data by authorized users. The proposed mechanism comprises a face region detection module, a virtual face features generation module, a virtual face feature vector generation module, a virtual face image data generation module, a face features recovery module, a recovered face feature vector generation module and a face image data recovery module. Figure 5 lists the sequences in which the mechanism is implemented. Table 1 shows the abbreviations used in the proposed mechanism.

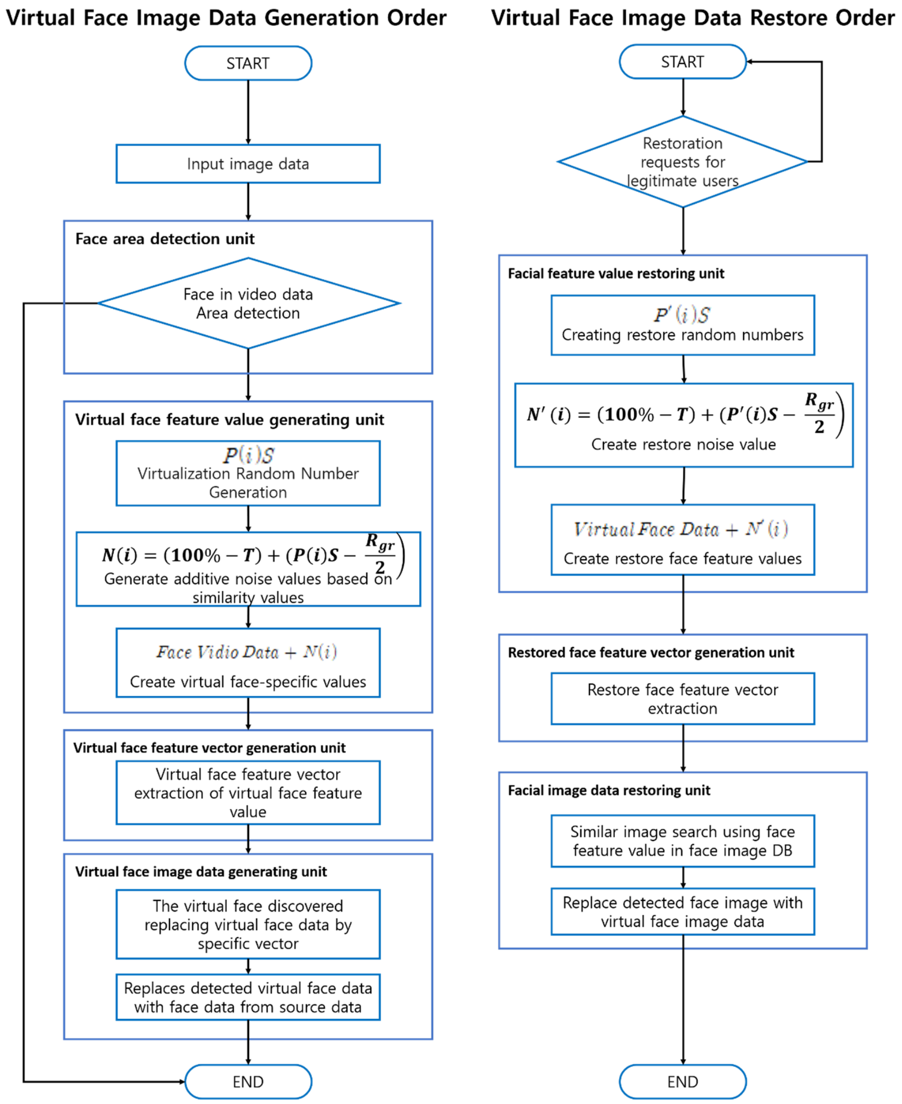

Virtual Face Image Data Generation Order
START
Input image data
Face area detection unit
Face in video data Area detection
Virtual face feature value generating unit
$P(i)S$
Virtualization Random Number Generation
$N(i) = (100\% - T) + (P(i)S - \frac{R_{gr}}{2})$
Generate additive noise values based on similarity values
Face Vidio Data + $N(i)$
Create virtual face-specific values
Virtual face feature vector generation unit
Virtual face feature vector extraction of virtual face feature value
Virtual face image data generating unit
The virtual face discovered replacing virtual face data by specific vector
Replaces detected virtual face data with face data from source data
END
Virtual Face Image Data Restore Order
START
Restoration requests for legitimate users
Facial feature value restoring unit
$P'(i)S$
Creating restore random numbers
$N'(i) = (100\% - T) + (P'(i)S - \frac{R_{gr}}{2})$
Create restore noise value
Virtual Face Data + $N'(i)$
Create restore face feature values
Restored face feature vector generation unit
Restore face feature vector extraction
Facial image data restoring unit
Similar image search using face feature value in face image DB
Replace detected face image with virtual face image data
END

Figure 5. Flowchart showing the proposed mechanism for generating and recovering virtual faces.

Table 1. Abbreviations.

Abbreviation	Content
P	Random number generation function
(i)	Random number generation circuit
S	Seed value defined by user
N	Face virtualization noise
R_{gr}	Random number generation range

3.1. Face Region Detection Module

This module detects a face region(s) found in the video data recorded by more than one camera. In response to the sequences in which more than one set of visual data are recorded, the module allocates an order or sequence of recording image data sets (more than one) according to the cameras used. In the case of no face regions detected in the visual data, the module may not allocate the recording order or sequence.

3.2. Virtual Face Features Generation Module

This module generates virtualization of random numbers for each image data set (more than one), which corresponds to the order/sequence of recording allocated by the face region detection module, and then uses the random numbers to convert the face features (parameters) that indicate the feature coordinates capable of identifying the person with the face information extracted from the face region into the virtual face feature parameters, by adding noise values based on the virtualization random numbers. As the initial values for the random number generation function, the virtual face features generation module uses pre-determined seed values.

$$P(i)S \tag{1}$$

Formula (1) is an expression of virtualization random number generation function that aims at generating virtual faces. "*P*" refers to the random number generation function; "*i*" is the order/sequence of issuing the random number generation circuit; and "*S*" is the pre-defined seed value. Using the formula above to generate virtualization random numbers according to the generation orders/sequences, one can use the random number generation function, where the pre-determined seed value and virtualization random number generation range are the same, to generate the random numbers in the generation sequences. This results in the same virtualization random numbers.

If an example is shown for Formula (1), it is assumed that the random number generation range is in the range of 0 to 100, and the generation cycle is two times. It is also assumed that the first and second rounds have an "S" of 0.2 and 0.5, respectively. In the mechanism, the first round randomly generates random numbers in the range of 0 to 100, and multiplies the number by a fixed seed so that random numbers can be generated within the range intended by the user. Accordingly, the random number generated in the first round will be generated within the range desired by the user by multiplying the fixed seed value by 0.2, and the range will be limited to 0 to 20. Thus, in the second round, it will range from 0 to 50.

The virtualization random numbers so obtained will be used to generate noise values by adding the differences between the median values of the pre-defined virtualization random number generation ranges, to the pre-determined similarity values.

$$N(i) = (100\% - T) + \left(P(i)S - \frac{R_{gr}}{2}\right) \tag{2}$$

Formula (2) calculates the noise values for generating virtual face features (parameters). Here, "$N(i)$" refers to the noise value for generating a virtual face; "*T*" the pre-defined similarity value; "$P(i)S$" the virtualization random numbers; and "R_{gr}" the pre-set virtualization random number generation range.

$$T = 50\%,\ \mathrm{Rgr} = 50 \tag{3}$$

$$P(1)S = 15,\ P(2)S = 40,\ P(3)S = 25 \tag{4}$$

Formula (3) with a 50% similarity and a 50 virtualization random number generation range is an example of how to generate the ranges. Calculating noise values to the three recordings will result in three virtualization random numbers generated within the range of 50, according to Formula (1). In this section, Formula (4) is adopted to arbitrarily set the random numbers at 15, 40 and 25.

$$N(1) = (100\% - 50\%) + \left(15 - \frac{50}{2}\right) = 40 \tag{5}$$

$$N(2) = (100\% - 50\%) + \left(40 - \frac{50}{2}\right) = 65 \tag{6}$$

$$N(3) = (100\% - 50\%) + \left(25 - \frac{50}{2}\right) = 50 \tag{7}$$

Formulas (5)–(7) show the process in which noise values are calculated according to the order/sequence of video recording. As shown, the noise values are in inverse proportion to the pre-defined similarity values—with an increased similarity value resulting in a closer rendition of the original data. In reverse, a lower similarity value will lead to the opposite rendition of the original.

3.3. Virtual Face Feature Vector and Data Generation Module

This module extracts virtual face feature vectors from the virtual face feature parameters, i.e., two-dimensional array-based coordinates. The virtual face image data generation module searches the virtual face information based on the virtual face feature vectors that are obtained from the vectorized virtual face feature parameters converted by the virtual face features generation module, as well as with the use of principal component analysis (PCA), which analyzes the inter-vector similarities using a covariance matrix in the face information DB; linear discriminant analysis (LDA), which maximizes the proportions of between-class distributions and within-class distributions; and such other techniques. Then the virtual face image data generation module inserts the virtual face information (search result) to substitute the face region in the video data and generate the virtual face image data. Even if the inserted virtual face information has the same level of similarity as the original face information, each recording may insert a different version of virtual face information. In this process, the real face data extracted from the original image is encrypted and recorded ona blockchain, and by implementing a blockchain in a private environment, it minimizes external access and prevents forgery. The images to be synthesized generate an identification key for the image and include the identification key of the image for the additional extracted face area so that if the restoration of the image is required, the images matched using the identifier for the image can be applied and restored.

3.4. Face Features Recovery Module

This module adds the recovery random number SDs to the pre-determined similarity values to calculate back the noise values that were used to generate the virtual face feature parameters. Then model deducts the counter-calculated noise values to generate the recovered face feature parameters that were restored from the face feature parameters. The pre-defined seed values are used to calculate the initial values for the recovery random number generation function.

$$P'(i)S \tag{8}$$

Formula (8) is an expression for the recovery random number generation function, which is to recover the virtual face. Here, "*P*" refers to the random number generation function; "*i*" is the issue order/sequence of the random number generation function; and "*S*" the pre-defined seed values. This formula can calculate the noise values by adding the similarity values that were used during the generation of virtual image data, to the recovery random number SDs. Formula (8) takes advantage of the value of $P(i)$ in the random number function $P(i)S$, which is carried out for face virtualization. This allows the restore the random number to have the same value as the random number generated in Formula (1). The formula can also generate the recovered face feature parameters by adding the virtual face feature parameters to the noise values.

$$N'\ (i) = (100\% - T) + \left(P'(i)S - \frac{R_{gr}}{2}\right) \tag{9}$$

Formula (9) is a noise calculation expression that aims to counter-calculate the virtual face information. "$N'(i)$" refers to the noise value for recovering virtual faces, and "T" is the pre-determined similarity value. "$P'(i)S$" indicates the recovery random number, and "R_{gr}" the pre-set virtualization random number generation range.

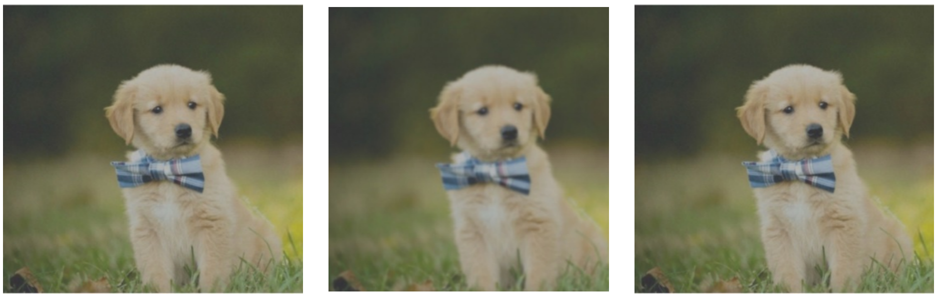
3.5. Face Image Restore Module

This module vectorizes the recovered face feature parameters that were generated from the face features recovery module, and uses the vectorized values to generate the recovered face feature vectors. After this, based on the recovered face feature vectors (i.e., a result of the vectorization of recovered face feature parameters), this module searches the recovered face information in the face information DB, and then inserts the recovered face information (search result) into the face regions in the virtual face image data to generate the recovered image data. The information necessary for the recovery will be offered only to the authorized administrators, hence the recovery by unauthorized third parties is prevented.

4. Comparison and Analysis of the Proposed Mechanism and the Existing Techniques

The proposed mechanism detects one or more pieces of face information contained in the recorded image data and replaces the face information with virtual face information so that the objects recorded in the video cannot be specified. The face information of the object(s) recorded is stored in the face information DB. When an authorized user requests access to the information, later on, the image information that has been replaced with virtual face information can be recovered to its original state and sent to the requesting user. The proposed mechanism (face image modification) was compared with four other techniques currently in use, namely, blurring, image mosaic, removal and transformation, and encryption:

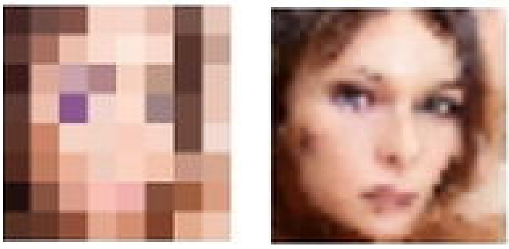
Blurring is a technique that applies pixel values of the visual data to weighted masks to obfuscate the image information. The technique allows recovery via deblurring but not the identification of the public information regarding the targeted object (2). Figure 6 lists the original image, blurring-applied image, and recovered image.



(**A**) Original (**B**) Blurring (**C**) Restored

Figure 6. Sequential change of images using blurring.

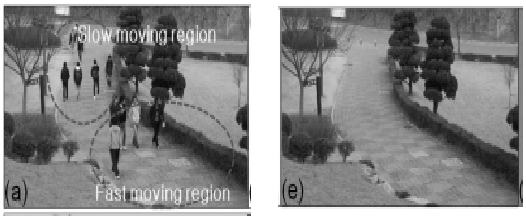
Image mosaic sums the pixel values in a particular region(s) in the video data that includes face information upon which de-identification is requested. The technique then obtains the average/mean from the sum and applies it unilaterally to all pixels in the region(s) so the target object(s) cannot be recognized. In the foregoing, the public information of the target(s) is unrecognizable; however, it is being challenged by, for example, a recent Google publication reported on the development and use of an artificial intelligence (AI)-enabled technology to guess images similar to those in the original data and recover them. Figure 7 shows images of mosaic techniques and restored images. In general, it is impossible to restore the mosaic technique if it has been seriously changed, but recently, the restoration technique using artificial intelligence has been unveiled, showing the possibility of restoration, although it is not perfect.



(**A**) Mosaic (original) (**B**) Restored

Figure 7. Mosaic-applied and recovered images.

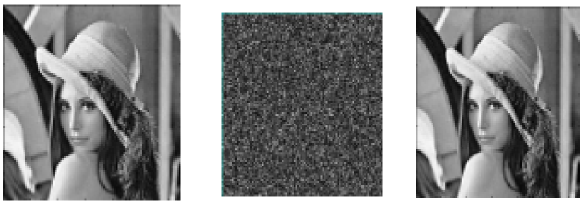
Removal and transformation refers to a method that renders the face information unrecognizable by removing or altering the pixel values for a face region(s) contained in the visual data. With the alteration of the visual data, identifying the public information regarding the target(s) is infeasible. Recovering the data is also difficult due to the deletion of the data contained in the visual information. Figure 8 shows the images prior to the application of Removal and Transformation and the images applied. Although it is possible to change the identification information to a completely different image by composing it with another image, the problem is that restoration is impossible because the image value is completely changed.



(**A**) Original (**B**) Removed

Figure 8. Removed and recovered image.

Lastly, encryption uses encryption keys on visual data to convert them into ciphertexts. The original data are offered only after authorized users decrypt the visual data. Depending on the type of encryption, techniques vary—from converting the original information and obfuscating it for preventing identification to modifying visual data units/frames to render visual data identification itself infeasible. These techniques, however, allow recovery, in which case the public information regarding the objects can be recognized. Access to the recovered information or to the decryption keys is granted only to authorized users. Figure 9 shows the encryption process for images. (A) shows the original image, (B) shows the result of encrypting the image, and (C) is the restored image. In the case of such encryption techniques, the original is unidentifiable so as not to be leaked to the outside world, and images can be restored by legitimate users.



(**A**) Original (**B**) Encryption (**C**) Restored

Figure 9. Encrypted and recovered images.

The process of conducting a similarity test on subjects will be carried out through a confrontation with virtual facial information stored in the DB for virtualization. At this time, each virtual face has a vector value for the face, and in the proposed mechanism, it is replaced by a virtual face with the most similar value by comparing the generated noise value with the vector value based on the extracted face. Figure 10 shows the vector value extracted through the face, the resulting random value, the generated random value and the name of the face with the most similar figure.

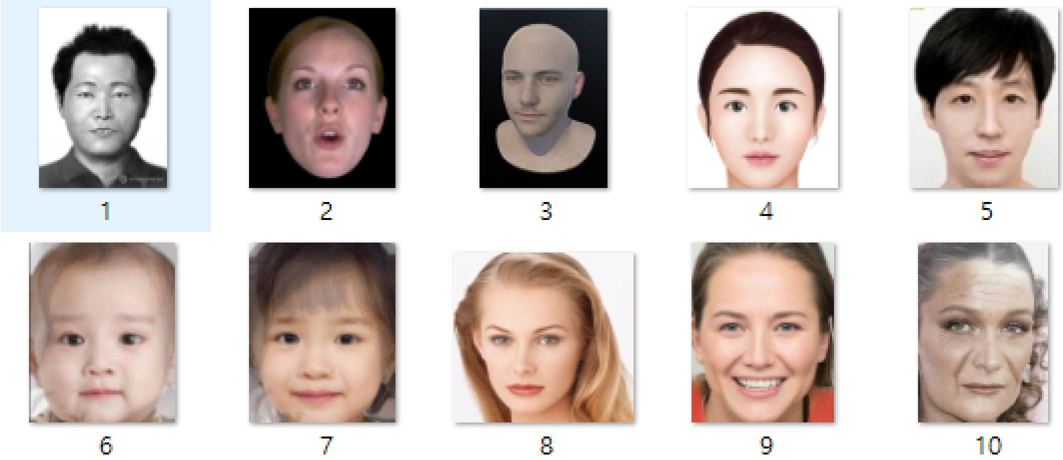
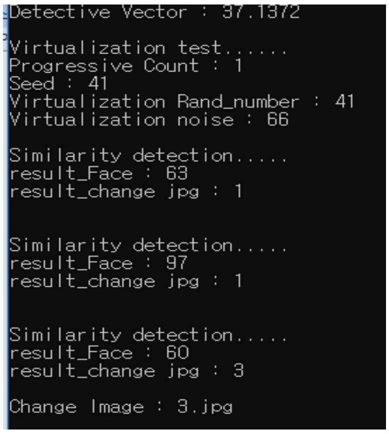


Figure 10. Virtual face list.

Figure 11 shows the process of extracting the most similar face by comparing vector values for each face image included in the Virtualized Face DB with vector values for the extracted face image for the face extracted from the image. In the proposed mechanism, the region for the face is first extracted and the vector value is extracted. Face data in the virtualization area have their own vector values and only generate vector values at the beginning of the first occurrence, and then use only vector values as a comparison target to reduce the image processing process, minimizing resource consumption for image processing.



```
Detective Vector : 37.1372

Virtualization test......
Progressive Count : 1
Seed : 41
Virtualization Rand_number : 41
Virtualization noise : 66

Similarity detection.....
result_Face : 63
result_change jpg : 1


Similarity detection.....
result_Face : 97
result_change jpg : 1


Similarity detection.....
result_Face : 60
result_change jpg : 3

Change Image : 3.jpg
```

Figure 11. Computational result.

The proposed mechanism substitutes the face information contained in the visual data with virtual faces. Hence, the approach allows the identification of visual data as well as the distinction of the basic public information regarding objects. An authorized user can have the virtual face image data recovered to its original state. The face information of the objects is stored in the face information DB, which offers the merit of stronger security in that the original face information is difficult to be identified from the converted virtual face image information even when both the face information DB and the virtual face image data are leaked.

Figure 12 is computed as shown in Figure 11 in the virtual face list, replacing the face with the original face with the most similar noise value. As such, replacing a face image with a value that is most similar to the noise value extracted on a vector basis can prevent a subject from leaking into the face image and provide basic information for identifying gender, etc.

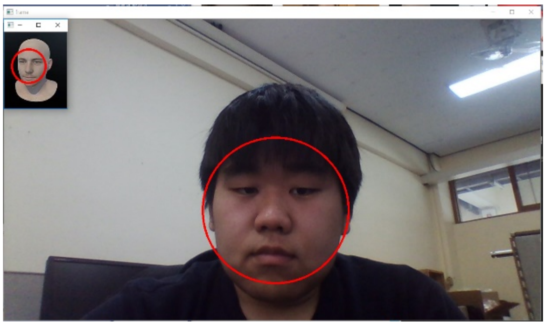
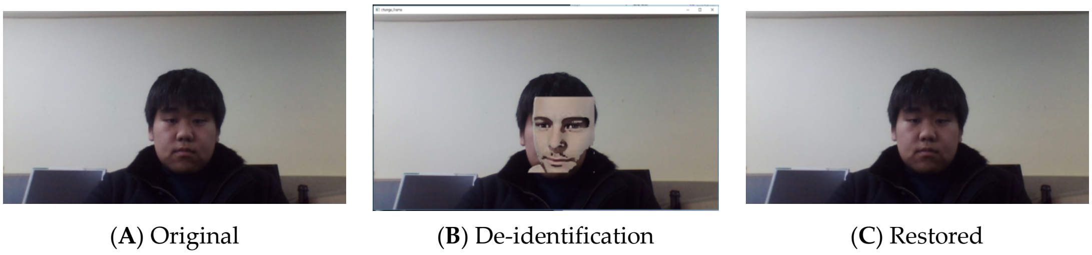


Figure 12. Computational result.

Figure 13 is the result of the proposed mechanism, and initially, the face identified as (B) is replaced with the virtual face by using the original image information that has not undergone a separate non-identification process as shown in (A). The information is recorded in a separate database, and when the original image is required, it is displayed as the original image as shown in (C) to show the face recognition process.



(**A**) Original (**B**) De-identification (**C**) Restored

Figure 13. Comparison of images from the proposed mechanism.

The experiment was conducted using the following CPU: Intel i3-6100 3.70GHz; random access memory (RAM): 4 GB environment. To verify the effectiveness of the mechanism proposed in the research process, the procedures required in the face virtualization phase and the time required for image virtualization were analyzed.

First, the procedures of face virtualization required four stages: face area detection, vector operation in the detection area, face vector-based image selection, image composition, face area detection, vector operation in the detection area, virtualization vector-based image extraction and image reconstruction, which required a total of eight steps. While traditional image privacy techniques proceed in three stages—face extraction, image processing algorithm operation and the resultant image production—it can be found that relatively many processes are required. The more processing required by the required items compared to the general images require more processes by adding face virtualization steps for the de-identification of the images in the existing procedures of face extraction, image processing and encrypted image generation. Secondly, we analyzed the time required to virtualize images, and in this process, only one person who was able to detect the face area was using the mp4 extension for the experiment. For the equality of time measurements, the process measured the time for frame extraction in images, face area detection, face virtualization and image production using only extracted frames, which typically took three times the time of the images taken. The images used for the experiment were 8 s long and took about 20 s to virtualize the face within the image and produce it. Time beyond the actual image was difficult to graft on the site, but it was conducted after a bitmap reinterpretation of the extracted images in the experimental environment, and the study of a simpler way to process this process is likely to reduce the required time, and it is assumed that the performance of the devices that virtualize the images will also be significantly different. Table 2 shows a comparative analysis of the existing non-identification techniques and the proposed mechanism.

Table 2. Comparison and analysis of the proposed mechanism and the existing techniques.

Comparison Item	Blurring	Mosaic	Removal and Transformation	Encryption	The Proposed Mechanism
De-identification	O	O	O	O	O
Unable to restore an image by the illegal user	X	X	O	X	O
Image reconstruction by legitimate users	O	X	X	O	O
Identification of de-identified public information	X	X	X	X	O

5. Conclusions

With increases in the number of video surveillance systems (VSSs), the number of per capita exposures to CCTV cameras is also increasing. Moreover, the ongoing advancement in technology is upgrading VSSs to the extent that they can recognize a particular individual or object or the actions of a moving object from the recordings and can understand the situations by sharing recorded information from and between the dispersed systems. Such capabilities are making VSS-enabled control easier. A concern over the violation of privacy of the objects being exposed to VSSs is nonetheless growing. In addition, the act of compromising the integrity of the system through image forgery of VSS causes a big problem. Images whose integrity is compromised cannot prove the justification of the images, and the surveillance of the place that is the purpose of the VSS may not produce sufficient results.

The main section of this study examined the differences between the existing data privacy preservation techniques and the mechanism herein proposed (the "proposed mechanism"). By recording the original image of the virtualized face on a blockchain, it is possible to prove the integrity of the image by lowering the possibility of face forgery. The proposed mechanism is capable of complementing problems shown in the existing data privacy protection methods. The unique feature of the mechanism includes its ability to detect face regions in the recorded visual information and locate the specific vectors in the detected face regions and substitute them with similarly constructed virtual faces. The identification information in the substituted visual data can be changed by adjusting similarity.

The market for VSSs is showing an upward trend each year, and the number of VSS control centers is also showing growth. These growths, however, are not being accompanied by the implementation of data privacy preservation techniques. Even in the cases where such techniques are applied, the original data cannot be inferred, depending on the methods used, or are susceptible to leakage via encryption key exposure. The violation of privacy due to video data leakage can be a serious issue in the coming years. As such, it is imperative that preventive methods be studied.

Author Contributions: J.K. implemented a methodology for interpreting the identification data of images and virtualizing the identification data. J.K. is currently working as a researcher at the Cybersecurity Human Resource Institute, Jeju National University. His current research interest includes Cloud System, Intelligent Video Surveillance System and Internet of Things.; N.P. contributed greatly to the conceptualization of methodology and overall project management and preparation of manuscripts. All authors have read and agreed to the published version of the manuscript. N.P. is currently working as a Professor in the Department of Computer Education, Teachers College, Jeju National University. His current research interest includes Convergence Technology Security, Computer Education, Smart Grid, IoT, Sea Cloud.

Funding: This work was supported by the Institute for Information & communications Technology Promotion (IITP) grant, funded by the Korea government (MSIT) (2019-0-00203, The Development of Predictive Visual Security Technology for Preemptive Threat Response). This work was also supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5C2A04083374).

Acknowledgments: Thanks to the editors and reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yuxi, P.; Luuk, J.S.; Raymond, N.J.V. Low-resolution face recognition and the importance of proper alignment. *IET Biom.* **2019**, *8*, 267–276.
2. Ling, J.; Zhang, K.; Zhang, Y.; Yang, D.; Chen, Z. A saliency prediction model on 360 degree images using color dictionary based sparse representation. *Signal Process. Image Commun.* **2018**, *69*, 60–68. [CrossRef]
3. Kim, J.; Park, N.; Kim, G.; Jin, S. CCTV Video Processing Metadata Security Scheme Using Character Order Preserving-Transformation in the Emerging Multimedia. *Electronics* **2019**, *8*, 412. [CrossRef]
4. Creasey, M.S.; Albalooshi, F.A.; Rajarajan, M. Continuous face authentication scheme for mobile devices with tracking and liveness detection. *Microprocess. Microsyst.* **2018**, *63*, 147–157. [CrossRef]
5. Chhokra, P.; Chowdhury, A.; Goswami, G.; Vatsa, M.; Singh, R. Unconstrained Kinect video face database. *Inf. Fusion* **2018**, *44*, 113–125. [CrossRef]
6. Kim, N.; Seo, D.; Lee, S.; Shin, C.; Cho, K.; Kim, S. Hierarchical image encryption with orthogonality. *J. KJOP* **2006**, *17*, 231–239.
7. Park, N. The Core Competencies of SEL-based Innovative Creativity Education. *Int. J. Pure Appl. Math.* **2018**, *118*, 837–849.
8. Arceda, V.E.M.; Fabián, K.M.F.; Laura, P.C.L.; Tito, J.J.R.; Cáceres, J.C.G. Fingertip Detection and Tracking for Recognition of Air-Writing in Videos. *Exp. Syst. Appl.* **2019**, *136*, 217–229.
9. Mutneja, V.; Singh, S. GPU accelerated face detection from low resolution surveillance videos using motion and skin color segmentation. *Optik* **2018**, *157*, 1155–1165. [CrossRef]
10. Rashedi, E.; Barati, E.; Nokleby, M.; Chen, X. "Stream loss": ConvNet learning for face verification using unlabeled videos in the wild. *Neurocomputing* **2019**, *329*, 311–319. [CrossRef]
11. Park, N.; Hu, H.; Jin, Q. Security and Privacy Mechanisms for Sensor Middleware and Application in Internet of Things (IoT). *Int. J. Distrib. Sens. Networks* **2016**, *12*, 2965438. [CrossRef]
12. Girod, B. The Information Theoretical Significance of Spatial and Temporal Masking in Video Signals. *OE/LASE '89* **1989**, *1077*, 178–189. [CrossRef]
13. Lee, K.; Yeuk, H.; Kim, J.; Hyungjoon Park, K.Y. An efficient key management solution for privacy masking, restoring and user authentication for video surveillance servers. *Comput. Stand. Interfaces* **2016**, *44*, 137–143. [CrossRef]
14. Milosavljević, A.; Rančić, D.; Dimitrijević, A.; Predić, B.; Mihajlović, V. Integration of GIS and video surveillance. *Int. J. Geogr. Inf. Sci.* **2016**, 1–19. [CrossRef]
15. Park, N.; Kim, M. Implementation of load management application system using smart grid privacy policy in energy management service environment. *Clust. Comput.* **2014**, *17*, 653–664. [CrossRef]
16. Wang, L.; Yu, X.; Bourlai, T.; Metaxas, N.D. A coupled encoder–decoder network for joint face detection and landmark localization. *Image Vis. Comput.* **2019**, *87*, 37–46. [CrossRef]
17. Hagmann, J. Security in the Society of Control: The Politics and Practices of Securing Urban Spaces. *Int. Political Sociol.* **2017**, *11*, 418–438. [CrossRef]
18. Essa, A.; Asari, V.K. Face recognition based on modular histogram of oriented directional features. In Proceedings of the 2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), Dayton, OH, USA, 25–29 July 2016.
19. Kose, N.; Dugelay, J.L. Countermeasure for the protection of face recognition systems against mask attacks. In Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, China, 22–26 April 2013.
20. Kim, H.A. Data Blurring Method for Dollaborative Filtering. Ph.D. Thesis, Department of Computer Engineering, The Graduate School of Dongguk University, Seoul, Korea, 2004.
21. Wang, Z.; Yang, X.; Cheng, K.T. Accurate face alignment and adaptive patch selection for heart rate estimation from videos under realistic scenarios. *PLoS ONE* **2018**, *13*, e0197275. [CrossRef] [PubMed]
22. Dimoulas, C.; Papanikolaou, G.; Petridis, V. Pattern Classification and Audiovisual Content Management techniques using Hybrid Expert Systems: A video-assisted Bioacoustics Application in Abdominal Sounds Pattern Analysis. *Exp. Syst. Appl.* **2011**, *38*, 13082–13093. [CrossRef]
23. Dimoulas, C.; Avdelidis, A.; Kalliris, G.; Papanikolaou, G. Joint Wavelet Video Denoising and Motion Activity Detection in multi-modal human activity analysis: Application to video—Assisted bioacoustic/psycho-physiological monitoring. *EURASIP J. Adv. Signal Process.* **2007**, *2008*, 1–19. [CrossRef]

24. Peng, Y.; Lin, C.; Sun, M.; Landis, C.A. Multimodality Sensor System for Long-Term Sleep Quality Monitoring. *IEEE Trans. Biomed. Circuits Syst.* **2007**, *1*, 217–227. [CrossRef] [PubMed]
25. Park, N.; Lee, D. Electronic identity information hiding methods using a secret sharing scheme in multimedia-centric internet of things environment. *Person. Ubiquitous Comput.* **2018**, *22*, 3–10. [CrossRef]
26. Jeon, B.; Shin, S.; Jung, K.; Lee, J.; Yoo, K. Reversible Secret Sharing Scheme Using Symmetric Key Encryption Algorithm in Encrypted Image. *J. KMS* **2015**, *18*, 1332–1341.
27. Liu, S.; Yu, M.; Li, M.; Xu, Q. The research of virtual face based on Deep Convolutional Generative Adversarial Networks using TensorFlow. *J. Phys. A SMA* **2019**, *521*, 667–680. [CrossRef]
28. Huh, J. PLC-based design of monitoring system for ICT-integrated vertical fish farm. *Hum.-Cent. Comput. Inf. Sci.* **2017**, *7*, 7. [CrossRef]
29. Yan, J.; Zhang, X.; Lei, Z.; Li, S.Z. Face detection by structural models, Image annotation: Then and now. *Image Vis. Comput.* **2014**, *32*, 790–799. [CrossRef]
30. Lee, D.; Park, N.; Kim, G.; Jin, S. De-identification of metering data for smart grid personal security in intelligent CCTV-based P2P cloud computing environment. *J. Peer-to-Peer Network. Appl.* **2018**, *11*, 1299–1308. [CrossRef]
31. Bhagat, P.K.; Choudhary, P. Image annotation: Then and now. *Image Vis. Comput.* **2018**, *80*, 1–23. [CrossRef]
32. Lee, D.; Park, N. ROI-based efficient video data processing for large-scale cloud storage in intelligent CCTV environment. *J. IJET* **2018**, *7*, 151–154.
33. Cucchiara, R.; Grana, C.; Prati, A.; Vezzani, R. Computer vision system for in-house video surveillance. *IEE Proc. Vis. Image Signal Process.* **2005**, *152*, 242–249. [CrossRef]
34. Park, N.; Kang, N. Mutual Authentication Scheme in Secure Internet of Things Technology for Comfortable Lifestyle. *Sensors* **2016**, *16*, 20. [CrossRef] [PubMed]
35. Azhar, H.; Amer, A. Classification of surveillance video objects using chaotic series. *IET Image Process.* **2012**, *6*, 919–931. [CrossRef]
36. Zhang, L.; Dou, P.; Kakadiaris, I.A. Patch-based face recognition using a hierarchical multi-label matcher, Image annotation: Then and now. *Image Vis. Comput.* **2018**, *73*, 28–39. [CrossRef]
37. Hu, J. Discriminative transfer learning with sparsity regularization for single-sample face recognition, Image annotation: Then and now. *Image Vis. Comput.* **2017**, *60*, 48–57. [CrossRef]
38. Teferi, D.; Bigun, J. Damascening video databases for evaluation of face tracking and recognition—The DXM2VTS database. *Pattern Recognit. Lett.* **2007**, *28*, 2143–2156. [CrossRef]
39. Wang, C.; Li, Y. Combine image quality fusion and illumination compensation for video-based face recognition. *Neurocomputing* **2010**, *73*, 1478–1490. [CrossRef]
40. Li, H.; Achim, A.; Bull, D. Unsupervised video anomaly detection using feature clustering. *IET Signal Process.* **2012**, *6*, 521–533. [CrossRef]
41. Leung, V.; Colombo, A.; Orwell, J.; Velastin, S.A. Modelling periodic scene elements for visual surveillance. *IET Comput. Vis.* **2001**, *37*, 20–21.
42. Loideain, N.N. Cape Town as a smart and safe city: Implications for governance and data privacy. *Int. Data Priv. Law* **2017**, *7*, 314–334. [CrossRef]
43. Lee, D.; Park, N. Geocasting-based synchronization of Almanac on the maritime cloud for distributed smart surveillance. *J. Supercomput.* **2017**, *73*, 1103–1118. [CrossRef]
44. Ayesha, C.; Santanu, C. Video analytics revisited. *IET Comput. Vis.* **2016**, *10*, 237–249.
45. Pons, J.; Prades-Nebot, J.; Albiol, A.; Molina, J. Fast motion detection in compressed domain for video surveillance. *Electron. Lett.* **2002**, *38*, 409–411. [CrossRef]
46. Park, N.; Bang, H. Mobile middleware platform for secure vessel traffic system in IoT service environment. *Secur. Commun. Netw.* **2016**, *9*, 500–512. [CrossRef]
47. Jeong, M.; Jeong, J. Uniform Motion Deblurring using Shock Filter and Convolutional Neural Network. *J. KSBE* **2018**, *23*, 484–494.
48. Park, N.; Kwak, J.; Kim, S.; Won, D.; Kim, H. WIPI Mobile Platform with Secure Service for Mobile RFID Network Environment, Advanced Web and Network Technologies, and Applications. *LNCS* **2006**, *3842*, 741–748.
49. Dahl, R.; Norouzi, M.; Shlens, J. Pixel Recursive Super Resolution. *J. ICCV* **2017**, *7*, 7. [CrossRef]

50. Raghavendra, R.; Busch, C. Novel presentation attack detection algorithm for face recognition system: Application to 3D face mask attack. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014.
51. Guo, J.M.; Lin, C.C.; Wu, M.F.; Chang, C.H.; Lee, H. Complexity Reduced Face Detection Using Probability-Based Face Mask Prefiltering and Pixel-Based Hierarchical-Feature Adaboosting. *IEEE Signal Process. Lett.* **2011**, *18*, 447–450. [CrossRef]
52. Niu, G.; Chen, Q. Learning an video frame-based face detection system for security fields. *J. Vis. Commun. Image Represent.* **2018**, *55*, 457–463. [CrossRef]
53. Du, L.; Zhang, W.; Fu, H.; Ren, W.; Zhang, X. An efficient privacy protection scheme for data security in video surveillance. *Vis. Commun. Image Represent.* **2019**, *59*, 347–362. [CrossRef]
54. Hu, X.; Peng, S.; Wang, L.; Yang, Z.; Li, Z. Surveillance video face recognition with single sample per person based on 3D modeling and blurring. *Neurocomputing* **2017**, *235*, 46–58. [CrossRef]
55. Hu, M.; Wang, H.; Wang, X.; Yang, J.; Wang, R. Video facial emotion recognition based on local enhanced motion history image and CNN-CTSLSTM networks. *Vis. Commun. Image Represent.* **2019**, *59*, 176–185. [CrossRef]
56. Hsu, C.Y.; Wang, H.F.; Wang, H.C.; Tseng, K.K. Automatic extraction of face contours in images and videos. *Future Gener. Comput. Syst.* **2012**, *28*, 322–335. [CrossRef]
57. Cui, D.; Zhang, G.; Hu, K.; Han, W.; Huang, G.B. Face recognition using total loss function on face database with ID photos. *Opt. Laser Technol.* **2019**, *110*, 227–233. [CrossRef]
58. Huh, J.H.; Seo, K. Smart Grid Test Bed Using OPNET and Power Line Communication. In Proceedings of the 2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS), Sapporo, Japan, 25–28 August 2016.
59. Naveen, S.; Fathima, R.S.; Moni, R.S. Face recognition and authentication using LBP and BSIF mask detection and elimination. In Proceedings of the 2016 International Conference on Communication Systems and Networks (ComNet), Ahmedabad, India, 19–20 February 2016.
60. Cakiroglu, O.; Ozer, C.; Gunsel, B. Design of a Deep Face Detector by Mask R-CNN. In Proceedings of the 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24–26 Nisan 2019.
61. Ramkumar, G.; Logashanmugam, E. An effectual face tracking based on transformed algorithm using composite mask. In Proceedings of the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, India, 15–17 December 2016.
62. Kim, D.; Park, S. A Study on Face Masking Scheme in Video Surveillance System. In Proceedings of the 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN), Prague, Czech Republic, 3–6 July 2018.
63. Uddin, M.; Alam, A.; Tu, N.; Islam, M.; Lee, Y. SIAT: A Distributed Video Analytics Framework for Intelligent Video Surveillance. *Symmetry* **2019**, *11*, 911. [CrossRef]
64. Brkić, K.; Hrkać, T.; Kalafatić, Z. Protecting the privacy of humans in video sequences using a computer vision-based de-identification pipeline. *Exp. Syst. Appl.* **2017**, *87*, 41–55. [CrossRef]
65. Gross, R.; Sweeney, L.; de la Torre, F.; Baker, S. Model-Based Face De-Identification. In Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), New York, NY, USA, 17–22 June 2006.