electronics

MDPI

Article

Verifying the Effectiveness of New Face Spoofing DB with Capture Angle and Distance

Jin Yeong Bok [1], Kun Ha Suh [1] and Eui Chul Lee [2,*]

[1] Department of Computer Science, Graduate School, Sangmyung University, Hongjimun 2-Gil 20, Jongno-Gu, Seoul 03016, Korea; bjywlsdud23@gmail.com (J.Y.B.); tjrjsgk@naver.com (K.H.S.)
[2] Department of Human-Centered Artificial Intelligence, Sangmyung University, Hongjimun 2-Gil 20, Jongno-Gu, Seoul 03016, Korea
* Correspondence: eclee@smu.ac.kr; Tel.: +82-2-781-7553

Received: 26 February 2020; Accepted: 15 April 2020; Published: 17 April 2020

Abstract: Face recognition is a representative biometric that can be easily used; however, spoofing attacks threaten the security of face biometric systems by generating fake faces. Thus, it is not advisable to only consider sophisticated spoofing cases, such as three-dimensional masks, because they require additional equipment, thereby increasing the implementation cost. To prevent easy face spoofing attacks through print and display, the two-dimensional (2D) image analysis method using existing face recognition systems is reasonable. Therefore, we proposed a new database called the "pattern recognition-face spoofing advancement database" that can be used to prevent such attacks based on 2D image analysis. To the best of our knowledge, this is the first face spoofing database that considers the changes in both the angle and distance. Therefore, it can be used to train various positional relationships between a face and camera. We conducted various experiments to verify the efficiency of this database. The spoofing detection accuracy of our database using ResNet-18 was found to be 96.75%. The experimental results for various scenarios demonstrated that the spoof detection performances were better for images with pinch angle, near distance images, and replay attacks than those for front images, far distance images, and print attacks, respectively. In the cross-database verification result, the performance when tested with other databases (DBs) after training with our DB was better than the opposite. The results of cross-device verification in terms of camera type showed negligible difference; thus, it was concluded that the type of image sensor does not affect the detection accuracy. Consequently, it was confirmed that the proposed DB that considers various distances, capture angles, lighting conditions, and backgrounds can be used as a training DB to detect spoofing attacks in general face recognition systems.

Keywords: face spoofing DB; face recognition; spoofing attack; RGB image sensor; angle and distance variations

1. Introduction

Nowadays, biometrics provide reliable indicators for individual recognition and authentication problems [1]. As the biometric identifiers are inherent to individuals, it is difficult to manipulate, share, or overlook these traits [2]. Therefore, these systems have been used in various fields such as cell phone encryption and internet banking authentication. Among biometric methods, the technique used by a face recognition system, which includes face detection and recognition, is one of the most convenient and useful practices [3–6]. The face recognition system uses a non-invasive method, and the face images have more complex biometric features compared to others. Various features that are used to detect fake face data can be extracted from each instance of data via local binary pattern (LBP), convolutional neural network (CNN), discrete cosine transform (DCT), and Laplacianfaces [7–10].

These reasons have led to the growth of the market size associated with face recognition, and the development of relevant robust systems is, therefore, required [11]. However, in the past few years, potentially vulnerable spoofing attacks have been reported [12]. These attacks occur when people attempt to pretend to be someone else by using fake data, thereby gaining illegitimate access and advantage [13]. Therefore, the face anti-spoofing task has attracted massive attention with the aim to assure reliability of security. In short, the necessity of detecting spoofing attacks in face recognition has increased. In conventional studies, to prevent spoofing attacks, its various types were divided into 2D attacks forged by displaying printed photos, replay attacks using recorded videos on mobile devices, and complex 3D facial mask attacks [14]. As known, there are several public databases where each has unique or characteristic data that has been collected in terms of various aspects. For example, the NUAA PI database and Yale Face Database B, which are among well-known face anti-spoofing databases, use only printed photo attacks [15,16]. Although the face images are obtained through various elements, such as movement, rotation, bending, and lighting manipulation of the photographs, practical requirements might not be satisfied for detection of counterfeit data because of the relatively small number of subjects, i.e., 15 and 10. Further, the Unicamp video-attack dataset (UVAD) prevents only the replay video attacks using 17,076 video clips by capturing 404 subjects from outdoor as well as indoor sites [17]. Additionally, databases such as CASIA-FASD, REPLAY-ATTACK, MSU-MFSD, MSU-USSA, REPLAY-MOBILE, and OULU-NPU include both printed photo and replay video attacks and can be used to consider more situations [18–23]. In particular, the MSU-USSA uses a unique factor that is not present in other databases. It has 1140 subjects that not only includes the face data collected by Wang [21] from web images but also data from REPLAY-ATTACK, CASIA-FASD, and MSU-MFSD public databases. Here, the images obtained from the web face database will have only one celebrity whose duplicate image does not exist. Finally, the ROSE-Youtu Face database has data that prevents masking attacks as well as printed photo and replay video attacks [24]. These public databases have significantly contributed to the field of face spoofing detection. Research results have currently given way to the use of commercialized face recognition systems that have anti-spoofing technologies for detecting fake data. Figure 1 shows the samples of public databases. Table 1 presents a comparison between previous face spoofing databases (DBs) and the proposed DB (Pattern Recognition-Face Spoofing Advancement Database (PR-FSAD)); to the best of our knowledge, the PR-FSAD is the only DB that considers the variations in both the distance and angle.

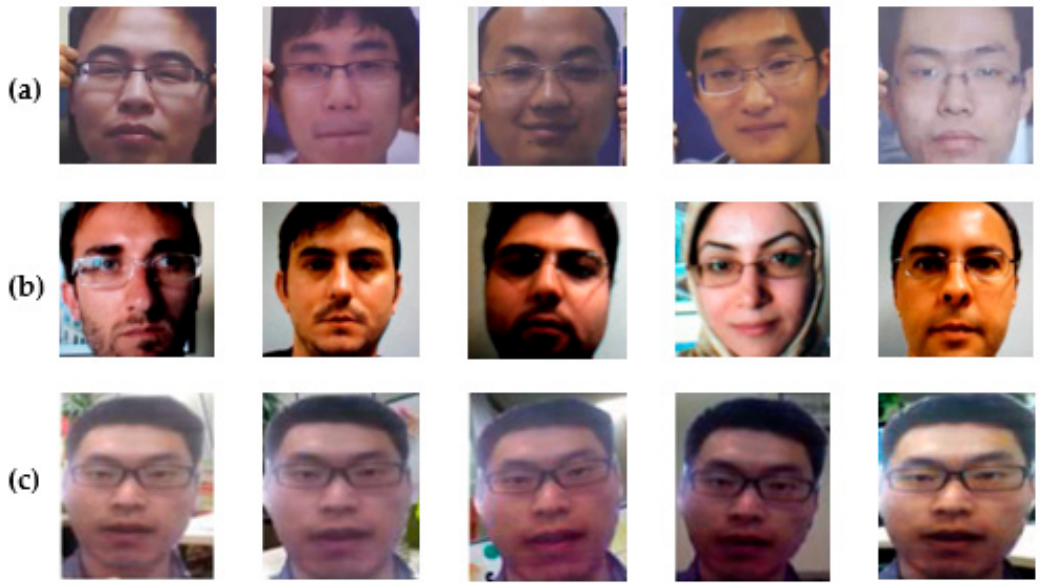

Figure 1. Fake image examples of public face databases: (**a**) printed photo attack samples; (**b**) replay video attack samples; (**c**) 3D facial mask attack samples.

Table 1. Comparison of previous face spoofing databases (DBs) and our DB (PR-FSAD).

Year	Database	# of Subjects	# of Samples (Real/Fake)	Attack Type (Medium)	Consideration of Positional Variation between Camera and Face
2010	NUAA-PI [15]	15	5105/7509	printed photo	angle (yaw)
2011	Yale-Recaptured [16]	10	640/1920	displayed photo	angle (yaw)
2012	CASIA-FASD [18]	50	150/450	printed photo, replayed video	distance
2012	REPLAY-ATTACK [19]	50	200/1000	displayed photo, replayed video, printed photo	distance
2014	MSU-MFSD [20]	35	70/210	printed photo, replayed video	distance
2015	UVAD [17]	404	808/16,268	replayed video	-
2016	MSU-USSA [21]	1000	1000/8000	printed photo, displayed photo	-
2016	REPLAY-MOBILE [22]	40	390/640	displayed photo, replayed video, printed photo	-
2017	OULU-NPU [23]	55	990/3960	printed photo, replayed video	-
2018	ROSE-Youtu [24]	20	899/2598	replayed video, printed photo	angle (yaw and pitch)
2019	Our DB (PR-FSAD)	30	42,480/84,960	printed photo, replayed video	angle (yaw and pitch), distance

Until now, most of the researches used only public face databases for fake detection. In such cases, as the face data missing from the public database might include new environmental factors, the performance of fake detection for new data might be lowered. Consequently, we created our own face database called PR-FSAD using RGB image sensor to prevent sophisticated spoofing attacks based on printed photo and replay video attacks. In this database, we considered the distance and angle conditions which were not applied in the previous public databases. Further, PR-FSAD consists of 30 subjects with an age range from 13 to 32 years regardless of gender. Thus, the requirements in terms of training and evaluation are met in a better manner compared to other databases. Once the entire data was obtained, the real and fake face data were preprocessed using a face detection algorithm. Four protocols were designed for performance evaluation, and classification was conducted using ResNet-18. Additionally, the cross-database scenarios were tested for evaluating detection accuracy among different DBs. Figure 2 shows the real and fake face data acquisition scheme of PR-FSAD using capture devices.

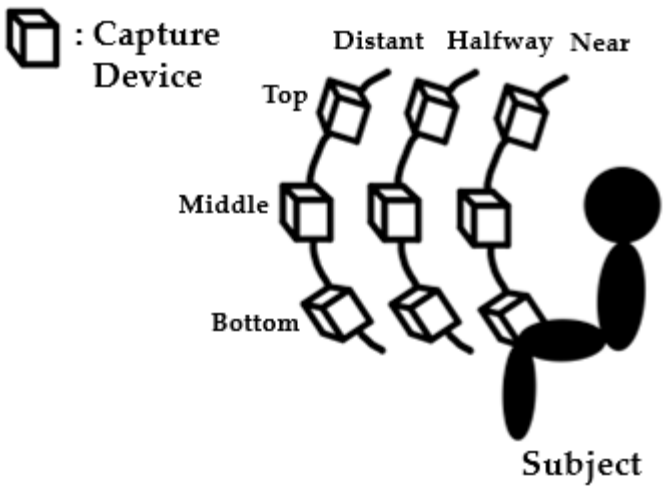

Figure 2. The capturing scheme of PR-FSAD with the three distances and three angles.

Our research had the following advantages over previous studies. First, a new face spoofing DB was constructed considering the distance and angle variations. Second, the efficiency of the proposed DB considering the deviations in distance and angle was verified by evaluating the spoofing detection accuracy through an algorithm based on deep neural networks. Third, the performance of the proposed DB was compared with those of previous face spoofing DBs by combining the training and test data from heterogeneous DBs to confirm the possibility of generalizing the proposed DB.

The rest of the paper is organized as follows. The detailed design of PR-FSAD and the four protocols for evaluation are described in Section 2. Section 3 shows the experimental results for the protocols for evaluation of PR-FSAD and cross-database and cross-device scenarios. The analysis of these results is shown in Section 4. Finally, Section 5 gives the conclusion, expected benefits, and future works.

2. Materials and Methods

In this section, we describe the camera devices, environmental conditions (such as posture, illumination, and background), and considerations for constructing PR-FSAD. Further, the real and fake face databases and protocols designed for evaluation are described in detail.

2.1. PR-FSAD

The PR-FSAD has various characteristics, such as unfixed backgrounds and three distance and angle cases, for capturing face images. These features can be distinguished from those of other conventional public face databases and may affect the process of real and fake face classification. To construct PR-FSAD, the real and counterfeit face images were obtained from 30 subjects. For the fake database, the printed photo and replay video attacks were used as attack methods to prevent spoofing attacks. The camera system, capture environment, consideration in terms of the PR-FSAD design, and the detailed information about real and fake face data are described in the following sections. Additionally, PR-FSAD is publicly available and can be obtained by submitting a request at our website [25].

2.1.1. Camera System

For capturing the robust real and fake face data, we used four photographing devices consisting of two smartphones and two tablets. The subjects captured the images using the front camera of the devices and recorded videos using the basic camera application that is built into each device. In addition, the camera was set to automatically adjust the brightness based on the change in illumination. The frames per second (fps) were set to be the same. The detailed information for all the devices is shown in Table 2.

Table 2. Information of the devices used in PR-FSAD.

Device	Category	Display Size	Pixel/Inch	Spatial Resolution	fps
Galaxy Tab 3	Tablet	7.0 Inch	170 ppi	1920 × 1080	30
iPad 6	Tablet	9.7 Inch	264 ppi	1280 × 720	30
iPhone X	Smartphone	5.8 Inch	463 ppi	1920 × 1080	30
Nexus 5X	Smartphone	5.2 Inch	424 ppi	1920 × 1080	30

2.1.2. Environmental Conditions

This section describes the considerations while capturing such as the pure-pose, expression, illumination, and background. All the images were captured using the camera system that was described in Section 2.1.1.

Firstly, each subject was asked to sit comfortably and look at the front, not toward the side, diagonal, or at the device. The capturing task was performed by keeping the subject's head at the

center of the subject's arm. Most of the subjects took pictures by holding the capturing device, while some of the subjects who found capturing difficult were assisted by the researcher. As is the case with most face recognition circumstances, the facial expressions were required to be natural, not that of laughing or frowning. Further, we explained the diversity of unfixed backgrounds, which is one of the distinct characteristics of PR-FSAD when compared to other existing public face databases. In a conventional face recognition system, the face data are completely segmented from the backgrounds during the preprocessing stage. This is the reason why previous face recognition systems might not have considered the system of not being affected by the backgrounds. However, the face data can often be modified by using various functions of the camera device such as automatic brightness balance or white balance applications. These factors can distort the facial appearance due to the varying illumination conditions of the background environment and can even affect facial recognition. Therefore, for similar real-world situations, the photographs were taken at various unfixed places, including cafes, restaurants, the lobby of a building, and lecture classroom. In other words, the subjects captured images in natural environments without arbitrarily fixed backgrounds.

2.1.3. Consideration of PR-FSAD

The two differentiated features of PR-FSAD were distance and angle. These two factors between the subject and camera may be applied differently depending on the environment where the face recognition system is being used. Furthermore, these differences may affect face recognition due to the factors such as changes in lighting or image texture. Therefore, we considered the abovementioned two factors. Firstly, for distances, as each subject has a different body, we used the relative ratio of the face occupying the display of the devices. Further, to apply the ratio accurately to all the subjects, the display was divided into 3×3 grids when the camera was used for capturing. One of the three distances, called near distance, fills the subject's face to approximately 90% of the screen. The halfway distance fills approximately two-thirds of the entire screen. In other words, the face occupies approximately 50% of the eight rectangles except the rectangle at the center of the screen and is located at the center of the display. Finally, the face of distant distance is photographed while keeping only the center of the 3×3 grids screen filled. Further, for angles, the top and bottom were positioned differently approximately 30° from the center angle. When capturing a face from various angles, to match a real-world situation, the subject's gaze will be in the same direction as the middle angle and not looking at the device. While acquiring the face data of PR-FSAD, each subject had to capture three preset distances and angles. In addition, as a face is rarely yawed or rolled in actual use-cases of face recognition, we captured the face images with different pitches. Once captured, the face data for distance and angle were stored using the tags "near", "halfway", and "distant" and "bottom", "middle", and "top". Figure 2 shows the capture method with the considerations.

2.1.4. Real Face Database

The PR-FSAD consists of 30 subjects (male: 19, female: 11). All the subjects except one captured face images in two sessions with the time interval set to at least six hours. As time difference is an important factor that decreases the classification performance, subjects took pictures during the first session at daytime and the other at night [26]. During each session, different backgrounds were applied to each subject. Other capture conditions were performed by the ones written in Section 2.1.3. In addition, the accuracy of face detection while constructing the PR-FSAD was checked using the multitask cascaded convolutional network (MTCNN) face detection algorithm [27]. This is because spoofing detection must be performed based on the accuracy of the detected face data to obtain a significant result. If face detection is not accurate, data are recaptured to construct precise real face data for the PR-FSAD. Figures 3 and 4 show the real face data of the PR-FSAD and the results of the MTCNN method at three angles.

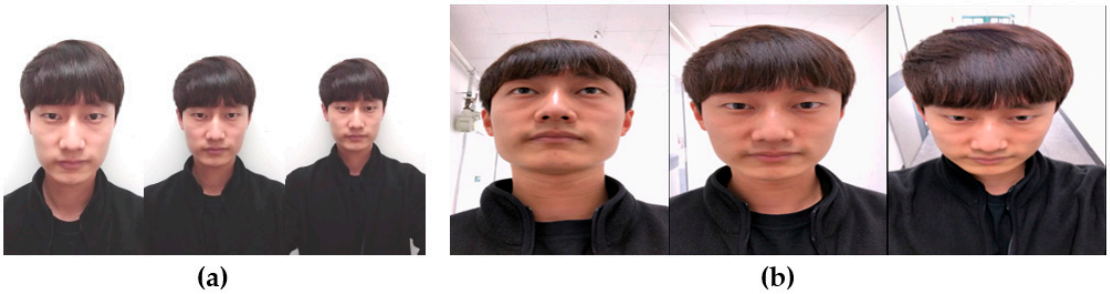

Figure 3. Samples from the real face database: (**a**) three distances (near, halfway, distant); (**b**) three angles (bottom, middle, top).

Figure 4. Results for the multitask cascaded convolutional network (MTCNN) using real face data.

2.1.5. Fake Face Database

When creating a fake face for PR-FSAD for spoofing attacks, we used two categories of attacks, namely, printed photo and replay video attacks. The capturing angles and distances used were the same as in the case of the real face data.

Firstly, we used the photographed real face images of all the subjects with four devices for the printed photo attack. For the counterfeit face data to be as similar as possible to the real face, the frame with the most natural look was chosen among the images taken at the halfway distance from the middle angle. The selected frame was printed using a high-quality color printer (Samsung SL-C483W, Fuji Xerox CP115W) to deceive the face recognition system with a high probability of spoofing attacks. While keeping the subject's gaze in the printed image at the front, the counterfeit face images were captured as the real face data. While capturing, the other conditions were performed by the ones written in Section 2.1.3. However, unlike the case for real people, detection of a face in printed photos might be difficult due to the reflection of unexpected light from behind the paper. Therefore, a printed photo has to be maintained as if it were the real face of a person holding the image. Once the fake data were captured, the procedure for normal face detection was also applied quite similar to as it is in the case of real face database. Figure 5 shows the printed photo attack of PR-FSAD and Figure 6 shows the results of the MTCNN method.

Figure 5. Samples of a printed photo attack: (**a**) three distances (near, halfway, distant); (**b**) three angles (bottom, middle, top).

Figure 6. Results for MTCNN using fake face data from a printed photo attack.

Further, we used photographed real face videos of all the subjects for the replay video attack. However, a drawback with smartphones was the relatively small face scale on the screen. Therefore, when the spoofing attacks were attempted at close distances, the focus often did not match with the device for face recognition. To prevent this problem, two tablet devices were used for the replay video attack. Further, as only the face of the subjects in the tablet's display had to be detected, the tablet was used by keeping it at approximately 0.1 m below the shoulder of the person holding the device. In contrast, the smartphone was kept next to the shoulder of the person holding the device. Other capture conditions and procedures were similar to that of the printed photo attack. Figure 7 shows the replay video attack of PR-FSAD, and Figure 8 shows the results for the MTCNN method.

Figure 7. Samples of a replay video attack: (**a**) three distances (near, halfway, distant); (**b**) three angles (bottom, middle, top).

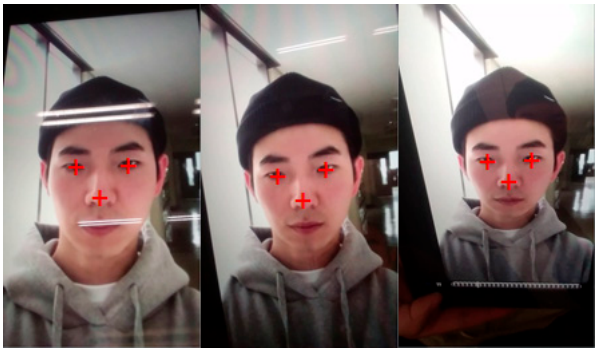

Figure 8. Results for MTCNN using fake face data from a replay video attack.

2.2. Evaluation Protocols

We considered various backgrounds, distances, and angles as the features of PR-FSAD. Protocols consisting of eight scenarios were designed to evaluate and verify the performance of face spoofing attack detection using PR-FSAD. For classification evaluation, the distances and angles were divided into three and two cases, respectively. The variables T, M, and B were used to represent top, middle, and bottom of the angle factor, and N, H, and D were used to represent near, halfway, and distant of the distance factor. The background, however, was not considered in the protocol, because it is configured differently and, hence, difficult to divide. In the test, 1, 2, and 3 indicate the real, printed photo, and replay video attacks. The detailed description of the protocols is as follows:

1. Angle test: At each of the three different angles, real and fake data for all the three distances are used:
    a. Top-angle protocol: use {TN1-3, TH1-3, TD1-3};
    b. Middle-angle protocol: use {MN1-3, MH1-3, MD1-3};
    c. Bottom-angle protocol: use {BN1-3, BH1-3, BD1-3}.
2. Distance test: To clarify the difference in the distances, the three angles of real and fake data are used for two of the distances, where the halfway distance is excluded:
    a. Near distance protocol: use {TN1-3, MN1-3, BN1-3};
    b. Distant distance protocol: use {TD1-3, MD1-3, BD1-3}.
3. Counterfeit face test: For all the angles and distances, two types of counterfeit face tests are used:
    a. Printed photo attack protocol: this uses real and printed photo attack data at all the angles and distances (or uses 1 and 2 at all the angles and distances);
    b. Replay video attack protocol: this uses real and replay video attack data at all the angles and distances (or uses 1 and 3 at all the angles and distances).
4. Overall test: The evaluation test is conducted using all the angles and distances of PR-FSAD:
    a. Entire data protocol: all the real and fake face data are used.

To perform the designed protocols, PR-FSAD was divided into training, validation, and test sets, which included 7, 10, and 13 subjects, respectively. The composition of the data used for the experiment and the processing time required for training and testing are presented in Table 3. The processing time was measured using Intel (R) Core (TM) i7-6700HQ quad-core CPU, 2.60 GHz with 16 GB RAM and NVIDIA GeForce GTX 1070 GPU with 16 GB RAM. In the training process, the early stopping strategy was adopted before 20 epochs to ensure that the training was completed relatively quickly.

Table 3. Data composition and processing time for the experiment.

Protocols 1–4		Number of Images			Total Processing Time (s)	
		Training (32.2%)	Validation (23.7%)	Test (44.1%)	Training	Test
1	Top	13,680	10,080	18,720	1898	374 (20.0 ms/image)
	Middle	13,680	10,080	18,720	1920	376 (20.1 ms/image)
	Bottom	13,680	10,080	18,720	1886	374 (20.0 ms/image)
2	Near	13,680	10,080	18,720	1869	372 (19.9 ms/image)
	Distant	13,680	10,080	18,720	1906	375 (20.0 ms/image)
3	Print	27,360	20,160	37,440	3895	751 (20.1 ms/image)
	Replay	27,360	20,160	37,440	3912	755 (20.2 ms/image)
4	Total	41,040	30,240	56,160	5985	1132 (20.2 ms/image)

2.3. Face Spoofing Detection Method

The PR-FSAD face data constructed for evaluating spoofing detection performances had similarities to adjacent frames. Therefore, to eliminate this unnecessary similarity and use varying data, sampling was performed. In particular, only 20 images were sampled by extracting images at intervals of 2 to 3 frames per video. Next, the extracted images were preprocessed to crop the face area except the background. In this study, we resized the images to 224 × 224 pixels. The preprocessed images are shown in Figure 9.

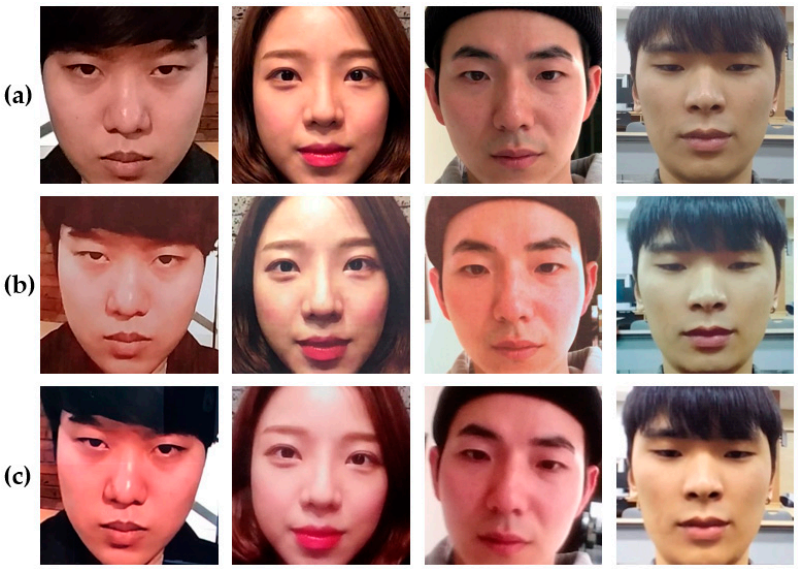

Figure 9. The results of face data preprocessing: (**a**) real; (**b**) printed photo; (**c**) replay video.

One of the deep neural network models, ResNet-18, was used for the final real and fake face classification based on the processed face data [28,29]. Compared with conventional neural network models, ResNet-18 did not cause problems in terms of gradient vanishing or exploding as the layer deepens. This effect was due to the shortcut connection that passes the input of a specific layer directly to the output, making it easier to find out and train fine-grained changes during a model's training process. The preprocessed image was normalized to 224 × 224 pixels and input into the ResNet-18 model. Because it is a three-channel color image, the input feature was defined in 150,528 dimensions (224 × 224 × 3). The feature vector output obtained by average pooling the ResNet-18 model had 512 dimensions. Finally, an output value that determines whether the input face image was real or fake was calculated using the sigmoid function. The procedure of counterfeit face detection using the proposed method is shown in Figure 10.

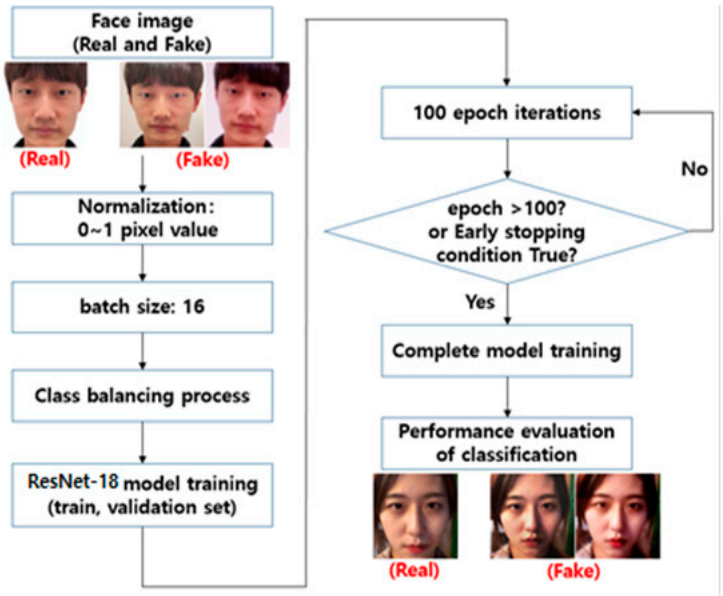

Figure 10. Flowchart of the proposed method for classification.

3. Results

In this study, the ResNet-18 model with 0.01 learning rate was used for face spoofing attack detection. Additionally, the half total error rate (HTER) was calculated to verify the performance of the test results. The HTER is the average error rate of the false acceptance rate (FAR) and the false rejection rate (FRR) of the validation set. In the HTER, a smaller value means that the classification performance is better. The HTER indicator was obtained using the confusion matrix which is one of the most intuitive and simple methods for measuring the performance of binary classification models. The four values returned by the confusion matrix are true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Here, assuming real face as positive and fake face as negative, TP and TN indicate correct classification, whereas FP and FN indicate incorrect classification. Table 4 lists the HTER for protocols 1–4. Among them, the result of the confusion matrix for protocol 4 using the total PR-FSAD is shown in Figure 11. In this case, the HTER is calculated by using Equation (1) which is approximately 3.25%.

$$HTER = (FP/(TN + FP) + FN/(FN + TP)) \times 0.5 \quad (1)$$

Table 4. The half total error rate HTER of the ResNet-18 model using PR-FSAD.

Protocol 1–4		HTER (%)
1	Top	2.34
	Middle	4.96
	Bottom	2.84
2	Near	1.41
	Distant	4.24
3	Print	5.36
	Replay	3.60
4	Total	3.25

It is evident from Table 4 that the replay attack of the protocol 3 experiment demonstrated better spoofing attack detection performance than print attack. In the experiment, we used a single frame image for face spoofing attack detection. Information related to the changes in the time series of the face video in the replay attack was not used. Therefore, the texture information might have been

used as the most important feature in both replay and print attacks. The screen of the smart device used in the replay attack was made of a glass material that could cause specular reflection from the ambient light source. Such specular reflection can be observed in Figures 7 and 8. This characteristic was markedly different from the actual skin surface. In contrast, paper materials did not produce specular reflection. This difference can be analyzed as a performance variation.

		Predicted Values: Negative (Fake)	Predicted Values: Positive (Real)
Actual Values	Negative (Fake)	36,226 (TN)	1,214 (FP)
	Positive (Real)	608 (FN)	18,112 (TP)

Figure 11. The result of confusion matrix for protocol 4 (TN: True Negative, TP: True Positive, FN: False Negative, FP: False Positive).

In addition, when examining the performance of protocol 1, it was evident that the performance for the front face (middle) was significantly lower than those for the top and bottom wherein the pitch angle difference existed. This could be analyzed because the top and bottom images were distorted by the vertical perspective of the face, whereas other features along with the three-dimensional features of the face were reflected in the training process. In other words, in the middle case, only the texture feature was reflected without considering the perspective property.

However, as the above results used only intra-database scenarios of PR-FSAD, it may not be sufficient to demonstrate the effectiveness of PR-FSAD which includes differentiated features compared to previous public databases. Therefore, we used public databases called MSU-MFSD and REPLAY-ATTACK for cross-database scenarios. In all the scenarios, the face spoofing attack detection was performed using the ResNet-18 model, and the HTER obtained by using the confusion matrix was used as an indicator of performance evaluation.

Table 5 lists the results of cross-database scenarios. To confirm that the PR-FSAD is also efficient with other spoofing detection algorithms, along with ResNet-18, we considered DenseNet [30] and LBP [31] for the cross-database scenario experiments. Experimental results revealed that the spoofing attack detection performance by training with the PR-FSAD was the best for ResNet-18, DenseNet, and LBP. This shows that although the PR-FSAD contains an angle variation element, it is feasible as training data that can be generalized and used in other face recognition systems. In addition, the relatively good performance for the LBP feature that only uses the texture property can be considered to explain the textural features of spoofing attack media independent of the image sensor.

Although the previous DBs (MSU-MFSD, REPLAY-ATTACK) used in the comparison included distance variations, our DB had both distance and angle variations. In the cross-database scenarios presented in Table 5, our DB used both distance and angle variations for training and testing. Therefore, we performed cross-database experiments for a fair comparison using images with only distance variation in the front (such as distant, halfway, and near in the middle position, as shown in Figure 2). The results are presented in Table 6.

Table 5. Results of cross-database scenarios using ResNet-18, DenseNet, and LBP (1st column: Trained DB, 2nd Column: Test DB).

Cross-Database Scenarios		HTER (%)		
		ResNet-18	DenseNet	LBP
PR-FSAD	MSU-MFSD	19.96	18.52	23.10
	REPLAY-ATTACK	20.80	19.67	25.12
MSU-MFSD	PR-FSAD	23.48	25.21	28.36
	REPLAY-ATTACK	17.58	18.36	21.53
REPLAY-ATTACK	PR-FSAD	34.41	32.23	35.36
	MSU-MFSD	28.44	30.81	33.29

Table 6. Results of cross-database scenarios using ResNet-18, DenseNet, and LBP when using only frontal face images from the PR-FSAD (1st column: Trained DB, 2nd Column: Test DB).

Cross-Database Scenarios		HTER (%)		
		ResNet-18	DenseNet	LBP
PR-FSAD	MSU-MFSD	13.45%	14.28%	21.15%
	REPLAY-ATTACK	14.93	16.50%	23.96%
MSU-MFSD	PR-FSAD	16.36%	17.13%	23.83%
	REPLAY-ATTACK	17.58%	18.36%	21.53%
REPLAY-ATTACK	PR-FSAD	18.76%	18.37%	30.21%
	MSU-MFSD	28.44%	30.81%	33.29%

In comparison to the experiments wherein the variations in both the angle and distance were included (Table 5), improved results were obtained as demonstrated in Table 6. These results indicate that our DB reflects the perspective and resolution characteristics via angle and distance variations, respectively, to generalize the data to other capturing environments. However, because Table 6 presents the results for training and testing performed only with frontal face images from the PR-FSAD, an absolute comparison in terms of the training dataset with the results of Table 5 and Figure 12 is not possible.

Next, we performed experiments to measure the effect of the PR-FSAD, which consists of nine times more data than only the frontal face images by considering the distance and angle variations, on the processing time and classification accuracy. When the training image is used for the front face only and when the nine times more images are used, the time required for training can be considered to be a computational cost. However, because training is performed only once, comparing the training time would be inconsequential. Instead, we measured the time required for spoofing detection with one face image. The processing time was measured using Intel (R) Core (TM) i7-6700HQ quad-core CPU 2.60 GHz with 16 GB RAM and NVIDIA GeForce GTX 1070 GPU with 16 GB RAM. The results of the measurement time, which represents the average time required for 500 images, are presented in Table 7. The measurements are expressed in terms of when only CPU was used and when it was used along with GPU. In addition, Table 7 includes the test accuracy results when we used the nine times extended DB considering the angle and distance and when only the front face was used for training.

Table 7. Spoofing detection processing times and classification accuracies for a single image based on distance and angle variations (using only CPU/using GPU together).

	Training with only Front Face Image	Training with Total Image
Processing time (ms)	320/20	321/20
Accuracy (HTER (%))	5.12	3.25

The results show that the time difference in the actual face spoofing detection process is insignificant, but the accuracy is significantly improved. In other words, the benefit of using the image with nine times more data considering the distance and angle variations was confirmed.

Finally, we assessed the performance of face spoofing attacks for cross-device scenarios on images captured with the four types of capturing devices specified in Table 2. The experimental results are presented in Table 8. In the experiment, the images acquired for each device were divided into training and test sets. This test was performed using ResNet-18.

Table 8. HTER results for cross-device scenarios of PR-FSAD using ResNet-18 (unit: %).

Training \ Test	Galaxy Tab 3	iPad 6	iPhone X	Nexus 5X
Galaxy Tab 3	3.32	3.30	3.29	3.25
iPad 6	3.18	3.17	3.20	3.21
iPhone X	3.18	3.25	3.23	3.19
Nexus 5X	3.32	3.28	3.26	3.28

As evident from Table 8, the intra-device and inter-device facial spoofing detection performances were not significantly different. In some cases, the HTER of the intra-device was larger than that of the inter-device. Thus, it can be concluded that the characteristics of the media (paper or display) used for the spoofing attack and the geometric positional relationships were reflected accurately in the training process of ResNet-18 as the main feature of spoofing detection, instead of the differences in the image sensor for each device.

4. Discussion

Results of forgery detection methods using only the face data from PR-FSAD showed excellent performance with an HTER value of less than 5% for all the protocols except one that had an HTER value of 5.36%. As shown by the result of the confusion matrix for protocol 4, which uses the entire PR-FSAD, the ratios of misclassification for real and fake face data were almost identical. Although the number of misclassified data was different, the ratio was the same because PR-FSAD had a 1:2 ratio between the real and fake data. Moreover, this means that the ResNet-18 model training was performed accurately without any biases. Next, significant results were obtained for cross-database scenarios using three public databases.

Firstly, the best spoofing detection result was obtained for MSU-MFSD using the ResNet-18 model trained with PR-FSAD. For the test using REPLAY-ATTACK, the classification result using a model trained with MSU-MFSD, which had a similar data distribution, was the best, and the performance using the model trained with PR-FSAD was the second best. Although a difference of only 3% was observed, because the PR-FSAD demonstrated its training effect in heterogeneous DBs, it can be considered advantageous as a generalized DB that can overcome the variations in face image capturing conditions. Further, for the other two public databases, although the classification results for each other were good, the results for PR-FSAD were relatively poor. In contrast, our database is considered to have a better generalization performance because the HTER values for the public databases were approximately 20%. In other words, it can be noted that more face features were applied to PR-FSAD.

The main contribution of this study was the introduction of a new face spoofing DB that reflects distance and angle differences. We divided the distance and angle into three phases to construct facial spoofing images for nine combinations. In addition, the data were configured in various environments without controlling the lighting or background to reflect the actual environment. In our experiments with the proposed database, a face spoofing detection accuracy of 96.75% was observed. Using the proposed database, the resolution variations in the facial region can be consistently reflected as learned features in the deep neural network learning process by capturing the data at different distances and including perspective variations by including different angles in the images.

The limitation of our DB, however, is that data are obtained by dividing the positional relationship between the face and camera into nine types according to the angle and distance. This limitation was intended to provide a clear guide to the subject in the process of acquiring images. In the future, we plan to incorporate additional data by changing the positional relationship between the camera and face.

Consequently, PR-FSAD, which has a relatively good classification performance for different data alongside itself, is a meaningful face database to prevent spoofing attacks. For result visualization, we used the receiver operating characteristic (ROC) curve, which is useful for visualizing performances [32]. The wider the area under the curve (AUC), which indicates the bottom area of the ROC curve, the better the performance of the classification model. Figure 12 shows the ROC curve results using cross-database scenarios. The rate of sensitivity or recall, called the true positive rate, and the rate of specificity, called the false positive rate, are plotted on the *y*-axis and *x*-axis, respectively. Further, PR-FSAD showed the best classification performance among the four databases as can be seen in Figure 12.

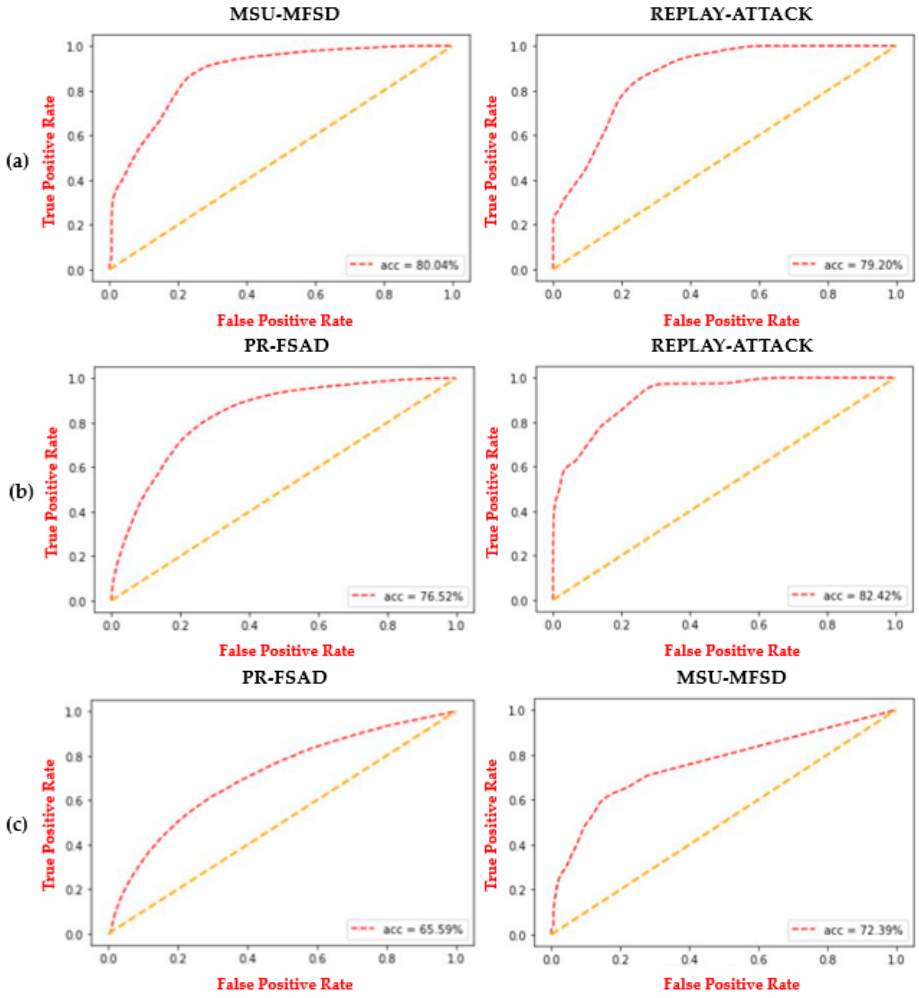

Figure 12. The result of ROC curve using the model trained with each of three databases: (**a**) PR-FSAD; (**b**) MSU-MFSD; and (**c**) REPLAY-ATTACK.

5. Conclusions

In this study, we reported a new face DB called the PR-FSAD and verified its effectiveness in terms of face spoofing detection accuracy. The real and fake face data were obtained using four capture devices, and the spoofing attacks primarily consisted of printed photo attack and replay video

attack. In particular, compared to public face databases, PR-FSAD is composed of two new factors which are distance and angle. To the best of our knowledge, a DB that considers both the distance and angle has not been proposed in the existing literature which is the main contribution of this study. A combination of three distances and angles were used to construct the real and fake face database, followed by sampling of the images from the captured videos. Finally, the face region was cropped, and the processed face data were applied to ResNet-18 neural network model for classification. To verify the effectiveness of PR-FSAD, 10, 7, and 13 subjects out of the total 30 subjects were used in the proposed method for training, validation, and test, respectively. Specifically, 41,040 (32.2%), 30,240 (23.7%), and 56,160 (44.1%) were applied to training, validation, and test. In addition, the classification was performed using RGB images without any additional equipment or sensors to detect the spoofing attacks. As a result, the HTER, which was used to measure the performance evaluation of the classification, was 3.25%. This result demonstrated a good classification performance in comparison to the existing DBs. For further effectiveness verification of PR-FSAD, we designed the cross-database scenarios using three face databases, including two public databases. The test performances of the three algorithms, namely, ResNet-18, DenseNet, and LBP, which were trained using the proposed DB, were the best. It is still lacking in terms of various algorithm comparisons, but it can be concluded that our DB was more applicable to face recognition systems in other environments.

In future studies, more accurate spoofing detection methods using PR-FSAD may be considered for other applications such as low-quality face data. Further, we will implement additional procedures, such as registering face data, with new feature elements. Furthermore, PR-FSAD can be used in released applications with face recognition systems for preventing counterfeit attacks. By acquiring more face images by further subdividing the environment of distance and angle variations, the proposed PR-FSAD will be improved to a more generalized face spoofing DB. Additionally, we plan to add the face spoofing DB generated from the frontal face images along with face spoofing images obtained through attack media that already include the reflected angle and distance variations. This may be an attack case that is more difficult to filter than face spoofing with angle variations in the frontal face image. Moreover, by comparing various deep neural network models, we will analyze the effects of texture and perspective on face spoofing detection. With respect to facial spoofing systems, we plan to conduct research on the prevention of disturbances related to deep learning or new attack models, such as Deepfake and adversarial perturbations, which are recently becoming issues.

Author Contributions: Conceptualization, E.C.L.; methodology, E.C.L.; software, J.Y.B.; validation, J.Y.B. and K.H.S.; formal analysis, E.C.L., J.Y.B.; investigation, K.H.S.; resources, E.C.L.; data curation, J.Y.B.; writing—original draft preparation, J.Y.B.; writing—review and editing, E.C.L. and K.H.S.; visualization, J.Y.B.; supervision, K.H.S.; project administration, E.C.L.; funding acquisition, E.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Bio&Medical Technology Development Program of the NRF funded by the Korean government, MSIP(Grants No. 2016M3A9E1915855).

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

References

1. Jain, A.K.; Ross, A.; Pankanti, S. Biometrics: A tool for information security. *IEEE Trans. Inf. Forensics Secur.* **2006**, *1*, 125–143. [CrossRef]
2. Ross, A.; Jain, A. Information fusion in biometrics. *Pattern Recognit. Lett.* **2003**, *24*, 2115–2125. [CrossRef]
3. Chen, L.F.; Liao, H.Y.M.; Ko, M.T.; Lin, J.C.; Yu, G.J. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognit.* **2000**, *33*, 1713–1726. [CrossRef]
4. Jeng, S.H.; Liao, H.Y.M.; Han, C.C.; Chern, M.Y.; Liu, Y.T. Facial feature detection using geometrical face model: An efficient approach. *Pattern Recognit.* **1998**, *31*, 273–282. [CrossRef]
5. Han, C.C.; Liao, H.Y.M.; Yu, G.J.; Chen, L.H. Fast face detection via morphology-based pre-processing. *Pattern Recognit.* **2000**, *33*, 1701–1712. [CrossRef]
6. Zhao, W.; Chellappa, R.; Phillips, P.J.; Rosenfeld, A. Face recognition: A literature survey. *ACM Comput. Surv. (CSUR)* **2003**, *35*, 399–458. [CrossRef]

7. Ahonen, T.; Hadid, A.; Pietikainen, M. Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 2037–2041. [CrossRef]
8. Parkhi, O.M.; Vedaldi, A.; Zisserman, A. Deep face recognition. *BMVC* **2015**, *1*, 6. [CrossRef]
9. Gunlu, G.; Bilge, H.S. Face Recognition by Using 3D Discrete Cosine Transform. In Proceedings of the 2007 IEEE 15th Signal Processing and Communications Applications, Eskisehir, Turkey, 11–13 June 2007; pp. 1–4. [CrossRef]
10. He, X.; Yan, S.; Hu, Y.; Niyogi, P.; Zhang, H.J. Face recognition using laplacianfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 328–340. [CrossRef]
11. Huang, J.; Heisele, B.; Blanz, V. Component-based face recognition with 3D morphable models. In Proceedings of the International Conference on Audio-and Video-Based Biometric Person Authentication, Guildford, UK, 9–11 June 2003; pp. 27–34. [CrossRef]
12. Biggio, B.; Akhtar, Z.; Fumera, G.; Marcialis, G.L.; Roli, F. Security evaluation of biometric authentication systems under real spoofing attacks. *IET Biom.* **2012**, *1*, 11–24. [CrossRef]
13. Hadid, A. Face biometrics under spoofing attacks: Vulnerabilities, countermeasures, open issues, and research directions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 113–118. [CrossRef]
14. Erdogmus, N.; Marcel, S. Spoofing face recognition with 3D masks. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 1084–1097. [CrossRef]
15. Tan, X.; Li, Y.; Liu, J.; Jiang, L. Face liveness detection from a single image with sparse low rank bilinear discriminative model. In Proceedings of the European Conference on Computer Vision, Heraklion, Greece, 5–11 September 2010; pp. 504–517. [CrossRef]
16. Georghiades, A.S.; Belhumeur, P.N.; Kriegman, D.J. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 643–660. [CrossRef]
17. Pinto, A.; Schwartz, W.R.; Pedrini, H.; de Rezende Rocha, A. Using visual rhythms for detecting video-based facial spoof attacks. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 1025–1038. [CrossRef]
18. Zhang, Z.; Yan, J.; Liu, S.; Lei, Z.; Yi, D.; Li, S.Z. A face antispoofing database with diverse attacks. In Proceedings of the 2012 5th IAPR International Conference on BIOMETRICS (ICB), New Delhi, India, 29 March–1 April 2012; pp. 26–31. [CrossRef]
19. Chingovska, I.; Anjos, A.; Marcel, S. On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the 2012 BIOSIG International Conference of Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 6–7 September 2012; pp. 1–7.
20. Wen, D.; Han, H.; Jain, A.K. Face spoof detection with image distortion analysis. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 746–761. [CrossRef]
21. Patel, K.; Han, H.; Jain, A.K. Secure face unlock: Spoof detection on smartphones. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 2268–2283. [CrossRef]
22. Costa-Pazo, A.; Bhattacharjee, S.; Vazquez-Fernandez, E.; Marcel, S. The replay-mobile face presentation-attack database. In Proceedings of the 2016 International Conference of the Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 21–23 September 2016; pp. 1–7. [CrossRef]
23. Boulkenafet, Z.; Komulainen, J.; Li, L.; Feng, X.; Hadid, A. OULU-NPU: A mobile face presentation attack database with real-world variations. In Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 30 May–3 June 2017; pp. 612–618. [CrossRef]
24. Li, H.; Li, W.; Cao, H.; Wang, S.; Huang, F.; Kot, A.C. Unsupervised domain adaptation for face anti-spoofing. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1794–1809. [CrossRef]
25. Available online: http://pr.smu.ac.kr/patent (accessed on 17 April 2020).
26. Gao, W.; Cao, B.; Shan, S.; Chen, X.; Zhou, D.; Zhang, X.; Zhao, D. The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Trans. Syst. Manand Cybern. Part A Syst. Hum.* **2007**, *38*, 149–161. [CrossRef]
27. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* **2016**, *23*, 1499–1503. [CrossRef]

28. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
29. Wu, S.; Zhong, S.; Liu, Y. Deep residual learning for image steganalysis. *Multimed. Tools Appl.* **2018**, *77*, 10437–10453. [CrossRef]
30. Gao, H.; Shichen, L.; Laurens, M.; Kilian, Q.W. CondenseNet: An efficient DenseNet using learned group convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 2752–2761.
31. Lei, L.; Xiaoyi, F.; Zhaoqiang, X.; Xiaoyue, J.; Abdenour, H. Face spoofing detection with local binary pattern network. *J. Vis. Commun. Image Represent.* **2018**, *54*, 182–192. [CrossRef]
32. Powers, D.M. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *J. Mach. Learn Technol.* **2011**, *2*, 37–63.