Article

Face Spoof Attack Detection using Deep Background Subtraction

Azeddine Benlamoudi[1], Salah Eddine Bekhouche [2], Maarouf Korichi [1], khaled Bensid [1], Abdeldjalil Ouahabi [3]*, Abdenour Hadid [4], and Abdelmalik Taleb-Ahmed [4]

[1] Univ Ouargla, Fac. des nouvelles technologies de l'information et de la communication. Lab. de Génie Électrique, Ouargla 30 000, Algeria.
[2] University of the Basque Country UPV/EHU, San Sebastian, Spain.
[3] UMR 1253, iBrain, INSERM, Université de Tours, Tours 37000, France.
[4] Institut d'Electronique de Microélectronique et de Nanotechnologie (IEMN), UMR 8520, Université Polytechnique Hauts de France, Université de Lille, CNRS, 59313, Valenciennes, France.
* Correspondence: ouahabi@univ-tours.fr

Abstract: Currently, face recognition technologies are the most widely used methods for verifying an individual's identity. Nevertheless, it has increased in popularity, raising concerns about face spoofing attacks, in which a photo or video of an authorized person's face is used to get access to services. Based on a combination of Background Subtraction (BS) and Convolutional Neural Networks (CNN), as well as an ensemble of classifiers, we propose an efficient and more robust face spoof detection algorithm. This algorithm includes a Fully Connected (FC) classifier with a Majority Vote (MV) algorithm, which uses different face spoof attacks (e.g., printed photo and replayed video). By including a majority vote to determine whether the input video is genuine or not, the proposed method significantly enhances the performance of the Face Anti-Spoofing (FAS) system. For evaluation, we considered the MSU MFSD, REPLAY-ATTACK, and CASIA-FASD databases. The obtained results by our proposed approach are better than those obtained by state of the art methods. On the REPLAY-ATTACK database, we were able to attain a Half Total Error Rate (HTER) of 0.62% and an Equal Error Rate (EER) of 0.58%. It was possible to attain an EER of 0% on both the CASIA-FASD and the MSU FAS databases.

Keywords: Biometrics; Face spoofing; CNN; BS; ResNet-50

1. Introduction

Individuals can be successfully identified and authenticated using biometric features and traits. Hence, it is appropriate for access control and global security systems that depend on person recognition which is achieved through the use of a variety of biometric modalities, ranging from the classic fingerprint through the face, iris, ear [1–4] and, more recently, vein and blood flow. Furthermore, a number of spoofing methods have been developed in order to overcome such biometric systems. [5]. When someone tries to get around a face biometric system by placing a fake face in front of the camera, this is known as a spoofing attack. Nevertheless, compared to other modalities, the abundance of still face images or video sequences on the internet has made it exceptionally easy to obtain a person's facial data.

The spoofing detection literature discusses two types of spoofing attacks namely print and replay. The print attack spoofs 2D face recognition systems by using printed photographs of a subject, whereas the replay attack presents a video of a live person to avoid liveness detection. Furthermore, the low cost of launching a face spoof attack has increased the prevalence of the problem. Face recognition system spoofing media ranges from low-quality paper prints to high-quality photographs, as well as video streams played in front of the biometric authentication system sensor.

Table 1. Definition of the main acronymes.

Acronym	Description
AIM-FAS	Adaptive Inner-update Meta Face Anti-spoofing
AIU	Adaptive Inner-Update
AUC	Area Under Curve
BASN	Bipartite Auxiliary Supervision Network
BS	Background Subtraction
BiFPN	Bi-Directional Feature Pyramid Network
CDC	Central Difference Convolution
CDCN	Central Difference Convolutional Network
CDP	Central Difference Pooling
CM	CASIA MSU
CNN	Convolutional Neural Networks
CR	CASIA Replay
DET	Detection Error Tradeoff
DL	Deep Learning
DMD	Dynamic Mode Decomposition
DSIFT	Dense Scale Invariant Feature Transform
DTN	Deep Tree Network
EER	Equal Error Rate
FAR	False Acceptance Rate
FAS	Face Anti-Spoofing
FC	Fully Connected
FDML	Frame Difference and Multi-Level
FL	Feature Learning
FRR	False Rejection Rate
GT	Ground Through
HAR	Human Activity Recognition
HTER	Half Total Error Rate
IDA	Image Distortion Analysis
IQA	Image Quality Assessment
LBP	Local Binary Patterns
LBP-TOP	Local Binary Pattern on Three Orthogonal Planes
MBSIF-TOP	Multiscale Binarized Statistical Image Features on Three Orthogonal Planes
MC	MSU CASIA
MDA	Marginal Distribution Alignment
MEGC	Multiple Explainable and Generalizable
MFRM	Multi-level Feature Refinement Module
MLPQ-TOP	Multiscale Local Phase Quantization on Three Orthogonal Plane
MR	MSU Replay
MV	Majority Vote
OFFB	Optical Flow guided Feature Block
RC	Replay CASIA
RM	Replay MSU
ROC	Receiver Operating Characteristic
STASN	Spatio-Temporal Anti-Spoof Network
SVM	Support Vector Machine
TL	Transfer Learning
USDAN	Unsupervised and Semi-supervised Domain Adaptation Network

Table 1 defines the main acronymes used in this paper.

Feature extraction is a critical component of the face anti-spoofing task when using a classical machine learning classifier. CNN can be used also to predict scores. This latter is a crucial component of deep learning algorithms such as the ResNet-50 pre-trained model, which has been studied for a few years under a variety of conditions and scenarios. In our work, we use BS with CNN to predict each frame in the input video and rank the score using the MV algorithm to determine whether the input video is real or fake.

Inspired by the work of Frame Difference and Multilevel Representation (FDML) [6], we propose effective biometrics systems based on the detection of face spoofing. To do this, we suggest using the background substruction method in the preprocessing step to adjust the face's motion. The MV algorithm is used to improve the performance rate as well as the decision of the input video after predicting the score of each frame by ResNet-50. To test our system, we used videos from numerous public face spoof databases with varying quality, resolutions, and dynamic ranges. We also compared our results to those of a number of current state-of-the-art approaches. The following are the main contributions of this work:

- Improving face spoofing attack detection using BS that discriminates the motion of real face from a fake one.
- Fine-tuning ResNet-50 model for face spoofing detection task to extract meaningful deep facial features..
- Using the MV algorithm to increase the classification rate of the system which is clearly observed when the methodology outperformed previous methodologies in the literature, according to the results of our experiments.
- Tackling the sensor interoperability problem by including the experiments of inter-database and intra-database tests.

The remainder of the paper is structured as follows. Section 2 describes related work on face anti-spoofing. Then, our approach is described in detail in Section 3. Section 4 summarizes the experimental results and provides a comparative analysis. The section also describes the databases that we used in our tests. Section 5 draws some conclusions and highlight some future directions.

2. Related Work

Spoof attacks can be detected in a variety of ways. In this paper, we will only look at two types of face anti-spoofing methods: handcrafted and deep learning-based methods. In this section, We present most previous work in face anti-spoofing approaches. However, we only focus on those that are thematically closer to our goals and contributions.

2.1. Handcraft based techniques

Texture features, which can describe the contents and details of a specific region in an image, are an important low-level feature in face anti-spoofing methods. Therefore, the analysis of image texture information is used in many techniques such as compressed sensing which preserves texture information and denoising at the same time [7,8]. These techniques based on handcrafted features provide accurate features that increase the detection rate of a spoofing system. Smith *et al.* [9] proposed a method for countering attacks on face recognition systems by using the color reflected from the user's face as displayed on mobile devices. The presence or absence of these reflections can be utilized to establish whether or not the images were captured in real time. The algorithms use simple RGB images to detect spoof attacks. These strategies can be classified into two categories: static and dynamic approaches. The static is used on a single image, whilst dynamic is used on the video.

The majority of approaches for distinguishing between real and synthetic faces are focused on texture analysis. Arashloo *et al.* [10] combined two spatial-temporal descriptors using kernel discriminant analysis fusion. They are Multiscale Binarized Statistical Image Features on Three Orthogonal Planes (MBSIF-TOP) and Multiscale Local Phase Quantization on Three Orthogonal Planes (MLPQ-TOP). To distinguish between real and fake individuals, Pereira *et al.* [11] also experimented with a dynamic texture that was based on Local Binary Pattern on Three Orthogonal Planes (LBP-TOP). The good results of LBP-TOP are due to the fact that temporal information is crucial in face anti-spoofing. Tirunagari *et al.* [12] used Local Binary Patterns (LBP) for dynamic patterns and Dynamic Mode Decomposition (DMD) for visual dynamics. Wen *et al.* [13] proposed an Image Distortion Analysis-based method (IDA). To represent the face images, four different features were used: blurriness, color diversity, specular reflection and chromatic moments, also relying on the features that can detect differences between real image and fake one without capturing any information about the user's identity. Patel *et al.* [14]investigated the impact of different RGB color channels (R, G, B, and Gray Scale) and different facial regions on the performance of LBP and Dense Scale Invariant Feature Transform (DSIFT) based algorithms. Their investigations have revealed that extracting the texture from the red channel produces the best results. Boulkenafet *et al.* [15] proposed a color texture analysis-based face anti-spoofing approach. They employed the LBP descriptor to extract texture features from each channel after encoding the RGB images in two color spaces: HSV and YCbCr, and then concatenated these features to distinguish between real and fake faces.

Some methods, such as [16], have recently used user-specific information to improve the performance of texture-based FAS techniques. Garcia *et al.* [17] proposed face spoofing detection by looking for Moiré patterns caused by digital grid overlap where their detection is based on frequency domain peak detection. For classification, they used Support Vector Machine (SVM) with an radial basis function kernel. They started to run their tests on the Replay Attack Corpus and Moiré databases. Other face anti-spoofing solutions are based on textures on 3D models, such as those used in [18]. Because the attacker in 3D models utilizes a mask to spoof the system, the introduction of wrinkles might be extremely helpful in detecting the attack. The presented work in [18] examines the viability of performing low-cost assaults on 2.5D and 3D face recognition systems using self-manufactured three-dimensional (3D) printed models.

2.2. Deep learning based techniques

Actually, deep Learning is used in a variety of systems and applications for biometric authentication[19] where the deep network can be trained using a number of patterns. After learning all of the dataset's unique features, the network can be used to identify similar patterns. Deep learning approaches have mostly been used to learn facial spoofing detection features. Also, Deep Learning is efficient at classification (supervised learning) and clustering tasks (unsupervised learning). Thus, the system assigns class labels to the input instances in a classification task, but the instances in clustering approaches are clustered based on their similarity without the usage of class labels.

To train models with significant discriminative abilities, Yang *et al.*[20] used a deep CNN rather than manually constructing features from the scratch. Quan *et al.* proposed a semi-supervised learning-based architecture to fight face spoofing threats using only a few tagged data, rather than depending on time-consuming data annotations. They assess the reliability of selected data pseudo labels using a temporal consistency requirement. As a result, network training is substantially facilitated. Also, by progressively increasing the contribution of unlabeled target domain data to the training data, an adaptive transfer mechanism can be implemented to eliminate domain bias. According to the authors in [21], they use a type of Ground Through (GT) termed appr-GT in conjunction with the identity information of the spoof image to generate a genuine image of the appropriate subject in the training set. A metric learning module constrains the generated genuine images from the spoof images to be near the appr-GT and far from the input images. This reduces the effect of changes in the imaging environment on the appr-GT and GT of a spoof image.

Jia *et al.* [22] proposed a Unified unsupervised and Semi-supervised Domain Adaptation Network (USDAN) for cross-scenario face anti-spoofing, with the purpose of reducing the distribution mismatch between the source and target domains. The marginal Distribution Alignment Module (MDA) and the conditional distribution alignment module (CDA) are two modules that use adversarial learning to find a domain-invariant feature space and condense features of the same class.

Raw optical flow data from the clipped face region and the complete scene were used to train a neural network by feng's team *et al.* [23]. Motion-based anti-spoofing does not need a scenic model or motion assumption to generalize. They present an image quality-based and motion-based liveness framework that can be fused together using a hierarchical neural network.

In their work [24], Liu *et al.* proposed a Deep Tree Network (DTN) that learns characteristics in a hierarchical form and may detect unanticipated spoofing attacks by identifying the features that are learned.

Yu *et al.* [25] introduces two new convolution and pooling operators for encoding fine-grained invariant information: Central Difference Convolution (CDC) and Central Difference Pooling (CDP). CDC outperforms vanilla convolution in extracting intrinsic spoofing patterns in a number of situations.

As described in Qin *et al.* [26], Adaptive Inner-Update (AIU) is a novel meta learning approach that uses a meta-learner to train on zero- and few-shot FAS tasks utilizing a newly constructed Adaptive Inner update Meta Face Anti spoofing (AIM-FAS).

According to Yu *et al.* [27], the Multi-level Feature Refinement Module (MFRM) and material-based multi-head supervision can help increase BCN's performance. In the first approach, local neighborhood weights are reassembled to create multi-scale features, while in the second, the network is forced to acquire strong shared features in order to perform tasks with multiple heads.

CDC-based frame-level FAS approaches, proposed by the authors in [28], have been developed. These patterns can be captured by aggregating information about intensity and gradient. In comparison to a vanilla convolutional network, the Central Difference Convolutional Network (CDCN) built with CDC has a more robust modeling capability. CDCN++ is an improved version of CDCN that incorporates the search backbone network with the Multiscale Attention Fusion Module (MAFM) for collecting multi-level CDC features effectively.

Spatio-Temporal Anti-Spoof Network (STASN) is a new attention mechanism invented by the Yang *et al.* [29] that combines global temporal and local spatial information, allowing them to examine the model's understandable behaviors.

To improve CNN generalization, Liu *et al.* [30] proposed to use innovative auxiliary information to supervise CNN training. A new CNN-RNN architecture for learning the depth map and rPPG signal from end to end is also proposed.

Wang *et al.* [31] proposed a depth-supervised architecture that can efficiently encode spatiotemporal information for presentation attack detection and develops a new approach for estimating depth information from several RGB frames. Short-term extraction is accomplished through the use of two unique modules: the Optical Flow guided Feature Block (OFFB) and the convolution gated recurrent units (ConvGRU). Jourabloo *et al.* [32] proposed a new CNN architecture for face de-spoofing, with appropriate constraints and supplementary supervisions. to discern between living and fake faces, as well as long-term motion. In order to detect presentation attacks effectively and efficiently, Kim *et al.* [33] introduced Bipartite Auxiliary Supervision Network (BASN), an architecture that learns to extract and aggregate auxiliary information.

Huszár *et al.* [34] proposed a Deep Learning (DL) approach to address the problem of spoof attacks occurring from video. The approach was tested in a new database made up of several videos of users juggling a football. Their algorithm is capable of running in parallel with the Human Activity Recognition (HAR) in real-time. Roy *et al.* [35] proposed an approach called Bi-Directional Feature Pyramid Network (BiFPN) to detect spoof attacks because the approach containing high-level information demonstrates negligible improvements. Ali *et al.* [36] based on stimulating eye movements by using the use of visual stimuli with randomized trajectories to detect spoof attacks. Ali, *et al.* [37] by the combination of two methods which are head-detection algorithm and deep neural network-based classifiers. The test was various face presentation attacks in thermal infrared in various conditions.

3. PROPOSED APPROACH

Figure 1 describes the overall structure of our proposed approach, which is divided into three modules: background subtraction, feature learning, and data classification. To begin, we use the background subtraction between consecutive frames to extract motion, we can also call this technique BS. Then, the features are extracted using the ResNet-50 Transfer Learning model on the foreground of BS. Finally, to distinguish between real and fake faces of each frame we use a classification layer that employs a fully connected layer. After that, we use MV to predict the input video is real or not. In the subsections that follow, all subsystems (modules) will be discussed.

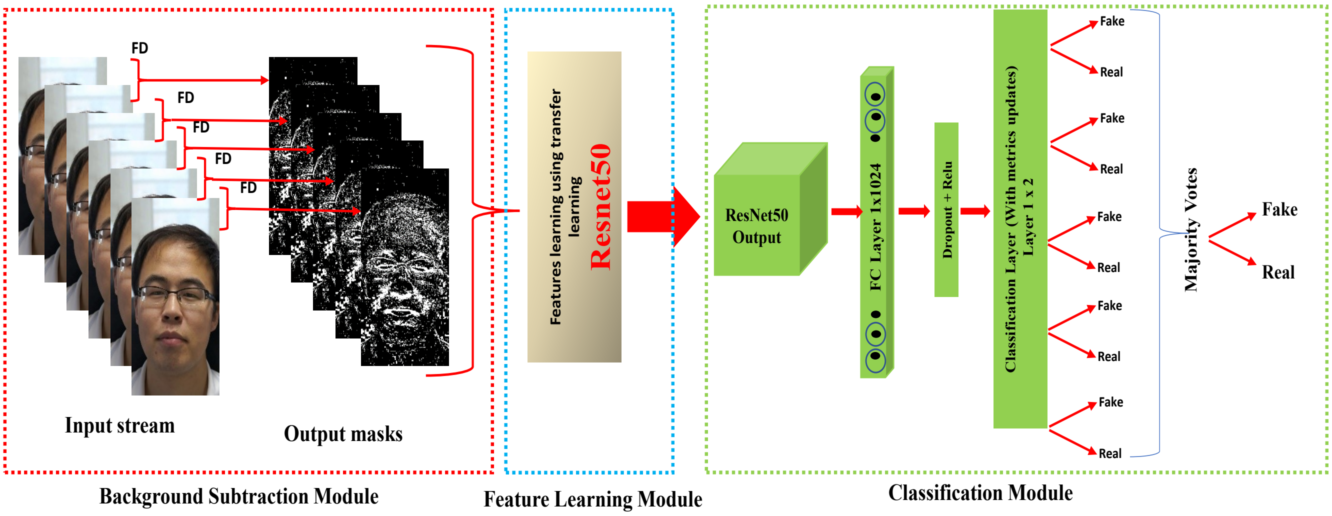


Figure 1. Framework of our proposed approach.

3.1. Background Subtraction Module

In our research work, we use a face spoofing detection system based on an extended BS algorithm. The Background Subtraction approach, which is based on the premise of getting the pixels in the image sequence difference operation to do two or three continuous frames, is the most commonly used action target detection measure. Using an image pixel value obtained by subtracting the difference image and the binarized difference image, if the pixel value change threshold is less than a predefined one, we can feel this as a background pixel in the adjacent frame. If the pixel value of an image area changes dramatically, it is possible to deduce that this is due to the action of detecting spoof in the image caused by these symbols as foreground pixel regions. While taking dynamic information into account, a pixel region based on symbolic actions can determine the position of the target in the image.

Background Subtraction is applied to images and the thresholded result is displayed as a foreground image. Figure 2 shows an example of the output. This is a low-cost and ineffective method of detecting motion in a video stream. The image P_t is transformed into a grey-scale (intensity) image I_t. Then, given the image I_t and the previous image I_{t-1}, the current output is R_t, where:

$$R_t(x,y) = \begin{cases} I_t(x,y) & \text{if } |I_t(x,y) - I_{t-1}(x,y)| > T \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

T is the value of the threshold parameter. In our situation, we just utilized a threshold to remove the pixels with the same values across the two frames. The foreground pixels take the value of the current frame if there is motion. The foreground pixel is set to zero if there is no motion.

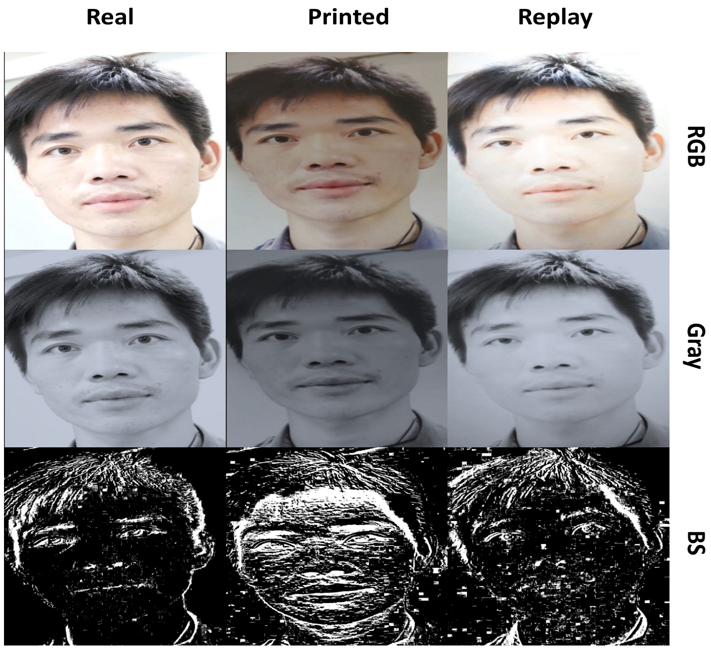


Figure 2. Example of a genuine face and corresponding print and replay attacks in grey-scale and BS.

3.2. Feature Learning Module

Feature learning (FL) is a set of approaches in machine learning that allows a system to discover the representation needed for feature detection, prediction, or classification from a preprocessed dataset automatically. This enables a machine to learn the features and apply them to a specific task like classification and prediction. Feature learning can be achieved in deep learning by either creating a complete CNN to train and test the collection of images or adapting a pre-trained CNN for classification or prediction for the new images-set. Transfer learning is the latter strategy used in the deep learning domain. Transfer learning is a machine learning technique in which a model created for one task is utilized as the basis for a model on a different task.

Transfer Learning (TL) is commonly used in DL applications to allow you to use a pre-trained network for solving new classification tasks. To meet the new learning tasks, the learning parameters of the pre-trained network with randomly initialized weights must be fine-tuned. Transfer learning is typically considerably faster and easier to learn/train than building a network from the initial concept. Transfer learning is an optimization and a quick way that can save time or improve efficiency.

In this section, a transfer learning technique is applied by fine-tuning a pretrained ResNet-50 model on ImageNet dataset using multiple spoofing datasets where the output of the last FC layer is changed to output two classes (real/fake). The network called ResNet-50 due to the fact that it has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer, and it introduced the use of residual blocks.

3.3. Classification Module

Data classification is a vital process for separating large datasets into classes for decision-making, pattern detection, and other purposes. For multi-class classification problems with mutually exclusive classes, a classification layer uses a fully connected layer to compute the cross-entropy loss.

The features from ResNet-50 are passed via a FC layer made of 1024 neurons with a 40% dropout to prevent over-fitting in the classification module. Having followed that, the units were activated with a rectification mechanism called $ReLU$. $MAX(X, 0)$ is the $ReLu$ function, which sets all negative values in the matrix X to zero while keeping all other values constant. The reason for choosing $ReLU$ is that deep network training with $ReLU$ tended to converge considerably faster and more reliably than deep network training with $sigmoid\ activation$. Finally, the output layer consisted of one neuron unit configured with the $Sigmoid\ function$ to generate probabilities for the classes (Binary classifier). $Sigmoid$ is a mathematical function that takes a vector of k real values as an input and converts it to a probability distribution with two probabilities.

We employ Voting Ensemble in our tests to classify each subject (video) as real or fake. A voting ensemble (sometimes known as a "majority voting ensemble") is a machine learning model that incorporates predictions from several other models, such as multiple predictions in each frame after the input video's last layer (Classification layer). The predictions for each label are combined, and the label with the majority vote is forecasted (See Fig. 1, classification module) to determine if the input video belongs to a real or fake one. The majority voting ensemble creates forecasts based on the most common one. It's a strategy that can be utilized to boost performance, with the goal of outperforming every frame used independently in the ensemble.

4. Experimental results and analysis

In this section, the employed benchmark datasets will be introduced first, followed by a brief description of the evaluation criteria. After that, we present and analyze a series of experiments that we assume demonstrate the efficacy of the proposed BS-CNN+MV based face spoofing detection technique.

4.1. Database and protocol

In order to assess of the effectiveness of our proposed anti spoofing technique, we performed a set of experiments on a well known databases where most three challenging databases were used: The CASIA-FASD [1] Face Anti-Spoofing database, Replay-Attack [2] database and MSU [3] Mobile Face Spoofing databases. Those databases contain video recording of real and fake attacks. A brief description of these databases is given as fellow:

The CASIA-FASD database[38] is a data-set for face anti spoofing detection. This database contains 50 genuine subjects in total and the corresponding fake faces are captured with high quality from the original ones.Therefore each subject contains 12 videos (3 genuine and 9 fake) under three different resolutions and light conditions namely the low quality, normal quality and high quality. Also, three fake face attacks are designed, which include warped photo attack, cut photo attack and video attack. The overall database contains 600 video clips and the subjects are divided into subsets for performing training and test in which 240 videos of 20 subjects are used for training and 360 videos of 30 subjects for testing. Test protocol is provided, which consists of 7 scenarios for a thorough evaluation from all possible aspects see fig 3.

Among the popular databases designed for the anti spoofing application, one can find the Replay-Attack database [39]. This database consists of 1300 video of real-access and attack attempts to 50 subjects, (See Fig 5). However, These video were taken using a built-in webcam on a Macbook laptop under two separate scenarios (controlled and adversed). In addition, Two cameras were used to create the faked facial attack for each person in high-resolution images and videos: a Canon PowerShot SX150 IS and an iPhone 3GS camera. Also, Fixed attacks and hand attacks are the two types of attacks. There are ten videos in each subsets: 4 mobile attacks with a resolution of 480 × 320 pixels on

[1] http://www.cbsr.ia.ac.cn/english/FaceAntiSpoofDatabases.asp
[2] https://www.idiap.ch/dataset/replayattack
[3] https://drive.google.com/drive/folders/1nJCPdJ7R67xOiklF1omkfz4yHeJwhQsz

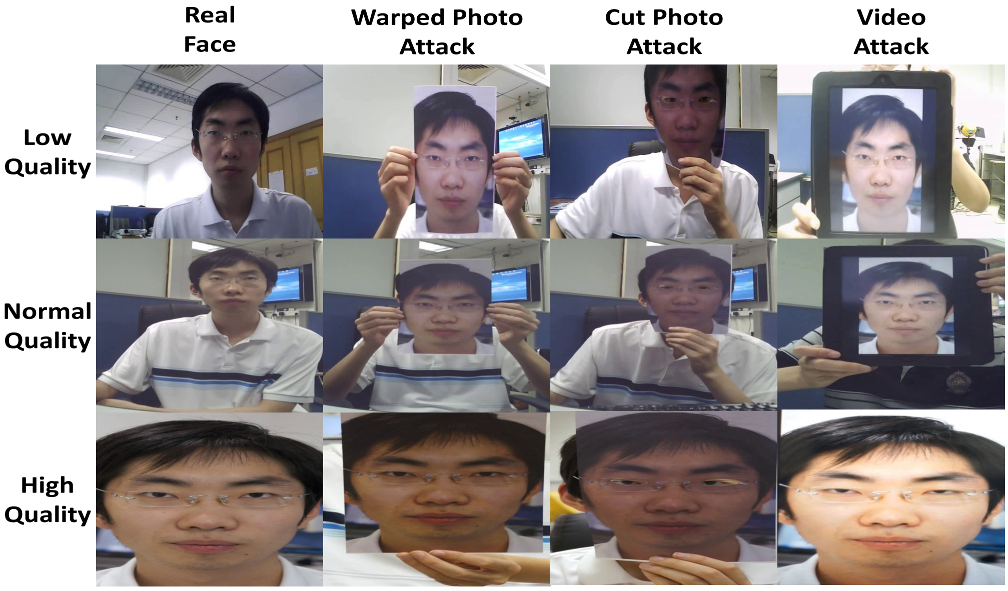


Figure 3. Samples from the CASIA FAS database.

an iPhone 3GS screen, then, by using an iPad first generation with a screen resolution of 1024 × 768 pixels, four high-resolution screen attacks were performed. On A4 paper, two hard-copy print attacks (printed on a Triumph-Adler DCC 2520 colour laser printer) occupied the whole available printing surface. It will be noted that the complete set of video is divided into three non-overlapping subsets for training, development, and testing in order to evaluate them.

The Patterns Recognition and Image Processing (PRIP) group at Michigan State University developed a publicly available MSU-MFSD database for face spoof attacks. The database contains 280 video clips of attempted photo and video attacks on 35 clients. It was created using a mobile phone to capture both genuine and spoof attacks. This was accomplished using two types of cameras: 1) the built-in camera in the MacBook Air 13 inch (640 × 480) and 2) the front-facing camera on the Google Nexus 5 Android phone (720 × 480). Each subject received two video recordings, the first of which was taken using a laptop camera and the second with an Android camera (See Fig 4). High-resolution video was recorded for each subject utilizing two devices to create the attacks:1) Canon PowerShot 550D SLR camera, which captures 18.0 Megapixel photos and 1080p high-definition video clips; 2) iPhone 5S back-facing camera, which captures 1080p video clips. There are three types of spoof attack, the first one 1) high-resolution replay videoThe first type of spoof attack is a high-resolution replay video attack using an iPad Air screen, with a resolution of 2048 × 1536, the second is a mobile phone replay video attack using an iPhone 5S screen, with a resolution of 1136 × 640, and the third is a printed photo attack using an A3 paper with a fully-occupied printed photo of the client's biometry, with a paper size of: 11 × 17 ($279mm \times 432mm$), printed with an HP Colour Laserjet CP6015xh printer at a resolution of 1200 × 600 dpi. Finally, to assess performance, the 35 subjects in the MSU-MFSD database were divided into two subsets: 15 for training and 20 for testing.

4.2. Evaluation metrics

The comparative results for cross-scenario testing are expressed in terms of the HTER, which is the mean of the False Acceptance Rate (FAR) and False Rejection Rate (FRR). On the development set, we first compute the EER and the corresponding threshold, and then use the threshold to determine the HTER on the testing set. Additionally, the Receiver Operating Characteristic (ROC) is reported to assess the method's performance. We use HTER for the Idiap Replay-Attack dataset and EER for the CASIA-FASD and MSU-MFSD

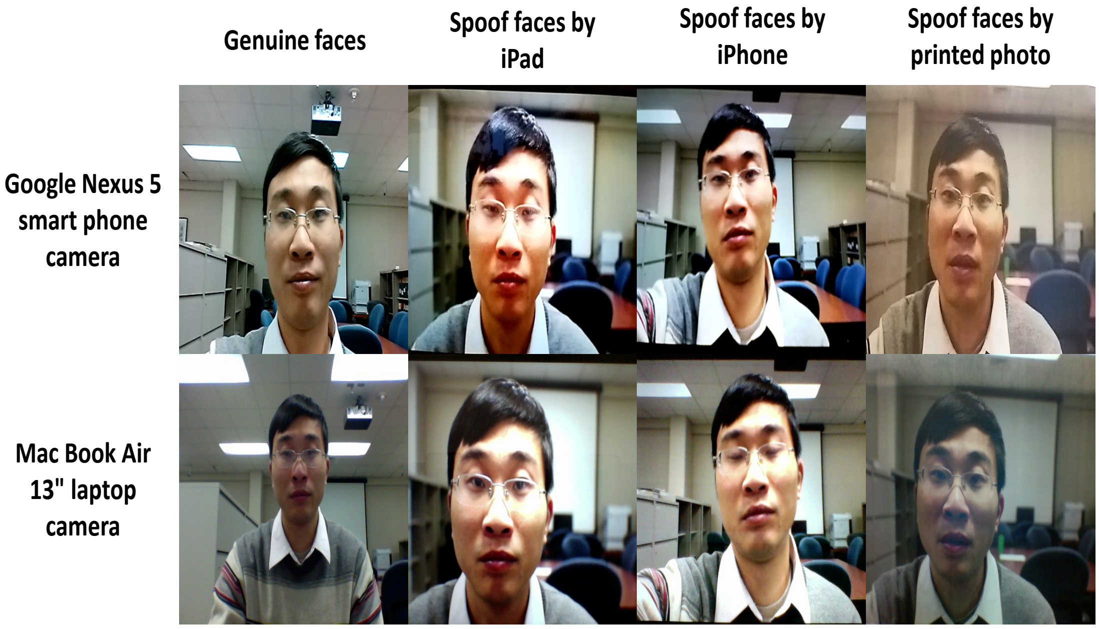

Genuine faces
Spoof faces by iPad
Spoof faces by iPhone
Spoof faces by printed photo
Google Nexus 5 smart phone camera
Mac Book Air 13" laptop camera

Figure 4. Example images of genuine and spoof faces of one of the subjects in the MSU-MFSD database.

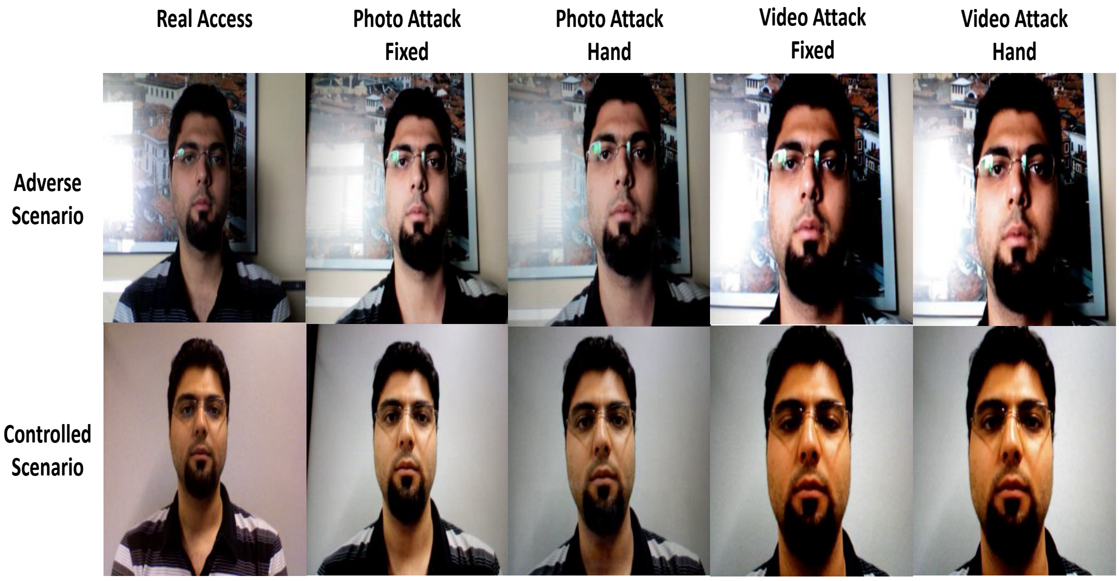


Figure 5. Examples of real accesses and attacks in different scenarios.

datasets for intra-scenario testing. We employ Area Under Curve (AUC) as a performance metric for type-scenario testing. 332 333

4.3. *Performance comparison on intra-database* 334

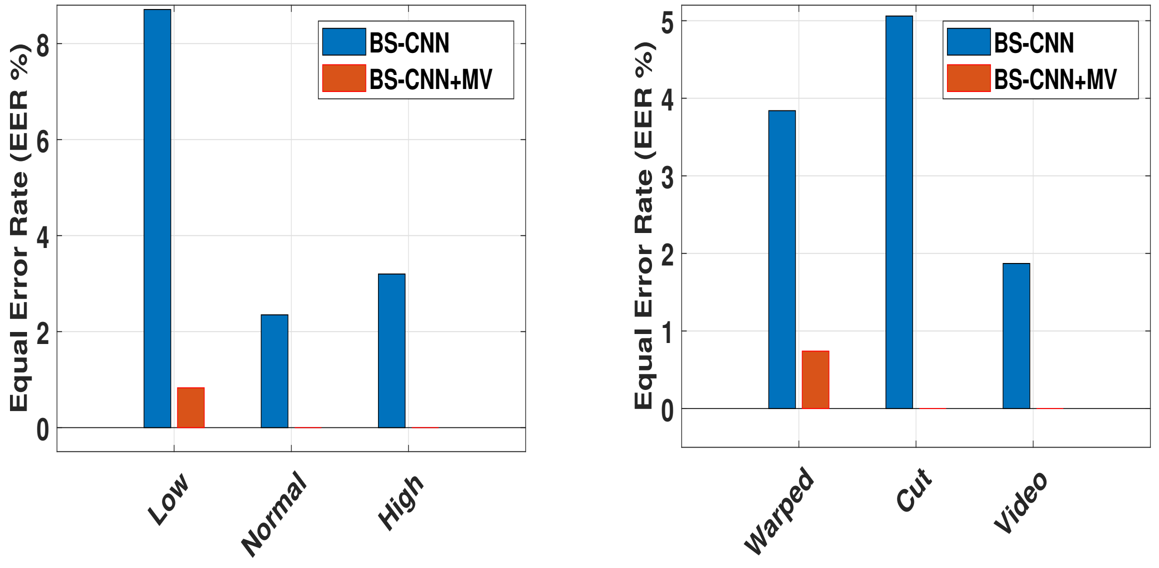
We computed the EERs for the seven scenarios, including different qualities and media, to meet the official CASIA Face Anti-Spoofing test protocol. Low, normal, and high-quality image sequences are provided as quality descriptors, and the used media for spoofing attacks are warped images, chopped photos, and videos played on an iPad. The last scenario is the overall test, which will look at how image quality and spoofing media affect system performance. In this part, we computed two tests, the first of which was performed per-frame and the second of which was performed per-video. The first test uses 335 336 337 338 339 340 341

Table 2. Comparison EER (in %) between the proposed approach and the state-of-the-art methods on different scenario on CASIA FAS.

	Scenarios						
Methods	**Low**	**Normal**	**High**	**Warped**	**Cut**	**Video**	**Overall**
IQA [40]	31.70	22.20	05.60	26.10	18.30	34.40	32.40
DoG baseline [38]	13.00	13.00	26.00	16.00	06.00	24.00	17.00
visual codebooks [41]	10.00	17.78	13.33	07.78	22.22	08.89	14.07
LBP-overlapping+fisher [42]	07.20	08.80	14.40	12.00	10.00	14.70	13.10
CDD [43]	01.50	05.00	02.80	06.40	04.70	00.30	11.80
ML-LPQ fisher [44]	12.49	08.96	05.22	13.62	09.66	10.10	11.39
LBP-TOP [11]	10.00	12.00	13.00	06.00	12.00	10.00	10.00
FD-ML-BSIF-FS [6]	07.93	11.85	12.42	05.85	03.11	15.84	09.96
MLLBP + MLBSIF [45]	006.56	09.93	007.36	09.98	03.45	10.04	09.81
Kernel Fusion [10]	00.70	08.70	13.00	01.40	10.10	04.30	07.20
YCbCr+HSV-LBP [15]	07.80	10.10	06.40	07.50	05.40	08.10	06.20
Identity-DS [21]	-	-	-	-	-	-	03.30
USDAN-Norm [22]	-	-	-	-	-	-	01.10
S-CNN+PL+TC [46]	-	-	-	-	-	-	00.69
BS-CNN+MV (Ours)	**00.83**	**00.00**	**00.00**	**00.74**	**00.00**	**00.00**	**00.00**

BS-CNN to determine whether an image is real or fake, while the second test uses MV to determine whether a video is real or not (See Table 2). In addition, we discovered that the proposed method (BS-CNN+MV) improves the performance.

Moreover, we have found that combining the MV with BS-CNN yields the best results for picture quality (low, normal, and high), as well as with spoof media (warped photo, cut photo and video attacks) (See Fig 6). This can be explained by the fact that MV improves decision-making performance. Table 2 shows that our proposed technique outperforms the CASIA baseline in all scenarios when compared to the database created by CASIA [38].



(a) Quality

(b) Spoofing Media

Figure 6. Effect of Quality and Spoofing Media on the Performance on the CASIA-FASD. (a) Quality and (b) Spoofing Media

The suggested method is compared to state-of-the-art methods in Table 4 using data from the Replay-Attack database. Despite the fact that our EER and HTER are similar to the previous approaches. The following are the types of attacks on replay databases that we compute the performance of using our method: Depending on the method used to hold the

Table 3. Testing our proposed countermeasure using all scenarios of the REPLAY-ATTACK database.

		BS-CNN+MV (our)	
		EER	**HTER**
Scenarios	**Digitalphoto**	01.25	01.87
	Highdef	01.42	03.43
	Mobile	00.00	00.31
	Photo	00.53	02.50
	Print	00.83	00.62
	Video	00.00	01.56
	Overall	00.58	00.62

Table 4. Comparison between the proposed countermeasure and the state-of-the-art methods on REPLAY-ATTACK database

Methods	overall	
	EER	**HTER**
IQA [40]	00.00	15.20
LBP [47]	13.90	13.87
MotionCorrelation [48]	11.78	11.79
LBP-TOP [11]	07.90	07.60
IDA [13]	08.58	07.41
Motion+LBP [49]	04.50	05.11
FD-ML-LPQ-Fisher [6]	05.62	04.80
DMD [12]	05.30	03.75
Colour-LBP [15]	00.40	02.90
Spectral cubes [41]	-	02.75
CNN [20]	06.10	02.10
USDAN-Norm [22]	-	00.30
Bottleneck Feature Fusion + NN [23]	00.83	**00.00**
Identity-DS [21]	00.20	00.00
S-CNN+PL+TC [46]	0.36	-
BS-CNN+MV (our)	00.58	00.62

attack replay device (paper, mobile phone, or tablet), the three attack subsets (print, mobile, and highdef) were recorded in two different modes: i) fixed-support and ii) hand-based (See Tables 3). We also put our method to the test using the MSU-MFSD database. (See Table 5). It will be noted that there is no articles have been published that detail the results of various sorts of attacks on this database. We can see that our results are better to the state of the art, with our BS-CNN+MV providing the best results.

The final comparison results on intra-database that are shown in (Tables: 2, 4 and 5), which indicates that our proposed method achieves much lower errors on all three datasets than other state-of-the-art methods. Meanwhile, from Figure 7 it can be observed that the BS-CNN+MV perform better than the BS-CNN, which further verify the effectiveness of the proposed background subtraction based convolution neural network.

Table 5. Comparison EER (in %) between the proposed approach and the state-of-the-art methods on different scenario on MSU-MFSD.

	Scenarios						
Methods	**HD Android**	**HD Laptop**	**Mobile Android**	**Mobile Laptop**	**Print Android**	**Print Laptop**	**Overall**
IDA [13]	-	-	-	-	-	-	08.58
Identity-DS [21]	-	-	-	-	-	-	08.58
FD-ML-BSIF-FS [6]	-	-	-	-	-	-	02.10
S-CNN+PL+TC [46]	-	-	-	-	-	-	00.64
USDAN-Norm [22]	-	-	-	-	-	-	00.00
BS-CNN+MV (our)	**00.00**	**00.00**	**00.00**	**00.00**	**00.00**	**00.00**	**00.00**

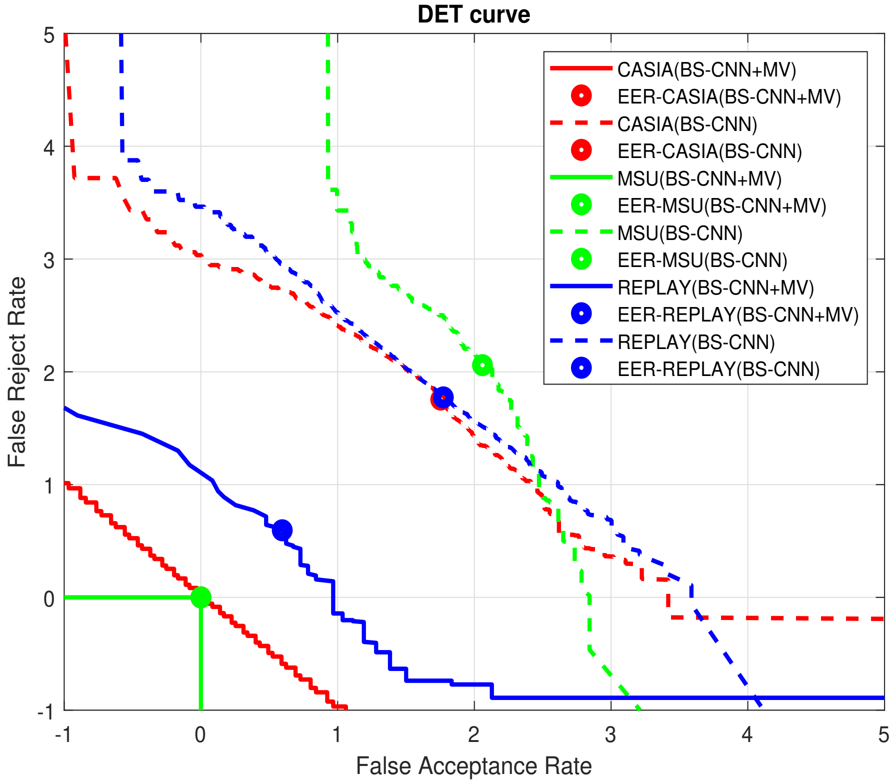


Figure 7. DET curve of the proposed approach on CASIA, MSU and REPLAY databases.

Table 6. AUC (%) of the model cross-type testing on CASIA-FASD, Replay-Attack, and MSU-MFSD.

Methods	CASIA-FASD			Replay-Attack			MSU-MFSD			overall
	Video	Cut photo	Wrapped	Video	Digital Photo	Printed	Printed	HR Video	Mobile Video	
OC-SVM+BSIF [50]	70.74	60.73	95.90	84.03	88.14	73.66	64.81	87.44	74.69	78.68 ± 11.74
NN+LBP [51]	94.16	88.39	79.85	99.75	95.17	78.86	50.57	99.93	93.54	86.69 ± 16.25
SVM+LBP [52]	91.94	91.70	84.47	99.08	98.17	87.28	47.68	99.50	97.61	88.55 ± 16.25
NAS-Baseline [25]	96.32	94.86	98.60	99.46	98.34	92.78	68.31	99.89	96.76	93.90 ± 09.87
DTN [24]	90.00	97.30	97.50	99.90	99.90	99.60	81.60	99.90	97.50	95.90 ± 06.20
AIM-FAS [26]	93.6	99.7	99.1	99.8	99.9	99.8	76.3	99.9	99.1	96.40 ± 07.80
CDCN [28]	98.48	99.90	99.80	100.00	99.43	99.92	70.82	100.00	99.99	96.48 ± 09.64
CDCN++ [28]	98.07	99.90	99.60	99.98	99.89	99.98	72.29	100.00	99.98	96.63 ± 09.15
BCN [27]	99.62	100.00	100.00	99.99	99.74	99.91	71.64	100.00	99.99	96.77 ± 09.99
NAS-FAS [25]	99.62	100	100	99.99	99.89	99.98	74.62	100.00	99.98	97.12 ± 08.94
BS-CNN+MV (our)	100	100	**99.98**	**100**	**100**	**100**	**100**	**100**	**100**	**99.99 ± 0.0067**

4.4. Inter-Dataset Cross-Type Testing

In this part of experiments, the CASIA-FASD, Replay-Attack, and MSU-MFSD are used to perform the intra-dataset cross-type testing between replay and print attacks . As shown in Table 6, our proposed method outperforms state-of-the-art methods in terms of overall performance (99.99% AUC), indicating that learned features extended well to unknown attacks. As a result of this, it appears that our method can learn intrinsic material patterns from a wide range of materials and so generalizes well to previously unexplored types of material.

4.5. Inter-Dataset Cross-dataset Testing

This experiment includes six cross-dataset testing protocols. The first is that we perform CASIA-FASD training and testing on Replay-Attack, which is known as protocol CR; the second is that we perform CASIA-FASD training and testing on MSU-MFSD, which is known as protocol CM; and the third is that we exchange the training and testing datasets that we have in the first, which is known as protocol RC. The rest of the protocols

Table 7. The results of cross-dataset testing between CASIA-FASD, MSU-MFSD and Replay-Attack. The evaluation metric is HTER(%)

Method	**Protocol CR**		**Protocol CM**		**Protocol RC**		**Protocol RM**		**Protocol MC**		**Protocol MR**	
	Training	**Testing**	**Training**	**Testing**	**Training**	**Testing**	**Training**	**Testing**	**Training**	**Testing**	**Training**	**Testing**
	Casia	**Replay**	**Casia**	**MSU**	**Replay**	**Casia**	**Replay**	**MSU**	**MSU**	**Casia**	**MSU**	**Replay**
FD-ML-LPQ-FS [6]	50.25		50.41		42.59		38.00		50.00		48.00	
Motion-Mag [53]	50.10		NP		47.00		NP		NP		NP	
LBP-TOP [54]	49.70		NP		60.60		NP		NP		NP	
LBP [15]	47.00		NP		39.60		NP		NP		NP	
Spectral cubes [41]	34.40		NP		50.00		NP		NP		NP	
STASN [29]	31.50		NP		30.90		NP		NP		NP	
Color Texture [55]	30.30		NP		37.70		NP		NP		NP	
FaceDs [32]	28.50		NP		41.10		NP		NP		NP	
Auxiliary [30]	27.60		NP		28.40		NP		NP		NP	
MEGC [56]	20.20		NP		27.90		NP		NP		NP	
FAS-TD [31]	17.50		NP		24.00		NP		NP		NP	
BASN [33]	17.50		NP		24.00		NP		NP		NP	
Patch+BCN+MFRM [27]	16.60		NP		36.40		NP		NP		NP	
CDCN [28]	**15.50**		NP		32.60		NP		NP		NP	
BS-CNN+MV (our)	17.62		**23.75**		**20.35**		**24.16**		**35.45**		**44.33**	

are identical, with the exception that we utilize the data to train one time and test the next; the protocols are protocol RM, protocol MC, and protocol MR. As may be seen in Table 7. On protocol CR, our suggested BS-CNN+MV has 17.62% HTER, exceeding the previous state-of-the-art by a convincing margin of 2%. Increasing the size of the training set with data augmentation could increase performance even more. For protocol RC, we also outperform state-of-the-art frame-level techniques (see Table 7, third column). Furthermore, we can see in the same table for our suggestion that the convincing margin between protocols RC and CR is 3% in the same technique, compared to other most convincing methods in the same protocol, such as in [28] (17%). As a result, we can presume that our approach outperforms current approaches.

5. CONCLUSION AND FUTURE DIRECTIONS

Fake face detection is a problem that has been addressed in this work. We analyzed seven scenarios from the MSU-MFSD, the REPLAY-ATTACK, and the CASIA-FASD databases. In fact, texture and motion-based characteristics were used by the majority of authors in the field of face spoof detection. However, BS and a CNN with a majority vote seems to determine well if a person is using a fake face. In our paper, we evaluated our approach under different protocols. Firstly, we used all three types of databases to evaluate if they produced satisfactory results when compared to the current state of the art. The proposed technique is then put to the test using Cross-Type Testing to ensure that it can handle all attributes and attacks across the three databases. In the final test, we used Cross-dataset Testing to compare each train of any data with the test to other data in order to improve the validity of our approach. The obtained results have shown that our proposed methods outperform the current state-of-the-art. As a future direction, face spoof detection research could focus on making the system more robust across all databases. It is also of interest to create a common training model for each face spoof detection using the transformer method.

Author Contributions “Conceptualization, A.B. and S.B.; methodology, A.B. and A.T-A.; software, A.B. and SE.B.; validation, A.B., A.T-A. and SE.B.; formal analysis, A.B. and M.K.; investigation, A.H., A.T-A. and A.B.; data curation, M.K. and K.B.; writing—original draft preparation, A.B., M.K. and SE.B.; writing—review and editing, A.O. and A.O.; visualization, A.B., SE.B. and A.T-A.; supervision, A.T-A.; project administration, A.B.; All authors have read and agreed to the published version of the manuscript.”

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, Present, and Future of Face Recognition: A Review. *Electronics* **2020**, *9*. doi:10.3390/electronics9081188.
2. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Jacques, S. Multi-Block Color-Binarized Statistical Images for Single-Sample Face Recognition. *Sensors* **2021**, *21*. doi:10.3390/s21030728.
3. El Morabit, S.; Rivenq, A.; Zighem, M.E.n.; Hadid, A.; Ouahabi, A.; Taleb-Ahmed, A. Automatic Pain Estimation from Facial Expressions: A Comparative Analysis Using Off-the-Shelf CNN Architectures. *Electronics* **2021**, *10*. doi:10.3390/electronics10161926.
4. Khaldi, Y.; Benzaoui, A.; Ouahabi, A.; Jacques, S.; Taleb-Ahmed, A. Ear Recognition Based on Deep Unsupervised Active Learning. *IEEE Sensors Journal* **2021**, *21*, 20704–20713. doi:10.1109/JSEN.2021.3100151.
5. Benlamoudi, A. Multi-Modal and Anti-Spoofing Person Identification. PhD dissertation, University of Kasdi Merbah, Ouargla, Faculty of New Technologies of Information and Communication (FNTIC), Department of electronics and telecommunications, 2018.
6. Benlamoudi, A.; Aiadi, K.E.; Ouafi, A.; Samai, D.; Oussalah, M. Face antispoofing based on frame difference and multilevel representation. *Journal of Electronic Imaging* **2017**, *26*, 043007. doi:10.1117/1.JEI.26.4.043007.
7. Mahdaoui, A.E.; Ouahabi, A.; Moulay, M.S. Image Denoising Using a Compressive Sensing Approach Based on Regularization Constraints. *Sensors* **2022**, *22*. doi:10.3390/s22062199.
8. Haneche, H.; Boudraa, B.; Ouahabi, A. A new way to enhance speech signal based on compressed sensing. *Measurement* **2020**, *151*, 107117. doi:https://doi.org/10.1016/j.measurement.2019.107117.
9. Smith, D.F.; Wiliem, A.; Lovell, B.C. Face recognition on consumer devices: Reflections on replay attacks. *Information Forensics and Security, IEEE Transactions on* **2015**, *10*, 736–745.
10. Arashloo, S.R.; Kittler, J.; Christmas, W. Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features. *Information Forensics and Security, IEEE Transactions on* **2015**, *10*, 2396–2407.
11. de Freitas Pereira, T.; Komulainen, J.; Anjos, A.; De Martino, J.M.; Hadid, A.; Pietikäinen, M.; Marcel, S. Face liveness detection using dynamic texture. *EURASIP Journal on Image and Video Processing* **2014**, *2014*, 1–15.
12. Tirunagari, S.; Poh, N.; Windridge, D.; Iorliam, A.; Suki, N.; Ho, A.T. Detection of face spoofing using visual dynamics. *Information Forensics and Security, IEEE Transactions on* **2015**, *10*, 762–777.
13. Wen, D.; Han, H.; Jain, A.K. Face spoof detection with image distortion analysis. *Information Forensics and Security, IEEE Transactions on* **2015**, *10*, 746–761.
14. Patel, K.; Han, H.; Jain, A.K.; Ott, G. Live face video vs. spoof face video: Use of moiré patterns to detect replay video attacks. In Proceedings of the Biometrics (ICB), 2015 International Conference on. IEEE, 2015, pp. 98–105.
15. Boulkenafet, Z.; Komulainen, J.; Hadid, A. face anti-spoofing based on color texture analysis. In Proceedings of the Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015, pp. 2636–2640.
16. Chingovska, I.; dos Anjos, A.R. On the use of client identity information for face antispoofing. *IEEE Transactions on Information Forensics and Security* **2015**, *10*, 787–796.
17. Garcia, D.C.; de Queiroz, R.L. Face-spoofing 2D-detection based on Moiré-pattern analysis. *IEEE transactions on information forensics and security* **2015**, *10*, 778–786.
18. Galbally, J.; Satta, R. Three-dimensional and two-and-a-half-dimensional face recognition spoofing using three-dimensional printed models. *IET Biometrics* **2016**, *5*, 83–91.
19. Bekhouche, S.E.; Dornaika, F.; Benlamoudi, A.; Ouafi, A.; Taleb-Ahmed, A. A comparative study of human facial age estimation: handcrafted features vs. deep features. *Multimedia Tools and Applications* **2020**, *79*, 26605–26622.
20. Yang, J.; Lei, Z.; Li, S.Z. Learn convolutional neural network for face anti-spoofing. *arXiv preprint arXiv:1408.5601* **2014**.
21. Xu, Y.; Wu, L.; Jian, M.; Zheng, W.S.; Ma, Y.; Wang, Z. Identity-constrained noise modeling with metric learning for face anti-spoofing. *Neurocomputing* **2021**, *434*, 149–164. doi:https://doi.org/10.1016/j.neucom.2020.12.095.
22. Jia, Y.; Zhang, J.; Shan, S.; Chen, X. Unified unsupervised and semi-supervised domain adaptation network for cross-scenario face anti-spoofing. *Pattern Recognition* **2021**, *115*, 107888. doi:https://doi.org/10.1016/j.patcog.2021.107888.
23. Feng, L.; Po, L.M.; Li, Y.; Xu, X.; Yuan, F.; Cheung, T.C.H.; Cheung, K.W. Integration of image quality and motion cues for face anti-spoofing: A neural network approach. *Journal of Visual Communication and Image Representation* **2016**, *38*, 451–460.
24. Liu, Y.; Stehouwer, J.; Jourabloo, A.; Liu, X. Deep tree learning for zero-shot face anti-spoofing. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4680–4689.
25. Yu, Z.; Wan, J.; Qin, Y.; Li, X.; Li, S.Z.; Zhao, G. Nas-fas: Static-dynamic central difference network search for face anti-spoofing. *arXiv preprint arXiv:2011.02062* **2020**.
26. Qin, Y.; Zhao, C.; Zhu, X.; Wang, Z.; Yu, Z.; Fu, T.; Zhou, F.; Shi, J.; Lei, Z. Learning meta model for zero-and few-shot face anti-spoofing. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2020, Vol. 34, pp. 11916–11923.
27. Yu, Z.; Li, X.; Niu, X.; Shi, J.; Zhao, G. Face anti-spoofing with human material perception. In Proceedings of the European Conference on Computer Vision. Springer, 2020, pp. 557–575.
28. Yu, Z.; Zhao, C.; Wang, Z.; Qin, Y.; Su, Z.; Li, X.; Zhou, F.; Zhao, G. Searching central difference convolutional networks for face anti-spoofing. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5295–5305.

29. Yang, X.; Luo, W.; Bao, L.; Gao, Y.; Gong, D.; Zheng, S.; Li, Z.; Liu, W. Face anti-spoofing: Model matters, so does data. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 3507–3516.
30. Liu, Y.; Jourabloo, A.; Liu, X. Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 389–398.
31. Wang, Z.; Zhao, C.; Qin, Y.; Zhou, Q.; Qi, G.; Wan, J.; Lei, Z. Exploiting temporal and depth information for multi-frame face anti-spoofing. *arXiv preprint arXiv:1811.05118* **2018**.
32. Jourabloo, A.; Liu, Y.; Liu, X. Face de-spoofing: Anti-spoofing via noise modeling. In Proceedings of the Proceedings of the European conference on computer vision (ECCV), 2018, pp. 290–306.
33. Kim, T.; Kim, Y.; Kim, I.; Kim, D. Basn: Enriching feature representation using bipartite auxiliary supervisions for face anti-spoofing. In Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019, pp. 0–0.
34. Huszár, V.D.; Adhikarla, V.K. Live Spoofing Detection for Automatic Human Activity Recognition Applications. *Sensors* **2021**, *21*. doi:10.3390/s21217339.
35. Roy, K.; Hasan, M.; Rupty, L.; Hossain, M.S.; Sengupta, S.; Taus, S.N.; Mohammed, N. Bi-FPNFAS: Bi-Directional Feature Pyramid Network for Pixel-Wise Face Anti-Spoofing by Leveraging Fourier Spectra. *Sensors* **2021**, *21*. doi:10.3390/s21082799.
36. Ali, A.; Hoque, S.; Deravi, F. Directed Gaze Trajectories for Biometric Presentation Attack Detection. *Sensors* **2021**, *21*. doi:10.3390/s21041394.
37. Kowalski, M. A Study on Presentation Attack Detection in Thermal Infrared. *Sensors* **2020**, *20*. doi:10.3390/s20143988.
38. Zhang, Z.; Yan, J.; Liu, S.; Lei, Z.; Yi, D.; Li, S.Z. A face antispoofing database with diverse attacks. In Proceedings of the Biometrics (ICB), 2012 5th IAPR international conference on. IEEE, 2012, pp. 26–31.
39. Chingovska, I.; Anjos, A.; Marcel, S. On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the 2012 BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), 2012, pp. 1–7.
40. Galbally, J.; Marcel, S. Face anti-spoofing based on general image quality assessment. In Proceedings of the 2014 22nd International Conference on Pattern Recognition (ICPR). IEEE, 2014, pp. 1173–1178.
41. Pinto, A.; Pedrini, H.; Robson Schwartz, W.; Rocha, A. Face Spoofing Detection Through Visual Codebooks of Spectral Temporal Cubes. *Image Processing, IEEE Transactions on* **2015**, *24*, 4726–4740.
42. Benlamoudi, A.; Samai, D.; Ouafi, A.; Bekhouche, S.E.; Taleb-Ahmed, A.; Hadid, A. Face spoofing detection using Local binary patterns and Fisher Score. In Proceedings of the Control, Engineering & Information Technology (CEIT), 2015 3rd International Conference on. IEEE, 2015, pp. 1–5.
43. Yang, J.; Lei, Z.; Liao, S.; Li, S.Z. Face liveness detection with component dependent descriptor. In Proceedings of the Biometrics (ICB), 2013 International Conference on. IEEE, 2013, pp. 1–6.
44. Benlamoudi, A.; Samai, D.; Ouafi, A.; Bekhouche, S.; Taleb-Ahmed, A.; Hadid, A. Face spoofing detection using Multi-Level Local Phase Quantization (ML-LPQ). In Proceedings of the Proceeding of the first International Conference on Automatic Control, Telecommunication and signals ICATS'15, 2015. doi:10.13140/RG.2.1.3335.6241.
45. Benlamoudi, A.; Bougourzi, F.; Zighem, M.; Bekhouche, S.; Ouafi, A.; Taleb-Ahmed, A. Face Anti-Spoofing Combining MLLBP and MLBSIF. In Proceedings of the 10ème Conférence sur le Génie Electrique, 2017.
46. Quan, R.; Wu, Y.; Yu, X.; Yang, Y. Progressive Transfer Learning for Face Anti-Spoofing. *IEEE Transactions on Image Processing* **2021**, *30*, 3946–3955. doi:10.1109/TIP.2021.3066912.
47. Chingovska, I.; Anjos, A.; Marcel, S. On the effectiveness of local binary patterns in face anti-spoofing. In Proceedings of the Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the. IEEE, 2012, pp. 1–7.
48. Anjos, A.; Marcel, S. Counter-measures to photo attacks in face recognition: a public database and a baseline. In Proceedings of the Biometrics (IJCB), 2011 International Joint Conference on. IEEE, 2011, pp. 1–7.
49. Komulainen, J.; Hadid, A.; Pietikainen, M.; Anjos, A.; Marcel, S. Complementary countermeasures for detecting scenic face spoofing attacks. In Proceedings of the Biometrics (ICB), 2013 International Conference on. IEEE, 2013, pp. 1–7.
50. Arashloo, S.R.; Kittler, J.; Christmas, W. An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol. *IEEE access* **2017**, *5*, 13868–13882.
51. Xiong, F.; AbdAlmageed, W. Unknown presentation attack detection with face rgb images. In Proceedings of the 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS). IEEE, 2018, pp. 1–9.
52. 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). IEEE, 2017, pp. 612–618.
53. Bharadwaj, S.; Dhamecha, T.I.; Vatsa, M.; Singh, R. Computationally efficient face spoofing detection with motion magnification. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2013, pp. 105–110.
54. Freitas Pereira, T.; Anjos, A.; De Martino, J.M.; Marcel, S. Can face anti-spoofing countermeasures work in a real world scenario? In Proceedings of the Biometrics (ICB), 2013 International Conference on. IEEE, 2013, pp. 1–8.
55. Boulkenafet, Z.; Komulainen, J.; Hadid, A. Face spoofing detection using colour texture analysis. *IEEE Transactions on Information Forensics and Security* **2016**, *11*, 1818–1830.

56. Bian, Y.; Zhang, P.; Wang, J.; Wang, C.; Pu, S. Learning Multiple Explainable and Generalizable Cues for Face Anti-spoofing. *arXiv preprint arXiv:2202.10187* **2022**.