Fusion of Handcrafted and Deep Learning Features for Large-scale Multiple Iris Presentation Attack Detection

Daksha Yadav[1], Naman Kohli[1], Akshay Agarwal[2], Mayank Vatsa[2], Richa Singh[2], Afzel Noore[3]
[1]West Virginia University, [2]IIIT-Delhi, [3]Texas A&M University-Kingsville
[1]{dayadav, nakohli}@mix.wvu.edu,[2]{akshaya, mayank, rsingh}@iiitd.ac.in, [3]afzel.noore@tamuk.edu

Abstract

Iris recognition systems may be vulnerable to presentation attacks such as textured contact lenses, print attacks, and synthetic iris images. Increasing applications of iris recognition have raised the importance of efficient presentation attack detection algorithms. In this paper, we propose a novel algorithm for detecting iris presentation attacks using a combination of handcrafted and deep learning based features. The proposed algorithm combines local and global Haralick texture features in multi-level Redundant Discrete Wavelet Transform domain with VGG features to encode the textural variations between real and attacked iris images. The proposed algorithm is extensively tested on a large iris dataset comprising more than 270,000 real and attacked iris images and yields a total error of 1.01%. The experimental evaluation demonstrates the superior presentation attack detection performance of the proposed algorithm as compared to state-of-the-art algorithms.

1. Introduction

Iris recognition systems are being rapidly deployed in several commercial applications for authentication purposes. In the literature, tremendous growth has been observed and several covariates such as off-angle and quality variations are well-researched [1]. However, challenges such as ophthalmic disorders [2, 3] still require additional research. Similarly, the introduction of vulnerabilities such as presentation or spoofing attacks to gain unauthorized access or evade one's own identity is a major risk factor. Due to its negative impact, detection of presentation attacks has become a key research topic. Several presentation attack detection (PAD) frameworks have been developed in the literature which focus on developing algorithms for one specific presentation attack.

As shown in Figure 1, typical iris presentation attacks include textured contact lens [4, 5], print attack [6, 7], and synthetically generated iris images [8]. These attacks are described below:

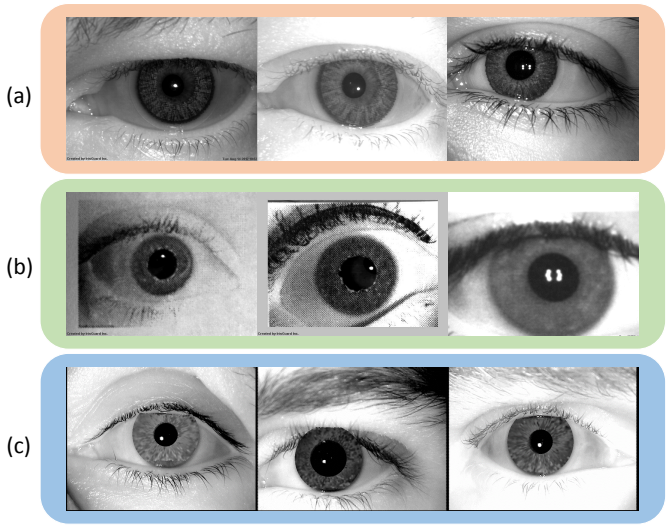


Figure 1. Samples with iris presentation attacks: (a) textured contact lenses, (b) printed iris images, and (c) synthetic iris images. Image sources: [4], [6], [7], [8], and [9].

- Textured contact lenses: The artificial texture pattern in cosmetic contact lenses can significantly conceal the original iris texture of the eyes and hence, may be utilized for intentional or unintentional identity evasion. To detect iris images containing such contact lenses, Zhang et al. [10] developed an altered local binary pattern based algorithm termed as weighted LBP (WLBP). Yadav et al. [4] demonstrated the performance of modified LBP on two different databases containing contact lens iris images. Several other algorithms have been proposed which utilize different handcrafted features [9] and deep learning features [11] to encode variations between real iris images and textured contact lens iris images.

- Printed iris attack: In this attack, a printed iris image is presented to the iris sensor for impersonating another

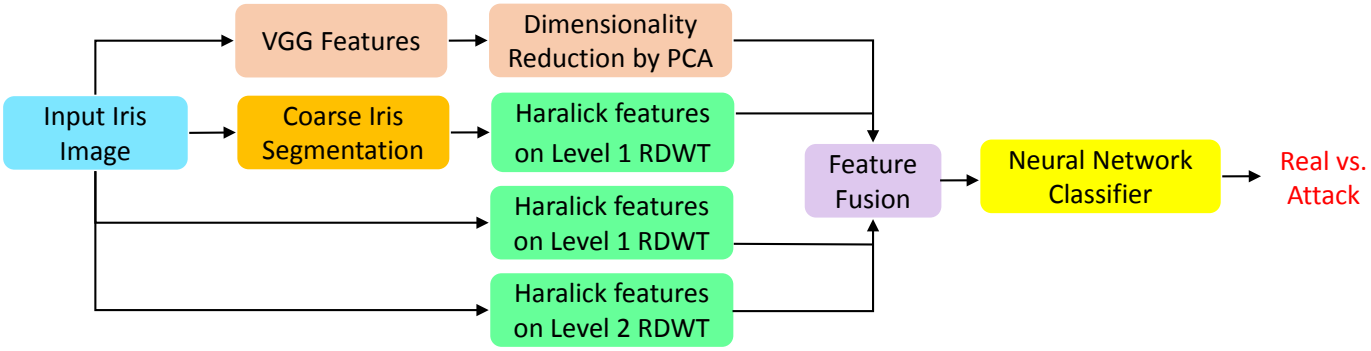


Figure 2. Proposed Multi-level Haralick and VGG Fusion (MHVF) algorithm for iris presentation attack detection.

individual's identity. Gupta et al. [6] demonstrated the efficacy of different handcrafted features in detecting such attacks. LivDet2013 [7] and LivDet2017 [12] competitions also contained samples of print iris attack and different participants showcased the performance of their approaches in classifying printed iris images.

- Synthetic iris: Galbally et al. [8] developed a genetic algorithm based approach to create synthetic iris-like texture patterns. Recently, Kohli et al. [13] utilized the generative adversarial network to produce realistic looking synthetic iris images.

As mentioned previously, existing presentation attack detection techniques generally focus on accurately detecting a specific type of iris presentation attack but in real-world scenarios, iris recognition systems should be able to handle and detect multiple types of iris presentation attacks. In the literature, there is a lack of algorithms which are designed to detect multiple iris presentation attacks. Kohli et al. [14] proposed an algorithm to detect a medley of iris presentation attacks and demonstrated the performance on a combined database of 21,525 iris images. To simulate the realistic environment, it is critical to develop presentation attack detection algorithms which are evaluated on a larger database with variations across acquisition sensors and multiple presentation attacks. With these objectives, the key contributions of this paper are:

- Proposed a novel framework utilizing representation encoding using handcrafted and deep learning based features for iris presentation attack detection.
- Demonstrated results on the Combined Iris database, prepared by combining the images from multiple publicly available databases. It contains more than 270,000 iris images with multiple presentation attacks.
- Compared the performance with several state-of-the-art PAD algorithms.

2. Proposed Multi-level Haralick and VGG Fusion (MHVF) Algorithm

Figure 2 illustrates the proposed **M**ulti-level **H**aralick and **V**GG **F**usion (MHVF) algorithm for iris presentation attack detection. The proposed algorithm has two components which are described in detail subsequently:

1. Iris representation encoding using handcrafted features
2. Iris representation encoding using deep learning based features

2.1. Encoding iris Representation using Multi-Level Haralick Features

Haralick features [15] are popular statistical feature descriptors for encoding textural information in images. They have been successfully utilized in various applications such as texture classification [16], medical imaging classification [17], and face presentation attack detection [18]. Haralick features are based on computation of gray level co-occurrence matrices (GLCM) which are defined as the distribution of co-occurring pixel intensity values in an image (I) at a specific offset $(\Delta p, \Delta q)$ at location (x, y). Thus, $GLCM_{\Delta p,\Delta q}(x, y)$ is computed as:

$$\sum_{p=1}^{n}\sum_{q=1}^{m}\begin{cases}1,\text{if } I(p,q)=x\ ;\ I(p+\Delta p,q+\Delta q)=y\\ 0,\text{otherwise.}\end{cases} \quad (1)$$

After computation of GLCM, Haralick statistical features are calculated from the GLCM for encoding the textural information in the image. The 13 Haralick features correspond to angular second moment, contrast, correlation, inverse difference moment, sum variance, sum of squares variance, sum average, entropy, sum entropy, difference variance, difference entropy, and two information measures of correlation.

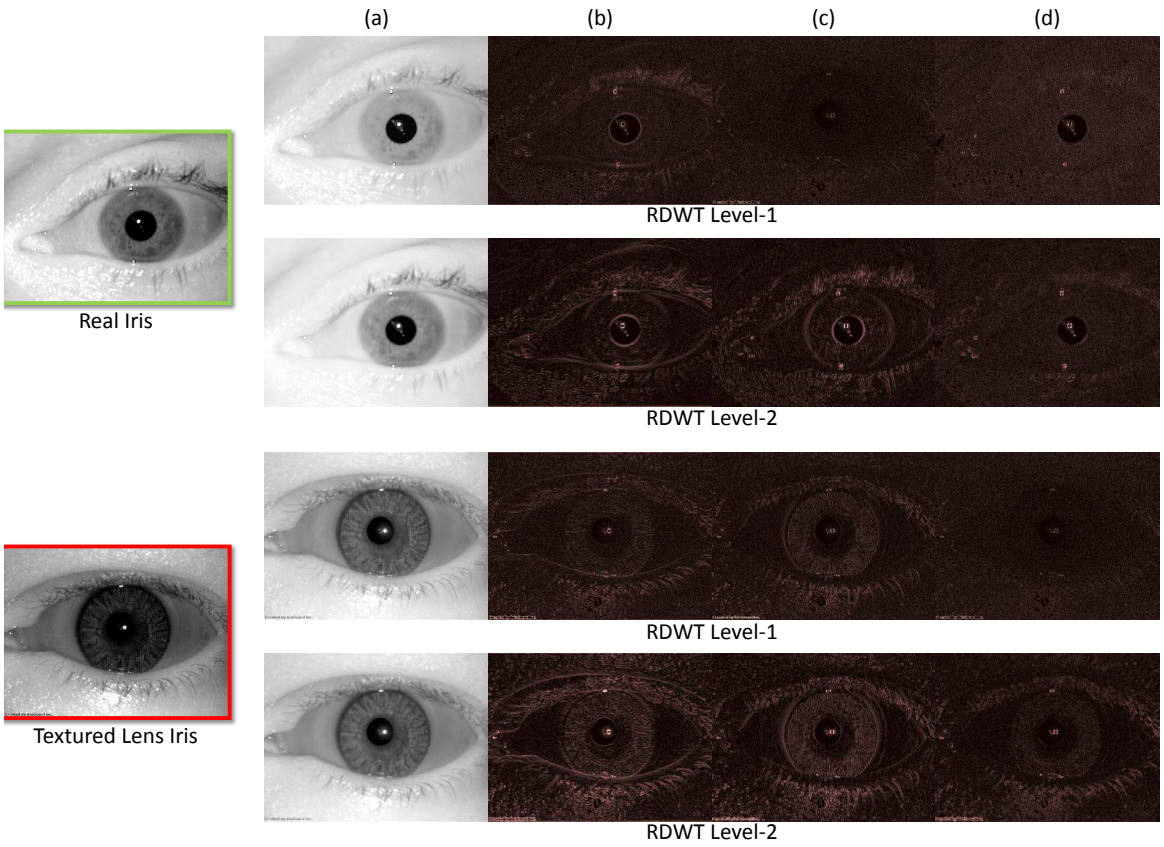


Figure 3. RDWT decomposition of a real iris image and textured contact lens iris image: (a) approximation subband, (b) horizontal subband, (c) vertical subband, and (d) diagonal subband.

In the proposed MHVF algorithm, Haralick features are extracted from redundant discrete wavelet transform (RDWT). Along with its shift-invariance property, RDWT achieves multi-scale decomposition along with over-complete encoding. RDWT decomposition is performed to obtain four sub-bands with the size same as the original image: horizontal (H_h), vertical (H_v), diagonal (H_d), and approximation (H_{approx}). n-level RDWT decomposition is performed by utilizing the previous level's H_{approx} output. Multi-level RDWT decomposition provides complementary information about image features at different scales. Figure 3 shows sample RDWT decomposition of real and attacked iris images.

In the proposed MHVF algorithm, the input iris image is resized to 640×480 and is converted to the RDWT domain. Given the resized iris image, Haralick features are computed on the four sub-bands of the first and second level RDWT domain. Haralick features are also computed on the original grayscale image. Thus, the input iris image produces 117-dimensional Haralick feature descriptor. Next, to learn local textural information, features are extracted from segmented iris region. For faster computation, only a coarse iris segmentation is performed by first detecting the pupil and then approximating the iris region around it. The pupil detection utilizes retinex based image normalization and finding circular contours to detect pupil. Haralick features of the segmented iris region are computed on the grayscale intensity values and in first level RDWT domain since the textural information of the iris region is highlighted in the first level. This produces the local Haralick feature vector of size 65. Thus, Haralick features in multi-level RDWT of the global iris and Haralick features in single-level RDWT of the coarsely segmented localized iris region encode the minuscule textural variations between real and attacked iris images.

2.2. VGG Iris Feature Extraction

Convolutional neural networks (CNNs) have demonstrated the ability to effectively encode image representation in various visual recognition tasks. VGG-16 [19] is a popular 16-layer CNN architecture which utilizes 3×3 filters and achieves exceptional performance on the large ImageNet database. As the architecture is trained on more than a million images, it is capable of learning rich feature encodings of various objects. In the proposed MHVF algorithm, the input iris image is resized to 224×224 and

Table 1. Characteristics of the Combined Iris Database and its constituent databases.

Database	No. of Iris Images	Type(s) of Iris Images
LivDet2013 (Warsaw Subset) [7]	1,667	Print and Real
Combined Spoofing Database [14]	21,525	Real, Print, Textured Contact Lens, and Synthetic Iris
NDCLD-2013 [4]	4,200 (LG4000) and 900 (AD100)	Real and Textured Contact Lens
NDCLD-2015 [9]	7,300	Real and Textured Contact Lens
ND-Iris-0405	64,980	Real
ND-CrossSensor-Iris-2013	117,106 (LG2200) & 29,845 (LG4000)	Real
ND-TimeLapseIris-2012	6,796	Real
CASIA-Iris-Thousand	20,000	Real
Combined Iris Database	**274,319**	**Real, Print, Textured Contact Lens, and Synthetic Iris**

pre-trained VGG model is utilized to extract deep learning based features. Principal component analysis is performed to reduce the resultant 4096 feature vector to an underlying low dimensional space of 117 size.

2.3. Feature Fusion and Classification

Multi-level Haralick and VGG features encode inherent textural, edge and shape information of the input iris image. Thus, feature-level fusion is performed to combine the features by concatenating 117-dimensional multi-level Haralick feature of the global iris, 65-dimensional single-level Haralick feature of the segmented iris, and reduced dimension VGG feature. This fused feature vector is utilized as the input to a three-layer artificial neural network (ANN) to classify the input iris image as real or attacked.

3. Experimental Evaluation

3.1. Combined Iris Database

Various iris presentation attack databases exist in the literature which contain iris images pertaining to a specific presentation attack. However, there is a scarcity of iris databases consisting of multiple presentation attacks. Kohli et al. [14] created Combined Spoofing Dataset (CSD) which was a combination of multiple publicly available spoofing databases. It contained 11,368 attack iris images and 9,325 real iris images from different iris databases. Inspired by CSD, we present the largest iris presentation attack database, Combined Iris database, containing more than 270,000 real and attacked iris images. A key feature of this database is that it comprises iris images acquired from different sensors, subjects of different ethnicities, and multiple presentation attacks. The constituent databases of the Combined Iris database are:

- LivDet2013 (Warsaw Subset) [7] contains 1,667 real and printed iris images.
- Combined Spoofing Database (CSD) [14] comprises various iris presentation attack images such as textured contact lenses, print attack, and synthetic iris images as well as real iris images from different databases [4, 6, 8, 20]. The total number of images is 21,525.
- NDCLD-2013 [4] consists of real and attacked images acquired from two sensors: LG4000 and AD100. 4,200 images have been captured with LG4000 iris sensor while 900 images are acquired with AD100 iris sensor is 900. This database contains iris images with textured contact lens iris images from different manufacturers.
- NDCLD-2015 [9] contains iris images with textured contact lens and real iris images. The total number of images in the database is 7,300.

The above-mentioned databases contain both, real and attacked iris images. For increasing the real samples in the Combined Iris database, the following databases are also added which contain only real iris images: ND-Iris-0405 [21], ND-CrossSensor-Iris-2013 [22], ND-TimeLapseIris-2012 [23], and CASIA-Iris-Thousand [24]. The characteristics of the Combined Iris database and its constituent databases are summarized in Table 1. Figure 4 shows sample iris images from different constituent databases of the Combined Iris database along with their image type (real or attack).

3.2. Experimental Setup

The performance of the proposed MHVF algorithm is evaluated on the Combined Iris database. For experimental evaluation, each constituent database is split into five cross-validation folds and merged to form the five folds of the Combined Iris database. It is ensured that the subjects in the training and testing partitions are disjoint in every fold. Table 2 shows the number of real and attacked iris samples in

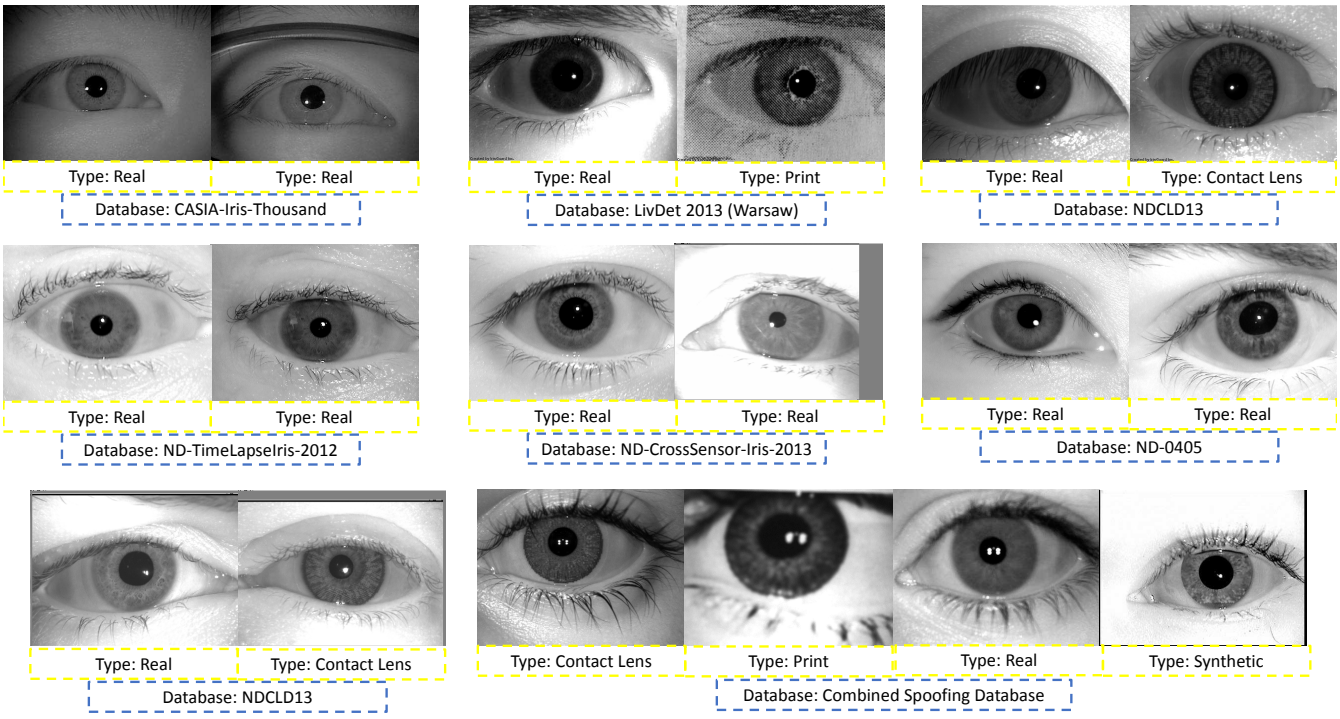


Figure 4. Sample iris image types from various constituent databases of the Combined Iris database.

Table 2. The number of real and attacked iris images in each fold for the Combined Iris database.

Fold	Training Partition		Testing Partition	
	Real	Attack	Real	Attack
1	210,339	9,760	51,698	2,522
2	208,973	9,513	53,064	2,769
3	209,282	9,324	52,755	2,958
4	209,633	9,427	52,404	2,855
5	211,601	9,424	50,436	2,858

the training and testing partitions of each fold. Comparative analysis is performed with existing PAD algorithms: LBP [6] , WLBP [10], DESIST [14], and the two components of MHVF algorithms: Multi-Level Haralick (MH) features and VGG features. The presentation attack detection performance is evaluated using the following metrics:

- Total Error: Error rate of all misclassified iris images.
- Attack Presentation Classification Error Rate (APCER): Error rate of misclassified attacked iris images.
- Bonafide Presentation Classification Error Rate (BPCER): Error rate of misclassified real iris images.

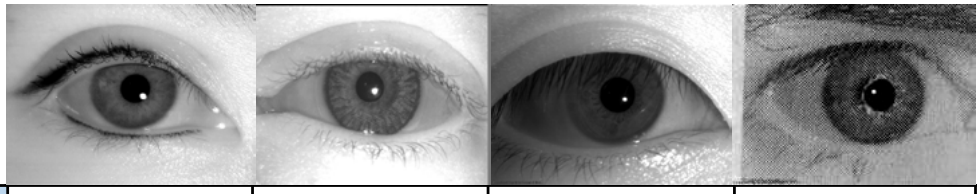
Table 3. Iris presentation attack detection performance (%) of the proposed MHVF and existing algorithms on the Combined Iris database.

Algorithm	Total Error	APCER	BPCER
LBP [6]	22.94	80.00	20.00
WLBP [10]	52.75	48.18	53.00
DESIST [14]	4.13	77.48	0.20
MH	3.36	59.96	0.32
VGG [19]	1.53	21.83	0.44
Proposed MHVF	**1.01**	**18.58**	**0.07**

3.3. Results on Combined Iris Database

The average total error, APCER, and BPCER values across the five cross-validation folds of the Combined Iris database are summarized in Table 3. It is observed that the proposed MHVF algorithm yields the best results with a minimum total error of 1.01%. It achieves 18.58% error in detecting presentation attack iris images and 0.07% error in detecting real iris images.

Comparative analysis of the proposed MHVF algorithm is also performed with existing iris PAD algorithms: LBP [6], WLBP [10], and DESIST [14]. It is observed that the MHVF algorithm outperforms the existing algorithms on the Combined Iris database. It achieves $3-21\%$ lower total error as compared to LBP, WLBP, and DESIST. With respect to detecting attacked iris images, MHVF achieves



MH	Attack	Attack	Real	Real
VGG	Real	Real	Real	Attack
MHVF	Real	Attack	Real	Real
True Label	Real	Attack	Real	Attack

Figure 5. Sample classification outputs by Multi-level Haralick (MH), VGG, and the proposed Multi-level Haralick and VGG Fusion (MHVF) algorithm.

Table 4. Iris presentation attack detection performance (%) of the proposed MHVF and existing algorithms on various constituents of the Combined Iris database.

Database	**Metric**	**Presentation Attack Detection Algorithms**					
		LBP	**WLBP**	**DESIST**	**MH**	**VGG**	**MHVF**
LivDet2013 (Warsaw) [7]	Total Error	4.86	7.28	7.20	4.05	0.49	**0.00**
	APCER	4.74	6.86	2.71	0.32	0.98	0.00
	BPCER	4.97	7.70	10.83	7.69	0.00	0.00
Combined Spoofing Database [14]	Total Error	8.01	19.04	12.49	5.73	1.33	**1.23**
	APCER	12.27	24.94	25.58	8.11	2.11	1.82
	BPCER	4.62	14.39	2.47	3.87	0.75	0.78
NDCLD-2013 (LG4000) [4]	Total Error	0.25	1.33	0.50	0.08	0.00	**0.00**
	APCER	0.00	2.00	0.50	0.25	0.00	0.00
	BPCER	0.38	1.00	0.50	0.00	0.00	0.00
NDCLD-2013 (AD100) [4]	Total Error	7.67	12.33	1.67	0.33	0.33	**0.33**
	APCER	0.00	9.00	2.00	0.00	1.00	1.00
	BPCER	11.50	14.00	1.50	0.50	0.00	0.00
NDCLD-2015 [9]	Total Error	25.58	23.02	17.52	14.57	1.08	**1.01**
	APCER	6.15	50.58	29.81	21.73	1.54	1.92
	BPCER	38.70	4.41	9.22	9.74	0.78	0.39

29−61% lower APCER as compared to existing algorithms. Similarly, in detecting real iris images, it yields at least 0.13% lower BPCER as compared to the other algorithms.

The proposed MHVF algorithm involves feature-level fusion of local and global handcrafted multi-level Haralick features and deep learning based VGG features. Thus, for analyzing the contribution of each constituent, the performance of multi-level Haralick (MH) features with ANN classifier and VGG features with ANN classifier is reported in Table 3. Sample classification outputs by the three PAD algorithms are shown in Figure 5. It is observed that MH with ANN classifier yields 2.35% higher total error as compared to the proposed MHVF algorithm. On the other hand, VGG with ANN classifier yields 0.52% higher total error as compared to the proposed MHVF algorithm. These results highlight the improvement in the performance of the proposed algorithm due to feature fusion.

3.4. Results on Constituent Databases

The performance of the proposed MHVF algorithm is also analyzed on different constituent databases of the Combined Iris database. For this experiment, the databases containing both types of iris images (real and attack) are selected, i.e. LivDet2013, Combined Spoofing Database, NDCLD-2013, and NDCLD-2015. The training of the proposed MHVF algorithm is performed using the respective training partition of each constituent database and the results are summarized in Table 4. For comparative analysis, the performance of LBP, WLBP, DESIST, MH, and VGG is also reported.

With respect to the total error, it is observed that the proposed MHVF algorithm demonstrates improved perfor-

mance as compared to the other PAD algorithms on all the databases. It also achieves the lowest APCER value on LivDet2013, CSD, and LG4000 subset of NDCLD-2013. Likewise, it yields the lowest BPCER value on LivDet2013, NDCLD-2013 (both subsets), and NDCLD-2015 databases. Additionally, deep learning based VGG features showcase the second best PAD performance on different constituent databases.

4. Conclusion

Existing studies in the iris presentation attack detection literature are based on the assumption that the system encounters a specific iris presentation attack. However, this may not be true in real-world scenarios where the iris recognition system may have to handle multiple kinds of presentation attacks. To address this challenging scenario, we propose a framework for detecting multiple presentation attacks. For this, we develop the largest iris presentation attack database by combining several databases. The Combined Iris database contains more than 270,000 real and attacked iris images. The proposed MHVF algorithm is developed by fusing handcrafted and deep learning based features to encode variations between real and attacked iris images. Experimental evaluation reveals that the proposed algorithm outperforms state-of-the-art iris presentation attack detection algorithm by achieving 1.01% total error on the Combined Iris database.

Acknowledgment

The authors acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40 and Titan X Pascal GPUs used for this research.

References

[1] I. Nigam, M. Vatsa, and R. Singh, "Ocular biometrics: A survey of modalities and fusion approaches," *Information Fusion*, vol. 26, pp. 1–35, 2015. 1

[2] S. S. Arora, M. Vatsa, R. Singh, and A. Jain, "Iris recognition under alcohol influence: A preliminary study," in *IEEE International Conference on Biometrics*, 2012, pp. 336–341. 1

[3] I. Nigam, M. Vatsa, and R. Singh, "Ophthalmic disorder menagerie and iris recognition," in *Handbook of Iris Recognition*. Springer, 2016, pp. 519–539. 1

[4] D. Yadav, N. Kohli, J. S. Doyle, R. Singh, M. Vatsa, and K. W. Bowyer, "Unraveling the effect of textured contact lenses on iris recognition," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 851–862, 2014. 1, 4, 6

[5] N. Kohli, D. Yadav, M. Vatsa, and R. Singh, "Revisiting iris recognition with color cosmetic contact lenses," in *IEEE International Conference on Biometrics*, 2013, pp. 1–7. 1

[6] P. Gupta, S. Behera, M. Vatsa, and R. Singh, "On iris spoofing using print attack," in *IEEE International Conference on Pattern Recognition*, 2014, pp. 1681–1686. 1, 2, 4, 5

[7] D. Yambay, J. S. Doyle, K. W. Bowyer, A. Czajka, and S. Schuckers, "LivDet-iris 2013-iris liveness detection competition 2013," in *IEEE International Joint Conference on Biometrics*, 2014, pp. 1–8. 1, 2, 4, 6

[8] J. Galbally, A. Ross, M. Gomez-Barrero, J. Fierrez, and J. Ortega-Garcia, "Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms," *Elsevier Computer Vision and Image Understanding*, vol. 117, no. 10, pp. 1512–1525, 2013. 1, 2, 4

[9] J. S. Doyle and K. W. Bowyer, "Robust detection of textured contact lenses in iris recognition using BSIF," *IEEE Access*, vol. 3, pp. 1672–1683, 2015. 1, 4, 6

[10] H. Zhang, Z. Sun, and T. Tan, "Contact lens detection based on weighted LBP," in *IEEE International Conference on Pattern Recognition*, 2010, pp. 4279–4282. 1, 5

[11] R. Raghavendra, K. B. Raja, and C. Busch, "ContlensNet: Robust iris contact lens detection using deep convolutional neural networks," in *IEEE Winter Conference on Applications of Computer Vision*, 2017, pp. 1160–1167. 1

[12] D. Yambay, B. Becker, N. Kohli, D. Yadav, A. Czajka, K. W. Bowyer, S. Schuckers, R. Singh, M. Vatsa, A. Noore, D. Gragnaniello, C. Sansone, L. Verdoliva, L. He, Y. Ru, H. Li, N. Liu, Z. Sun, and T. Tan, "Livdet Iris 2017 - Iris liveness detection competition 2017," in *IEEE International Joint Conference on Biometrics*, 2017, pp. 733–741. 2

[13] N. Kohli, D. Yadav, M. Vatsa, R. Singh, and A. Noore, "Synthetic iris presentation attack using iDCGAN," in *IEEE International Joint Conference on Biometrics*, 2017, pp. 674–680. 2

[14] ——, "Detecting medley of iris spoofing attacks using DESIST," in *IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2016, pp. 1–6. 2, 4, 5, 6

[15] R. M. Haralick, K. Shanmugam *et al.*, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 6, pp. 610–621, 1973. 2

[16] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, 1995. 2

[17] M. B. Nagarajan, M. B. Huber, T. Schlossbauer, G. Leinsinger, A. Krol, and A. Wismüller, "Classification of small lesions in breast MRI: evaluating the role of dynamically extracted texture features through feature selection," *Journal of Medical and Biological Engineering*, vol. 33, no. 1, 2013. 2

[18] A. Agarwal, R. Singh, and M. Vatsa, "Face anti-spoofing using Haralick features," in *IEEE International Conference on Biometrics Theory, Applications and Systems*, 2016, pp. 1–6. 2

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014. 3, 5

[20] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern Recognition*, vol. 43, no. 3, pp. 1016–1026, 2010. 4

[21] K. W. Bowyer and P. J. Flynn, "The ND-IRIS-0405 iris image dataset," *arXiv preprint arXiv:1606.04853*, 2016. 4

[22] R. Connaughton, A. Sgroi, K. W. Bowyer, and P. Flynn, "A cross-sensor evaluation of three commercial iris cameras for iris biometrics," in *IEEE Computer Vision and Pattern Recognition Workshops*, 2011, pp. 90–97. 4

[23] S. E. Baker, K. W. Bowyer, P. J. Flynn, and P. J. Phillips, "Template aging in iris biometrics," in *Handbook of Iris Recognition*. Springer, 2013, pp. 205–218. 4

[24] [Online], "CASIA Iris image database," http://biometrics.idealtest.org. 4