Face Anti-Spoofing using Speeded-Up Robust Features and Fisher Vector Encoding

Zinelabidine Boulkenafet, Jukka Komulainen and Abdenour Hadid
Center for Machine Vision and Signal Analysis, University of Oulu, Finland

***Abstract*—The vulnerabilities of face biometric authentication systems to spoofing attacks have received a significant attention during the recent years. Some of the proposed countermeasures have achieved impressive results when evaluated on intra-tests i.e. the system is trained and tested on the same database. Unfortunately, most of these techniques fail to generalize well to unseen attacks e.g. when the system is trained on one database and then evaluated on another database. This is a major concern in biometric anti-spoofing research which is mostly overlooked. In this paper, we propose a novel solution based on describing the facial appearance by applying Fisher Vector encoding on Speeded-Up Robust Features (SURF) extracted from from different color spaces. The evaluation of our countermeasure on three challenging benchmark face spoofing databases, namely the CASIA Face Anti-Spoofing Database, the Replay-Attack Database and MSU Mobile Face Spoof Database, showed excellent and stable performance across all the three datasets. Most importantly, in inter-database tests, our proposed approach outperforms the state of the art and yields in very promising generalization capabilities, even when only limited training data is used.**

I. Introduction

It is well known nowadays that face biometric systems are vulnerable to spoofing attacks e.g. when presenting fake faces using printed photos, video displays and masks. In a recent study [1], six commercial face recognition systems (Face Unlock, Facelock Pro, Visidon, Veriface, Luxand Blinkand FastAccess) were easily fooled with crude photo attacks using images of the targeted person downloaded from social networks.

To overcome the problem of spoofing attacks, many non-intrusive software-based countermeasures have been proposed [2], [3]. While it is possible to exploit different visual cues for face spoofing detection such as motion [4], [5], [6], [7] and scene context [4], [8], an approach solely based on single images of the face region is more appealing and also more challenging. It is appealing because the same information (i.e. the facial region) that is used for face recognition will also be used for spoofing detection. So, the two tasks can easily be coupled.

The methods solely based on single images of the face region exploit the fact that fake face images captured from printed photos, video displays and masks usually suffer from various quality and texture issues related to the spoofing medium or the manufacturing process. This includes lack of details, printing artifacts, specular reflections, or differences in shading. Assuming that these inherent disparities between real and fake faces can be observed in single visual spectrum images, the proposed methods in the literature analyze the facial appearance properties like texture [9], [10] and quality [11], [12] for face spoofing detection from single images of face region.

The existing face anti-spoofing techniques analyzing motion, facial texture content and image quality have already achieved impressive results particularly when trained and evaluated on the same database (i.e. intra-test protocols). As all the existing benchmark publicly available datasets lack variations in the collected data (e.g. user demographics, application scenarios, illumination conditions and input cameras), the reported anti-spoofing results may unfortunately not reflect the real uncontrolled operating conditions that the methods will be definitely faced in real world applications such as in mobile authentication.

To gain insight into the generalization performance of face anti-spoofing techniques, de Freitas Pereira *et al.* [13] suggested an inter-database evaluation in which the anti-spoofing models are trained and tuned on one database and then tested on other databases. The experiments have revealed that the performance of the state-of-the-art methods drastically drops as the methods failed to cope with new spoofing conditions that have not been seen during training and development phases. Even the popular convolutional neural networks (CNN) have failed to derive well-generalizing features for face anti-spoofing [14].

It is indeed impossible to cover all possible variations related to spoofing operating conditions in the training data. Instead of augmenting the training data, a possible direction towards more robust software-based solutions is to design novel feature representations that are less sensitive to different environmental and subject-specific factors. In order to improve the generalization of texture based anti-spoofing methods, we have proposed the use of color texture analysis in [15], exploiting the fact that the color gamut of printing and display devices is limited. In order to map the out of gamut colors into the color gamut of different devices, color mapping algorithms are applied. Since the human eye is more sensitive to the luminance than the chrominance information, these mapping algorithms give a huge importance to the preservation of the spatially local luminance variations at the cost of the chroma information. These inherent disparities can be captured by analyzing the texture content of the chrominance channels. Our preliminary investigations in [15] suggested that color texture when extracted separately from the luminance and the chrominance channels are more stable in many (unknown) conditions compared to their RGB and gray-scale counterparts.

The generalization capability of our color texture analysis method [15] was dependent on the diversity of the training

data. The method was performing very well when trained on the CASIA Face Anti-Spoofing Database [16] containing different imaging qualities and then tested on the more constrained Replay-Attack Database [17]. However, the performance of the method was less satisfactory when trained on constrained data (Replay-Attack Database) and then tested on more diverse data (CASIA Face Anti-Spoofing Database). This can be partially explained by the fact that only the basic local binary patterns (LBP) [18] were considered for exploring the facial appearance. While LBP is indeed a simple and powerful texture descriptor that has shown to be very effective in many applications including face anti-spoofing, we argue that more advanced feature descriptors and encoding methods are needed to further enhance the generalization capability of face anti-spoofing.

In this present work, we propose a novel face representation for a well-generalizing anti-spoofing method using Speeded-Up Robust Features (SURF) and Fisher Vector encoding [19]. The color information is exploited for discriminating real from fake faces by extracting dense SURF descriptions from different color spaces. The SURF features extracted from the different band images are concatenated and encoded using the Fisher Vector method. The face representation is then fed into a Softmax classifier. Our experiments on three challenging benchmark face spoofing databases, namely the CASIA Face Anti-Spoofing Database [16], the Replay-Attack Database [17] and MSU Mobile Face Spoof Database [12], showed robust and stable performance across all these datasets. Most importantly, in the inter-database tests, our approach outperforms all the state of the art and yields in promising generalization capabilities, even when only limited training data is used.

II. Proposed Countermeasure

A. The Speeded-Up Robust Features (SURF)

The Speeded-Up Robust Features (SURF) [20] is a fast and efficient scale and rotation invariant descriptor. It was originally proposed to reduce the computational complexity of the Scale Independent Feature Transform (SIFT) descriptor [21]. Instead of using the Difference of Gaussian (DoG) filters to approximate the Laplacian of Gaussian, the SURF descriptor uses the Haar box filters. A convolution with these box filters can be computed rapidly by utilizing integral images.

The SURF descriptor is obtained using the Wavelet responses in the horizontal and vertical directions. The region around each interest point is first divided into 4×4 sub-regions. Then, for each sub-region j, the horizontal and vertical Wavelet responses are used to form a feature vector V_j as follows:

$$V_j = [\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|]. \quad (1)$$

Where d_x and d_y are the Haar wavelet responses in the horizontal and vertical directions, respectively. The feature vectors extracted from each sub-region are concatenated to from a SURF descriptor with 64 dimensions:

$$SURF = [V_1, ..., V_{16}]. \quad (2)$$

The SURF descriptor was originally proposed for gray-scale images. Inspired by our previous finding [15], [22] showing the importance of the color texture in face anti-spoofing, we propose to extract the SURF features from the color images instead of the gray-scale representation. First, the SURF descriptor is applied on each color band separately. Then, the obtained features are concatenated to form a single feature vector (referred to as CSURF). Finally, Principal Component Analysis (PCA) [23] is applied to de-correlate the obtained feature vector and reduce the dimensionality of the face description.

B. Fisher Vector (FV)

Extracting dense features has shown to be an essential component in many computer vision applications [24], [25]. In [26], Fisher Vector (FV) encoding was shown to perform very well in many image recognition benchmarks. FV embeds a set of feature vectors into a high dimensional space more amenable to linear classification. The feature vectors are obtained by fitting a generative parametric model, e.g. Gaussian Mixture Model (GMM), to the features to be encoded. Let $X = \{x_t, t = 1, ..., T\}$ be a D-dimensional local descriptors extracted from a face Image I and let $\lambda = \{\mu_k, \sigma_k, w_k, k = 1, ..., M\}$ be the means, the covariance matrices and the weights of the GMM model λ trained with a large set of local descriptors. The derivations of the model λ with respect of the mean and the covariance parameters (Equation 3 and 4) capture the first and the second order differences between the features X and each of the GMM components.

$$\phi_k^1 = \frac{1}{T\sqrt{w_k}} \sum_{t=1}^{T} \alpha_t(k)(\frac{x_t - \mu_k}{\sigma_k}) \quad (3)$$

$$\phi_k^2 = \frac{1}{T\sqrt{2w_k}} \sum_{t=1}^{T} \alpha_t(k)[\frac{(x_t - \mu_k)^2}{\sigma_k^2} - 1], \quad (4)$$

where, $\alpha_t(k)$ is the soft assignment weight of the feature x_t to the GMM component k:

$$\alpha_t(k) = \frac{w_k u_k(x_t)}{\sum_{j=1}^{M} w_j u_j(x_t)} \quad (5)$$

Here, u_i denote the probability density function of the Gaussian component i. The concatenation of these two order differences $[\phi_1^1, ..., \phi_M^1, \phi_1^2, ..., \phi_M^2]$ represent the Fisher Vector of the image I described by its local descriptors X. The dimensionality of this vector is $2MD$. A Fisher vector represents how the distribution of the local descriptors X differ from the distribution of the GMM model trained with all the training images. To further improve the performance, the Fisher vectors are normalized using a square rooting followed by L_2 normalization [27]. Figure 1 depict the general block diagrm of our face spoofing detection method.

III. Experimental data and setup

To assess the generalization capability of our proposed countermeasure, we used three public face anti-spoofing databases: CASIA Face Anti-Spoofing Database (CASIA FA) [16], Replay-Attack Database [17] and MSU Mobile Face

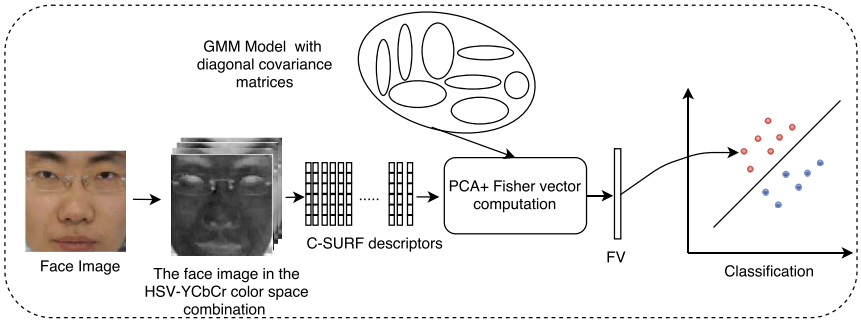

Fig. 1. An overview of our proposed face anti-spoofing method

Spoof Database (MSU MFS)[12]. These three datasets are the most challenging face anti-spoofing benchmark databases that consist of recordings of real client accesses and various spoofing attack attempts captured with different imaging qualities, including mobile phones, webcams and digital system cameras.

To allows a fair comparison with other methods proposed in the literature, we followed the official overall test protocols of the three databases. For CASIA FA and MSU MFS the model parameters are trained and tuned using a subject-disjoint cross-validation on the training set and the results are reported in terms of Equal Error Rate (EER) on the test set. The Replay-Attack database provides also a separate development set for tuning the model parameters. Thus, the results are given in terms of EER on the development set and the Half Total Error Rate (HTER) on the test set following the official test protocol.

In all our experiments, The dense $SURF$ features were extracted from 64×64 face images with a stride of two pixels and block size of 11 pixels. The frame images, were taken from each video every 320 ms. The Fisher Vectors were estimated using a GMM model with diagonal covariance matrices computed using training set of each database. Finally, the normalized Fisher Vector were fed into a Softmax classifier with a cross-entropy loss function [28].

In addition to the intra-database evaluation, we have also conducted a cross-database evaluation. Where, we used the training set of each database to train the countermeasure model and the testing set to estimate the threshold τ which will be used on the other databases to compute the Half Total Error Rate (HTER):

IV. Results and discussion

A. Effect of the color information

We begin our experiments by first evaluating the importance of the color SURF features (referred to as CSURF) compared to the gray-scale SURF features (referred to as SURF). In these experiments, we extracted the CSURF features from three color spaces: RGB, HSV and YCbCr. To show the effect of the color information, the extracted features were concatenated then fed into the softmax classifier without any feature encoding technique. The results in both intra-database and cross-database scenarios are presented in Table I and Table II, respectively. These results clearly indicate the importance of CSURF descriptions compared to the original SURF descriptions extracted from the gray-scale images. Comparing the results obtained using the different color spaces, we observe that using HSV and YCbCr color spaces yields in better performance compared to the RGB color space. This confirms the importance of using separated luminance and chrominance color spaces. As the color (luminance and the chrominance) information in the HSV and YCbCr color spaces are different, we propose to fuse the features extracted from these two color spaces in order to benefit from their potential complementarity. As shown in tables I and II, this fusion improves the performance in both intra-database (except on MSU database where the use of the HSV color space gives the best performance) and inter-database scenarios compared to the performance obtained using each color space separately. Even the dimension of the concatenated CSURF features is very high (393216) the obtained results are competitive to the state-of-the-art methods. To boost these performances, We applied the FV encoding methods on these features before the classification step. The next part shows the effect of this encoding method on the performances of CSURF features extracted from HSV-YCbCr color space combination.

TABLE I
PERFORMANCE OF SURF VERSUS CSURF IN INTRA-DATABASE TESTS

Method	Replay-Attack		CASIA	MSU
	EER	HTER	EER	EER
SURF(Gray)	19.5	21.2	17.8	18.8
CSURF (RGB)	11.3	13.5	14.1	17.3
CSURF (HSV)	6.2	11.5	7.1	7.0
CSURF (YCbCr)	5.2	8.9	7.8	9.2
CSURF (HSV+YCbCr)	**3.3**	**8.2**	**5.7**	7.1

TABLE II
PERFORMANCE OF SURF VERSUS CSURF IN INTER-DATABASE TESTS

Train on:	CASIA		Replay		MSU		Average
Test on:	Replay	MSU	CASIA	MSU	CASIA	Replay	
SURF	52.3	35.1	52.6	43.8	43.1	48.2	45.8
CSURF(RGB)	50.7	32.2	49.4	44.1	44.6	47.8	44,8
CSURF(HSV)	50.5	26.1	44.5	44.3	38.9	54.6	43.1
CSURF(YCbCr)	40.0	26.2	36.9	31.8	31.7	53.8	36.7
CSURF(HSV+YCbCr)	37.9	20.5	36.2	33.0	34.8	50.6	**35.5**

B. Effect of the Fisher Vector encoding method

The FV encoding capture the first and the second order differences between the image features and the center of the GMM components. Applying the PCA method before the encoding step was found to improve the performances and decrease the dimension of the final descriptors. Tables III and

IV show the effect of using different principal components on both the intra-database and inter-database scenarios. The results in these tables are obtained using 128 GMM components. From these tables, we can see that applying the PCA method before the FV encoding improves the performances in the two scenarios. In the intra-database scenario, the use of 100 principle components gives the best performances. However, on the cross-database scenario, the best results are obtained using 300 principal components. Since we are focusing more on the generalization capability, the CSURF features were projected into 300 principle components. In addition to the number of the principal components, the number of the GMM components has also an effect on the FV performances. Tables V and VI show that using 256 Gaussian components gives the best performance on both intra-database and inter-database tests.

TABLE III
EFFECT OF DIMENSIONALITY REDUCTION ON INTRA-DATABASE PERFORMANCE

Method	Replay-Attack		CASIA	MSU
	EER	HTER	EER	EER
Without	0.1	3.9	3.8	3.1
350	0.1	1.9	3.3	2.6
300	0.1	1.7	2.9	2.8
200	0.1	1.7	2.9	2.2
100	0.9	1.4	2.9	1.9

TABLE IV
EFFECT OF DIMENSIONALITY REDUCTION ON INTER-DATABASE PERFORMANCE

Train on:	CASIA		Replay		MSU		Average
Test on:	Replay	MSU	CASIA	MSU	CASIA	Replay	
Without	33.1	18.9	33.4	29.6	25.5	46.5	31.2
350	33.1	21.2	28.4	33.4	33.1	29.8	28.8
300	27.3	20.8	27.7	33.1	26.7	28.4	**27.4**
200	34.1	19.1	29.1	32.3	26.7	32.1	28.9
100	42.6	19.8	33.4	29.6	25.5	46.5	34.1

TABLE V
EFFECT OF THE NUMBER OF GMM COMPONENTS ON THE INTRA-DATABASE PERFORMANCE

Method	Replay-Attack		CASIA	MSU
	EER	HTER	EER	EER
64	0.2	2.0	3.2	2.8
128	0.2	2.3	2.9	2.8
256	**0.1**	**1.8**	**2.8**	**2.2**
512	0.3	2.0	3.0	2.5

TABLE VI
EFFECT OF THE NUMBER OF GMM COMPONENTS ON THE CROSS-DATABASE PERFORMANCE

Train on:	CASIA		Replay		MSU		Average
Test on:	Replay	MSU	CASIA	MSU	CASIA	Replay	
64	28.1	22.8	18.8	36.9	28.4	31.3	27.7
128	27.3	20.8	27.7	33.1	26.7	28.4	27.4
256	26.9	19.1	23.2	31.7	24.2	29.7	**25.8**
512	25.3	24.3	25.3	32.1	26.4	32.1	27.6

C. Comparison with the state of the art

Tables VII and VIII provide a comparison between the results of our proposed approach and those of the state-of-the-art methods in both intra-database and cross-database evaluation. In intra-database evaluation (Table VII), our proposed approach achieves the best performance on two databases: CASIA FASD and MSU MFSD. On the Replay-Attack Database, our obtained results are very competitive compared to the state-of-the-art methods. Note that the best performing method on the Replay-Attack Database (i.e. Motion mag+LBP [6]) gives low performance on the CASIA FASD whereas our proposed methods is able to perform equally well across all the three datasets.

Most importantly, the inter-database evaluation (Table VIII) demonstrates that our proposed CSURF approach outperforms all the state-of-the-art methods. The CSURF based face description yields in very promising generalization capabilities, even when only limited training data is used. Hence, our new features and encoding methods seems to better describe the inherent disparities in color information across various conditions.

TABLE VII
COMPARISON BETWEEN OUR PROPOSED COUNTERMEASURE AND STATE-OF-THE-ART METHODS IN INTRA-DATABASE TESTS

Method	Replay-Attack		CASIA	MSU
	EER	HTER	EER	EER
Motion [4]	11.6	11.7	26.6	-
LBP [17]	13.9	13.8	18.2	
LBP-TOP [5]	7.9	7.6	10.0	-
Motion mag+LBP [6]	0.2	**0.0**	14.4	-
IQA[11]	-	15.2	32.4	-
CNN [14]	6.1	2.1	7.4	-
DMD [7]	5.3	3.8	21.8	-
IDA [12]	-	7.4	-	8.5
Motion+LBP [29]	4.5	5.1	-	-
Color texture [15]	0.4	2.9	6.2	-
Color texture [22]	0.0	3.5	3.2	3.5
Proposed method	0.1	2.2	**2.8**	**2.2**

TABLE VIII
COMPARISON BETWEEN OUR PROPOSED COUNTERMEASURE AND STATE-OF-THE-ART METHODS IN CROSS-DATABASE TESTS

Train on video:	CASIA		Replay		MSU		Average
Test on:	Replay	MSU	CASIA	MSU	CASIA	Replay	
Motion [13]	45.2	-	47.9	-	-	-	46.5
LBP [13]	45.9	-	57.6	-	-	-	51.7
LBP-TOP [13]	49.7	-	60.6	-	-	-	55.1
Motion-Mag [6]	50.1	-	47.0	-	-	-	48.5
CNN [14]	48.5	-	45.5	-	-	-	47.0
Color texture [15]*	37.9	21.0	35.4	32.9	45.7	44.8	36.3
Color texture [22]	30.3	20.4	37.7	34.1	46.0	33.9	33.7
Our method	26.9	19.1	23.2	31.8	24.3	29.7	**25.8**

* the results are re-computed using the frame based scenario instead of the video based scenario.

V. CONCLUSION

We proposed a face anti-spoofing scheme based on color SURF (CSURF) features and Fisher Vector encoding. We extracted the SURF features from two different color spaces (HSV and YCbCr). Then, we applied PCA and Fisher Vector encoding on the concatenated features. The proposed approach based on fusing the features extracted from the HSV and YCbCr was able to perform very well on three most challenging face spoofing datasets, outperforming state of the art results. More importantly, our proposed approach yielded in very interesting generalization performance in the inter-database experiments even when only limited training data was used. As a future work, we plan to investigate other strategies for creating more robust feature spaces for spoofing detection, including person-specific adaptation of anti-spoofing models [30].

REFERENCES

[1] Y. Li, K. Xu, Q. Yan, Y. Li, and R. H. Deng, "Understanding OSN-based facial disclosure against face authentication systems," in *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security*, ser. ASIA CCS '14. ACM, 2014, pp. 413–424.

[2] A. Anjos, J. Komulainen, S. Marcel, A. Hadid, and M. Pietikäinen, "Face anti-spoofing: visual approach," in *Handbook of biometric anti-spoofing*, S. Marcel, M. S. Nixon, and S. Z. Li, Eds. Springer, 2014, ch. 4, pp. 65–82.

[3] J. Galbally, S. Marcel, and J. Fiérrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[4] A. Anjos and S. Marcel, "Counter-measures to photo attacks in face recognition: a public database and a baseline," in *Proceedings of IAPR IEEE International Joint Conference on Biometrics (IJCB)*, 2011.

[5] T. de Freitas Pereira, J. Komulainen, A. Anjos, J. M. De Martino, A. Hadid, M. Pietikäinen, and S. Marcel, "Face liveness detection using dynamic texture," *EURASIP Journal on Image and Video Processing*, 2013.

[6] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, and S. Richa, "Computationally efficient face spoofing detection with motion magnification," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Biometrics*, 2013.

[7] S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, and A. T. S. Ho, "Detection of face spoofing using visual dynamics," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 762–777, 2015.

[8] J. Komulainen, A. Hadid, and M. Pietikäinen, "Context based face anti-spoofing," in *Proc. International Conference on Biometrics: Theory, Applications and Systems (BTAS 2013)*, 2013.

[9] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using micro-texture analysis," in *Proceedings of International Joint Conference on Biometrics (IJCB)*, 2011.

[10] J. Yang, Z. Lei, S. Liao, and S. Z. Li, "Face liveness detection with component dependent descriptor," in *IAPR International Conference on Biometrics, ICB*, June 2013.

[11] J. Galbally and S. Marcel, "Face anti-spoofing based on general image quality assessment," in *Proc. IAPR/IEEE Int. Conf. on Pattern Recognition, ICPR*, 2014, pp. 1173–1178.

[12] D. Wen, H. Han, and A. Jain, "Face spoof detection with image distortion analysis," *Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 746–761, 2015.

[13] T. de Freitas Pereira, A. Anjos, J. De Martino, and S. Marcel, "Can face anti-spoofing countermeasures work in a real world scenario?" in *International Conference on Biometrics (ICB)*, June 2013, pp. 1–8.

[14] J. Yang, Z. Lei, and S. Z. Li, "Learn convolutional neural network for face anti-spoofing," *CoRR*, vol. abs/1408.5601, 2014. [Online]. Available: http://arxiv.org/abs/1408.5601

[15] Z. Boulkenafet, J. Komulainen, and A. Hadid, "Face anti-spoofing based on color texture analysis," in *IEEE International Conference on Image Processing (ICIP2015)*, 2015.

[16] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face antispoofing database with diverse attacks," in *5th IAPR International Conference on Biometrics (ICB)*, 2012, pp. 26–31.

[17] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *International Conference of the Biometrics Special Interest Group (BIOSIG)*, Sept 2012, pp. 1–7.

[18] T. Ojala, M. Pietikäinen, and T. Mäenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 24, no. 7, pp. 971–987, Jul 2002.

[19] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 143–156.

[20] H. Bay, T. Tuytelaars, and L. Gool, *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, ch. SURF: Speeded Up Robust Features, pp. 404–417. [Online]. Available: http://dx.doi.org/10.1007/11744023_32

[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[22] Z. Boulkenafet, J. Komulainen, and A. Hadid, "Face spoofing detection using colour texture analysis," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 1818–1830, 2016.

[23] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[24] K. Simonyan, O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Fisher Vector Faces in the Wild," in *British Machine Vision Conference*, 2013.

[25] J. Sánchez and F. Perronnin, "High-dimensional signature compression for large-scale image classification," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1665–1672.

[26] A. V. Ken Chatfield, Victor Lempitsky and A. Zisserman, "The devil is in the details: an evaluation of recent feature encoding methods," in *Proceedings of the British Machine Vision Conference*, 2011, pp. 76.1–76.12.

[27] F. Perronnin, J. Sánchez, and T. Mensink, *Computer Vision – ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, ch. Improving the Fisher Kernel for Large-Scale Image Classification, pp. 143–156. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15561-1_11

[28] C. M. Bishop, "Pattern recognition," *Machine Learning*, vol. 128, 2006.

[29] J. Komulainen, A. Anjos, A. Hadid, S. Marcel, and M. Pietikäinen, "Complementary countermeasures for detecting scenic face spoofing attacks," in *IAPR International Conference on Biometrics*, 2013.

[30] J. Yang, D. Yi, and S. Z. Li, "Person-specific face anti-spoofing with subject domain adaptation," *IEEE Transactions on Information Forensics and Security*, 2015.