

An Auxiliary Method Based on Hyperspectral Reflectance for Presentation Attack Detection

Shiwei Li, Mohsen Ardabilian, Abdel-Malek Zine

▶ **To cite this version:**

Shiwei Li, Mohsen Ardabilian, Abdel-Malek Zine. An Auxiliary Method Based on Hyperspectral Reflectance for Presentation Attack Detection. The 15th International Conference on Signal Image Technology & Internet based Systems, Nov 2019, Sorrento, Italy. 10.1109/SITIS.2019.00042 . hal-02875939

HAL Id: hal-02875939
https://hal.science/hal-02875939

Submitted on 19 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Auxiliary Method Based on Hyperspectral Reflectance for Presentation Attack Detection

1st Shiwei Li
LIRIS UMR5205
Ecole Centrale de Lyon
Lyon, France
shiwei.li@ec-lyon.fr

2nd Mohsen Ardabilian
LIRIS UMR5205
Ecole Centrale de Lyon
Lyon, France
Mohsen.Ardabilian@ec-lyon.fr

3rd Abdel-Malek Zine
ICJ UMR5208
Ecole Centrale de Lyon
Lyon, France
abdel-malek.zine@ec-lyon.fr

***Abstract*—Face recognition has reached a high accuracy in recent years by adopting convolutional neural networks. However, it suffers from presentation attacks such as 2D face photos, and 3D masks. The vulnerability of face recognition and presentation attacks detection (PAD) attract numerous researchers in recent years. Most studies have only focused on PAD algorithms by analyzing texture information, depth information or thermal images. On the other hand, hyperspectral reflectance, which benefits from the development of line-scan HSI sensors, makes it possible to detect information about the inner structure of materials. Our research proposes an auxiliary method to support face recognition by analyzing hyperspectral reflectance. Combined with biological facts of human skin, we trained a neural network with pigmentation fractions inside human skin and corresponding reflectance. The results show high accuracy in identifying skin and non-skin.**

***Index Terms*—presentation attack detection, hyperspectral reflectance, neural network**

I. Introduction

Face recognition is widely used in biometric domain considering its nearly perfect performance. This technique brings much convenience to people's daily life, such as entrance authentication, or paying with your face. The earliest study on face recognition started in the 1950s in psychology and the classification method has evolved from PCA, SVM to ANN [1]. Particularly, deep learning techniques are applied like FaceNet [2], VGG-Face [3], and LightCNN [4]. These face recognition systems achieve high recognition accuracy. However, presentation attacks that can spoof face recognition systems have been developed at the same time and become major security issues. To our best knowledge, the main presentation attacks are printed face photos, face videos and 3D masks of authenticated users.

The vulnerability of face recognition against these spoofing attacks makes PAD especially important. In recent years, PAD has gained increasing interest from researchers in order to guarantee the safety and suitability of face recognition systems. Ramachandra and Busch provided a complete review of face PAD techniques, available databases and extensive analysis on state-of-the-art approaches [5]. Generally speaking, PAD algorithms can be divided into two types in terms of data acquisitions: 2D and 3D. 2D-PAD algorithms aim to identify printed face photos or videos from bona-fide presentations and are separated into three main categories based on the information analyzed: motion analysis, texture analysis, and life sign detection [6]. 3D-PAD algorithms address more complicated information involved with depth and reach higher accuracy while the devices are not cost-effective. A new countermeasure based on foreground and background motion correlation together with the optical flow is tested on the PHOTO-ATTACK database and reaches an equal-error rate of 1.52% [7]. Based on wavelet, Wild et al. propose an approach that extracts texture information from the grey level co-occurrence matrix of face photos for liveness detection without installing any extra equipment [8]. Some research considering eyes blinking, head movement or dynamic frequency information as visual rhythms [9], also shows a superb performance in PAD. Wen et al. study on an efficient and robust PAD algorithm based on image distortion analysis to separate genuine and spoof faces [10]. Despite these significant progress, presentation attacks techniques are also changing rapidly. With the appearance of custom made face masks and deepfake videos trained by generative adversarial network [11], PAD meets new challenges. A recent study shows that CNN based face recognition systems are vulnerable to custom silicone masks [12] and several PAD techniques are validated on custom masks database [13].

On the other hand, spectral reflectance is an efficient indicator to detect presentation attacks due to its difference among materials. Zhang et al. propose a novel liveness detection method by analyzing multispectral reflectance based on the Lambertian model [14]. Two spectral bands 850nm and 1450nm are chosen to differentiate between genuine faces and artefacts. Furthermore, Kose and Dugelay decompose the image into illumination and reflectance components using the variational retinex algorithm, then the reflectance component is reshaped as a feature vector for classification [15]. Since existing research only considers several independent spectral bands and the variance between genuine faces and artefacts in long-wavelength bands especially NIR, we propose a novel method derived from light-skin interaction and skin structure. A skin model has been built in the viewpoint of tissue optics. Subsequently, skin hyperspectral reflectance (SHSR) database was generated by Monte-Carlo simulations and a skin-network was trained. By fitting the reflectance curve over the ensemble of spectral bands of objects and analyzing the performance,

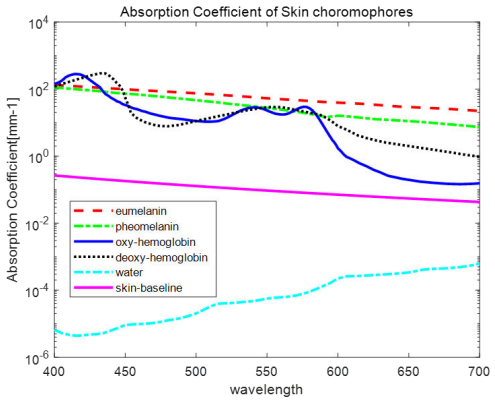


Fig. 1. Absorption coefficients of several pigments inside skin

which differentiates real skin from other materials, we can effectively identify genuine faces.

II. Skin Model And Database

Human skin consists of the epidermis, the dermis, and the subcutaneous tissue generally. And inside skin, pigments such as melanin, play the main role in the color appearance. The absorption coefficients of these pigments change as the wavelength changes as shown in Fig. 1, and that is the key factor why we can find out real skin among artefacts. The uniqueness of these absorption coefficients makes it possible to build a skin model interacting with light. Considering the effect of wavelength on penetration in light-skin interaction [16], a two-layered skin model that consists of the epidermis and the dermis has been built. 8 parameters including melanin volume fraction C_m, pheomelanin and eumelanin ratio $ratio_{ph_eu}$, blood volume fraction C_{bl}, oxygen saturation S, water volume fraction in both the epidermis and the dermis C_{w_epi}, C_{w_der}, and thickness of the epidermis and the dermis d_{epi}, d_{der} are chosen to describe this skin model. With a more complicated skin model, it comes more details and more computational resources. Here, a two-layered skin structure is adequate to distinguish from other materials. Afterward, we need to figure out the relation between these parameters and the optical coefficients including the absorption coefficient and the scattering coefficient in each layer.

A. Epidermis

As the outermost layer of skin, the epidermis has melanin, a broad term for a group of natural pigments that causes human skin to darken. Hence, the absorption coefficient of the epidermis μ_{a_epi} is dominated by melanin:

$$\begin{aligned}\mu_{a_epi} = C_m * (\mu_{a_ph} * ratio_{ph_eu} \\ + \mu_{a_eu} * (1 - ratio_{ph_eu})) \\ + (1 - C_m) * C_{w_epi} * \mu_{a_water} \\ + (1 - C_m) * (1 - C_{w_epi}) * \mu_{a_baseline}\end{aligned} \tag{1}$$

where μ_{a_ph} and μ_{a_eu} are the absorption coefficients of pheomelanin and eumelanin respectively; $\mu_{a_baseline}$ is the absorption coefficient of skin tissue without any pigments. The melanin volume fraction C_m varies from 0.013 (lightly pigmented) to 0.43 (darkly pigmented) [17]. The ratio between the fraction of two types of melanin $ratio_{ph_eu}$ varies from 0.049 to 0.36 which is much too diverse among individuals [18].

The scattering coefficient is defined by a wavelength dependent equation:

$$\begin{aligned}\mu_{s_epi} = 66.7(cm^{-1}) * (f_{Ray} * (\frac{\lambda}{500(nm)})^{-4} \\ + (1 - f_{Ray}) * (\frac{\lambda}{500(nm)})^{-0.689})\end{aligned} \tag{2}$$

where f_{Ray} is the fraction of Rayleigh scattering.

B. Dermis

The dermis that consists of dense irregular connective tissue is adjacent to the epidermis. Blood vessels are deposited in this layer, which means hemoglobin, another important pigment inside human skin, primarily locates here. The absorption and scattering coefficients are shown respectively below:

$$\begin{aligned}\mu_{a_der} = C_{bl} * (\mu_{a_oxy} * S + \mu_{a_deoxy} * (1 - S)) \\ + (1 - C_{bl}) * C_{w_der} * \mu_{a_water} \\ + (1 - C_{bl}) * (1 - C_{w_der}) * \mu_{a_baseline}\end{aligned} \tag{3}$$

$$\begin{aligned}\mu_{s_der} = 43.6(cm^{-1}) * (f_{Ray} * (\frac{\lambda}{500(nm)})^{-4} \\ + (1 - f_{Ray}) * (\frac{\lambda}{500(nm)})^{-0.562})\end{aligned} \tag{4}$$

where μ_{oxy}, μ_{deoxy} represent the absorption coefficients of oxy-hemoglobin and deoxy-hemoglobin.

C. Database

With the help of the skin model, Monte Carlo simulations are carried out to acquire the hyperspectral reflectance. Fig. 2 explains how it works in a single photon cycle. 8 parameters are randomly assigned values within the scope of Table I which is in line with biological facts. In total, 50000 samples are generated in SHSR database with wavelength varying from 450nm to 700nm in about 18.58 hours.

TABLE I
Lower bounds (LB) and upper bounds (UB) of 8 parameters

	C_m	$ratio_{ph_eu}$	C_{w_epi}	C_{bl_der}
LB	1.3%	4.9%	10%	0.2%
UB	43%	36%	20%	7%
	S	C_{w_der}	d_{epi}	d_{der}
LB	70%	40%	$0.027mm$	$0.6mm$
UB	95%	90%	$0.15mm$	$3mm$

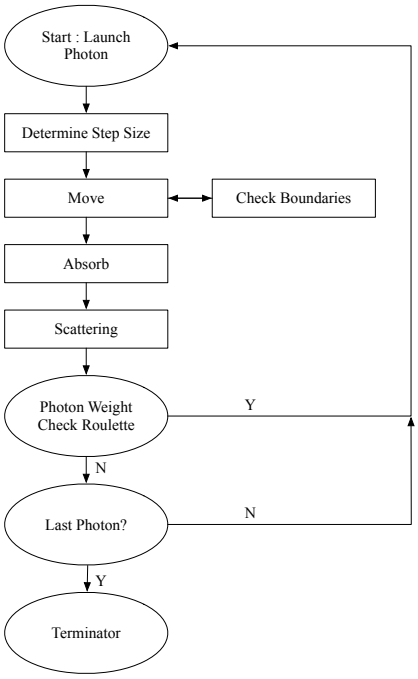


Fig. 2. Flowchart of Monte Carlo simulations

III. Experiments And Results

While the database has been built, we train a neural network between 8 parameters and the hyperspectral reflectance. The Skin-Net is trained respectively by varying hidden layers from 1 to 3 with alternate neurons from 35 to 55 (an increment 10). To verify that it matches well with tissue optics, 30 hyperspectral reflectance samples of genuine skin are validated. All 8 parameters must be assigned within the realistic range to ensure its applicability. Fig. 3 shows the validation flowchart, and the error is computed as the mean square error (MSE):

$$MSE = \frac{\sum_{\lambda_{initial}}^{\lambda_{final}} (Pre_\lambda - Tar_\lambda)^2}{Nb_{wavelength}} \quad (5)$$

where Pre_λ and Tar_λ are hyperspectral reflectance predicted and targeted;$Nb_{wavelength}$ is the number of spectral bands. The validation process is iterated to reach the minimum error by adjusting 8 parameters and fitting reflectance predicted into genuine skin reflectance. This curve-fitting process is executed using "fmincon" built-in interior-point algorithm in MATLAB 2018a.

The results show that Skin-Net can estimate very similar reflectance. The mean error of all 30 samples validations is $3.66e^{-05}$. In Fig. 4, hyperspectral reflectance of genuine skin no.2 and the one predicted by Skin-Net are illustrated. These two reflectance curves are quite close among 450nm-700nm. Skin-Net characteristics are listed below:

- Skin model is assumed to be two-layered structure
- Reflectance predicted only concerns in the visible light range
- Network is trained in database generated by Monte Carlo simulations

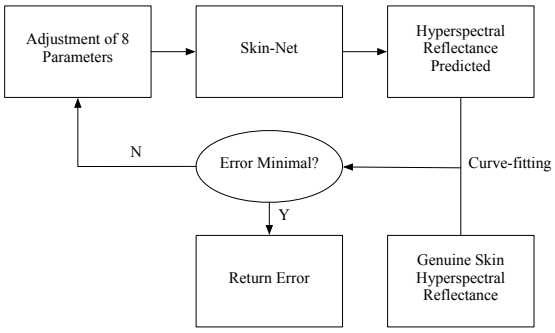


Fig. 3. Flowchart of Skin-Net validation: fitting the hyperspectral reflectance predicted to genuine skin

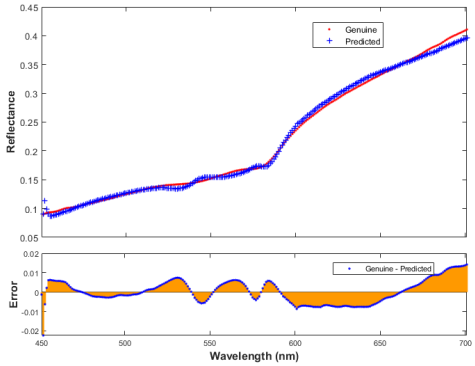


Fig. 4. Validation of genuine skin no.2

As we mentioned above, other materials cannot interact with light in the same way skin does. Therefore, we take Skin-Net and curve-fitting process to do liveness detection tasks. 270 hyperspectral reflectance of objects including colored papers and newsprint which comes from the public database of the University of Eastern Finland together with 30 samples of genuine skin are classified. All MSE and STD are computed as the basis when the curve-fitting process reaches an optimal solution. For the reason that Skin-Net has a great performance in fitting the genuine skin reflectance curve when it refers to other materials, the results are significantly different. Here, we identify genuine skin by setting thresholds of MSE and STD. We define two categories: (a) at least one of them is less than the threshold; (b) MSE and STD are less than the thresholds at the same time. 0.001 and 0.05 are separately assigned to MSE and STD. Table II reveals the classification results. For category (a), skin objects are all classified in the correct group, while 24 objects (near 8.9%) are wrongly accepted. And we find that even in a more strict situation, there is a paper object accepted. Fig. 5 shows the STD-MSE distribution of 300 samples. Compared to the existing research, the spectral band 930nm has shown the lowest error rate: 6% and 8.5% attack presentation classification error rate (APCER), 3% and 0% bona-fide presentation attack classification error rate (BPCER) with the classification method LBP-SVM and BSIF-SVM [19]. Our results still need to be improved. By analyzing the fitting

hyperspectral reflectance of the paper object accepted, we decide to focus on the spectral bands 500nm-600nm. As shown in Fig. 6, the main variance occurs from 500nm to 600nm. The wavelength scope for calculating MSE is adjusted, and "fmincon" is applied once again to get new results.

TABLE II
CLASSIFICATION RESULTS IN TWO CATEGORIES

MSE $\leq$ 0.001 or STD $\leq$ 0.05	**Acceptance**	**Rejection**
Non-skin	24	246
Skin	30	0
MSE $\leq$ 0.001 and STD $\leq$ 0.05	**Acceptance**	**Rejection**
Non-skin	1	269
Skin	28	2

The mean error of all 30 skin samples validations has decreased to $3.51e^{-06}$. In contrast, MSE of non-skin samples augments. Because one important pigment of skin, hemoglobin, has a unique curve of the absorption coefficients from 500nm to 600nm. It is similar to "W" in shape, which is

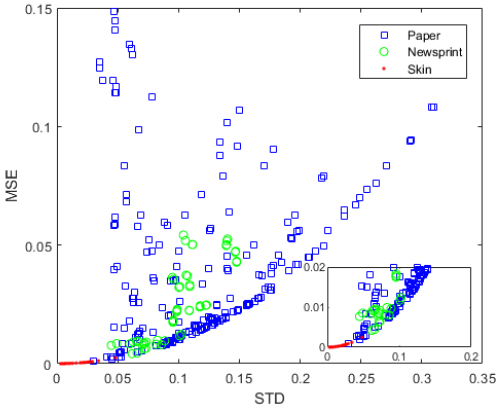


Fig. 5. MSE and STD values of 300 samples are calculated after finishing curve-fitting process. Blue squares, green circles and red points denote paper, newsprint and skin respectively. The bottom right sub-figure is a zoom in for MSE varying from 0 to 0.02.

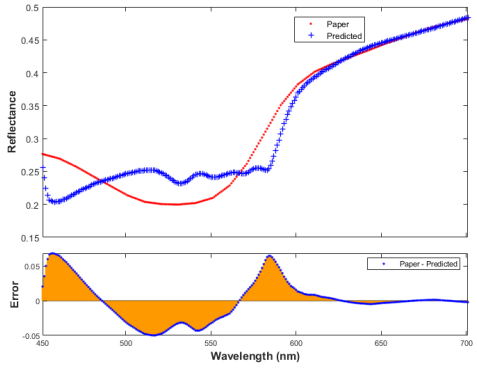


Fig. 6. Fitting process on paper object no.59



Fig. 7. face photo and genuine face at 400nm, 500nm, 600nm, 700nm

called "W" pattern [20]. After applying this new wavelength range, both the APCER and the BCER reach 0%. With an appropriate wavelength range, the validation performance becomes much better. The wavelength range selection means that finding the range in which the significant difference between genuine skin and other materials locates. From the view of the skin, we simply need to check the unique patterns like "W" shape in hyperspectral reflectance curve. On the other hand, multi-validation with multi-wavelength ranges can reach satisfying results as well.

Limited to the spectra database, it needs more hyperspectral images data to validate. To our best knowledge, there isn't yet a public database about hyperspectral face images. We found a set of hyperspectral face images that only concern one person from CAVE [21] as shown in Fig. 7. These images vary from 400nm to 700nm with a step 10nm captured by a cooled CCD camera (Apogee Alta U260) under the light source CIE Standard Illuminant D65. As a start, we apply our Skin-Net to identify these images. All 31 images are in format PNG (16bit), and consists of 512*512 pixels. We denote a tensor $I_{i,j,k}$ which represents these 31 images, where i,j,k are the height (H), the width (W), and the spectra bands (S) respectively. Considering the hyperspectral reflectance, k is expanded from 31 to 301 by using the interpolation algorithm. Then, $I_{i,j,k}$ should locate on faces, in other words, i and j are restricted. Several points are randomly taken in the faces range including face photo and genuine face defined by $P_{h,w}$. The metric to validate these points is shown below:

$$MSE = \frac{\sum\sum_{h-1,w-1}^{h+1,w+1}(Ref(P_{h,w}) - Pre(P_{h,w}))}{9} \quad (6)$$

where Ref() is the hyperspectral reflectance of point $P_{h,w}$; Pre() represents the reflectance predicted by Skin-Net of point $P_{h,w}$.

TABLE III
MSE RESULTS IN TWO WAVELENGTH RANGE

	P1	P2	P3
450nm-700nm	$1.80e^{-05}$	$1.40e^{-05}$	$1.37e^{-05}$
500nm-600nm	$2.70e^{-06}$	$1.38e^{-05}$	$3.30e^{-06}$
	P4	P5	P6
450nm-700nm	$1.91e^{-04}$	$1.66e^{-04}$	$1.58e^{-04}$
500nm-600nm	$6.31e^{-05}$	$5.43e^{-05}$	$5.17e^{-05}$

That is to say, a 3*3*301 reflectance tensor is fitted by Skin-Net, and the mean square error is returned back to validate the performance. Pixel points in both face photo and genuine face are taken into the fitting process. Results are calculated in two wavelength range: 450nm-700nm and 500nm-600nm. From

Table III, we can find that the former three points have fewer errors than the latter three. Generally, the errors are in different orders of magnitude, like the latter three are about 10 times larger than the former three. The fact is that the former three points are taken from the genuine face, and the latter three belong to face photos. Regardless of from which sources these points are taken, the errors for range 450nm-700nm show signs of having greater difference than range 500nm-600nm. This may be not only because of a larger wavelength range adopted, but also the degree of skin pigmented. This means that the degree of skin pigmented influences fitting results. For skin lightly-pigmented, the melanin volume fraction varies between 1.3% and 3%, and the effect of blood to skin appearance becomes more significant. Thus, it is better to choose the range 500nm-600nm where the absorption coefficient of hemoglobin changes rapidly. As for skin darkly-pigmented, due to the substantial melanin volume fraction and its smooth change over the spectra, it's necessary to adopt a larger wavelength range.

IV. Conclusions

Our research proposes an auxiliary method for PAD by analyzing the hyperspectral reflectance. We firstly trained a network Skin-Net with a spectra database generated by Monte Carlo simulations. Unlike the Lambertian model, Monte Carlo simulations are closer to reality. Then, Skin-Net is applied to detect if the spectra belong to the skin or not with the help of curve-fitting. Two spectra ranges are studied for better performance and robustness. Based on tissue optics and biological facts of skin, it's much efficient to detect artefacts. Due to a lack of hyperspectral face image database, we only check this method on hyperspectral photos of one person. The results between face photos and genuine face are quite different. That would be used as an indicator to detect presentation attacks. We also discuss the wavelength range selection. Our proposed method has the potential to develop, and it requires more hyperspectral face images to validate the suitability. Our perspective is to collect hyperspectral face images with a snap-scan VNIR range system integrating a linescan HSI sensor (IMEC CMV2K LS150+VNIR) and an image sensor (CMOSIS CMV2000). With more images data to validate, we'll apply our method to face recognition system for supporting PAD.

References

[1] S. G. Bhele and V. Mankar, "A review paper on face recognition techniques," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 1, no. 8, pp. 339–346, 2012.

[2] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 815–823.

[3] O. M. Parkhi, A. Vedaldi, A. Zisserman *et al.*, "Deep face recognition." in *bmvc*, vol. 1, no. 3, 2015, p. 6.

[4] X. Wu, R. He, Z. Sun, and T. Tan, "A light cnn for deep face representation with noisy labels," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2884–2896, 2018.

[5] R. Ramachandra and C. Busch, "Presentation attack detection methods for face recognition systems: A comprehensive survey," *ACM Computing Surveys (CSUR)*, vol. 50, no. 1, p. 8, 2017.

[6] O. Kähm and N. Damer, "2d face liveness detection: An overview," in *2012 BIOSIG-Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*. IEEE, 2012, pp. 1–12.

[7] A. Anjos, M. M. Chakka, and S. Marcel, "Motion-based counter-measures to photo attacks in face recognition," *IET biometrics*, vol. 3, no. 3, pp. 147–158, 2013.

[8] P. Wild, P. Radu, L. Chen, and J. Ferryman, "Robust multimodal face and fingerprint fusion in the presence of spoofing attacks," *Pattern Recognition*, vol. 50, pp. 17–25, 2016.

[9] A. Pinto, W. R. Schwartz, H. Pedrini, and A. de Rezende Rocha, "Using visual rhythms for detecting video-based facial spoof attacks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 1025–1038, 2015.

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

[11] P. Korshunov and S. Marcel, "Deepfakes: a new threat to face recognition? assessment and detection," *arXiv preprint arXiv:1812.08685*, 2018.

[12] S. Bhattacharjee, A. Mohammadi, and S. Marcel, "Spoofing deep face recognition with custom silicone masks," in *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, 2018, pp. 1–7.

[13] R. Ramachandra, S. Venkatesh, K. B. Raja, S. Bhattacharjee, P. Wasnik, S. Marcel, and C. Busch, "Custom silicone face masks: Vulnerability of commercial face recognition systems & presentation attack detection," in *2019 7th International Workshop on Biometrics and Forensics (IWBF)*. IEEE, 2019, pp. 1–6.

[14] Z. Zhang, D. Yi, Z. Lei, S. Z. Li *et al.*, "Face liveness detection by learning multispectral reflectance distributions." in *FG*, 2011, pp. 436–441.

[15] N. Kose and J.-L. Dugelay, "Reflectance analysis based countermeasure technique to detect face mask attacks," in *2013 18th International Conference on Digital Signal Processing (DSP)*. IEEE, 2013, pp. 1–6.

[16] C. Ash, M. Dubec, K. Donne, and T. Bashford, "Effect of wavelength and beam width on penetration in light-tissue interaction using computational methods," *Lasers in medical science*, vol. 32, no. 8, pp. 1909–1918, 2017.

[17] S. L. Jacques, "Origins of tissue optical properties in the uva, visible, and nir regions," *OSA TOPS on advances in optical imaging and photon migration*, vol. 2, pp. 364–369, 1996.

[18] G. V. Baranoski and A. Krishnaswamy, *Light and skin interactions: simulations for computer graphics applications*. Morgan Kaufmann, 2010.

[19] R. Raghavendra, K. B. Raja, S. Venkatesh, F. A. Cheikh, and C. Busch, "On the vulnerability of extended multispectral face recognition systems towards presentation attacks," in *2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. IEEE, 2017, pp. 1–8.

[20] W. Chen, M. Ardabilian, A. Zine, and H. Zahouani, "Reflectance spectra based skin and non-skin classification," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 755–759.

[21] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE transactions on image processing*, vol. 19, no. 9, pp. 2241–2253, 2010.