Dermoscopy Image Analysis: Overview and Future Directions

M. Emre Celebi, Noel Codella, and Allan Halpern

***Abstract*—Dermoscopy is a non-invasive skin imaging technique that permits visualization of features of pigmented melanocytic neoplasms that are not discernable by examination with the naked eye. While studies on the automated analysis of dermoscopy images date back to the late 1990s, because of various factors (lack of publicly available datasets, open-source software, computational power, etc.), the field progressed rather slowly in its first two decades. With the release of a large public dataset by the International Skin Imaging Collaboration in 2016, development of open-source software for convolutional neural networks, and the availability of inexpensive graphics processing units, dermoscopy image analysis has recently become a very active research field. In this paper, we present a brief overview of this exciting subfield of medical image analysis, primarily focusing on three aspects of it, namely, segmentation, feature extraction, and classification. We then provide future directions for researchers.**

***Index Terms*—Skin cancer, melanoma, dermoscopy, computer-aided diagnosis, dermoscopy image analysis.**

I. Overview of the Field

SKIN cancer is the most common cancer in the United States, with over 5 million cases diagnosed each year [1]. Melanoma is the deadliest form of skin cancer, with roughly 91,000 new cases every year in the US and more than 9,000 deaths [2]. The treatment of melanoma costs over $3 billion per year in the US alone [3]. Skin cancer poses a major public health threat internationally as well. In Australia, more than 14,000 new cases of melanoma are reported yearly, leading to nearly 2,000 deaths [4]. In Europe, over 100,000 new melanoma cases and 22,000 melanoma related deaths are reported annually [5]. What is perhaps more alarming is that unlike many other cancer types, the incidence rate of melanoma has been steadily increasing over the past several decades. For example, between 1990 and 2018, a 225% increase has been observed in the US [2], [6].

In the past, the primary form of diagnosis for melanoma was unaided clinical examination, which has limited and variable accuracy, leading to significant challenges both in the early detection of disease and the minimization of unnecessary biopsies. In recent years, dermoscopy, a high-resolution skin imaging technique that allows visualization of deeper skin structures by reducing surface reflectance, has improved the diagnostic capability of trained specialists [7]. Unfortunately, dermoscopy remains difficult to learn [8] and several studies have demonstrated limitations of dermoscopy when proper training is not administered [9], [10]. In addition, even with sufficient training, visual analysis remains subjective [11].

Newer imaging technologies such as infrared imaging, multispectral imaging, and confocal microscopy, have recently come to the forefront in providing the potential for higher diagnostic accuracy [12]–[14]. In addition, various studies since the late 1990s have been focused on developing algorithms for the automated analysis of dermoscopy images [15]–[21]. Combinations of such technologies have the potential to serve as adjuncts to physicians, improving clinical management, especially for high-risk patients.

In this paper, we primarily focus on dermoscopy because of its growing use in clinical practice throughout the world [22]–[24]. The popularity of dermoscopy is also readily seen in the computational literature, where an overwhelming majority of the papers use dermoscopy as their primary modality. The term *dermoscopy image analysis* (DIA) appears to have been coined by Menzies *et al.* [25] and popularized by Celebi *et al.* [15], [26], [27]. For historical surveys of DIA, the reader is referred to [27]–[30].

In this section, we will discuss three primary aspects of DIA: segmentation, feature extraction, and classification. In the next section, we will talk about other, relatively less explored, issues as well. Segmentation (a.k.a. lesion border detection) essentially involves the localization of lesions in the image. Along with preprocessing, segmentation commonly forms the first step in the DIA pipeline. In fact, segmentation is one of the few steps in DIA that requires only a limited amount of domain knowledge, which partly explains the abundance of segmentation related studies in the DIA literature [31], [32]. The great variety of lesion shapes, sizes, and colors; skin types and textures; intrinsic cutaneous features (blood vessels, hair [33], etc.) as well as inconsistent imaging conditions [34] make it difficult to develop a robust segmentation algorithm. An accurate segmentation may be important for the extraction of informative features, although whether this is definitely true or not has not been established conclusively yet [31], [35], [36]. Historically, clustering [37] and thresholding [38] based approaches

Manuscript received January 22, 2019; accepted January 24, 2019. Date of publication January 28, 2019; date of current version March 6, 2019. *(Corresponding author: M. Emre Celebi.)*

M. E. Celebi is with the Department of Computer Science, University of Central Arkansas, Conway, AR 72035 USA (e-mail: ecelebi@uca.edu).

N. Codella is with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: nccodell@us.ibm.com).

A. Halpern is with the Dermatology Service, Memorial Sloan Kettering Cancer Center, New York, NY 10065 USA (e-mail: halperna@mskcc.org).

Digital Object Identifier 10.1109/JBHI.2019.2895803

have dominated the dermoscopy image segmentation literature [31], [32]. With the advent of deep learning [39], [40], however, various segmentation-free classification systems have started appearing in the literature. It should be noted that segmentation may still be important for temporal analysis of skin lesions (see the next section).

In feature extraction (a.k.a. dermoscopic attribute detection) [36], [41], the goal is to automatically detect and localize clinical dermoscopic criteria previously established in the medical literature, such as pigment networks [42], blue-white areas [43], streaks [44], milia-like cysts, globules [45], global patterns [46]–[49], and number of clinically significant colors [50]–[53]. With varying degrees, these patterns have been found to be correlated with disease states, and incorporated into medical diagnostic algorithms [54].

In classification, early efforts using dermoscopy images were restricted to lesions that were obviously melanocytic [55] (i.e., originating in the pigment cells of the skin). Such lesions were typically categorized into one of two diagnostic classes: benign and melanoma. With the recent availability of larger datasets, studies with more refined categorization of melanocytic lesions (e.g., benign, atypical, and melanoma) have started appearing in the DIA literature. For example, Part 3 of the ISIC 2018 challenge involved seven classes: melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma, and vascular lesion. In early studies, traditional classifiers (e.g., decision trees [43], artificial neural networks [56], and support vector machines [15]) and their ensembles [57], [58] were used. Since 2015 [18], convolutional neural networks (CNNs) have become the classifiers of choice in DIA. In addition to providing high classification accuracy, CNNs alleviate the machine learning expert's burden of "feature engineering" by automatically discovering high-level abstractions from low-level (e.g., pixel) data [59].

II. Future Directions

Early dermoscopy image classification systems typically used low-level visual features (shape, texture, color, etc.) extracted from the lesion of interest (e.g., area/diameter of the lesion, statistical texture features extracted from the gray-level co-occurrence matrix, and mean/variance of select color channels). These low-level features were adopted mostly from the computer vision literature [15] and lacked clinical meaning. Before deep learning became popular in DIA, the feature extraction literature was moving towards high-level (a.k.a. clinically oriented [51], [52]) features. With a push from the European Unions recent General Data Protection Regulation [60], we expect the deep learning literature to move eventually towards transparent, understandable, and explainable approaches [61] as well.

An important issue that has been stifling advances in DIA has been the lack of color standardization. Various color normalization approaches have been developed to address this problem [62]–[65], but the development of approaches that consistently produce natural looking results, while leaving intact those input images which do not require any modifications, have remained elusive. In deep learning applications, the training data could be augmented using various color modifications [66]. Acquisition-time color standardization [67]–[70], however, is likely to be more effective than post-acquisition color normalization. In some medical imaging domains [71], color standardization has been partly addressed. In DIA, we expect the ISIC to have a similar standardizing influence on the image acquisition procedure.

For many years, the lack of large and diverse dermoscopy image sets has hindered the progress in DIA. The publication of the *Interactive Atlas of Dermoscopy* by Argenziano *et al.* (EDRA Medical Publishing & New Media) in 1999, partially mitigated this problem with its accompanying CD-ROM that contained nearly 1,000 dermoscopy images. Another dermoscopy atlas entitled *An Atlas of Surface Microscopy of Pigmented Skin Lesions: Dermoscopy* by Menzies *et al.* (McGraw-Hill) published in 2003, contained about 200 images. These two atlases served as the primary source of images for DIA studies for over a decade. The next breakthrough came with the Ph2 dataset, which contained 200 images, released by Mendonca *et al.* [72] in 2013; see also Mendonca *et al.* [73]. Unfortunately, in total, these datasets contained only about 1,400 images, which was grossly insufficient for certain types of data-hungry classifiers (e.g., CNNs). The ISIC Archive dataset addressed this problem by making over 10,000 images publicly available. This large and heterogeneous dataset, and the three ISIC Challenges organized as part of the IEEE International Symposium on Biomedical Imaging (ISBI 2016 [74] and 2017 [20]) and the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2018 [75]) that made use of various snapshots of the archive, greatly accelerated DIA research. For example, the ISIC 2018 Challenge had over 1,000 registered users and well over 100 papers that used the ISIC Archive datasets have been published over the past three years. Currently, the archive dataset contains over 24,000 images, and this valuable resource continues to expand rapidly since its first release in 2016.

As the ISIC Archive spearheads the advancement of DIA, and the dataset continues to grow, several outstanding issues inhibiting the integration of AI technology in the clinical workflow should become addressable. These include 1) the comprehensive representation of data from an international population in a centralized repository, both for the purposes of training, as well as transparent and thorough testing and validation; 2) the development of explainable AI systems that provide evidence for decisions in order to enable clinical staff of appropriate skill to verify the validity of decisions; and 3) the development of AI systems that are able to recognize lesions or disease states for which either sufficient training data was not available to render a reliable decision, or for which no training data exists at all (this is referred to as *out-of-distribution* detection).

As mentioned in the previous section, the present paper is predominantly about dermoscopy image analysis. Dermoscopy is commonly used by melanoma specialists and has been shown in multiple studies to improve diagnostic accuracy for skin cancer. Dermoscopy also improves the quality and standardization of imaging, as it uses commercial devices that eliminate surface glare and, when used in direct contact with the skin, control lighting. That said, simple close up photography of skin lesions has been shown to permit the successful application of diagnostic AI. In theory, combination of routine photography that

highlights surface texture together with dermoscopy that visualizes subsurface features through magnification and elimination of surface texture, should be complementary and improve diagnostic accuracy. Other subsurface imaging techniques such a spectral imaging, optical coherence tomography, and reflectance confocal microscopy should theoretically further complement the diagnostic information present in dermoscopy images.

Content-based image retrieval (CBIR) has been an active area of research for the past two decades. Starting with a user-supplied query image, a typical CBIR system displays database images in descending order of visual similarity [76] to the query image. Medicine is often cited as one of the principal application domains of CBIR [77]. Potential applications of CBIR in medicine include training, diagnostics (classification), image data mining, and automated annotation. For example, lecturers can use CBIR systems to train medical students to distinguish between cases that have similar visual appearance, but different diagnoses. Alternatively, physicians can diagnose borderline cases more easily by visually comparing them with similar cases. Despite the numerous applications of CBIR in medicine [77], [78], there are very few studies on the content-based retrieval of dermoscopy images [79]–[83], which is most likely because large dermoscopy image datasets have been publicly unavailable until the release of the ISIC Archive dataset in 2016. With this ever-growing dataset, we expect to see more CBIR applications in dermoscopy.

Dysplastic nevi are both potential precursors of melanomas and markers for melanoma-prone individuals [84]. Total body photography (TBP), a.k.a. total body imaging, is a valuable imaging technique that allows monitoring of the evolution of dysplastic nevi, especially in high-risk patients. TBP provides baseline documentation of all of the patient's skin surface that can be used in future visits to facilitate the identification of new lesions, lesions that are undergoing significant morphologic changes, and lesions that are growing at a faster rate than the remainder of the patient's nevi [85]. Photographic monitoring can also reduce unnecessary biopsies by eliminating the need to biopsy lesions that are erroneously perceived by the patient to have changed. TBP can also be combined profitably with other skin imaging modalities. For example, TBP can be first used to detect new lesions as well as changes in previously discovered nevi, and then dermoscopy can be used to examine each nevus in more detail [86]. Short term monitoring of lesions using digital dermoscopy has also been shown to increase diagnostic specificity. So far, there have been a limited number of studies on automated analysis of total body images [87]–[90]. This is most likely because of lack of standardization of the TBP technology as well as privacy issues associated with images of naked patients. Once these problems are addressed, we expect TBP to play a more significant role in skin lesion image analysis.

When diagnosing cases, physicians consider the physical appearance of the case as well as clinical covariates (age, gender, familial history, etc.) associated with the patient. As described earlier, physicians also consider the morphologic appearance of any given lesion in the context of the appearance of the individual's other skin lesions, looking for outliers. Another very important consideration is change over time in the appearance of a lesion relative to the overall changes in the individual's lesions. Such metadata has been largely ignored in the DIA literature, where researchers relied primarily on visual features extracted from images. Future DIA systems should use an integrated approach combining visual features and clinical metadata to achieve higher diagnostic accuracy and wider clinical acceptance.

References

[1] American Cancer Society, Cancer Facts & Figures 2018, 2018. [Online]. Available: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-fact s-figures/cancer-facts-figures-2018.html

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: Cancer J. Clin.*, vol. 68, no. 1, pp. 7–30, 2018.

[3] G. P. Guy, Jr., S. R. Machlin, D. U. Ekwueme, and K. R. Yabroff, "Prevalence and costs of skin cancer treatment in the U.S., 2002–2006 and 2007–2011," *Amer. J. Preventive Med.*, vol. 48, no. 2, pp. 183–187, 2015.

[4] Australian Government, Melanoma of the Skin Statistics, 2019. [Online]. Available: https://melanoma.canceraustralia.gov.au/statistics

[5] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018.

[6] E. Silverberg, C. C. Boring, and T. S. Squires, "Cancer Statistics, 1990," *CA: Cancer J. Clin.*, vol. 40, no. 1, pp. 9–26, 1990.

[7] M. E. Vestergaard, P. Macaskill, P. E. Holt, and S. W. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting," *Brit. J. Dermatol.*, vol. 159, no. 3, pp. 669–676, 2008.

[8] A. M. Forsea *et al.*, "Factors driving the use of dermoscopy in Europe: A pan-European survey," *Brit. J. Dermatol.*, vol. 175, no. 6, pp. 1329–1337, 2016.

[9] M. Binder *et al.*, "Epiluminescence microscopy. a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists," *Arch. Dermatol.*, vol. 131, no. 3, pp. 286–291, 1995.

[10] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," *Lancet Oncol.*, vol. 3, no. 3, pp. 159–165, 2002.

[11] M. L. Bafounta, A. Beauchet, P. Aegerter, and P. Saiag, "Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests," *Arch. Dermatol.*, vol. 137, no. 10, pp. 1343–1350, 2001.

[12] E. Ruocco, G. Argenziano, G. Pellacani, and S. Seidenari, "Noninvasive imaging of skin tumors," *Dermatol. Surgery*, vol. 30, no. 2, pp. 301–310, 2004.

[13] A. Esmaeili, A. Scope, A. C. Halpern, and A. A. Marghoob, "Imaging techniques for the in vivo diagnosis of melanoma," *Seminars Cutaneous Med. Surgery*, vol. 27, no. 1, pp. 2–10, 2008.

[14] R. M. Bakos, T. P. Blumetti, R. Roldan-Marin, and G. Salerni, "Noninvasive imaging tools in the diagnosis and treatment of skin cancers," *Amer. J. Clin. Dermatol.*, vol. 19, pp. S3–S14, 2018.

[15] M. E. Celebi *et al.*, "A methodological approach to the classification of dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 31, no. 6, pp. 362–373, 2007.

[16] H. Iyatomi *et al.*, "An improved Internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm," *Comput. Med. Imag. Graph.*, vol. 32, no. 7, pp. 566–579, 2008.

[17] C. Barata, M. Ruela, M. Francisco, T. Mendonca, and J. S. Marques, "Two systems for the detection of melanomas in dermoscopy images using texture and color features," *IEEE Syst. J.*, vol. 8, no. 3, pp. 965–979, Sep. 2014.

[18] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2015, pp. 118–126.

[19] N. C. F. Codella *et al.*, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM J. Res. Develop.*, vol. 61, no. 4, pp. 5:1–5:15, 2017.

[20] N. C. F. Codella *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC)," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2018, pp. 168–172.

[21] M. A. Marchetti *et al.*, "Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging Challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images," *J. Amer. Acad. Dermatology*, vol. 78, no. 2, pp. 270–277, 2018.

[22] P. Piliouras, P. Buettner, and H. P. Soyer, "Dermoscopy use in the next generation: A survey of Australian dermatology trainees," *Australas. J. Dermatol.*, vol. 55, no. 1, pp. 49–52, 2014.

[23] E. C. Murzaku, S. Hayan, and B. K. Rao, "Methods and rates of dermoscopy usage: A cross-sectional survey of US dermatologists stratified by years in practice," *J. Amer. Acad. Dermatol.*, vol. 71, no. 2, pp. 393–395, 2014.

[24] P. Bahadoran and J. Malvehy, "Dermoscopy in Europe: Coming of age," *Brit. J. Dermatol.*, vol. 175, no. 6, pp. 1132–1133, 2016.

[25] S. W. Menzies *et al.*, "The performance of solarscan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma," *Arch. Dermatol.*, vol. 141, no. 11, pp. 1388–1396, 2005.

[26] M. E. Celebi, "Development of algorithms for dermoscopy image analysis," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Texas, Arlington, TX, USA, 2006.

[27] M. E. Celebi, T. Mendonca, and J. S. Marques, Eds., *Dermoscopy Image Analysis*. Boca Raton, FL, USA: CRC Press, 2015.

[28] K. Korotkov and R. Garcia, "Computerized analysis of pigmented skin lesions: A review," *Artif. Intell. Med.*, vol. 56, no. 2, pp. 69–90, 2012.

[29] J. Scharcanski and M. E. Celebi, Eds., *Computer Vision Techniques for the Diagnosis of Skin Cancer*. New York, NY, USA: Springer, 2013.

[30] N. K. Mishra and M. E. Celebi, "An overview of melanoma detection in dermoscopy images using image processing and machine learning," 2016. [Online]. Available:https://arxiv.org/abs/1601.07843

[31] M. E. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 33, no. 2, pp. 148–153, 2009.

[32] M. E. Celebi, Q. Wen, H. Iyatomi, K. Shimizu, H. Zhou, and G. Schaefer, "A state-of-the-art survey on lesion border detection in dermoscopy images," in *Dermoscopy Image Analysis*, M. E. Celebi, T. Mendonca, and J. S. Marques, Eds. Boca Raton, FL, USA: CRC Press, 2015, pp. 97–129.

[33] Q. Abbas, M. E. Celebi, and I. F. Garcia, "Hair removal methods: A comparative study for dermoscopy images," *Biomed. Signal Process. Control*, vol. 6, no. 4, pp. 395–404, 2011.

[34] C. Barata, M. E. Celebi, and J. S. Marques, "Towards a robust analysis of dermoscopy images acquired under different conditions," in *Dermoscopy Image Analysis*, M. E. Celebi, T. Mendonca, and J. S. Marques, Eds. Boca Raton, FL, USA: CRC Press, 2015, pp. 1–22.

[35] R. Garnavi, M. Aldeen, M. E. Celebi, G. Varigos, and S. Finch, "Border detection in dermoscopy images using hybrid thresholding on optimized color channels," *Comput. Med. Imag. Graph.*, vol. 35, no. 2, pp. 105–115, 2011.

[36] C. Barata, M. E. Celebi, and J. S. Marques, "A survey of feature extraction in dermoscopy image analysis of skin cancer," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2018.2845939.

[37] H. Zhou, G. Schaefer, A. H. Sadka, and M. E. Celebi, "Anisotropic mean shift based fuzzy C-means segmentation of dermoscopy images," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 1, pp. 26–34, Feb. 2009.

[38] M. E. Celebi, Q. Wen, S. Hwang, H. Iyatomi, and G. Schaefer, "Lesion border detection in dermoscopy images using ensembles of thresholding methods," *Skin Res. Technol.*, vol. 19, no. 1, pp. e252–e258, 2013.

[39] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, 2016.

[40] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[41] A. Madooei and M. S. Drew, "Incorporating colour information for computer-aided diagnosis of melanoma from dermoscopy images: A retrospective survey and critical analysis," *Int. J. Biomed. Imag.*, vol. 2016, 2016, Art. no. 4868305.

[42] C. Barata, J. S. Marques, and J. Rozeira, "A system for the detection of pigment network in dermoscopy images using directional filters," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2744–2754, Oct. 2012.

[43] M. E. Celebi *et al.*, "Automatic detection of blue-white veil and related structures in dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 32, no. 8, pp. 670–677, 2008.

[44] M. Sadeghi, T. K. Lee, D. McLean, H. Lui, and M. S. Atkins, "Detection and analysis of irregular streaks in dermoscopic images of skin lesions," *IEEE Trans. Med. Imag.*, vol. 32, no. 5, pp. 849–861, May 2013.

[45] M. G. Fleming *et al.*, "Techniques for a structural analysis of dermatoscopic imagery," *Comput. Med. Imag. Graph.*, vol. 22, no. 5, pp. 375–389, 1998.

[46] H. Iyatomi *et al.*, "Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin," *J. Investigative Dermatol.*, vol. 128, no. 8, pp. 2049–2054, 2008.

[47] C. Serrano and B. Acha, "Pattern analysis of dermoscopic images based on Markov random fields," *Pattern Recognit.*, vol. 42, no. 6, pp. 1052–1057, 2009.

[48] Q. Abbas, M. E. Celebi, C. Serrano, I. F. Garcia, and G. Ma, "Pattern classification of dermoscopy images: A perceptually uniform model," *Pattern Recognit.*, vol. 46, no. 1, pp. 86–97, 2013.

[49] A. Saez, C. Serrano, and B. Acha, "Model-based classification methods of global patterns in dermoscopic images," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1137–1147, May 2014.

[50] M. E. Celebi and A. Zornberg, "Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification," *IEEE Syst. J.*, vol. 8, no. 3, pp. 980–984, Sep. 2014.

[51] C. Barata, M. E. Celebi, J. S. Marques, and J. Rozeira, "Clinically inspired analysis of dermoscopy images using a generative model," *Comput. Vis. Image Understanding*, vol. 151, pp. 124–137, 2016.

[52] C. Barata, M. E. Celebi, and J. S. Marques, "Development of a clinically oriented system for melanoma diagnosis," *Pattern Recognit.*, vol. 69, pp. 270–285, 2017.

[53] A. Saez, B. Acha, A. Serrano, and C. Serrano, "Statistical detection of colors in dermoscopic images with a texton-based estimation of probabilities," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2018.2823499.

[54] G. Argenziano *et al.*, "Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet," *J. Amer. Acad. Dermatol.*, vol. 48, no. 5, pp. 679–693, 2003.

[55] K. Shimizu, H. Iyatomi, M. E. Celebi, K. A. Norton, and M. Tanaka, "Four-class classification of skin lesions with task decomposition strategy," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 274–283, Jan. 2015.

[56] H. Iyatomi *et al.*, "Quantitative assessment of tumour extraction from dermoscopy images and evaluation of computer-based extraction methods for an automatic melanoma diagnostic system," *Melanoma Res.*, vol. 16, no. 2, pp. 183–190, 2006.

[57] A. Sboner *et al.*, "A multiple classifier system for early melanoma diagnosis," *Artif. Intell. Med.*, vol. 27, no. 1, pp. 29–44, 2003.

[58] G. Schaefer, B. Krawczyk, M. E. Celebi, and H. Iyatomi, "An ensemble classification approach for melanoma diagnosis," *Memetic Comput.*, vol. 6, no. 4, pp. 233–240, 2014.

[59] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.

[60] B. Goodman and S. Flaxman, "European Union regulations on algorithmic decision-making and a "right to explanation"," *AI Mag.*, vol. 38, no. 3, pp. 50–57, 2017.

[61] R. Goebel *et al.*, "Explainable AI: The New 42?" in *Proc. 2nd IFIP TC 5, TC 8/WG 8.4, 8.9, TC 12/WG 12.9 Int. Cross-Domain Conf.*, 2018, pp. 295–303.

[62] H. Iyatomi, M. E. Celebi, G. Schaefer, and M. Tanaka, "Automated color calibration method for dermoscopy images," *Comput. Med. Imag. Graph.*, vol. 35, no. 2, pp. 89–98, 2011.

[63] G. Schaefer, M. I. Rajab, M. E. Celebi, and H. Iyatomi, "Colour and contrast enhancement for improved skin lesion segmentation," *Comput. Med. Imag. Graph.*, vol. 35, no. 2, pp. 99–104, 2011.

[64] Q. Abbas, I. F. Garcia, M. E. Celebi, W. Ahmad, and Q. Mushtaq, "A perceptually oriented method for contrast enhancement and segmentation of dermoscopy images," *Skin Res. Technol.*, vol. 19, no. 1, pp. e490–e497, 2013.

[65] C. Barata, M. E. Celebi, and J. S. Marques, "Improving dermoscopy image classification using color constancy," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 1146–1152, May 2015.

[66] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation for skin lesion analysis," in *Proc. Int. Workshop Skin Image Anal.*, 2018, pp. 303–311.

[67] Y. V. Haeghen, J. M. A. D. Naeyaert, I. Lemahieu, and W. Philips, "An imaging system with calibrated color image acquisition for use in dermatology," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 722–730, Jul. 2000.

[68] C. Grana, G. Pellacani, and S. Seidanari, "Practical color calibration for dermoscopy applied to a digital epiluminescence microscope," *Skin Res. Technol.*, vol. 11, no. 4, pp. 242–247, 2005.

[69] P. Wighton, T. K. Lee, H. Lui, D. McLean, and M. S. Atkins, "Chromatic aberration correction: An enhancement to the calibration of low-cost digital dermoscopes," *Skin Res. Technol.*, vol. 17, no. 3, pp. 339–347, 2011.

[70] J. Quintana, R. Garcia, and L. Neumann, "A novel method for color correction in epiluminescence microscopy," *Comput. Med. Imag. Graph.*, vol. 35, no. 7/8, pp. 646–652, 2011.

[71] A. Badano *et al.*, "Consistency and standardization of color in medical imaging: A consensus report," *J. Digit. Imag.*, vol. 28, no. 1, pp. 41–52, 2015.

[72] T. Mendonca, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "PH^2—A dermoscopic image database for research and benchmarking," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 5437–5440.

[73] T. F. Mendonca *et al.*, "PH^2—A public database for the analysis of dermoscopic images," in *Dermoscopy Image Analysis*, M. E. Celebi, T. Mendonca, and J. S. Marques, Eds. Boca Raton, FL, USA: CRC Press, 2015.

[74] D. Gutman *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," 2016. [Online]. Available: http://arxiv.org/abs/1605.01397

[75] N. C. F. Codella, M. E. Celebi, K. Dana, and A. Halpern, *Proc. 3rd Int. Skin Imag. Collab. Workshop*, 2018.

[76] Y. Smith, A. Weinberg, S. Klauss, D. Soffer, and A. Ingber, "Improving screening for melanoma by measuring similarity to pre-classified images," *Melanoma Res.*, vol. 10, no. 3, pp. 265–272, 2000.

[77] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions," *Int. J. Med. Inform.*, vol. 73, no. 1, pp. 1–23, 2004.

[78] C. B. Akgul, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar, "Content-based image retrieval in radiology: Current status and future directions," *J. Digit. Imag.*, vol. 24, no. 2, pp. 208–222, 2011.

[79] A. Baldi *et al.*, "Definition of an automated Content-Based Image Retrieval (CBIR) system for the comparison of dermoscopic images of pigmented skin lesions," *BioMed. Eng. OnLine*, vol. 8, 2009, Art. no. 18.

[80] M. M. Rahman and P. Bhattacharya, "An integrated and interactive decision support system for automated melanoma recognition of dermoscopic images," *Comput. Med. Imag. Graph.*, vol. 34, no. 6, pp. 479–486, 2010.

[81] N. C. F. Codella, C. C. Lin, A. Halpern, M. Hind, R. Feris, and J. R. Smith, "Collaborative Human-AI (CHAI): Evidence-based interpretable melanoma classification in dermoscopic images," in *Proc. 1st Int. Workshop Interpretability Mach. Intell. Med. Image Comput.*, 2018, pp. 97–105.

[82] M. Sadeghi, P. K. Chilana, and M. S. Atkins, "How users perceive content-based image retrieval for identifying skin images," in *Proc. 1st Int. Workshop Interpretability Mach. Intell. Med. Image Comput.*, 2018, pp. 141–148.

[83] P. Tschandl, G. Argenziano, M. Razmara, and J. Yap, "Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features," *Brit. J. Dermatol.*, to be published.

[84] W. Slue, A. W. Kopf, and J. K. Rivers, "Total-Body photographs of dysplastic nevi," *Arch. Dermatol.*, vol. 124, no. 8, pp. 1239–1243, 1988.

[85] A. C. Halpern, "Total body skin imaging as an aid to melanoma detection," *Seminars Cutaneous Med. Surgery*, vol. 22, no. 1, pp. 2–8, 2003.

[86] J. Malvehy and S. Puig, "Follow-up of melanocytic skin lesions with digital total-body photography and digital dermoscopy: A two-step method," *Clin. Dermatol.*, vol. 20, no. 3, pp. 297–304, 2002.

[87] T. K. Lee, M. S. Atkins, M. A. King, S. Lau, and D. I. McLean, "Counting moles automatically from back images," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 11, pp. 1966–1969, Nov. 2005.

[88] K. Korotkov, J. Quintana, S. Puig, J. Malvehy, and R. Garcia, "A new total body scanning system for automatic change detection in multiple pigmented skin lesions," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 317–338, Jan. 2015.

[89] H. Mirzaalian, T. K. Lee, and G. Hamarneh, "Skin lesion tracking using structured graphical models," *Med. Image Anal.*, vol. 27, pp. 84–92, 2016.

[90] K. Korotkov *et al.*, "An improved skin lesion matching scheme in total body photography," *IEEE J. Biomed. Health Inform.*, to be published.