UC Irvine

UC Irvine Previously Published Works

Title

A Clinical Perspective on the Automated Analysis of Reflectance Confocal Microscopy in Dermatology.

Permalink

<https://escholarship.org/uc/item/8tg6t94s>

Journal

Lasers in surgery and medicine, 53(8)

ISSN

0196-8092

Authors

Mehrabi, Joseph N
Baugh, Erica G
Fast, Alexander
[et al.](#)

Publication Date

2021-10-01

DOI

10.1002/lsm.23376

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

A Clinical Perspective on the Automated Analysis of Reflectance Confocal Microscopy in Dermatology

Joseph N. Mehrabi,¹ Erica G. Baugh,¹ Alexander Fast,² Griffin Lentsch,² Mihaela Balu, PhD,² Bonnie A. Lee, MD,¹ and Kristen M. Kelly, MD^{1,2*}

¹Department of Dermatology, University of California, Irvine, California 92697

²Beckman Laser Institute, University of California Irvine, Irvine, California 92612

Background and Objectives: Non-invasive optical imaging has the potential to provide a diagnosis without the need for biopsy. One such technology is reflectance confocal microscopy (RCM), which uses low power, near-infrared laser light to enable real-time *in vivo* visualization of superficial human skin from the epidermis down to the papillary dermis. Although RCM has great potential as a diagnostic tool, there is a need for the development of reliable image analysis programs, as acquired grayscale images can be difficult and time-consuming to visually assess. The purpose of this review is to provide a clinical perspective on the current state of artificial intelligence (AI) for the analysis and diagnostic utility of RCM imaging.

Study Design/Materials and Methods: A systematic PubMed search was conducted with additional relevant literature obtained from reference lists.

Results: Algorithms used for skin stratification, classification of pigmented lesions, and the quantification of photoaging were reviewed. Image segmentation, statistical methods, and machine learning techniques are among the most common methods used to analyze RCM image stacks. The poor visual contrast within RCM images and difficulty navigating image stacks were mediated by machine learning algorithms, which allowed the identification of specific skin layers.

Conclusions: AI analysis of RCM images has the potential to increase the clinical utility of this emerging technology. A number of different techniques have been utilized but further refinements are necessary to allow consistent accurate assessments for diagnosis. The automated detection of skin cancers requires more development, but future applications are truly boundless, and it is compelling to envision the role that AI will have in the practice of dermatology. Lasers Surg. Med. © 2020 Wiley Periodicals LLC

Key words: reflectance confocal microscopy; artificial intelligence; machine learning; photo-aging; skin stratification; pigmented lesions; melanocytic lesions

INTRODUCTION

Artificial intelligence (AI) is utilized by physicians to acquire, analyze, and apply large amounts of data to aid

in diagnosis and to process images in various fields of medicine, including dermatology [1–3]. Much of the current attention on AI has focused on machine learning, deep neural networks, and computer vision, each with different algorithms to achieve human-level processing of complex imaging data [4]. Augmented intelligence (AuI) is a relatively newer term that refers to the integration of clinical medicine with AI tools and systems to enhance medical care [5]. The American Academy of Dermatology recently released a position statement on AuI that outlines quality assurance measures and laws to ensure adequate professional, legal, and accountable use [5]. However as noted by Fried et al. [6], there is an alarming lack of oversight on AI in the area of mobile application regulation for skin cancer screening, education, mole mapping, and diagnosis.

A number of technologies have been developed to aid in the diagnosis and management of skin conditions. Basic dermoscopy and digital photography were among the first to be developed and are now standard for the diagnosis [7,8] and follow-up of skin conditions [9,10]. Lesions can be screened, but invasive biopsies are often required to provide a definitive diagnosis, especially in the case of pigmented lesions. Non-invasive optical imaging has the potential to provide a diagnosis without the need for biopsy. One such technology is reflectance confocal microscopy (RCM) which uses low power, near-infrared laser light to enable real-time, *in vivo* visualization of superficial human skin from the epidermis down to the papillary dermis. In this technique, laser light is focused and raster-scanned at different depths within the tissue to create grayscale images. These images are created from the back-scattering reflectance of light from tissue constituents with different refractive indices. The reflected light is detected through a confocal pinhole located on the front of the detector to attain a lateral resolution of 1 μm

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

*Correspondence to: Kristen M. Kelly, MD, Department of Dermatology, University of California, 118 Medical Surg I, Irvine, CA 92697. E-mail: kmkelly@hs.uci.edu

Accepted 25 December 2020

Published online 00 Month 2021 in Wiley Online Library (wileyonlinelibrary.com).

DOI 10.1002/lsm.23376

Fig. 1. A reflectance confocal microscopy slice of the skin. Note the highly reflective cells near the dermal papillae representing melanin in the epidermis. The depth of this slice is about 75 μm below the skin surface.

[11], which is comparable to histological resolution. RCM images are captured as a sequence of images at 1–5 μm separation in depth from the epidermis to the dermis. This set of images is known as a *stack*, with each image in the stack known as a *slice* [12]. RCM allows visualization of skin *in vivo* to a depth of approximately 200 μm . A typical RCM image of human skin is depicted and labeled in Figure 1. RCM has been demonstrated to have a sensitivity and specificity above 70% for detecting melanocytic and non-melanocytic skin cancers [13–17] and has also been used for diagnosis and evaluation of pathogenesis for psoriasis, acne, seborrheic keratosis, and other skin conditions [17–21].

The combination of dermoscopy and RCM has improved patient care by identifying high-risk lesions and minimizing biopsies of benign lesions [22,23]. There is a commercially available RCM device along with dedicated billing codes. However, while RCM has great potential as a diagnostic tool, there is still a need for RCM image analysis programs, as acquired grayscale images can be difficult and very time-consuming to visually assess due to poor visual contrast within the RCM images themselves and difficulty navigating between image stacks. A number of algorithms have been used such as image segmentation, statistical methods, and machine learning techniques, and are defined in Table 1.

The clinical applications for these algorithms are vast. For example, accurate localization of identified changes is an important aspect of diagnosis. For example, skin cancer diagnosis can be aided by identification of tumor aggregates in the dermis for basal cell carcinoma and identification of melanocytes or their dendrites in the

upper layers of the epidermis for melanoma. Correctly identifying the malignant potential of pigmented lesions is an important goal of histopathologic diagnosis and should be an important application of light-based imaging including RCM. Diagnosis of melanoma is complex and relies on evaluation of multiple different lesional characteristics including symmetry (or lack thereof), circumscription, and upward scattering of melanocytes above the DEJ.

Accurate identification of the DEJ is also clinically valuable. As with identifying specific skin layers, localizing the DEJ within grayscale images can be challenging and time-consuming due to poor contrast. This could be mitigated with the use of automated algorithms. The ability to accurately identify changes at the DEJ can have broad utility. For example, inflammatory cells present in and above the DEJ could help identify lichenoid inflammatory conditions as well as malignancies such as mycosis fungoides. Additionally, identification of a sub-epidermal split at the level of the DEJ versus intra-epidermal separation of the keratinocytes can aid in distinguishing between various blistering diseases. Mandel et al. conducted an observational, retrospective study to assess the accuracy of RCM and optical coherence tomography (OCT) in diagnosing bullous pemphigoid and pemphigus vulgaris, and found improved diagnostic accuracy when bedside diagnostic RCM was combined with OCT [39]. The authors noted similar limitations in interpreting RCM images as previously discussed, adding that inflammatory phenomena and laser power attenuation can render the blister difficult to distinguish [39].

Assessment of the epidermal thickness and characterization of the stratum corneum in terms of thickening (hyperkeratosis) or retained nuclei (parakeratosis) can aid in the diagnosis of inflammatory conditions such as psoriasis and eczema or in neoplasms such as actinic keratoses. Differences in specific epidermal strata pre and post-therapy can also determine the impact of skin treatments. RCM allows visualization of the above features of inflammation as well as hypogranulosis, papillomatosis, dilated blood vessels surrounded by inflammatory cells, spongiosis, and melanophages [40].

Finally, automatic evaluation of the dermis through analysis of the RCM images facilitates the non-invasive assessment of photoaging. On RCM images, epidermal keratinocytes are visualized as a honeycomb-like arrangement and rete ridges can also be identified [41,42].

In this review, we examine the different approaches to automating the quantitative analysis of RCM images and their clinical applications. The intent is to define the current state of clinical RCM image analysis and to consider opportunities for future research.

METHODS

A systematic PubMed search was conducted to identify relevant clinical articles addressing AI and RCM. Relevant keywords and Message-Subject-Headings (MeSH) were applied and following search strategy was created: confocal

TABLE 1. AI Terms and Definitions

Term	Definition
<i>Image segmentation and pattern recognition</i>	
Texton	Texture representations of characteristic structural features found in a two-dimensional image [24–27]
Texton-based algorithm	Algorithm that relies on the generation of a library of textons based on the image. This library permits the classification and separation of images via texture, which differed between layers of the skin [24–27]
Bag-of-features approach	Method that represents images as random collections of local features, which may consist of textons in the context of skin imaging [28]
K-means clustering	Algorithm commonly used to automatically partition data into a set of k groups [29]
Wavelet transformation, Fourier transformation, and other transformation analyses	All decompose signal into a set of basic functions like contractions, expansions, and translations of a parent function known as a wavelet [30]
Markov chain	Algorithm which uses an independent random search for each step
Monte Carlo	Algorithm which runs millions of paths to determine the average value from random trials
Image segmentation	The process of simplifying the representation of an image into something more meaningful and easier to analyze by partitioning it into smaller components, such as pixels [31]
Tree-of-shapes	Mathematical method of merging multiple image functions into one [32]
<i>Statistical approaches</i>	
Regression analysis	Estimates relationships between independent and dependent variables
Logistic regression	Measures the relationship between one or more independent variables and a categorical dependent variable by estimating the probabilities using a logistic function, for example, the relative appearance of textural features in epidermal strata
Poisson point process	A mathematical model in Bayesian probability statistic to explain seemingly random processes; in this context it was used to explain the undulating pattern of the dermal-epidermal junction (DEJ)
<i>Computer vision and machine learning</i>	
Conditional random fields	Class of statistical modeling methods often applied to pattern recognition and machine learning [33]
Classification and regression trees (CART)	Non-parametric decision tree learning technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively [34]
Random forest classifier	Learning method that involves the composition of multiple decision trees, which are predictive models consisting of a series of Boolean tests to eventually lead to outcomes [35]
Neural network	Set of algorithms designed to recognize patterns [36] modeled after the organization of neurons in the brain and the basis for machine learning algorithms
Machine learning	Allows algorithms to learn from data and make unsupervised decisions
Deep learning	Subset of machine learning by which neural networks automatically learn from unlabeled or unstructured data [37]
Support vector machine (SVM)	Supervised learning model with algorithms for data classification and regression [38]

[All Fields] AND (“microscopy”[MeSH Terms] OR “microscopy”[All Fields]) AND (“skin”[MeSH Terms] OR “skin”[All Fields]) AND (“automated”[MeSH Terms] OR “automated”[All Fields]). A filter for only human subjects was applied. This search returned 38 articles.

RESULTS

One independent reviewer analyzed the results of the search algorithm. Of the 38 articles, 11 satisfied our criteria of discussing automated RCM image analysis *in vivo*, and an additional 16 articles, including one PhD thesis, were found via reference lists. Three of the additional articles were not included in this review because they studied cervical epithelium and oral mucosa instead of the skin. Ten additional articles were provided by reviewers and outside sources. A total of 34 studies were deemed eligible for our review of computational methods of RCM image analysis.

Articles were grouped based on the application of interest: uninvolved versus lesional skin, skin stratification, DEJ delineation, analysis of pigmented lesions and quantification of photoaging. Table 2 details the method used for each skin stratification algorithm. A summary of the published algorithms and their accuracies, average errors, sensitivity, and specificity are presented below.

Uninvolved Versus Lesional Skin

One important study examined RCM image quality using a machine learning algorithm with the goal of understanding whether automated algorithms could identify uninvolved skin surrounding a lesion [45]. The algorithm was able to detect uninvolved skin with $82 \pm 9.8\%$ sensitivity and $93 \pm 2.4\%$ specificity [45]. The algorithm was also able to detect artifacts that might obscure dermoscopic images such as bubbles, particulates, hair, insufficient illumination, and pixel saturation [45].

Skin Stratification

The majority of the published algorithms described automated methods of stratifying skin layers in RCM images by labeling each of the epidermal strata in groups from the stratum corneum to the basal layer [12,24–26,43,44,46] and delineating the DEJ from the epidermis and dermis

[47–55]. Table 2 summarizes the methods described for epidermal stratification. The *accuracy* of these algorithms referred to the rate of correctly labeled skin slices against ground-truth expert labeling and was found to maximize at 88.18% with implementations of convolutional neural networks by Bozkurt et al. [12,44]. Sensitivity and specificity were maximized with these techniques as well, achieving 87% and 94%, respectively [12,44]. Algorithms stratifying the epidermal layers with texture and feature analysis combined with logistic regression, conditional random fields, or support vector machines (SVMs) [24,25,43] were found to have inferior accuracies, sensitivities, and specificities when compared with neural network-based algorithms [12,26,44]. One algorithm attempting to delineate only the stratum corneum using texture analysis and wavelet transformation performed at 82.89% accuracy compared to ground-truth expert labeling of slices [46].

Delineation of the DEJ

Locating the DEJ boundary is difficult in lighter-skinned individuals due to the relative absence of melanin and reduced contrast of the basal layer of the epidermis from the dermis. Extensive work from Kurugol et al. using a bag of features approach and SVMs provided an average of 60% accuracy for the labeling of epidermal, DEJ, and dermal layers for lighter skin phototypes when compared to manual labeling of the layers by experts with an average error of $8.0 \mu\text{m}$ per slice, less than the size of a single basal cell [47–52]. However, 88% accuracy with an average error of $7.9 \mu\text{m}$ was achieved when analyzing the RCM images of darker skin phototypes. Various statistical and machine learning methods achieved correct identification of the DEJ [53–55]. Thus, a spatial Poisson point process approach identified the DEJ of lighter skin phototypes with a $12.1 \mu\text{m}$ average error and the DEJ of darker skin phototypes with $5.41 \mu\text{m}$ average error when compared to expert labeling [53]. Machine learning-based statistical methods using random forest classifiers and two-dimensional (2D) or three-dimensional (3D) conditional random fields achieved about 54%–86% accuracy, 55%–97% sensitivity, and 89%–98% specificity depending on whether the skin layer being identified was the epidermis, DEJ, or dermis [54,55]. Kose et al. developed and

TABLE 2. Skin Stratification Algorithms

Author	Year	Method	Average accuracy	Correlation	Sensitivity	Specificity
Somoza [43]	2014	Unsupervised texton-based approach	54%	0.84–0.95	0.527	0.9
Hames [24]	2016	Bag of features and logistic regression	85.60%	0.54–0.74	0.839	0.921
Hames [25]	2015	Patch-based bag-of-features approach followed by Support vector machine classifier	85.70%		0.3811	0.902
Kaur [26]	2016	Hybrid deep learning	82%		0.7174	0.962
Bozkurt [12]	2017	Recurrent convolutional neural network	87.97%		0.8711	0.93957
Bozkurt [44]	2017	Recurrent convolutional neural network	88.18%		0.87327	0.94070

trained a recurrent neural network to identify three skin strata (epidermis, DEJ, and dermis) with 88.17% accuracy [56,57].

Analyzing Pigmented Lesions

The applicability of automated RCM image analysis based on machine learning algorithms for diagnostic discrimination of melanocytic lesions has been investigated in several studies [58–60]. An algorithm based on CART analysis applied by Koller et al. to RCM images from melanocytic lesions showed moderate outcome with malignancy classification for 55% of melanomas, but also for 47% of the benign melanocytic nevi [58]. The authors suggested the non-standardized RCM image acquisition as a possible cause for the low performance of the algorithm. Gareau et al. applied a pattern recognition algorithm to identify pagetoid melanocytes and DEJ disruption in RCM images from superficial spreading melanoma. The algorithm was able to identify pagetoid melanocytes in all melanomas and none in the nevi group, while DEJ disruption was significantly more pronounced in melanomas compared to nevi [61]. Although the study was performed on a limited number of patients ($n = 5$ in each group), the results showed promise for developing quantitative metrics for melanoma diagnosis based on RCM images. For lentigo maligna, RCM images of 135 malignant tumors (115 lentigo maligna or lentigo maligna melanoma and 20 basal cell carcinoma) and 88 benign tumors were analyzed using a series of extraction methods, textural interpretations, and machine learning algorithms [62]. All tested algorithms achieved sensitivities and specificities greater than 75% [62].

More recently, a machine learning algorithm that can imitate the clinician's qualitative and visual analysis of RCM images developed by Kose et al. [59] identified 5–6 different morphological patterns in the DEJ on RCM mosaics of melanocytic lesions. The sensitivity and specificity of this algorithm ranged from 55%–81% and 81%–89%, respectively, using SVMs [59]. An even more recent implementation of a multiresolution convolutional neural network to identify the same morphological patterns averaged 76.8% sensitivity and 94.7% specificity [60]. The aforementioned algorithm by Kose et al. [56,57] used the accurate classification of the skin strata and an additional multiscale convolutional neural network to segment cellular-morphological patterns in 117 RCM image mosaics of melanocytic lesions at the DEJ. This segmentation across six classes achieved a mean sensitivity of 70% and specificity of 94% [56,57].

In addition to distinguishing various patterns of melanocytic neoplasms, another important clinical application for this technology would be distinguishing melanoma from a non-melanocytic pigmented lesion such as a solar lentigo. Multiple studies have attempted to classify solar lentigines using automated algorithms of RCM mosaics, based on observed alterations in the cellular network at the DEJ that result in irregularly shaped papillae surrounded by bright borders as compared with healthy controls. Halimi et al. [63] developed a hierarchical

Bayesian model to quantify image reflectivity with Markov chain Monte Carlo algorithm to classify images based on their reflectivity. On 45 patients, classification of solar lentigines had an accuracy of 97.7%, sensitivity of 96.2%, and specificity of 100%. On the same number of patients, the same group proposed a multiresolution wavelet decomposition method at each skin depth based on observed variations in papillae height and shape, then SVM classifiers analyzed a Gaussian distribution of image pixels at each depth allowing for real-time detection of lentigo [64]. Sensitivity and specificity were 81.4% and 83.3% respectively at depths from 50 to 60 μm , approximately the depth of the DEJ [64]. Another group used a convolutional neural network to classify solar lentigines and achieved 98% accuracy, 96% sensitivity, and 100% specificity [65].

Quantifying Photo-Aging

RCM has the ability to capture the flattening of the DEJ and deviations in the size and shape of epidermal keratinocytes from their traditional honeycomb pattern, changes consistent with skin aging [42]. These characteristics have been quantified with image segmentation and transformation analysis [66,67] and machine learning approaches [41,68]. Calculations from image segmentation and 2D Fourier transformation analysis by Raphael et al. [67] correlated positively with clinician ratings of photo-aging, analysis of clinical photographs, and expert assessment of RCM images. Building from a promising algorithm from Gareau [66] that identifies epidermal keratinocytes via RCM, a machine learning algorithm involving image segmentation and a random forest classifier found a positive correlation between younger patients and the number of regular keratinocytes, exhibiting minimal variability in size and shape and a consistent honeycomb pattern [41,68]. Older patients were found to have higher numbers of irregular keratinocytes and sun-exposed areas displayed a 22% increase in irregularity of the honeycomb pattern as identified by the random forest classifier [41,68].

Using a random forest classifier, another study developed a scoring system to qualify the degree of photo-aging based upon the degree of irregularity in the honeycomb pattern of the epidermis, the predominant shapes of the dermal papillae as being ringed or polycyclic, and the predominance of collagen fiber types as either coarse or reticulated [69]. Based on expert assessment, the algorithm correctly identified qualities of the three characteristics with 80%–83% accuracy, 63%–81% sensitivity, and 81%–89% specificity, and the skin age assessment scores correlated with those of the expert evaluators [69].

In an attempt to delineate the DEJ with 3D reconstruction and to quantify photo-aging by calculating the number and volume of dermal papillae and rete ridges, RCM was employed in conjunction with micro-computed X-ray tomography (microCT) [70]. A semi-automatic, near real-time segmentation program created a 3D image of the DEJ and dermal papillae from stacks of RCM images. An algorithm then found a horizontal

median plane at the DEJ and designated the dermis above it as dermal papillae and the epidermis below it as rete ridges. MicroCT was better at discerning age-related differences between young and elderly skin than was RCM, where no significant differences were observed between volumes of dermal papillae and rete ridges [70]. Robic et al. [71] described a tree-of-shapes algorithm that mapped the topography of the DEJ and assessed photoaging in seven patients aged 18–25 and eight patients aged 55–65. The surface area of the topographical representation of the DEJ was flatter among the cheeks of older patients [71].

DISCUSSION

AI has been implemented in various practices in dermatology [72] and the algorithms discussed in this review can serve as clinician support tools which may allow increased utilization of RCM in clinics. AI in RCM image processing has evolved over the past decade from crude texture analysis to enhanced machine learning algorithms to optimize efficiency and clinical utility. This review discussed different applications of published AI algorithms for RCM image analysis, which are summarized in Table 3. Most algorithms in this review relied on image texture analysis and machine learning algorithms.

Poor visual contrast within RCM images and difficulty navigating image stacks were addressed with several skin stratification algorithms which label the different layers of the epidermis [12,24–26,43,44] and provide measurement of the stratum corneum [1]. All demonstrated promising accuracy and correlation with expert labeling of skin strata, but sensitivity and specificity of strata identification were maximized among those implementing machine learning [12,25,26,44]. Additionally, AI algorithms focused on delineating the DEJ in RCM images and detecting mild differences in contrast using machine learning and texture-based classifiers [54,55].

Diagnosis of melanoma is complex; of the four algorithms discussed in this review, only one offered sufficient sensitivity and specificity for the diagnosis of melanoma using a highly sophisticated machine learning algorithm [60]. There is much potential for the future use of AI as an adjunct to clinical analysis of melanoma and suspicious lesions, but current technology remains far from replacing clinician expertise. As Fried et al. [6] noted, in the case where dermatologists and AI disagree, there is no way to identify the point of discrepancy to improve knowledge on either end. Additionally, if machine learning algorithms rely on photo datasets that contain a majority of lighter-skinned examples, there may already be inherent bias due to a lack of lesions in skin of color [73]. Thus, more work will be required on this front.

Algorithms have also been developed to track photoaging by detecting irregularities in the keratinocyte arrangement or DEJ flattening and reliably correlate with patient age, expert scores and ratings of skin age, and histological analysis. AI-directed photo-aging assessment could be clinically useful to determine skin cancer risk and to track improvement with photoaging treatments.

Finally, patients have perceived the implementation of AI support tools as helpful for improving diagnostic accuracy without replacing a true physical examination [74]. A recent report explored patient perspectives on the use of AI support tools to help clinicians and patients in skin cancer screening [74]. In contrast to the clinician support tools, patients viewed direct-to-patient implementations of AI as beneficial for increasing diagnostic speed, patient activation, and health care access, but there was concern that they could increase patient anxiety with the loss of human interaction and discovery of a suspicious skin tumor [74]. Overall, 79% of patients saw utility in AI-powered clinician support tools, a form of AuI [5,74].

This review is not without limitations. The search method relied heavily on reference lists from relevant

TABLE 3. Summary of Algorithms Used for Image Processing and Interpretation for Each Application

Epidermal stratification	Dermal-epidermal junction (DEJ) delineation	Classifying pigmented lesions	Quantification of photoaging
Texton-based algorithm [43]	Texton-based algorithm [47–52]	Computation of DEJ roughness [61]	Transformation analysis [67]
Logistic regression [24]	Conditional random fields [54,55]	Classification and Regression Trees [58,62]	Support vector machine [41,68]
Support vector machine [25]	Spatial Poisson process [53]	Convolutional neural network [60,62,65]	Random Forest Classifier [69]
Hybrid deep learning [26]	Support vector machine [47–52]	Support vector machine [59,62]	3D reconstruction [70]
Convolutional neural network [12,44]	Recurrent neural network [56,57]	Recurrent neural network plus Convolutional neural network [56,57]	Tree-of-shapes [71]
Wavelet transform [46]		Bayesian Model [63] Wavelet decomposition with support vector machine [64]	

articles that analyze automatic image processing of RCM images. Technical details of each algorithm were not analyzed in detail, as the goal of this review was to give clinicians an idea of how RCM image interpretation could be applied to the clinical setting using developed software mechanisms. Additionally, skin of color was only analyzed by a single group for DEJ delineation [47,51,52], and given the distinct contrast in reading and processing RCM images between light and dark skin phototypes, the accuracy of AI algorithms should be tested on images of all Fitzpatrick skin types.

In summary, RCM and other optical imaging technologies have great potential to provide rapid non-invasive skin condition diagnosis. However, analysis of images is a current limitation. AI has begun to address this issue but will need enhancements to guarantee appropriate sensitivity and specificity. Future advancements are truly boundless but could include the combination of all described algorithms into one software system for a broader range clinical application. Real-time skin stratification would greatly assist in image navigation and interpretation. In addition, similar algorithms could be applied to other modalities for non-invasive imaging of human skin, such as multiphoton microscopy [75] and OCT [76].

REFERENCES

- Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl* 2004;86(5):334–338.
- Tang A, Tam R, Cadrin-Chênevert A, et al. Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J* 2018;69(2):120–135.
- Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;172(5):1122–1131.
- Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: Promises and perils. *Ann Surg* 2018;268(1):70–76.
- Kovarik C, Lee I, Ko J, et al. Commentary: Position statement on augmented intelligence (AuI). *J Am Acad Dermatol* 2019;81(4):998–1000.
- Fried L, Tan A, Bajaj S, Liebman TN, Polsky D, Stein JA. Technological advances for the detection of melanoma: Advances in diagnostic techniques. *J Am Acad Dermatol* 2020;83(4):983–992.
- Guitera P, Pellacani G, Longo C, Seidenari S, Avramidis M, Menzies SW. In vivo reflectance confocal microscopy enhances secondary evaluation of melanocytic lesions. *J Invest Dermatol* 2009;129(1):131–138.
- Guitera P, Menzies SW, Longo C, Cesinaro AM, Scolyer RA, Pellacani G. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. *J Invest Dermatol* 2012;132(10):2386–2394.
- Schneider SL, Kohli I, Hamzavi IH, Council ML, Rossi AM, Ozog DM. Emerging imaging technologies in dermatology: Part II: Applications and limitations. *J Am Acad Dermatol* 2019;80(4):1121–1131.
- Schneider SL, Kohli I, Hamzavi I, Council ML, Rossi AM, Ozog DM. Emerging imaging technologies in dermatology: Part I: Basic principles. *J Am Acad Dermatol* 2019;80(4):1114–1120.
- Que SKT, Fraga-Braghiroli N, Grant-Kels JM, Rabinovitz HS, Oliviero M, Scope A. Through the looking glass: basics and principles of reflectance confocal microscopy. *J Am Acad Dermatol* 2015;73(2):276–284.
- Bozkurt A, Gale T, Kose K, et al. Delineation of skin strata in reflectance confocal microscopy images with recurrent convolutional networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017: 777–785. 10.1109/CVPRW.2017.108.
- Nori S, Rius-Díaz F, Cuevas J, et al. Sensitivity and specificity of reflectance-mode confocal microscopy for in vivo diagnosis of basal cell carcinoma: A multicenter study. *J Am Acad Dermatol* 2004;51(6):923–930.
- Astner S, Gonzalez E, Cheung A, Rius-Díaz F, Gonzalez S. Pilot study on the sensitivity and specificity of in vivo reflectance confocal microscopy in the diagnosis of allergic contact dermatitis. *J Am Acad Dermatol* 2005;53(6):986–992.
- Pellacani G, Guitera P, Longo C, Avramidis M, Seidenari S, Menzies S. The impact of in vivo reflectance confocal microscopy for the diagnostic accuracy of melanoma and equivocal melanocytic lesions. *J Invest Dermatol* 2007;127(12):2759–2765.
- Gerger A, Hofmann-Wellenhof R, Langsenlehner U, et al. In vivo confocal laser scanning microscopy of melanocytic skin tumours: Diagnostic applicability using unselected tumour images. *Br J Dermatol* 2008;158(2):329–333.
- Rajadhyaksha M, Marghoob A, Rossi A, Halpern AC, Nehal KS. Reflectance confocal microscopy of skin in vivo: From bench to bedside. *Lasers Surg Med* 2017;49(1):7–19.
- Manfredini M, Greco M, Farnetani F, et al. In vivo monitoring of topical therapy for acne with reflectance confocal microscopy. *Skin Res Technol* 2017;23(1):36–40.
- Lacarrubba F, Ardigo M, Di Stefani A, Verzi AE, Micali G. Dermatoscopy and reflectance confocal microscopy correlations in nonmelanocytic disorders. *Dermatol Clin* 2018;36(4):487–501.
- Agozzino M, Gonzalez S, Ardigo M. Reflectance confocal microscopy for inflammatory skin diseases. *Actas Dermosifiliográficas* 2016;107(8):631–639.
- Guo A, Chen J, Yang C, Ding Y, Zeng Q, Tan L. The challenge of diagnosing seborrheic keratosis by reflectance confocal microscopy. *Skin Res Technol* 2018;24(4):663–666.
- Borsari S, Pampena R, Lallas A, et al. Clinical indications for use of reflectance confocal microscopy for skin cancer diagnosis. *JAMA Dermatol* 2016;152(10):1093–1098.
- Pellacani G, Witkowski A, Cesinaro AM. Cost-benefit of reflectance confocal microscopy in the diagnostic performance of melanoma. *J Eur Acad Dermatol Venereol* 2016;30(3):413–419.
- Hames SC, Ardigo M, Soyer HP, Bradley AP, Prow TW. Automated segmentation of skin strata in reflectance confocal microscopy depth stacks. *PLoS One* 2016;11(4):e0153208.
- Hames SC, Ardigo M, Soyer HP, Bradley AP, Prow TW. Anatomical skin segmentation in reflectance confocal microscopy with weak labels. In 2015 International Conference on Digital Image Computing: Techniques and Applications, DICTA 2015 [7371231] Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/DICTA.2015.7371231>
- Kaur P, Dana KJ, Cula GO, Mack MC Hybrid deep learning for reflectance confocal microscopy skin images. In 2016 23rd International Conference on Pattern Recognition (ICPR) (pp. 1466–1471). IEEE.
- Cula OG, Dana KJ, Murphy FP, Rao BK. Skin texture modeling. *Int J Comput Vis* 2005;62(1–2):97–119.
- O'Hara S, Draper BA. Introduction to the bag of features paradigm for image classification and retrieval. *arXiv* 2011. 11013354. <https://arxiv.org/abs/1101.3354>
- MacQueen J. Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 281–297, Berkeley, Calif: University of California Press, 1967. <https://projecteuclid.org/euclid.bsm/1200512992>
- Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans Inf Theor* 1990;36(5):961–1005.
- Shapiro LG, Stockman GC. *Computer Vision*. Upper Saddle River, NJ: Prentice Hall; 2001.
- Ballester C, Caselles V, Monasse P. The tree of shapes of an image. *ESAIM Contr Optim Ca* 2003;9:1–18.

33. Lafferty J, McCallum A, Pereira FC. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: ICML '01: Proceedings of the Eighteenth International Conference on Machine Learning. 2001: 282–289.
34. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression Trees. Boca Raton, FL: CRC Press; 1984. <https://doi.org/10.1201/9781315139470>
35. Ho TK. Random decision forests. Proceedings of 3rd International Conference on Document Analysis and Recognition, Canada: Montreal, Quebec, 1995;1: 278–282. 10.1109/ICDAR.1995.598994
36. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci* 1982;79(8):2554–2558.
37. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv 2014. 1412.6980. <https://arxiv.org/abs/1412.6980v9>
38. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995;20(3):273–297.
39. Mandel VD, Cinotti E, Benati E, et al. Reflectance confocal microscopy and optical coherence tomography for the diagnosis of bullous pemphigoid and pemphigus and surrounding subclinical lesions. *J Eur Acad Dermatol Venereol* 2018;32(9):1562–1569.
40. Ilie MA, Caruntu C, Lixandru D, et al. In vivo confocal laser scanning microscopy imaging of skin inflammation: Clinical applications and research directions. *Exp Ther Med* 2019;17(2):1004–1011.
41. Robic J, Nkengne A, Perret B, Couprie M, Talbot H. Automated quantification of the epidermal aging process using in-vivo confocal microscopy. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, 2016, pp. 1221–1224, doi:10.1109/ISBI.2016.7493486
42. Wurm E, Longo C, Curchin C, Soyer H, Prow T, Pellacani G. In vivo assessment of chronological ageing and photoageing in forearm skin using reflectance confocal microscopy. *Br J Dermatol* 2012;167(2):270–279.
43. Somoza E, Cula GO, Correa C, Hirsch JB. Automatic localization of skin layers in reflectance confocal microscopy. Paper presented at: International Conference Image Analysis and Recognition 2014.
44. Bozkurt A, Kose K, Coll-Font J, et al. Delineation of skin strata in reflectance confocal microscopy images using recurrent convolutional networks with Toeplitz attention. arXiv 2017. 1712.00192. https://link.springer.com/chapter/10.1007/978-3-319-11755-3_16
45. Kose K, Bozkurt A, Alessi-Fox C, et al. Utilizing machine learning for image quality assessment for reflectance confocal microscopy. *J Invest Dermatol* 2020;140(6):1214–1222.
46. Bozkurt A, Kose K, Alessi-Fox C, Dy JG, Brooks DH, Rajadhyaksha M. Unsupervised delineation of stratum corneum using reflectance confocal microscopy and spectral clustering. *Skin Res Technol* 2017;23(2):176–185.
47. Kurugol S, Kose K, Park B, Dy JG, Brooks DH, Rajadhyaksha M. Automated delineation of dermal-epidermal junction in reflectance confocal microscopy image stacks of human skin. *J Invest Dermatol* 2015;135(3):710–717.
48. Kurugol S, Rajadhyaksha M, Dy JG, Brooks DH. Validation study of automated dermal/epidermal junction localization algorithm in reflectance confocal microscopy images of skin. Proceedings of SPIE—the International Society for Optical Engineering, 2012;8207. doi:10.1117/12.909227
49. Kurugol S, Dy JG, Rajadhyaksha M, Gossage KW, Weissman J, Brooks DH. Semi-automated algorithm for localization of dermal/epidermal junction in reflectance confocal microscopy images of human skin. *Proc SPIE Int Soc Opt Eng* 2011;7904:7901a.
50. Kurugol S, Dy JG, Brooks DH, Rajadhyaksha M. Pilot study of semiautomated localization of the dermal/epidermal junction in reflectance confocal microscopy images of skin. *J Biomed Opt* 2011;16(3):036005.
51. Kurugol S, Dy JG, Rajadhyaksha M, Brooks DH. Localizing the dermis/epidermis boundary in reflectance confocal microscopy images with a hybrid classification algorithm. 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009: 1322–1325.
52. Kurugol S, Dy J, Rajadhyaksha M, Brooks DH. Detection of the dermis/epidermis boundary in reflectance confocal images using multi-scale classifier with adaptive texture features. 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, 2008, 492–495. <https://doi.org/10.1109/ISBI.2008.4541040>
53. Ghanta S, Jordan MI, Kose K, Brooks DH, Rajadhyaksha M, Dy JG. A marked poisson process driven latent shape model for 3D segmentation of reflectance confocal microscopy image stacks of human skin. *IEEE Trans Image Process* 2017;26(1):172–184.
54. Robic J, Perret B, Nkengne A, Couprie M, Talbot H. Three-dimensional conditional random field for the dermal-epidermal junction segmentation. *J Med Imaging (Bellingham)* 2019;6(2):024003.
55. Robic J, Perret B, Nkengne A, Couprie M, Talbot H. Classification of the dermal-epidermal junction using in-vivo confocal microscopy. Paper presented at: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) 2017.
56. Kose K, Bozkurt A, Fox CA, et al. Detection of the DEJ and Segmentation of Its Morphological Patterns in RCM Images of Melanocytic Skin Lesions. In *Biophotonics Congress: Biomedical Optics 2020 (Translational, Microscopy, OCT, OTS, BRAIN)*, OSA Technical Digest (Optical Society of America, 2020), paper MW2A.1. <https://doi.org/10.1364/MICROSCOPY.2020.MW2A.1>
57. Kose K, Bozkurt A, Alessi-Fox C, et al. Segmentation of cellular patterns in confocal images of melanocytic lesions in vivo via a multiscale encoder-decoder network (MED-Net). arXiv 2020. 2001.01005.
58. Koller S, Wiltgen M, Ahlgrimm-Siess V, et al. In vivo reflectance confocal microscopy: Automated diagnostic image analysis of melanocytic skin tumours. *J Eur Acad Dermatol Venereol* 2011;25(5):554–558.
59. Kose K, Alessi-Fox C, Gill M, Dy JG, Brooks DH, Rajadhyaksha M. A machine learning method for identifying morphological patterns in reflectance confocal microscopy mosaics of melanocytic skin lesions in-vivo. Paper presented at: Photonic Therapeutics and Diagnostics XII 2016. <https://www.scopus.com/inward/record.uri?eid=s2.0-84973334714&partnerID=40&md5=439aebcd2e540c99b6a301a1bd121bf>
60. Bozkurt A, Kose K, Alessi-Fox C, et al. A multiresolution convolutional neural network with partial label training for annotating reflectance confocal microscopy images of skin. Paper presented at: International Conference on Medical Image Computing and Computer-Assisted Intervention 2018.
61. Gareau DS, Hennessey RJ, Wan E, Pellacani G, Jacques SL. Automated detection of malignant features in confocal microscopy on superficial spreading melanoma versus nevi. *J Biomed Opt* 2010;15(6):061713.
62. Cendre R, Mansouri A, Perrot J-L, Cinotti E, Marzani F. Classification of Lentigo Maligna at patient-level by means of reflectance confocal microscopy data. *Appl Sci* 2020;10(8):2830.
63. Halimi A, Batatia H, Le Digabel J, Josse G, Tourneret J-Y. An unsupervised Bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images. Paper presented at: 2017 25th European Signal Processing Conference (EUSIPCO) 2017, pp. 241–245, doi:10.23919/EUSIPCO.2017.8081205
64. Halimi A, Batatia H, Le Digabel J, Josse G, Tourneret JY. Wavelet-based statistical classification of skin images acquired with reflectance confocal microscopy. *Biomed Opt Express* 2017;8(12):5450–5467.
65. Zorgui S, Chaabene S, Bouaziz B, Batatia H, Chaari L. A convolutional neural network for lentigo diagnosis. Paper presented at: International Conference on Smart Homes and Health Telematics 2020.
66. Gareau D. Automated identification of epidermal keratinocytes in reflectance confocal microscopy. *J Biomed Opt* 2011;16(3):030502.
67. Raphael AP, Kelf TA, Wurm EM, Zvyagin AV, Soyer HP, Prow TW. Computational characterization of reflectance confocal microscopy features reveals potential for automated photoageing assessment. *Exp Dermatol* 2013;22(7):458–463.

68. Robic J. Automated Characterization of Skin Aging Using In Vivo Confocal Microscopy. Lyon, France: HAL open archive; 2018. <https://tel.archives-ouvertes.fr/tel-01935610>
69. Robic J, Nkengne A, Perret B, Couprie M, Talbot H, Pellacani G, Vie K. Clinical validation of a computer-based approach for the quantification of the skin ageing process of women using in vivo confocal microscopy. *J Eur Acad Dermatol Venereol* 2020. jdv.16810.
70. Newton VL, Bradley RS, Seroul P, et al. Novel approaches to characterize age-related remodelling of the dermal-epidermal junction in 2D, 3D and in vivo. *Skin Res Technol* 2017;23(2):131–148.
71. Robic J, Perret B, Nkengne A, Couprie M, Talbot H. Self-dual pattern spectra for characterising the dermal-epidermal junction in 3D reflectance confocal microscopy imaging. Paper presented at: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing 2019. Germany: Saarbrücken, pp. 508–519. <https://hal.archives-ouvertes.fr/hal-02169702>
72. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. *J Dermatol Treat* 2020;31(5):496–510.
73. Adamson AS, Smith A. Machine learning and health care disparities in dermatology. *JAMA Dermatol* 2018;154(11):1247–1248.
74. Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: A qualitative study. *JAMA Dermatol* 2020;156(5):501–512.
75. Balu M, Kelly KM, Zachary CB, et al. Distinguishing between benign and malignant melanocytic nevi by in vivo multiphoton microscopy. *Cancer Res* 2014;74(10):2688–2697.
76. Olsen J, Holmes J, Jemec GB. Advances in optical coherence tomography in dermatology—A review. *J Biomed Opt* 2018;23(4):1–10.