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, Erica G. Baugh, Alexander Fast, Griffin Lentsch, Mihaela Balu, PhD, Bonnie A. Lee, MD, and Kristen M. Kelly, MD

1Department of Dermatology, University of California, Irvine, California 92697
2Beckman Laser Institute, University of California Irvine, Irvine, California 92697

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
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), circumscript, 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 subepidermal split at the level of the DEJ versus intraepidermal 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 characteristics 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>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image segmentation and pattern recognition</strong></td>
<td></td>
</tr>
<tr>
<td>Texton</td>
<td>Texture representations of characteristic structural features found in a two-dimensional image [24–27]</td>
</tr>
<tr>
<td>Texton-based algorithm</td>
<td>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]</td>
</tr>
<tr>
<td>Bag-of-features approach</td>
<td>Method that represents images as random collections of local features, which may consist of textons in the context of skin imaging [28]</td>
</tr>
<tr>
<td>K-means clustering</td>
<td>Algorithm commonly used to automatically partition data into a set of ( k ) groups [29]</td>
</tr>
<tr>
<td>Wavelet transformation, Fourier transformation, and other transformation analyses</td>
<td>All decompose signal into a set of basic functions like contractions, expansions, and translations of a parent function known as a wavelet [30]</td>
</tr>
<tr>
<td>Markov chain</td>
<td>Algorithm which uses an independent random search for each step</td>
</tr>
<tr>
<td>Monte Carlo</td>
<td>Algorithm which runs millions of paths to determine the average value from random trials</td>
</tr>
<tr>
<td>Image segmentation</td>
<td>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]</td>
</tr>
<tr>
<td>Tree-of-shapes</td>
<td>Mathematical method of merging multiple image functions into one [32]</td>
</tr>
<tr>
<td><strong>Statistical approaches</strong></td>
<td></td>
</tr>
<tr>
<td>Regression analysis</td>
<td>Estimates relationships between independent and dependent variables</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>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</td>
</tr>
<tr>
<td>Poisson point process</td>
<td>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)</td>
</tr>
<tr>
<td><strong>Computer vision and machine learning</strong></td>
<td></td>
</tr>
<tr>
<td>Conditional random fields</td>
<td>Class of statistical modeling methods often applied to pattern recognition and machine learning [33]</td>
</tr>
<tr>
<td>Classification and regression trees (CART)</td>
<td>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]</td>
</tr>
<tr>
<td>Random forest classifier</td>
<td>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]</td>
</tr>
<tr>
<td>Neural network</td>
<td>Set of algorithms designed to recognize patterns [36] modeled after the organization of neurons in the brain and the basis for machine learning algorithms</td>
</tr>
<tr>
<td>Machine learning</td>
<td>Allows algorithms to learn from data and make unsupervised decisions</td>
</tr>
<tr>
<td>Deep learning</td>
<td>Subset of machine learning by which neural networks automatically learn from unlabeled or unstructured data [37]</td>
</tr>
<tr>
<td>Support vector machine (SVM)</td>
<td>Supervised learning model with algorithms for data classification and regression [38]</td>
</tr>
</tbody>
</table>
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 ± 9.8% sensitivity and 93 ± 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 μm per slice, less than the size of a single basal cell [47–52]. However, 88% accuracy with an average error of 7.9 μm 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 μm average error and the DEJ of darker skin phototypes with 5.41 μm 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>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Method</th>
<th>Average accuracy</th>
<th>Correlation</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somoza [43]</td>
<td>2014</td>
<td>Unsupervised texton-based approach</td>
<td>54%</td>
<td>0.84–0.95</td>
<td>0.527</td>
<td>0.9</td>
</tr>
<tr>
<td>Hames [24]</td>
<td>2016</td>
<td>Bag of features and logistic regression</td>
<td>85.60%</td>
<td>0.54–0.74</td>
<td>0.839</td>
<td>0.921</td>
</tr>
<tr>
<td>Hames [25]</td>
<td>2015</td>
<td>Patch-based bag-of-features approach followed by Support vector machine classifier</td>
<td>85.70%</td>
<td>0.381</td>
<td>0.802</td>
<td></td>
</tr>
<tr>
<td>Kaur [26]</td>
<td>2016</td>
<td>Hybrid deep learning</td>
<td>82%</td>
<td>0.7174</td>
<td>0.962</td>
<td></td>
</tr>
<tr>
<td>Bozkurt [12]</td>
<td>2017</td>
<td>Recurrent convolutional neural network</td>
<td>87.97%</td>
<td>0.8711</td>
<td>0.93957</td>
<td></td>
</tr>
<tr>
<td>Bozkurt [44]</td>
<td>2017</td>
<td>Recurrent convolutional neural network</td>
<td>88.18%</td>
<td>0.87327</td>
<td>0.94070</td>
<td></td>
</tr>
</tbody>
</table>
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, textual 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 publications. 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].

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

<table>
<thead>
<tr>
<th>Epidermal stratification</th>
<th>Dermal-epidermal junction (DEJ) delineation</th>
<th>Classifying pigmented lesions</th>
<th>Quantification of photoaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet transform [46]</td>
<td></td>
<td>Bayesian Model [63]</td>
<td>Wavelet decomposition with support vector machine [64]</td>
</tr>
</tbody>
</table>
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 noninvasive 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 advances 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


