Received: 18 December 2022 | Accepted: 22 June 2023

DOI: 10.1111/ddg.15194

ORIGINAL ARTICLE

Artificial intelligence-based PRO score assessment in actinic keratoses from LC-OCT imaging using Convolutional Neural Networks

Janis R. Thamm[1] | **Fabia Daxenberger**[2] | **Théo Viel**[3] | **Charlotte Gust**[2] | **Quirine Eijkenboom**[2] | **Lars E. French**[2] | **Julia Welzel**[1] | **Elke C. Sattler**[2] | **Sandra Schuh**[1]

[1]Department of Dermatology and Allergology, University Hospital, University of Augsburg, Augsburg, Germany

[2]Department of Dermatology and Allergology, University Hospital, LMU Munich, Munich, Germany

[3]DAMAE Medical Paris, Paris, France

Correspondence
Janis Thamm, MD, Department of Dermatology and Allergology, University Hospital Augsburg, Sauerbruchstrasse 6, 86179 Augsburg, Germany.
Email: janis.thamm@uk-augsburg.de

Summary

Background and Objectives: The histological PRO score (I–III) helps to assess the malignant potential of actinic keratoses (AK) by grading the dermal-epidermal junction (DEJ) undulation. Line-field confocal optical coherence tomography (LC-OCT) provides non-invasive real-time PRO score quantification. From LC-OCT imaging data, training of an artificial intelligence (AI), using Convolutional Neural Networks (CNNs) for automated PRO score quantification of AK *in vivo* may be achieved.

Patients and Methods: CNNs were trained to segment LC-OCT images of healthy skin and AK. PRO score models were developed in accordance with the histopathological gold standard and trained on a subset of 237 LC-OCT AK images and tested on 76 images, comparing AI-computed PRO score to the imaging experts' visual consensus.

Results: Significant agreement was found in 57/76 (75%) cases. AI-automated grading correlated best with the visual score for PRO II (84.8%) vs. PRO III (69.2%) vs. PRO I (66.6%). Misinterpretation occurred in 25% of the cases mostly due to shadowing of the DEJ and disruptive features such as hair follicles.

Conclusions: The findings suggest that CNNs are helpful for automated PRO score quantification in LC-OCT images. This may provide the clinician with a feasible tool for PRO score assessment in the follow-up of AK.

KEYWORDS
Actinic keratoses, artificial intelligence, Convolutional Neural Networks, LC-OCT, non-invasive diagnostics, PRO score

INTRODUCTION

Keratinocyte skin cancer (KSC) is one of the most common skin cancers in elderly patients and is typically related to extensive solar damage of the skin and/or immunosuppression beside other co-stimulatory factors.[1]

In actinic keratosis (AK) lesions, being early stages of KSC, keratinocyte dysplasia is limited to the epidermis. In contrast, loss of the dermo-epidermal junction (DEJ) can be observed in invasive KSC and defines its invasive proliferation.[2] Despite the fact that the DEJ remains intact in AK lesions, its basal growth pattern changes under the

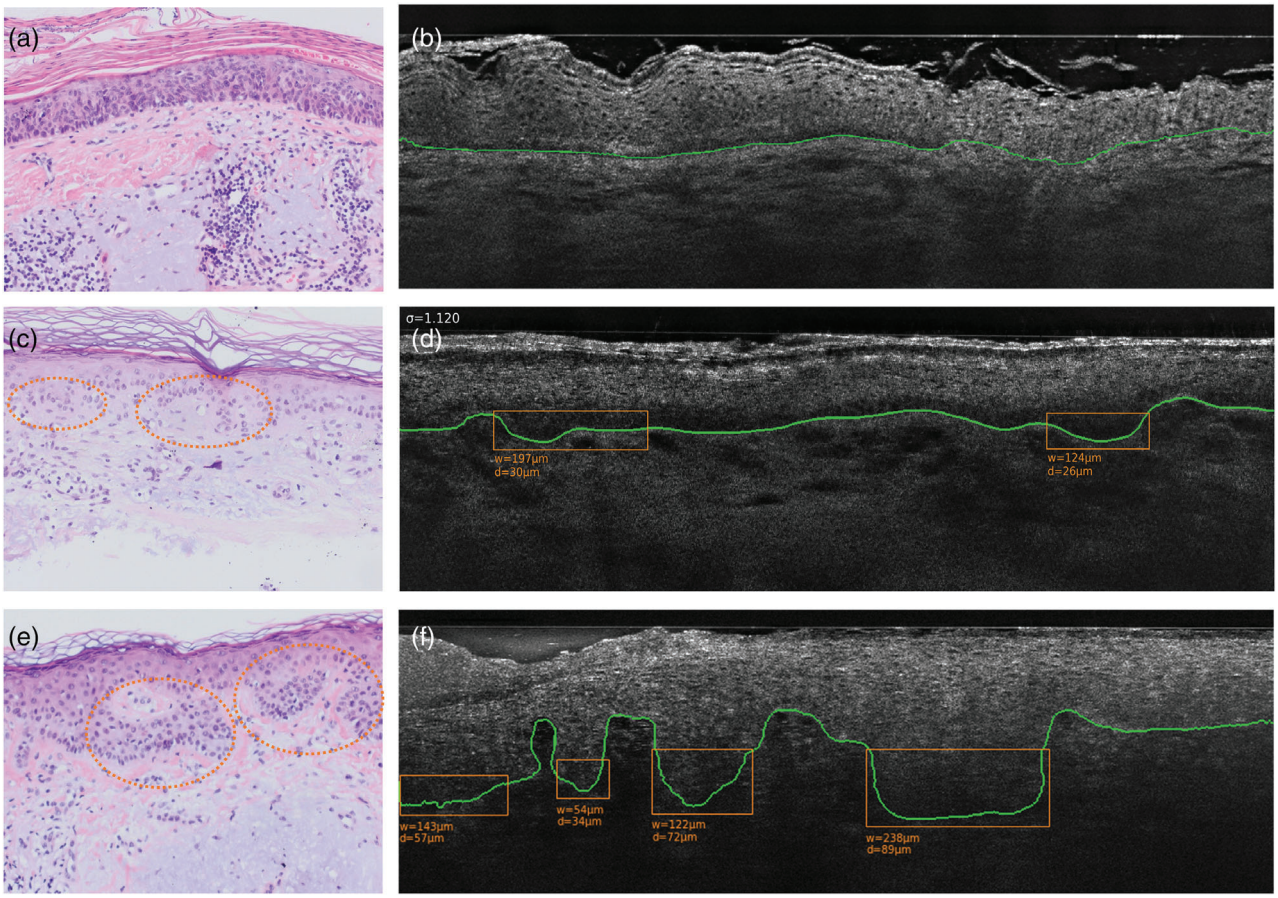


FIGURE 1 The histological PRO score and the PRO score detected by CNN algorithm from LC-OCT images (exemplary pictures). (a) For PRO I no protrusions can be observed, while for PRO II the basal epidermis protrudes (orange circles) into the dermis. (e) For PRO III cone-like protrusions are appreciated. (b, d and f) depicts the classified PRO score on LC-OCT images (image size: 1.2×0.5 mm^2, lateral and axial resolution: 1.1×1.3 μm) using CNN. (b) The green line resembles the intact DEJ, with no detected protrusions by the algorithm for PRO I. (d) For PRO II the algorithm detects slight protrusions (orange squares) and for (f) PRO III cone-like protrusions are detected comparable to histology.

transformation process towards invasive KSC.[3] Macroscopically, AK lesions appear as solitary pink to brown macules, usually accompanied by hyperkeratosis on sun-exposed areas of the skin.[1] In the clinical setting, clinical and dermatoscopical examinations are frequently used to diagnose AK lesions and for therapy monitoring.[4]

Recently, distinct morphological changes in the basal growth patterns were shown to predict the transformation into invasive KSC. They can be categorized histologically using the PRO score I-III.[3,5] Early-stage PRO I is characterized by the crowding of atypical keratinocytes in the basal layer (Figure 1a), in PRO II round nests or protrusions into the upper papillary dermis thinner than the overlying epidermis are seen (Figure 1c). In PRO III, spikes of atypical keratinocytes protruding into the dermis that are thicker than the overlying epidermis can be observed (Figure 1e).[5] Hence, from the clinical inspection of solitary and extensive AK lesions alone, no valid prediction can be made about the likelihood of progression into invasive KSC.[6–8] While histology is still considered the gold standard in diagnosis of AK, it is often not used routinely for first-line or follow-up diagnostics due to its invasiveness. For this reason, AK lesions are often not assessed for their initial PRO score and are also not monitored for changes of the PRO score under treatment on a routine basis. Therefore, the assumption of therapy response oftentimes depends on clinical inspection alone, bearing the risk that transformation of the basal growth patterns towards invasiveness may proceed, while the clinically apparent skin surface seems to recover.

Line-field confocal optical coherence tomography (LC-OCT) is a non-invasive imaging tool, providing real-time in vivo imaging of the skin. LC-OCT creates higher resolution scans of the epidermis and the upper dermis than conventional optical coherence tomography (OCT), while reaching a higher detection depth compared to reflectance confocal microscopy (RCM).[9] Previous studies conducted by Ruini et al. could show that LC-OCT can provide *in vivo* real-time imaging of AK lesions and allows evaluation of the downward proliferation pattern of keratinocytes in AK lesions in agreement with histology.[9,10] Manually practiced real time evaluation of the PRO score in AKs in LC-OCT images may considerably depend on the observers' subjective assessment and experience and therefore may show a

high interobserver variability. Since LC-OCT imaging is used only occasionally in the context of AK diagnosis and follow up so far, clinicians may not be well trained in the use of such technologies and might feel more comfortable with making their diagnosis based on histology or clinical findings. Therefore, providing the clinician with an automated analysis of the images may help to reduce the expertise gap between invasive and non-invasive technologies and may also lead to a more frequent use of LC-OCT imaging in the diagnosis and follow-up of AK in daily routine. Automated evaluation of visual data can be achieved using Convolutional Neural Networks (CNN).[11] Convolutional Neural Networks are the most largely adopted deep learning architectures nowadays for computer vision. UNet is a specific architecture of CNNs designed for medical image segmentation. It allows a segmentation at pixel level, using a series of convolutional layers with down sampling then a series of up-sampling layers.[12] In the medical field, CNNs already demonstrated their significant value in the evaluation of diabetic retinopathy, lymph node metastasis detection or skin lesion classification on expert level.[13–15]

The aim of this study was to find out whether the use of a CNN, that can automatically segment each skin layer in LC-OCT images, could lead to an automated PRO score assessment of AKs. Hypothesizing that training of an artificial intelligence (AI) on LC-OCT imaging data using CNN may eventually allow automated real-time PRO score quantification from live-imaging data, this would provide the clinician with an adequate tool for automatized PRO score detection and allow real-time assessment of epidermal and dermal skin health.

PATIENTS AND METHODS

Dataset and annotation

The machine learning model was trained on a database of LC-OCT vertical section images (histologic-like orientation) acquired using the deepLive device (deepLive™ DAMAE Medical, Paris, France) on healthy skin (HS) volunteers and patients with AK. The HS dataset included 3,701 LC-OCT images obtained by DAMAE Medical from 142 individuals. 3,092 images were acquired from the face and 285 from the upper arm of volunteers with skin phototype II (Fitzpatrick) or less. 324 images were acquired from the face and upper arm of volunteers with skin phototype VI. The AK data set contained a total of 534 images, gathered from 142 AK lesions of a cohort of 84 patients, recruited at the University Hospital of the Ludwig Maximilian University (LMU), Munich, from December 2019 to December 2021. In 56 cases AK diagnosis was histopathologically confirmed. Prior to analysis, the AK dataset was divided into two separate study cohorts, namely a training and a test set. The training set comprised a total of 458 images, extracted from 104 AK lesions of 59 patients, obtained between December 2019 and June 2021. The test set consisted of 76 images extracted from 38 AK lesions of 25 patients collected from March 2021 to December 2021. For each lesion, a subset of the most representative LC-OCT images was selected, based on the visibility of the DEJ and protrusions. In cases which exhibited a significant variation of the epidermis and DEJ, multiple distinct images (up to 4 images per lesion on the test set) were selected from the same lesion. For each image of the training set the ground truth contour of the DEJ was annotated manually by an operator at DAMAE Medical, experienced in LC-OCT image reading. The ground truth for PRO score assessment (PRO I–III) was obtained from the consensus of two dermatologists and noninvasive imaging experts (J.W., E.S.) and a resident (J.T.) at LMU und Augsburg University Hospitals. A subset of 76 images of the test set was graded for the visual PRO score, applying the same criteria as for histopathological sections. The data pipeline used is depicted in Figure 2a.

Segmentation task

Segmentation models were first trained to segment skin layers on the HS dataset (3,701 images, 142 individuals) and subsequently finetuned on images from the AK dataset (458 images, AK lesions, 59 patients). The detailed protocol used can be followed in the supplement section.

PRO score models

PRO score models were trained on a subset of 237 LC-OCT AK images (67 lesions, 41 patients). Protrusions were detected by comparing the position of the DEJ to a reference estimator of a flat DEJ position. The skin surface was used as the estimator, shifted by the average thickness of the epidermis. Areas with a DEJ below the estimator were classified as protrusions. Once these areas were detected, the width of the protrusion was quantified as the distance between the two points of intersection where the DEJ crossed the estimator. The depth was computed as the difference between the highest and lowest point of the DEJ, in the interval described by the two crossing points expanded by a small margin of 10 μm. This definition was congruent with the definition of protrusions by Schmitz et al. and is robust to small imperfections of the DEJ segmentation (Figure 4).[16] We used the protrusion depth (d) and DEJ undulation metrics (σ) to subsequently compute the PRO score, adapting the pipeline reported by Schmitz et al. to our metrics.[16] The following rules were learned using a shallow decision tree: PRO I was characterized by a low DEJ undulation: $\sigma < \tau 1$, PRO II was characterized by a moderate DEJ undulation: $\tau 1 \leq \sigma < \tau 2$, PRO III is characterized by a high DEJ undulation: $\tau 2 \leq \sigma$ and a high maximum protrusion depth: $d > \tau d$, otherwise the image was graded PRO II. The three thresholds were learned on the validation data to maximize the accuracy of the obtained decision tree classifier (Figure 2c–f).

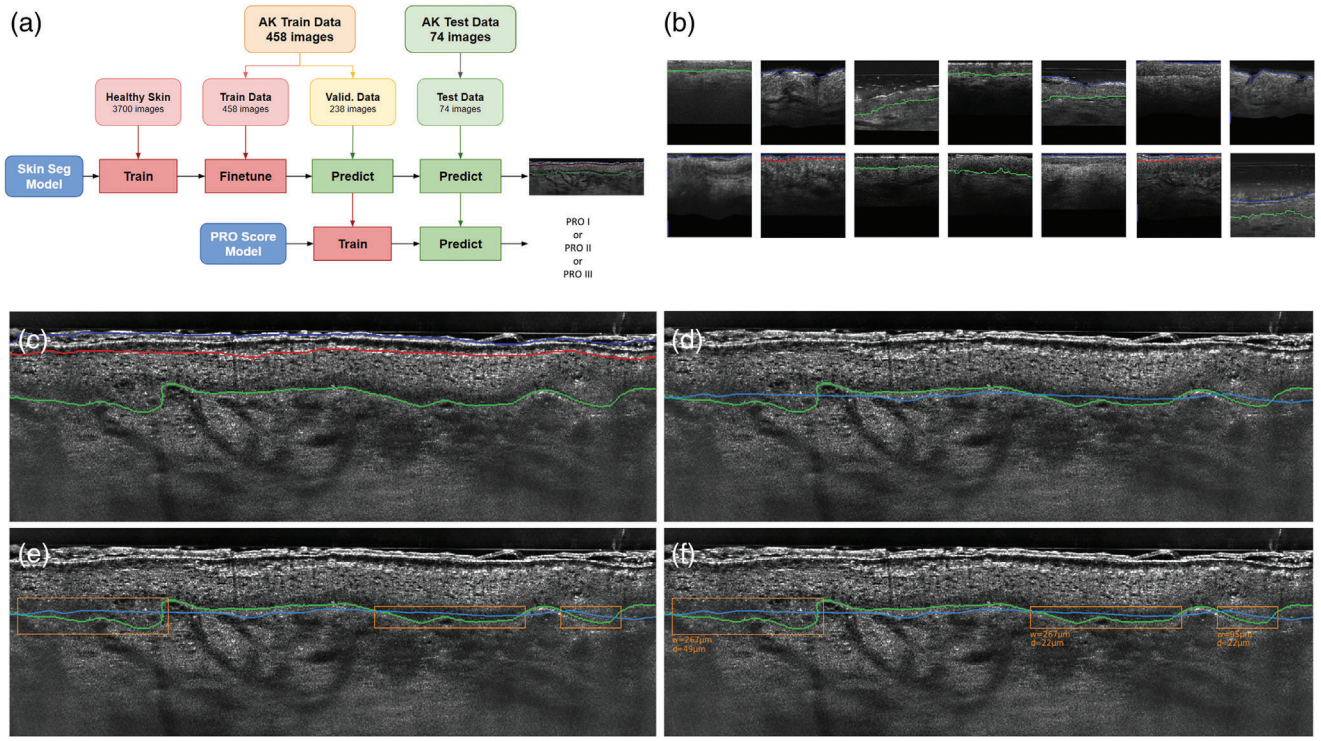


FIGURE 2 Training of the CNN algorithm for correct DEJ segmentation. (a) depicts the pipeline used for data annotation for the training and test set. (b) shows exemplary LC-OCT images fed to the CNN algorithm for DEJ segmentation. (c–f) Development of PRO score models in accordance with histology. The segmentation was performed on LC-OCT images (image size: 1.2 × 0.5 mm^2, lateral and axial resolution: 1.1 × 1.3 μm) as a 2D semantic segmentation task using a U-Net with a SEResNeXt-50 backbone pretrained on ImageNet. Protrusions were detected by comparing the position of the DEJ (green line) to a reference estimator (bright blue line) of a flat DEJ position. The skin surface (blue line) was used as the estimator, shifted by the average thickness of the epidermis. (e, f) Areas with a DEJ below the estimator were classified as protrusions (orange squares) and their width/depth were automatically computed.

Statistics

Statistical testing was performed using Python's SciPy library.[17] Experimental results were reported as means and standard deviations for normally distributed data and medians and interquartile ranges (IQR) for non-normally distributed data. Numerical data were analyzed using Student's *t* test or Mann-Whitney Wilcoxon test, depending on the normality of the distribution. The correlation of the outcome of the segmentation model (DEJ undulation and maximum protrusion depth) and the PRO score were measured using Spearman's correlation coefficient. Linearly weighted Cohen Kappa coefficient with the attached 95% CI was used to assess the agreement between visual grading and predicted PRO score. The level of significance was set to $p < 0.05$ for all statistical analyses performed.

Ethical approval

Ethical approval was obtained from the Ethics Committee of the LMU Munich under project no. 22–0781 for AI-based training of diagnostic algorithms and project no. 17–699 for diagnosis of skin lesions using LC-OCT.

RESULTS

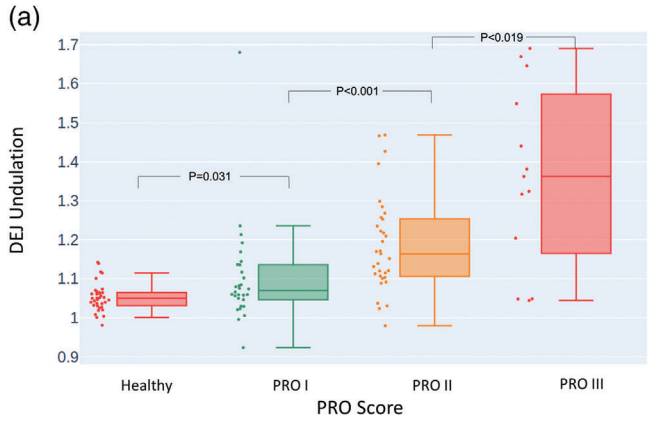
Segmentation task

The performance, as measured by IoU metric, was 0.979 ± 0.016 on the healthy skin dataset and 0.968 ± 0.028 on the AK dataset set ($p < 0.001$).

The average absolute distance εn (μm) in between the ground truth contour of the DEJ and the prediction of the model was 6.0 ± 3.8 μm on HS and 7.8 ± 6.5 μm on AK ($p < 0.001$). An increase of this distance from AK PRO I to PRO II was observed (4.7 ± 4.6 μm vs. 8.0 ± 7.3 μm, $p < 0.001$), similar to PRO II to PRO III (8.0 ± 7.3 μm vs. 12.0 ± 6.8 μm, $p = 0.021$) but not in between HS and AK PRO I (6.0 ± 3.8 μm vs. 4.7 ± 4.6 μm, $p = 0.15$) (Figure 3a-c).

Visual PRO score grading and PRO score models

For the clinical visual PRO score grading, 30/76 images were identified as PRO I, 33/76 as PRO II and 13/76 as PRO III. On this dataset, the AI undulation index and maximum protrusion depth correlated with the visual protrusion



(c)

Dataset	Subset	Class	IoU	ε_n (μm)
Healthy		Surface	0.996	1.5
		SC	0.988	4.1
		DEJ	0.979	6.0
AK	PRO I	DEJ	0.980	4.7
	PRO II		0.969	8.0
	PRO III		0.951	12.0

(d)

Ground truth \ *Predicted Label*	PRO I	PRO II	PRO III
PRO I	20	10	0
PRO II	4	28	1
PRO III	3	1	9

FIGURE 3 DEJ segmentation and PRO score models. (a) Significant increase of the DEJ undulation index and (b) the maximum protrusion depth comparing the metrics of healthy skin to PRO I, PRO II and PRO III. (c) depicts the metrics on the segmentation task, revealing a significant increase of average absolute distance εn (μm) in between the ground truth contour of the DEJ. (d) The AI-automated grading agreed with the experts' visual grading in 57/76 (75%) of the cases.

grading ($r = 0.49$ and $r = 0.47$ respectively, $p << 0.001$). The undulation index increased from healthy skin to PRO I (1.050 vs. 1.070, $p < 0.05$), from PRO I to PRO II (1.070 vs. 1.164, $p < 0.05$) and from PRO II to PRO III (1.164 vs. 1.362, $p < 0.05$) (Figure 3a). The maximum protrusion depth increased from PRO I to PRO II (0 vs. 38 μm, $p < 0.05$) and from PRO II to PRO III (38 vs. 73 μm, $p < 0.05$) (Figure 3b). The learned thresholds from the decision tree of the automated PRO score were $\tau 1 = 1.09$, $\tau 2 = 1.3$ and $\tau d = 64$. Using these thresholds, the AI automated grading derived from the undulation index and maximum protrusion depth agreed with the expert visual grading in 57/76 (75%) of the cases (Figure 3d). The weighted kappa for visual and AI classification was $\kappa = 0.60$ ($p = 6 \times 10^{-8} < 0.001$, 95% CI = [0.43, 0.77]). 20/30 (66.7%) were identified successfully as PRO I (Figure 1b), 28/33 (84.8%) as PRO II (Figure 1d) and 9/13 (69.2%) as PRO III (Figure 1f). Overall, in 11/76 (14.5%) of cases AI overestimated protrusions, while in 8/76 (10.5%) protrusions were underestimated. For PRO I, 10/30 were falsely overestimated as PRO II. In the group of PRO II, 4/33 were underestimated as PRO I, while 1/33 was assigned to PRO III. For PRO III, four underestimations occurred, 3/13 were falsely detected as PRO I, and 1/13 as PRO II (Figure 3d).

DISCUSSION

This study was able to present the successful training of an AI-based PRO score quantification algorithm for AK lesions using CNNs. The AI-based PRO score quantification correlated well with the experts' visual grading (ground truth) in 75% of the cases, with a statistically significant weighted kappa $\kappa = 0.60$ ($p = 6 \times 10^{-8} < 0.001$, 95 % CI = [0.43, 0.77]). Therefore, incidental accordance between AI-based and visual grading was ruled out and implies effective training of the algorithm, close to the experts' consensus. This level of accuracy was judged satisfactory, considering the inherent difficulty to establish PRO score ground truth, with an inter-rater agreement Kappa of respectively 0.793 and 0.84 on histopathological and LC-OCT images reported by Ruini et al. and Schmitz et al.[9,16] Although a high performance of the algorithm was accomplished using the current database of more than 4,000 images, a larger training database would consecutively improve correct DEJ detection. Given the fact that the training database will increase overtime, it is estimated that feeding the algorithm with a larger data set of diverse skin types and AK lesions, will lead to further improvement of the AI algorithm over time accordingly. Convolutional Neural Networks were used successfully to explicitly segment the skin layers. Their output was used to quantify the undulations of the DEJ and the depth of protrusions with two simple and interpretable features. Moreover, validation and test metrics justified the reliability of the pipeline. The approach was voluntarily based on segmentation results, although it would be feasible to train a CNN classifier directly on the images. But this would probably lead to two critical drawbacks, which are elaborated in the following. The training to predict the PRO score directly from images would require a lot of training images which are hard to obtain, and the

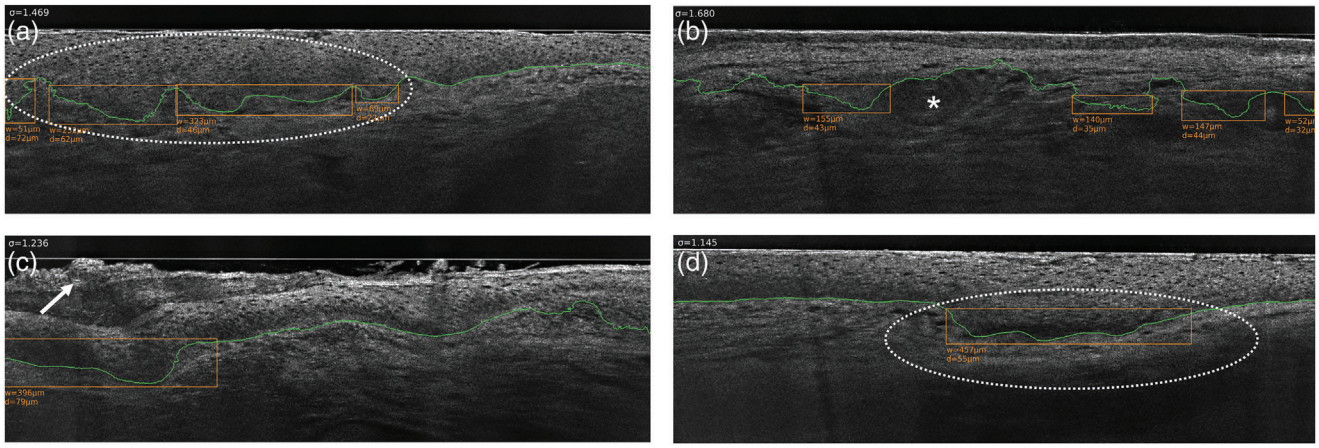


FIGURE 4 Overestimation of the PRO score. (a) Errors in LC-OCT images (image size: 1.2 × 0.5 mm^2, lateral and axial resolution: 1.1 × 1.3 μm) come from poor segmentations, (b) other pathologies affecting the epidermis such as dermal cysts (white asterisk), (c) crust on the skin (white arrow) and (d) ambiguous cases (AI detected a protrusion, which was not found characteristic in the experts' evaluation).

final predictions made by the algorithm would be hard to explain, since the model can learn from any feature of the image and the algorithm might not use the DEJ undulation contrary to the histopathological gold standard. In contrast the segmentation-based approach is performed by a very simple machine learning algorithm on top of a complex CNN, that segments the DEJ and the indicators used to predict the PRO score are based on interpretable features such as number of protrusions, protrusion depth and DEJ undulation.

For the most challenging cases, namely identification of PRO III, the average distance between the ground truth DEJ segmentation and the model prediction was 12 μm, which is lower than the average protrusion magnitude (approximately 50 μm) for this score. This finding revealed that the implemented model correctly captured protrusions, enabling automated quantification. Both metrics inferred from the segmentation did strongly correlate with the PRO score, hence allowing for the use of simple rules to classify it.

Convolutional Neural Networks for automatic PRO score detection seemed to perform best in the detection of PRO II with an accuracy of 28/33 (84.8%) vs. PRO III 9/13 (69.2%) and PRO I 20/30 (66.6%) (Figure 3d).

Still, misinterpretation occurred in a quartile of the cases, overall (Figure 3d). Overestimation of the PRO score was found in 11/76 (14.5%) of the cases and predominantly occurred in the group with ground truth PRO I. In this group, 10/30 cases were falsely assigned to PRO II. For ground truth PRO II, only a single case was overestimated as PRO III.

Different tint in static LC-OCT images was made responsible for poor DEJ segmentation, leading to the AI's assumption of DEJ deterioration. Therefore, a PRO I was classified falsely as PRO III (Figure 4a). In other cases, the overestimation of the PRO score (PRO II instead of I) was based upon an atrophic epidermis and an underlying cyst, but also due to a thick stratum corneum, leading to an overassessment of the underlying undulation (Figure 4c). In contrast to the experts' inspection, it was clear that the thick stratum corneum truly covered a homogenous undulation of the remaining upper and basal epidermis and therefore a real protrusion of the basal epidermal layer must not be assumed. Moreover, a single, slight but broad basal protrusion, also led to the AI's assumption of extensive undulation (Figure 4d), whereas the expert consensus did not identify the single protrusion to be characteristic for PRO II. From those cases the main two error sources for overestimation of the PRO score were identified as either false DEJ segmentation mainly related to the limited image quality or a falsely claimed thick epidermis, leading to misinterpretation of basal undulation and therefore assuming false protrusion.

Underestimation of the PRO score was found in 8/76 (10.5%) cases. For ground truth PRO II, 4/33 were underestimated as PRO I. More critical, for ground truth PRO III, 3/13 were falsely detected as PRO I, while a single case was assigned to PRO II. Underestimation of the PRO score was mostly due to a lack of sufficient image quality or epidermal artifacts, such as the presence of deep adjacent protrusion, only separated by thin papillae. The DEJ segmentation tended to miss these protrusions and therefore negatively affected the undulation metrics (Figure 5a, b). In cases where no undulation was detected the AI assumed a PRO I instead of III in three cases. In cases where the ground truth was ambiguous and the small protrusions were challenging to detect, a PRO I was assumed instead of PRO II (Figure 5c, d). Predominantly, misinterpretation was related to limited image quality, but this is an error source, which can easily be avoided. To solve the problem of over- or underestimation of the true PRO score, it is crucial that the imaging data acquired needs to explicitly show the DEJ and that possible artifacts, such as hair follicles or hair structures overlapping the epidermis do not lead to obliteration of the underlying DEJ structures. During live imaging the expert can correct limited image quality by averting artifacts leading to DEJ obliteration and thereby avoiding misinterpretation of the latter.

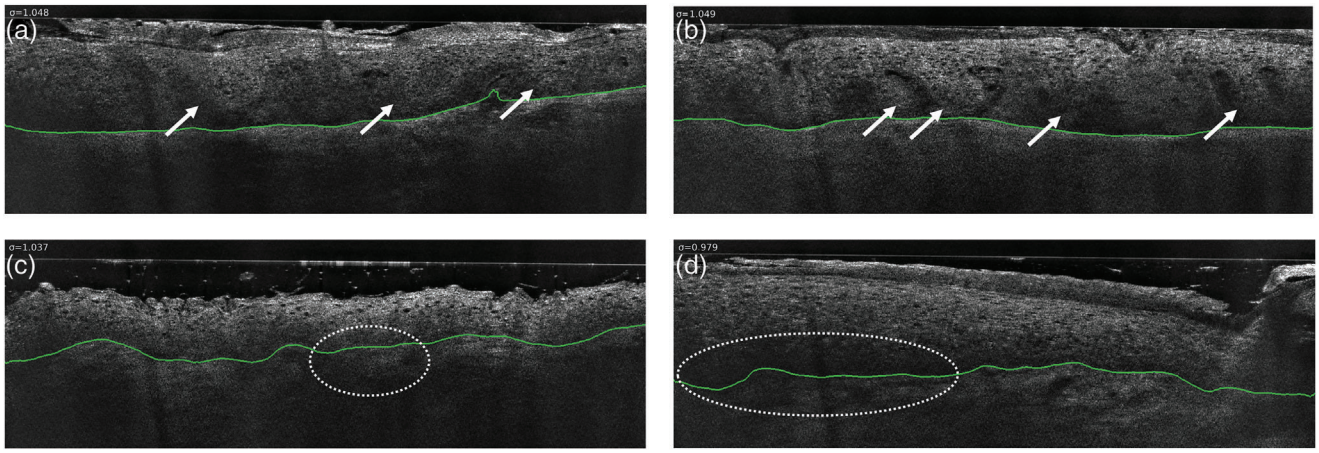


FIGURE 5 Underestimation of the PRO score. (a, b) Errors in LC-OCT images (image size: 1.2×0.5 mm^2, lateral and axial resolution: 1.1×1.3 μm) come from protrusions missed by the segmentation model (white arrows show true protrusions), (c) not entirely visible protrusions (white circle) or (d) images in which small protrusions are hard to detect and the ground truth is ambiguous (white circle).

While the overestimation of PRO score might be negligible in the clinical setting, underestimation of the true PRO score might be significant, especially in the few cases where a PRO III was misinterpreted as PRO I.

This might be critical, since histopathological findings imply that the downward proliferation in AK is associated with transformation towards KSC, which is reported in up to 20% of the cases per annum.[3,18] Hence, early detection of PRO III using non-invasive imaging tool might be crucial to implement an adequate treatment regime upon diagnosis. In a previous study, Schmitz et al. concluded, that from the visual appearance of AK no clinical valid assumption of the underlying malignant potential of AK can be assumed.[7] Due to these findings, it was stated that AKs should be admitted to an adequate treatment upon diagnosis and should be monitored carefully.[7] Other studies found that Olsen III correlated with invasiveness towards KSC, while Fernandez-Figueras et al. reported that KSC often emerges from atypia in the lower third of the epidermis (AK I according to Roewert-Huber) with normally appearing epidermal upper layers.[8,19] Hence, the value of clinical appearance assessment of AK remains unclear and is discussed controversially. The latest German S3 guidelines on AK and KSC do consider that clinical appearance does not correlate with the clinical risk or the histological grading, still there is no advice on possible risk factors for malignant transformation as only limited follow-up data on downward proliferation and keratinocyte dysplasia exists.[6]

PRO scoring and classification of atypia in AK lesions provide a histological grading of the malignant potential of AK, but invasiveness of skin biopsy often prevents the histological assessment and re-assessment, leading to the dilemma of a possible under-assessment of AK downward proliferation in the clinical setting.

LC-OCT was able to prove itself as a helpful non-invasive imaging tool for epidermal and papillary dermal structures, capable of PRO Score quantification of AK lesions *in vivo*. As AK show a broad intralesional histological grade heterogenicity in even small lesions, LC-OCT guided PRO score detection may allow a more holistic approach to lesion monitoring, whereas histology only allows a glimpse underneath the biopsy site of interest.[20] Therefore, LC-OCT may tackle the need for a non-invasive imaging tool in the follow-up of downward proliferation of AK. A possible future use of AI-automated PRO score quantification may be provision of an assessment of the whole lesion of interest, from which a minimum and maximum PRO score, as well as an average score may be assessed, and information of keratinocyte atypia may be gathered. These characteristics may provide a more comprehensive follow-up for the potential of malignant transformation in AK. The implementation of a follow-up should not be time-consuming, since it usually takes only a few minutes to cover the entire lesion in live video mode, using LC-OCT. Moreover, the AI computation time for the PRO score is compatible with live display and the integration in clinical routine would provide a live estimation of the PRO score from each LC-OCT image at eight frames per seconds. In conclusion, the possible integration of the AI computed PRO score would not increase the examination time, but would provide additional value by displaying real-time evaluation of the computed PRO score. To achieve this goal, additional data from three-dimensional LC-OCT images as well as reduction of label noise are of need for the training and test sets, to finally increase robustness and performance of the segmentation and PRO score prediction in order to quantify the protrusions' morphology more precisely.

Quantification of the downward proliferation from LC-OCT imaging in AK will allow a much more comprehensive follow-up of AK. It may help to identify the most aggressive area within the lesion or to recognize lesions resistant to topical non-invasive therapy and may allow the assessment of the potential for malignant transformation.

CONCLUSIONS

In conclusion, the findings presented, suggest that CNNs are a helpful tool for automatic segmentation of skin layers in LC-OCT images and may be used in the process of AI-based PRO score quantification. Use of an AI-based PRO Score quantification tool to assess the PRO Score on LC-OCT imaging data may provide the clinician with a feasible way to integrate PRO Score quantification as a future standardized diagnostic assessment in the diagnosis and follow-up of AKs.

ACKNOWLEDGMENT

We thank DAMAE Medical for making the LC-OCT system available for this study.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

None.

REFERENCES

1. Schmitz L, Oster-Schmidt C, Stockfleth E. Nonmelanoma skin cancer – from actinic keratosis to cutaneous squamous cell carcinoma. *J Dtsch Dermatol Ges*. 2018;16(8):1002-1013.
2. Cockerell CJ. Histopathology of incipient intraepidermal squamous cell carcinoma ("actinic keratosis"). *J Am Acad Dermatol*. 2000;42(1Pt 2):11-17.
3. Schmitz L, Gambichler T, Kost C, et al. Cutaneous squamous cell carcinomas are associated with basal proliferating actinic keratoses. *Br J Dermatol*. 2019;180(4):916-921.
4. Zalaudek I, Giacomel J, Argenziano G, et al. Dermoscopy of facial nonpigmented actinic keratosis. *Br J Dermatol*. 2006;155(5):951-956.
5. Schmitz L, Gambichler T, Gupta G, et al. Actinic keratoses show variable histological basal growth patterns – a proposed classification adjustment. *J Eur Acad Dermatol Venereol*. 2018;32(5):745-751.
6. Heppt MV, Leiter U, Steeb T, et al. S3 guideline for actinic keratosis and cutaneous squamous cell carcinoma – short version, part 1: diagnosis, interventions for actinic keratoses, care structures and quality-of-care indicators. *J Dtsch Dermatol Ges*. 2020;18(3):275-294.
7. Schmitz L, Kahl P, Majores M, et al. Actinic keratosis: correlation between clinical and histological classification systems. *J Eur Acad Dermatol Venereol*. 2016;30(8):1303-1307.
8. Fernandez-Figueras MT, Carrato C, Saenz X, et al. Actinic keratosis with atypical basal cells (AK I) is the most common lesion associated with invasive squamous cell carcinoma of the skin. *J Eur Acad Dermatol Venereol*. 2015;29(5):991-997.
9. Ruini C, Schuh S, Gust C, et al. In-vivo LC-OCT evaluation of the downward proliferation pattern of keratinocytes in actinic keratosis in comparison with histology: first impressions from a pilot study. *Cancers (Basel)*. 2021;13(12).
10. Ruini C, Schuh S, Gust C, et al. Line-field confocal optical coherence tomography for the in vivo real-time diagnosis of different stages of keratinocyte skin cancer: a preliminary study. *J Eur Acad Dermatol Venereol*. 2021;35(12):2388-2397.
11. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9(4):611-629.
12. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation2015 May 01, 2015:[arXiv:1505.04597 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2015arXiv150504597R [Last accessed December 17, 2022].
13. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
14. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199-2210.
15. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-2410.
16. Schmitz L, Gupta G, Stucker M, et al. Evaluation of two histological classifications for actinic keratoses – PRO classification scored highest inter-rater reliability. *J Eur Acad Dermatol Venereol*. 2019;33(6):1092-1097.
17. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-272.
18. Marks R, Rennie G, Selwood TS. Malignant transformation of solar keratoses to squamous cell carcinoma. *Lancet*. 1988;1(8589):795-797.
19. Ahmady S, Jansen MHE, Nelemans PJ, et al. Risk of invasive cutaneous squamous cell carcinoma after different treatments for actinic keratosis: a secondary analysis of a randomized clinical trial. *JAMA Dermatol*. 2022;158(6):634-640.
20. Schmitz L, Stucker M, Gambichler T, et al. Histological intralesional heterogeneity of actinic keratoses relates to field cancerization. *J Dtsch Dermatol Ges*. 2018;16(10):1211-1217.
21. Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks. 2016 November 01, 2016:[arXiv:1611.05431 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2016arXiv161105431X [Last accessed December 17, 2022].
22. Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation Networks. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(8):2011-2023.
23. Deng J, Dong W, Socher R, et al. editors. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition; 200920–25June2009.
24. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization2014 December 01, 2014:[arXiv:1412.6980 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6980K [Last accessed December 17, 2022].

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Thamm JR, Daxenberger F, Viel T, et al. Artificial intelligence-based PRO score assessment in actinic keratoses from LC-OCT imaging using Convolutional Neural Networks. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft*. 2023;21:1359–1366. https://doi.org/10.1111/ddg.15194