OPEN ACCESS
Journal of Innovative Optical Health Sciences
Vol. 14, No. 1 (2021) 2140005 (7 pages)
© The Author(s)
DOI: 10.1142/S1793545821400058

World Scientific
www.worldscientific.com

# A measurement of epidermal thickness of fingertip skin from OCT images using convolutional neural network

Yongping Lin*, Dezi Li†, Wang Liu†, Zhaowei Zhong†, Zhifang Li†,‡, Youwu He† and Shulian Wu†

*Fujian Provincial Key Laboratory of Optoelectronic Technology and Devices
School of Optoelectronic and Communication Engineering
Xiamen University of Technology, Xiamen 361024, P. R. China

†Key Laboratory of Optoelectronic Science and Technology for Medicine
Ministry of Education, Fujian Provincial Key Laboratory of Photonics Technology
Fujian Provincial Engineering Technology Research Center
of Photoelectric Sensing Application
College of Photonic and Electronic Engineering
Fujian Normal University
Fuzhou, Fujian 350007, P. R. China
‡lizhifang@fjnu.edu.cn

Received 21 September 2020
Accepted 26 November 2020
Published 21 December 2020

In this study, we proposed a method to measure the epidermal thickness (ET) of skin based on deep convolutional neural network, which was used to determine the boundaries of skin surface and the ridge portion in dermal–epidermis junction (DEJ) in cross-section optical coherence tomography (OCT) images of fingertip skin. The ET was calculated based on the row difference between the surface and the ridge top, which is determined by search the local maxima of boundary of the ridge portion. The results demonstrated that the region of ridge portion in DEJ was well determined and the ET measurement in this work can reduce the effect of the papillae valley in DEJ by 9.85%. It can be used for quantitative characterization of skin to differentiate the skin diseases.

*Keywords*: Epidermal thickness; cross-section OCT images; convolutional neural network.

## 1. Introduction

Skin analysis is one of the most popular and interesting tasks, since skin is the outermost part of the human body. Epidermis and dermis are two major layers of the skin which are separated by the dermal-epidermal junction (DEJ). Many skin

‡Corresponding author.

This is an Open Access article. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.



J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com
by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com
by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

diseases occur along with changes of the epidermal thickness (ET), such as atopic dermatitis,[1] psoriasis[2] and lichen sclerosus.[3] Thus, the ET is an important indicator for diagnosing skin disease.

Two noninvasive optical imaging techniques, multiphoton laser tomography[2,4,5] and optical coherence tomography (OCT),[6–8] were applied for measuring ET of skin. In the multiphoton laser tomography, the calculating method of ET is based on the profile of the reflected light in terms of depth. However, the method usually was used for the murine skin, not human, since the imaging depth is lower than the ET of human skin.

In OCT, an approximate depth penetration of 1 mm is sufficient to provide high-resolution images of the human epidermis *in vivo*. There are two leading methods for estimating the ET: the profile of A-scan signals[7,9,10] and image segmentation.[6,11] The profile of A-scan signals was used for estimating the ET by measuring the distance between the entrance peak and the second peak of the A-scan signals. Intensity feature was used for epidermis segmentation and the distance between upper and lower boundary was used for ET measurement. Recently, a deep-learning algorithm for automatic segmentation of several layers of mouse skin in OCT image data was developed using a deep convolutionl neural network (CNN), and the thickness of different layers was estimated.[8] These two methods were similar because they were all based on the characteristics of optical scattering intensity of top and bottom boundaries of epidermis.

The DEJ is a three-dimensional (3D) complex structure and seen as a thick transition band.[12] Such a junction area can be further divided into papillae ridge and papillae valley portions.[13] The ridge portion is defined as the structure from the tops of the papillae ridges to the bottoms of the papillae valleys. Neerken *et al.*[14] demonstrated that OCT–ET was larger than confocal laser scanning microscopy ET because the OCT–ET measurement includes the whole upper portion of the papillary dermis. In the cross-sectional image of a volunteer's thumb in Yu *et al.*'s research,[13] tops of papillae ridge were determined to calculate the ET since that the epidermis layer is defined from the skin surface to the DEJ binding epidermis and dermis tightly together. The accuracy of OCT measurement for ET can be improved by using the accurate upper boundary of papillary. However, it is difficult to determine the top and bottom of individual ridges for use in automatically segmented algorithms. Thus, the measurement of human skin ET *in vivo* is still challenging.

Deep-learning methods have been employed for automatic identification and segmentation of biomedical images.[15–17] U-Net is a generic deep-learning solution in biomedical image data. It can be trained end-to-end from very few images and perform a good result.[18] Furthermore, it is a tool to apply individual labeling rules to large data sets and thereby can save manual annotation effort in a vast variety of quantification tasks.[19] R. Asgari etc. employed an extensive U-Net model to the task of segmenting drusen and got a sensible result.[20]

In this work, we focus on automatic segmentation of cross-section OCT images of fingertip skin based on U-Net. DEJ is a wave-like structure as shown in Fig. 1. There is an obvious difference of the ridge and valley in intensity. Thus making a automatic CNN model to determine both boundaries of skin surface and ridge portion is feasible. The top of ridge portion could be determined by search the local maxima of boundary of the ridge portion.

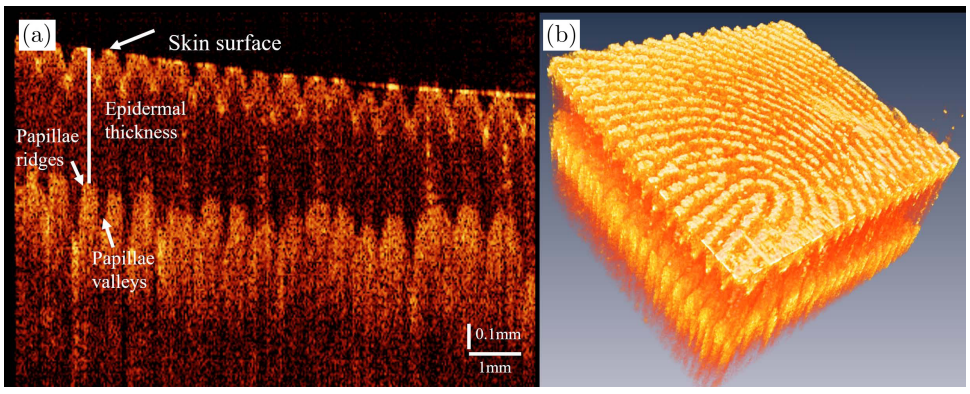



Fig. 1. (a) Typical cross-sectional OCT image and (b) 3D OCT image of fingertip skin.

J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com
by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

Therefore, the ET was calculated based on the row difference between the ridge top and surface.

## 2. Methods

### 2.1. *OCT system and sample*

In this work, the composition of the home-built spectral domain OCT is similar to the previous work.[21] The light source is a 12 mW PM-coupled superluminescent diode (SLD) with a FWHM bandwidth of 85 nm centered at 1310 nm (S5FC1021P, Thorlabs). The spectrometer (C-1235-1385, Wasatch Photonics) is integrated with a 2048-pixel CCD camera. Its maximum line scanning speed is 76 kHz the maximum imaging depth is 5.8 mm. The main difference is that the nonpolarized light are used in the sample and reference arms. Furthermore, only one reference arm is used in the OCT system. The axial and lateral resolutions in the air are 8.9 $\mu$m and 25 $\mu$m, respectively. As for the construction of a 3D image, 400 B-mode OCT images are acquired, in increments of 25 $\mu$m of the position of the light beam. The scanning range of the sample is $1 \times 1\,\text{cm}^2$. The human fingertip skin is chosen for the study.

### 2.2. *Definition of ET*

Figure 1 shows the typical cross-sectional and 3D OCT image of fingertip skin. The ridge portion is composed of the structure from the tops of papillae ridges to the bottoms of papillae valleys. The ET is determined by the pixel difference between the top of papillae ridge and the segmented skin surface at the same column. Once the top of papillae ridge was determined, the corresponding position was recorded in order to calculate the ET of fingertip skin. The distance of one pixel is equal to 6 $\mu$m. Hence, the actual thickness can be calculated in the following equation: $\text{ET}(x, y) = (n_{\text{epi}} - n_{\text{surf}}) \times 6\,\mu\text{m}$, where $n_{\text{epi}}$ is the row value at the column of ridge top, $n_{\text{surf}}$ is the row value at the column of skin surface. This calculated method of ET is based on distances between epidermal and dermal regions expressed in pixels for cross-sectional OCT images.

### 2.3. *Segmentation by U-Net*

In order to calculate the ET of the skin, the boundaries of both surface and the ridge portion were segmented employing U-Net which is shown in Table 1. A total of 70 original input OCT images containing various skin depths and shapes were obtained. To increase the number training samples, the images were augmented to 100 images by randomly sheared, zoomed, rotated, stretched, horizontally and vertically flipped versions of the original images. These images were randomly split into training (70%) and test (30%) sets. To avoid overfitting, we used dropout on the recurrent connections and an L1 regularization penalty on the recurrent weights with coefficients of 0.1 and 0.05, respectively. The model was implemented by Keras (see https://keras.io/) which is a high-level application programming interface written in python. We perform our computation using the Theano

Table 1. U-Net architecture.

| Layer | Parameters | Activation | Output |
|---|---|---|---|
| Input | | | 1*512*512 |
| Conv 1 | 64*3*3 | ReLU | 512*512*64 |
| Conv 2 | 64*3*3 | ReLU | 512*512*64 |
| MAX POOL 1 | 64*2*2, stride 2 | | 256*256*64 |
| Conv 3 | 128*3*3 | ReLU | 256*256*128 |
| Conv 4 | 128*3*3 | ReLU | 256*256*128 |
| MAX POOL 2 | 128*2*2, stride 2 | | 128*128*128 |
| Conv 5 | 256*3*3 | ReLU | 128*128*256 |
| Conv 6 | 256*3*3 | ReLU | 128*128*256 |
| MAX POOL 3 | 256*2*2, stride 2 | | 64*64*256 |
| Conv 7 | 512*3*3 | ReLU | 64*64*512 |
| Conv 8 | 512*3*3 | ReLU | 64*64*512 |
| Dropout 1 | 0.5 | | 64*64*512 |
| MAX POOL 4 | 512*2*2, stride 2 | | 32*32*512 |
| Conv 9 | 1024*3*3 | ReLU | 32*32*1024 |
| Conv 10 | 1024*3*3 | ReLU | 32*32*1024 |
| Dropout 2 | 0.5 | | 32*32*1024 |
| UPConv 11 | 512*2*2 | | 64*64*512 |
| Concat | 512*3*3 | | 64*64*1024 |
| Conv 12 | 512*3*3 | ReLU | 64*64*512 |
| Conv 13 | 512*3*3 | ReLU | 64*64*512 |
| UPConv 14 | 256*2*2 | ReLU | 128*128*256 |
| Concat | 256*3*3 | | 128*128*512 |
| Conv 15 | 256*3*3 | ReLU | 128*128*256 |
| Conv 16 | 256*3*3 | ReLU | 128*128*256 |
| UPConv 17 | 128*2*2 | ReLU | 256*256*128 |
| Concat | 128*3*3 | | 256*256*256 |
| Conv 18 | 128*3*3 | ReLU | 256*256*128 |
| Conv 19 | 128*3*3 | ReLU | 256*256*128 |
| UPConv 20 | 64*2*2 | ReLU | 512*512*64 |
| Concat | 64*3*3 | | 512*512*128 |
| Conv 21 | 64*3*3 | ReLU | 512*512*64 |
| Conv 22 | 64*3*3 | ReLU | 512*512*64 |
| Conv 23 | 2*3*3 | ReLU | 512*512*2 |
| Output | 1*1*1 | Sigmoid | 512*512*1 |

backend on a single NVIDIA Tesla K40GPU. As a result, the testing accuracy is 0.95.

## 3. Results and Discussions

Figure 2(a) demonstrated that both boundaries of surface and ridge portion were determined correctly to segment the epidermis using U-Net, although the structure of skin is complex. The region inside two boundaries contains the papillae valleys since the DEJ is a wave-like structure. Then the ridge tops were located and labeled with numbers as shown in Fig. 2(b) to help the classification of papillae valley in epidermis. The measurement of ET is accurate due to the row position difference between the ridge tops and surfaces.

Figure 3(a) shows the two boundaries of surface and ridge portion in terms of pixels, and the ridge tops were found by locating the peak of boundary curve. The ET distribution is shown in Fig. 3(b). The mean of ET is $550.2 \pm 38.8\,\mu$m. In addition, Fig. 3(b) also demonstrates the ET distribution based on the ET calculation using row difference between the two boundaries. The mean is $604.4 \pm 63.3\,\mu$m, which is overestimated by 9.85% for the ET, mainly due to classification of papillae valley in epidermis.

The two envelop curves of local maxima and minima constructed the region of ridge portion. The ridge height was defined as the difference between the local maxima an minima[22] and used for characterizing DEJ structure. Figure 3(c) showed the distribution of ridge height. The mean is $129.8 \pm 51.4\,\mu$m which is larger than that presented in Newton study.[22] This is primarily due to different measurement methods and sample and regional difference.

Typical cross-sectional OCT images of the back of palm, palm and fingertip skin are shown in Fig. 4. The dermal-epidermal junction (DEJ) is a big obvious difference between fingertip skin and the back of the palm and palm. The papillary structure is

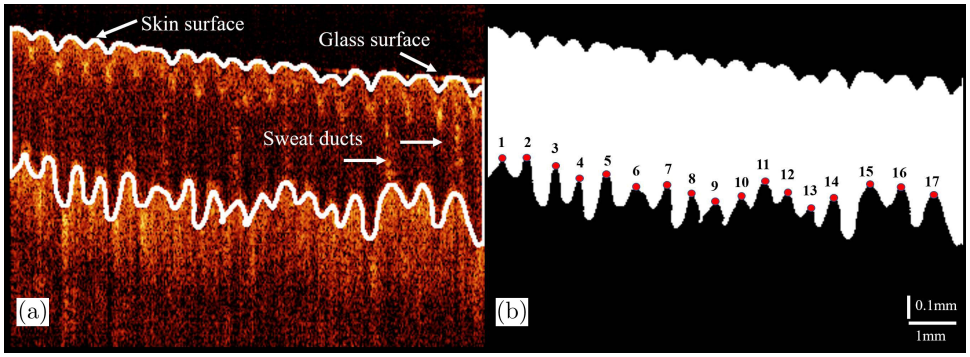
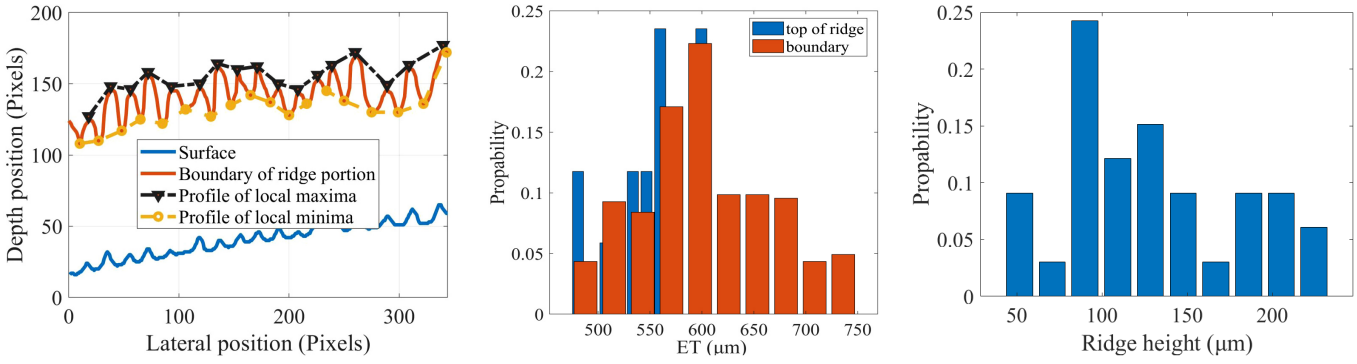



Fig. 2. (a) boundaries of skin surface and ridge portion and (b) all tops of ridge labeled with number.




(a) (b) (c)

Fig. 3. (a) Two boundaries of surface and ridge portion as shown in Fig. 2(a), (b) ET distribution based on two methods and (c) ridge height distribution.



J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

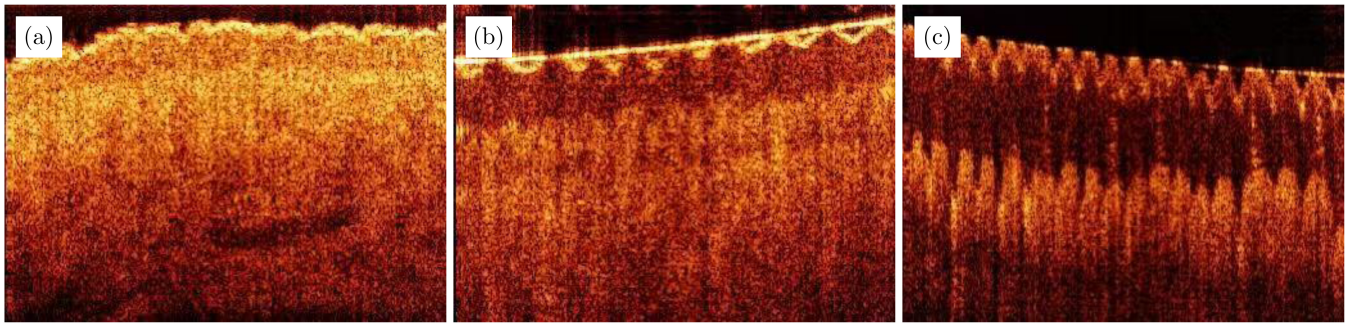



Fig. 4. Typical cross-sectional OCT image of the (a) back of the palm, (b) palm and (c) fingertip skin.

J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

easily observed in the fingertip skin but hard to determine by traditional methods. Thus, we proposed a method based on CNN to extract the DEJ in the fingertip skin via expert labeling. Envelop curves of local maxima were used to construct the upper boundary of the ridge portion, which can produce a more accurately ET. The result is in accordance with the Yu *et al.* research results.[13]

A previous work based on the profile of A-scan signals was used for estimate the ET,[23] which is widely applied for ET measurement.[7,9,10] The method is limited by the accuracy of the entrance peak and the second peak of the A-scan signals. Figure 5 demonstrates that there are several peaks in the region around the second peak, and the curve become smoothly with increasing average size. The main reason is due to the complexity of the structure of skin, especially the surface undulation which is wave-like DEJ and has sweat ducts. The complex structure includes the multiple peaks of A-scan signal so that it is difficult to determine the position of peaks to calculate ET accurately. In addition, it is not possible to locate the DEJ based on the profile of A-scan signal.

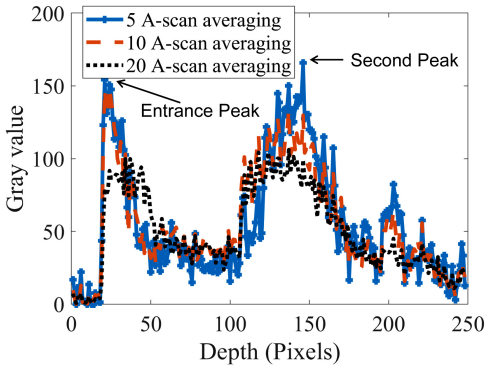



Fig. 5. OCT intensity in terms of depth with different-size A-scan averaging.

There were different segmentation methods of epidermis, such as the graph cut algorithm[24] and shapelet analysis.[6] The flowcharts of these segmentation methods are as shown in Fig. 6. For the graph cut algorithm, even small parameter value changes lead to very different segmentation results. Besides, the segmentation accuracy is not always sensitive to parameter changes.[25] The shapelet analysis fails to closely follow all of the contours of a wavy dermal–epidermal junction because the shapelet kernel is only processed in one direction.[6] Comparing the other segmentation methods, the CNN-based segmentation takes the raw B-scans of the OCT as input and segment the skin surface and

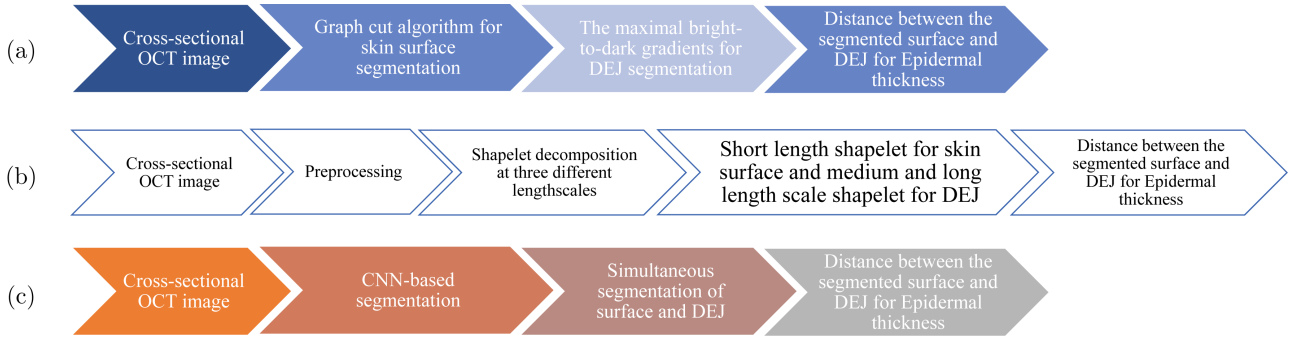



Fig. 6. (a) Gradient segmentation, (b) shapelet segmentation and (c) CNN-based segmentation for ET.

DEJ simultaneously with a pixel-wise prediction for each class label. Furthermore, the CNN-based approaches can be more easily adapted to new segmentation problems with no adjustments. The image segmentation for ET measurement by U-Net has solved the problems based on the profile of A-scan signal, such as surface undulation and multiple peaks in wave-like DEJ. However, this method was difficult to avoid the classification of papillae valley in epidermis and result in overestimation of ET.

## 4. Conclusions

In summary, we focus on automatic segmenting cross-section OCT images of fingertip skin using U-Net to determine the boundaries of skin surface and the ridge portion. The results show the top of ridge portion was determined by locating the local maxima of boundary of the ridge portion. Thus, the ET was calculated based on the row difference between the ridge top and the corresponding surface. The automatic ET measurement in this work can reduce the effect of the papillae valley in DEJ by 9.85%. In addition, the region of ridge portion in DEJ was determined by the envelop curves of local maxima and minima of the ridge boundary, and the ridge height was estimated for characterizing the DEJ. As a result, the segmentation of ridge portion in DEJ can improve the accuracy of ET. It is a potential application for quantitative characterization of skin to differentiate the skin diseases.

## Conflict of Interest

The authors have no conflicts of interest relevant to this article.

## Acknowledgments

This work has been supported by the National Natural Science Foundation of China (Nos. 61875038 and 81901787), Fujian Provincial National Science Foundation (No. 2019J01268) and the Fund for High-level Talents of Xiamen University of Technology (No. YKJ19010R).

## References

1. D. D. Hou, Z. H. Di, R. Q. Qi, H. X. Wang, S. Zheng, Y. X. Hong, H. Guo, H. D. Chen, X. H. Gao, "Sea buckthorn (*Hippophaë rhamnoides* L.) oil improves atopic dermatitis-like skin lesions via inhibition of NF-$\kappa$B and STAT1 activation," *Skin Pharmacol. Physiol.* **30**(5), 268–276 (2017).
2. H. L. Rizzo, S. Kagami, K. G. Phillips, S. E. Kurtz, S. L. Jacques, A. Blauvelt, "IL-23-mediated psoriasis-like epidermal hyperplasia is dependent on IL-17A," *J. Immunol.* **186**(3), 1495–1502 (2011).
3. J. Carlson, R. Ambros, J. Malfetano, J. Ross, R. Grabowski, P. Lamb, H. Figge, M. C. Mihm, "Vulvar lichen sclerosus and squamous cell carcinoma: A cohort, case control, and investigational study with historical perspective; implications for chronic inflammation and sclerosis in the development of neoplasia," *Hum. Pathol.* **29**, 932–948 (1998).
4. M. J. Koehler, T. Vogel, P. Elsner, K. König, R. Bückle, M. Kaatz, "*In vivo* measurement of the human epidermal thickness in different localizations by multiphoton laser tomography," *Skin Res. Technol.* **16**(3), 259–264 (2010).
5. E. Decencière, E. Tancrède-Bohin, P. Dokládal, S. Koudoro, A. M. Pena, T. Baldeweck, "Automatic 3D segmentation of multiphoton images: A key step for the quantification of human skin," *Skin Res. Technol.* **19**(2), 115–124 (2013).
6. J. Weissman, T. Hancewicz, P. Kaplan, "Optical coherence tomography of skin for measurement of epidermal thickness by shapelet-based image analysis," *Opt. Express* **12**(23), 5760 (2004).
7. T. Gambichler, S. Boms, M. Stücker, A. Kreuter, G. Moussa, M. Sand, P. Altmeyer, K. Hoffmann, "Epidermal thickness assessed by optical coherence tomography and routine histology: Preliminary results of method comparison," *J. Eur. Acad. Dermatol. Venereol.* **20**(7), 791–795 (2006).
8. T. Kepp, C. Droigk, M. Casper, M. Evers, G. Hüttmann, N. Salma, D. Manstein, M. P. Heinrich, H. Handels, "Segmentation of mouse skin layers in optical coherence tomography image data using deep convolutional neural networks," *Biomed. Opt. Express* **10**(7), 3484 (2019).
9. J. Welzel, E. Lankenau, R. Birngruber, R. Engelhardt, "Optical coherence tomography of the human skin," *J. Am. Acad. Dermatol.* **37**(6), 958–963 (1997).
10. J. Delacruz, J. Weissman, K. Gossage, Automated measurement of epidermal thickness from optical coherence tomography images using line region growing, *Photonic Therapeutics and Diagnostics VI*, Vol. 7548, p. 75480E (International Society for Optics and Photonics, 2010).
11. A. Li, J. Cheng, A. P. Yow, C. Wall, D. W. K. Wong, H. L. Tey, J. Liu, Epidermal segmentation in high-definition optical coherence tomography, *Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS)*, Milan, Italy, pp. 3045–3048 (2015).

J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

J. Innov. Opt. Health Sci. 2021.14. Downloaded from www.worldscientific.com
by 171.83.9.102 on 02/08/21. Re-use and distribution is strictly not permitted, except for Open Access articles.

12. S. Kurugol, K. Kose, B. Park, J. G. Dy, D. H. Brooks, M. Rajadhyaksha, "Automated delineation of dermal–epidermal junction in reflectance confocal microscopy image stacks of human skin," *J. Invest. Dermatol.* **135**(3), 710–717 (2015).
13. X. Yu, Q. Xiong, Y. Luo, N. Wang, L. Wang, H. L. Tey, L. Liu, "Contrast enhanced subsurface fingerprint detection using high-speed optical coherence tomography," *IEEE Photonics Technol. Lett.* **29**, 70–73 (2017).
14. S. Neerken, G. W. Lucassen, M. A. Bisschop, E. Lenderink, T. A. M. Nuijs, "Characterization of age-related effects in human skin: A comparative study that applies confocal laser scanning microscopy and optical coherence tomography," *J. Biomed. Opt.* **9**(2), 274 (2004).
15. X. Yang, W. T. Tang, G. Tjio, S. Y. Yeo, Y. Su, "Automatic detection of anatomical landmarks in brain MR scanning using multi-task deep neural networks," *Neurocomputing* **396**, 514–521 (2020).
16. X. Yang, L. Wu, W. Ye, K. Zhao, Y. Wang, W. Liu, J. Li, H. Li, Z. Liu, C. Liang, "Deep learning signature based on staging CT for preoperative prediction of sentinel lymph node metastasis in breast cancer," *Acad. Radiol.* **27**(9), 1226–1233 (2020).
17. C. J. Lynch, C. Liston, "New machine-learning technologies for computer-aided diagnosis," *Nat. Med.* **24**, 1304–1305 (2018).
18. O. Ronneberger, P. Fischer, B. Thomas, U-Net: Convolutional networks for biomedical image segmentation, *Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol. 9351, Munich, Germany, pp. 234–241 (2015).
19. T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, O. Ronneberger, "U-Net: Deep learning for cell counting, detection, and morphometry," *Nat. Methods* **16**, 67–70 (2019).
20. R. Asgari, S. Waldstein, F. Schlanitz, M. Baratsits, U. Schmidt-Erfurth, H. Bogunović, U-Net with spatial pyramid pooling for drusen segmentation in optical coherence tomography, *Int. Workshop on Ophthalmic Medical Image Analysis*, Lect. Notes Comput. Sci. (including subseries Lect. Notes Artif. Intell. and Lect. Notes Bioinf.), Vol. 11855, pp. 77–85, Springer, Cham (2019).
21. Y. He, Z. Li, Y. Zhang, H. Li, "Single camera spectral domain polarization-sensitive optical coherence tomography based on orthogonal channels by time divided detection," *Opt. Commun.* **403**, 162–165 (2017).
22. V. L. Newton, R. S. Bradley, P. Seroul, M. Cherel, C. E. Griffiths, A. V. Rawlings, R. Voegeli, R. E. Watson, M. J. Sherratt, "Novel approaches to characterize age-related remodelling of the dermal-epidermal junction in 2D, 3D and *in vivo*," *Skin Res. Technol.* **23**(2), 131–148 (2017).
23. S. Wu, H. Li, X. Zhang, Z. Li, "Optical features for chronological aging and photoaging skin by optical coherence tomography," *Lasers Med. Sci.* **28**, 445–450 (2013).
24. A. P. Yow, J. Cheng, A. Li, R. Srivastava, J. Liu, D. W. K. Wong, H. L. Tey, Automated *in vivo* 3D high-definition optical coherence tomography skin analysis system, *Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS)*, Lake Buena Vista (Orlando), Florida USA, pp. 3895–3898 (2016).
25. H. Chen, X. Pan, X. Lu, Q. Xie, "A modified graph cuts image segmentation algorithm with adaptive shape constraints and its application to computed tomography images," *Biomed. Signal Process. Control* **62**, 102092 (2020).