# Classification of the dermal-epidermal junction using in-vivo confocal microscopy

Julie Robic, Benjamin Perret, Alex Nkengne, Michel Couprie, Hugues Talbot

# CLASSIFICATION OF THE DERMAL-EPIDERMAL JUNCTION USING IN-VIVO CONFOCAL MICROSCOPY

*J. Robic*[⋆†]*, B. Perret*[†]*, A. Nkengne*[⋆]*, M. Couprie*[†]*, and H. Talbot*[†]

[⋆] Clarins Laboratories, Pontoise, France.
[†] Université Paris-Est, LIGM UMR 8049, ESIEE Paris, France.

## ABSTRACT

Reflectance confocal microscopy (RCM) is a powerful tool to visualize the skin layers at cellular resolution. The dermal-epidermal junction (DEJ) is a thin complex 3D structure. It appears as a low-contrasted structure in confocal en-face sections, which is difficult to recognize visually, leading to uncertainty in the classification. In this article, we propose an automated method for segmenting the DEJ with reduced uncertainty. The proposed approach relies on a 3D Conditional Random Field to model the skin biological properties and impose regularization constraints. We improve the restitution of the epidermal and dermal labels while reducing the thickness of the uncertainty area in a coherent biological way from 16.9 $\mu$m (ground-truth) to 10.3 $\mu$m.

***Index Terms***— Conditional Random Fields, Reflectance confocal microscopy, image segmentation, skin modeling

## 1. INTRODUCTION

The characterization of skin conditions is a major challenge for skin aging understanding. Skin aging is defined by a set of cumulative alterations of its various components over the years. The two major layers of the skin, the epidermis and the dermis, are separated by the dermal-epidermal junction (DEJ), which is also affected by skin aging.

The DEJ is a complex, wave-like, 3D structure. Its peaks and troughs, called dermal papillae, are due to projections of the dermis into the epidermis. With skin aging, the DEJ's appearance flattens, which could have important consequences such as lower epidermal adhesion.

Reflectance confocal microscopy (RCM) is a powerful tool for non-invasively assessing the skin architecture and cytology. RCM images provide a representation of the skin at the cellular level, with melanin and keratin working as natural contrast agents [1]. Our data consists of depth-oriented (en-face), in-vivo, confocal sections of the skin from the surface to 200 $\mu$m. Several approaches to automatize confocal image analysis have been proposed focusing on quantifiying the epidermal state [2], performing computer-aided diagnostic of skin lesions [3] and identifying the layers of human skin. In fair skin, the DEJ detection is a difficult task since it appears as an amorphous and low-contrasted structure. Its segmentation could further improve skin aging quantification.

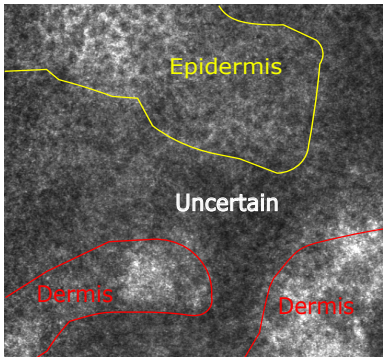



**Fig. 1**. Expert segmentation in 3 labels : Epidermis, Uncertain and Dermis. The epidermis exhibits a well defined honeycomb pattern, the dermis is textured by the collagen fiber representation. One can notice the low-contrasted appearance of the uncertain area.

In the skin, the DEJ architecture is a thin layer of cells but state-of-art methods define it as a thick transition zone due to the difficulty in precisely locating it by visual inspection (Fig. 1). Somoza et al. [4] used an unsupervised clustering method to classify whole en-face image as a single distinct layer, resulting in a good correlation between human classification and automated assessment. Kurugol et al. [5] developed a hierarchical method on small tiles to segment the DEJ as a transition zone from the epidermis to the dermis. In [6, 7], Hames et al. proposed a supervised classification method based on a one-layer per depth strategy for the labeling of their training set. They have extended their method to perform pixel-wise classification, allowing them to distinguish several skin strata in one image. Their algorithm uses a 1D linear chain Conditional Random Field and structured Support Vector Machine to model the skin structure.

In this paper, we provide a method to automatically segment the dermo-epidermal junction in 3 classes : epidermis (ED), uncertain area (U), and dermis (D) using in vivo reflectance confocal microscopy on fair skin. Our proposed ap-

proach consists of a 3D Conditional Random Field, which allows us to provide a spatial regularization on label distribution and to model skin biological properties. The CRF model predicts a pixel label according to 1) its probabilities to belong to one of the skin layers, and to 2) the labels of its neighbors. The probability of a pixel to belong to one of the three classes is estimated by a Random Forest classifier trained on relevant features. The relations between pixel neighbors mimic the skin layers behavior in 3D by imposing the transition order between en-face sections. Such a model is represented in Fig. 2. Our main result is a DEJ segmentation which respects the high-confidence expert label while reducing the thickness of the uncertain area.

## 2. DATABASE

Our database consists of 23 annotated stacks of confocal images acquired from fifteen healthy volunteers with fair skin. Volunteers were assigned to two groups: a 7-persons group aged from 18 to 25 and another 8-persons group aged from 55 to 65. Image acquisition was carried out on the cheek to further assess chronological aging. No cosmetic products nor skin treatment were allowed on the day of the acquisitions. Appropriate consent was obtained from all subjects before imaging. RCM images are acquired using a near-infrared reflectance confocal laser scanning microscope (Vivascope 1500; Lucid Inc, Rochester, NY, USA) [8]. Each image corresponds to a horizontal section with a 500 $\times$ 500 $\mu$m field of view and a resolution of 1000 $\times$ 1000 pixels. On each imaged site, stacks are acquired from the skin surface to the reticular dermis with a step of 5 $\mu$m. Visual labeling of the DEJ is not easy to perform even for experts, therefore they were asked to delineate the stacks in the 3 zones described in Sec. 1 (see Fig. 1).

## 3. CONDITIONAL RANDOM FIELDS

An image $\mathbf{y}$ consists of M pixels i $\in \mathbb{S} = [\![1, M]\!]$ with observed data $y_i$, i.e, $\mathbf{y}$ = ($y_1$, $y_2$, ..., $y_M$) organized in layers (en-face images) forming a 3D structure. We want to assign a discrete label $x_i$ to a each pixel $i$ from a given set of classes $\mathcal{C} = \{\textit{Epidermis}, \textit{Uncertain}, \textit{Dermis}\}$. The classification problem can be formulated as finding the configuration $\hat{x}$ that maximizes $p(x \mid y)$, the posterior probability of the labels given the observations.

A CRF is a model of $p(x \mid y)$with an associated graph $G = (V, E)$ where $V$ is the set of vertices representing the image pixels and $E$ the set of edges modeling the interaction between neighbors [9]. Here, $E$ is the usual 3D 6-connectivity. The CRF model is represented in Fig. 2.

We use a model with pairwise interactions defined by :

$$p(x \mid y) \propto \prod_{i \in \mathbb{S}} \varphi_i(x_i, \mathbf{y}) \times \prod_{i,j \in \mathbb{E}} \psi_{ij}(x_i, x_j, \mathbf{y}), \qquad (1)$$

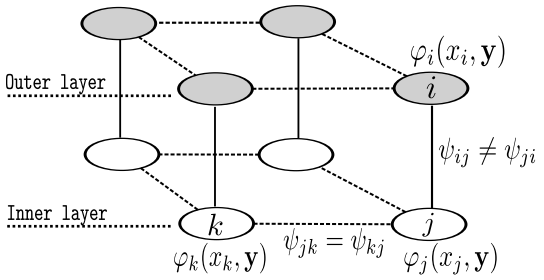



**Fig. 2**. 3D CRF modelisation. The sets of nodes in gray and in white belong to two different en-face sections. The edge potentials of each en-face sections $\psi_{jk}$ (Eq. 3) are learned for each depth. Edge potentials between en-face sections $\psi_{ij}$ (Eq.3) impose biological transition constraints.

where $\varphi_i(x_i, \mathbf{y})$ is the *node potential* linking the observations to the class label at pixel $i$ and $\psi_{ij}(x_i, x_j, \mathbf{y})$ the *interaction potentials* modeling the dependencies between the labels of two neighboring pixels $i$ and $j$.

### 3.1. Node potential

The node potential is defined as the probability of a label $x_i$ to take a value $c$ given the observed data $\mathbf{y}$ by :

$$\varphi_i(x_i, \mathbf{y}) = p(x_i = c \mid \mathbf{f_i}(\mathbf{y})) \qquad (2)$$

with $\mathbf{f_i}(\mathbf{y})$ a feature vector computed at pixel $i$ from the observed data.

In our case, each node potential is associated with the predicted class probability vector produced by a Random Forest (RF) classifier. We apply the following classical features from the literature [5, 7] : first and second order statistics, gray level co-occurrence matrix contrast, energy and variance, Gabor filter output and Laplacian variance.

We propose new features to estimate the distance of the current pixel to the DEJ. The DEJ is an amorphous structure compared to the epidermis, which appears as a honeycomb pattern, and the dermis, which contains collagen fibers. Thus, we expect low values of Laplacian variance in en-face sections around the DEJ location. For a pixel $i$ at a given en-face section $p$, we calculate the feature vector for every en-face section at its location. We define $D_{\text{DEJ}}$ as the distance to the closest minimum. The Laplacian variance of the closest minima is also added to the set of features. An example is presented in Fig.3.

### 3.2. Interaction potential

The interaction potential describes how likely $x_i$ is to take the value $c$ given the label $c'$ of one of its neighboring pixel $j$:

$$\psi_{ij}(x_i, x_j, \mathbf{y}) = p(x_i = c \mid x_j = c') \qquad (3)$$

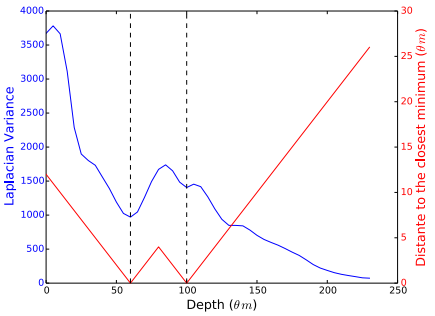



**Fig. 3**. Laplacian variance and distance to the closest minimum.

Prior information on skin structure is essential to determine efficiently the interaction potentials in our CRF model. We know that several skin layers can co-exist in a single en-face section. However, the skin layers follow a specific order from the surface to inner layers : ED $\rightarrow$ DEJ (contained in the uncertain area) $\rightarrow$ D. We define an incoherent transition as a transition not following such a specific order.

In an en-face section, edges are modeled symmetrically i.e $\psi_{ij} = \psi_{ji}$ (dashed arrows in Fig. 2). Between en-face sections, only coherent transitions are allowed, the edge potentials thus depend of their direction, i.e $\psi_{ij} \neq \psi_{ji}$ (plain arrows in Fig 2).

The interaction potentials are modeled by a 3×3 matrix representing the transition probabilities between classes. We estimate such transition probabilities from the frequency of co-occurrence of classes ($c$, $c'$) between neighboring pixels $i$ and $j$ in the ground-truth images. Co-occurrence frequencies are learned for each depth of en-face sections.

For the 3D case, constraints are added to the transition matrix according to the edge direction. Between two pixels $i$ and $j$ belonging to adjacent en-face sections, if $i$ is above $j$, $\psi_{ij}$ is equal to $H(label_i, label_j)$, the transition matrix in Table 1, and to $H(2 - label_i, 2 - label_j)$ otherwise.

| $label_i$ \ $label_j$ | Epidermis (label 0) | Uncertain (label 1) | Dermis (label 2) |
|---|---|---|---|
| Epidermis (label 0) | 0.58 | 0.38 | 0.04 |
| Uncertain (label 1) | 0 | 0.91 | 0.08 |
| Dermis (label 2) | 0 | 0 | 1 |

**Table 1**. Transition matrix $H(label_i, label_j)$ with $i$ above $j$ where values are the learned probabilities of transition between depth 80 and 85 $\mu$m. The null values ensure that incoherent transitions are impossible.

## 4. RESULTS & DISCUSSION

Our goal is to produce a precise information about the DEJ shape, i.e. to simultaneously provide an accurate restitution of the epidermal and dermal labels while reducing the uncertain area.

We want to evaluate our set of features and the regularization strategy. We consider three cases: 1) $RF_{lit}$ in which we train the RF classifier on the classical set of features from the literature, 2) $RF_{prop}$ in which we train the RF classifier on our proposed set of features, and 3) $CRF3D_{prop}$ in which we train the RF classifier on our proposed set of features followed by the CRF modelisation.

To evaluate our proposed set of features, we compare the mean accuracy of $RF_{lit}$ and $RF_{prop}$. Then, to validate our 3D model, we compare the sensitivity for ED and D, to assess the accurate restitution of those layers, and the specificity for U that penalizes the classification of the epidermal or dermal layers as uncertain area.

Each labeling is evaluated using a 10-fold cross-validation. The optimal labels for Eq. (1) were computed with a LBP algorithm using DGM Lib [10].

The mean accuracies of the RF classifications are presented in Table 2. Using features from the literature [5, 7], we achieve a recall score of 90% on the high confidence labels, epidermis and dermis, and 57% on the uncertain area.

Our proposed set of features allows us to increase the accuracy on the uncertain area classification to 66%. These results suggest that our set of features is relevant to identify the three skin labels according to the experts' visual inspection. However, the result of $RF_{prop}$ contains 7.1% of incoherent transitions motivating the introduction of spatial constraints with $CRF3D_{prop}$.

| | ED | U | D |
|---|---|---|---|
| $RF_{lit}$ | 0.89 | 0.57 | 0.97 |
| $RF_{prop}$ | 0.90 | 0.66 | 0.97 |

**Table 2**. Results for the unregularized experiments. Mean accuracy of the RF classifications of the three labels.

| | Sensitivity | | Specificity | | |
|---|---|---|---|---|---|
| | ED | D | ED | U | D |
| $RF_{prop}$ | 0.90 | 0.97 | 0.99 | 0.96 | 0.98 |
| $CRF3D_{prop}$ | 0.96 | 0.98 | 0.99 | 0.98 | 0.98 |

**Table 3**. Sensitivity and specificity of the three labeling in the unregularized vs. regularized cases.

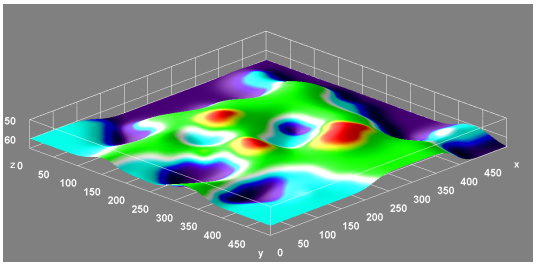
Our regularization promotes spatial consistency and forbids incoherent layer transitions. The results of the regularized CRF model, presented in Table 3, show that the epidermal and dermal sensitivity are increased. The misclassification scores of high confidence labels improve with the 3D

CRF regularization from 9% to 4% for the epidermis and 3% to 1% for the dermis.

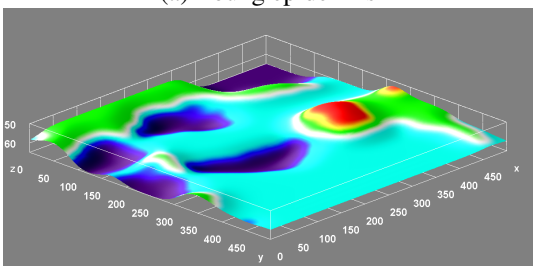
The specificity of the uncertain detection is enhanced suggesting a coherent expansion, in a biologically-coherent way, of the high-confidence labels into the uncertain area.

We also compute the average thickness of our uncertain area defined as the average distance between its two borders for every pixel. In the manual annotation, the uncertain area thickness is measured as 16.4 $\mu$m. We estimate the thickness of this area as 16.9 $\mu$m in the RF model. With our CRF 3D model, the uncertain region thickness is reduced to 10.3 $\mu$m.

3D visualizations of the DEJ segmentations, obtained with $\text{CRF3D}_{\text{prop}}$, are presented in Fig. 4.



(a) Young epidermis



(b) Aged epidermis

**Fig. 4**. Visual appearance of the lower border of the epidermal layer. Notice that the aged epidermis appears flatter than the young epidermis, as expected. Colors encode the depth.

## 5. CONCLUSION

In this article, we have proposed an improved set of features and a 3D Conditional Random Field regularization for the modeling of skin structures. We have shown that the use of a 3D CRF has a positive impact in terms of improved sensitivity and specificity for the Epidermis and Dermis labeling. We have also shown that the thickness of the uncertain region is reduced significantly. With our method, the DEJ segmentation is nearly always (more than 98% specificity) included within the ground-truth. It allows us to get closer to the true shape of the dermal epidermal junction. Our next aim will be to develop measures for the shape of the DEJ to quantify its changes within skin aging.

## 6. REFERENCES


[1] P. Calzavara-Pinton, C. Longo, M. Venturini, R. Sala, and G. Pellacani, "Reflectance confocal microscopy for in vivo skin imaging," *Photochemistry and photobiology*, vol. 84, no. 6, pp. 1421–1430, 2008.

[2] J. Robic, A. Nkengne, B. Perret, M. Couprie, and H. Talbot, "Automated quantification of the epidermal aging process using in-vivo confocal microscopy," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 1221–1224.

[3] K. Kose, C. Alessi-Fox, M. Gill, J. G Dy, D. H Brooks, and M. Rajadhyaksha, "A machine learning method for identifying morphological patterns in reflectance confocal microscopy mosaics of melanocytic skin lesions in-vivo," in *SPIE BiOS*. International Society for Optics and Photonics, 2016, pp. 968908–968908.

[4] E. Somoza, G. O. Cula, C. Correa, and J. B Hirsch, "Automatic localization of skin layers in reflectance confocal microscopy," in *International Conference Image Analysis and Recognition*. Springer, 2014, pp. 141–150.

[5] S. Kurugol, K. Kose, B. Park, J. G Dy, D. H Brooks, and M. Rajadhyaksha, "Automated delineation of dermal–epidermal junction in reflectance confocal microscopy image stacks of human skin," *Journal of Investigative Dermatology*, vol. 135, no. 3, pp. 710–717, 2015.

[6] S. C Hames, M. Ardigò, H P. Soyer, A. P Bradley, and T. W Prow, "Automated segmentation of skin strata in reflectance confocal microscopy depth stacks," *PloS one*, vol. 11, no. 4, pp. e0153208, 2016.

[7] S. C. Hames, M. Ardigo, H. P. Soyer, A. P. Bradley, and T. W. Prow, "Anatomical skin segmentation in reflectance confocal microscopy with weak labels," in *Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on*, Nov 2015, pp. 1–8.

[8] M. Rajadhyaksha, M. Grossman, D. Esterowitz, R. H Webb, and R R. Anderson, "In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast," *Journal of Investigative Dermatology*, vol. 104, no. 6, pp. 946–952, 1995.

[9] S. Kumar and M. Hebert, "Discriminative random fields," *International Journal of Computer Vision*, vol. 68, no. 2, pp. 179–201, 2006.

[10] S. Kosov, "Direct graphical models C++ library," http://research.project-10.de/dgm/, 2013.