Facilitating the Adoption of Reflectance Confocal Microscopy (RCM) in Clinical Cancer Care Practice with Machine Learning

Skin cancer is the most common type of cancer with incidence rates of more than 5 million new cases every year in the USA[1] and another million in other parts of the world. Clinical (visual) examination and dermoscopy (skin surface microscopy that reveals lesion color, textural patterns and borders) are performed, followed by biopsy and pathology for diagnosis. While visual examination and dermoscopy provide high sensitivity (~80-90%), the specificity can be lower and quite variable (~30-60%), resulting in benign-to-malignant biopsy ratios of 2 to 47 and millions of biopsies of benign skin lesions every year[2]. Several non-invasive optical imaging modalities have been developed to address this challenge[3-6]. One of them, reflectance confocal microscopy (RCM)[7,8] is at the forefront with increasing clinical utility. RCM imaging combined with dermoscopy provides 2 times higher diagnostic specificity (~70-80%) compared to dermoscopy alone, without sacrificing sensitivity (~90-100%), resulting in a reduction in the benign-to-malignant biopsy ratio by two[2,9]. In 2016, RCM imaging crossed a major milestone in the USA, receiving current procedural terminology (CPT) reimbursement codes from the Centers for Medicare and Medicaid Services. *Adoption and implementation of RCM imaging in routine clinical practice in the USA, Europe and Australia is now in progress. Dermoscopy combined with RCM imaging is being implemented to noninvasively triage skin lesions into benign versus suspicious, guiding dermatologists to rule out malignancy and biopsy, and sparing patients from biopsies of benign lesions*.

However, in order for RCM imaging to fully achieve its promise to impact widespread clinical practice, the imaging must reach beyond its current early adoption in tertiary healthcare centers and be integrated into clinical practice by dermatologists in primary and secondary care settings. But RCM images are very different in appearance from standard pathology. The images are in *en face* (parallel to skin surface) orientation and appear with grayscale contrast that is quite different than the conventional purple-and-pink stained patterns seen in vertically sectioned (perpendicular to skin surface) pathology. Consequently, the cellular morphologic patterns in RCM images are challenging for dermatologists to read and interpret and requires specialized training. Thus, training and guiding novice dermatologists to read RCM images with a standardized and consistent approach is vital to achieving this goal. Just as in traditional pathology, diagnostic examination of RCM images today relies solely on visual reading and thus diagnostic accuracy greatly depends on the level of training and experience[3]. *Development of appropriate machine learning methods is essential to provide powerful quantitative analysis tools and a training and education platform for dermatologists to accelerate widespread adoption and integration.*

With this objective in mind, we report here on the design and testing of 2 classification and segmentation algorithms for detecting benign and malignant cellular morphologic patterns in RCM images of melanocytic skin lesions (benign moles and malignant melanoma). A particular challenge for machine learning in this field is the lack of datasets that are considered large by modern standards, and the cost of acquiring labeled training data. Thus, our methods were developed to work with, and be robust to, relatively small amounts of sparsely/partially labelled training data. In clinical practice, depth-stacks of images with small fields-of-view (0.5 by 0.5 mm) are acquired to probe skin to depths of ~150-200 um, followed by the capture of *en face* mosaics of images that display up to 12 mm x 12 mm, to enable examination of large fields of view, as needed by dermatologists. We have developed and tested two distinct algorithms to address the analytical needs of each of these imaging formats.

We first developed a recurrent neural network (RNN) to identify the three skin strata (epidermis, dermal-epidermal junction (DEJ), and dermis) in image stacks. We first train an Inception-v3 model[10] for classification of each image independently. We use this initial model as a feature extractor and train an RNN model with an attention mechanism to utilize the correlation between consecutive images, since the labels in a stack of RCM images collected at increasing depths should follow a specific sequence (epidermis, DEJ, dermis). The attention mechanism helps to impose this condition on the image labels by dynamically adjusting the influence of each neighbor in the final classification. Our experiments show that the RNN-with-Attention model successfully produces image labels that are consistent with the known layered strata in skin. Testing our model on an expert labeled dataset of 504 RCM stacks, we achieved 88.17% image-wise classification accuracy to classify epidermis versus DEJ versus dermis.

Since most skin cancers originate in and spread from the DEJ, diagnosis is based on examining cellular morphologic patterns at this strata. Thus, we concentrated our analysis of mosaics on classifying cellular morphologic patterns at the DEJ[11]. The patterns are encountered at multiple scales and varying shapes. Inspired by the practice of dermatologists in starting their examination with low magnification (low resolution, large field of view, looking at the entire mosaic) followed by closer inspection of suspicious areas with higher magnification (higher resolution, smaller fields of view, looking at sub-mosaics), we developed a novel multiscale convolutional neural network, called "Multiscale Encoder Decoder Network (MED-Net)", for semantic segmentation of six clinically significant cellular morphological patterns. *The novelty in our approach includes our modeling of textural patterns at multiple scales, while using the segmentation outcome at each scale as a prior estimate for the subsequent segmentation at the next scale. The translational novelty is its application to noninvasive microscopic images of living skin directly on humans in vivo (whereas deep learning approaches are currently being applied mostly to traditional pathology slides.)*

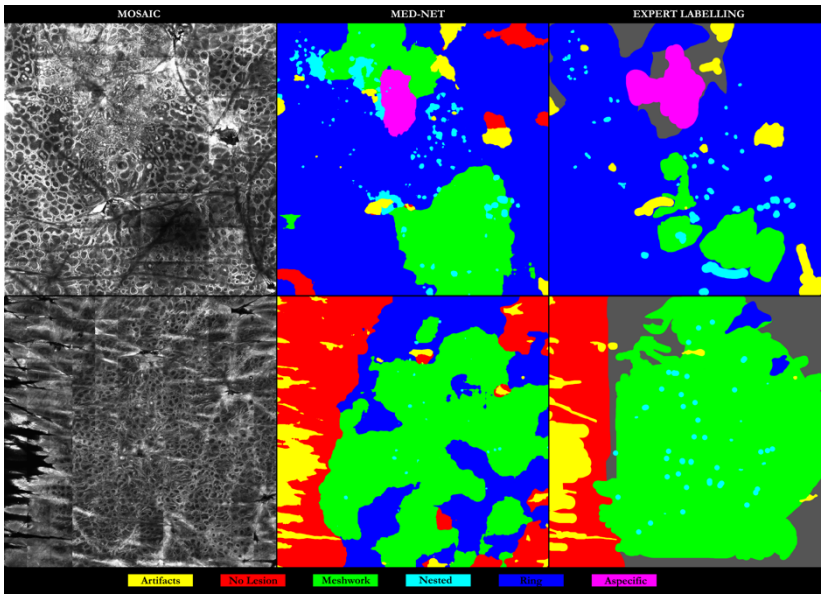

Figure 1: Sample semantic segmentation results for in-vivo RCM mosaics of melanocytic lesions for 2 cases (column - 1) Sample semantic segmentation results from MED-Net (column 2) trained and tested over a dataset of 117 mosaics with 5-fold cross-validation. The mosaics are consensus labeled into 6 different classes by two expert readers (column - 3).

Our model features sub-networks for each scale that explicitly cooperate with each other, utilizing the correlation between scales. To prevent the cascaded subnetworks in the model from becoming too deep and hard to train, we also provide direct feedback at intermediate scales (i.e. intermediate sub-networks) by comparing their outcomes against ground truth segmentations at each scale and backpropagating the intermediate errors. This gives the algorithm direct access to the early layer coefficients of the network and allows efficient propagation of the error gradients. MED-Net was tested on 117 RCM mosaics of melanocytic lesions, collected at two clinical centers, one in the US and the other in Italy, that had been labeled by consensus by two expert clinical image readers. (We note, as mentioned earlier, that given the relatively early stage of clinical adoption of RCM, and the laboriousness of manual expert labeling, this is both an unprecedentedly large labeled dataset for this application on living skin in vivo and yet a challengingly small one for the training of deep neural networks.) The network achieved pixel-wise segmentation of these mosaics across the 6 classes with mean sensitivity and specificity of 70% and 94% respectively, with a ~0.7 mean Dice coefficient, over 5-fold cross-validation. The next translational step toward entering the clinic is now in progress: the MED-NET algorithm is being integrated into one of our commercial scopes and we will start testing on RCM mosaics of melanocytic skin lesions on patients.

We also note that the multiscale process for reading cellular and morphologic textural patterns seen in RCM images of melanocytic skin lesions will be similar to that for reading patterns seen in other tissues and conditions (for example, non-melanocytic skin lesions, skin pre-cancers, oral pre-cancers and cancers, benign and inflammatory conditions) in image mosaics from RCM and also from other emerging optical microscopic imaging modalities (e.g. optical coherence tomography, multiphoton, photoacoustic - and optical coherence microscopy). In this sense, with MED-NET we have introduced a flexible model that can be readily trained to serve as a segmentation tool for other cellular imaging modalities, spur the adoption of *in vivo* optical microscopy in clinical practice, which will help facilitate noninvasive diagnoses and, ultimately, spare patients from millions of biopsies of benign conditions in diverse settings.

REFERENCES

[1] https://cancerstatisticscenter.cancer.org/
[2] Rajadhyaksha M., Marghoob A., Rossi A., Halpern AC, and Nehal KS. Reflectance confocal microscopy of skin in vivo: From bench to bedside. Lasers in surgery and medicine, 49(1):7–19, 2017.
[3] March J, Hand M, Grossman D. Practical application of new technologies for melanoma diagnosis: Part I. Non-Invasive Approaches. J Am Acad Dermatol 2015; 72:929-41:941–942.19.
[4] Menge TD, Pellacani G. Advances in noninvasive imaging of melanoma. Semin Cutan Med Surg 35:18–24 2016
[5] Hibler BP, Qi Q, Rossi AM. Current state of imaging in dermatology. Semin Cutan Med Surg 35:2–8.21, 2016.
[6] Giavedoni P, Puig S, Carrera C. Noninvasive imaging for non-melanoma skin cancer. Semin Cutan Med 3:31–41, Surg 2016
[7] Guitera P, Pellacani G, Longo C, Seidenari S, Avramidis M, Menzies SW. In vivo reflectance confocal microscopy enhances secondary evaluation of melanocytic lesions. J Invest Dermatol, 129:131–138, 2009.
[8] Guitera P, Menzies SW, Longo C, Cesinaro AM, Scolyer RA, Pellacani G. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: Analysis of 710 consecutive clinically equivocal cases. J Invest Dermatol, 32:2386–2394, 2012.
[9] Alarcon I, Carrera C, Palou J, Alos L, Malvehy J, Puig S.Impact of in vivo reflectance confocal microscopy on the number needed to treat melanoma in doubtful lesions. Br J Dermatol 170:802–808.3, 2014.
[10] Szegedy C., et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[11] Scope A., Guitera P. and Pellacani G., Reflectance Confocal Microscopy of Cutaneous Tumors, 2nd Ed. Boca Raton, CRC Press, 2017, ch. RCM Diagnosis of Melanocytic neoplasms: Terminology, algorithms and their accuracy and clinical integration., pp. 168–186.