Dual Stream Network with Selective Optimization for Skin Disease Recognition in Consumer Grade Images

Krishnam Gupta

Jaiprasad

Dr. Monu Krishnan

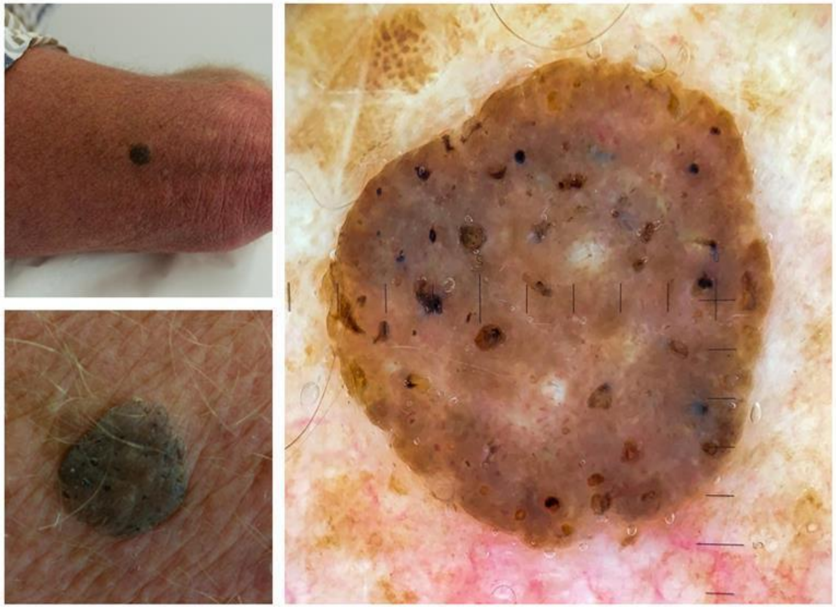
Ajit Narayanan

Nikhil S. Narayan

25th International Conference on Pattern Recognition - ICPR 2020

mfine

Consumer Grade Images in Dermatology



Image[1] of same skin lesion : on a phone (left) vs on a dermatoscope (right)

Images acquired by users on a non-clinical mobile imaging device, usually smartphones, are popularly referred to as consumer grade images.

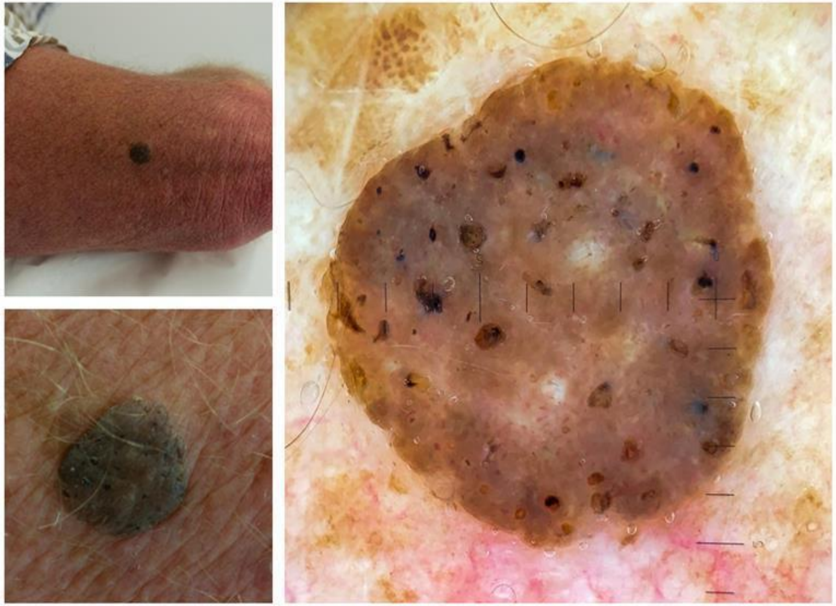
Dermatology is one of the top specialities being opted for in teleconsultations.

This is opening up new data funnels, giving rise to more opportunities for AI in consumer-grade images research.

In our platform, teleconsultations typically involve the patients describing their condition and uploading one or more images of the skin condition of interest acquired using mobile phones. This information is presented to the specialist doctor along with AI generated differential diagnosis

[1] https://www.guidelinesinpractice.co.uk/skin-and-wound-care/top-tips-getting-started-with-dermoscopy/455038.article

Consumer Grade Images in Dermatology



Image[1] of same skin lesion : on a phone (left) vs on a dermatoscope (right)

Automatically localizing, classifying and quantizing the skin conditions on consumer grade images is challenging due to:

(1) the boundaries of skin conditions being diffuse, irregular and fuzzy

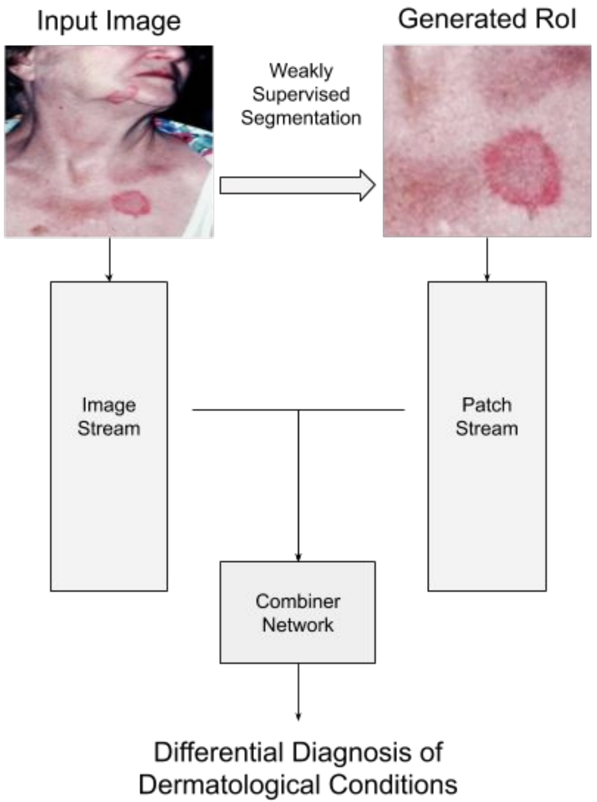
(2) the contrast being relatively low between the lesion and the surrounding skin

(3) the occurrence of fragmentation or variegated colouring inside the lesion

(4) The skin condition of interest usually occupying a relatively small area in the image

[1] https://www.guidelinesinpractice.co.uk/skin-and-wound-care/top-tips-getting-started-with-dermoscopy/455038.article

Methodology

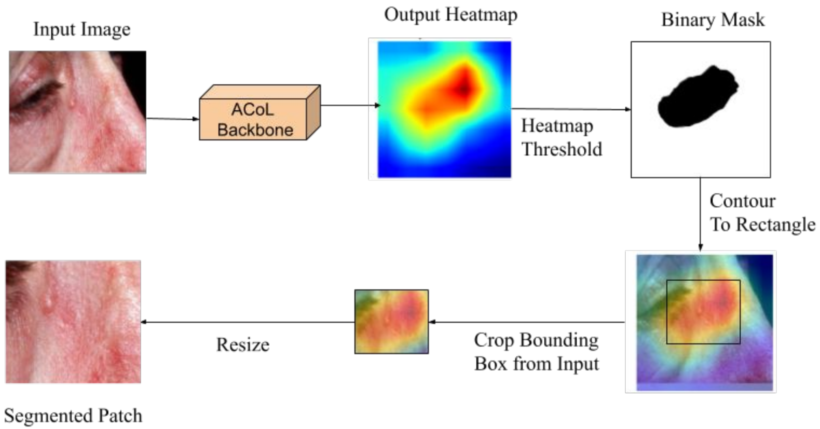


In our work, a novel dual stream deep network is proposed which is trained in a multi-phase manner that employs both global and local cues about the skin condition of interest to perform skin condition recognition with implicit segmentation.

We do not have Ground truth segmentation maps for the skin condition. So, we employ weakly supervised segmentation techniques and crop out the Region of Interest from the full resolution image.

While the first stream of the proposed dual stream network learns global cues from the input image, the second stream learns the local cues from the cropped-out high resolution segmentation masks.

Generating RoI-Cropped Regions



Adversarial Complementary Learning(ACoL)[1] is first used to generate a heatmap using classification labels.

This heatmap is thresholded into a binary mask. The largest contour is retained, rest are discarded. The value of threshold is chosen empirically.

A rectangular bounding box is then fitted into this contour. This bounding box is then cropped out from the full resolution input image.

Thus we get the prominent region from the image in high resolution, which now acts as second input to the dual stream neural network.

[1] Zhang, Xiaolin, et al. "Adversarial complementary learning for weakly supervised object localization." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

Dual Stream Network

$D(w, x) = S_c(w_c, (S_i(w_i, X_{image}), S_p(w_p, X_{patch})))$

$L_t = L_c + \beta L_s$

Selective Optimization

mfine

The dual stream network is optimized in an alternate manner over multiple phases.

In first phase of learning, the network is optimised using Stream loss L_s which helps it to learn independent features from stream. When training loss stops to decrease, the loss is switched to Total loss L_t. In this second phase of learning, the network now learns combined features from both the streams with the help of combiner sub-network S_c, as L_t contains combiner sub-network's loss. When training loss stops decreasing in second phase, we again switch back to optimizing L_s.

Thus the architecture is designed to keep on alternating between L_t & L_s until the network stops to learn, and training loss does not decrease further.

Dataset	**Stream**	**Opt.**	**Accuracy**	**F1**
SD-198	Image	Full	66.5 ± 1.4	63.9 ± 1.6
SD-198	Patch	Full	65.2 ± 1.4	63.1 ± 1.7
SD-198	Dual	Full	67.2 ± 1.3	66.5 ± 1.5
SD-198	**Dual**	**Alternate**	71.4 ± 1.1	70.9 ± 1.2
SkiD-75	Image	Full	46.6 ± 1.9	44.9 ± 1.9
SkiD-75	Patch	Full	43.7 ± 2.1	42.5 ± 1.9
SkiD-75	Dual	Full	47.2 ± 1.9	46.6 ± 1.8
SkiD-75	**Dual**	**Alternate**	51.8 ± 1.7	50.9 ± 1.8

Results

mfine

(a) Angular Cheilitis (b) Stasis Edema (c) Basal Cell Carcinoma
(d) Skin Tag (e) Drug Eruption (f) Keratosis Pilaris

Conclusion & Future Work

mfine

- In this paper, we propose a dual stream network & an optimization strategy for the same to jointly learn from global and local features for better classification of skin diseases.
- Future work on dual stream networks may focus on simultaneously segmenting the RoI patch from the image stream, and jointly learning features from it and the image stream in an end-to-end manner.
- We also help in advancing future research in consumer grade dermatological images by releasing a new dataset SkiD-75 of more than 12,000 images containing 75 dermatological conditions.
- The results indicate that proposed network and optimization strategy are effective and demonstrate its advantage over the current state-of-the-art methods.