Detection and Quantification of Hand Eczema by Visible Spectrum Skin Pattern Analysis

Christoph Suter[1] and **Alexander Navarini**[2] and **Marc Pouly**[1]
and **Ruedi Arnold**[1] and **Florian S. Gutzwiller**[3] and **René Meier**[1] and **Thomas Koller**[1]

Abstract. Hand eczema is a frequent dermatosis with severe health and financial consequences to patients and society. It follows a chronic course and persists up to 15 years after onset. Early detection of an exacerbation followed by the application of specific drugs for a few days can considerably reduce disease activity and avoid temporary disability. However, dermatitis patients usually rely on their own perception in assessing their skin condition and therefore often miss the time point for effective treatment. In this paper we present a prototype-based feasibility study of automated detection and quantification of hand eczema using texton-based imaging and machine-learning techniques.

1 INTRODUCTION

Hand eczema is one of the most frequent dermatoses, affecting up to 14% of the population [4], and more than 20% have moderate-to-severe disease requiring intensive treatment. It can occur in all ages, commonly follows a chronic course and persists 10-15 years after onset [2]. Hand eczema manifests by redness of the skin, scaling, fissures, pain, itching, papules and blisters, as well as clear discharge and sometimes bacterial infection with subsequent pustules. Although the area it usually affects is small, it still can have grave consequences for the patient, leading to impairment or disability to work in many professions. Costs caused by occupational hand dermatitis (OCD) consist of direct medical costs and indirect costs associated with lost workdays and loss of productivity. Data on the economic impact of OCD reveals that in the US alone, annual direct costs for physician, clinic services and prescription drugs were $1.6 billion, with indirect costs of approximately $566 million for lost productivity [2]. If hand eczema cannot be prevented, e.g. by changing profession, early detection of exacerbation is key to an effective treatment. While moisturizers and gloves are often helpful, more severe or recalcitrant cases require the application of topical steroids or other drugs for a few days to reduce disease activity. In this process, dermatitis patients rely on their own perception whether their skin condition is stable or worsening. Consequently, they often miss the time point when early treatment of a relapse would be efficacious and successfully prevent full-fledged eczema.

Our ultimate project goal consists of an objective detection and quantification of eczematous changes by a mobile device application in order to allow patients to react early on worsening disease states and physicians to use tele-dermatology on a novel level. Here we summarize first results obtained from a prototype-based feasibility study under idealized conditions using imaging and machine learning techniques applied to high-resolution pictures on desktop computers.

2 FEATURE EXTRACTION

Image processing methods based on the concept of *textons* [1] are a standard approach for texture classification and have been proven successful in applications targeted to other dermatological defects with different appearance such as detection of carcinoma or cases of urticarial vasculitis [6]. In particular, Varma and Zisserman [5] claim superior results using a single imaging method with rotationally invariant filter banks (*VZ algorithm*). This builds the foundation of our prototype implementation.

The VZ algorithm consists of a learning and classification phase. In the learning phase, texture characteristics (subsequently called *textons*) are identified, learned and assembled in a so-called *texton dictionary*. This is achieved by first applying a filter bank with image filters of different scale and orientation to texture images and then taking the maxima over the orientation filters to achieve rotation invariance. We therefore obtain for every pixel a vector of responses and apply the *k-means* algorithm to partition these vectors into k clusters. The centers of these clusters form the textons. Repeating the process for n images produces a texton dictionary of $n \cdot k$ textons. We next use the created dictionary together with training images annotated by a dermatologist to learn models of healthy and eczema skin. However, because we aim to classify local skin regions rather than the entire image, we first divide each training image into fixed-size sub-images called *patches*. For each patch we again apply the filter bank, create a vector of responses for every pixel and determine the closest texton with respect to Euclidean distance. This gives us a histogram of the closest textons for this patch. Because the patches are taken from annotated images, we can consider each histogram as a representative for either eczema or healthy skin. These labeled histograms provide the models for the subsequent classification process.

During classification we divide the image into overlapping patches, apply the filter bank and create the texton histogram. A *chi-square-test* then finds the model that correlates best with the current histogram. The label of this model becomes the classification at the center pixel of the patch.

3 EVALUATION

The above classification applies to the center pixels of fixed-size image patches. Centering a patch at every image pixel is computationally too expense, so we use a predefined gap between patch cen-

ters and interpolate the remaining values. This produces a *prediction value* in $[0, 1]$ for every pixel to belong to an eczema region. Finally, a threshold is determined by *ROC curve analysis* as pointed out below in order to decide on each pixel's label.

Figure 1. Labeled training image with hand photograph and eczema mask.

For training and evaluation of our prototype we have access to photographs of hands with eczema and healthy skin, labeled by the same dermatologist who created an overlay bit-mask for every image to point out the eczema regions, see Figure 1. We feed one such image to the already trained algorithm, obtain a prediction value for every pixel and apply a certain threshold for the final classification. The last step is repeated for increasing threshold values from 0 to 1, and the *true-positive rate* (TPR) and *false-positive rate* (FPR) are calculated with respect to the expert labeling. Contrasting TPR and FPR for the different threshold values defines a *ROC curve* for every picture, see Figure 2. We can now determine the optimum threshold by drawing a $45°$-tangent to the curve. This reflects the assumption that a high TPR is equally important as a low FPR, which, in practice, is not always the case.
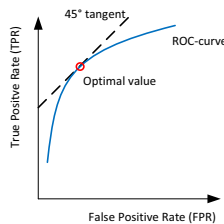

Figure 2. ROC curve with optimum threshold value.

As training data is limited, our evaluation is based on *k-fold cross validation*. All available (labeled) pictures are randomly partitioned into k folds; one fold is set aside for testing and the other $k - 1$ folds are used as training data in the classification algorithm. This process is repeated k times so that the algorithm is tested against every fold exactly once, and the final result is obtained by taking the average over the k runs. Calculation of the *Youden index* of the optimum threshold allows comparison of the algorithm's performance under different configurations.

There are many parameters and thus configurations with a direct influence on performance. Our experiments showed best results with square patches of edge length 40 pixels. A patch is labeled as eczema skin if at least 95% of its pixels carry this label. Concerning the number of models learned for classification, we could observe that increasing the amount of models also improves classification, but naturally for the price of higher computational costs. The two criteria seem to be best balanced taking 5 textons per training image and 10 training images, i.e. 50 textons.

4 RESULTS & OUTLOOK

In this setting with 18 labeled pictures and fold size $k = 9$ the algorithm showed an average accuracy of 0.78 with precision 0.35.

However, it also turned out that these results do not correctly reflect reality. During assessment of hand eczema, a practicing dermatologist naturally focuses on the most distinct parts of an eczema. Consequently, many smaller eczema patches have not been labeled as they are clinically less relevant. When the classification results were presented to the expert, many false-positive results were actually approved to correctly indicate eczema skin. Such an example is shown on the left-hand picture in Figure 3. Another difficulty in measuring the algorithm's performance is that the evaluation currently works on a per-pixel basis, although labeling of training data can never be done with such precision. Also, many false-positive results are due to hairs on the patient's wrist as a consequence of the line and edge filters used, see right-hand picture in Figure 3. Distinguishing hairs from eczema fissures is one of the main challenges.
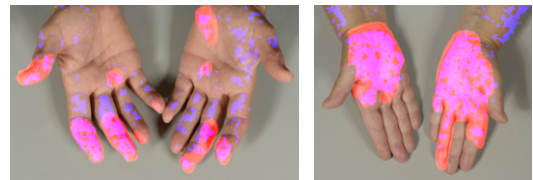

Figure 3. Result of classification including expert labeling (orange color) and automated classification (blue color).

Finally, the current approach is based on grayscale pictures and therefore leaves color information (e.g. skin redness) out of the classification process. Using a color-aware extension of the VZ-algorithm such as *VZ-dipoles* [3] failed to improve the results. We will conduct experiments with pictures that include a color chart for normalization and aim to develop new extensions to the VZ-algorithm that uses color information more locally.

With this first prototype we could show that reliable, texton-based skin eczema detection on high-resolution pictures is possible. One of the main challenges identified is to find an image labeling method that is efficient for dermatologists, precise enough for evaluation and complete in the sense that it also includes small eczema patches. We will further investigate whether labelling of healthy skin in addition to eczema skin can improve classification.

ACKNOWLEDGEMENTS

REFERENCES

[1] Julesz B., 'Textons, the elements of texture perception, and their interactions', *Nature*, **290**(5802), 91–97, (1981).

[2] Meding B., Wrangsjo K., and Jarvholm B., 'Fifteen-year follow-up of hand eczema: persistence and consequences', *The British journal of dermatology 2005*, **152**, 975–980, (2005).

[3] Burghouts G.J. and Geusebroek J.-M., 'Color textons for texture recognition', in *BMVC*, eds., Mike J. Chantler, Robert B. Fisher, and Emanuele Trucco, pp. 1099–1108. British Machine Vision Association, (2006).

[4] Hald M., Berg ND., Elberling J., and Johansen JD., 'Medical consultations in relation to severity of hand eczema in the general population', *The British journal of dermatology 2008*, **158**, 773–777, (2008).

[5] Varma M. and Zisserman A., 'A statistical approach to texture classification from single images', *Int. J. Comput. Vision*, **62**(1-2), 61–81, (2005).

[6] Cula O.G., Dana K.J., Murphy F.P., and Rao B.K., 'Bidirectional imaging and modeling of skin texture', *IEEE Transactions on Biomedical Engineering*, **51**(12), 2148–2159, (2004).