Image Content-Based Navigation of Skin Conditions
Jeremy Kawahara and Ghassan Hamarneh
Medical Image Analysis Lab, Department of Computing Science, Simon Fraser University
In the 23rd World Congress of Dermatology - Vancouver 2015
Abstract

Background: Resorting to the web for diagnosing conditions prior to contacting a doctor is becoming increasingly common. For skin disorders in particular, people can examine their lesions and attempt to find conditions that exhibit similar appearance using publicly available online datasets of skin lesion images with their corresponding diagnoses (class labels). While there is an abundance of online labeled images, for an untrained user, the amount of possible diagnoses and images to browse through is overwhelming.

Objective: Our objective is to help a person with an unknown skin condition find similar images with known diagnoses by leveraging a skin lesion image "selfie". To this end, our goal is an intuitive visualization that captures the degree of similarity of the subject's novel image with lesion images where the correct diagnosis is known - essentially a content-based search engine for skin lesions images.

Methods: To realize this objective, we follow these steps: 1) Given a dataset of skin lesion images with known diagnoses, we crop a lesion area. 2) Borrowing computer vision techniques, we compute features capturing colour and texture image statistics. 3) Given a new user provided image with an unknown diagnosis, we compute the same image features. 4) Applying machine learning concepts, we compare the similarity between the user's unknown image and the known lesions, and retrieve similar images with known diagnoses. 5) The user is shown their unknown lesion image in the context of other similar lesions images using a network graph visualization where similar known images with their corresponding diagnosis are displayed closer in proximity to the user's image. This facilitates a more efficient search in the space of skin lesion images.

Results: We built a prototype system using 1050 publicly available lesion images with known diagnoses: acne, atopic dermatitis, basal cell carcinoma, psoriasis and dermatophytosis. To test our system, we chose 50 images from each disorder (250 total) and for each produced a network visualization. A baseline method that randomly selects and display images (from the remaining 800) has a 20% chance (1 out of 5) of picking an image with the same skin condition as the test image. Using our proposed method to select a single image improves the accuracy substantially: more than doubling the chance of displaying an image with the correct class (from 20% to 47%).

Limitations: Our dataset is currently limited to only five disorders and many more are needed for a practical system. This requires gathering a much larger labeled dataset.