CALIFORNIA STATE SCIENCE FAIR
2017 PROJECT SUMMARY

Name(s)
Akhil Arun; Sahana Srinivasan

Project Number
S1502

Project Title
A Machine Learning Based Approach to Skin Lesion Segmentation Using Superpixels

Abstract

Objectives/Goals
Lesion segmentation plays an important role in the early identification and treatment of melanoma, a deadly skin cancer. Our objective was to develop an end-to-end methodology for segmenting skin lesions that was independent of dermoscopic image resolution and size in order to allow for increased efficiency and greater flexibility in implementation.

Methods/Materials
We employed the SLIC pre-processing algorithm to split images into smaller, spatially coherent areas called superpixels, and we extracted two feature sets that both included texture properties and distance metrics, with one set also containing RGB values and the other a color histogram. Random forest classifiers, dense neural networks, and two distinct convolutional neural networks were designed and tested to find their optimal configurations, after which we employed a post-processing, hole-filler algorithm to improve segmentation predictions.

Results
Our color histogram-based random forest achieved the highest accuracy of 89.83%, and our color histogram-based dense neural network obtained an accuracy of 89.01% but represents a more efficient methodology. Both of these superpixel-based segmentation techniques demonstrated comparable accuracy levels to prior studies done on a pixel level while processing 145 to 4,070 fewer pixels, making the process significantly more efficient and allowing it to operate on a wide array of images.

Conclusions/Discussion
We propose a pipeline that includes superpixels and our color histogram-based feature set, a random forest for machines with sufficient computing power or our dense neural network for those without, and the implementation of our post-processing algorithm to achieve efficient and high-accuracy segmentation of dermoscopic images regardless of resolution. Our results indicate that segmenting dermoscopic images using a superpixel-based approach can perform comparably to machine learning techniques on a pixel basis, even at the faster speed and that superpixels can potentially be used in a wide variety of medical image analyses to increase efficiency and flexibility while maintaining accuracy.

Summary Statement
We developed a computational model for segmenting skin lesions that is independent of dermoscopic image resolution and size, achieving a 90% accuracy comparable to that of previous studies while increasing efficiency and flexibility.

Help Received
Our fathers, software engineers, helped us set up our PC and helped in the initial selection of the programming language.