Automated Bias Reduction in Deep Learning Based Melanoma Diagnosis Using a Semi-Supervised Algorithm

Das, Sauman (School: Thomas Jefferson High School for Science and Technology)

Melanoma is one of the most fatal forms of skin cancer and is often very difficult to differentiate from other benign skin lesions. However, if detected at its early stages, it can almost always be cured. Researchers and data scientists have studied this disease in-depth with the help of large datasets containing high-quality dermoscopic images, such as those assembled by the International Skin Imaging Collaboration (ISIC). However, these images often lack diversity and over-represent patients with very common skin features such as light skin and having no visible body hair. In this study, we introduce a novel architecture called LatentNet which automatically detects over-represented features and reduces their weights during training. We tested our model on four distinct categories - three skin color levels corresponding to Type I, II, and III on the Fitzpatrick Scale, and images containing visible hair. Then, we compared the accuracy against the conventional Deep Convolutional Neural Network (DCNN) model trained using the standard approach (i.e. without detecting over-represented features) and containing the same hyperparameters as the LatentNet. The LatentNet showed significant performance improvement over the standard DCNN model with accuracy of 89.52%, 79.05%, 64.31%, and 64.35% compared to the DCNN accuracy of 90.41%, 70.82%, 45.28%, 56.52% in the corresponding categories, respectively. Differences in the average performance between the models were statistically significant ($p<0.05$), suggesting that the proposed model successfully reduced bias amongst the tested categories. The LatentNet is the first architecture that addresses racial bias (and other sources of bias) in deep-learning based Melanoma diagnosis.

Awards Won: