Design, implementation and evaluation of deep learning prototype to classify non-pigmented malignant skin cancer from dermatoscopical images

Popular science summary of Master's thesis in Mathematics for Engineers,
Lund University 2022 by María del Pilar Aguilera Manzanera

This master thesis investigates the possibility of using machine learning to classify non-pigmented skin cancer lesions on patients in clinics. The most common non-pigmented types of skin cancer are Basal Cell Carcinoma, Squamous Cell Carcinoma, the last one possibly starting from a lesion called Actinic Keratosis. In Sweden, the number of cases is estimated to grow in the next 20 years.

To contain and possibly reverse these trends, more primary and secondary prevention initiatives are required. The Skåne University Hospital in Lund in collaboration with Lund University has started a project to create an algorithm capable of classifying between the three aforementioned types of lesions.

To find a solution using machine learning, a present problem is the lack of sufficient data which in this project was solved by studying the different sources of data available online that could be used. The data sets chosen were based on the availability of the lesions studied in the data set and the proximity of the ethnicity of the skin to the fair skin of Swedes. Furthermore, during 3 weeks more data was collected at Skåne University Hospital in Lund, and this was used for further network testing.

Multiple networks were trained with the different data sets and compared in terms of performance. With the best network of each data set, we tested the versatility of it to be used in a different group of images not used during training. From this study, we obtain poor versatility results that could be explained by the fact that the images from one data set to another varied in terms of which device was used to capture the image and the different skin types. Additionally, we studied the parts of the image in which the network was focussing when making a prediction, to give more feedback to the dermatologist. It was found that the model does not look directly at the lesion but rather looks at the surrounding skin.

To conclude, the results of the networks obtained are promising and give hope for a future solution with good prediction capabilities that could be deployed in the clinics to help dermatologists reach a diagnosis.