

Abstract

Karim E. Zaki

Machine Learning Model for Melanoma Detection

Millions of people die because of skin cancers, and the number of patients is increasing around the world. Skin cancer has a lot of types and shapes. Melanoma is the most dangerous form of skin cancer. According to the American Cancer Society, melanoma has many risk factors. One of these factors is having a non-cancerous lesion. As we are living in the era of artificial intelligence, an intelligent system is needed to be developed for helping dermatologists in their job, and data science engineers are needed to develop systems that can distinguish between the malignant and benign lesions. The International Skin Imaging Collaboration (ISIC) website is a great source of data that contains over 10,000 dermoscopic images with different diagnoses and available online for research. So, it is a good source for collecting the needed data. Lesion diagnosis became one of the important research topics in the field of intelligent computing and especially in the area of computer vision. Many machine learning and deep learning models were proposed for lesion classification. The purpose of this research is to contribute with a multiclass machine learning model that can be used as a computer-aided design for melanoma detection. This research methodology depends on developing a non-costly machine learning model to solve the problem in a small-time without using high-performance machines. The model works on classifying three types of lesions melanoma as a type of malignant lesion, nevus as benign, and seborrheic keratosis which is diagnosed as a benign lesion, but it has some malignant features, and that makes the classification more complex. The previous deep learning research that worked on classifying those lesion types scored accuracy of 83% after eleven days of training. Because the proposed model is a machine learning model, all the phases took three hours only. The proposed was model applied on a large-scale dataset rather than the previous models and scored 67.4% accuracy on a high noisy data after executing many experiments with different classifiers. Finally, future work was proposed for enhancing the results. This future work is a suggestion of using different models such as ensemble model, and deep learning model.