

Abstract

This projects, focuses on models which can identify, types of skin cancer. Making a correct identification, is extremely important, since skin cancer is one of the most common types of cancer and is highly treatable if detected early. So, the aim of the project is to find models which can perform image classification, for skin cancer images. Finding accurate models is key as they can lead to faster diagnosis.

The method followed, was to train three models with grayscale and colored images and assess their performance. The models chosen, are the CNN, SVM, SVM + LDA, and RF + LDA. All were trained using 28x28 images. For evaluation, the factors considered were the testing accuracy, the computational time, and the advantages/disadvantages of each model.

Results showed that the model with the lowest computational time was the RF + LDA model on colored images, which took approximately 112s to train and gave a testing accuracy of around 92%. The SVM model fitted on grayscale images had the best accuracy of 98%, but the highest computational time (673s), while the SVM + LDA on colored images had an accuracy of 97% and a computational time of 578s. Finally, the CNN for grayscale and colored images had 75% and 95% accuracy and 178s and 383s training time respectively.

The choice of the models, and the combination of techniques have been used before for image classification. However, for this dataset, the most common model used for classification is CNN. This project investigates if there are other models that can be more accurate and efficient.

Key images

Model	Testing Accuracy	Time (seconds)
CNN (grayscale)	75%	178.7 s
CNN (RGB)	95%	383.88 s
SVM (grayscale/sample)	98%	673.24 s
SVM + LDA(RGB)	97%	578 s
RF + LDA(RGB)	92%	112 s