

The Effectiveness of Data Augmentation for Melanoma Skin Cancer Prediction Using Convolutional Neural Networks

ABSTRACT

Melanoma skin cancer has been a serious threat due to its high fatality. For this reason, early detection and treatments are given more attention as countermeasures. In recent years, skin cancer detection has been utilizing artificial intelligence techniques, specifically deep convolutional neural network. However, the performance of the convolutional neural network is highly vulnerable to different data constraints, such as the quality and quantity of the data. Therefore, this study explores the synthetization of training data using different data augmentation methods. The work presented in this paper utilizes four different categories of data augmentation methods, which include geometrical transformation, noise addition, colour transformation, and image mix. Multiple layers data augmentation approach is also explored. Dataset expansion strategies and optimized dataset expansion scale are determined to improve the performance of the skin cancer classification. The core findings in our study revealed that single-layer augmentation has better performance than multiple layers augmentation approaches, where region of interest (ROI) image mix has the highest effectiveness compared to other methods. In addition, the best dataset expansion strategy is random ROI image mix. Finally, the optimized dataset expansion is determined at 300%, which yielded the best overall test accuracy at 82.9%, 4.6% improvement compared to unprocessed raw dataset.