ABSTRACT

Skin lesions are a widely known problem caused due to unhealthy lifestyle, skin reactions, prolonged exposure to direct sunlight, hereditary, etc. But such diseases are often ignored due to scarcity of dermatologists and inaccessibility to pathology lab for diagnostics. These diseases remain undiagnosed with the naked eye. To examine such disease dermatologists uses technique called Dermatoscopy which uses surface microscopy to examine skin surface. Dermatoscopy outputs closely rendered images of skin epidermis layer for diagnosis. With improved mathematical models using neural networks, LSTMS, these skin lesions can be classified with high accuracy and reach the sections of population that don't have access to pathology lab facility. Now, Convolutional neural network (CNN) models have been widely used for skin disease diagnosis and many of them are able to diagnose comparable or even superior to those of dermatologists. But still there are certain limitations for CNN implementation like availability of less diverse and bias publicly available datasets towards melanomas.

This paper devises a novel solution for Classification of pigmented skin lesions using image augmentation and CNN technique to handle such imbalances in dataset with high accuracy and low losses. In the paper first, the dataset is pre-processed to remove the noise, impulse features by apply bilateral filtering and gray scaling. Secondly, applying image augmentation techniques like rescale, shearing, shifting (horizontal, vertical, right, left), flipping, padding and filling to create more diverse features in the existing dataset for better accuracy and prevention of overfitting. Third, CNNs are trained on processed and augmented datasets to verify whether the model is able to generate desirable output with moderate complexity. Lastly, Dropout layers are added to reduce overfitting and two optimization algorithms i.e., EarlyStopping and ReduceOnPlateau are added to achieve optimum computation efficiency as well as updating the model's weights and biases to more optimum point by reducing the learning rate. Finally, a specific Sparse Categorical Cross Entropy(SCCE) loss function is introduced. Therefore, combining all the functions into our CNN model helped to achieve comparatively high accuracy than multiple ensembling models on different skin lesions dermoscopic datasets. Our paper represents how our model achieved such optimum results with low computational power.