

DC-GAN-based synthetic X-ray images augmentation for increasing the performance of EfficientNet for COVID-19 detection

ABSTRACT

Currently, many deep learning models are being used to classify COVID-19 and normal cases from chest X-rays. However, the available data (X-rays) for COVID-19 is limited to train a robust deep-learning model. Researchers have used data augmentation techniques to tackle this issue by increasing the numbers of samples through flipping, translation, and rotation. However, by adopting this strategy, the model compromises for the learning of high-dimensional features for a given problem. Hence, there are high chances of overfitting. In this paper, we used deep-convolutional generative adversarial networks algorithm to address this issue, which generates synthetic images for all the classes (Normal, Pneumonia, and COVID-19). To validate whether the generated images are accurate, we used the k-mean clustering technique with three clusters (Normal, Pneumonia, and COVID-19). We only selected the X-ray images classified in the correct clusters for training. In this way, we formed a synthetic dataset with three classes. The generated dataset was then fed to The EfficientNetB4 for training. The experiments achieved promising results of 95% in terms of area under the curve (AUC). To validate that our network has learned discriminated features associated with lung in the X-rays, we used the Grad-CAM technique to visualize the underlying pattern, which leads the network to its final decision.