Data Augmentation and Lightweight Convolutional Neural Networks in Classification of Thoracic Diseases

Ana-Maria HAMZ*, Diogen BABUC, and Todor IVAŞCU

Computer Science Department, West University of Timişoara, Vasile Pârvan Blvd., no. 4, 300223 Timişoara, Romania.
E-mails: ana.hamz04@e-uvt.ro; diogen.babuc@e-uvt.ro; todor.ivascu@e-uvt.ro

* Author to whom correspondence should be addressed;

Abstract
The medical field has seen a tremendous transformation due to technological breakthroughs, with advanced medical imaging techniques becoming indispensable for diagnosis and treatment. Convolutional neural network (CNN) models have demonstrated remarkable accuracy in analyzing and classifying medical images, often surpassing human performance. In this study, we contrast two important methods for classifying medical images: lightweight CNN models (that are tailored for devices with limited resources) and data augmentation (to improve model generalization). Evaluating these models' effectiveness and performance in identifying thoracic illnesses, such as breast cancer, COVID-19 effects, and pneumonia, is the main goal. To enhance model performance, the study used preprocessed and augmented publicly available chest X-ray scans and computer tomography images. Specific CNN models used in the experiments are MobileNet, EfficientNet-B0, ResNet50, and DenseNet121. The state-of-the-art for these models show that despite lowering the danger of overfitting, data augmentation greatly increases model accuracy. Lightweight models provided the best possible compromise between accuracy and resource efficiency, performing on par with complicated models. The suitability of lightweight models for portable medical equipment was validated by testing on devices with limited resources, allowing for quick and precise pre-diagnosis. In addition to highlighting the potential of lightweight CNN to increase diagnostic accessibility, these findings emphasize the importance of striking balancing performance and efficiency in medical applications, particularly in resource-limited settings.

Keywords: Convolutional Neural Networks (CNN); Data Augmentation; Lightweight Models; Medical Images; Classifications.