A Unified Comparative Analysis of Attention Mechanisms for Pneumonia Detection in Chest X-Ray Images

T. Selvarasa*, A. Ramanan

Department of Computer Science, University of Jaffna, Jaffna, 40000, Sri Lanka
**thushani05s@gmail.com*

Lower respiratory tract infections continue to represent a significant global health burden, particularly among children in low-resource settings. Limited access to expert radiological interpretation further increases the need for reliable automated diagnostic systems using chest X-ray (CXR) imaging. This study presents a systematic and controlled comparative analysis of three widely used attention mechanisms: Squeeze-and-Excitation (SE), Spatial Attention (SA), and Convolutional Block Attention Module (CBAM) within a unified deep learning framework for pediatric pneumonia detection. Unlike existing studies, many approaches report results using inconsistent datasets, preprocessing methods, and training strategies, making fair comparison difficult. To address this limitation, the present study ensures fairness and reproducibility by integrating each attention mechanism into the same ResNet50 backbone architecture and training all models under identical experimental conditions. The study utilises a publicly available Chest X-ray dataset (https://tinyurl.com/kaggle-CXR) comprising 5,856 pediatric CXR images, categorised into pneumonia and normal classes, using the predefined patient-wise train, validation, and test splits to prevent data leakage and ensure fair evaluation. Consistent preprocessing and data augmentation techniques are applied to improve model generalisation. Each attention module is incorporated across all hierarchical stages of the network, enabling progressive refinement of feature representations from low-level visual patterns to high-level disease-specific characteristics. Model evaluation is conducted using accuracy, precision, recall, and F1-score, along with Grad-CAM visualisations to assess interpretability and highlight clinically relevant lung regions. An ablation study is conducted to analyse the contribution of attention placement across network stages and evaluate the impact of channel and spatial attention components. The experimental findings indicate that all attention-based models outperform the ResNet50 models, confirming the effectiveness of attention mechanisms in improving classification performance. ResNet50 achieved an accuracy of 91.03%, while ResNet50+SE, ResNet50+SA, and ResNet50+CBAM achieved 91.99%, 92.31%, and 92.95%, respectively. ResNet50+CBAM improved accuracy by 1.92 percentage points over ResNet50. Test results confirm this gain stems specifically from combined channel-spatial refinement across all stages, not either component alone. These findings offer the first controlled evidence that CBAM's advantage holds specifically for pediatric pneumonia detection.

Keywords: pneumonia, chest X-Ray Image, deep learning, attention mechanisms, CBAM