

WeCAViT: A Weighted CNN Model for Pneumonia Detection in Chest X-rays

Maryam Yashtini

Georgetown University

Pneumonia is a severe respiratory disease with high global mortality rates, particularly impacting children under 5 and the elderly. Early and accurate detection of pneumonia is critical for improving patient outcomes; however, diagnosing pneumonia through chest X-rays (CXR) is often challenging and requires substantial expertise. This paper presents WeCAViT—a Weighted Ensemble CNN with Attention and Visual Transformer—a novel AI-driven model designed to enhance pneumonia detection in CXR images by automating and improving diagnostic accuracy. The WeCAViT model integrates an ensemble of Convolutional Neural Networks (CNNs), a Spatial and Channel Attention block, and a Visual Transformer (ViT), each contributing uniquely to feature extraction. CNN encoders capture local features at varying scales (24×24 to 112×112), allowing multi-resolution analysis. The attention block highlights the most relevant spatial and channel details, while the Visual Transformer captures broader, global characteristics within the image. Evaluated on the PneumoniaMNIST dataset, WeCAViT achieved a test accuracy of 93.6%, outperforming or matching current state-of-the-art models. These results suggest that WeCAViT could serve as an effective tool in clinical settings, supporting timely and reliable pneumonia diagnosis.

In Special Session 58 on Recent Advances in Convex and Riemannian Optimization at the Mathematical Congress of the Americas 2025.