Project Proposal: Automating Chest X-Ray Analysis with CNNs and Vision Transformers

Team Members

This project, developed collaboratively team members

1. Aditya Manickam Arumugam (251448854) (aarumug6@uwo.ca)
2. Christian Arvin Castelo (251448854) (ccastelo@uwo.ca)
3. Semal Shastri (251448774) (sshastr@uwo.ca)

Dataset Description (Chest X-Ray Images (Pneumonia)**)**

Our model is trained on a dataset of pediatric chest X-ray images from Guangzhou Women and Children's Medical Center, Guangzhou. The dataset comprises 5,863 labeled images, divided into three folders—train, test, and validation—and organized into subfolders based on diagnostic categories: Pneumonia and Normal. These images, in JPEG format, were selected from pediatric patients aged one to five years and underwent rigorous quality control. Low-quality or unreadable scans were excluded, and two expert physicians reviewed each image to ensure accurate diagnosis, with a third expert verifying the evaluation set. This careful curation guarantees a high-quality, balanced dataset for training and evaluation.

Idea and Implementation

This project explores the use of Convolutional Neural Networks (CNNs) and Vision Transformers (ViT) to automate chest X-ray analysis, supporting radiologists in diagnosing thoracic diseases such as pneumonia and tuberculosis. Using a curated dataset of 5,863 pediatric X-ray images, we aim to identify abnormalities reliably and accurately.

The approach combines ResNet-50, a CNN architecture known for its feature extraction capabilities, with ViT, which enhances spatial relationship analysis by treating image patches as sequences. This fusion allows the model to capture both fine-grained details and broader spatial patterns, critical for accurate diagnosis. Pre-processing steps like contrast normalization and data augmentation will further strengthen model robustness.

To assess performance, metrics such as precision, recall, F1-score, and AUC will evaluate the model's diagnostic accuracy and clinical reliability. Ultimately, this project aims to demonstrate the effectiveness of CNNs and ViTs in reducing diagnostic time, supporting early disease detection, and enhancing patient care through reliable, automated analysis.

Evaluation Metrics

To rigorously assess the model's performance, metrics such as precision, recall, F1-score, and AUC will be employed. These metrics will help evaluate diagnostic accuracy, sensitivity, and specificity, ensuring the model's clinical reliability and robustness.

Conclusion

This project aims to demonstrate the diagnostic potential of CNNs and Vision Transformers in automating chest X-ray analysis. By reducing diagnostic time and supporting early detection, the model offers a promising tool for radiologists, ultimately enhancing patient care through reliable and efficient diagnostic assistance.