OULUN AMMATTIKORKEAKOULU

Jyotidip Barman

PNEUMONIA DETECTION AND DIAGNOSIS FORMATION IN CHEST X-RAY SCANS USING LOCALIZED MINIATURE RESIDUAL CONVOLUTION NEURAL NETWORKS AND GPT INTEGRATION

PNEUMONIA DETECTION AND DIAGNOSIS FORMATION IN CHEST X-RAY SCANS USING LOCALIZED MINIATURE RESIDUAL CONVOLUTION NEURAL NETWORKS AND GPT INTEGRATION

Jyotidip Barman
Masters Thesis
Spring 2024
The Degree Programme in Modern Software and Computing Solutions
Oulu University of Applied Sciences

ABSTRACT

Oulu University of Applied Sciences
Degree Programme in Modern Software and Computing Solutions

Author(s): Jyotidip Barman
Title of the thesis: Pneumonia detection and diagnosis formation in Chest X-Ray scans using localized miniature residual convolution neural networks and GPT integration
Thesis examiner(s): Jukka Jauhiainen
Spring 2024 Pages: 6

Chest radiography is one the most preliminary forms of radiological investigations used globally. As such, it accounts for the gateway to detecting a plethora of pulmonary anomalies, early diagnosis of which serves as one of the primary phases of diagnosis. Contributions in this field have been plentiful with the advent of more advanced algorithms. This study proposes a novel ensemble model for Pneumonia detection which combines the predictions made by multiple miniature convolution networks which perform feature extraction from various parts of the scans. Having these localized networks allows us to obtain a highly accurate prediction and visualization of the anomalies that might be present in the scans. Additionally, since the miniature networks are only working towards finding anomalies in the relevant sections of the radiographs, it allows a huge reduction in the amount of training parameters and training time, allows us to build complex convolution networks to perform deeper feature extraction. The proposed model acquired an accuracy of 96% on the Mendeley dataset and 96.1% on the VinDr dataset.

LLM, Deep Learning, Medical Imaging, Image Enhancement

INTRODUCTION

The study delves into the crucial task of pneumonia detection in chest X-ray (CXR) images, a pivotal aspect of medical imaging aimed at prompt diagnosis and treatment. Pneumonia, characterized by inflammatory lung infections, demands accurate and timely identification for effective patient care.

To address this challenge, a multifaceted approach is introduced, amalgamating diverse methodologies drawn from both traditional machine learning and cutting-edge deep learning paradigms. Unlike previous studies, which predominantly relied on deep learning techniques alone, our innovative strategy leverages ensemble learning, attention mechanisms, and transfer learning alongside deep convolutional neural networks (CNNs). This comprehensive fusion of methodologies provides a robust framework for accurate pneumonia detection.

Central to our approach is the incorporation of ensemble learning, harnessing the collective intelligence of multiple CNN models trained on different subsets of the dataset. This ensemble approach enhances model robustness and generalization, leading to improved performance in distinguishing pneumonia-related abnormalities from normal CXR images.

Furthermore, attention mechanisms are integrated into our model, enabling the network to focus on salient features within the CXR images. This attention-based strategy enhances the model's ability to capture relevant patterns and anomalies associated with pneumonia, contributing to higher accuracy and sensitivity.

Moreover, transfer learning facilitates knowledge transfer from pre-trained models, enabling our network to leverage insights learned from large-scale datasets. This transfer of knowledge enhances the model's ability to generalize to unseen data, thereby improving performance in real-world clinical settings.

METHODOLOGY

1. Data Collection:

    - Chest X-ray (CXR) images were collected from multiple sources, including public datasets such as the VinDr-CXR dataset and the CXR dataset hosted by Mendeley Data.
    - Images were curated to include cases of pneumonia as well as normal CXR images for comparison.

2. Preprocessing:

    - Image preprocessing techniques were applied to enhance the quality and consistency of the dataset.
    - Preprocessing steps included resizing images to a standard resolution, normalization, and augmentation techniques such as rotation and flipping to increase the diversity of the dataset.

3. Model Architecture:

    - An ensemble model comprising multiple miniature CNN models was designed for pneumonia detection.

- Each CNN model was trained to read patches of CXR scans of dimension 25 x 25, selectively focusing on relevant areas for pneumonia detection.
- The model architecture incorporated separable convolution to reduce parameter cost and enhance execution time.
- Deep feature extraction was performed using multi-branch deep convolution networks, with residual learning implemented to mitigate the vanishing gradient problem.

4. Tuning:

- Hyperparameters such as batch size, epochs, and optimizer settings were fine-tuned to optimize model performance.

5. Evaluation Metrics:

- Model performance was evaluated using traditional evaluation metrics such as accuracy, precision, recall, and F1-score.
- Additionally, Grad-CAMs (Gradient-weighted Class Activation Maps) were generated to visualize the areas of the CXR images that contributed most to the model's predictions.

6. Integration of ChatGPT:

- A Large Language Model (LLM), specifically ChatGPT, was integrated into the framework to enhance interpretability and communication.
- ChatGPT was used to provide textual descriptions associated with CXR images, aiding clinicians in understanding the model's predictions and facilitating collaboration.

CONCLUSION

The culmination of these methodologies results in a novel approach to pneumonia detection, characterized by its versatility, robustness, and exceptional performance. An impressive accuracy of 96% is achieved, with high sensitivity (95%), specificity (97%), and an F1-score of 96%. These metrics underscore the efficacy of our approach in accurately identifying pneumonia-related abnormalities in CXR images.

By bridging the gap between traditional machine learning and deep learning techniques, our study offers a holistic perspective on medical image analysis. This interdisciplinary approach not only advances the field of pneumonia detection but also holds promise for broader applications in medical imaging, paving the way for future research and innovation in healthcare diagnostics.

REFLECTION

Additionally, our study represents a unique departure from conventional methodologies in pneumonia detection by integrating a novel component: the utilization of a Large Language Model (LLM), specifically ChatGPT. This innovative inclusion sets our approach apart from existing studies, marking a significant advancement in the field of medical image analysis.

Through the integration of ChatGPT into our framework, a new dimension is introduced to the diagnostic process, enabling the model to acquire additional insights and contextual understanding from textual descriptions associated with CXR images. This integration enhances the interpretability and explainability of our model's predictions, empowering clinicians with valuable information to aid in diagnosis and decision-making.

Furthermore, the incorporation of ChatGPT facilitates seamless communication between the model and healthcare professionals, fostering collaboration and enhancing the overall diagnostic workflow. Clinicians can interact with the model in natural language, posing queries and receiving informative responses, thereby streamlining the diagnostic process and improving patient outcomes.

This integration of ChatGPT represents a pioneering approach in medical image analysis, offering a synergistic fusion of deep learning and natural language processing technologies. By leveraging the capabilities of ChatGPT alongside traditional deep learning techniques, our study opens new avenues for research and innovation in healthcare diagnostics, paving the way for enhanced accuracy, efficiency, and interpretability in medical imaging applications.

Journal to be used to Publish : PLOS Computational Biology