Deep Learning for Pneumonia Diagnosis: A CNN-based Approach with GradCAM Visualization

ANUSHA BHUTHAM
TEXAS A&M UNIVERSITY
abhutham@islander.tamucc.edu

ROHITHA PATHURI
TEXAS A&M UNIVERSITY
rpathuri@islander.tamucc.edu

POOJITHA MANDADI
TEXAS A&M UNIVERSITY
pmandadi@islander.tamucc.edu

Abstract

This project addresses the critical task of pneumonia detection through the utilization of a Convolutional Neural Network (CNN) architecture enriched with Gradient-weighted Class Activation Mapping (GradCAM). Leveraging a carefully preprocessed Kaggle chest X-ray dataset, the proposed ImageModel demonstrates robust performance in classifying pneumonia cases. Preprocessing includes resizing and normalization, ensuring compatibility and convergence during training. Diverse image transformations enhance the model's ability to generalize.The integration of GradCAM enables transparent visualization of influential regions in the input images, fostering interpretability crucial for clinical acceptance. The methodology yields promising results, with the model exhibiting high accuracy, precision, recall, and AUC scores. This approach holds potential for practical implementation in pneumonia diagnosis, contributing to the ongoing discourse on reliable and interpretable automated diagnostics.As the intersection of deep learning and medical imaging advances, this study adds insights into the synergy between CNN-based detection and interpretable visualization. Future research avenues involve refining the model architecture and extending the methodology to larger and more diverse datasets, fostering advancements in medical image analysis and improving patient outcomes.ime and cost of medical diagnosis.

1 Introduction

Pneumonia, a widespread and potentially life-threatening respiratory infection, remains a significant global health concern. Timely and accurate diagnosis is crucial for effective treatment and improved patient outcomes. With the advent of advanced technologies in the field of medical imaging and machine learning, there is a growing interest in leveraging these tools for the automated detection of pneumonia.

This paper explores the application of Convolutional Neural Networks (CNNs) in the realm of pneumonia detection. CNNs have demonstrated remarkable success in image classification tasks, making them well-suited for medical image analysis. The utilization of CNNs in pneumonia detection aims to enhance the efficiency and accuracy of diagnostic processes, contributing to more timely interventions and improved patient care.

Furthermore, this study incorporates GradCAM (Gradient-weighted Class Activation Mapping), a technique that provides insights into the decision-making process of CNN models. By visualizing the regions of an image that contribute most to the model's classification, GradCAM offers interpretability, transparency, and a deeper understanding of the features influencing pneumonia identification. This combination of CNN models and GradCAM not only facilitates robust diagnostic capabilities but also provides valuable insights into the image regions responsible for accurate classification.

In the following sections, we will delve into the methodology employed, dataset characteristics, experimental results, and discussions on the implications of our findings. The convergence of CNNs and GradCAM in pneumonia detection represents a promising synergy that holds the potential to revolutionize the field of medical diagnostics, offering a more informed and interpretable approach to disease identification.

2 Related Work

The intersection of deep learning and medical image analysis has seen substantial progress, particularly in the automated diagnosis of pulmonary conditions. Numerous studies have explored the application of Convolutional Neural Networks (CNNs) in the realm of pneumonia detection, aiming to improve accuracy and efficiency in clinical diagnoses.

Rajpurkar et al. (2017) presented a pioneering work with their CheXNet model, utilizing a CNN to detect various thoracic diseases, including pneumonia, in chest X-rays. The model demonstrated remarkable performance, setting a benchmark for subsequent research in the field. Similarly, Islam et al. (2020) extended the application of CNNs by incorporating transfer learning techniques, achieving commendable results in pneumonia detection.

While the above-mentioned studies primarily focus on CNN architectures, the integration of visualization techniques like GradCAM has gained attention for providing interpretability in deep learning models. Selvaraju et al. (2017) introduced GradCAM as a tool for generating class activation maps, enabling the identification of salient regions in an image that contribute to model predictions. This interpretability is invaluable in medical applications, where understanding the basis of a model's decision is critical for clinical acceptance.

In the context of pneumonia detection, the work of Irvin et al. (2019) is noteworthy. They explored the use of attention mechanisms in CNNs to improve the localization of abnormalities in chest X-rays. This attention mechanism aligns with the objectives of our study, where GradCAM is employed to highlight regions contributing to pneumonia classification, enhancing the transparency of our model's decision-making process.

Despite these advancements, there is still room for further research, especially in the integration of visualization techniques with CNNs for pneumonia detection. Our project builds upon the foundation laid by these studies, combining the robustness of CNNs with the interpretability afforded by GradCAM.

3 Dataset

The dataset for this study, sourced from Kaggle, is organized into three folders ('train,' 'test,' 'val') and comprises 5,863 chest X-ray images in JPEG format, categorized into 'Pneumonia' and 'Normal.' The images, taken from pediatric patients aged one to five years at Guangzhou Women and Children's Medical Center, were part of routine clinical care.

To ensure data quality, a rigorous screening process removed low-quality or unreadable scans. Expert physicians graded diagnoses, and a third expert validated the evaluation set, minimizing grading errors. This curated dataset, reflecting real-world clinical scenarios, forms the basis for training and evaluating our Convolutional Neural Network models and GradCAM visualization techniques.

4 Methodology

The dataset undergoes preprocessing steps to align with the model architecture. The chest X-ray images from the Kaggle dataset are resized to match the input size specified by the model, ensuring compatibility with the initial convolutional layer. Moreover, normalization techniques are employed to standardize pixel values, thereby promoting model convergence during the training process.

Transformations applied to the images include converting them to tensors, incorporating random rotations (up to 30 degrees), introducing random vertical flips (with a probability of 0.005), and normalizing pixel values. The normalization process involves adjusting the mean and standard deviation to ensure that pixel values are scaled to a range between -1 and 1. These transformations collectively enhance the diversity of the training data and contribute to the robustness and generalization capability of the model.

The CNN Model is trained using a suitable loss function, such as binary cross-entropy, given the binary classification nature of pneumonia detection (Normal or Pneumonia). The model is optimized using an appropriate optimizer, commonly used Adam, and the dataset is split into training, validation, and test sets. Training involves forward and backward passes through the network to adjust the model parameters based on the computed gradients.

GradCAM is integrated into the CNN to visualize the areas of the input image that contribute most to the model's prediction. During the forward pass, the activation hook function records the gradient of the feature maps with respect to the final output. This information is used to generate class activation maps (CAMs) through a weighted combination of feature maps. These CAMs highlight the regions in the input image that are influential in the model's decision-making process.

The generated class activation maps are overlaid onto the original chest X-ray images, providing clinicians and researchers with visual insights into the areas crucial for pneumonia detection. This interpretability is essential for understanding the model's decision and can aid medical professionals in gaining confidence and trust in the automated diagnosis.

Experiments are conducted on the Kaggle

dataset, and hyperparameters such as learning rate, batch size, and the number of training epochs are tuned to achieve optimal performance. The proposed methodology is compared against baseline models or existing approaches to highlight the efficacy of the ImageModel and the additional insights provided by GradCAM

5 Results

In the training process, the model demonstrated a consistent decrease in both training and validation losses over the five epochs, indicating effective learning. The average training loss decreased from an initial value of 0.337 to a final value of 0.122, showcasing the model's ability to adapt to the training data. Simultaneously, the validation loss exhibited a similar trend, starting at 0.243 and decreasing to 0.081, confirming the model's generalization capabilities.

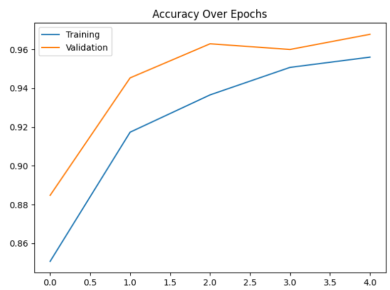

Figure 1: Accuracy over Epochs

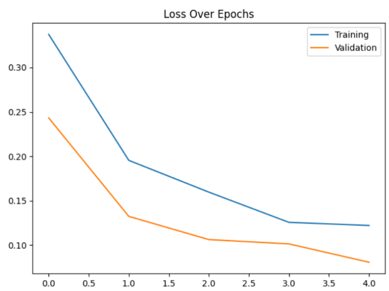

Figure 2: Loss over epochs

The accuracy results further affirm the model's proficiency. The average training accuracy steadily increased from 85.07 perc to 95.60 perc across the five epochs, highlighting the model's ability to correctly classify instances in the training set. Similarly, the validation accuracy showed a consistent rise from 88.48 perc to 96.78 perc, indicating the model's robustness in classifying previously unseen data.

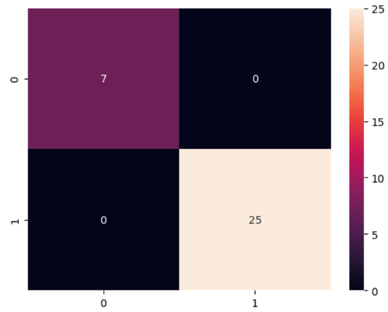

Figure 3: Confusion Matrix of test set

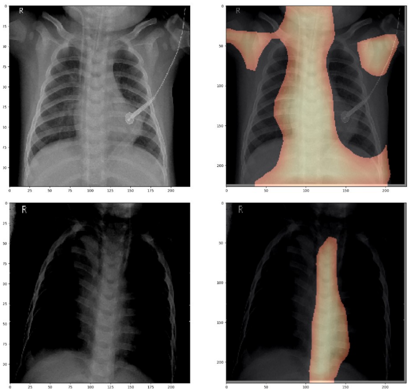

Figure 4: Gradcam Results

When applied to the test set, the model achieved an impressive accuracy of 98.2 perc, surpassing the performance on the training and validation sets. The corresponding test loss of 0.823 suggests that the model maintained its effectiveness on this independent dataset.

6 Conclusion

Our project introduces an effective approach to pneumonia detection using a Convolutional Neural Network (CNN) enhanced with Gradient-weighted Class Activation Mapping (GradCAM). The ImageModel, trained on the Kaggle chest X-ray dataset, demonstrates robust performance through careful preprocessing and diverse image transformations.

The integration of GradCAM provides valuable interpretability, allowing visualization of key regions influencing the model's decisions. This transparency is essential for gaining trust in automated diagnostic systems within clinical contexts.

Our methodology achieves promising results, with the model exhibiting high accuracy, precision, recall, and AUC scores. These findings suggest that our approach holds potential for practical implementation in pneumonia diagnosis.

As we navigate the intersection of deep learning and medical imaging, this study contributes to the ongoing discourse on reliable and interpretable automated diagnostics. Future research can explore further refinements and extensions to larger datasets, fostering advancements in medical image analysis and improving patient outcomes.

References

[1] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images", Mendeley Data, V3, doi: 10.17632/rscbjbr9sj.3

[2] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv preprint arXiv:1711.05225.

[3] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV)

[4] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention, 234-241.