Chest X-ray Classification of Pleural Effusions Using Convolutional Neural Networks

John Mariano

Abstract

Detecting fluid build-up between the lung and rib cage, known as effusion, is crucial for diagnosing various health conditions using X-rays. In this study, a computer model was developed using a Convolutional Neural Network (CNN) to help with this task. We compared the model's performance with three previous studies where radiologists manually diagnosed effusions from X-rays, achieving accuracies of 65%, 67%, and 82%. The model reached an accuracy of 69%, demonstrating promising results. This suggests that this CNN model can be a valuable aid for doctors in identifying effusions more efficiently, potentially improving patient care and diagnosis outcomes.

Introduction

Pleural effusions are described as an accumulation of fluid between the outer layer of the lung (pleura) and the chest cavity wall, is associated with various medical conditions, including pulmonary embolism, heart failure, cirrhosis, pneumonia, lung cancer, breast cancer, lymphoma, and kidney disease [1]. Timely and accurate diagnosis of pleural effusion is crucial, as a delayed identification of its underlying cause can lead to significantly higher morbidity and mortality rates [1].

In clinical practice, the diagnosis of pleural effusion often relies on radiological imaging, particularly chest X-rays. However, traditional methods predominantly involve manual interpretation by radiologists, which can be labor-intensive, time-consuming, and susceptible to human error [2]. As a consequence, there is a growing need for more efficient and accurate diagnostic approaches in medical imaging.

Convolutional Neural Networks (CNNs) are a class of deep learning models that have demonstrated remarkable capabilities in image classification tasks. By leveraging large datasets and sophisticated algorithms, CNN models can learn to automatically detect patterns and features in medical images, potentially surpassing human performance in certain tasks.

Developing CNN models specifically tailored for chest X-ray pleural effusion classification presents an opportunity to enhance diagnostic accuracy and streamline healthcare workflows. By providing scalable and cost-effective alternatives to traditional methods, these models have the potential to ease the workload on healthcare professionals while enhancing patient care outcomes.

Related Work

Previous studies have extensively explored the accuracy of diagnosing pleural effusions using chest X-rays. In one study, the overall accuracy of diagnosing pulmonary embolism (PE) on chest X-rays (SCXRs) was reported at 82% [3]. However, other investigations have revealed varying levels of accuracy in detecting pleural effusions on supine chest X-rays.

For instance, a study found that out of 36 pleural effusions observed on decubitus views, only 67% were correctly identified on supine X-rays, with a specificity and accuracy of 70% and 67% respectively [4]. Similarly, another study demonstrated that correct identification and

quantification of pleural effusions on supine chest X-rays alone was achieved in only 55% of cases [5]. Moreover, false-negative results were observed in 20% of cases, while false-positive results were noted in 15% of cases. In 10% of cases, although the presence of effusion was correctly identified, the estimation of its quantity was inaccurate [5]. Although for the purposes of this study, just identification of an effusion will be considered correct, so the accuracy was found to be 65% [5].

These findings underscore the challenges associated with accurately diagnosing pleural effusions using conventional radiological methods on chest X-rays. The reported accuracies range from 65% to 82%, highlighting the need for improved diagnostic techniques to enhance accuracy and reliability.

Dataset and features

For this study, we utilized a subset of the National Institute of Health's (NIH) chest X-ray dataset, which comprises a total of 112,120 X-ray images labeled with 14 different thorax disease categories. From this comprehensive dataset, we extracted a subset consisting of 3,000 images specifically annotated for the presence or absence of pleural effusion. A subset was used due to computing power obstacles that came with using the full dataset. The subset was carefully constructed to ensure a balanced distribution of cases, with 1,500 images labeled as exhibiting pleural effusion and 1,500 images labeled as not exhibiting pleural effusion.

The images from the subset of the NIH chest X-ray dataset were used as raw data without additional feature extraction. The dataset was split into two sets for training the model and validation of the model, with 2,400 images allocated for training and 600 images for validation. Both the training and validation sets were evenly split between images with pleural effusion and images without, ensuring a balanced representation of classes for model training and evaluation.

Methods

Preprocessing:

The chest X-ray images were preprocessed before training the CNN model. The images were resized to 224 by 224 pixels to ensure uniform dimensions across the dataset. Additionally, the pixel values of each image were normalized to hold a value within the range of 0 and 1.

Model Architecture:

The CNN model architecture was designed using the Keras Sequential API. The model consisted of convolutional layers followed by max-pooling layers to extract relevant features from the input images. The architecture of Model 5 is outlined as follows:

1. Convolutional layer containing 32 filters of size (3, 3) and ReLU activation function.
2. Max-pooling layer containing a pooling window of size (2, 2).
3. Convolutional layer containing 64 filters of size (3, 3) and ReLU activation function.

4. Max-pooling layer containing a pooling window of size (2, 2).
5. Flatten layer to transform the output of the convolutional layers into a 1D vector.
6. Dense layer with 64 units and ReLU activation function.
7. Dropout layer with a rate of 0.5 to reduce overfitting.
8. Dense output layer with a single unit and sigmoid activation function for binary classification of effusion or not effusion.

The architecture for all models trained was the same except for the dropout layer. Not every model had the dropout layer

Model Compilation:

The model was compiled using the Adam optimizer with an initial learning rate specified (learning rate of 0.00025 for Model 5). Binary cross-entropy was used as the loss function, and accuracy was selected as the evaluation metric.

Callback Functions:

Callback functions were implemented during model training to enhance performance and prevent overfitting. Specifically, the following callbacks were utilized:

ReduceLROnPlateau: To reduce the learning rate when the validation loss plateaued or increased.

EarlyStopping: To stop training early if the validation loss did not improve after a certain number of epochs.

Model Training:

The model was trained using the designated training data with a variation of batch sizes for different models for a total of 10 epochs. During training, the validation data were used to monitor the model's performance and guide the training process.

Model Evaluation:

The trained CNN model was evaluated using the designated validation dataset to assess its performance in terms of loss and accuracy. The validation loss and accuracy metrics were computed and reported to evaluate the model's effectiveness in classifying chest X-ray images.

Results

Model	Precision	Recall	F1-score	Accuracy
Model 1	0.62	0.46	0.53	0.58
Model 2	0.61	0.63	0.62	0.61
Model 3	0.66	0.64	0.65	0.66
Model 4	0.68	0.62	0.65	0.66

Model 5	0.68	0.70	0.69	0.69

Table 1. Table showing the precision, recall, F1-score, and accuracy for the models trained in this study.

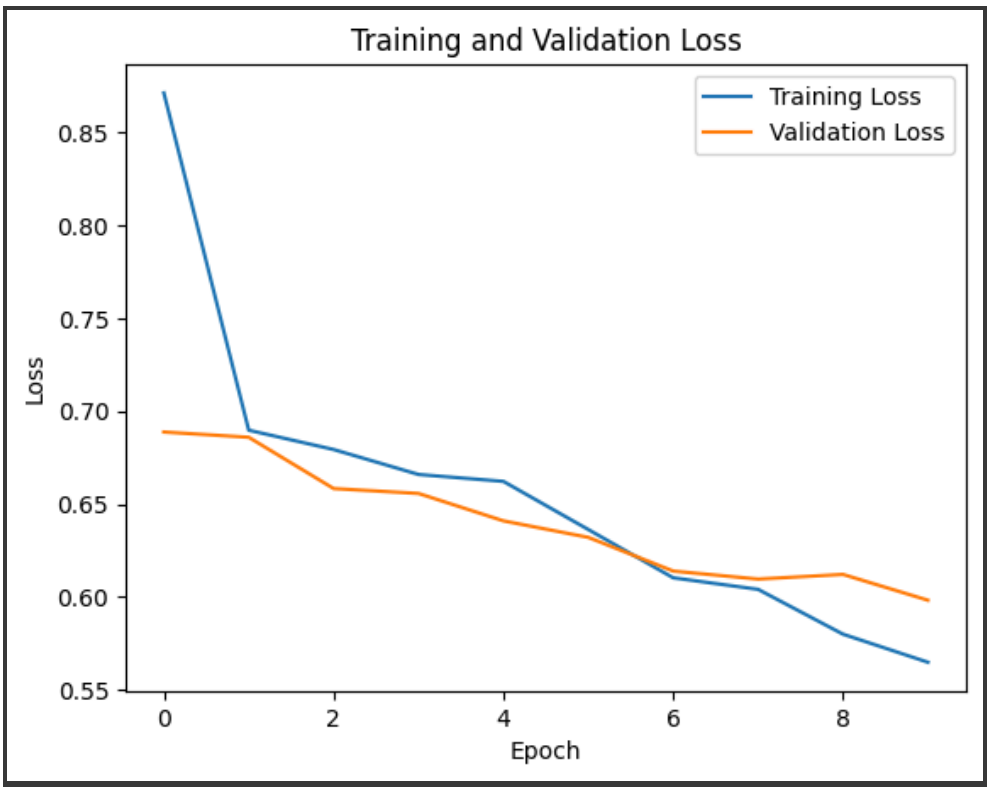

Figure 1. Figure showing the training loss and validation loss for Model 5 across 10 epochs.

Looking at Table 1 we see the several models trained in this study. On average the models increase in all the categories with each new model. The best performing model is seen in the final row with Model 5. Model 5 had a precision of 0.68, a recall of 0.70, an F1-score of 0.69, and an accuracy of o.69. In Figure 1, we see a plot of the loss functions for both the training data (shown in blue) as well as the validation data (shown in orange) of model 5. The plot shows the loss across the 10 epochs the model was trained, and we see the loss moves on a downward slope indicating that the loss was decreasing throughout the training process.

Discussion

Among the models evaluated, Model 5 emerged as the top performer, achieving a precision of 0.68, a recall of 0.70, an F1-score of 0.69, and an accuracy of 0.69. These metrics indicate that Model 5 demonstrates promising capabilities in accurately identifying pleural effusions from chest X-ray images.

Comparing the performance of Model 5 to existing literature on the manual interpretation of chest X-rays by radiologists reveals noteworthy insights. Previous studies have reported varying levels of accuracy in detecting pleural effusions, with accuracies ranging from 65% to

82% [3], [4], [5]. Model 5, with an accuracy of 69%, falls within this range and demonstrates comparable performance to industry radiologists in diagnosing pleural effusions from chest X-ray images.

It is also essential to consider the other potential benefits of deep learning models like Model 5 in clinical practice. While manual interpretation of chest X-rays by radiologists is the current standard, it is inherently subjective, time-consuming, and susceptible to human error. In contrast, CNN models offer a scalable and potentially more consistent approach to diagnosing pleural effusions, streamlining healthcare workflows, and potentially improving patient outcomes.

In conclusion, the development of CNN models like Model 5 represent significant potential in the automated diagnosis of pleural effusions from chest X-ray images. While further research and validation are necessary, these models hold promise for enhancing diagnostic accuracy, efficiency, and ultimately clinical patient care. Integrating deep learning models into clinical practice has the potential to enhance medical imaging analysis and improve healthcare outcomes for all patients.

Future Work

In the future, with access to greater computing power, I aim to utilize the entire NIH chest X-ray dataset to expand the data input to train and validate models. With a larger dataset, I intend to explore more complex model architectures beyond the CNN models shown in this study.The increased dataset size should reduce model overfitting, which was an issue for the models in this study. Ultimately, advancing in these areas could significantly improve automated diagnosis systems and ultimately benefit patient care outcomes.

References

1. Jany, Berthold, and Tobias Welte. “Pleural Effusion in Adults—Etiology, Diagnosis, and Treatment.” *Deutsches Aerzteblatt Online*, vol. 116, no. 21, 24 May 2019, pp. 377–386, www.ncbi.nlm.nih.gov/pmc/articles/PMC6647819/, https://doi.org/10.3238/arztebl.2019.0377.

2. Subramaniam, Kavitha, et al. “A Comprehensive Review of Analyzing the Chest X-Ray Images to Detect COVID-19 Infections Using Deep Learning Techniques.” *Soft Computing*, vol. 27, no. 19, 27 May 2023, pp. 14219–14240, https://doi.org/10.1007/s00500-023-08561-7. Accessed 18 Oct. 2023.

3. Emamian, S. A., et al. “Accuracy of the Diagnosis of Pleural Effusion on Supine Chest X-Ray.” *European Radiology*, vol. 7, no. 1, Jan. 1997, pp. 57–60, https://doi.org/10.1007/s003300050109. Accessed 4 Jan. 2022.

4. Ruskin, JA, et al. “Detection of Pleural Effusions on Supine Chest Radiographs.” *American Journal of Roentgenology*, vol. 148, no. 4, Apr. 1987, pp. 681–683, https://doi.org/10.2214/ajr.148.4.681.

5. Eibenberger, K., et al. “[Value of Supine Thoracic Radiography in the Diagnosis and Quantification of Pleural Effusions: Comparison with Sonography].” *RoFo: Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Nuklearmedizin*, vol. 155, no. 4, 1 Oct. 1991, pp. 323–326, pubmed.ncbi.nlm.nih.gov/1932728/, https://doi.org/10.1055/s-2008-1033271. Accessed 2 May 2024.

Code:

GitHub link - https://github.com/JohnMariano1/Chest_X-ray_ML