Dean&Francis

ISSN 2959-6157

A literature review of pneumonia detection algorithms based on deep learning and chest X-ray imaging

Chenyu Wang

School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, 410000, China
alessioliz@foxmail.com

Abstract:

Pneumonia is a serious disease that poses a threat to people's health of all ages. It could happen when people are infected by viruses, fungi, bacteria etc. Typically, Chest X-rays are the first and foremost imaging approach to implement pneumonia detection. This paper introduces the latest research achievements to help those who are new in this field to have basic intuition about AI in pneumonia detection., including Vision Transformers on chest X-ray, a novel model based on RetinaNet, CP_DeepNet, and a novel Efficient NetV2L model. The last part gives some suggestions about the future study of pneumonia detection using Deep learning.

Keywords*: Deep Learning, Pneumonia detection, Vision Transformer, literature review*

1. Introduction

Pneumonia is responsible for approximately 16% of deaths in children under five globally, making it the primary cause of death among young children worldwide. [1]. Each year in the United States, approximately 1 million adults require hospital care for pneumonia, resulting in 50,000 deaths from the disease [1].

The process of detecting and classifying lung diseases through chest X-ray images is intricate for radiologists. As a result, researchers have placed considerable emphasis on developing automatic techniques for lung disease detection[2].

Deep learning constitutes a vital segment within the realm of artificial intelligence, and its appearance opens up a new research direction in modern computer field. It was proposed by Hinton et al in 2006. Its essence is supervised learning, which uses raw or pre-processed data with clear labels as the input of convolutional neural network algorithm. Then, the algorithm abstracts the original input data layer by layer into the target features required by its own task. Finally, map the learned features to the task goal to finish.[3]

In this paper, the research status of deep learning in pneumonia detection is discussed in detail, the achievements and problems in pneumonia detection are fully described, and further research suggestions are given.

2. Review of Deep Learning algorithms for pneumonia detection

2.1 Vision Transformers

This paper presents an efficient method for chest X-ray imaging pneumonia detection using vision Transformer (ViT) architecture [4]. Pneumonia is a global and serious respiratory disease that affects people of all ages. Timely identification and treatment of pneumonia are crucial to prevent complications and enhance clinical results. By designing and deploying effective tests, we can decrease mortality rates, enhance medical effectiveness, and play a role in the worldwide battle against illnesses that have afflicted humanity for centuries. The identification of pneumonia is not just a medical requirement, but also a frontier in technology. Chest X-ray is one of the commonly used imaging modes to diagnose pneumonia. This paper discusses in detail a cutting-edge approach to pneumonia detection based on Vision Transformer (ViT) architecture implemented on public datasets on Kaggle. To capture global context and spatial relationships from chest X-ray images, the proposed framework uses a ViT model that integrates a self-attention mechanism and Transformer architecture. According to our experiments with the proposed Vision Transformer based framework, it achieves a much higher accuracy in detecting pneumonia in chest X-ray images, achieving 97.61% accuracy, 95% sensitivity, and 98% specificity.

The Vision Transformer (ViT) model mentioned in this study is able to efficiently capture global context and spatial relationships from chest X-ray images by integrating a self-attention mechanism and converter architecture. Compared with traditional methods, ViT achieves higher accuracy in image classification tasks. Its self-attention mechanism can capture complex patterns and relationships in images, which helps to improve the accuracy and reliability of pneumonia detection. By leveraging ViT's ability to extract global and local image features, combined with a self-attention mechanism, ViT is able to effectively capture complex patterns and relationships in X-ray images, thereby improving the accuracy and reliability of pneumonia detection.

While the authors' study demonstrates the effectiveness of the Vision Transformer (ViT) architecture in detecting pneumonia on chest X-rays, there are still some shortcomings. First, the data set used in the study is publicly available, but there is no mention of its size, source, or potential bias. Second, while ViT models perform well in global context and spatial relationships, they have not been adequately validated for robustness with small samples or noisy data. Furthermore, there is no mention of further validation of the model, such as generalization performance on different medical institutions or different devices, and comparative experiments with clinical experts. Finally, for the diagnosis of pneumonia, in addition to X-ray images, it may be necessary to consider other clinical information, such as medical history, clinical symptoms, etc., which were not considered in this study. Hence, upcoming research can broaden the dataset, assess the model's generalization capabilities, and integrate additional clinical data to enhance the accuracy and dependability of pneumonia diagnosis.

2.2 Model based on RetinaNet

Gabruseva T's research team's contribution of this study is the development of a computational method based on deep learning for the automatic detection of pneumonia regions. By combining techniques such as single detector, deep convolutional neural network, data enhancement, and multi-task learning, the method has achieved remarkable results in pneumonia detection, successfully applied to the Radiological Society's pneumonia Detection Challenge, and achieved one of the best results. By improving the accuracy and efficiency of diagnosis, this study provides a powerful tool for improving early diagnosis of pneumonia. The model proposed by Gabruseva's team utilizes an SSD RetinaNet with SE-ResNext101 encoder pre-trained on ImageNet [5].

The study used the mean precision mean (mAP) to evaluate the model's performance at different intersection ratio (IoU) thresholds. Threshold values range from 0.4 to 0.75, with a step size of 0.05, and the predicted target is considered a "hit" if its intersection ratio with the ground real target is greater than 0.4. The mAP index changes with the training rounds and the non-maximum suppression (NMS) threshold by observing. In addition, the test set prediction box was resized to 87.5% of its original size to reflect differences in the test and training set labeling processes. Through the ablation study, it is proved that the proposed method can improve the accuracy of the model. This approach deliberately avoids data enhancement during testing, striking a balance between accuracy and resource allocation. The method detailed in the study delivered outstanding results, ranking among the top performers in the challenge.

Although the author's work has achieved good results, there are still some shortcomings. First, there is no mention of verifying the robustness of the model, such as the ability to generalize images collected from different data sources or different devices. Second, although the label and category distribution of the data set is mentioned, it is not clear whether there is a category imbalance problem and how to deal with this situation. In addition, the evaluation index only considers the detection accuracy, but does not involve the speed and practicality of the model. In terms of improvements, it is possible to increase the validation of the generalization ability of the model, adopt more data enhancement techniques to improve the robustness of the model, and explore other evaluation metrics such as speed and resource consumption to evaluate the performance of the model more comprehensively.

2.3 CP_DeepNet

The world has been significantly affected by the COVID-19 pandemic in recent years. Its symptoms, such as fever and cough, resemble those of the common flu and have rapidly become a leading cause of death. China is currently facing a new wave of COVID-19. This virus poses a serious threat to people of all ages, especially the elderly due to their weaker immune systems. While real-time polymerase chain reaction (RT-PCR) tests are commonly used for identifying coronaviruses, they are expensive, time-consuming, and have a high rate of false negatives. Therefore, there is an urgent need for a cost-effective, rapid, and reliable method to detect COVID-19. Chest X-ray images are commonly used to detect various respiratory diseases, such as lung infections, dyspnea syndrome, lung cancer, and gas accumulation in lung cavities. In this study, chest X-ray datasets were utilized to identify COVID-19 and pneumonia. The authors proposed a novel deep learning model called CP_DeepNet, which is based on pre-trained models like SqueezeNet. To evaluate its classification effectiveness, three additional convolutional layer blocks were incorporated, and data augmentation techniques were employed to generate more images and overcome overfitting. The performance of the proposed model was assessed using the COVID-19 radiographic dataset[6].

CP_DeepNet model uses deep learning technology and adds three convolutional layer blocks based on SqueezeNet model to classify COVID-19 and pneumonia for chest X-ray images. In the study, data enhancement methods were used to generate more images to overcome overfitting problems, and COVID-19 radiographic data sets were used to evaluate model performance. The model has excellent performance on multi-class data sets, and can efficiently and accurately identify multiple diseases such as COVID-19 and pneumonia, providing support for clinical diagnosis.

The RT-PCR method has some problems in the detection of COVID-19, such as expensive, long time, and complicated sample processing. The CP_DeepNet model proposes a fast and accurate alternative through deep learning techniques based on X-ray images. The model uses data-enhanced, innovative preprocessing methods combined with multi-class datasets to effectively identify COVID-19, pneumonia, and normal conditions. Compared to RT-PCR, CP_DeepNet has higher accuracy and precision, and can handle multiple diseases simultaneously, providing reliable support for rapid diagnosis of COVID-19.

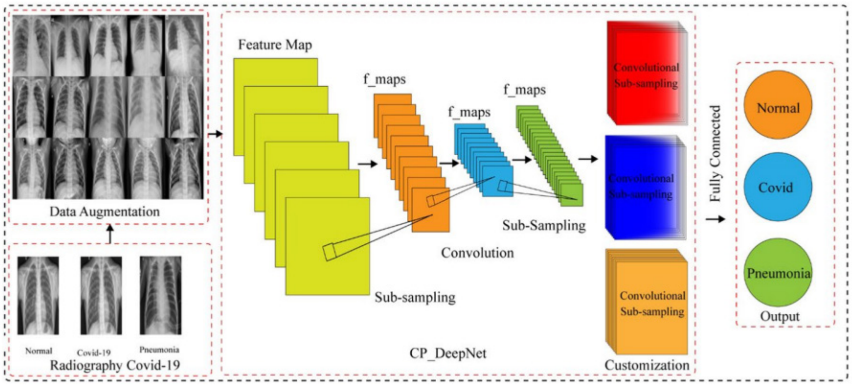

Figure 1 Architecture of proposed system [5]

The CP_DeepNet model demonstrated 99.32% accuracy, 100% accuracy, 99% recall, 99.2% specificity, 99.78% area under the curve (AUC), and 99.49% F1 score in the binary classification of COVID-19 and normal classes. The CP_DeepNet model achieved 99.62% accuracy, 99.79% accuracy, 99.52% recall, 99.69% specificity, 99.62% AUC, and 99.72% F1 scores for COVID-19, pneumonia, and normal human subjects, respectively. These remarkable results show that the CP_DeepNet model has high accuracy and precision and can effectively identify pneumonia and COVID-19.

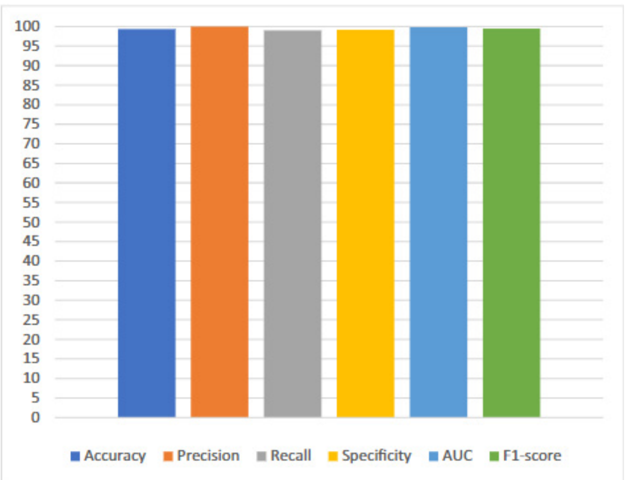

Figure 2 Performance evaluation of CP_DeepNet on Covid and Normal [5]

2.4 Novel EfficientNetV2L Model

Ali M et al propose a novel EfficientNetV2L model for the detection of pneumonia by chest X-ray[7]. Pneumonia is a potentially life-threatening infectious disease that is usually diagnosed by physical examination and diagnostic imaging techniques such as chest x-rays, ultrasound, or lung biopsies. An accurate diagnosis is crucial for patients, as an incorrect diagnosis, inadequate treatment, or lack of treatment can have serious consequences for patients, even fatal. Advances in deep learning have significantly facilitated the diagnosis of pneumonia by medical professionals by aiding the decision-making process.

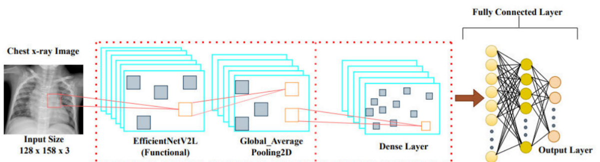

Figure 3 Architecture of EfficientNetV2L Model [6]

The model proposed by the authors, EfficientNetV2L, performed well in the evaluation of ROC-AUC and PR-AUC indicators. Outperformed other models including CNN, InceptionResNetV2, VGG16, ResNet50, and EfficientNetV2L. The model achieved excellent performance on the test set, achieving 94.02% accuracy, 94.40% accuracy, 97.24% recall and 95.80% F1 score. Through hyperparameter tuning, the model demonstrated excellent performance, not only generating accurate prediction results, but also achieving a good balance between accuracy and recall, improving classification accuracy. The EfficientNet-V2L model has excellent scalability and efficiency in the face of complex classification challenges, and performs well in various evaluation indicators.

There are still some shortcomings and unfinished work in the author's study. First, they mention that data set size can lead to overfitting problems, so they recommend collecting more data to improve the generalization of the model. Secondly, the paper mentions the importance of exploring various pre-processing techniques and convolutional neural network configurations in deep learning models, but does not elaborate on the specific implementation and effects of these techniques. In addition, although data enhancement techniques are proposed to increase the size of data sets, the concrete implementation of these techniques and the effect evaluation of data sets after enhancement are not given. Finally, the authors recommend the inclusion of additional X-ray datasets, including data for different pathology labels, to improve the accuracy and reliability of the model, but do not indicate how these additional datasets will be obtained and how their quality and consistency will be ensured. Therefore, future work can focus on addressing these issues and further refining the research results.

3. Conclusion

With the surge of deep learning, many literatures are studying the use of deep learning methods to detect and diagnose pneumonia. Although certain progress has been made, its effect still cannot fully meet the actual diagnosis of pneumonia. The main reason is the rigor of medical disease diagnosis. In the case of the current immature technology, professional doctors still need to check again, and the final result still needs doctors to make the final judgment according to their previous clinical diagnosis experience and professional knowledge. Most researchers have only looked at images of the lungs to determine whether they are pneumonia, and there are many types of pneumonia, such as lobar pneumonia, bronchopneumonia, and interstitial pneumonia, so there are some suggestions for the future study.

Multimodal data fusion: Combining medical data of different modes, such as X-ray images, CT scans, clinical data, etc., can improve the accuracy and reliability of pneumonia detection.

Transfer learning and reinforcement learning: The use of transfer learning can solve the problem of insufficient data and uneven data distribution, and improve the generalization ability of models in new fields. Reinforcement learning enables the model to continuously optimize through interaction with the environment and adapt to pneumonia detection tasks in different situations.

Model explainability: Researchers can explore how to increase the explainability of deep learning models, for example by visualizing the model's attention mechanisms, or by leveraging interpretive deep learning model structures.

Privacy protection technology: Privacy protection technology such as state encryption, federated learning, etc., can help model training without exposing sensitive data and protect the privacy of patient data.

Automated screening systems: Combining deep learning models with automated screening systems can improve the efficiency and accuracy of pneumonia detection and reduce the workload of doctors.

4. Reference

[1]American Thoracic Society. (2019). Top 20 Pneumonia Facts. https://www.thoracic.org/patients/patient-resources/resources/top-pneumonia-facts.pdf

[2]Goyal, S., Singh, R. Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques. J Ambient Intell Human Comput 14, 3239–3259 (2023). https://doi.org/10.1007/s12652-021-03464-7

[3] Li Xin, Chen Fan, Hao Haijiang, et al. Research review of deep learning in pneumonia detection [J]. Journal of Guilin University of Technology, 2020, 40(4).

[4]Singh, S., Kumar, M., Kumar, A. et al. Efficient pneumonia detection using Vision Transformers on chest X-rays. Sci Rep 14, 2487 (2024). https://doi.org/10.1038/s41598-024-52703-2

[5]Gabruseva T, Poplavskiy D, Kalinin A. Deep learning for automatic pneumonia detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020: 350-351.

[6] Mehmood M H, Hassan F, Rahman A U, et al. CP_DeepNet: a novel automated system for COVID-19 and pneumonia detection through lung X-rays[J]. Multimedia Tools and Applications, 2024: 1-18.

[7] Ali M, Shahroz M, Akram U, et al. Pneumonia Detection Using Chest Radiographs With Novel EfficientNetV2L Model[J]. IEEE Access, 2024.