폐렴 이미지 분류를 위한 Swin과 ResNet 모델 비교

Comparing Swin and ResNet model in Determining Chest Pneumonia

Alabdulwahab Abrar
Department of AI Software
Gachon University
Seongnam, S. Korea
abrarssw@gachon.ac.kr

Sang-Woong Lee *
Department of AI Software
Gachon University
Seongnam, S. Korea
slee@gachon.ac.kr

Abstract

Pneumonia is a respiratory disease that causes infection in both the upper respiratory tract and the lungs. It is considered one of the leading causes of infection-related deaths in children. Chest X-ray images have proven helpful in diagnosing pneumonia. It is essential for early diagnosis of pneumonia to control the spread of the disease and save the patient. Therefore, there is a need for deep learning artificial intelligent systems to assist clinicians in early and better diagnosis. In this study, Residual Neural Network (ResNet) and Swin Transformer are used to classify pneumonia and healthy chest X-ray images from the Chest X-Ray Images dataset. Experimental results show that the ResNet achieved a maximum accuracy of 99.00% in detecting pneumonia after ten epochs. Whereas the Swin transformer achieved a maximum accuracy of 98.46% in detecting pneumonia after ten epochs.

Keywords: ResNet, Swin Transformer, deep learning, pneumonia detection, chest X-ray.

1. Introduction

Residual Neural Network (ResNet) and Swin Transformer have proven effective in disease diagnosis and classification [1,2]. They can differentiate between normal chest X-rays and those indicating pneumonia.

Pneumonia is a respiratory infectious disease that affects the lungs and can cause death, particularly in children. It involves alveoli -the small sacs in the lungs- being filled with pus and fluid, making breathing painful and hindering oxygen intake. It is one of the leading causes of infection-related deaths in children.

Pneumonia can be diagnosed by radiological imaging, such as X-rays and computed tomography (CT). Chest X-ray images usually show blurry parts of the bronchial tubes and alveoli, indicating suspected pneumonia. Chest X-ray images from different databases have been accumulated over the years to help detect pneumonia and further improve the deep-learning model performance. Recently, studies on detecting and predicting diseases with deep learning have been thoroughly conducted in Computer Aided Diagnosis (CAD) and obtained good performance in several medical fields, including radiology [3]. It is used to help medical professionals to save time, and effort, and improve accuracy in pneumonia diagnosis. Recent studies showed that ResNet and Swin Transformer give various degrees of accuracy in diagnosing pneumonia [4,5]. In fact, it is unclear which of them is superior to the other. Therefore, this study is aimed to compare the performance of ResNet and Swin Transformer in reading chest X-rays and diagnosing pneumonia. Figure 1 depicts the proposed scheme of the experiment[6].

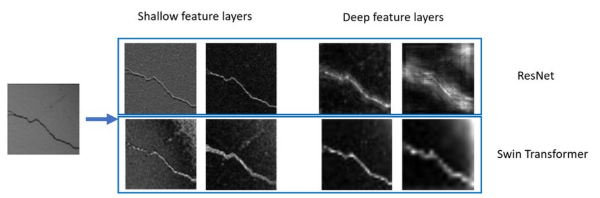

Figure 1 ResNet and Swin extract the feature.

2. Related works

ResNet and Swin Transformer have been intensively used to improve pneumonia diagnosis [1,2]. In one study, ResNet is employed to classify X-ray chest images into three classes: Normal, COVID-19, and pneumonia. The study have shown that ResNet performs well on the image recognition task and were helpful in pneumonia diagnostics [7]. While Swin Transformer is a type of neural network architecture that produces a hierarchical feature representation and has linear computational complexity with respect to input image size [2]. It has been used in the field of medical imaging to identify pneumonia [3].

The ResNet yielded an accuracy of 78.73%, for classifying an X-ray image on its possibility of exhibiting the early stages of pneumonia [8]. ResNet reported classifying pneumonia better than other backbone networks [4]. Moreover, integrating ResNet with the empirical wavelet transform and temporal convolutional neural network produced 99.5% accuracy in classifying pneumonia [5].

Furthermore, the Swin Transformer backbone network optimized for chest X-rays achieved an accuracy of 97.2% in detecting pneumonia after five epochs of training [3].

Therefore, it is obvious that both ResNet and Swin transformer are efficient in detecting pneumonia chest images. However, it is unclear which is the best for detecting pneumonia. Hence, this study will compare the performance of ResNet and Swin transformer in classifying chest pneumonia.

3. Methods

3.1. Dataset

The dataset used is released from Guangzhou Women and Children’s Medical Center, Guangzhou and available in Kaggle[8]. Figure 2 depicts examples of the dataset. The dataset contains 5,863 children's X-Ray chest images saved in JPEG format. The dataset includes two categories, “normal” and “pneumonia”.

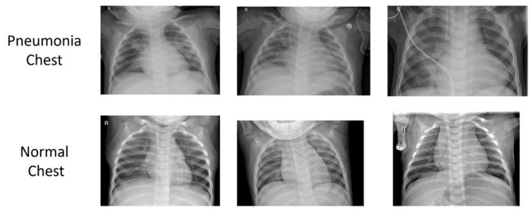

Figure 2 Data Samples

3.2. Experiment setup

Both ResNet and Swin were given 512 x 512 pixel input images. The dataset is split into 3:1 ratio for train and test. The batch size is set for 4. ResNet used a LR of 1e-4, while Swin used 1e-6. Both models were trained for ten epochs.

4. Experiment result

During this experiment, both ResNet and Swin demonstrated high levels of accuracy. However, according to the loss function scores ResNet model achieved lowest loss of 0.0001 with an accuracy of 99.00% in the tenth epoch, while Swin got the lowest score of 0.002 with an accuracy of 98.31% in the nineth epoch. Table 1 depicts the results of all the test data loss per epoch. One reason for this difference between ResNet and Swin performance is that Swin may attribute to the factors such as the dataset used, the model configuration, and the complexity of the task [2].

Table 1 ResNet and Swin loss score

Epoch	ResNet	Swin
1	0.003	0.99
2	0.01	0.05
3	0.004	0.019
4	0.0008	0.018
5	0.008	0.005
6	0.001	0.004
7	0.002	0.02
8	0.0007	0.0027
9	0.006	**0.0020**
10	**0.0001**	0.02

5. Conclusions

From the experiments above we can safely conclude that both ResNet and Swin can accurately classify pneumonia from X-ray images. In our results ResNet outperformed Swin in terms of accuracy. This may be due to factors such as the dataset used, the model configuration, and the complexity of the task. To find the exact reason for these differences we are currently working on additional experiments with different configurations.

Acknowledgement

This work was supported by the GRRC program of Gyeonggi province. [GRRC-Gachon2020(B02), Analysis of AI-based Medical Image].

References

1. Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.
2. Ma Y and Lv W. “Identification of Pneumonia in Chest X-Ray Image Based on Transformer.” *Hindawi International Journal of Antennas and Propagation* Vol. 2022, Article ID 5072666
3. Wang, Xiaosong, et al. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2017.
4. El-Dahshan, El-Sayed A., et al. “RESCOVIDTCNnet: A Residual Neural Network-Based Framework for COVID-19 Detection Using TCN and EWT with Chest X-Ray Images.” *Expert Systems with Applications*, vol. 204, no. 117410, 2022, p. 117410.

5. Zhou, Zhong, et al. “Hybrid Semantic Segmentation for Tunnel Lining Cracks Based on Swin Transformer and Convolutional Neural Network.” *Computer-Aided Civil and Infrastructure Engineering,* 2023
6. Ansari, Naseem, et al. "Effective Pneumonia Detection using Res Net based Transfer Learning." *Test Engineering and Management*, vol. 82 , 2020, pp. 15146-15153.
7. Saul, Can Jozef, et al. Early Diagnosis of Pneumonia with Deep Learning. 2019.
8. Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification*. Mendeley data*, 2(2), 651.