**Benchmark Test for Pneumonia Classification Solutions**

Yong Woo Kim\*

1. Introduction

In medically underserved regions, the accessibility to advanced diagnostic tools becomes crucial for improving healthcare outcomes. Pneumonia, a prevalent respiratory infection, poses a significant challenge in such areas where limited medical resources may hinder timely and accurate diagnoses. The development of an efficient pneumonia classification solution holds promise in bridging this healthcare gap by providing valuable support to medical professionals.

The significance of an accurate pneumonia diagnostic solution lies in its potential to enhance the effectiveness of medical professionals in identifying and treating respiratory conditions. By leveraging advanced deep learning techniques, these solutions can analyze medical imaging data with precision, aiding medical professionals in making timely and informed decisions. This not only contributes to the overall efficiency of healthcare delivery but also addresses the pressing need for accessible and reliable diagnostic tools in regions with limited healthcare infrastructure.

This paper aims to present a comprehensive benchmark test of pneumonia classification solutions, evaluating their performance against established standards. By exploring and comparing the capabilities of these solutions, we seek to contribute valuable insights to the ongoing efforts in leveraging artificial intelligence for medical diagnosis. The goal is to demonstrate the potential impact of pneumonia diagnostic solutions in supporting medical professionals, particularly in medically underserved areas, where accurate and prompt diagnosis is important for improving patient outcomes.

2. Dataset

The dataset comprises approximately 5000 chest X-ray images, including both pneumonia-infected and non-infected individuals. This dataset is governed by the CC BY 4.0 license, allowing for unrestricted sharing, replication, modification, and distribution of the dataset. Adherence to the terms of the CC BY 4.0 license, including proper attribution and acknowledgment of the data source, permits researchers to freely utilize and build upon this dataset for various purposes. **Figure 1.** is example of dataset.

엑스레이 필름, 의료 영상, 방사선과, 방사선 촬영이(가) 표시된 사진

자동 생성된 설명

Figure 1. Dataset example

3. Experiments

**3-1. Use case**

If a doctor classifies and diagnoses for abnormal findings using the solution before diagnosing a disease by looking at an X-ray of the patient's chest, we think it will help doctor diagnose effectively and reduce diagnostic deviations in underserved areas.

**3-2. Anchor solutions**

We have three anchor solution. To benchmark the performance of our proposed pneumonia classification solutions, we compared it against three anchor solutions widely recognized for their excellence in the field. These anchor solutions were selected based on their prominence in the data science community.

1. **"Deep Learning-Based Pneumonia Classification Model Research Using Defense Medical Data"**

Source: DBpia – Research on a classification model of deep learning-based pneumonia using defense medical data.

Brief Description: This study, recognized as the top-cited paper on Korean papers, investigates a deep learning-based pneumonia classification model utilizing defense medical data.

2. **"Top-Rated Kaggle Solution for Pneumonia Classification"**

Source: Kaggle Data Analysis Community - Pneumonia Classification

Brief Description: It is the most accurate solution among the top 5 recommended solutions among kaggle deep learning solutions for pneumonia classification.

3. **"Leading GitHub Starred Pneumonia Classification Solution"**

Source: GitHub - Pneumonia Classification Model

Brief Description: Among GitHub's deep learning solutions for pneumonia classification, it ranks first in accuracy among the top three and top five recommended solutions.

These anchor solutions serve as robust benchmarks, allowing for a comprehensive evaluation of the proposed solution’s efficacy in pneumonia classification.

엑스레이 필름, 스크린샷, 방사선과, 의료 영상이(가) 표시된 사진

자동 생성된 설명

Figure 2. Solution flow chart

**Table 2.** is also brief description of solutions. Based on the solution presented in the paper, the Defense Medical Solution implemented PyTorch to fit the train data. The other two solutions used the code that was implemented.

|  |  |  |  |
| --- | --- | --- | --- |
| **Solutions** | **Defense Medical**  **Solution** | **Kaggle**  **Solution** | **GitHub Solution** |
| Framework | PyTorch | Tensorflow | Tensorflow |
| Backbone | ResNet34 | DenseNet121 | InceptionV3 |

Table 1. Solutions

**3-3. Evaluation**

We evaluated accuracy, F1 Score, recall and processing time. F1 score was also considered because the validation data were biased to 390 chest X-ray photos of patients with pneumonia and 234 chest X-ray photos of non-pneumonia. Processing time was all measured with the maximum use of memory on the T4 GPU.

**3-4. Results**

Accuracy, F1 score, Recall, and Processing timing for validation data in Anchor solutions are in table 2. F1 Score was also considered because the validation data were somewhat biased to 390 chest X-ray photos of pneumonia patients and 234 chest X-ray photos of non-pneumonia patients. Processing time was all measured using memory to the maximum on the T4 GPU.

|  |  |  |  |
| --- | --- | --- | --- |
| **Solutions** | **Defense Medical**  **Solution** | **Kaggle**  **Solution** | **GitHub Solution** |
| Accuracy | 97.41% | **97.89%** | 97.51% |
| F1 score | 98.38% | **98.58%** | 98.13% |
| Recall | 98.83% | 98.83% | 98.83% |
| Processing Time | 4312ms | 5023ms | 4882ms |

Table 2. Result

The Kaggle Solution was the highest in Accuracy, F1 score, and Recall. As for the processing time, the Kaggle Solution with the highest parameter in the backbone network was the largest.

**3-5. Demonstration**

We created demo to make the solution available directly to Kaggle Solution, which performed the best. **Figure 3.** is the screenshot of demo.

엑스레이 필름, 의료 영상, 방사선과, 의료 방사선이(가) 표시된 사진

자동 생성된 설명

Figure 3. Demo

**4. Conclusion**

Benchmark tests on three Anchor solutions confirmed that Kaggle solutions had the highest accuracy, F1 score, recall.

If you use this solution in conjunction with Vuno's mobile X-ray imaging device, it will help you identify patients in medically underserved areas.

To make further suggestions for our solution, the result value of the current solution derives the result value as having or not having abnormal findings. However, we think that deriving the result of the solution as a probability value will be more helpful for the doctor's sophisticated diagnosis. All three of our current solutions derive one result (one of the abnormal findings or no abnormal findings) by selecting the maximum value among the results that come out as probability values through the softmax function. At this time, if the process of selecting the maximum value is eliminated so that the result comes out as a probability value, the suggestions mentioned above can be easily implemented.

REFERENCES

[1] Tae-Hwan Lim, & Seung-Chul Han (2021). Research on a diagnostic model of deep learning-based pneumonia using defense medical data. *Journal of Digital Contents Society*, *22*(3), 509-517, 10.9728/dcs.2021.22.3.509

[2] Vuno demonstrates medical artificial intelligence solution at vision declaration ceremony to strengthen AI/SW capabilities of military personnel, BioTimes, 2022.07.05

[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).