Pneumonia detection

Emma Carlsson, Roman Ermolaev, Oskar Karlsson

Luleå University of Technology, Sweden
emacar-8@student.ltu.se
romerm-8@student.ltu.se
oskkar-8@student.ltu.se

***Abstract*—The aim of this project was to identify pneumonia infection in chest x-ray images using deep learning methods. The chosen method was to apply transfer learning on a CNN. The architectures ResNet50 and ResNet152 were pretrained on ImageNet with different optimizers and learning rates. This yielded results of 77% accuracy and 0.84 in F1-score with ResNet50 using Adam as optimizer and a learning rate of 0.001.**

I. Introduction

This project was assigned as part of the course D7047E where the the aim was to identify Pneumonia with the help of deep learning methods focusing on computer vision. Pneumonia is an infection that causes inflammation in one or both of the lungs and may be caused by a virus, bacteria, fungi or other germs. A physician may conduct a physical exam and use chest x-ray, chest CT, chest ultrasound, or needle biopsy of the lung to help diagnose the condition. Typically, An x-ray image exam will allow a radiologist to check on lungs, heart, and blood vessels to help determine if one have pneumonia. When interpreting the x-ray, the radiologist will look for white spots in the lungs (called infiltrates) that identify an infection. Such exams will also help determine if one have any complications related to pneumonia such as abscesses or pleural effusions (fluid surrounding the lungs). Having a neural network that could do this instead of a human opens the possibility of helping a radiologist determine if a patient has pneumonia more efficiently and more accurately. For this project pretrained ResNet architectures were chosen as the method of identifying Pneumonia.

II. Experimental Setting

A. Task

The task was to use deep learning methods to build a binary classifier capable of classifying chest x-ray images as either Normal or Pneumonia with an accuracy of at least 70%.

B. Datasets

The pneumonia dataset was collected from Guangzhou Women and Children's Medical Center, Guangzhou, by Kermany et al. from the University of California San Diego [1]. It contains 5860 images of chest x-rays divided into two categories: Normal and Pneumonia. The images are divided into three parts; *train*, which contains 1342 Normal images and 3876 Pneumonia images for a total of 5218 images; *validation*, which contains 18 images, 9 of each class; and *test*, which contains 234 Normal images and 390 Pneumonia images for a total of 624 images. The gray-scale images were of various sizes and of patients of one to five years old. The images are from the patients anterior and all images have the same rotation. Before being compiled the x-rays were screened and their diagnoses were graded by professional physicians.

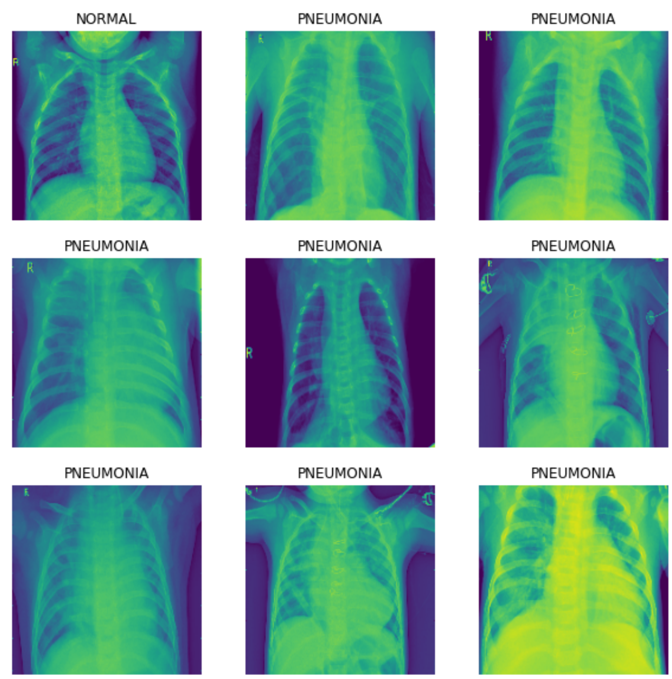

Fig. 1. Nine sample images showing one Normal and eight Pneumonia x-rays.

C. Model Architecture

The models used for comparisons were the ResNet50 and ResNet152 architectures [2].

D. Training Procedure

To begin, the model was pretrained on ImageNet. Then transfer learning was used to retrain the model. This was done by feature extraction where the convolutional base was frozen and then a new classifier was added on top of the pretrained model and trained it on the pneumonia dataset. The loss function used was Binary Cross Entropy. As optimizer functions both Adam and Rectified Adam was used and compared. All this was implemented in Jupyter Notebooks with the help of the Tensorflow libraries.

III. Results

The results presented in tables I and II were reached by training the models for 10 epochs on the whole of the train-set in smaller batches. The times reported in table II are all from the same computer with the code running on CPU.

TABLE I
TESTING SCORES FOR ACCURACY AND F1-SCORE FOR DIFFERENT COMBINATIONS OF MODEL ARCHITECTURE, LEARNING RATE AND OPTIMIZER.

	Optimizer	LR	Accuracy	F1-Score
ResNet50	RAdam	0.01	74%	0.82
ResNet50	RAdam	0.001	74%	0.82
ResNet50	RAdam	0.0001	65%	0.78
ResNet50	Adam	0.01	75%	0.83
ResNet50	Adam	0.001	77%	0.84
ResNet50	Adam	0.0001	69%	0.79
ResNet152	RAdam	0.01	70%	0.81
ResNet152	Adam	0.01	70%	0.80

TABLE II
TIME REQUIRED FOR TRAINING THE MODELS WITH DIFFERENT DEPTHS AND OPTIMIZERS.

	Optimizer	Training time
ResNet50	RAdam	19 min
ResNet50	Adam	18 min
ResNet152	RAdam	87 min
ResNet152	Adam	73 min

IV. Analysis

ResNet50 performed better than ResNet152 in terms of accuracy, F1-score, and training time required to train it. While ResNet50 was able to reach an accuracy of 77%, and F1-score of 0.84 using Adam with a learning rate of 0.001, ResNet152 only got a highest accuracy of 70% with a F1-score of 0.80, see table I. As shown in table II, the time required to train ResNet50 was slightly lower than 20 minutes while the ResNet152 took roughly 80 minutes.

One could also see the difference in performance between the optimizers. In terms of accuracy and F1-score Adam performed better in all tests pertaining to ResNet50. In terms of training time Adam was slightly faster than Rectified Adam. For ResNet152 the results were approximately the same, Adam and Rectified Adam produced the same accuracy, but Adam resulted in a slightly higher F1-score. In terms of training time Adam was notably faster than Rectified Adam.

An important observation of the results is that that there are more tests done on ResNet50, which introduces the risk of missing an important discovery, this was done because of the limited computational resources available. Since the ResNet152 was considerably more computationally expensive, and with limited time to finish the project, fewer tests were made with that architecture. Therefore, when observing the results this should be taken into consideration.

V. Conclusion

In conclusion, our results indicate that for a simple problem, it is better to use a less complicated solution. In this case that meant a standard pretrained ResNet50 with Adam as optimizer rather than the deeper ResNet152 or more complicated optimizer Rectified Adam.

The takeaway message from this article should be that a simpler problem often requires a simpler network.

References

[1] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," CoRR, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385