Pneumonia classification using convolutional neural networks

Olle Dyberg*, **Alex Jensen Sasafuchi***

*Luleå University of Technology, Sweden

Abstract—**In this article, we present a project aimed at developing a Convolution Neural Network (CNN) for the detection of pneumonia from X-ray images. In this project we made use of both a CNN, trained from scratch and compared it to a pre-trained ResNet18 model. Both the CNN and ResNet18 model achieved a test accuracy above 80%. This finding highlights the efficiency of transfer learning in this particular subject.**

I. Introduction

Pneumonia, characterized by inflammation of the lungs, affects millions of individuals worldwide, presenting a significant challenge for medical professionals. Early and accurate diagnoses are crucial for improved patient outcomes. While X-ray imaging has long been a standard tool for pneumonia assessment, it heavily relies on the expertise of radiologists to correctly interpret the X-ray images. This manual process can be subjective, and prone to human error, especially when dealing with a high volume of cases. With the advent of artificial intelligence (AI) this manual process can be outsourced to deep learning models specifically designed to excel at image classification. Convolutional Neural Networks (CNNs) are a class of deep learning models that excel at precisely this, which enables them to make accurate classifications based on patterns and characteristics that might escape the human eye. In this paper we will present our CNN, that is trained to classify cases of pneumonia based on X-ray images.

II. Theory

A. Convolution neural networks

Convolutional neural networks, also known as CNNs are a type of neural network that make use of the convolution operation rather than matrix multiplication which is used in other neural networks. CNNs are used when dealing with data where spatial properties are of interest, such as images and time series. Each layer in a CNN is a rectangular grid of neurons where the convolution of one layer will be the input for another. The convolutional operation is defined as

$$\int x(a)w(t-a)da \tag{1}$$

where $x(t)$ is the input and $w(t)$ is the kernel [1].

B. Residual networks

Residual networks (ResNets) are a type of neural network architecture that add the output of one layer to the input of a subsequent layer. This process forms a shortcut connection, known as a residual connection which facilitates gradient flow during training and enables easier propagation through the network.

III. Experimental Setting

A. Task

In this project we wanted to achieve a relatively high accuracy (80% +) on a classification task involving x-ray images.

B. Datasets

We used an x-ray dataset containing pre-determined train, test and validation sets. The dataset has 5863 images and is divided into two classes labeled Pneumonia and Normal. The images are chest x-rays taken from Guangzhou Women and Children's Medical Center, Guangzhou of one to five year old patients [2]. The average size of the images is around 1000x1000 but it varies greatly.

C. Model Architecture

We chose to create our own simple convolutional neural network (CNN) with two convolutional layers and compare this to a fine-tuned ResNet18 model [3]. We chose the ResNet18 model since we are familiar with it's architecture and because it's the smallest version of the pre-trained ResNets.

D. Training Procedure

The dataset was first augmented using some transformations to help with training. Firstly all the images were downscaled to 512x512 in order to tackle the problem of them being different sizes. Then some random horizontal and vertical flips were applied to make our model more robust and generalized. The model was trained on 10 epochs with an initial learning rate of 0.001, decaying exponentially with a coefficient of 0.9 every epoch. We chose a small batch size of 10 in order to increase the generalization of the model and reduce over-fitting. We also introduced weight decay in order to generalize the model further. The validation and training loss was plotted for every epoch and is shown in section IV.

IV. Results

The test set accuracy achieved with the final model is 84% which exceeded our goal of eighty percent. The training and validation set losses can be seen in FigureS 1 and 3. The fine-tuned ResNet18 model reached an accuracy of 83% and the training/validation loss can be seen in Figures 2 and 4.

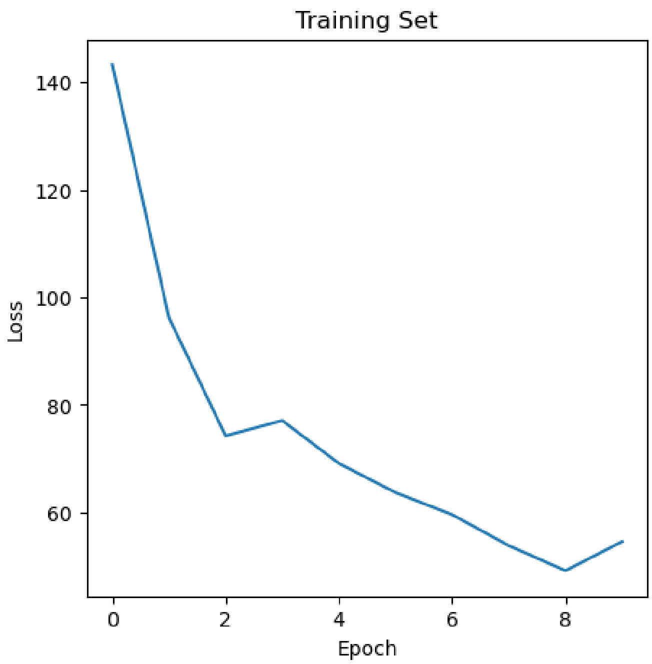

Fig. 1. Training loss of the custom CNN every epoch of training.

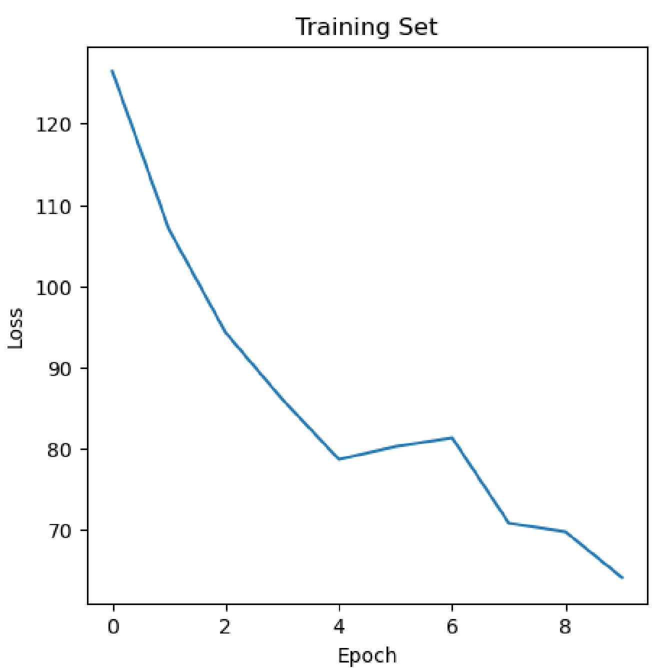

Fig. 2. Training loss of the ResNet18 model every epoch of training.

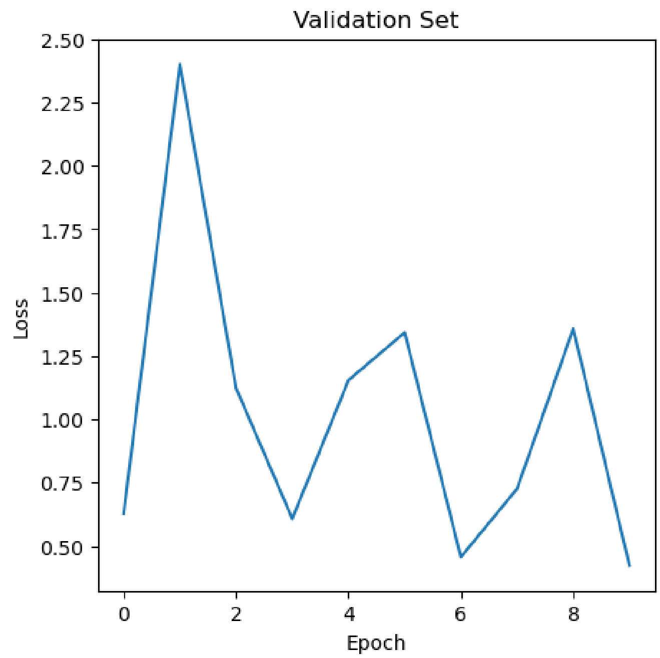

Fig. 3. Validation loss of the custom CNN every epoch of training.

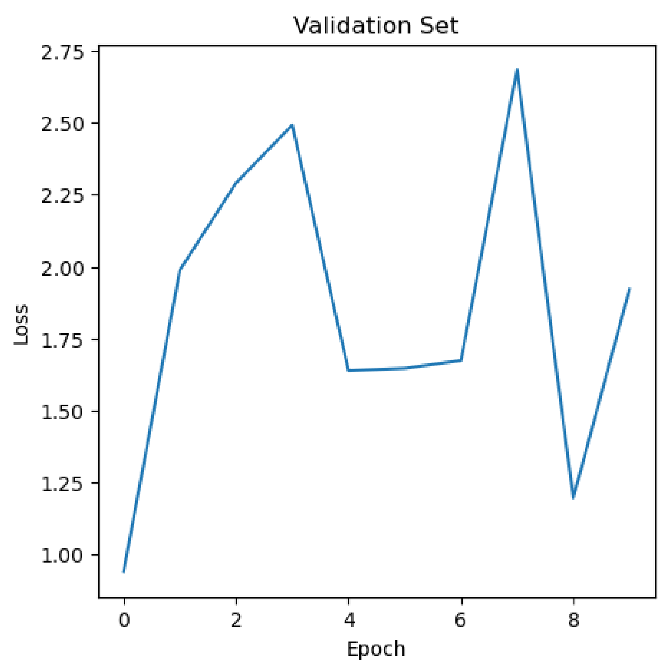

Fig. 4. Validation loss of the ResNet18 model every epoch of training.

V. Analysis

That the two models we looked at came out at around the same accuracy was surprising to us. We expected the fine-tuned model to do worse. This goes to show how well feature layers inside of pre-trained networks can generalize to other problems. To improve the accuracy we could alter the hyperparameters of the network, like the amount of layers of feature maps. We could also have increased the epochs, since it looked like the validation set loss hadn't converged yet. However, we chose not to continue improving our network after we got an accuracy above 80%.

VI. Conclusion

We have successfully built two models which achieve high accuracy in a classification task involving x-ray images of different sizes. One model was custom built from scratch with few layers and the other is a fine-tuned ResNet18 model. We learned the power of using pre-existing models and how simple models can lead to already good results!

References

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[2] K. G. M. Kermany, Daniel; Zhang, "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," in *Mendeley Data, V2, doi: 10.17632/rscb- jbr9sj.2*, 2018.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015.