ISSN 2350-109X
www.indianscholar.co.in

Indian Scholar

An International Multidisciplinary Research e-Journal

PNEUMONIA DETECTION USING TRANSFER LEARNING AND CONVOLUTIONAL NEURAL NETWORKS

Aarav Jain
Prabhavati Padamshi Soni
International Junior College
aarav8304@gmail.com

Abstract

This paper outlines the advantages of using transfer learning on pre-trained convolutional neural networks and the problems transfer learning solves for Pneumonia detection. Seven pre-trained models are used for Pneumonia detection along with one convolutional neural network trained from scratch. This paper also concludes why ResNet-50 [1] performs best for detecting pneumonia given an x-ray scan, and the reasons why this particular model works best with respect to its architecture, data it was initially trained with and the number of parameters the pretrained model has among many other things. ResNet-50 [1] achieves the highest validation accuracy of 93.27% with just 5216 files of image training data.

Keywords-*Transfer learning, Pneumonia, Convolutional Neural Networks, Image Classification, Resnet-50, Supervised Learning*

Introduction

Pneumonia killed more than 800,000 children in 2017 [2] who were under the age of 5. Pneumonia also accounts for 15% of all deaths of children under 5 years. It transmites in many ways, air-borne droplets from coughing and inhalation of bacteria found commonly on one's throat are just a few ways Pneumonia can spread. Additionally, Pneumonia is one of the most misdiagnosed conditions.

Given the high number of covid-19 cases, Pneumonia being caused due to covid-19 have also seen an alarming rise in recent times. To detect Pneumonia given an x-ray scan of the patient's lungs, radiologists search for white spots near the lungs which are called infiltrates. However, to speed up this process of searching for these infiltrates while reducing the error rate in detecting them, an image classifier made using convolutional neural networks could be trained to detect these infiltrates instead.

However, few problems arise. Often enough data is not available, especially tasks related to medical and healthcare. This lack of data, could lead to unacceptable accuracies and overfitting of the training dataset for detecting pneumonia using convolutional neural networks. Also, when high accuracies are needed, it is common to increase the total number of layers and hidden units of a neural network. This increases computational costs and time taken to train the neural network significantly. Hence, a solution to this problem is a technique called "transfer learning". A pretrained model is taken with its weights already set according to a dataset with a

large number of image files, such as the ImageNet dataset [3], and then training on our personalized dataset takes place on the last few layers of the neural network, which help classify the given input image into different classes. Of course, this idea of transfer learning could be applied to different problems in this domain apart from image classification too, such as image segmentation and object detection.

This method of diagnosis of Pneumonia can replace the need for patients to undergo several tests, such as Blood tests, Pulse oximetry, Sputum test, CT scans and Pleural fluid culture. Additionally, as these tests may not be needed to diagnose Pneumonia, this will speed up the diagnosis process, hence reducing the strain on hospital resources.

Theory

A convolutional neural network used for classification of an input image to a particular class can be thought of by dividing the neural network into two different sections: Feature learning and Classification. A convolutional neural network is used to respect the spatiotemporal features of an image, as flattening the image into a vector in the first layer itself, without performing convolutions, will lead to the loss of the spatiotemporal features.

One of the main problems with medical imaging analysis is the lack of data available. However, due to the lack of data it could be difficult to train the data without overfitting. Overfitting leads to low validation accuracy as it memorizes specific attributes from the training dataset which do not generalize well to the rest of the data. However, transfer learning avoids the problem of overfitting by training a network on a very large dataset, one which generalizes very well. Hence, we can use these layers which do not overfit for predictions for another dataset, leading to higher validation accuracies. This can be done as the first few layers of the neural network learn generic features which are commonly required by the network to detect and perform tasks. These generic features, for example, are lines, colours, edges and shapes amongst many others.

Using such a method has multiple benefits, such as faster training times, higher validation accuracies with a low data size and lower computational costs.

Firstly, using transfer learning, the model can be trained faster. The first few layers of the neural network do not require to be trained as they are pre-trained to recognize lower level generic features of an image like edges, curves and lines. Hence backpropagation computations do not need to be done on these layers of the neural network and hence as a result of smaller number of computations, the training time reduces.

Additionally, we can get very accurate results even if small amounts of data is available. Transfer learning reduces the problem of overfitting as it has already learnt to recognize generic features of an image such as lines, curves and edges and depending on the data the pretrained model was trained with, it would also have learnt to recognize higher level features of an image, such as being able to detect colours, shapes, patterns, sizes and even objects. This happens as the pretrained weights used to recognize these generic features are a lot more accurate in doing so due to the high dataset size which prevents from overfitting. Thus, by using these pretrained weights, we can customize the last few layers to our data as those layers usually learn features which are a lot more specific to a particular dataset and hence get more accurate results.

Lastly, convolutional neural networks can have many parameters that need to be trained, as a result, the computational costs are quite high. Transfer learning reduces the computational costs due as the pretrained parameters do not need to be trained. Typically, only the last layer of the neural network, which is used to generate a probability distribution for the possible classes that the input image could belong to, needs to be trained when we use transfer learning, hence avoiding the need to train multiple hidden convolutional and fully connected layers.

Experimental

Binary cross entropy was selected as the appropriate loss function for this task as there are only two class labels: Pneumonia and no Pneumonia.

The Adam optimization algorithm was used for the training of the neural networks due to the faster convergence of the neural network especially after the loss becomes lower compared to the results obtained using Stochastic Gradient Descent.

Eight different models were trained for this task, seven pretrained models (DenseNet [4], EfficientNet [5], InceptionNet [6], MobileNet [7], ResNet-50 [1], VGG16 [8] and Xception [9]) used by transfer learning whose weights were pre-trained on ImageNet [3] and one Convolutional Neural Network trained from scratch. The results below show which models give the most accurate results. The models were trained on one GPU on Google Colaboratory.

The dataset used was "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification" distributed by Mendeley Data [10]. The training dataset consists of 5216 files belonging to 2 classes (pneumonia and no pneumonia), while the validation dataset consists of 624 files belonging to 2 classes (pneumonia and no pneumonia). The neural networks were trained for 10 epochs. A batch size of 32 was selected. Additionally, all images were resized into 250 * 250 size with 3 channels.

However, due to the low amount of image files available, the model started to overfit very quickly, both for the pre-trained neural networks as well as the one trained from scratch. Hence, data augmentation was used to generate additional data which were modified using random rotation, width shift, brightness change, zoom, shear and vertical and horizontal flips. Also regularization such as Dropout with a rate of 0.4 and Batch Normalization was used.

Results

Training and validation loss along with training and validation accuracy for each of the eight neural networks were computed.

The results obtained clearly showed a vast improvement in detecting Pneumonia when transfer learning is used. After training the eight different neural networks, the best performance for neural networks trained using transfer learning have a validation accuracy on average 4.35% better than the neural network trained from scratch. It can also be observed using the help of the graphs below that the training of the neural network trained from scratch is a lot more unstable due to the high fluctuations of the validation loss.

The best performing neural network, Resnet-50 [1], had a validation accuracy of 93.27%, which is 7.21% higher than the neural network trained from scratch, which had a validation accuracy of 86.06%.

ISSN 2350-109X
www.indianscholar.co.in
Indian Scholar
An International Multidisciplinary Research e-Journal

Discussion

The results below are of image classification models which detects if a given x-ray scan has Pneumonia or does not have Pneumonia.

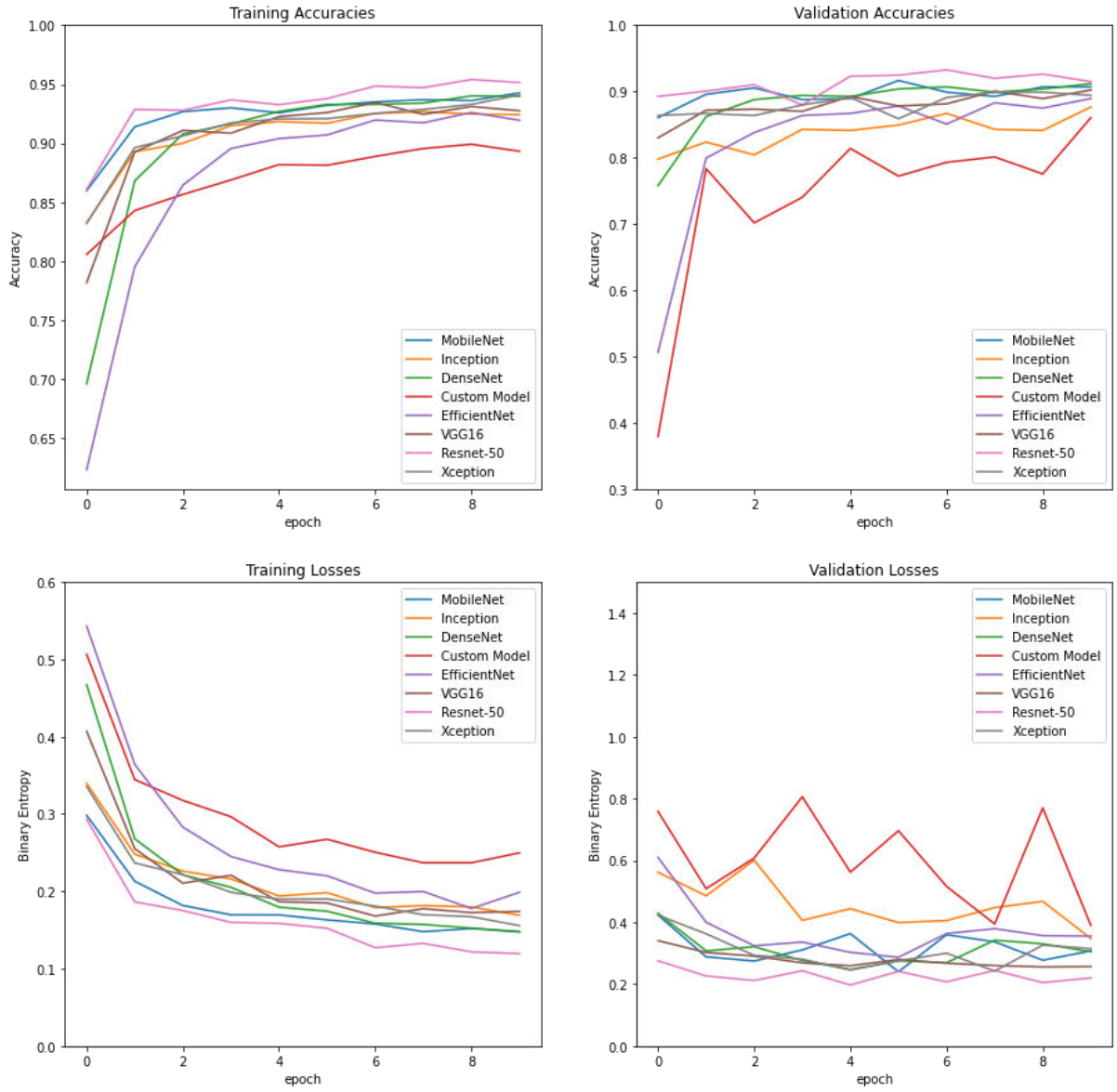


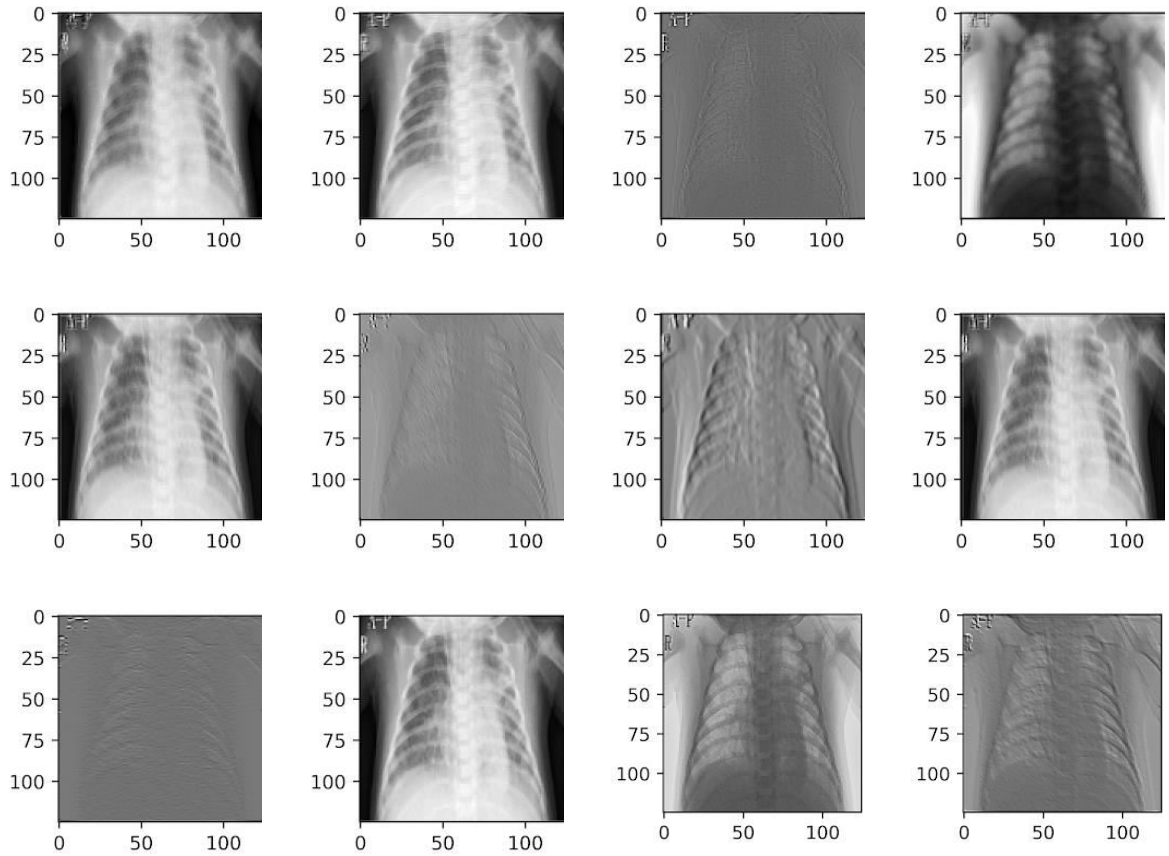
Fig. 1 Shows the results of training the eight different convolutional neural networks. The lower the (validation and training) losses, the better. Higher the (validation and training accuracies) the better.

ISSN 2350-109X
www.indianscholar.co.in
Indian Scholar
An International Multidisciplinary Research e-Journal

From the results above, it is clear to see that models trained using transfer learning have much better performance compared to a convolutional neural network trained from scratch. The neural network trained from scratch seems to start overfitting very quickly as well, as the validation accuracy is a lot lower than the training accuracy.

However, amongst all of the models, the ResNet-50 [1] model performs the best, achieving a validation accuracy of 93.27% on the 7th epoch. The MobileNet [7] model comes a close second, achieving a validation accuracy of 89.58% on the 10th epoch.

One of the main reasons that Resnet-50 [1] performs so well on the dataset is due to the number of parameters it has in its network. For example, Resnet-50 [1] comprises a total of 23,587,712 parameters, whereas EfficientNet [5] comprises a total of 17,673,823 parameters. A higher number of parameters that are pretrained leads to the neural network having learnt more features to detect in an image. Hence, Resnet-50 [1] learns features to detect the small infiltrates and increase in white density around the lungs of the x-ray scan image to classify whether a patient has Pneumonia or does not have Pneumonia. Additionally, due to the use of transfer learning using Resnet-50 [1], the neural network does not need to relearn all of its parameters, which not only would take a lot of time and be computationally expensive, but also would overfit due to the limited data. However, as the weights have been pre trained on a much larger dataset such as ImageNet [3], which has more than 14 million images, hence the neural network had not overfit on that.



ISSN 2350-109X
www.indianscholar.co.in
Indian Scholar
An International Multidisciplinary Research e-Journal

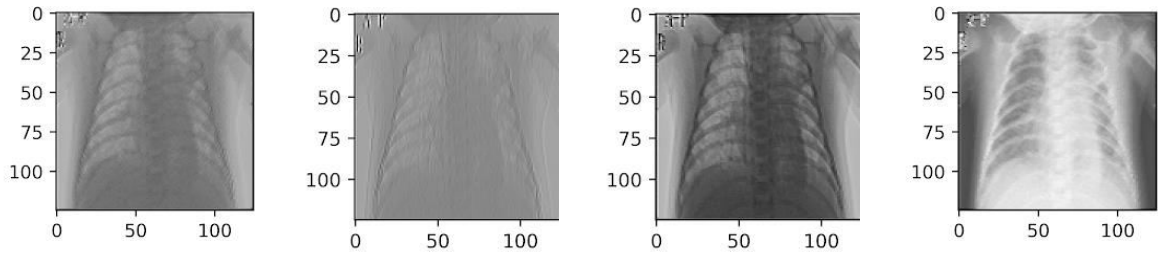


Fig. 2 Shows a visualization of what different convolutional layers in the neural network in Resnet-50 [1] learnt after training.

The pictures above show how the convolutional layers learn the different features needed to detect if the given x-ray scan has pneumonia or does not have pneumonia. As we can see, these are a few visualizations of the original image after the convolutional operations.

Conclusion

The use of transfer learning was critical in achieving such high validation accuracies with only 5216 images of x-ray scans available. With a small dataset, a neural network trained from scratch would overfit very quickly, and hence not allow it to converge to the global minima and keep a low training loss whilst maintaining a high validation accuracy. Additionally, as the weights from the pretrained network are freezed apart and training is carried out on the classification head, computational costs reduce as the trainable parameters have reduced in number. However, these results could further be improved by using a technique called "finetuning", where along with training the custom feedforward neural networks the layers of the original pre-trained neural network undergo training as well. Additionally, more data could be collected to gain higher validation accuracy. In the future, I plan to use transfer learning to achieve high accuracy and make it usable in the real world for instance segmentation on x-ray scans. This will help give doctors exact locations where the neural network predicts infiltrates are present that suggest Pneumonia.

Acknowledgements

I would like to thank the teachers from my school, Ms Varsha Kusnur, Ms Shanta Chowdhury, Mr Bhushan Pisat and Ms Ameena Fernandes for their guidance in the research paper.

References

1. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, June 27-30, 2016
2. "Pneumonia." World Health Organization. https://www.who.int/news-room/fact-sheets/detail/pneumonia (accessed July. 1, 2021)

3. J. Deng, W. Dong, R. Socher, L. Li, K. Li, L. Fei-Fei. "ImageNet: A large-scale hierarchical image database" August 18, 2009. Distributed by IEEE. 10.1109/CVPR.2009.5206848 .
4. G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, July 21-26, 2017
5. M. Tan, Q. V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". Arxiv. arXiv:1905.11946 (accessed June. 19, 2021)
6. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, June 7-12, 2015
7. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications". Arxiv. arXiv:1704.04861 (accessed June. 20, 2021)
8. K. Simonyan, A. Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition". Arxiv. arXiv:1409.1556 (accessed June. 29, 2021)
9. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, July 21-26, 2017
10. D. Kermany, K. Zhang, M. Goldbaum. "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification." January 06, 2018. Distributed by Mendeley Data, v2. http://dx.doi.org/10.17632/rscbjbr9sj.2