Enhanced Pneumonia Detection Using Transfer Learning With Fine Tuning In CNN

Sankhya Gudivada
Information Technology
Gayatri Vidya Parishad
College of Engineering
for Women
Visakhapatnam, India

Pediredla Sri Varsha
Information Technology
Gayatri Vidya Parishad
College of Engineering
for Women
Visakhapatnam, India

Abstract: Pneumonia is an infection that causes inflammation in one or both of the lungs. It causes the air sacs, or alveoli, of the lungs to fill up with fluid or pus. Pneumonia poses significant health risks, notably in India, where it contributes to 23% of the global burden with high mortality rates. As the diagnosis of pneumonia depends on skilled radiologists reading chest x-ray, delay in diagnosis, particularly in distant places, impede prompt treatment. The effectiveness of deep learning, in particular Convolutional Neural Networks (CNNs), in automating pneumonia detection is examined in this article. Through transfer learning we investigate pre-trained models such as VGG16, ResNet-50V2, InceptionV3, MobileNetV2, and Dense Net as feature extractors to improve the performance of the CNN architectures. Through experimentation and training on Chest X-ray dataset, the optimal CNN model for pneumonia diagnosis can be found by accurately classifying lung X-rays as normal or pneumonia infected

Keywords - Pneumonia, Chest X-ray, Transfer Learning, Convolutional Neural Networks (CNNs), VGG16, ResNet-50V2, InceptionV3, MobileNetV2, Dense Net.

I. INTRODUCTION

Pneumonia is a severe respiratory illness that typically affects the alveoli in the lungs. Pneumonia remains a global health concern. Its clinical arrangement consists of a variety of symptoms that differ in severity from person to person, such as cough, fever, discomfort in breathing, and other respiratory problems [1] [2]. While bacterial and viral infections are the key causes of pneumonia, other microbial sources, such as fungi, might trigger the onset of the illness. The different pathogenic pathways involved can lead to a proper diagnosis [1]. Physical scan results of lungs and auxiliary check-ups play a pivotal role in understanding lung health, which remains a global health concern in clinical practice to date. X-ray images are essential for finding out the involvement of the lung as they can identify prominent and well defined indicators of inflammation [3]. These learnings provide informative insights about the location of the parts affected by the disease and the dimensions of pulmonary illnesses. This helps professionals make accurate diagnoses and offer appropriate treatment [4]. Chest X-rays show opacities, which are commonly referred to as patchy or lobar infiltrates, and these are a sign of inflammation in the lung [5].

In addition, chest X-rays are necessary for evaluating the development of an infection and the efficacy of a treatment plan, permitting medical professionals to alter therapeutic actions as and when required [6]. Blood tests, such as comprehensive blood counts, support radiological analysis by helping to evaluate the body's immunological response and general state of health [7]. The effects of pneumonia go further away from individual health; they have quite a substantial impact worldwide in terms of the death rate in children.

Pneumonia is predominantly fatal in children under five years old and adds significantly to global death rates, emphasizing the need for reachable healthcare services and precautionary measures [8].

II. RELATED WORK

The usage of machine learning (ML) algorithms to categorize thoracic diseases has attained a lot of interest lately [9], predominantly in the area of medical image classification. The potential of deep learning approaches, specifically convolutional neural networks (CNNs), has been studied in this field in a number of publications [10]. Sundaram and Lakhani for example, in 2017, proposed the idea of employing deep convolution neural network (DCNN) architectures such as AlexNet and GoogleNet to detect Pulmonary Tuberculosis [11]. Huang et al. experimented on lung nodule classification for diagnosis of lung cancer using deep learning techniques [12].

Transfer learning (TL) techniques, which use CNN architecture have already been trained to increase the effectiveness in medical image analysis which is a substantial innovation in this field [13]. Numerous transfer learning architectures, including AlexNet, VGG, ResNet, and GoogleNet, have been studied by researchers in an effort to improve the accuracy of disease diagnosis [14]. According to the previous studies, DenseNet based models are more effective compared to the radiologists at diagnosing pneumonia. The same was demonstrated by Pranav Rajpurkar et al. (2017) [15].

Latest studies have emphasized the significance of transfer learning (TL) in overcoming challenges such as lack of training data and over fitting [16]. The ability of architectures such as VGG, InceptionNet, ResNet, and DenseNet to generalize well to medical imaging problems has led to their pervasive implementation [17]. VGG, for instance, is well-known for its composite model structure and ease of use, has validated favourable results in a variety of medical image classification tasks [18]. In a similar way, InceptionV3 has proven to provide improved computational efficiency and performance. This is done by employing its inventive inception component [19]. Furthermore, precision in the processing of medical images have been made possible by DenseNet's dense connectivity and ResNet's skip connections [20].

The Convolutional Neural Network is mainly utilized for image classification tasks. This is because, CNNs can effortlessly extract important features from the images and provide outputs based on these features. The CNN architecture consists of three layers as shown in fig (1). Namely, they are, the convolutional layer, pooling layer, and fully connected layer. Each of these layers has its own purpose in performing a specific task. The hidden layers are utilized for feature extraction and down sampling. Fully connected layers are used for flattening the output received by the previous layers. They have the ability to independently identify relevant features.
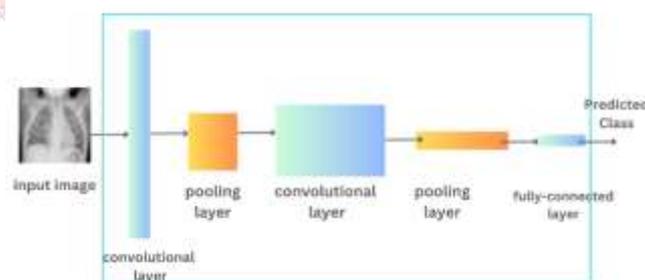

Fig (1): CNN

III. PROPOSED SYSTEM

Transfer learning is a method within Machine Learning (ML) where pre-trained models, equipped with prior knowledge are employed to tackle new and diverse problems. Instead of starting from scratch, they have randomly initialized weights, these pre-trained models serve as a foundation for classification tasks. Typically trained on extensive datasets, transfer learning can be adapted even for smaller datasets through techniques like data augmentation. This study utilizes three renowned pre-trained deep learning convolutional neural networks (CNNs), namely VGG16, ResNet-50V2, InceptionV3, MobileNet-V2, DenseNet for pneumonia detection.

ResNet-50v2

Resnet-50v2, an enhanced version of ResNet50, is utilized for pneumonia detection. It addresses issues of vanishing gradients and degradation, offering improved performance.

Vgg16

VGG, a CNN model, demonstrated an impressive accuracy of 83% when tested on over fourteen million images encompassing a thousand classes. Its simplicity lies in its use of 3* 3 convolutional layers stacked on top of each other, with size reduction facilitated by max-pooling, followed by fully-connected layers and Soft Max layer.

InceptionV3

Inception, a Google– developed convolutional neural network, integrates diverse filters like 1x1, 3x3, 5x5, convolutions for comprehensive image analysis. Its concatenated outputs form rich input for subsequent layers, crucial for accurate pneumonia detection.

MobileNetv2

MobileNetv2, designed with depth wise separable convolutions and linear bottlenecks, offers high efficiency and accuracy for pneumonia detection by efficiently extracting features while minimizing computational complexity. Its streamlined architecture facilitates swift processing of medical images aiding in prompt diagnosis.

DenseNet121

Densenet121 optimizes pneumonia detection by minimizing redundant feature maps and offering each layer direct access to original input and loss function gradient, thus enhancing computational efficiency and classification accuracy.

IV. DATASET

The dataset consists of 5836 chest X-ray images, divided into training and test sets, with both pneumonia and normal cases. The training set contains a little over 5000 images and the test set has 624. After splitting the train data, validation set has 1044 images, while the updated train set has 4172 necessitating shuffling for equitable distribution due to dataset's imbalance.
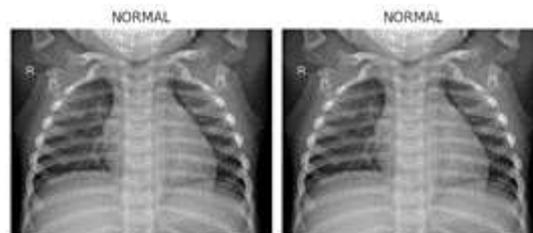

Fig (2) Chest x-ray images of Normal and Pneumonia affected.

| Category	Train	Test	Validation
	Train	Test	Validation
Normal	1072	234	269
Pneumonia	3100	390	775
Total	4172	624	1044

Table (1)

Data Pre-processing

Pre-Processing of the experimental dataset involved resizing and normalize chest x-rays images to meet input requirements of various neural networks, with dimensions ranging from 224x224 to 256x256 pixels. Each image underwent expert grading to ensure diagnosis accuracy before inclusion in the training process to ensure data quality. The dataset was then divided into a training set of 4172 images and a test set of 624 and validation set of 1044 images, with shuffling to achieve a balanced distribution.

Data Augmentation

Data Augmentation techniques, including Rotation Range = 10, Zoom range = 0.2, Shear range = 10, rescale = 1/255, horizontal flip = 10, Bright range = [0.5, 0.2], width shift range = 0.2, fill mode = nearest.
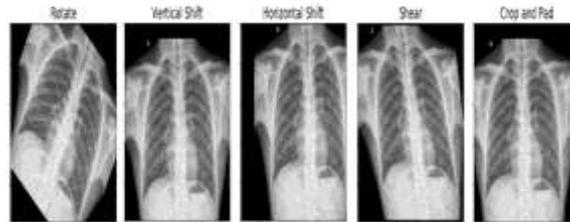
Fig(3)

V. ALGORITHM USED

Transfer learning (TL) is a machine learning (ML) technique that reuses information learned from tasks to improve the performance of related tasks. For example, for image classification, the knowledge gained when learning about cars can be used when trying to recognize trucks. This topic relates to psychological data on the change in learning, although the relationship between the two activities is limited. Learning efficiently may be greatly increased by applying knowledge from previously trained tasks to new ones.
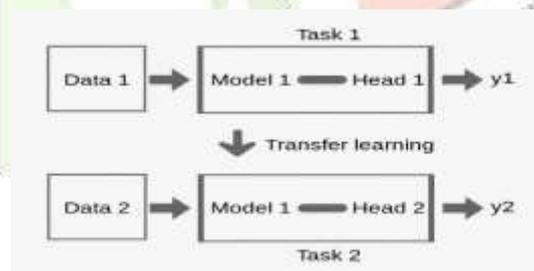
Fig (4) Transfer learning

Within the framework of this project, Transfer Learning was essential to the creation of precise and effective models for the diagnosis of pneumonia from chest x-ray images. This reduced the computational load of training sophisticated deep learning architectures from scratch by utilizing the representations gained by pre-trained models like InceptionV3, VGG16, MobileNetv2, Dense Net and ResNet50v2. Transfer learning was used to speed up the process of developing the models and make it easier to extract useful characteristics from the input photos, which improved the models' ability to predict outcomes.

ResNet50

ResNet-50v2 is a version of Microsoft Researcher's that integrates the modern developments in deep learning architecture and is intended to solve problems related to training neural networks which are less dense. ResNet-50v2's unique residual connections help in reducing the vanishing gradient challenge that often arises during training of deep networks. This makes it possible to create architectures that are prominently deeper. By using these residual or skip connections that wrap around some layers, it allows gradient to flow directly

during back propagation. Thus, ResNet-50v2 can train incredibly deep models without experiencing many failures. ResNet-50v2 was used in this study as one of the main architectures to extract features for the purpose of diagnosis pneumonia from chest X-ray images. Using the structured representations that ResNet-50v2 had learnt from a large dataset, the model performance best at identifying complex patterns that were suggestive of pneumonia compared to other models we have considered in this study. Custom dense layers added to the pre-trained Resnet-50v2 framework considerably improved the model's discriminative capability and generalization capacity by enabling fine-tuning.
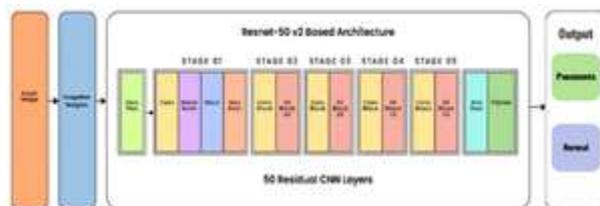

Fig (5) Architecture

ResNet-50v2's advanced architecture, with its deep layers and residual connections, excels in capturing intricate feature from chest X-ray images, enabling accurate and reliable detection of pneumonia, there by aiding in timely diagnosis and treatment.

VI. METHODOLOGY

This section provides the methodology for using a transfer learning framework to detect pneumonia from chest X-ray images. This section contains an overview of the techniques used to detect pneumonia. Here is a general outline of the steps that may be involved in this process:

A. Pre-Processing

The first step is to collect the dataset of chest X-ray images that contain both normal and pneumonia images. In the next step, we pre-process the dataset by resizing and cropping the images as needed. It may also be necessary to correct any errors or biases in the data.

They are divided into three parts: test, validation, and train. The original images are single-channeled, and are converted into 3-channelled images. In terms of pixel sizes, they are resized from 1024x1024 pixels to 224x224 pixels since processing pictures with a high resolution demands a lot of computer resources. This scaling procedure reduces computing costs and accelerates the subsequent stages of feature extraction and classification.

B. Model Architecture Selection and Training

The effectiveness of several pre-trained convolutional neural network (CNN) designs in pneumonia diagnosis was assessed. Of these designs, ResNet50v2 performed better in capturing relevant image characteristics; hence, it was the one that was chosen to undergo additional experimentation. The chosen ResNet50v2 model, pre-trained on the Image Net dataset, serves as the core for the pneumonia detection system.

Transfer learning is utilized to tune the ResNet50v2 model on the chest X-ray images specific to the pneumonia detection task. By freezing the weights of the pre-trained layers and adding fully connected custom layers for classification, the model is adapted to the target domain while retaining the knowledge learned from ImageNet dataset. With the use of methods like data augmentation, the modified ResNet50v2 model is trained on the augmented training dataset, improving model and generalization. Here, the model's performance depends on the parameters such as recall, precision, F1-score and ROC curve (AUC).Over-fitting can be avoided and efficiency can be maximized by applying learning rate scheduling and early stopping. Learning rate is set to 0.0010 and then reduced to half, that is, 5.0000e-04 after the first 5 epochs.

C. Model Validation and Testing

Model validation measures the performance of your model, so you can test and compare alternatives. The testing phase mostly deals with the testing data to evaluate proposed model using the trained model based on the test accuracy and other performance measures. Test data is unseen data, it is not used for training. The model performance is based on the test data and it includes various parameters namely recall, F1-score, precision and accuracy.

D. Fine Tuning

Fine tuning is process of taking pre trained model, which already worked on large dataset for general tasks like image classification. In fine tuning we make some modifications to our pre trained model as per our task. Depending on the complexity of our task, we freeze some layers in model and add some custom layers. The first 30 layers of the pre-trained model used are frozen, number of epochs are set to 20 and learning rate remains 0.0010. The main aim of fine tuning is to improve the model performance and give better results.

VII. RESULTS

We have assessed the performance of five well recognized convolutional neural network (CNN) architectures, namely ResNet50v2, MobileNetv2, DenseNet121, VGG16, and Inceptionv3, on an image classification task. In this research we applied a dataset of chest x-ray images that have been divided into two classes to the CNN architectures mentioned above. A portion of the dataset was used for training, while test and validation sets were utilized to determine the performance and generalization of the models.
The training results show how well each CNN architecture performs when it comes to image classification. ResNet outperformed the other designs in terms of test (92%), train (94%), and validation (93.30%). These models performed well in terms of recall, F1-score, and precision, demonstrating their ability to classify images accurately while reducing false positives and false negatives. The models performances improved after fine tuning, for instance, ResNet's accuracy has reached to 95.13 % for train data and 93.26 for test data. Below is the confusion matrix for Resnet model.
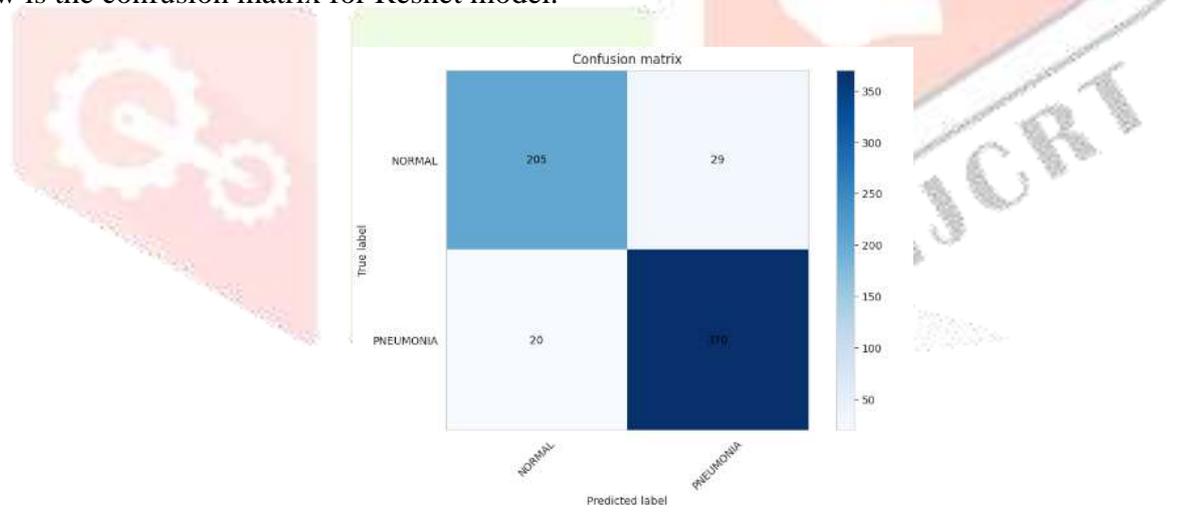

Fig (6) : Confusion Matrix for ResNet50v2

$$Accuracy = (TP+TN)/ (TP+TN+FP+FN)$$

Since the work aims to detect pneumonia based on chest X-ray, it is vital to reduce the model's false-negative rate. It is vital for such as a model to accurately detect which people are suffering from pneumonia, thus ensuring timely treatment. Recall's formula is:

$$Recall = TP/ (TP+FN)$$

Improving the model's recall decreases the false-negative rate. Thus for detecting pneumonia based on chest X-ray recall value closer to 1 ensures better performance. The precision of a model is:

Precision = TP/ (TP+FP)

However, it is crucial to take into account the trade-offs between performance and model complexity Resnet and Inceptionv3, while achieving great accuracy, also need more processing power than other architectures. But Mobilenet, with its lightweight architectures designed for embedded and mobile applications, performed admirably on the tests, with an accuracy of 90.38% after fine tuning. Additionally displaying worthy performance were DenseNet and VGG16, with test accuracies of 92.94% and 87.66%, respectively.

Four out of five architectures achieved test accuracies above 87%, indicating that the model's resilience was further substantiated by the evaluation on the test set. Furthermore, the model's ability to distinguish between the two classes is measured by the AUC values, which range from 92.42% to 97.78%. These scores were consistently high across all architectures.

Model	Accuracy	Precision	Recall	AUC
ResNet50V2	**0.9326**	**0.8189**	**0.8333**	**0.9777**
MobileNetV2	0.9038	0.8125	0.7781	0.9721
DenseNet	0.9282	0.7980	0.7852	0.9700
VGG16	0.8766	0.7532	0.7291	0.9612
InceptionV3	0.8157	0.7179	0.6666	0.9242

Table (2) : Metrics after fine tuning across all the models
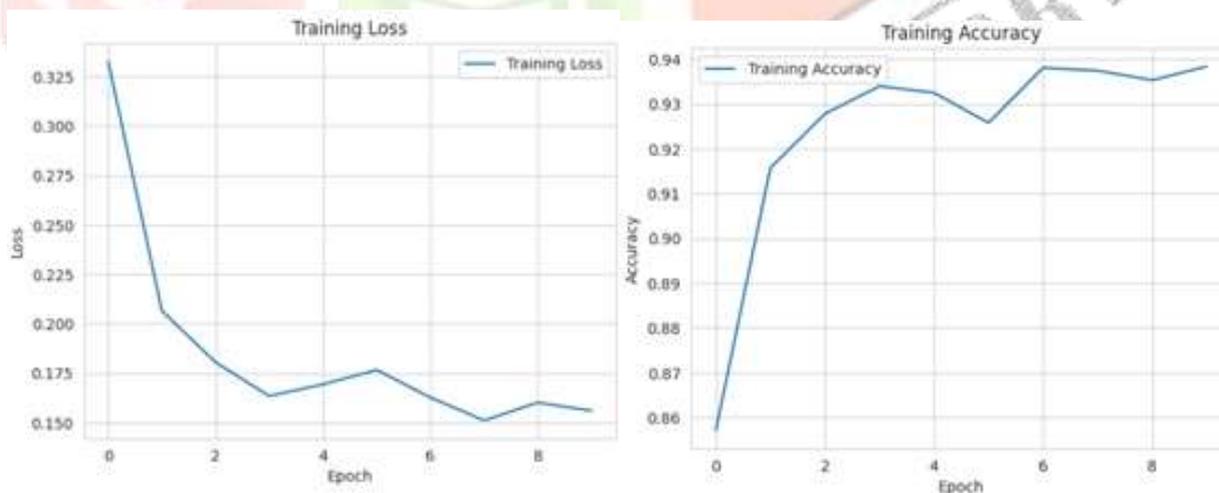
Fig(7) : Loss and Accuracy during model Training

VIII. CONCLUSION

Our research concludes by demonstrating the adaptability of CNN architectures of image classification applications. Although resnet and InceptionV3 are notable for their exceptional accuracy, other architectures, like Mobilenet, Densenet, and VGG16, provide good substitutes based on deployment needs and computing limitations. All things considered, these results offer insightful information about how to choose the best CNN

architecture for a given application. To increase performance and resilience, more study may look into model ensembles and fine-tuning strategies

IX. UPCOMING WORK

The next stages in developing convolutional neural networks (CNN) for classification of images in practical applications entail optimizing pre-trained models to perform well in certain fields, such as autonomous driving or medical imaging. We can push the limits of classification robustness and accuracy by investigating ensemble learning techniques and customizing CNN structures for specialized applications like remote sensing or personalized healthcare. Furthermore, it is necessary to make these models more interpretability in order to build confidence in important decision-making processes, especially in industries, such as finance or health care. We may close the app between state-of-the-art AI developments and their practical use in real-world situations by exploring these research pathways, which will ultimately benefit society as whole.

REFERENCES

[1] "Pneumonia -Symptoms | NHLBI, NIH". Nhlbi.nih.gov.

[2] Wikipedia contributors, "Pneumonia." Wikipedia, The Free Encyclopedia.

[3] "Chest X-ray." Mayo Clinic.

[4] Niedeman MS, Mandell LA, Anzueto A, et al. "Guidelines for the management of adults with community-acquired pneumonia." "American Journal of respiratory and critical Care Medicine 2001; 163 (7 Suppl): 1730.

[5] "chest radiography in pneumonia: the utility of rib views in diagnosing pleural effusion. "The American Journal of Emergency Medicine. 2019-37 (8): 1589-1590.

[6] Mandell LA, Wunderink RG, Anzueto A, et al. "Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. "Clinical Infectious Diseases. 2007; 44 Suppl 2:S27-S72.

[7] World Health Organisation. "Oxygen therapy for children. "WHO.

[8] World Health Organisation. "Pneumonia Key facts."

[9] Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography automated classification. Radiology, 284(2), 574-582.

[10] Huang, H., et al. (year). Lung nodule classification for diagnosing lung cancer using deep elarning techniques. [include full citation details]

[11] Rajpurkar, P., et al. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv: 1711.05225.

[12] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In proceeding of these IEE conference on computer vision and pattern recongition (pp. 770-778).

[13] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutioanl networks. In proceeding of the IEE conference on compuetr vision and pattern recongition (Vol. 1, No. 2).

[14] Szegedy, C., Vanhoucke, V., loffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Processings of the IEE conference on computer vision and Pattern recognition (pp. 2018-2826).

[15] Huang, G., & Liu, Z. (2017). Wide residual networks. arXiv preprint arXiv: 1605.07146

[16] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., … & Langlotz, C. (2018). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv: 1711-05225.

[17] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In advances in neural information processing systems (pp. 1097-1105).

[18] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional netwroks for large-scale iamge recogntiion. arXiv preprint arXiv: 1409.1556

[19] Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: an astounding baseline for recognition. In proceeding of the IEE conference on computer visison and patetrn recongition workshosp (pp. 806-813)

[20] Kumar Das, D., Ghosh, M., Pal, M., Maiti, A. K., & Chakraborty, C. (2013). Machine learning approach for automated screening of malaria parasite using microscope iamges. Micron 45, 97-106

[21] fig (4) By Biggerj1 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=143857678