Pneumonia Detection: An Efficient Approach Using Deep Learning

Ayush Pant
Final Year Student, Department of Computer Science
B.M.S. College of Engineering
Bengaluru, Karnataka
ayushpant1998@gmail.com

Akshat Jain
Final Year Student, Department of Computer Science
B.M.S. College of Engineering
Bengaluru, Karnataka
akshjain1512@gmail.com

Kiran C Nayak
Final Year Student, Department of Computer Science
B.M.S. College of Engineering
Bengaluru, Karnataka
kiran.nayak504@gmail.com

Daksh Gandhi
Final Year Student, Department of Computer Science
B.M.S. College of Engineering
Bengaluru, Karnataka
dakshgandhi1@gmail.com

Dr. B. G. Prasad
Professor, Department of Computer Science
B.M.S. College of Engineering
Bengaluru, Karnataka
bgprasad.cse@bmsce.ac.in

Abstract - Pneumonia is one of the largest infectious diseases that cause death in children and elderly people across the globe. Pneumonia impacts all the elderly and young people's families and children everywhere but is most prevalent in Sub-Saharan Africa and South Asia. In December 2019 Wuhan, a city of China was affected by deadly, gruesome Pneumonia which was declared a pandemic by World Health Organisation. But the reason for the outbreak was not clear to everyone. Later, the doctors identified the disease as a new species of coronavirus, also currently known as COVID-19. The main motivation behind this research was to identify Pneumonia just by using the X-Ray images of the patients. As doctors must do a lot of certain tests to identify if the patient has Pneumonia or not. To solve the cumbersome problem, an ensemble of two deep learning models is developed, to make the work of the doctors simpler. In this paper, a comparison between previously written relevant research papers is done and concluded with an ensembled deep learning model to achieve a remarkable test data accuracy or unseen data accuracy.

Keywords—Pneumonia, Deep Learning, Convolutional Neural Network, ResNet, U-Net, EfficientNet-B4.

I. Introduction

Pneumonia impacts all the elderly and young people everywhere but is most prevalent in Sub-Saharan Africa and South Asia. With the high growth in the popularity of neural networks, engineers and researchers have been able to find state-of-the-art products for computer vision. Artificial Intelligence helps us to automate analysis techniques, which is only possible now because of the technology of Deep Learning. The exposure to Pneumonia is quite high for many people, mainly in economically underdeveloped and developing countries where the majority are deprived of a nutritious diet. The World Health Organisation states that more than 4 million untimely deaths per year occur from diseases caused by air pollution.

The purpose of this project is to build an AI network, which takes the pixel values as input for a given X-Ray image and then proceeds to perform linear operations and activations on each of them. Then by taking all the above operations, and then multiplying them with each layer within the neural network, and the number of nodes. Suddenly you have millions of operations. By applying determination, there can be more efficiency in these tasks. The scope is to develop a model that will identify whether a patient is having Pneumonia or not by passing the chest X-Ray images through the Deep Learning model. The model should be highly precise as people's lives are at stake. In the past, it has been observed that the doctors undergo testing or get an X-Ray and give a false positive or a false negative result (Type 1 or Type 2 error), which resulted in bad medical conditions of the patients. Hence, such solutions can help in the medical fields to reduce these types of errors and save lives.

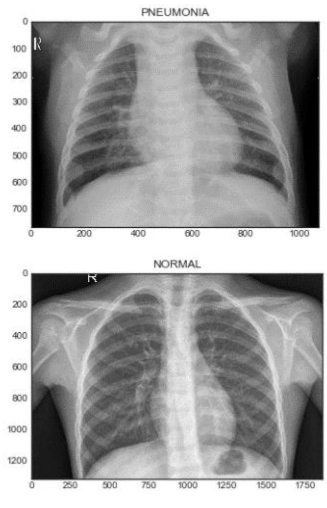


Figure 1: Pneumonia and Normal X-Ray Image

In Figure 1, it is difficult to determine which has Pneumonia and which is a normal X-Ray. The only difference which can be observed by anyone who is not a doctor is that the

X-Ray image with Pneumonia is a bit blurry as compared to the normal X-Ray image. Using just this knowledge target value cannot be determined. It is also difficult for a doctor to determine the outcome just by glancing at an X-Ray image.

II. Related Works

There are a lot of research papers trying to solve the problem involving the detection of Pneumonia and many of them have provided their improvement in the model. In this paper, an attempt is made to use these papers to analyse and get a subtle understanding of what solutions have been used and how a new model can be made.

A. Jointly Learning Convolutional Representations to Compress Radiological Images and Classify Thoracic Diseases in the Compressed Domain *[1]*.

This paper [1] has tackled the problem of information loss that occurs when the researcher downsamples the high-definition medical image data by using interpolation before feeding into the neural network. This paper [1] has presented a solution to the problem. They have used a Convolution Neural Network (CNN) based approach in which the resolution of the image using autoencoders and simultaneously classify it using a CNN. This method [1] helps in extracting essential features from the high dimension image. They have used high-definition images of a public dataset of a chest X-Ray and have outperformed state of the art solutions [1].

B. Pneumonia Detection with Supervised Learning *[2]*

To diagnose diseases in the proximity of chest area physicians use X-Rays. It is a cheaper alternative than a CT scan or MRI scan. It is difficult to diagnose with chest X-Rays compared to CT scan or MRI scan. Nowadays with the advent of technology physicians can diagnose diseases speedily and accurately by obtaining the chest X-Rays so the process of Pneumonia detection has become easy. Pneumonia, which is diagnosed only by chest X-Ray, takes around 50,000 people's lives each year. Input is again a front view Chest X-Ray image and the result is classified as Pneumonia free or not [2].

C. Deep Learning Approach to Pneumonia Classification *[3]*

In this paper [3], a Convolution Neural Network was designed from square one unlike in other papers where researches have used transfer learning to achieve a good score comparable to transfer learning techniques. The model proposed in the paper [3] is a CNN consisting of 2 primary components i.e. feature extractors and activation functions. The feature extractor comprises of multiple conv2D, conv3D layers with max-pooling layers stacked up with a RELU activator this output is flattened and then passed through a 2- layer ANN (Artificial Neural Network) which uses an activation function of RELU, a sigmoid function that performs classification tasks. After tuning the parameters and hyperparameters, the researchers were able to achieve a higher validation accuracy as compared to other approaches [3].

D. Dice Loss *[4]*

To understand dice loss let's take an example, let 90 percent of that data belong to class 1 and the remaining 10 percent of the data belonging to class 2. If a model is trained and it only predicts class 1 as an output of the model, then after predicting the accuracy, 90 percent accuracy is obtained, which is completely wrong. This is because of the class imbalancement. There is a possibility of getting high cross-entropy loss. Due to this, loss functions which tend to perform well even if there is imbalancement of data. One of such functions is Dice loss [4] (DL). It has been observed by many data scientists that using this loss can help the model towards imbalancement.

E. U-Net *[5]*

The U-Net [5] is a CNN (Convolution neural network) architecture that learns segmentation in an end to end-setting. This architecture is usually used for biomedical image segmentation. The reason why this architecture is called U-Net [5] is because of the shape of the network which looks like "U". The U-Net [5] consists of 2 parts which are, Encoder and Decoder. The encoder and decoder consist of convolution and max pool layers, but the size of the layers vary in both the Encoder and Decoder. The encoder and the decoder consist of the contraction path and the expansive path simultaneously. The U-Net [5] combines the location information from the encoder path with the contextual information in the decoder path to finally obtain a general information combining localization and context. This enables the architecture to achieve remarkable results for image segmentation.

F. Efficient Net *[6]*

This paper was written by two researchers at Google. This paper has compared other CNN models on ImageNet and it has shown how the efficient Net has outperformed other state of the art models such as Xception [7], AmeobaNet-C, ResNet and many other models. It has presented that the Efficient Net [6] can vary from EfficientNet-B0 to EfficientNet-B7 where, as the model number increases the more the accuracy of the model improves but the number of parameters increases simultaneously in millions, which will also increase computational power requirement. The reason why the Efficient Net can perform is because of the concept of Compound Model Scaling.

After evaluating the previous papers, it is evident that the outcome of some of the above papers has high precision but low recall & vice versa in case of other papers. To overcome the above problem, we came up with a solution of taking an ensemble of the two models in which one of the models should have high recall and the other with high

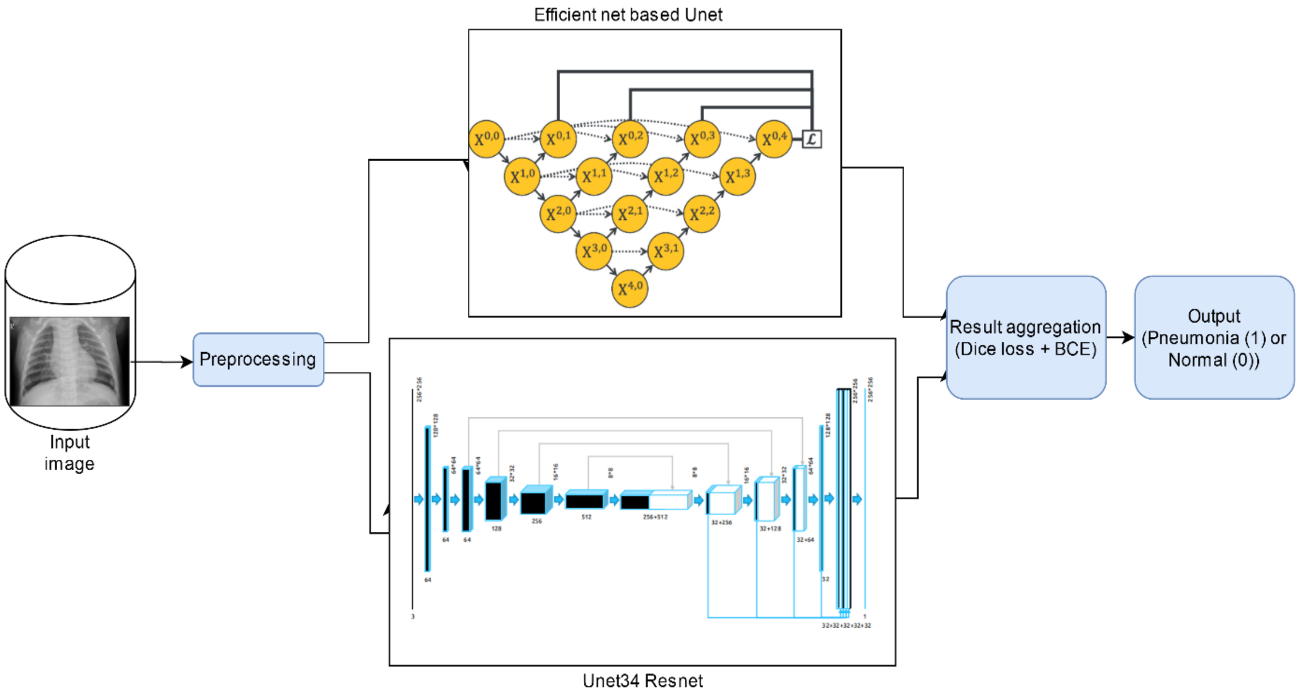


Figure 2: Model Architecture

precision. This might help in achieving a better result for our proposed overall model.

III. Proposed System

After a certain evaluation of the previous papers, a model is developed as seen in Figure 1. The problem that is supposed to be tackled while developing the model is class imbalancement i.e. the data is more biased towards the negative class (X-Ray images without Pneumonia) than the positive class (X-Ray images with Pneumonia). Initial planning was to solve the class imbalancement problem in two phases i.e. "the pre-processing" phase and in the "model architecture" as well. So at the pre-processing stage, the X-Ray image that needs to be classified as positive or negative Pneumonia must be converted to the required size/aspect ratio as the data consisted of different sizes of images. Since the data is imbalanced, data augmentation is needed on the X-Ray images i.e. artificially expand the size of the positive class (X-Ray images with Pneumonia). Many data augmentation techniques include such as Vertical and Horizontal Shift Augmentation, Vertical and Horizontal Flip Augmentation, Random Rotation Augmentation, Random Brightness Augmentation, and many more. In this project, such a model architecture will help in avoiding the demerits of class imbalancement. There is also a possibility of having noise in the image, to tackle this issue, Gaussian blur which is also known as Gaussian smoothing is used, which will help in reducing noise in the image. Once there are equal-sized classes of distribution and after pre-processing the data, images are passed through the model. The data must be simultaneously passed through both the models as shown in Figure 2. After having a look at the previous papers, it was concluded that an ensemble of ResNet-34 based U-Net and EfficientNet-B4 based U-Net would be a good way to tackle all the problems that were faced in the previously written papers.

IV. Algorithm

As discussed earlier the problem that was faced was due to Class imbalancement i.e. the X-Ray images of the non-Pneumonia class are more as compared to the Pneumonia X-Ray images. Therefore the architecture that is proposed is an ensemble of two deep learning models. The first model is a "ResNet-34 based U-Net" which is ensembled with the second model which is an "EfficientNet-B4 based U-Net". On this architecture, a range optimizer, BCE (Binary cross-entropy), and Dice Loss with progressive scaling are used.

The ResNet [8] is a serial collection of residual blocks (Res-Blocks) with skip connections, unlike CNN's. Resnet was released in 2015 which was a remarkable research, the advantage of using such a powerful model is the performance of the model applications. As seen in [8] each block is having a skip connection that acts like a memory with historical data. There is stability when ResNets are used instead of generic CNNs. This network is in parallel with EfficientNet-B4 based U-Net where each block in the U-Net is an EfficientNet. Here a new approach is proposed, that uses the compound model scaling feature of EfficientNets which supports depth, resolution, and width scaling, encoding and decoding feature from U-nets, historical memory capability from ResNets, which are ensembled together, to increase the efficiency of the model. A U-Net model is mainly used in biomedical Image classification, because of its ability to be used as Encoder-Decoder. U-Nets are most effective when used in applications that have the same input and output

dimensions. Therefore, U-nets are helpful in super-resolution or colorization. At the end of the U-Nets of both the models a dense layer followed by an output layer is added, to determine the output i.e. Pneumonia or Normal.
While training dense detectors, which have large class imbalancement, there is a possibility of getting high cross-entropy loss. Due to this, loss functions are used, which tend to perform well even if there is imbalancement of data, one such function is Dice loss (DL). It has been observed by many data scientists that using this loss can help the model towards imbalancement. The Dice loss as given in [4] is defined as (1).

$$\mathrm{DL}(p_t) = 1 - \frac{2\sum_{\text{pixeles}} Y_{\text{true}} Y_{\text{pred}}}{\sum_{\text{pixeles}} y^2{}_{\text{true}} + \sum_{\text{pixeles}} y^2{}_{\text{pred}}} \qquad (1)$$

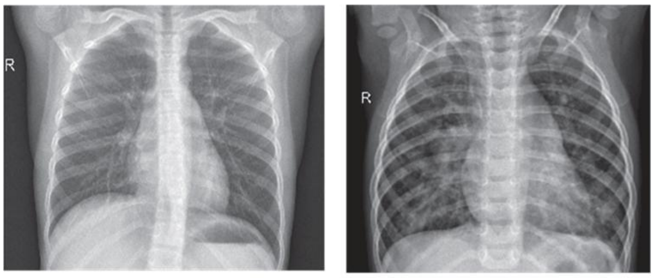
The dice coefficient is calculated for each class (positive or negative) mask. The above formula scoring is repeated all the classes, and, in the end, an average is considered.

$$\mathrm{CE}(p, y) = \begin{cases} -\log(p), & \text{if } y = 0 \\ -\log(1-p), & \text{otherwise} \end{cases} \qquad (2)$$

Equation (2) is defined as the Cross-Entropy Loss. Here p ∈ [0, 1] which is the probability of class with value 1 as its label. ResNet-34 based U-Net ensemble with EfficientNet-b4 based U-Net with ranger optimizer and uses BCE (Binary cross-entropy) using Dice loss with progressive scaling. An important point to note here is that the number of connections that exist between the input layer and first hidden layer should be less and similarly, the number of connections that exist between the last hidden layer and output layer should be less, this is done to maximize the efficiency.

V. DATA SET

A dataset from Kaggle was used. It contained 5863 Chest X-Ray Images in jpeg format and the images were categorized into a) Train b) Test and c) Validate categories. These were further broken into 2 categories, Pneumonia and Normal.
The dataset covers both Normal X-Ray images and X-Ray images with Pneumonia. As one can notice, X-Ray of a healthy human (left) would show a clear, translucent image of the lungs. Meanwhile, X-Ray of the lungs of a person suffering from Pneumonia (right) would be opacified in a focal manner as shown in Figure 3.



(a) Normal (b) Pneumonia

Figure 3: X-Ray Images

These images were collected from the Guangzhou Women and Children's Medical Centre in China and were a part of the normal routine check-ups of the patient. The quality of the dataset was verified by 2 expert physicians and all low-quality images were removed from the dataset. For any errors, the dataset was again verified by a third expert.

VI. RESULT

The first model that was trained, was ResNet based U-Net. In Figure 4 the result is aggregated and gives an idea about how the model is performing.

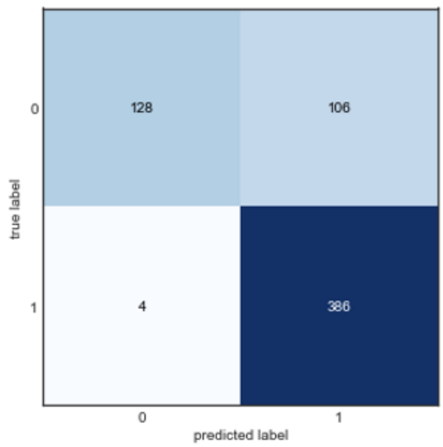


Figure 4: Test Data: Model 1 (ResNet based U-Net)

According to the results from Figure 4, an Accuracy of 0.82 is obtained, with a high Recall value of 0.99 which signifies that there are low false negative values i.e. the model had very low incorrect predictions for the negative class values (negative class here is Pneumonia X-Ray images). So, a clear conclusion can be made that this model is more biased towards the Pneumonia X-Ray images.
For the Model 2 (EfficientNet-B4 based U-Net) the accuracy and Precision are high values indicating that the False Positives and False Negatives are less after prediction. Hence, this model taken alone is reliable. The results for Model 2 can be seen in Figure 5.

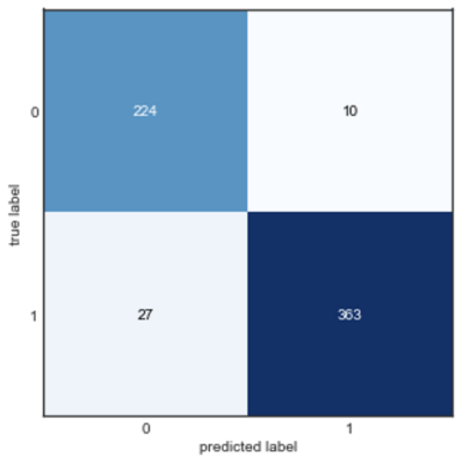


Figure 5: Test data: Model 2 (EfficientNet-B4 based U-Net)

Comparing this model with the previous model i.e. ResNet based U-Net, it can be seen that the recall of the current model is less as compared to the previous model. But simultaneously has high precision value.
As seen before, the recall of the ResNet based U-Net was high (0.99) and the precision of the EfficientNet-B4 based U-Net is also high (0.97). Thus it is decided to combine the benefits of both models i.e. take the ensemble of both models such that a better model can be obtained with results that are best of both models. The results for the Ensembled Model can be seen in Figure 6.

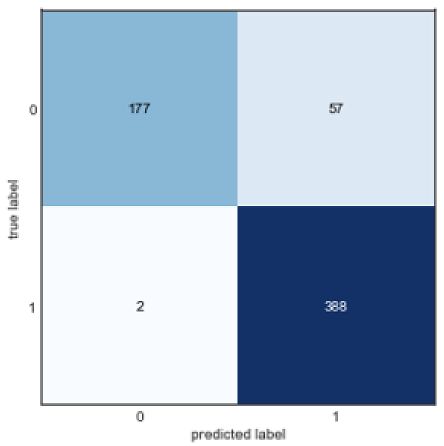


Figure 6 : Ensemble of EfficientNet-B4 based U-Net [6] and ResNet based U-Net [8]

Results for test data:

	Accuracy	**Precision**	**Recall**	**F1-score**
EfficientNet [6]**-B4 based U-Net** [8]	0.94	0.97	0.93	0.95
ResNet based U-Net [8]	0.82	0.78	0.99	0.87
Ensemble of EfficientNet-B4 based U-Net and ResNet based U-Net	0.90	0.87	0.99	0.92

VII. Conclusion

While this Pneumonia detector is not a fully deployable product that can change the world, it is clear that it is easy to get started with it. It is great to witness the growth and accuracy of deep learning in such real-world scenarios. This model is robust as it can work on any of the datasets that conform to the size of the image that is required for this model. We can observe that our model has given astounding results for the "Efficientnet-B4 based U-Net" model that has high precision and decent recall, but the other model, "ResNet based U-Net" had given high recall but low precision. The ensembled model uses the best of both worlds, in that the high Precision quality is drawn from EfficientNet-B4 based U-Net, and the high Recall quality is taken from the ResNet-34 based U-Net. The ensemble of both these models provides great results implicitly. But we can also see that our first model "Efficientnet-B4 based U-Net" individually performed better than our ensembled model in terms of the accuracy metric but the ensembled model has given a decently accurate result in the real case scenario.

VIII. References

[1] Ranjan, E. Paul, S. Kapoor, S. Kar, Aupendu, Sethuraman, R. Sheet and Debdoot, "Jointly Learning Convolutional Representations to Compress Radiological Images and Classify Thoracic Diseases in the Compressed Domain," in *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), ACM*, Hyderabad, 2018.

[2] B. Antin, J. Kravitz and E. Martayan, "Detecting Pneumonia in Chest X-Rays with Supervised Learning," 2017.

[3] O. Stephen, M. Sain, U. J. Maduh and D.-U. Jeong, "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare," *Journal of HealthCare Engineering,* vol. 2019, 2019.

[4] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin and M. J. Cardoso, "Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep*

Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2017.

[5] Brox, O. Ronneberger, P. Fischer and Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015.

[6] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *ICML,* 2019.

[7] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *IEEE conference on computer vision and pattern recognition*, 2017.

[8] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2015.

[9] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren and A. Y. Ng, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," 2017.

[10] T. Vijayakumar, "NEURAL NETWORK ANALYSIS FOR TUMOR INVESTIGATION AND CANCER PREDICTION," *Journal of Electronics,* pp. 89-98, 2019.

[11] D. T. Vijayakumar and M. R. Vinothkanna, "Mellowness detection in dragon fruit using deep learning strategy," *Journal of Innovative Image Processing (JIIP) 2,* pp. 35-43, 2020.

[12] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri and R. M. Summers, "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[13] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal loss for Dense Object Detection.," in *IEEE International Conference on Computer Vision (ICCV)*, 2017.