Multiclass Disease Classification using Modified CNN and Segmented Chest X-ray

Neelkant Newra, Sai Kumar, Saurab Gupta, Sumit Kumar Banchhor

Abstract

Multiclass Disease classification is always a challenge for any researcher. In the cases of pneumonia detection, which is a binary image classification task, many researchers have already achieved accuracy close to 99%. In multiclass classification, especially with 4 classes, the highest accuracy achieved was 93.42% with the AlexNet model. Although it was not possible for us to achieve the same accuracy with similar architecture, It can be due to the dataset distribution. In this study, we have provided an approach to using segmented CXR images, which are extracted through the U-Net model and further fed to the various CNN(Convolution Neural Network) Architecture. We have also kept the same data distribution to check different models' performance. Although our approach of using segmented images was not able to perform well compared to normal images, It can be due to not properly segmented data, since the IoU obtained for U-Net was 90%.

1. Introduction

Covid-19, Tuberculosis, Pneumonia, and other lung diseases share some very common image textures in CXR images, It is very hard for many physicians to determine and distinguish between multiple diseases through such images (Perumal, 2021). Many researchers have already achieved an accuracy of more than 99% in the case of Binary Classification (Pneumonia and Normal). And 93.42% accuracy is achieved in multiclass disease detection using CXR.

Covid-19 brought new challenges for many researchers as the texture of Covid-19 was overlapping with the images of Pneumonia and sometimes misclassify Covid-19 as Pneumonia (Pereira, 2020). In the pandemic, it is very important to have a proper detection of Covid. Classifying images for multiclass give the model more generalization, as it is now not only fixed to classify between two but is getting multiple probabilities for more than two class (Shotton, 2009). Based on that, whichever has the maximum probability is assigned that class. It can differentiate between the texture of Covid-19 and Pneumonia through the small margin of probability.

In this paper, we will be going to have a brief overview of the previous attempted works on multiclass disease classification. We will have a discussion on the Methodology we adopted, and the various layers of CNN. Furter in section 4, a detailed analysis is provided of the metrics achieved using segmented images and normal non-segmented images. Along with the description of data distribution, we kept for training, testing, and validating the model.

2. Literature review

Recently there are many studies conducted on multiclass image classification using the CNN model with CXR images data. Apostolopoulos et al. used MobileNet to classify Pneumonia, covid and normal and achieved an accuracy of 96.78%, and 94.72% for 2 class and 3 classes respectively. Basu et al. proposed a domain extension transfer learning approach and achieved an accuracy of 90%. Hasan et al. used the VGG-16 model to improve the performance of 3-class classification and have achieved an accuracy of 91.69%. Ibrahim et al. worked on the AlexNet model which

has already performed goof for binary classification problems. They tried for 4 classes, 3 classes, and 2 classes and achieved an accuracy of 93.42%,94%, and 99.62% respectively. Redie et al. improved the DarkCovidNet model and achieved an accuracy of 94.13% for 3-class classification.

3. Methodology

CNN mainly consists of three-layer, i.e. Convolution layer(Conv), Pooling layer(Pool), and Fully connected layer(FC). Additionally, if we are talking about U-Net Architecture, we see a Transposed Convolution layer. In this study, we have used the U-Net model to obtain the mask of CXR, further, we have performed bit-wise AND operation to get the segmented portion of the lung.

3.1. Convolution layer

Mathematically, Convolution is a method through which we combine two signals to form the third signal(Krizhevsky, 2012). CXR Images are also a form of signals, from which we have to obtain features to be fed into the fully connected layer as shown in figure 1. In the Convolution layer, the filter of (m,n,c) is moved around the images of (M, N, C) shape. Where m, n, f are the length of the filter, the width of the filter, and the number of filters respectively, and M, N, C are the length, width, and the number of channels in the images. It is often advised to take the filter width m=n as an odd number, to get the centroid, during the convolution. It helps to capture values from all directions which enhances features.

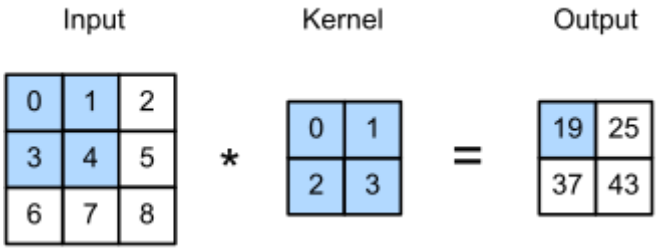

Figure 1: Convolution on 3×3 matrix, using 2×2 filter

3.2. Transposed Convolution layer

Transposed convolution layer is often confused with the de-convolution layer, basically, in Convolution, we do downsampling. It means the shape gets reduced. While in the case of Transposed Convolution, it is upsampling. The procedure of Transpose Convolution is very similar to that of convolution (Lai, 2017). Here we do some extra padding and carry out the normal convolution, which increases the shape of the next layer as shown in figure 2.

For a given size of the input (i), kernel (k), padding (p), and stride (s), the size of the output feature map (out) generated is given by:

$$out = (i - 1) \times s + k - 2p$$

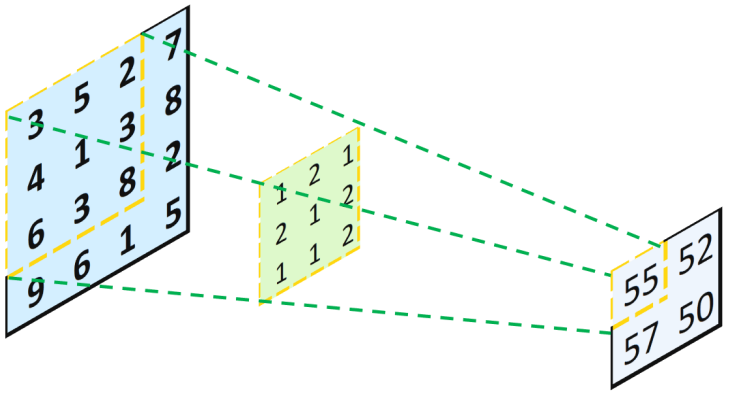

Figure 2: Transposed Convolution on 2×2 matrix, using 3×3 filter

3.3. Pooling layer

Pooling layer are used to reduce the dimension feature. In the pooling layer, there are generally no features to be trained. It helps in the reduction of the already available parameters to make the model more simplistic and computation feasible(Yu, 2014). It summarises the feature of the convolution layer. There are basically 3 types of pooling that are most used - Max Pooling, Average Pooling, and Global pooling.

3.3.1. Max-Pooling: In Maxpooling a kernel is moved in the layer, which pools out the max element in the covering layer as shown in Figure 3a.Max pooling is the most preferred pooling method used this day.

3.3.2. Average-Pooling: Initially, the Average pooling was highly used, as the name suggests, it takes an average of all elements covered in the kernel as shown in figure 3b.

3.3.3. Global-Pooling: It is itself classified into two categories- Global Average Pool and Global Max-Pool. In Global Average Pooling the average of all the cells is taken and given 1×1 output. Same with Global Max Pooling, it gives the max value as 1×1 output. As shown in Figure 3c.

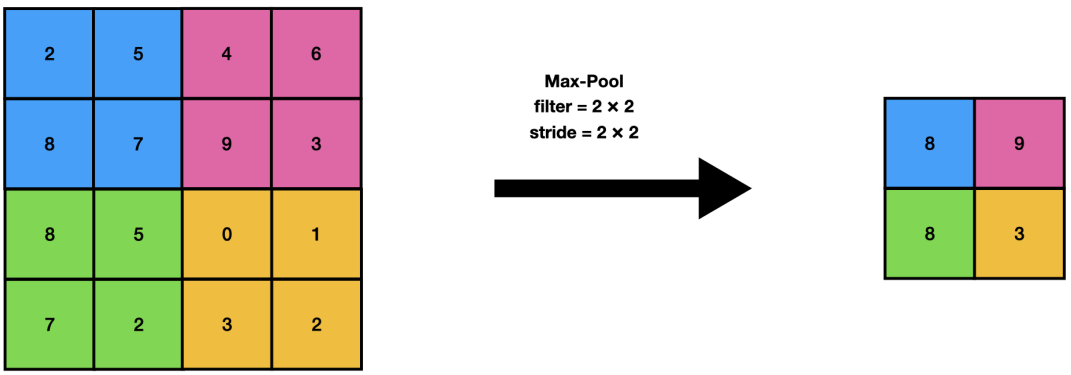

Figure 3a: Max-Pooling on 4×4 matrix using 2×2 filter and stride of 2×2

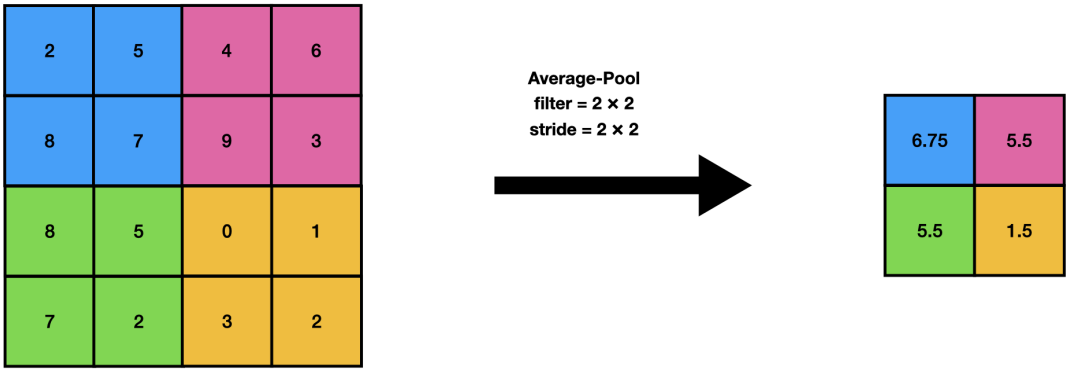

Figure 3b: Average-Pooling on 4×4 matrix using 2×2 filter and stride of 2×2

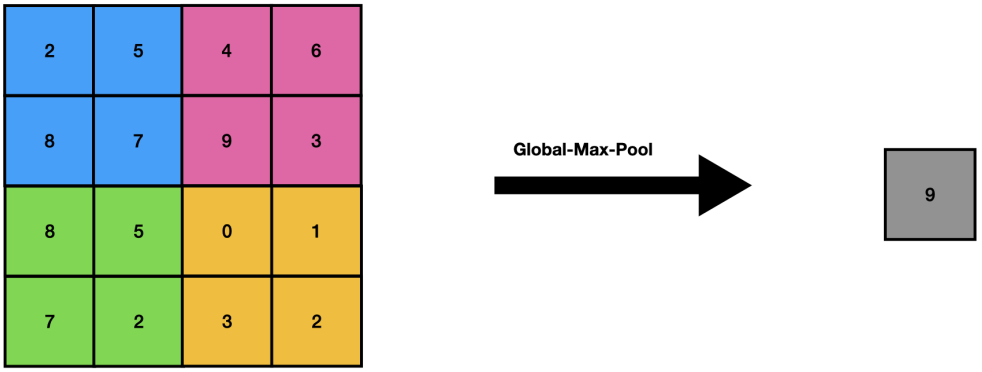

Figure 3c: Global-Max-Pooling on 4×4 matrix

3.4. Fully connected layer

As the name suggests every layer is fully connected to the previous layer as shown in Figure 2. The features obtained from the previous convolution or pooling layer are flattened and multiplied by a weight, a biased is also added. This layer can be assumed as the input of a simple neural network.

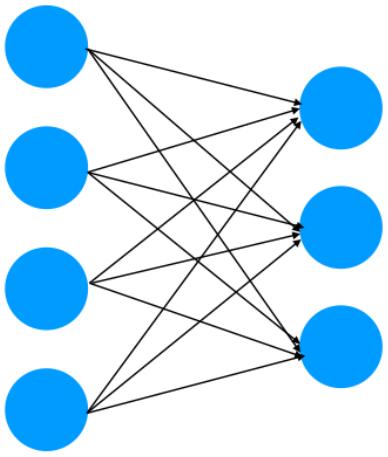

Figure 4: Fully connected layer

3.5. Segmented CXR images using U-Net

In this study we have used 512px × 512px, 650 CXR images, and a mask to develop the U-Net model (Ronneberger,2015) which was having IoU of 90%. Through this trained model the mask for our input sample is developed. We have used Bit-wise

AND between the original image and the mask to obtain the segmented region. As shown in Figure 3.

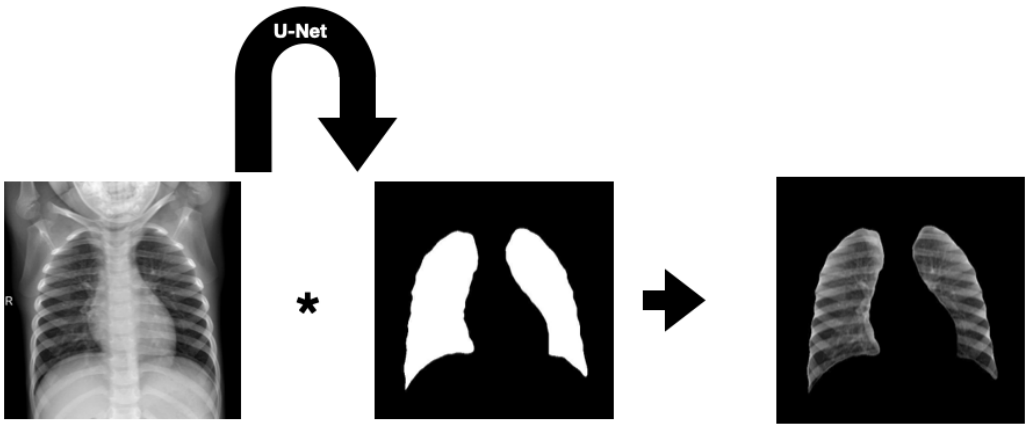

Figure 5: U-Net and Bitwise AND operation to obtain segmented CXR image

4. Result and discussion

4.1. Dataset description

In this study we have made two separate datasets for training. The first dataset consists of segmented images of 128px × 128px, obtained using the U-Net model. The second dataset consist of 128px × 128px original images. We have kept the same label for both datasets. We have used 7135 CXR images, having train, test, and validation distribution as 6326,771, and 38 respectively. All the datasets are randomized to avoid sequence pattern learning and reduce overfitting. The dataset used in this study is publicly available in kaggle(Newra , 2022).

4.2. Experimental Analysis

We have used the same dataset and evaluated accuracy, precision, recall, and F1 score for both segments as well as Original images on four different models. The achieved result for the Original and segmented CXR images is shown in Tables 1 and 2 respectively.

Model	Training				Testing			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
VGG-13	0.9512	0.9531	0.9485	0.9508	0.8560	0.8620	0.8508	0.8564
AlexNet	0.9683	0.9709	0.9658	0.9683	0.8197	0.8337	0.8132	0.8233
Mobile Net	0.9918	0.9924	0.9915	0.9919	0.8457	0.8449	0.8405	0.8427
Modified-DarkCovidNet	0.9813	0.9814	0.9810	0.9812	0.8313	0.8344	0.8300	0.8322

Table 1: Evaluation data with original CXR images for various model

Model	Training				Testing			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
VGG-13	0.9317	0.9367	0.9287	0.9327	0.7613	0.7668	0.7549	0.7608
AlexNet	0.9625	0.9652	0.9594	0.9623	0.7691	0.7781	0.7639	0.7709
Mobile Net	0.9344	0.9389	0.9301	0.9345	0.7756	0.7795	0.7704	0.7749
Modified-DarkCovidNet	0.9483	0.9497	0.9473	0.9485	0.7924	0.7935	0.7924	0.7929

Table 2: Evaluation data with segmented CXR images for various model

4.3. Confusion Matrix

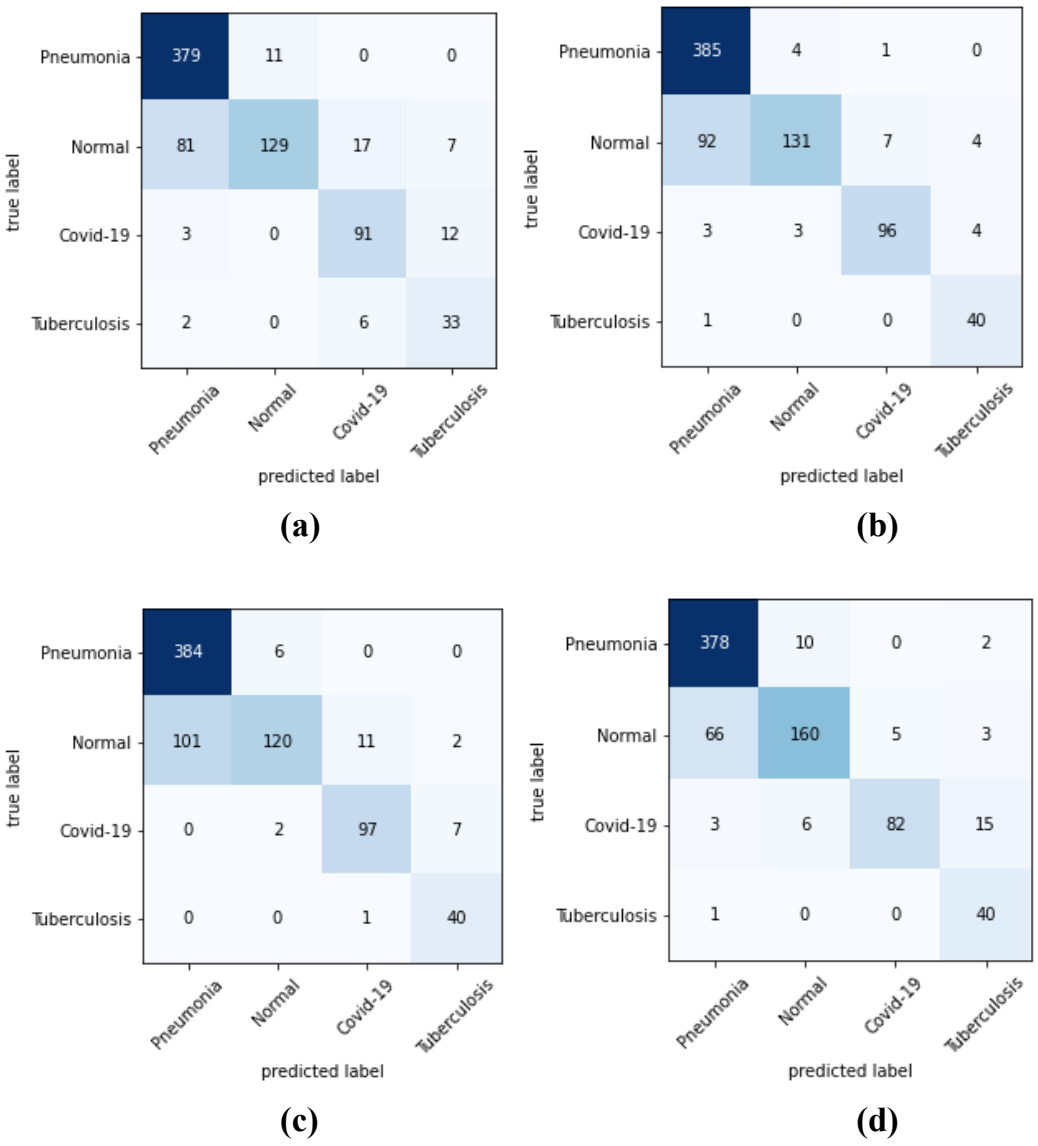

Figure 6: Confusion matrix of model evaluation on Original CXR images a)AlexNet b)MobileNet c)Modified-DarkCovidNet d)VGG13

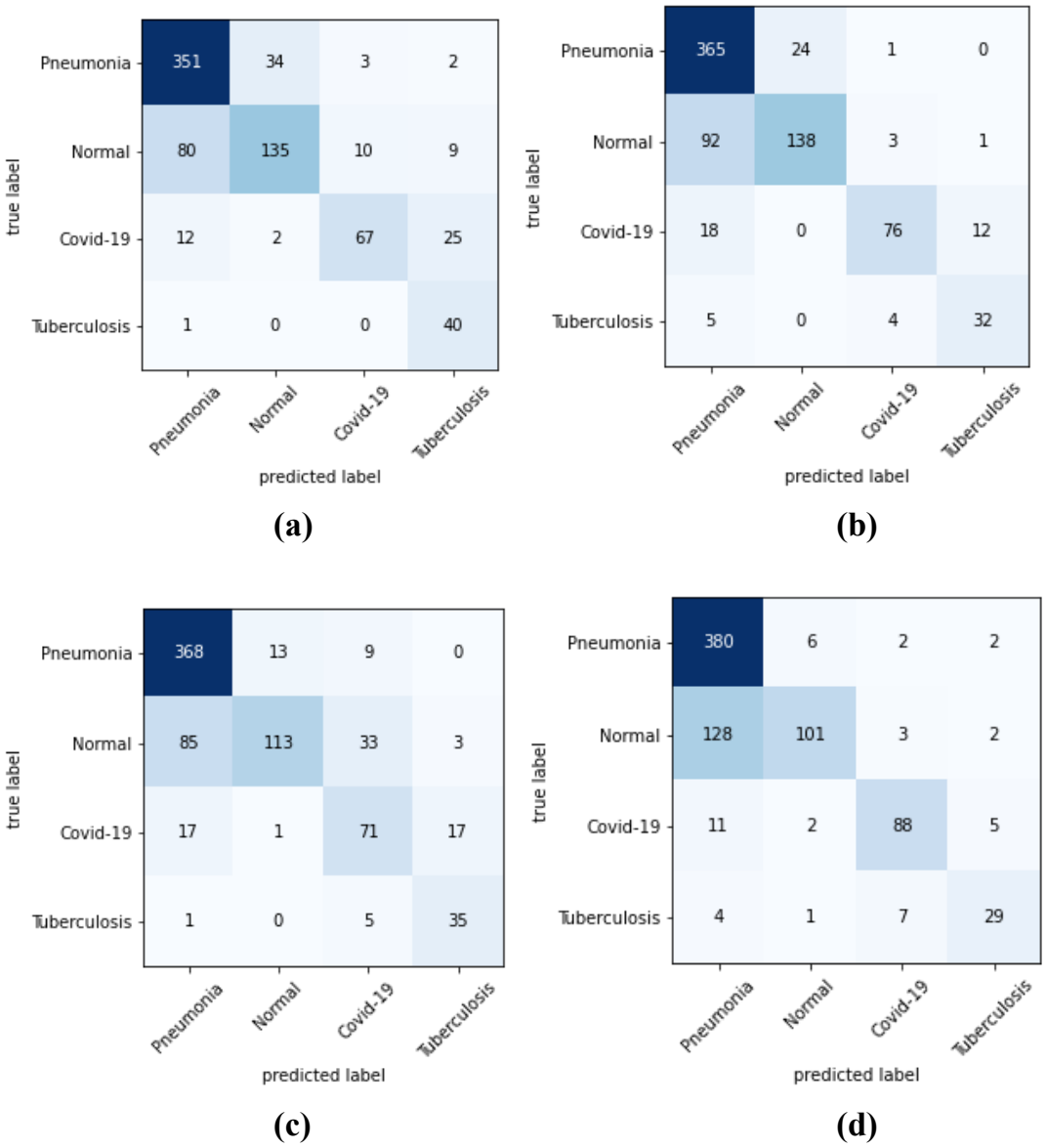

Figure 7: Confusion matrix of Model evaluation on Segmented CXR images a)AlexNet b)MobileNet c)Modified-DarkCovidNet d)VGG13

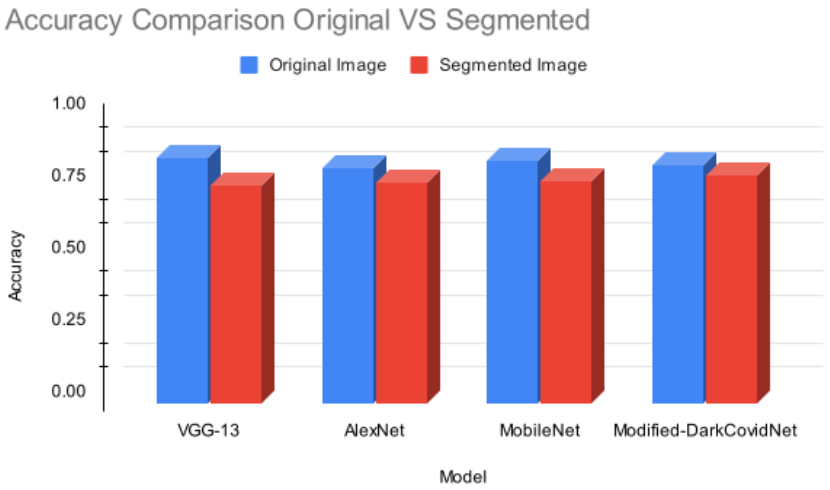

Figure 8: Accuracy Comparision between the original image and segmented image feeding

5. Conclusion

Multiple Disease detection through CXR will help the patient to evaluate their health and take precautions. While detection of Covid-19 will help to prevent its widespread. In this study, we discussed the segmented CXR data feeding method. Although it has not achieved a comparable accuracy compared to using the original CXR image. It can be due to not proper mask generation of U-Net. We can improve the IoU of U-Net and improve the segmented CXR data. This method ensures that our model is looking at the right resource in the CXR images to detect the disease.

Reference

1. Perumal V, Narayanan V, Rajasekar SJ. Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. Applied Intelligence. 2021 Jan;51(1):341-58.
2. Pereira RM, Bertolini D, Teixeira LO, Silla Jr CN, Costa YM. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. Computer methods and programs in biomedicine. 2020 Oct 1;194:105532.
3. Shotton J, Winn J, Rother C, Criminisi A. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. International journal of computer vision. 2009 Jan;81(1):2-3.
4. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and engineering sciences in medicine. 2020 Jun;43(2):635-40.
5. Basu S, Mitra S, Saha N. Deep learning for screening covid-19 using chest x-ray images. In2020 IEEE Symposium Series on Computational Intelligence (SSCI) 2020 Dec 1 (pp. 2521-2527). IEEE.
6. Hasan MD, Ahmed S, Abdullah ZM, Monirujjaman Khan M, Anand D, Singh A, AlZain M, Masud M. Deep learning approaches for detecting pneumonia in COVID-19 patients by analyzing chest X-ray images. Mathematical Problems in Engineering. 2021 May 19;2021.
7. Ibrahim AU, Ozsoz M, Serte S, Al-Turjman F, Yakoi PS. Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive Computation. 2021 Jan 4:1-3.
8. Redie DK, Sirko AE, Demissie TM, Teferi SS, Shrivastava VK, Verma OP, Sharma TK. Diagnosis of COVID-19 using chest X-ray images based on modified DarkCovidNet model. Evolutionary Intelligence. 2022 Mar 9:1-0.
9. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012;25.
10. Lai WS, Huang JB, Ahuja N, Yang MH. Deep laplacian pyramid networks for fast and accurate super-resolution. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 624-632).
11. Yu D, Wang H, Chen P, Wei Z. Mixed pooling for convolutional neural networks. InInternational conference on rough sets and knowledge technology 2014 Oct 24 (pp. 364-375). Springer, Cham.
12. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InInternational Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234-241). Springer, Cham.

13. Newra N, CXR DATA for Multiclass classification. 2022. doi: 10.34740/KAGGLE/DSV/3458562. url: https://www.kaggle.com/dsv/3458562.