



# Deep Learning models for Covid 19 Lesion Identification using CT Lung Images

Sai Krishna K<sup>1</sup>, Sasikala R<sup>2</sup>

2<sup>nd</sup> Year MTech CSE with specialization in AI & ML, Associate Professor

School of Computer Science and Engineering

Vellore Institute of Technology, Vellore, India

saikrishna.k2020@vitstudent.ac.in, sasikala.ra@vit.ac.in

**Abstract:** Corona Virus is an unexpected catastrophe that led to a devastating effect on human health. This paper aims to find the exact location of the lesion for estimating the disease progression using the Deep Learning technique like CNN as it is proven to be the best for the medical images. In this paper to enhance the quality and performance of the dataset in the pre-processing certain techniques are used namely CLAHE, Data Augmentation, cropping the images using the contour technique, and removal of images with no corresponding mask image. Later in this paper, a comparative analysis is done using K fold Cross Validation between different models like UNet segmentation, UNet segmentation with pre-trained InceptionResNetV2 as its backbone, UNet segmentation with pre-trained DenseNet201 as its backbone, and UNet ++ segmentation to find out the best model in identifying the exact location of the region of infection and for the best utilization of resources in this paper, Adam's Optimization is used with Binary Cross Entropy Dice loss function. After the experimental analysis, UNet segmentation with pre-trained DenseNet201 as its backbone got the best Dice Similarity Coefficient of 95.883.

**Index Terms -** Lesion segmentation, Transfer Learning, UNet, UNet++, DenseNet201, and InceptionResNetV2.

## I. INTRODUCTION

Corona Virus has raised several questions in terms of medical infrastructure. RT PCR is a standard diagnostic tool to diagnose Covid 19 by molecular testing of the respiratory tract by Nasopharyngeal swab Tracheal aspirant or Oropharyngeal swab. Though this technique is beneficial in diagnosis there are a few problems that are concerned with it i.e in different patients this disease has very high variability ultimately sensitivity is below 70%, and it is a time-consuming process i.e it takes nearly 6 – 8 hours to get the test results. Considering these facts into consideration in situations it is not correct to rely on RT PCR. Pathological imaging techniques like CT scan helps to overcome the problems that arise due to RT PCR. CT scan provides a 3-Dimensional view of the lung image helping the doctors in accessing the lung morphology which helps in diagonalizing the disease. But when the number of patients increases it is very difficult for the radiologists to manually annotate which ultimately slows down the CT analysis time and there are situations where due to lack of expertise covid 19 and pneumonia are often misinterpreted.

To reduce the burden of the radiologists this paper aims to find the exact location of the lesion for estimating the disease progression using the Deep Learning technique like CNN is used as it is proven to be the best for the medical images. [1][2][3]

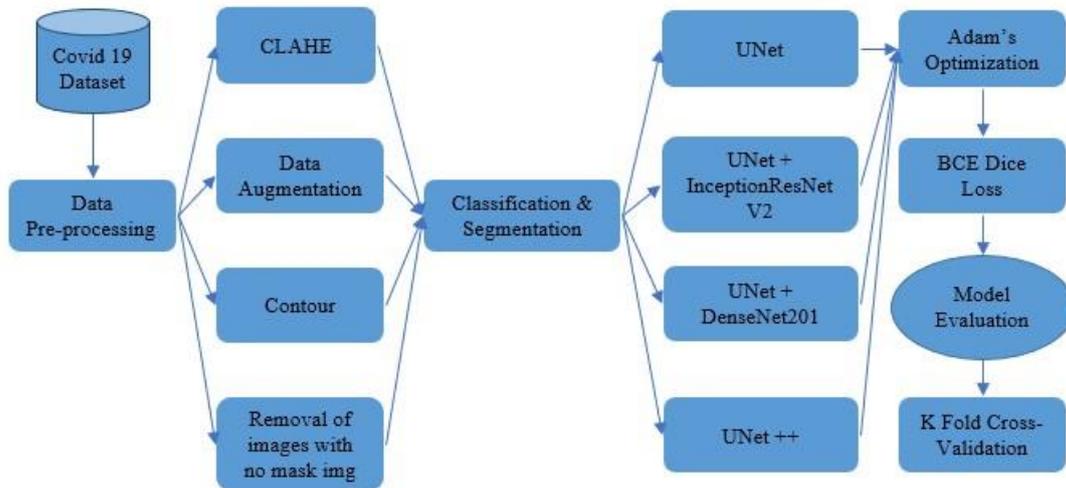
In this paper to enhance the quality and performance of the dataset in the pre-processing certain techniques are used namely CLAHE, Data Augmentation, cropping the images using the contour technique, and removal of images with no corresponding mask images. Medical images have Contrast Limited problems, CLAHE technique is used to enhance the quality of the images such that certain features are enhanced for the feature extraction. Deep learning models like CNN perform well only if there is a huge amount of training data to overcome overfitting. If a new disease emerges optimal training is not at all possible as initially there will be very few radiological images. Data augmentation creates "dummy data" which is unique in comparison to the actual data by performing operations like various rotations, width shifts, height shifts, and zooming. Segmentation requires normal lung images and corresponding mask images if the dataset contains lung images without mask images will affect the quality of training so the images that do not have corresponding mask images are removed. [4][5][6] The latest architectures are very huge and it consumes more RAM and takes more computing power as the number of images for training cannot be reduced as it leads to overfitting the size of the images can be reduced. Lung CT images possess a lot of black space containing nothing and the parts like the diaphragm i.e below the lungs are not necessary for the segmentation for this contour is drawn over the image and cropping the bigger area. Cropping the images will reduce the usage of RAM as well as computing power. [7][8][9]

Locating the exact pathological tissues that is been affected is very crucial for identifying the severity. So, the exact volumetry estimation of the lesion is very crucial so in this paper, a comparative analysis is done between UNet segmentation, UNet segmentation with pre-trained InceptionResNetV2 as its backbone, UNet segmentation with pre-trained DenseNet201 as its backbone, and UNet ++ segmentation to find out which technique is effective to identify the exact location of the Region of Infection, for the best utilization of resources in this paper, Adam's Optimization is used with Binary Cross Entropy Dice loss function. [10][11][12]

After the experimental analysis, UNet segmentation got a Dice Similarity Coefficient of 93.523, UNet segmentation with InceptionResNetV2 as its backbone got a Dice Similarity Coefficient of 95.575, UNet segmentation with DenseNet 201 as its backbone got a Dice Similarity Coefficient of 95.883 and UNet ++ got a Dice Similarity Coefficient of 92.385. To conclude UNet segmentation with pre-trained DenseNet201 as its backbone got the best Dice Similarity Coefficient of 95.883.

## II. PROPOSED METHODOLOGY

### 2.1 Proposed Architecture



### 2.2 Proposed Steps

#### Step 1: Dataset

In this step, a dataset is taken from Kaggle.

#### Step 2: Pre-Processing

In this step, 4 different operations are performed

##### i. CLAHE - Contrast Limited Adaptive Histogram Equalization

Medical images have contrast limited problems. To overcome this Histogram Equalization is done, it spreads the places that have more intensity, and thereby it creates images with increased contrast and this helps to show the existence of a few features & ultimately enhance the performance during the process of feature extraction. This HE can lead to too much lighter or darker regions as the contrast is not limited in order to overcome these kinds of problems in this paper CLAHE technique is used. This CLAHE transforms every pixel by making use of a transformation function that is obtained from the nearby region where the transformation of every pixel is done based on the histogram of a square, the subregion of the image surrounding the pixel.

##### ii. Data Augmentation

Deep learning models like CNN perform well in computer vision operations however it requires a huge amount of training data to avoid overfitting. If a new disease emerges optimal training is not at all possible as initially there will be very few radiological images. Since the dataset that is collected has less number of images in this paper data augmentation is done, it enables different techniques that enhance the size, as well as the quality of the training datasets by creating “dummy data” which is unique in comparison to the actual data by performing operations like various rotations, width shifts, height shifts, and zooming.

##### iii. Removing Images with No Corresponding Mask Images

Segmentation requires normal images and corresponding mask images, if there is no corresponding mask image for the lung images then it will affect the quality of training so in this paper, the images that do not have corresponding mask images are removed.

##### iv. Cropping the Images using the Contour technique

The latest architectures are very huge and it consumes more RAM and takes more computing power as the number of images for training cannot be reduced as it leads to overfitting the size of the images can be reduced. Lung CT images possess a lot of black space containing nothing and the parts like the diaphragm i.e below the lungs are not required for the segmentation. Cropping the images will reduce the usage of RAM as well as computing power. So, the images are cropped and only the Region of Infection is taken. For cropping, slicing the rows and columns using a trial-and-error process can be done but this process will be limited to a

specific dataset only. So, in this paper, a contour is drawn over the image and then the rectangle with the biggest contour i.e largest area is cropped. While cropping the CT images the respective mask images should also be cropped with the same limits otherwise pixel-level labeling will go wrong and the model might map a good area with an infected area or vice versa. Cropping the images will reduce the usage of RAM as well as computing power.

#### Step 4: Convolution Neural Network & Segmentation

Locating the exact pathological tissues that is been affected is very crucial for identifying the severity. So, the exact volumetry estimation of the lesion is very crucial so in this paper, a comparative analysis is done between UNet segmentation, UNet segmentation with pre-trained InceptionResNetV2 as its backbone, UNet segmentation with pre-trained DenseNet201 as its backbone, and UNet ++ segmentation to find out which technique is effective to identify the exact location of the Region of Infection.

##### A. UNet Segmentation

UNet segmentation was invented to work on biomedical images where the main moto is not to only classify the presence of infection but it aims to identify the region of infection.

It consists of two main parts

###### i. Encoder

In the image, the encoder is used to capture the context, it reduces the spatial dimension and later increases the number of channels i.e  $128*128*3$  to  $8*8*256$ . This encoder consists of convolutional and max-pooling layers with ReLU activation function and Batch Normalization. This max-pooling also called downsampling tells “WHAT” is present in the image, it is done by increasing the receptive field but there are chances of losing the information of “WHERE” that is present in the organ. After every down-sampling feature channel is doubled and the spatial dimensions are reduced to half. Finally, an image is converted from high resolution to low resolution.

###### ii. Decoder

The decoder is used to identify the exact location of the region of infection. It increases the spatial dimension and reduces the number of channels i.e  $8*8*256$  to  $128*128*3$ . The decoder consists of convolutional and upsampling layers with ReLU activation function and Batch Normalization. For Upsampling transposed convolution is preferred, normal convolution can be expressed as a matrix multiplication of the input image and filter to produce the output image whereas in the case of transposed convolution transpose of the filter matrix is taken. Finally, an image is converted from low resolution to high resolution.

##### B. UNet Segmentation with Pre-Trained Models

In this project transfer learning concept, i.e pre-trained model is used. Pre-trained model is a previously trained model on a different dataset that contains weights as well as a bias which ultimately represents the features of the already trained dataset. These learned features are frequently transferable to another dataset.

###### i. UNet Segmentation with pre-trained InceptionResNetV2 as its backbone

This model uses the concept of UNet segmentation where the encoder is changed with InceptionResNetV2 ultimately this model contains benefits of both InceptionResNetV2 as well as U-Net. InceptionResNetV2 is a CNN model with a combination of the Inception family as well as ResNet i.e it incorporates the residual connections. The residual function helps to overcome the degradation problem that arises in the deep structures as well as the learning time is minimized.

###### ii. UNet Segmentation with pre-trained DenseNet201 as its backbone

This model uses the concept of UNet segmentation where the encoder is changed with DenseNet 201 ultimately this model contains benefits of both DenseNet 201 as well as U-Net. This model is based on UNet++. DenseNet is a kind of CNN that trains the deeper network thereby vanishing gradient is prevented.

##### C. UNet++ Segmentation

The UNet++ segmentation model aims to bridge the semantic gap between the encoder as well as the decoder of the feature maps before concatenation. UNet++ is based on nested as well as dense skip connections. UNet++ captures fine-grained details of 2 Dimensional images effectively.

UNet++ has 3 addons compared to U-Net:

#### i. Redesigned Skip Pathways

It is added to bridge the semantic gap b/w the encoder as well as decoder sub paths. The use of convolutions layers is to reduce the semantic gap b/w the subnetwork's feature maps of the encoder & the decoder. Ultimately using UNet++ the output that we get from the preceding convolution layer that is of the same dense block is joined with their respective up-sampled output of the lower dense block.

#### ii. Dense Skip Connections

It is implemented to skip pathways b/w the encoder as well as the decoder. These dense blocks are motivated by DenseNet such that segmentation accuracy, as well as gradient flow, can be improved. It makes sure that all the previous feature maps are gathered as well as reach the current node as the dense convolution block with each skip pathway. Finally, this produces feature maps of full-resolution at different semantic levels.

#### iii. Deep Supervision

It is used such that the model can be pruned such that it adjusts to the model complexity and maintains balance b/w speed i.e the inference time as well as performance. For the accurate mode averaging the o/p's of all segmentation branches and for the fast mode from one of the segmentation branches, the final segmentation map is chosen.

#### Step 4: Optimization

Optimization is the process of making a model effective by using resources. In this paper Adam optimization is used, it is an extension to Stochastic gradient descent and it incorporates the momentum method & RMSprop, along with bias correction.

##### i. Momentum method

This technique is mainly used to accelerate by taking an exponentially weighted average of the gradients. By making use of averages, it converges towards the minima i.e in a faster way, as the gradients towards the dissimilar directions are ruled out.

##### ii. RMSprop - Root mean square

In this technique instead of the cumulative sum of squared gradients, the exponential moving average is taken. The main advantage of using this is that it is Computationally efficient, requires little memory, is suitable for problems that have many parameters and it requires little tuning.

The loss function calculates the entire error ultimately it adjusts the internal weights during backpropagation therefore the choice of loss function has a direct impact on the model's performance. For classification and segmentation, either Binary Cross Entropy or Dice Coefficient is chosen. In this paper, Binary Cross Entropy Dice loss is used and it is a combination of dice loss as well as Binary Cross-Entropy loss.

#### Step 5: Model Evaluation

In this step, the evaluation of the model is done. Cross-Validation is a technique for evaluating a deep learning model as well as testing its performance. Generally, Hold Out Cross-Validation is used it divides the dataset into two sets i.e training and testing but the main drawback is that if the dataset is not evenly distributed then the training set does not represent the testing set i.e both differ so much and either of them will be easier or harder. So, in this paper K Fold Cross-Validation is used, it splits the dataset into k-subsets and then the holdout method is repeated for k-times. Here every k subset is used as a test set and k-1 subsets are used for the training purpose.

### III. RESULTS AND DISCUSSION

#### 3.1 Pre-Processing

##### 3.1.1 CLAHE

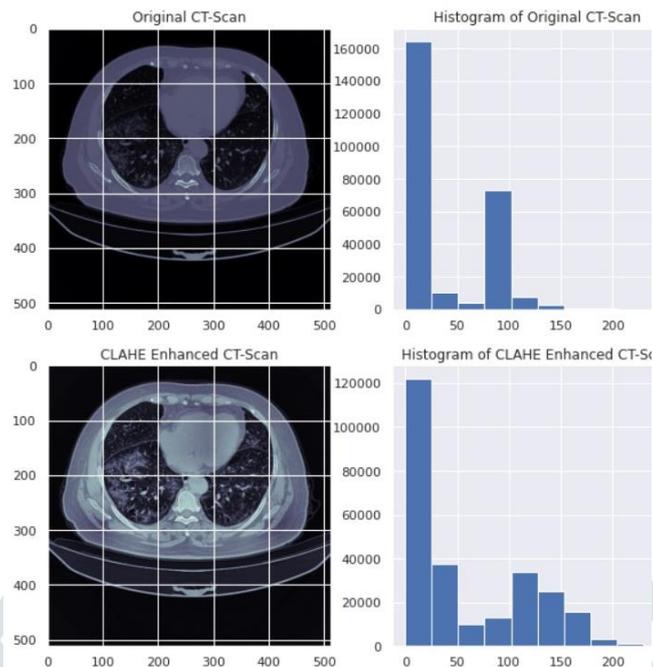


Figure 1: Lung CT image before and after pre-processing with a histogram plot

In the above figure comparison of random Lung CT images before and after pre-processing is done. On observation of the histogram, the x-axis represents pixel intensity value and the y-axis represents the number of pixels. It is very clear that certain features are enhanced. To conclude that CLAHE technique is more effective than the traditional methodology.

##### 3.1.2 Data Augmentation

After data augmentation, the training images is been increased ultimately the issue of overfitting is resolved.

##### 3.1.3 Cropping the images using the Contour technique

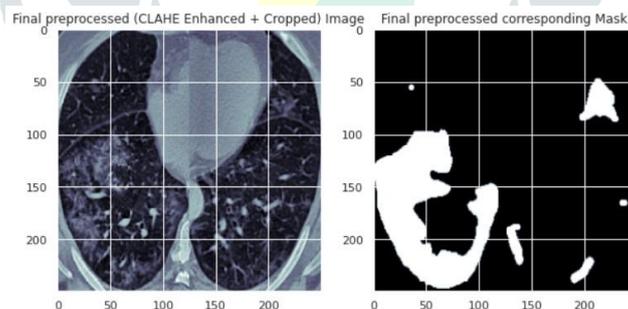


Figure 2: Cropped CT lung image

Observing figure 1 there are a lot of black spaces and the parts like the diaphragm are not needed for segmentation. In comparison with figure 1 with figure 2, it is clear that the unnecessary part is cropped by the contour technique ultimately the consumption of RAM as well as computational power is reduced.

##### 3.1.4 Removing images with no corresponding mask images

Initially, there were 2112 images, it was found that there are 498 images with no corresponding images after removing there are 1614 images. After removing images with no corresponding mask images improved the training performance.

#### 3.2 Model Evaluation

It is evaluated using K Fold Cross-Validation.

Classification and segmentation of medical images have class imbalance taking the default threshold leads to poor performance. So, in this paper to improve the performance the threshold is tuned to map the probabilities with the class label for that threshold value is set and then it is evaluated to obtain the optimal threshold.

Category	Actual lesion	Actual non-lesion
Predicted Lesion	TP	FP
Predicted non-lesion	FN	TN

A dice similarity coefficient is a statistical tool used to validate image segmentation.

$$DSC = 2 * \frac{Precision * Recall}{Precision + Recall}$$

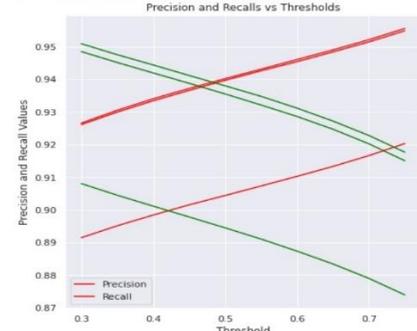
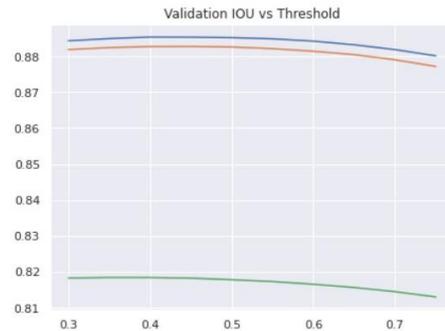
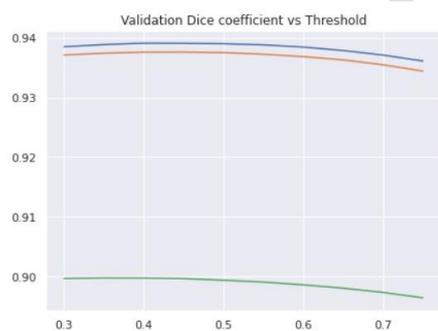
$$Precision = \frac{True_{positive}}{True_{positive} + False_{positive}} \quad Recall = \frac{True_{positive}}{True_{positive} + False_{negative}}$$

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + False_{positive} + True_{negative} + False_{negative}}$$

### 3.2.1 UNet Segmentation

3-fold Dices dataframe  
 Rows indices: Thresholds, Column indices: Dices

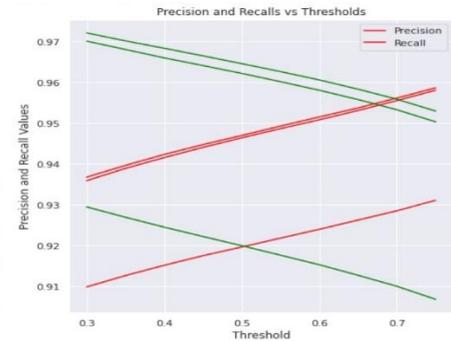
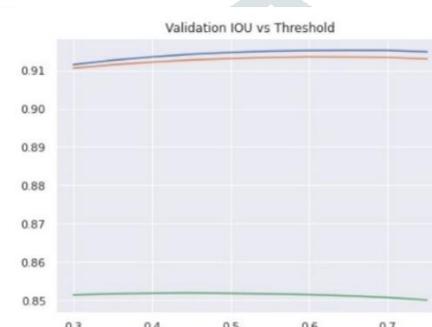
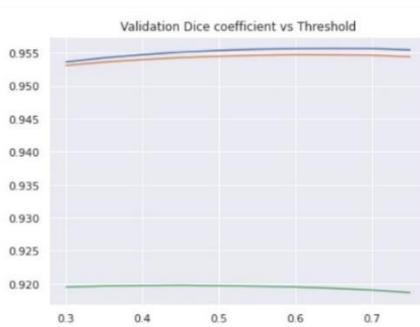
	1	2	3
0.30	0.932997	0.931911	0.896922
0.35	0.933737	0.932507	0.897210
0.40	0.934299	0.932929	0.897426
0.45	0.934714	0.933239	0.897496
0.50	0.935000	0.933482	0.897492
0.55	0.935189	0.933596	0.897462
0.60	0.935236	0.933592	0.897339
0.65	0.935199	0.933458	0.897138
0.70	0.935061	0.933196	0.896813
0.75	0.934723	0.932698	0.896359



### 3.2.2 UNet segmentation with pre-trained InceptionResNetV2 as its backbone

3-fold Dices dataframe  
 Rows indices: Thresholds, Column in

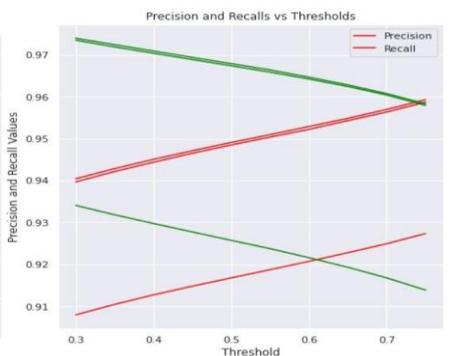
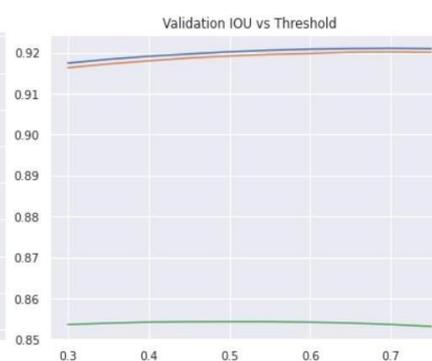
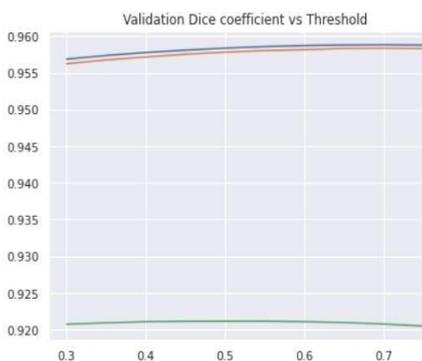
	1	2	3
0.30	0.954736	0.954132	0.916996
0.35	0.955190	0.954518	0.916927
0.40	0.955467	0.954762	0.916823
0.45	0.955672	0.954831	0.916673
0.50	0.955749	0.954864	0.916531
0.55	0.955714	0.954783	0.916225
0.60	0.955555	0.954612	0.915832
0.65	0.955282	0.954330	0.915336
0.70	0.954940	0.953882	0.914746
0.75	0.954396	0.953283	0.914040



### 3.2.3 UNet segmentation with pre-trained DenseNet201 as its backbone

3-fold Dices dataframe  
 Rows indices: Thresholds, Column in

	1	2	3
0.30	0.956120	0.954923	0.920440
0.35	0.956074	0.954860	0.920137
0.40	0.955909	0.954745	0.919818
0.45	0.955716	0.954579	0.919435
0.50	0.955498	0.954316	0.918997
0.55	0.955200	0.954020	0.918513
0.60	0.954852	0.953642	0.918035
0.65	0.954423	0.953202	0.917453
0.70	0.953848	0.952635	0.916732
0.75	0.953087	0.951896	0.915828



### 3.2.4 UNet ++ Segmentation

3-fold Dices dataframe  
 Rows indices: Thresholds, Column ir

	1	2	3
0.30	0.922808	0.920516	0.894204
0.35	0.923314	0.921073	0.894368
0.40	0.923632	0.921337	0.894442
0.45	0.923829	0.921513	0.894286
0.50	0.923853	0.921487	0.894107
0.55	0.923710	0.921349	0.893712
0.60	0.923374	0.921099	0.893227
0.65	0.922897	0.920686	0.892610
0.70	0.922255	0.920057	0.891774
0.75	0.921313	0.919088	0.890572

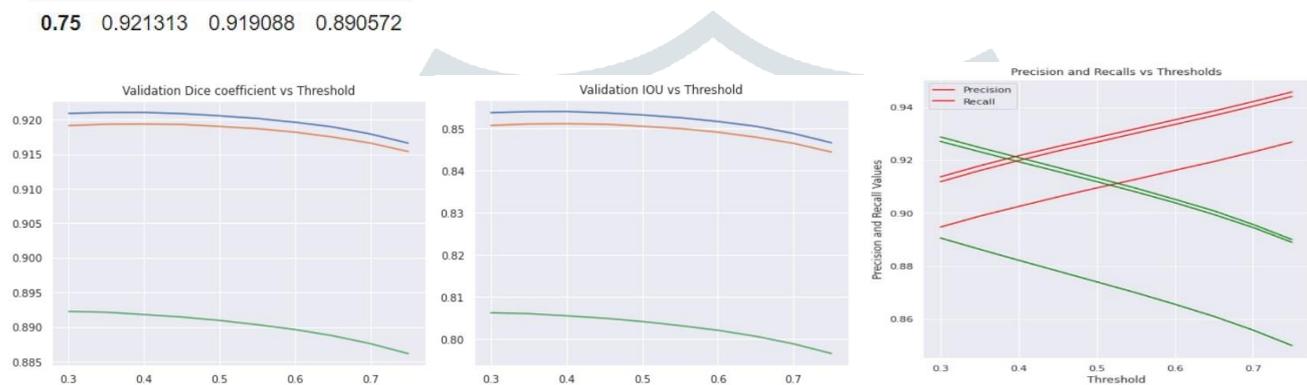


Table 1: Comparison of performance of different models

MODEL	PRECISION	RECALL	DC	IOU	ACCURACY
UNet	94.218	95.616	93.523	87.859	96.208
UNet + InceptionResNetV2	96.493	98.807	95.575	91.534	96.293
UNet + DenseNet 201	96.997	95.763	95.883	91.602	95.306
UNet ++	94.201	93.890	92.385	85.875	96.0912

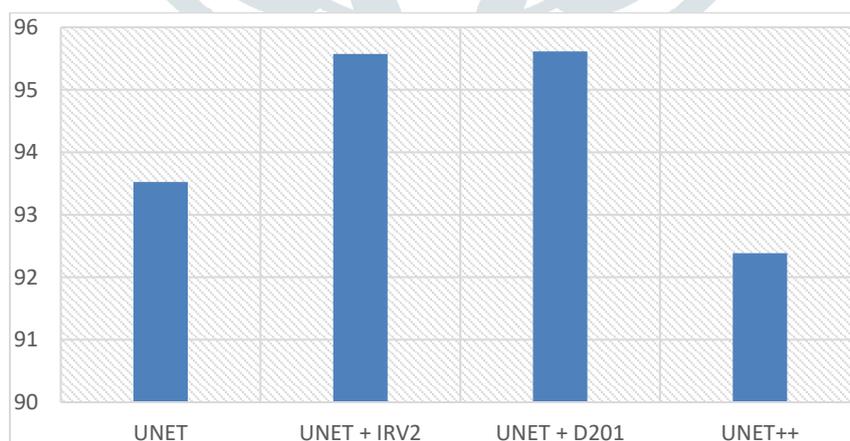


Figure 3: Comparison of best Dice Coefficient of UNet, UNet with InceptionResNetv2, UNet with densenet201 & UNet++

After the experimental analysis, UNet segmentation got a Dice Coefficient of 93.524, UNet segmentation with InceptionResNetV2 as its backbone got a Dice Coefficient of 95.575, UNet segmentation with DenseNet 201 as its backbone got a Dice Coefficient of 95.884 and UNet ++ got a Dice Coefficient of 92.385. To conclude UNet segmentation with pre-trained DenseNet201 as its backbone got the best Dice Coefficient of 95.884.

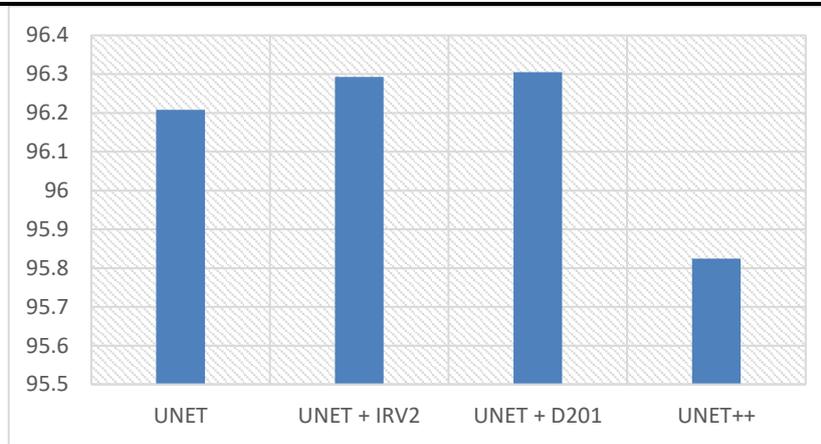
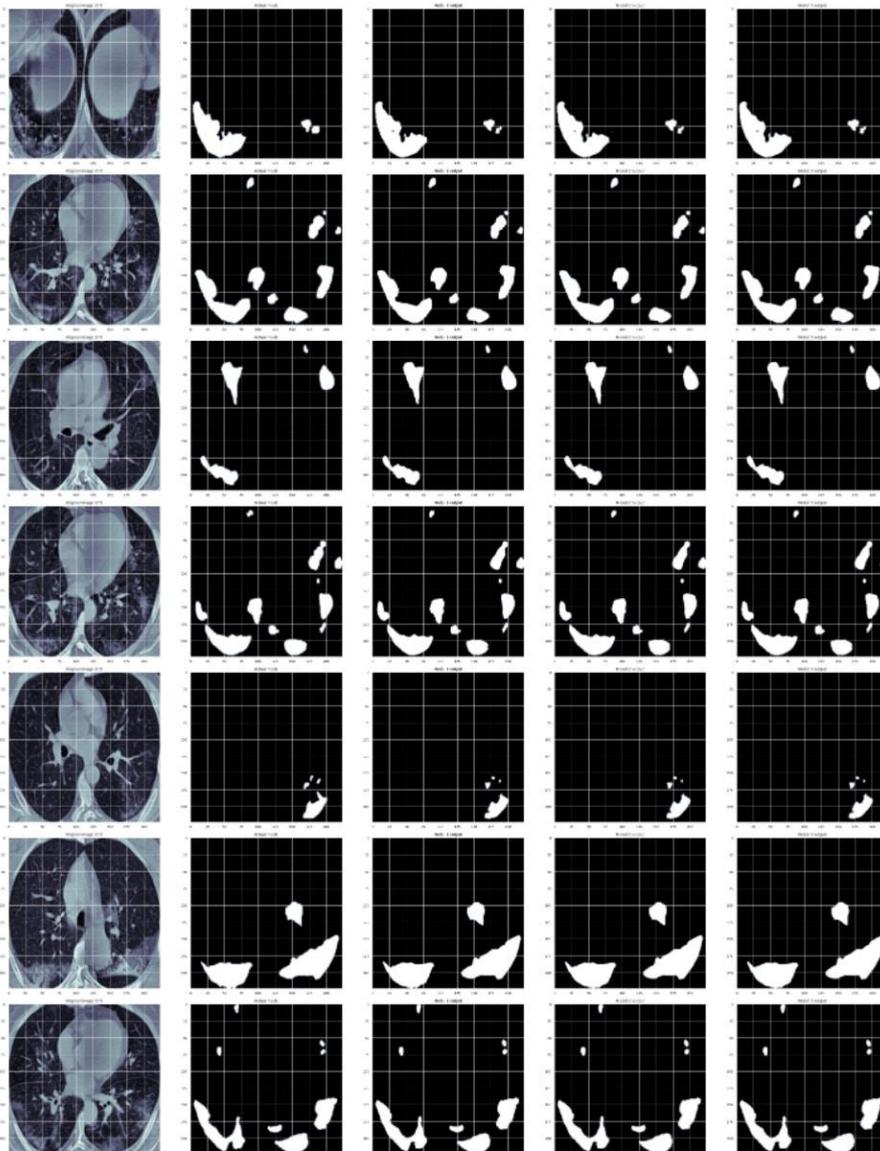


Figure 4: Comparison of best Dice Coefficient of UNet, UNet with InceptionResNetv2, UNet with densenet201 & UNet++

After the experimental analysis, UNet segmentation got an accuracy of 96.208, UNet segmentation with InceptionResNetV2 as its backbone got an accuracy of 96.293, UNet segmentation with DenseNet 201 as its backbone got an accuracy of 95.306 and UNet ++ got an accuracy of 96.091. To conclude UNet segmentation with pre-trained DenseNet201 as its backbone got the best accuracy of 95.612.

### 3.3 Output



### REFERENCES

- [1] S. V. Kogilavani, J. Prabhu, R. Sandhiya, M. Sandeep Kumar, Uma Shankar Subramaniam, Alagar Karthick, M. Muhibbullah, and Sharmila Banu Sheik Imam, "Covid-19 Detection Based on Lung CT Scan using Deep Learning Techniques", Hindawi, 2022.

- [2] Lucas O. Teixeira, Rodolfo M. Pereira, Diego Bertolini, Luiz S. Oliveira, Loris Nanni, George D. C. Cavalcanti and Yandre M. G. Costa, "Impact Of Lung Segmentation on the Diagnosis and Explanation of Covid-19 in Chest X-Ray Images", MDPI, 2022.
- [3] Daniel Arias-Garzón, Jesús Alejandro Alzate-Grisales, Simon Orozco-Arias, Harold Brayan Arteaga-Arteaga, Mario Alejandro Bravo-Ortiz, Alejandro Mora-Rubio, Jose Manuel Saborit-Torres, Joaquim Ángel Montell Serrano, Mariade la Iglesia Vayá, OscarCardona-Morales and ReinelTabares-Soto, "Covid-19 Detection in X-Ray Images using Convolutional Neural Networks", ELSEVIER, 2021.
- [4] RT Subhalakshmi, S Appavu alias Balamurugan, S Sasikala, "Deep Learning based Fusion Model for Covid-19 Diagnosis and Classification using Computed Tomography Images", SAGE journal, 2021.
- [5] Anuja Bhargava, Atul Bansal, and Vishal Goyal, "Machine Learning based Automatic Detection of Novel Coronavirus (Covid-19) Disease", MULTIMED TOOLS APPL, 2022.
- [6] Xiangyu Zhao, Peng Zhang, Fan Song, Guangda Fan, Yangyang Sun, Yujia Wang, Zheyuan Tian, Luqi Zhang, and Guanglei Zhang, "D2A U-Net: Automatic Segmentation of Covid-19 Lesions from CT Slices with Dilated Convolution and Dual Attention Mechanism", Arxiv, 2021.
- [7] Tongxue Zhou, Stéphanie Canu and Su Ruan, "Automatic Covid-19 CT Segmentation using U-Net Integrated Spatial and Channel Attention Mechanism", WILEY, 2020.
- [8] Alex Noel Joseph Raj, Haipeng Zhu, Asiya Khan, Zhemin Zhuang, Zengbiao Yang, Vijayalakshmi G. V. Mahesh, and Ganesan Karthik, "Detection of Covid-19 from CT Lung scans using Transfer Learning", HINDAWI, 2021.
- [9] Sahil Lawton and Serestina Viriri, "ADID UNet—A Segmentation Model for Covid-19 Infection from Lung CT Scans", PEER J COMPUT SCI, 2021.
- [10] Yazan Qiblawey, Anas Tahir, Muhammad E. H. Chowdhury, Amith Khandakar, Serkan Kiranyaz, Tawsifur Rahman, Nabil Ibtihaz, Sakib Mahmud, Somaya Al Maadeed, Farayi Musharavati and Mohamed Arselene Ayari, "Detection and Severity Classification of Covid-19 in CT images using Deep Learning", MDPI, 2021.
- [11] Sahil Lawton and Serestina Viriri, "Detection of COVID-19 from CT Lung Scans Using Transfer Learning", Computational Intelligence and Neuroscience, Hindawi 2021.
- [12] Mundher Mohammed Taresh, Ningbo Zhu, Talal Ahmed Ali Ali, Asaad Shakir Hameed, and Modhi Lafta Mutar, "Transfer Learning to Detect COVID-19 Automatically from X-Ray Images Using Convolutional Neural Networks", International Journal of Biomedical Imaging, Hindawi 2021.
- [13] Mohammad Shorfuzzaman, Mehedi Masud, Hesham Alhumyani, Divya Anand, and Aman Singh, "Artificial Neural Network-Based Deep Learning Model for COVID-19 Patient Detection Using X-Ray Chest Images", Journal of Healthcare Engineering, Hindawi, 2021.
- [14] Mohd Zulfaezal Che Azemin, Radhiana Hassan, Mohd Izzuddin Mohd Tamrin, and Mohd Adli Md Ali, "COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings", International Journal of Biomedical Imaging, Hindawi 2020.
- [15] Lin Li, Lixin Qin, Zeguo Xu, Youbing Yin, Xin Wang, Bin Kong, Junjie Bai, Yi Lu, Zhenghan Fang, Qi Song, Kunlin Cao, Daliang Liu, Guisheng Wang, Qizhong Xu, Xisheng Fang, Shiqin Zhang, Juan Xia, Jun Xia, "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy", RSNA, Radiology, 2020.
- [16] Vinay Arora, Eddie Yin-Kwee Ng, Rohan Singh Leekha, Medhavi Darshan, Arshdeep Singh "Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan", pubmed.gov, 2021.
- [17] Rekha Rajagopal, "Comparative Analysis of COVID-19 X-ray Images Classification Using Convolutional Neural Network, Transfer Learning, and Machine Learning Classifiers Using Deep Features", Springer, 2021.