2 Stage-UNET For Pneumothorax Segmentation

By Ayush Singhania

Problem Statement

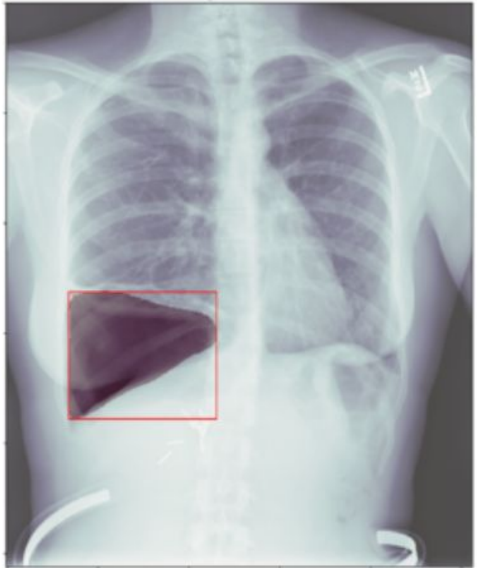
Given a chest X-Ray, we want to build an image processing pipeline that can automatically detect and segment out the regions of the X-Ray that are affected by Pneumothorax (collapsed lung). We will be using a 2 stage UNET for segmenting the X-Rays. This process of identifying the affected regions is primarily done manually by trained radiologists as of now, but such image processing pipelines can make the entire process much faster and economical for the patients and thus help in early diagnosis and saving precious lives of the patients.

Objectives

- Segmentation of X-Rays to find Pneumothorax affected regions.
- Perform Two Stage training on the chest X-ray training dataset from the SIIM-ACR Pneumothorax Segmentation Challenge.
- Testing the trained model on the test dataset and compare different pipelines to find the optimal segmentation pipeline.
- Document and report our findings and compare them with that of the reference paper.

Objectives

Image
Processing
Pipeline



Input X-Ray

Segmented Output

Dataset

- Chest X-ray dataset from the SIIM-ACR Pneumothorax Segmentation Challenge.
- Collected using the Cloud Healthcare API.
- Uses the Digital Imaging and Communications in Medicine (DICOM) format for encoding the metadata along with the X-Ray image in a single file.
- Training dataset has the X-Ray and annotations in the form of masks specifying the region affected by pneumothorax.
- 10,065 training and 1,830 testing images.

Method Overview

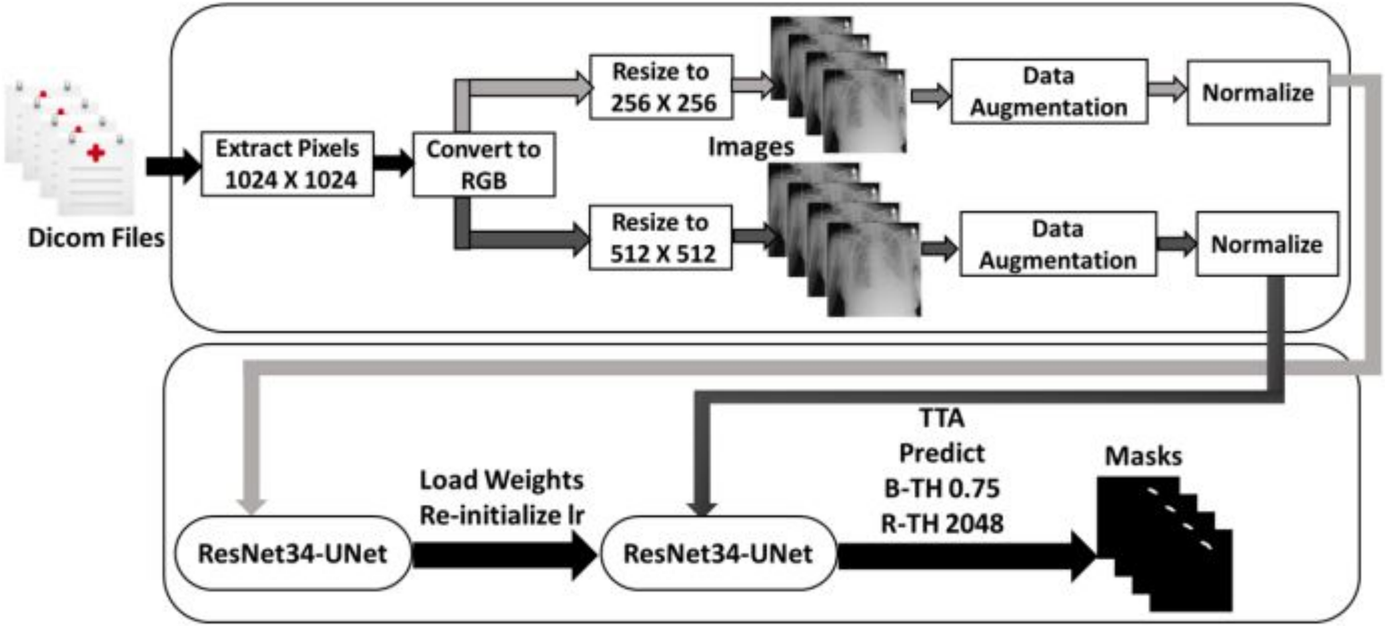
The entire project timeline can be broken down into 3 phases:

1. Data Processing and Augmentation : where we modify the given dataset to create more robust and diversified training data.

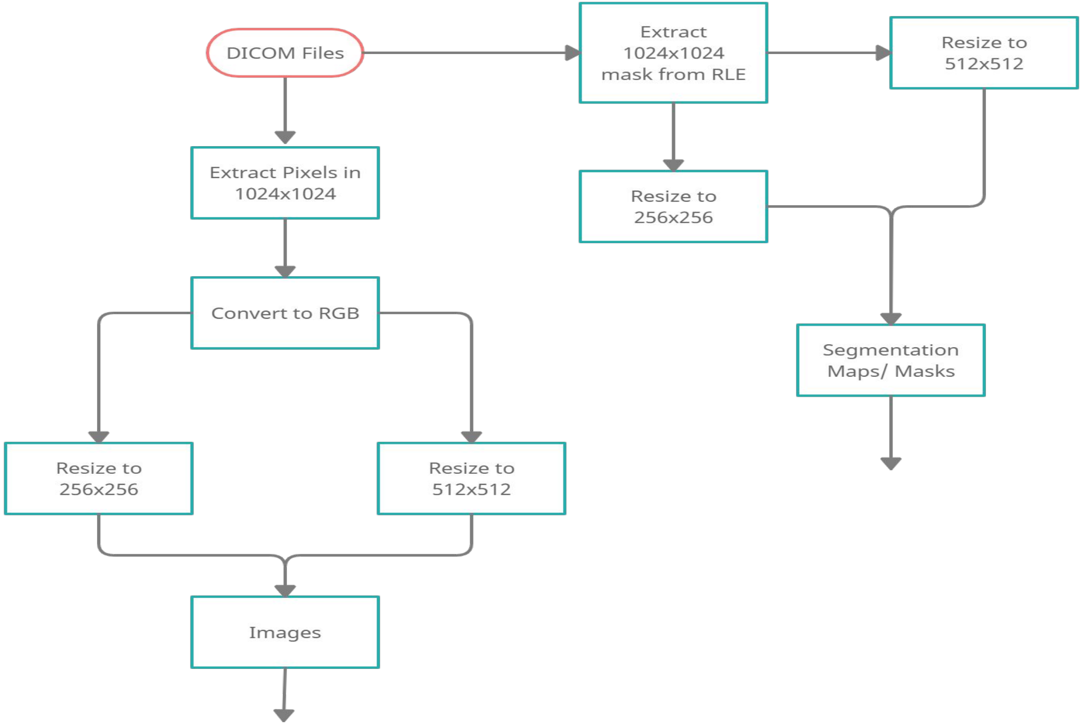
2. Training the UNET model

3. Testing the UNET model.

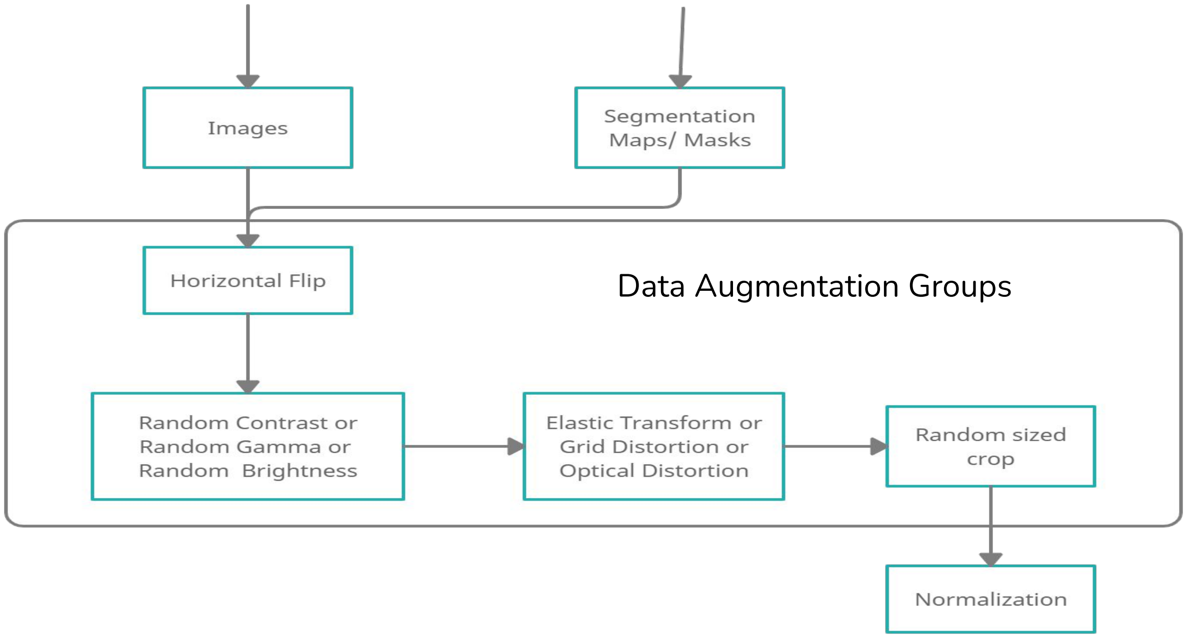
Method Overview



Method Overview (Data Preprocessing and Augmentation)



Method Overview (Data Preprocessing and Augmentation)



Data Augmentation and Normalization that helps in generalization of error.

After Normalization, both pixel intensity values and mask values are mapped to 0 or 1 instead of 0 or 255.

Method Overview (Training Phase)

- U-Net as a segmentation model.
- Transfer Learning with ResNet-34 pre-trained on ImageNet dataset as the backbone.
- 2 stage training.
- 1st stage : Train the model on lower resolution (256x256) images as obtained from Data Processing phase.
- 2nd stage: Reinitialise the learning rate and further train the model from the 1st stage on higher resolution images (512x512).

Method Overview (Training Phase)

- Combination of Binary Cross Entropy(BCE) and Dice Loss(DSCL) for the Loss function.

$$\mathrm{BCE} = -1/N \sum_{i=1}^{N} [y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

N = no. of samples, y = ground truth, p = predicted value

$$\mathrm{DSCL} = 1 - \frac{2 \times |\mathrm{X} \bigcap \mathrm{Y}|}{|\mathrm{X}| + |\mathrm{Y}|}$$

X = predicted set of pixels, Y = ground truth

Loss function is given by: $\text{BCE-Dice Loss} = \mathrm{BCE} + \mathrm{DSCL}$

Method Overview (Training Phase)

Test out 5 different training pipelines to optimise the model:

1. To train the network of grayscale images with pre-trained weights, 1-channel data has to be mapped to 3-channels. Adam optimizer and ReduceLROnPlateau9 scheduler are going to reduce the learning rate by a factor of 0.2 if no improvement is seen for 5 patience based on the validation loss. The Binarization Threshold (B-TH) is 0.5 for the prediction.

2. Segmentation network on the 256 × 256 resolution images with 3-channels for 35 epochs with 40 batch size is going to be trained. Moreover, the B-TH is 0.55 for the prediction.

3. The weights from the above model (pt. 2) is going to be loaded and the learning rate is going to be reinitialized. Then, we train the network on the 512 × 512 resolution images with 3-channels for 10 epochs with batch size 14. The B-TH is 0.55 for the prediction.

Method Overview (Training Phase)

4. Train on 256 × 256 resolution images, and 3-channels for 60 epochs with batch size 40 is trained. Then, the weights are going to be loaded and the learning rate is going to be reinitialized to retrain the network on the 512 × 512 resolution images and 3-channels for 29 epochs with batch size 14.

5. Train on 256 × 256 resolution images, and 3-channels for 100 epochs with batch size 40 is trained. Then, the weights are going to be loaded and the learning rate is going to be reinitialized to retrain the network on the 512 × 512 resolution images and 3-channels for 70 epochs with batch size 14.

Method Overview (Testing Phase)

1. Around 2k images were considered. The images were resized to 512×512, and converted to RGB colors with normalization.

2. Image post processing - Apply horizontal flip, Test time augmentation to test images.

3. Use the Removal Threshold (R-TH) for each predicted mask to reduce false positives. The components that are less than the minimum size of pixels are removed.

4. The final prediction is the average prediction for all images.

Method Overview (Testing Phase)

- IOU Score - Area of overlap between Ground truth and Predicted segmentation divided by union of between them.

$$\mathrm{IoU}(P_{\mathrm{true}}, P_{\mathrm{predicted}}) = \frac{P_{\mathrm{true}} \bigcap P_{\mathrm{predicted}}}{P_{\mathrm{true}} \bigcup P_{\mathrm{predicted}}}$$

- DSC Score - Pixel-wise agreement between a predicted segmentation and its corresponding ground truth. DSC is computed as follows. X - Predicted set of pixels, Y- Ground truth set of Pixels. Calculated Mean DSC score.

$$\mathrm{DSC}(X, Y) = \frac{2 \times |X \bigcap Y|}{|X| + |Y|}$$

Method Overview (Testing Phase)

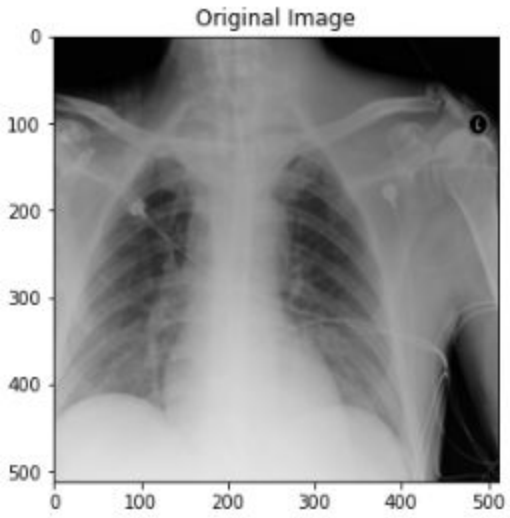
1. Performed TTA and Horizontal flip, random contrast, random brightness and random sized crop, elastic transform, grid distortion, optical distortion together along with R-TH and B-TH.

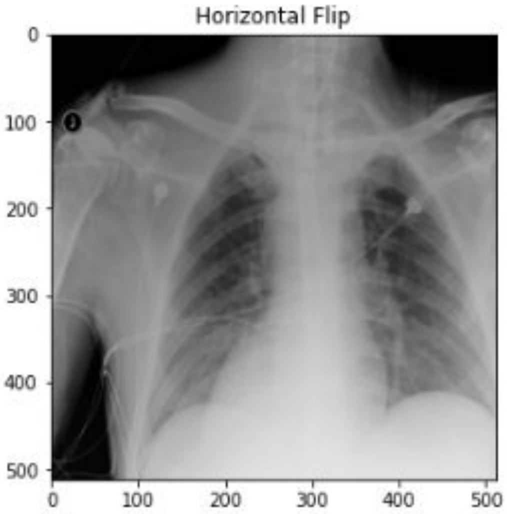
2. Test-time augmentation is the application of data augmentation techniques normally used during training when making predictions.

3. Specifically, it involves creating multiple augmented copies of each image in the test set, having the model make a prediction for each, then returning an ensemble of those predictions.

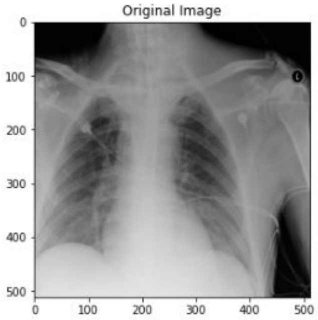
4. Testing was done on 2k images in batches of size 5. And IOU and DSC scores were calculated.

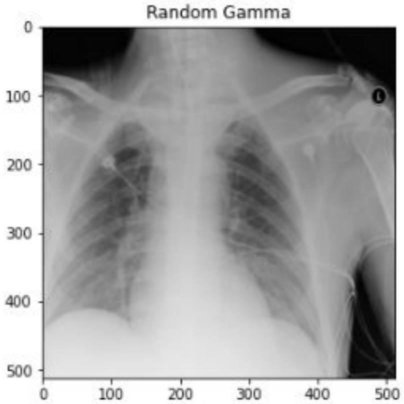
Results
Data Preprocessing and Augmentation



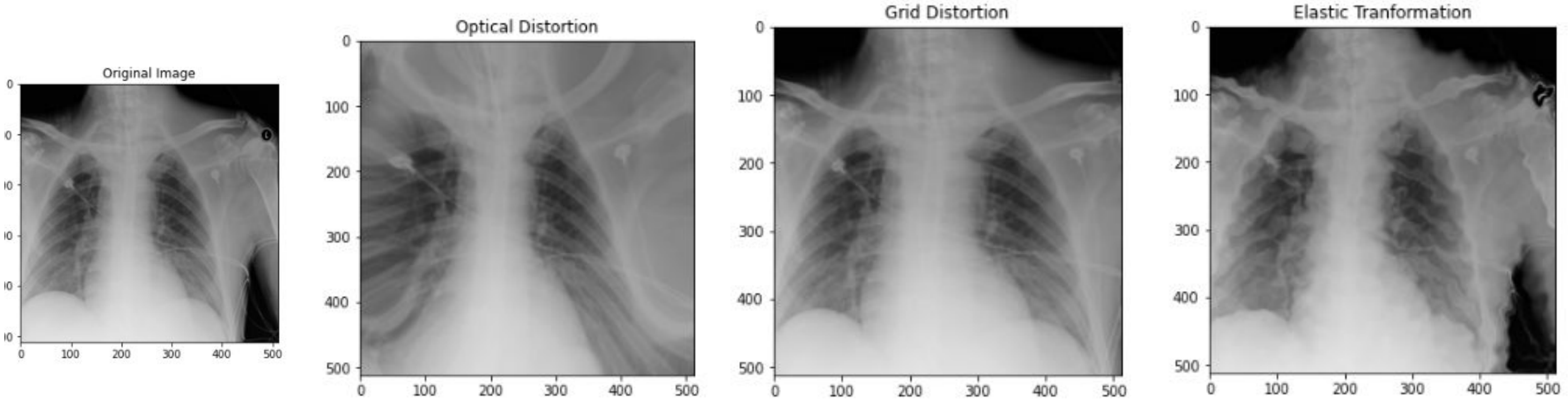


Results
Data Preprocessing and Augmentation





Results
Data Preprocessing and Augmentation



Results
Data Preprocessing and Augmentation

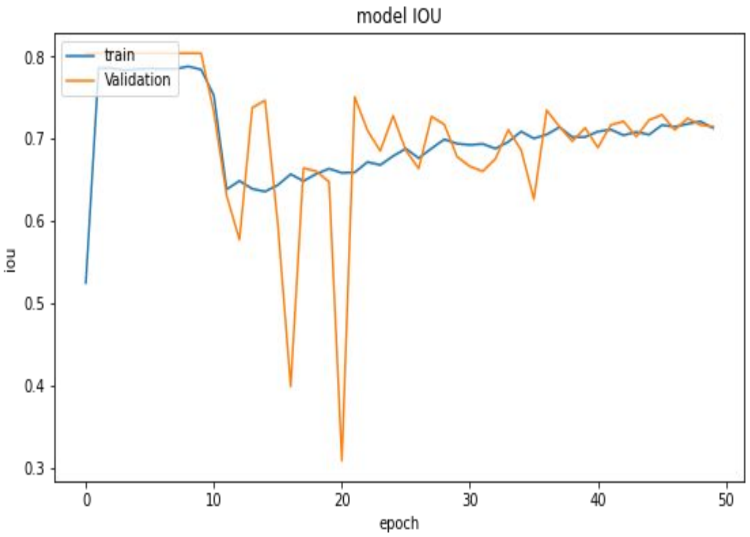
Results
Training Phase

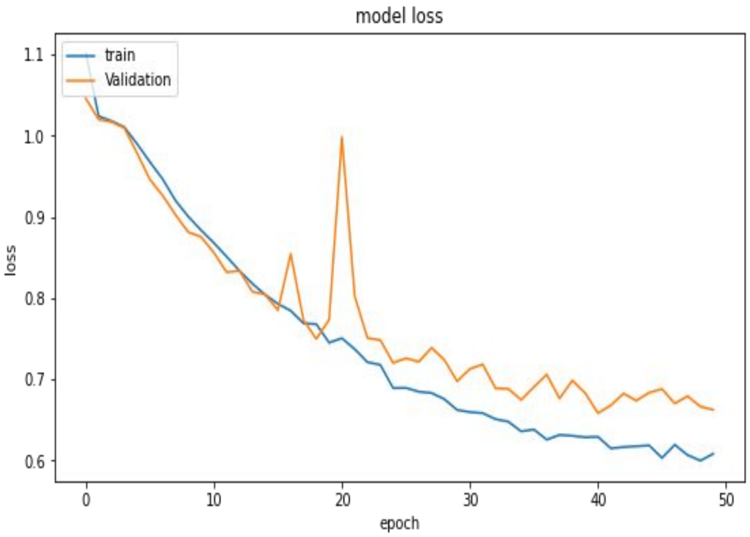
Experiment	SWA	B-TH	Channels	256 × 256		512 × 512	
				Batch size	Epochs	Batch size	Epochs
Exp1	✗	0.5	1	40	50	-	-
Exp2	✓	0.55	3	40	35	-	-
Exp3	✓	0.55	3	40	35	14	10
Exp4	✓	0.75	3	40	60	14	29
Exp5	✓	0.75	3	64	100	16	70

Results
Stage-1 Training IoU Scores

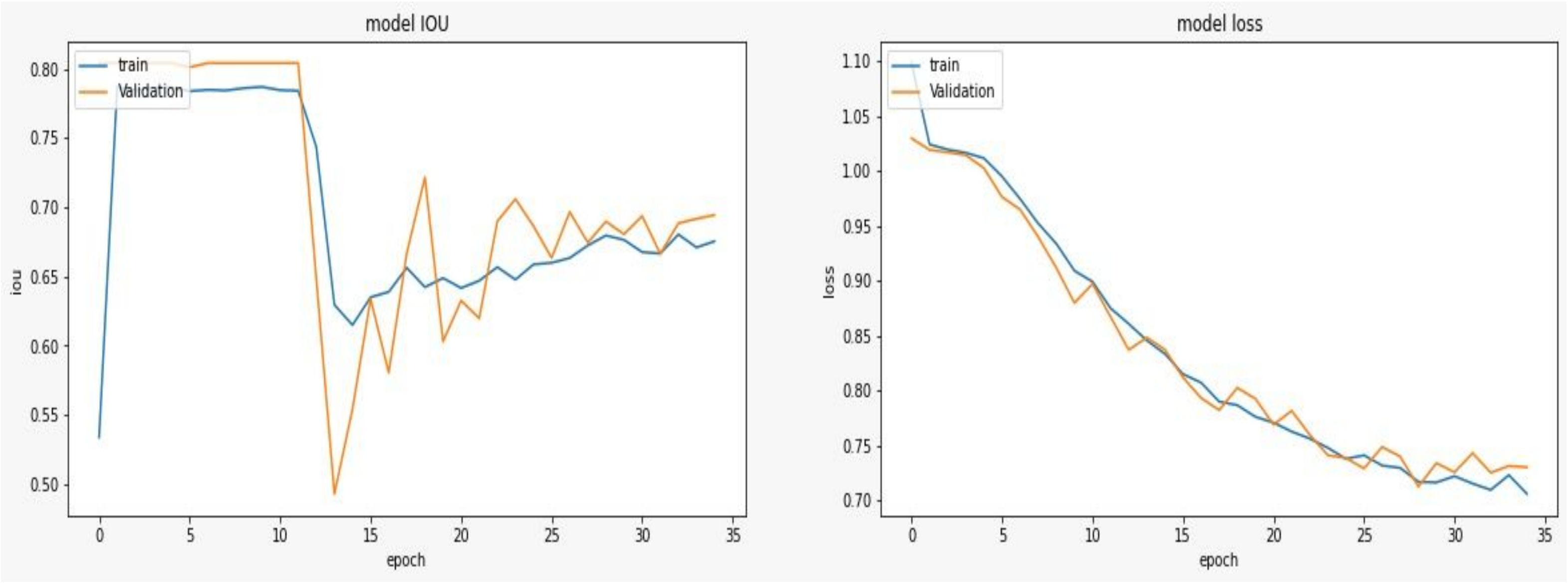
Exp No.	From the Study	From our experiments
Exp1	0.7242	0.7149
Exp2	0.7349	0.7223
Exp3	0.7105	0.7023
Exp4	0.7422	0.7366
Exp5	0.7638	0.7531

Results
Stage-1 Experiment-1

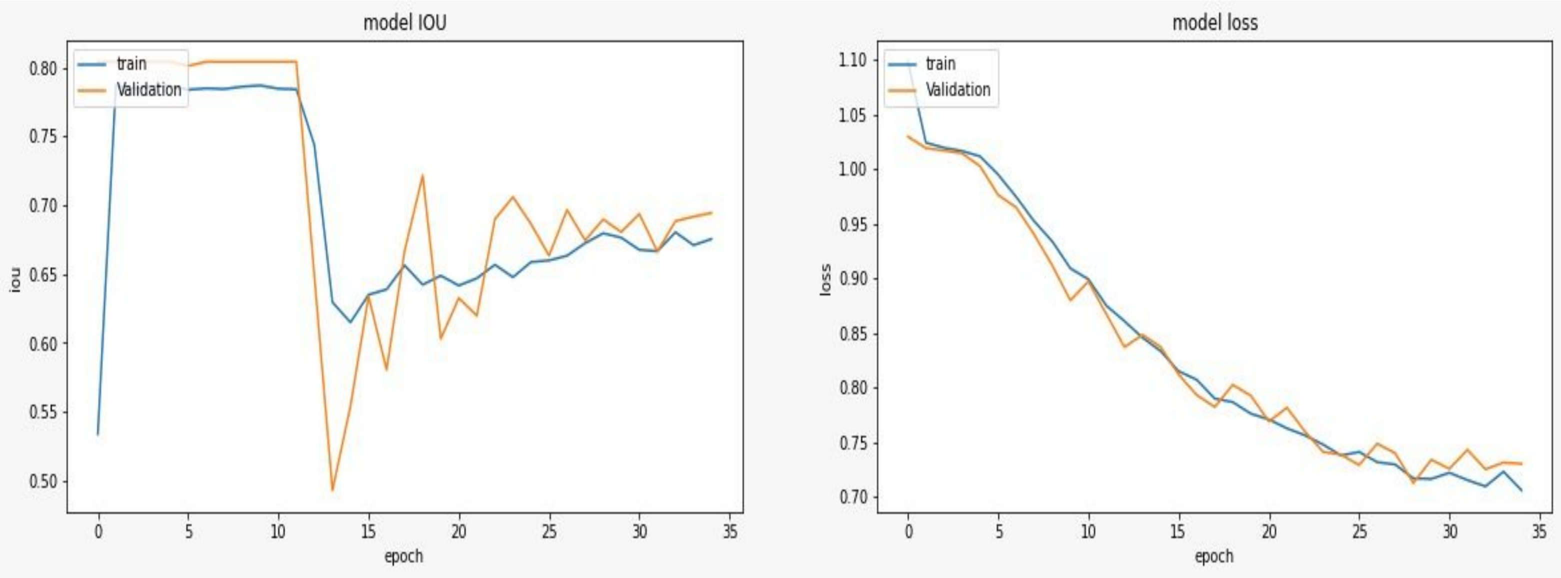




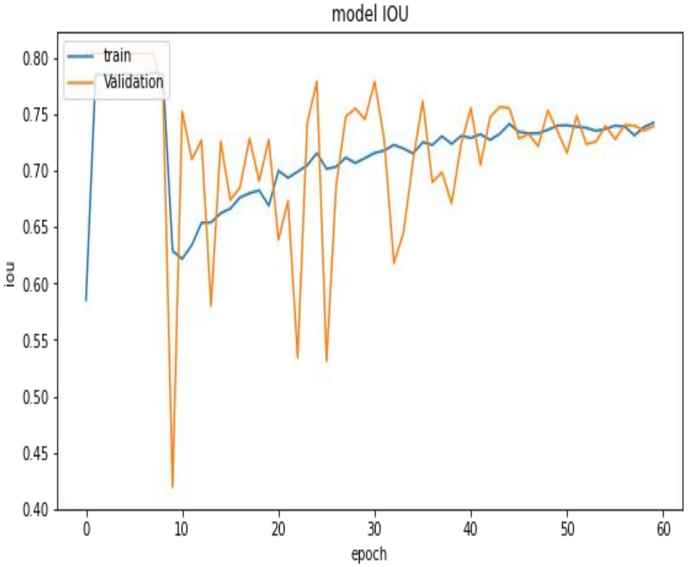
Results
Stage-1 Experiment-2

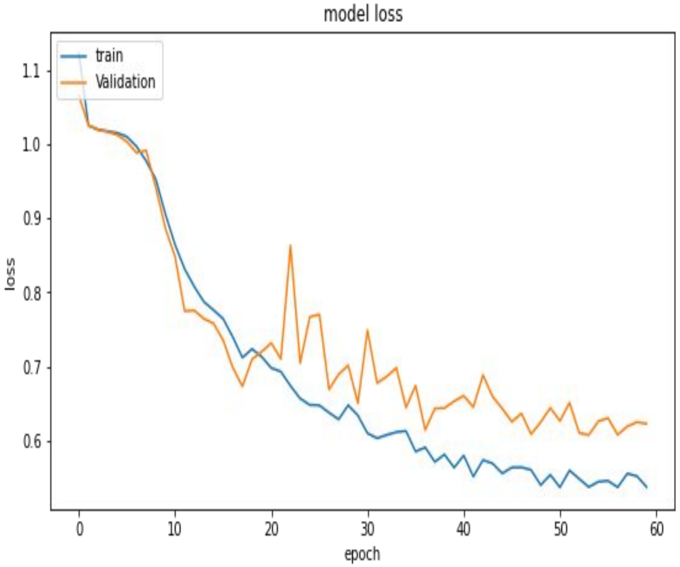


Results
Stage-1 Experiment-3

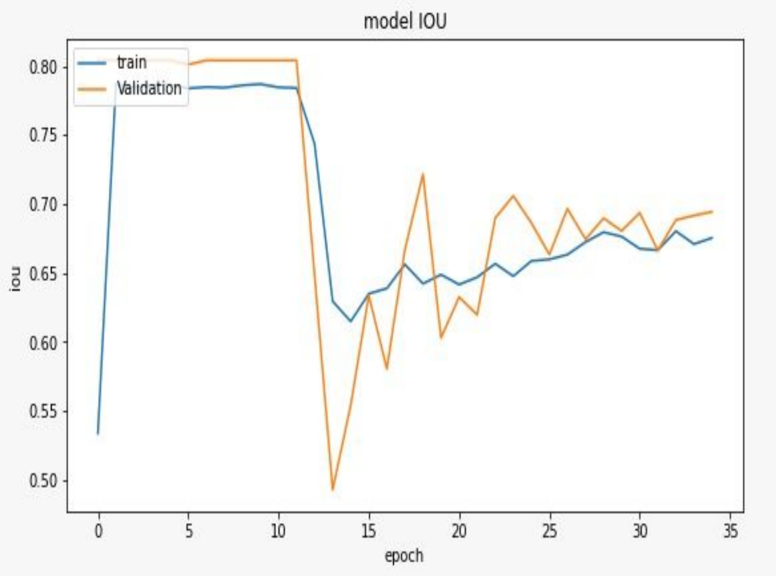


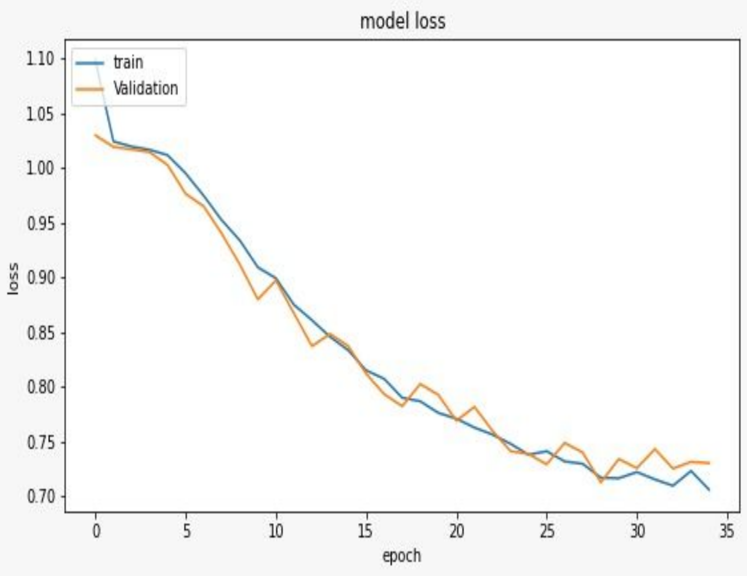
Results
Stage-1 Experiment-4





Results
Stage-1 Experiment-5

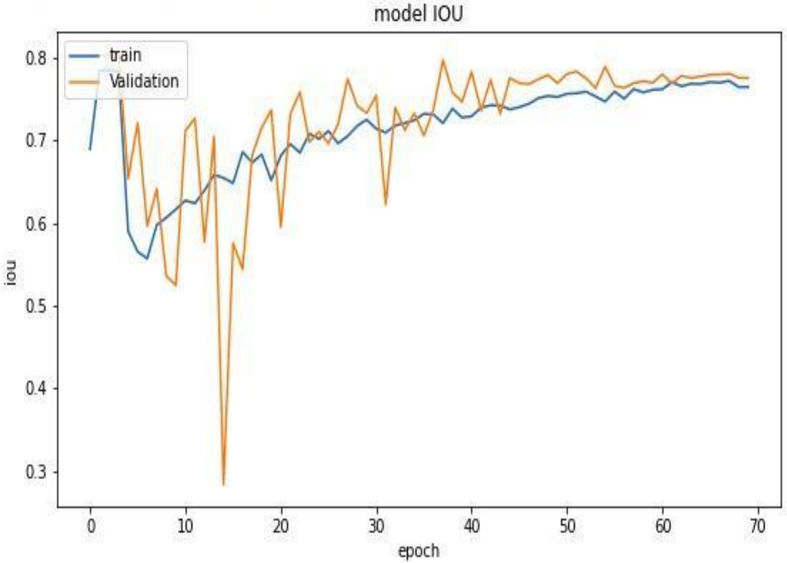


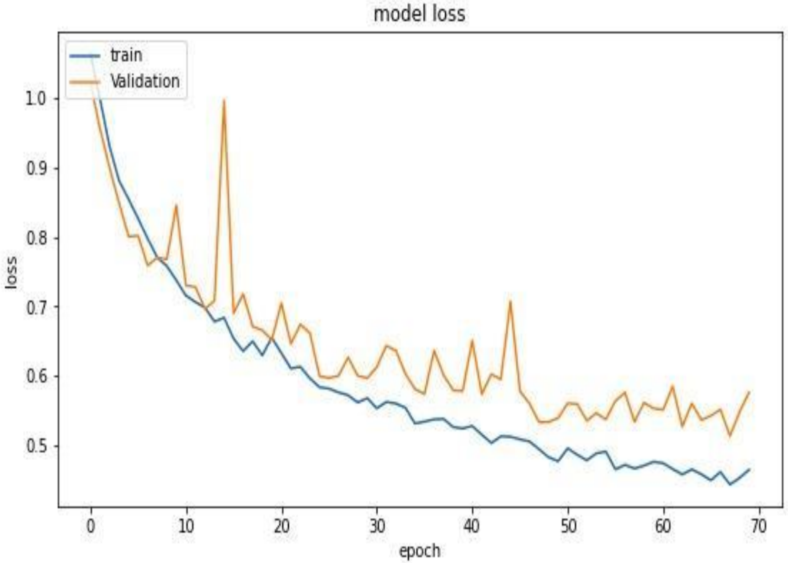


Results
Stage-2 Training IoU Scores

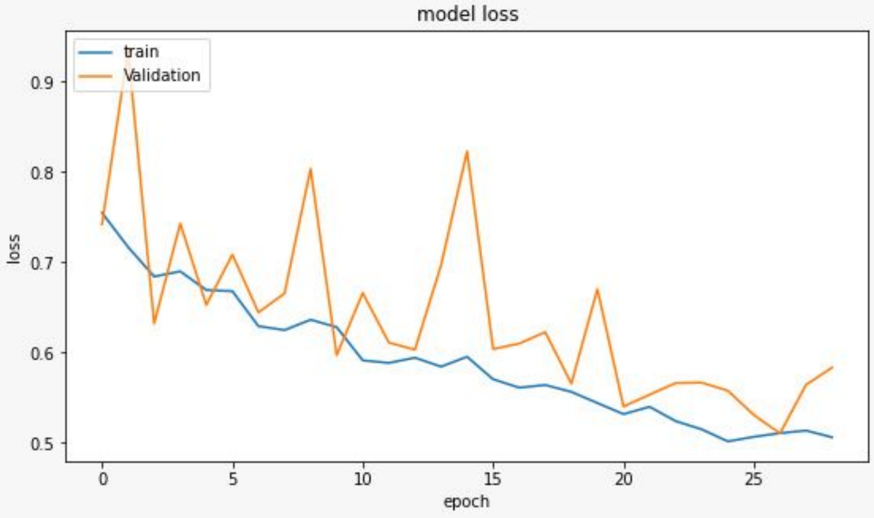
Exp No.	From the Study	From our experiments
Exp3	0.7844	0.7701
Exp4	0.7842	0.7724
Exp5	0.7850	0.7846

Results
Stage-2 Experiment-3

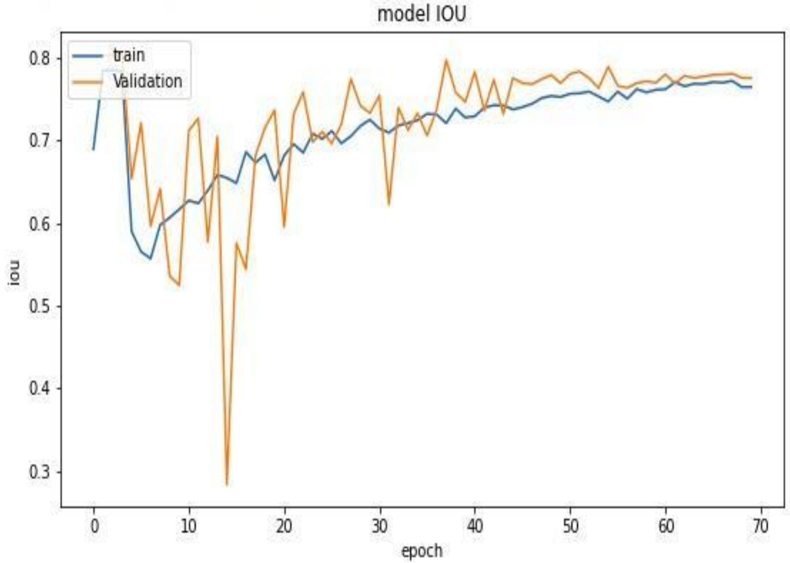


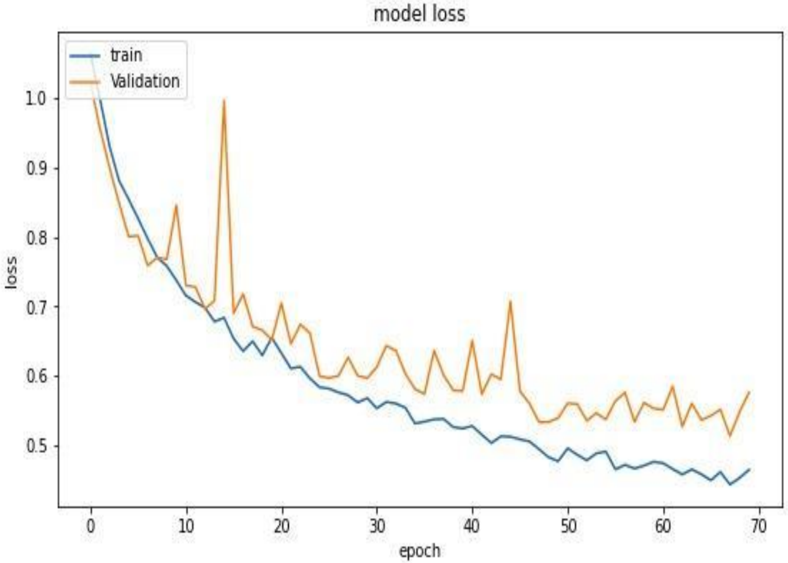


Results
Stage-2 Experiment-4



Results
Stage-2 Experiment-5





Testing Phase - Scores from experiments

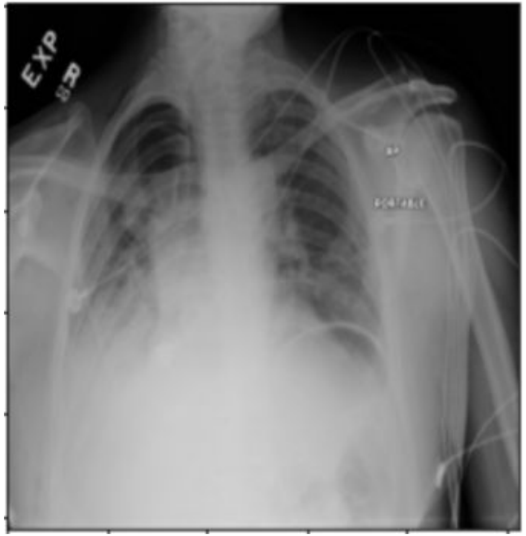
R-TH took values 1024, 2048, 3072, 4096. B-TH took values 0.2 to 0.9 with step size 0.001.

The best results were obtained for R-TH = 3072 and B-TH=0.48. Below listed are the scores obtained for the R-TH and B-TH.

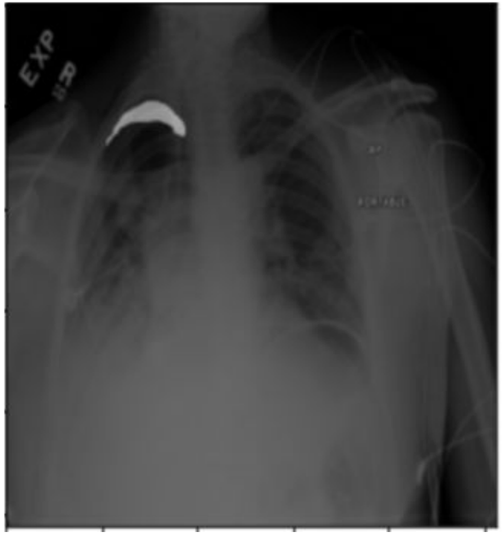
Metric	**From the Study**	**From our experiments**	**Image-Post Processing**
IOU	0.7850	0.7807	Horizontal Flip & R-TH and B-TH
DSC	0.8356	0.8347	Only Horizontal Flip.

Results- Test phase

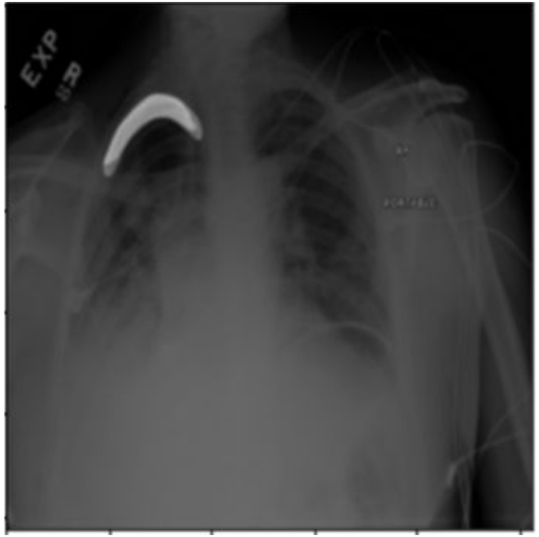
Original



Ground Truth



Predicted



Results- Test phase

Original

Ground Truth

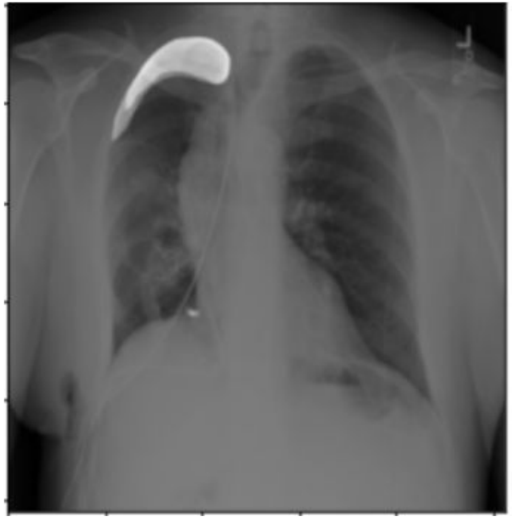
Predicted

Results- Test phase

Original

Ground Truth

Predicted



References

- The 2-ST UNET for Pneumothorax Segmentation in Chest X-Rays using ResNet34 as a backbone for UNET, Ayat Abedalla, Malak Abdullah, Mahmoud Al-Ayyoub , and Elhadj Benkhelifa, Jordan University of Science and Technology, Irbid, Jordan, September 9, 2020.

Thank You!