Report
Assignment 5b

Introduction

The problem is of multi class imbalanced detection, where one of the class has very small samples in the corpus. It was required to use focal loss with logits to make predictions with imbalance multi labelled classes. We used Focal loss with binary cross entropy with logits Loss for the purpose. Focal loss introduces two new parameters for the loss calculations. i.e., gamma and alpha. One major problem was faced during the implementation of focal loss that the loss gradients become infinite due to iterations. The loss gradient explodes after some iterations. It was figured out to be the gamma as with the backpropagation gamma, which was in the power in forward pass, multiplied and explodes the gradients. This problem was taken care by the logarithmic approach which binds the gradient and do not cause it to shoot.

Experimental Setup

Following experiments are performed,

Experiment No. 1

The experiment is with VGG loss, Focal Loss, with gamma =2, alpha = 0.25, learning rate = 0.01

The results of the experiments are as follows

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3800	200
Actual covid-19	400	1800

	Predicted Not normal	Predicted normal
Actual Not normal	1816	384
Actual normal	181	3819

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4109	91
Actual pneumonia	924	1076

F1 Score is : 0.859

Accuracy : 88.27

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	416	184
Actual covid-19	13	15

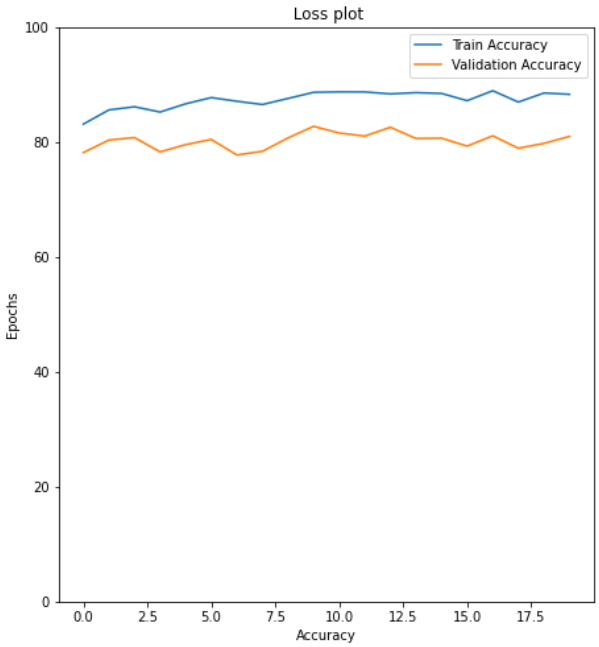
	Predicted Not normal	Predicted normal
Actual Not normal	188	40
Actual normal	16	384

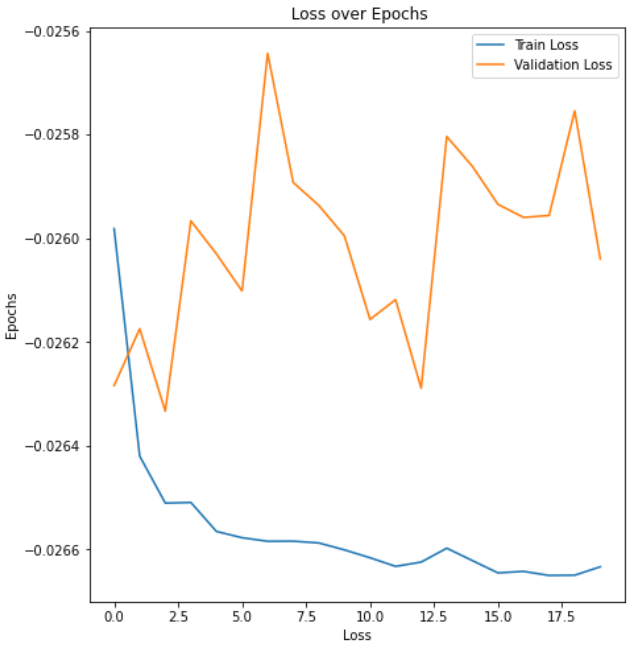
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	421	7
Actual pneumonia	86	114

F1 Score is : 0.747

Accuracy : 81.63

The loss and accuracy curves are as follows





Experiment 2

Now the loss is binary cross entropy with the same settings

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3589	411
Actual covid-19	252	1948

	Predicted Not normal	Predicted normal
Actual Not normal	1947	253
Actual normal	415	3585

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3675	525
Actual pneumonia	223	1777

F1 Score is : 0.875

Accuracy : 88.82

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	372	228
Actual covid-19	12	16

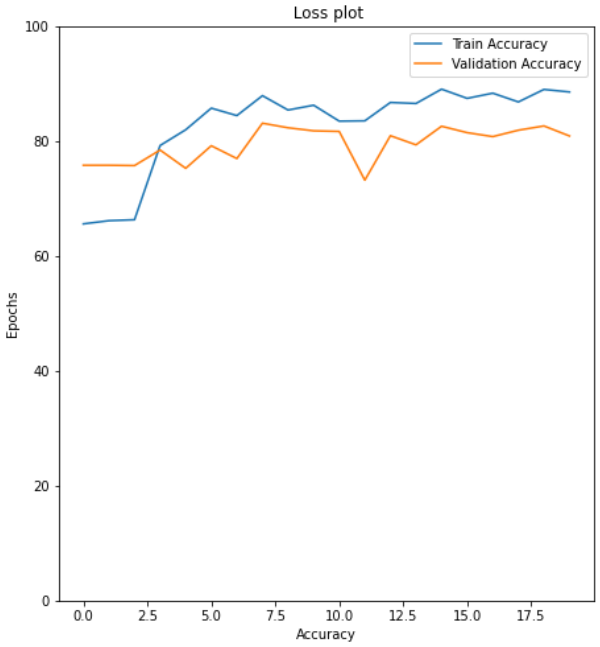
	Predicted Not normal	Predicted normal
Actual Not normal	198	30
Actual normal	47	353

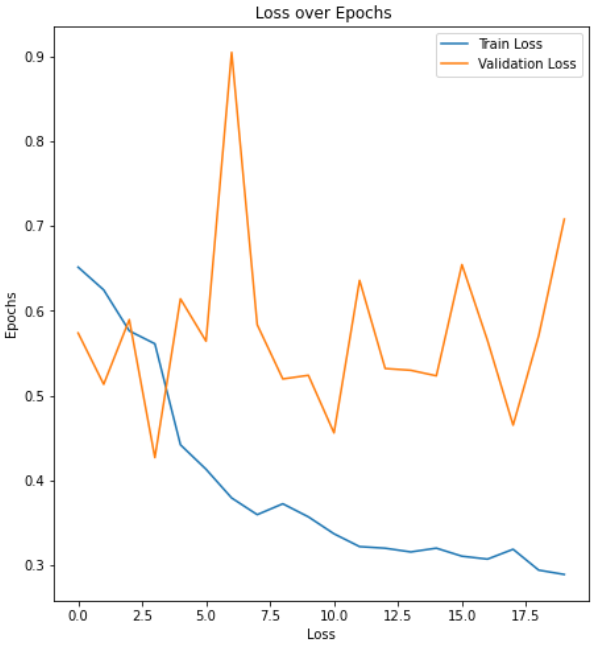
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	381	47
Actual pneumonia	21	179

F1 Score is : 0.740

Accuracy : 79.56

The loss and accuracy curves are as follows





Experiment 3

The following are the results for the ResNet 18with BCEloss

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3829	171
Actual covid-19	318	1882

	Predicted Not normal	Predicted normal
Actual Not normal	1885	315
Actual normal	176	3824

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3952	248
Actual pneumonia	293	1707

F1 Score is : 0.906

Accuracy : 91.82

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	396	204
Actual covid-19	10	18

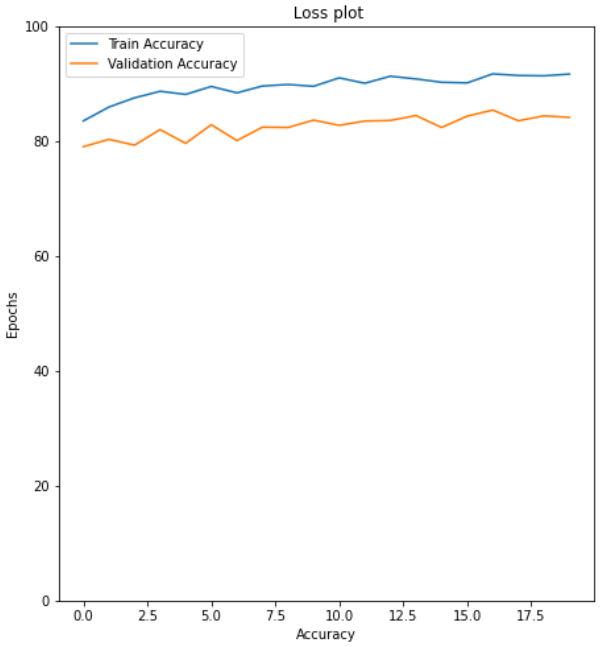
	Predicted Not normal	Predicted normal
Actual Not normal	206	22
Actual normal	19	381

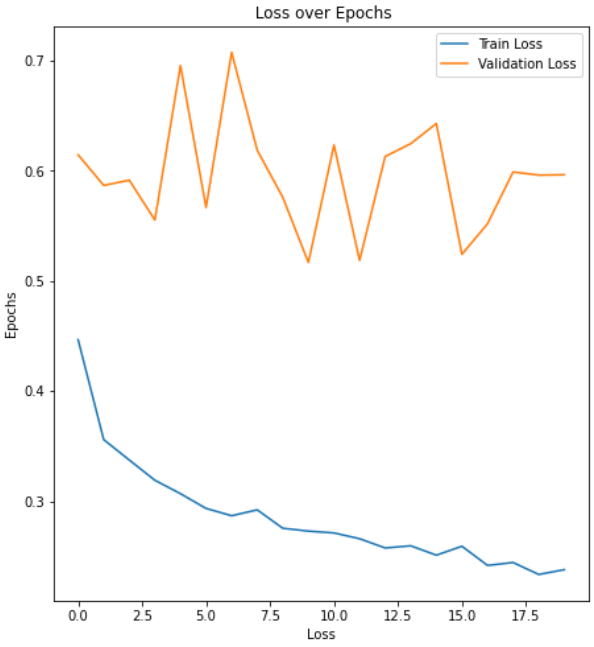
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	405	23
Actual pneumonia	15	185

F1 Score is : 0.799

Accuracy : 84.44

Following are the loss and accuracy curves





Experiment 4

This is the Focal loss with ResNet18

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3538	462
Actual covid-19	404	1796

	Predicted Not normal	Predicted normal
Actual Not normal	1820	380
Actual normal	462	3538

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4051	149
Actual pneumonia	898	1102

F1 Score is : 0.823

Accuracy : 85.18

Validation

Confusion Matrix

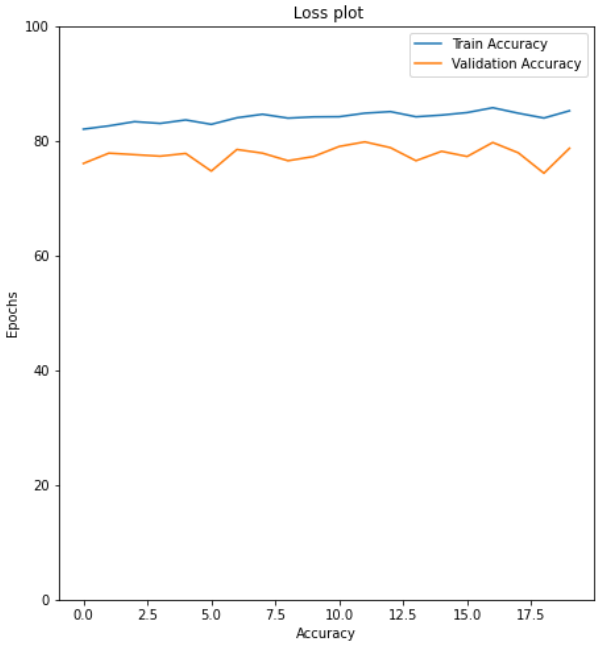
	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	387	213
Actual covid-19	14	14

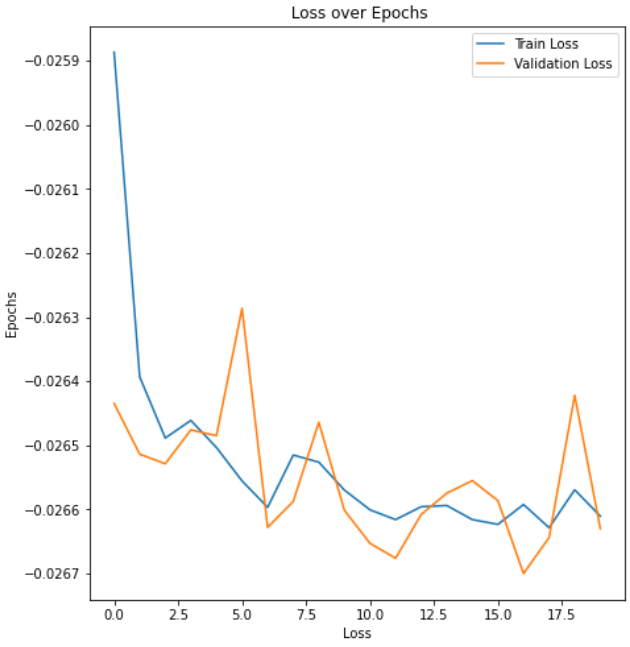
	Predicted Not normal	Predicted normal
Actual Not normal	180	48
Actual normal	45	355

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	413	15
Actual pneumonia	86	114

F1 Score is : 0.696

Accuracy : 77.65





Experiment 5

This is the testing with the low learning rate

VGG16 Focal Loss with learning rate 1e-4

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3172	828
Actual covid-19	786	1414

	Predicted Not normal	Predicted normal
Actual Not normal	1492	708
Actual normal	781	3219

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3666	534
Actual pneumonia	968	1032

F1 Score is : 0.711

Accuracy : 75.24

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	407	193
Actual covid-19	12	16

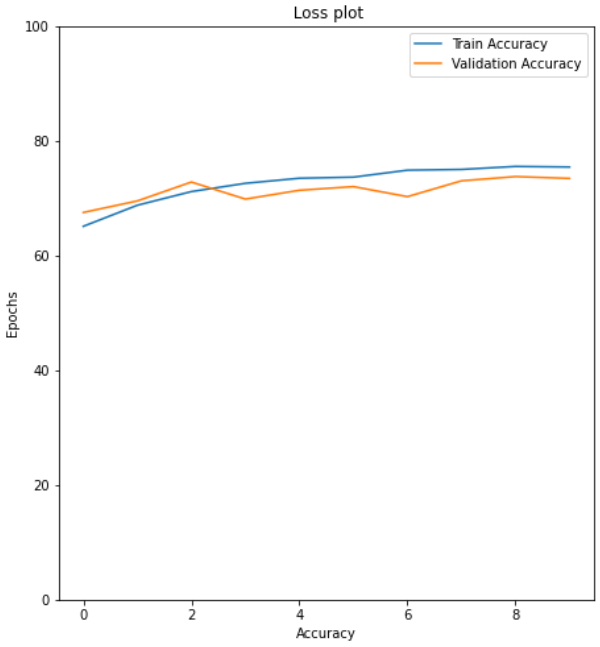
	Predicted Not normal	Predicted normal
Actual Not normal	160	68
Actual normal	76	324

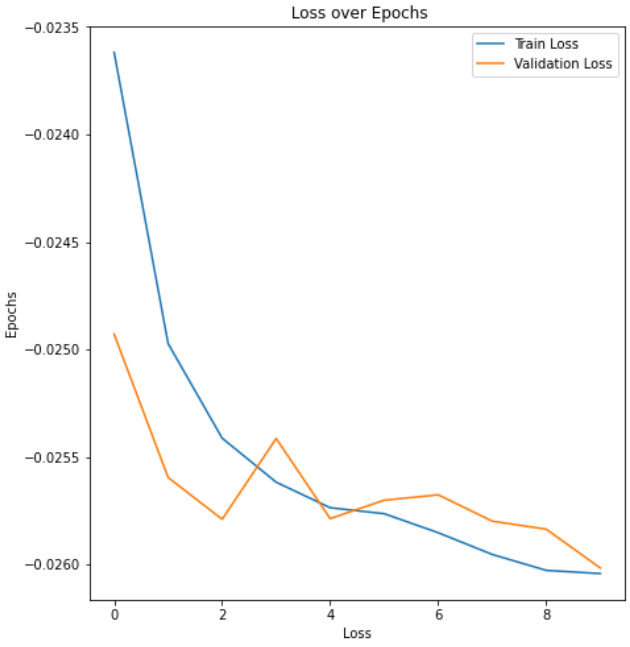
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	380	48
Actual pneumonia	98	102

F1 Score is : 0.641

Accuracy : 73.72

Following are loss and accuracy curves





Experiment 6

This is the experiment with the BCELoss with learning rate 0.00001

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3851	149
Actual covid-19	343	1857

	Predicted Not normal	Predicted normal
Actual Not normal	1855	345
Actual normal	148	3852

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3997	203
Actual pneumonia	347	1653

F1 Score is : 0.905

Accuracy : 91.74

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	407	193
Actual covid-19	11	17

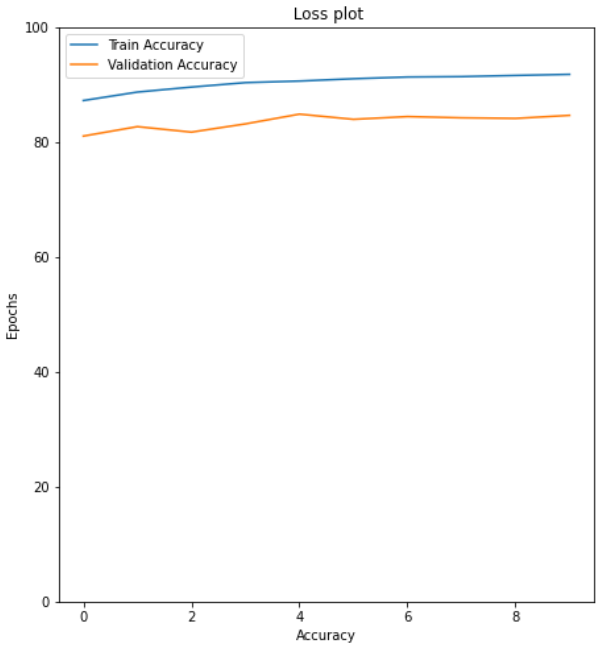
	Predicted Not normal	Predicted normal
Actual Not normal	189	39
Actual normal	21	379

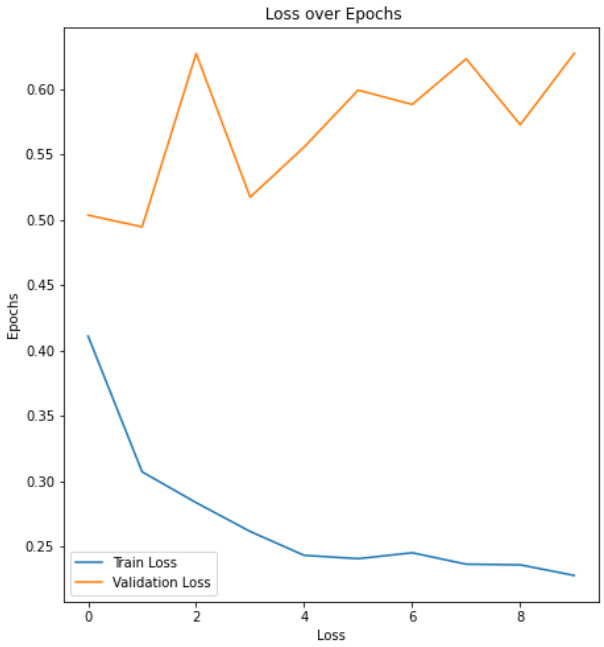
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	413	15
Actual pneumonia	29	171

F1 Score is : 0.786

Accuracy : 83.65

Following are the curves





Experiment 7

This is ResNet trained with the BCEloss learning rate 1e-4

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3689	311
Actual covid-19	432	1768

	Predicted Not normal	Predicted normal
Actual Not normal	1765	435
Actual normal	300	3700

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3894	306
Actual pneumonia	467	1533

F1 Score is : 0.861

Accuracy : 87.89

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	405	195
Actual covid-19	14	14

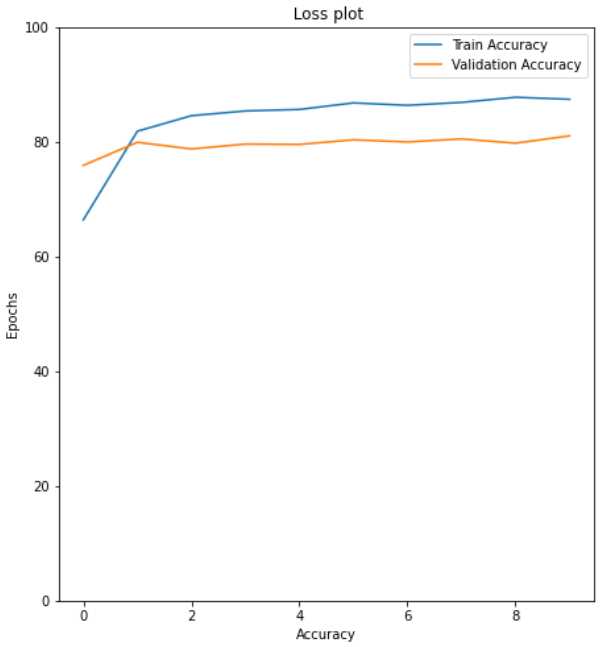
	Predicted Not normal	Predicted normal
Actual Not normal	182	46
Actual normal	27	373

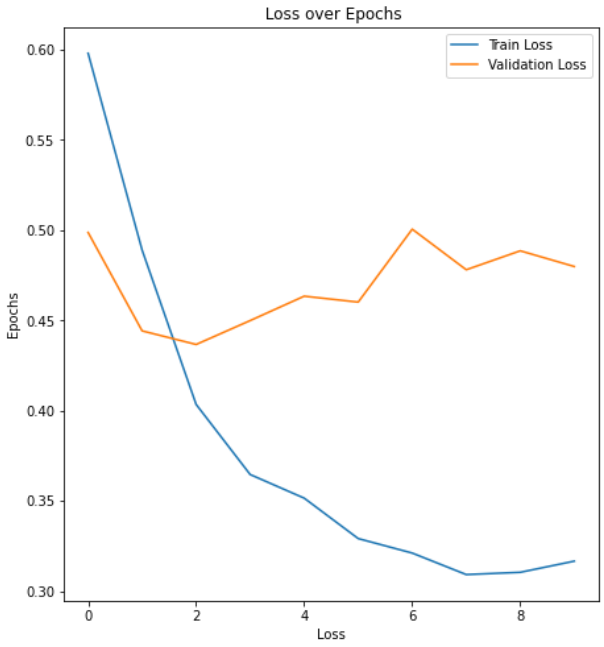
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	398	30
Actual pneumonia	41	159

F1 Score is : 0.755

Accuracy : 81.26

The loss and accuracy curves are as follows





Experiment 8

This experiment is with ResNet 18 focal loss learning rate 1e-4

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3853	147
Actual covid-19	1694	506

	Predicted Not normal	Predicted normal
Actual Not normal	483	1717
Actual normal	114	3886

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4197	3
Actual pneumonia	1960	40

F1 Score is : 0.611

Accuracy : 69.70

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	540	60
Actual covid-19	24	4

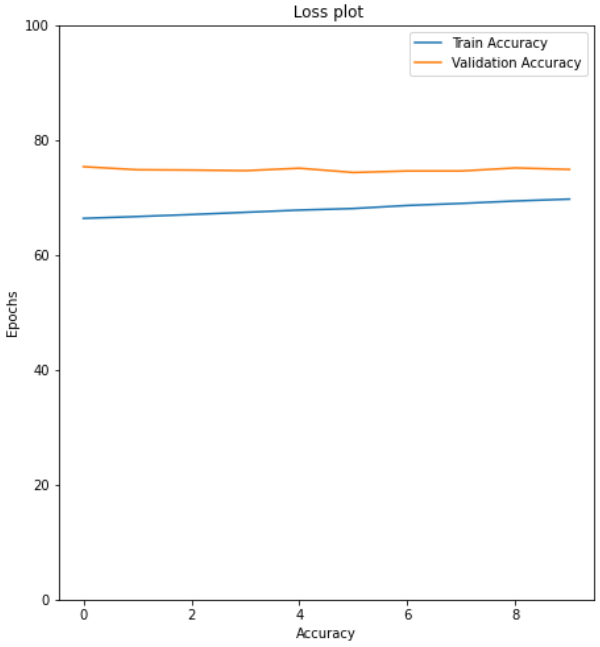
	Predicted Not normal	Predicted normal
Actual Not normal	37	191
Actual normal	12	388

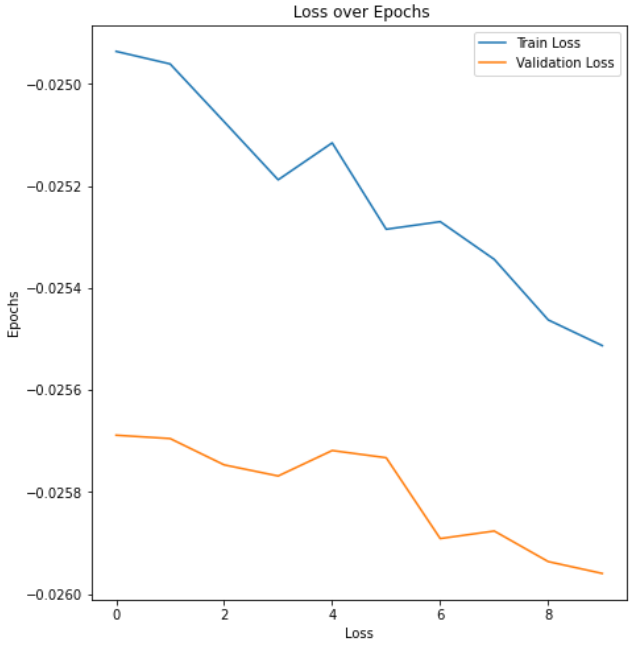
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	428	0
Actual pneumonia	193	7

F1 Score is : 0.624

Accuracy : 74.52

Loss and accuracy curves are as follows





Experiment

Some of the next experiments are on the Adam optimizers. On of the result of the experiment is as follows

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3579	421
Actual covid-19	716	1484

	Predicted Not normal	Predicted normal
Actual Not normal	1485	715
Actual normal	423	3577

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3774	426
Actual pneumonia	678	1322

F1 Score is : 0.790

Accuracy : 81.83

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	425	175
Actual covid-19	15	13

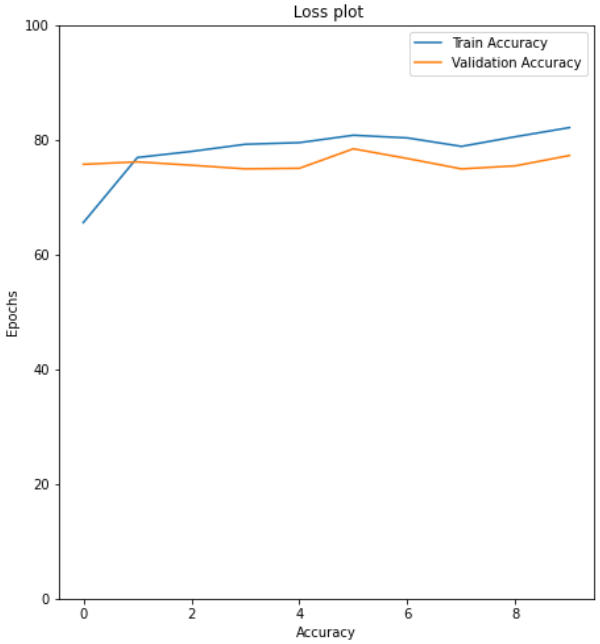
	Predicted Not normal	Predicted normal
Actual Not normal	152	76
Actual normal	37	363

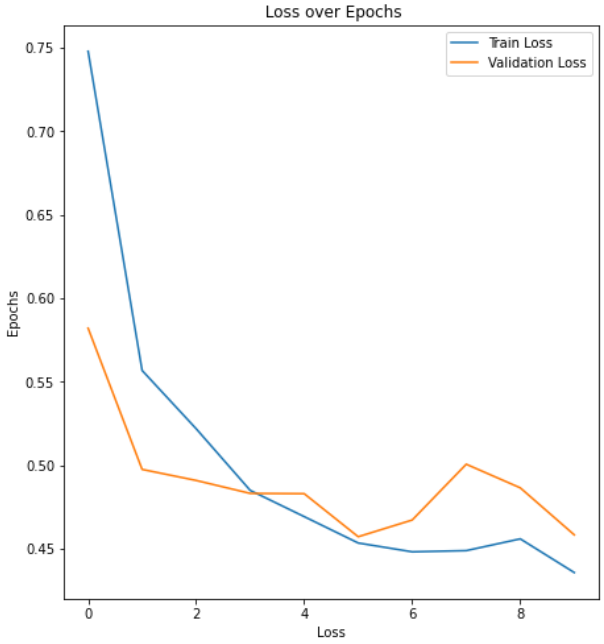
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	382	46
Actual pneumonia	67	133

F1 Score is : 0.709

Accuracy : 77.91

This is with the ResNet Architecture. Following are the loss and accuracy curves





Experiment 13

This is the experiment with the gamma value of focal loss . The experiment is performed on the VGG16

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3562	438
Actual covid-19	507	1693

	Predicted Not normal	Predicted normal
Actual Not normal	1728	472
Actual normal	456	3544

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3834	366
Actual pneumonia	530	1470

F1 Score is : 0.828

Accuracy : 85.11

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	390	210
Actual covid-19	12	16

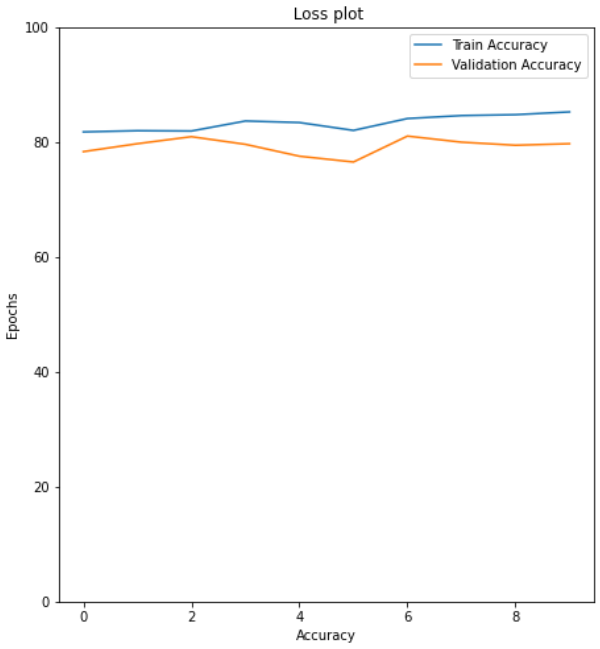
	Predicted Not normal	Predicted normal
Actual Not normal	172	56
Actual normal	45	355

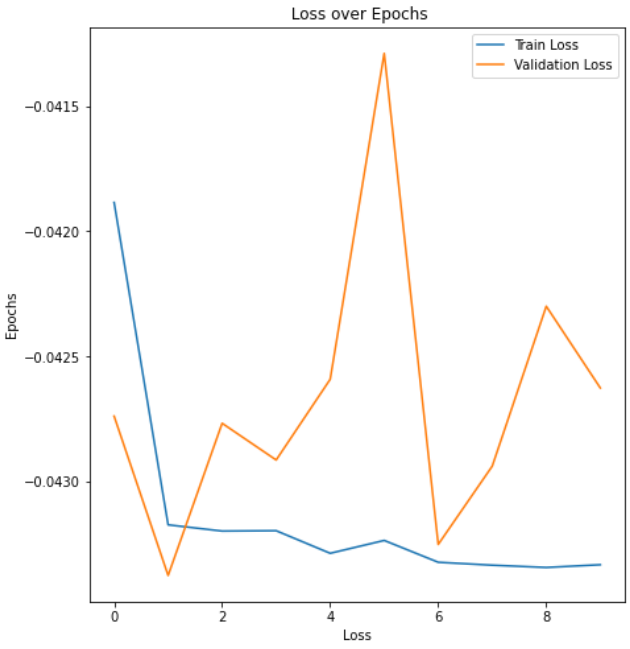
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	398	30
Actual pneumonia	51	149

F1 Score is : 0.720

Accuracy : 78.55

Loss and accuracy curves are as follows





Experiment 14

This is applied on the ResNet18with gamma 1.5

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3563	437
Actual covid-19	485	1715

	Predicted Not normal	Predicted normal
Actual Not normal	1721	479
Actual normal	472	3528

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3879	321
Actual pneumonia	589	1411

F1 Score is : 0.827

Accuracy : 85.03

```
100%|██████████| 10/10 [00:06<00:00,  1.50it/s]
```

Validation

Confusion Matrix

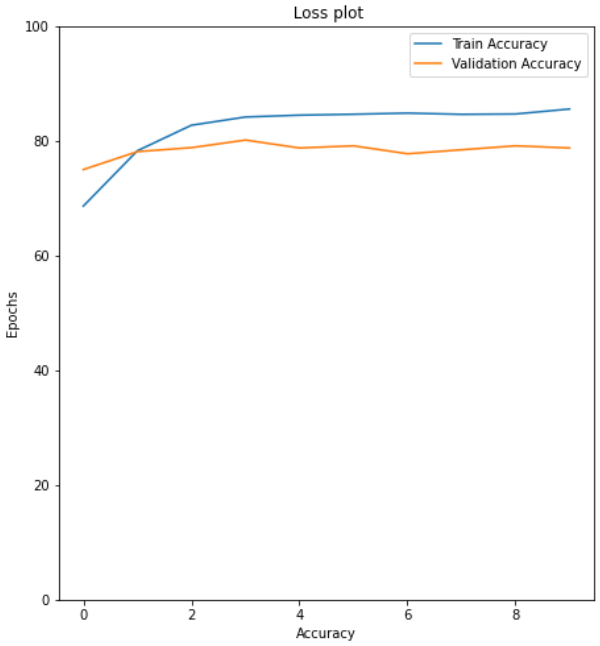
	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	391	209
Actual covid-19	14	14

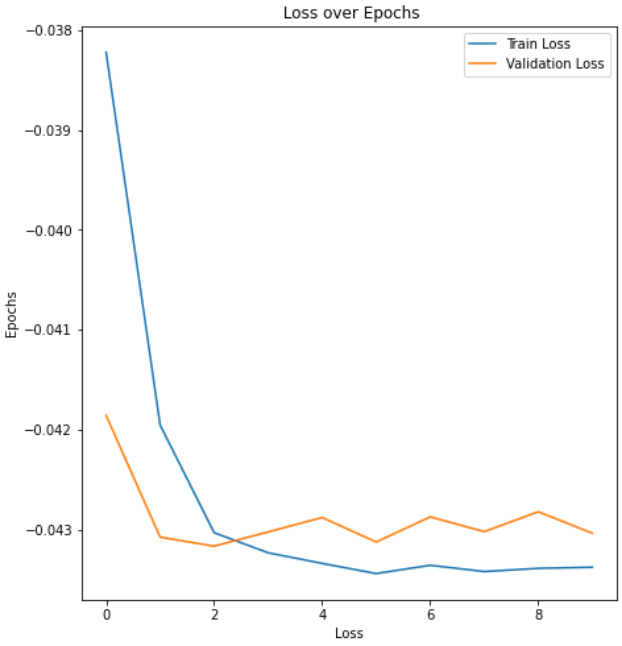
	Predicted Not normal	Predicted normal
Actual Not normal	179	49
Actual normal	51	349

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	398	30
Actual pneumonia	52	148

F1 Score is : 0.716

Accuracy : 78.50





Experiment 15

This is another fine tuned experiment results on VGG16

Fine tuned layers are `[0,2,5,10,24,28]`

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3832	168
Actual covid-19	258	1942

	Predicted Not normal	Predicted normal
Actual Not normal	1937	263
Actual normal	164	3836

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	3941	259
Actual pneumonia	236	1764

F1 Score is : 0.917

Accuracy : 92.75

Validation

Confusion Matrix

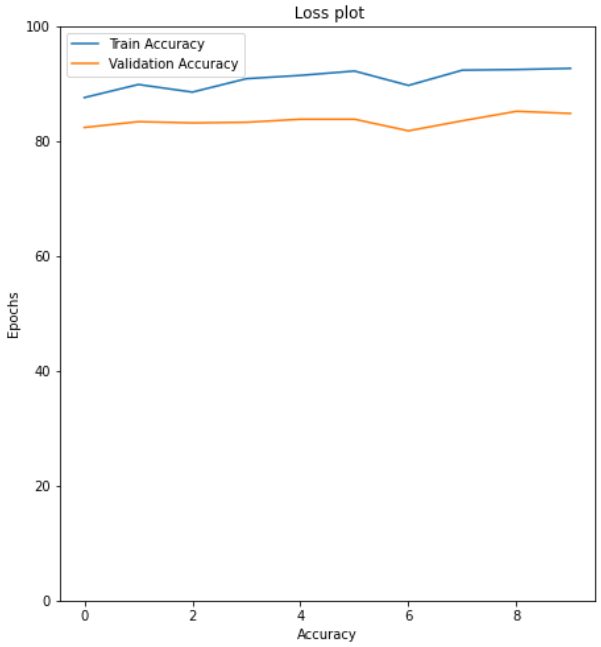
	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	405	195
Actual covid-19	8	20

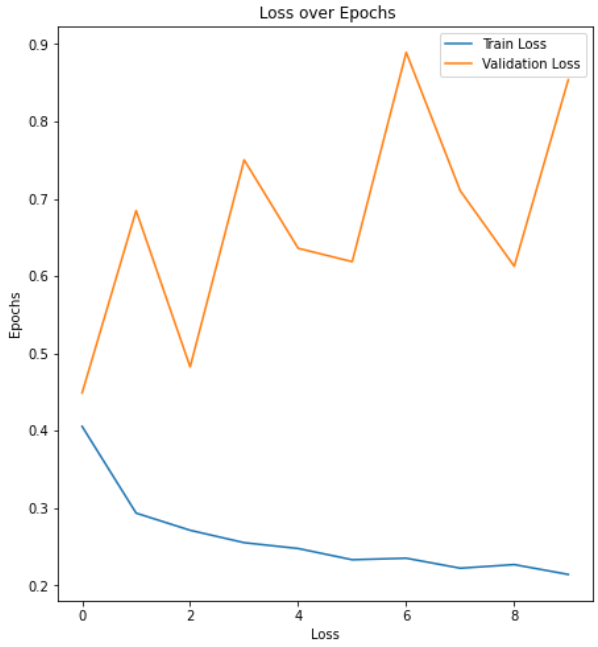
	Predicted Not normal	Predicted normal
Actual Not normal	200	28
Actual normal	13	387

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	409	19
Actual pneumonia	20	180

F1 Score is : 0.805

Accuracy : 84.97





Experiment 16

Following are the results obtained from the ResNet datasets

No Fine-tuned layers except the last two

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3844	156
Actual covid-19	1038	1162

	Predicted Not normal	Predicted normal
Actual Not normal	1133	1067
Actual normal	137	3863

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4149	51
Actual pneumonia	1372	628

F1 Score is : 0.747

Accuracy : 79.45

Validation

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	464	136
Actual covid-19	26	2

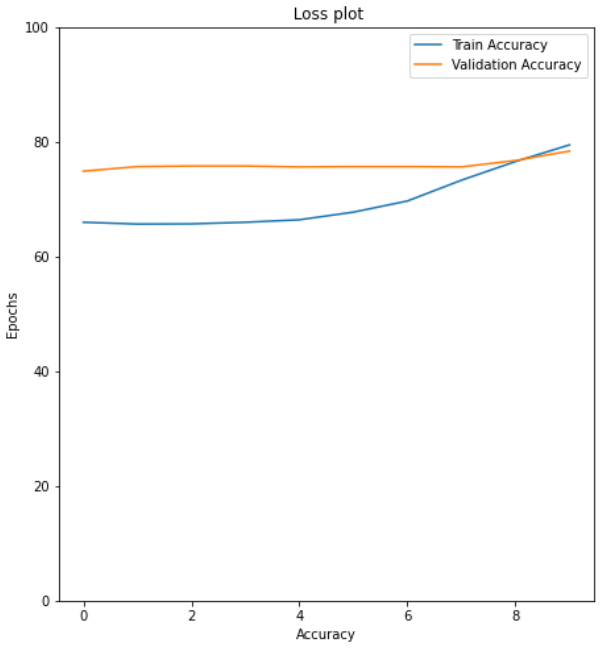
	Predicted Not normal	Predicted normal
Actual Not normal	115	113
Actual normal	7	393

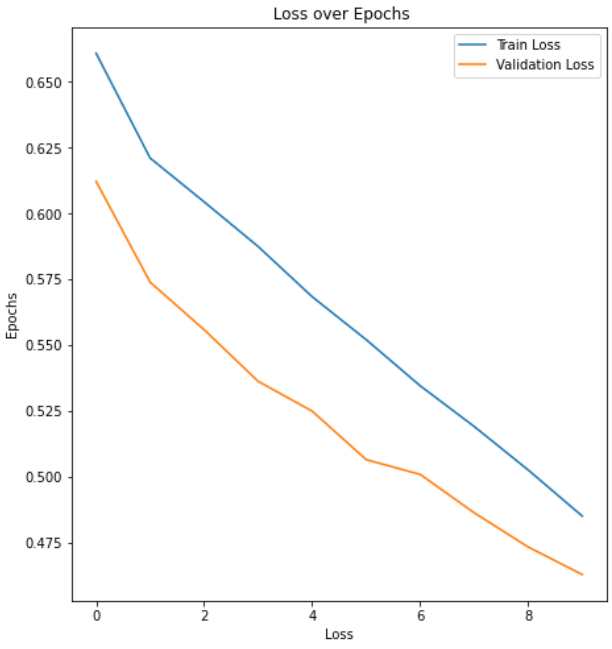
	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	426	2
Actual pneumonia	132	68

F1 Score is : 0.690

Accuracy : 77.91

Accuracy and loss curves are as follows





Experiment 17

This is the fined tuned experiment for the VGG16 architecture. Results are as follows.

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3844	156
Actual covid-19	1038	1162

	Predicted Not normal	Predicted normal
Actual Not normal	1133	1067
Actual normal	137	3863

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4149	51
Actual pneumonia	1372	628

F1 Score is : 0.747

Accuracy : 79.45

Validation

Confusion Matrix

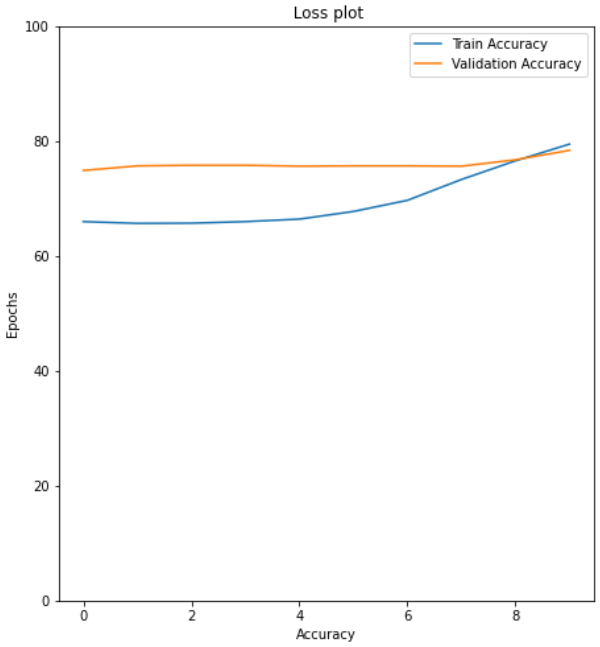
	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	464	136
Actual covid-19	26	2

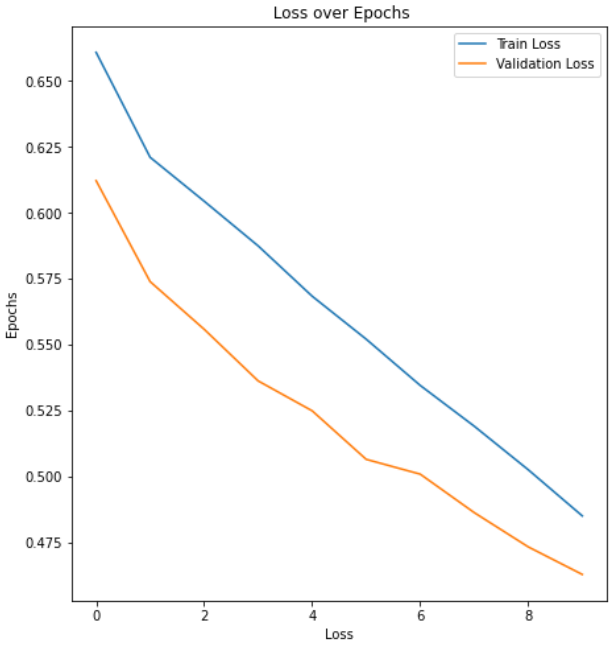
	Predicted Not normal	Predicted normal
Actual Not normal	115	113
Actual normal	7	393

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	426	2
Actual pneumonia	132	68

F1 Score is : 0.690

Accuracy : 77.91





Experiment 18

This is the fine tuned architecture for the ResNet18. Results are as follows.

Training

Confusion Matrix

	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	3370	630
Actual covid-19	374	1826

	Predicted Not normal	Predicted normal
Actual Not normal	1816	384
Actual normal	615	3385

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	4041	159
Actual pneumonia	867	1133

F1 Score is : 0.807

Accuracy : 83.71

Validation

Confusion Matrix

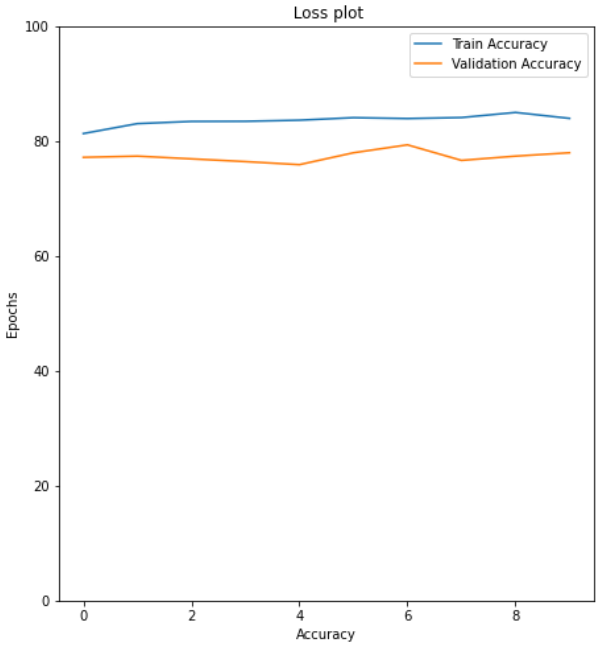
	Predicted Not covid-19	Predicted covid-19
Actual Not covid-19	367	233
Actual covid-19	15	13

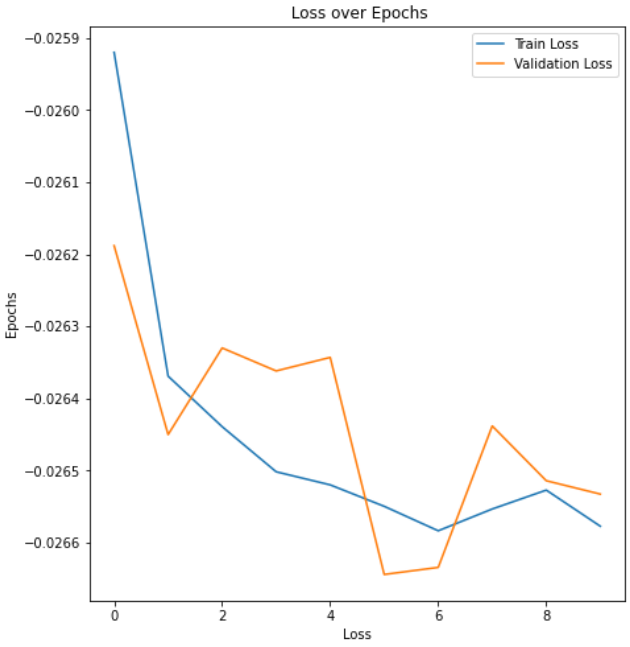
	Predicted Not normal	Predicted normal
Actual Not normal	187	41
Actual normal	56	344

	Predicted Not pneumonia	Predicted pneumonia
Actual Not pneumonia	411	17
Actual pneumonia	76	124

F1 Score is : 0.687

Accuracy : 76.75





Comparison and Analysis

Exp. No	Training		Validation		Details About the Experiment	Analysis
	Accuracy %	F1-Score	Accuracy %	F1-Score		
Block 1						
1.	88.27	0.859	81.63	0.747	VGG FLoss	From all the Architecture ResNet with Binary Cross Entropy gives the best results, apart from comparing with the Focal loss. But there's another thing that the loss curves and loss values are very small as compared to the BCELoss, and the curve is continuous going downwards. But the best results are obtained from the BCE ResNET.
2.	88.82	0.875	79.56	0.740	ResNet BCE	
3.	91.82	0.906	84.44	0.799	VGG BCE	
4.	85.18	0.823	77.65	0.696	ResNet FLoss	
Block 2						
5.	75.24	0.711	73.72	0.641	VGG16 Floss lr1e-4	In this part of the experiment it is evident that decreasing the learning rate has decreased the pace of learning of the model. Some of the results are reaching towards that of the other part, but there are some result which are resulting towards more better approach than the higher learning rate. Therefore, low learning rate although slows down the learning process. But it also gives better results.
6.	91.74	0.905	83.65	0.786	Res18 BCE lr1e-4	
7.	87.89	0.861	81.26	0.755	VGG16 BCE lr1e-4	
8.	69.7	0.611	74.52	0.624	Res18 Floss lr1e-4	
Block 3						
9.	65.59	0.555	75.79	0.636	VGG Floss AdamOpt	Changing the Optimizer has very bad affect on the learning as can be seen the Adam Optimizer has decreased the learning pace. Although, Adam produce better results than SGD but it is much slower. During experimentation in limited time constraint, Adam is not preferred.
10.	65.53	0.555	75.63	0.634	ResNet BCE AdamOpt	
11.	81.83	0.790	77.91	0.709	ResNet BCE AdamOpt	
12.	77.48	0.739	73.88	0.656	ResNet Floss AdamOpt	
Block 4						
13.	85.11	0.828	78.55	0.720	VGG gamma 1.5	Compare this blocks results with that of block1. It can be illustrated that decreasing the gamma has not been benefitted. It results in the approximation towards Cross Entropy loss with alpha affects on it also.
14	85.03	0.827	78.50	0.716	ResNet gamma 1.5	
Block 4b						
15.	92.75	0.917	84.97	0.805	Res18 Fine tuning BCELoss	Fine tuning experimentation after experimentation with the hyper parameters has resulted in good results. It can be seen that best results are obtained from the same experiment approach as were ontained in the part 1 of the same experiment. The best results obtained are from the experiment 15.
16.	79.45	0.747	77.91	0.690	VGG16 Fine tuning BCE loss	
17.	89.94	0.883	84.12	0.788	VGG16 Fine tuning FLoss	
18	83.71	0.801	76.75	0.687	Res18 Fine tuning BCE Loss	

From Comparing the confusion matrices for the experiments and looking at the most significant class is covid-19. Where the most important classification is False Negative. Moreover, True Positive is also very important class. And the validation data set is more significant the training dataset. By looking at the experiments results. We can see that the best results are obtained from the experiment 15. Where the False negative is minimum i.e.,8 and true positive are highest i.e.,20.

Overall, experiment number 15 has the best results.

GitHub Repository Link:

https://github.com/Basir-mahmood/msds19043_COVID19_DLSpring2020

Weights link:

https://drive.google.com/open?id=1jJ3VmSKyff-24oVgjyaM88usNoLMlcvo

Dataset link:

https://drive.google.com/file/d/1eytbwaLQBv12psV8I-aMkIli9N3bf8nO/view?usp=sharing