CHAPTER-5

Results and Discussion

5.1 Introduction

This section provides a comprehensive presentation and analysis of experimental outcomes, data, or investigations. Its primary purpose is to elucidate the significance of the findings, establish links to prior research, and offer insights into the broader implications of the study. The generalization of the proposed methodology is checked by taking two distinct datasets differ in terms of no. of total face images, unequal distribution of both (real & fake) samples. The effectiveness of the proposed methodology is evaluated by considering different train-test split ratio, as the different CNN don't give equal efficiency on all split ratio.

5.2 Result Analysis for Dataset 1

5.2.1 Result Analysis based on training phase parameters for TL VGG16

The Table 5.1 illustrates the performance of two different models, VGG16_ADAM and VGG16_RMSPROP, across various split ratios and epochs. Notably, for all split ratios and both optimizers (ADAM and RMSPROP), the training accur16acy consistently reaches a high level of approximately 98%, while the validation accuracy closely follows suit. This uniformity in high accuracy levels across different settings suggests that there is no evident overfitting occurring in the models. The models exhibit strong generalization capabilities, demonstrating consistent and robust performance across different training and validation datasets, as indicated by the matching training and validation accuracies.

cross-entropy loss plays a crucial role in binary classification tasks by effectively measuring the dissimilarity between predicted probabilities and true labels, guiding the model's training process, and promoting accurate, interpretable, and confident predictions. Table 5.2 illustrates the training and validation loss over epoch 1, 5, 15 and 20. The values shows that the loss is not reducing noticeably for 70%-30% split compared to other splits. The percentage improvement in the loss at epoch 30 as compared to epoch 1 higher in 50%-50% split ratio for both the optimizer (i.e. Adam and RMSprop). When Adam Optimizer is used during fine-tuning, the TL VGG16 model exhibits 0.1113% (min) loss over epoch 30 for 50%-50% train-test split ratio, which is 0.9219 at epoch 1. Whereas, for RMSprop, the minimum loss achieved is 0.1220 for the same split ratio over epoch 30.

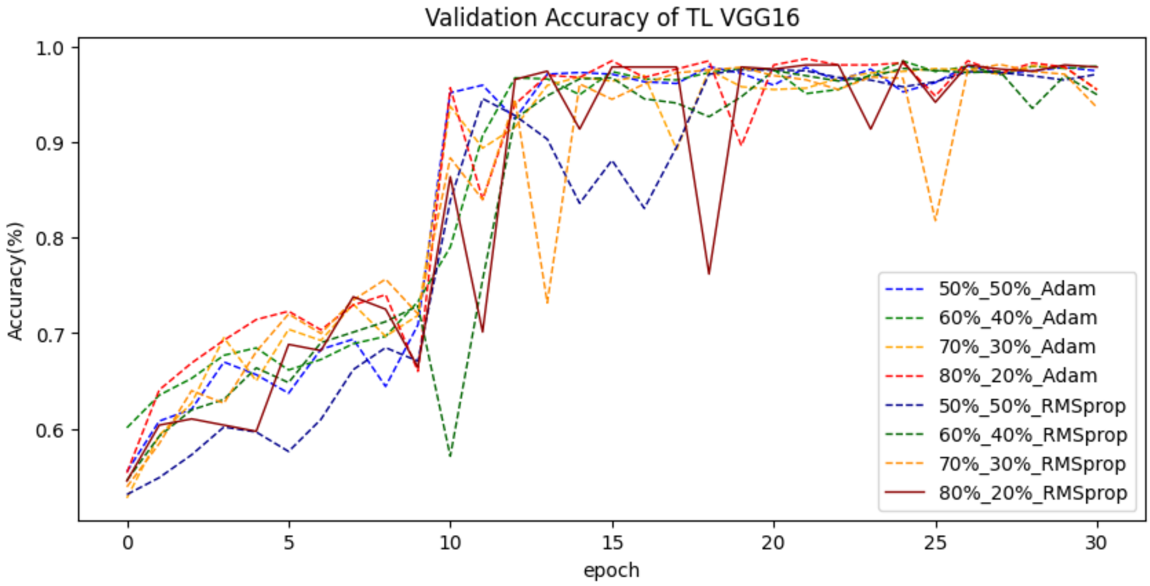
TABLE 5.1: Comparison of Accuracy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. VGG16_xxx: xxx represents the optimizer used during fine-tuning

		VGG16_Adam		VGG16_RMSProp	
SPLIT	**EPOCH**	**TRAINING ACC**	**VALIDATION ACC**	**TRAINING ACC**	**VALIDATION ACC**
50%50%	1	0.5205	0.5552	0.4857	0.5313
	10	0.6873	0.7087	0.6846	0.6708
	15	0.9669	0.9726	0.9263	0.8358
	30	0.9894	0.9744	0.9815	0.9709
60%-40%	1	0.5431	0.6011	0.5357	0.5462
	10	0.7030	0.7330	0.6828	0.7286
	15	0.9742	0.9495	0.9311	0.9659
	30	0.9934	0.9791	0.9834	0.9495
70%-30%	1	0.4965	0.5277	0.5208	0.5394
	10	0.6918	0.7187	0.6905	0.7201
	15	0.9701	0.9679	0.9339	0.9606
	30	0.9918	0.9548	0.9852	0.9359
80%-20%	1	0.4901	0.5541	0.5080	0.5455
	10	0.6929	0.6602	0.6970	0.6645
	15	0.9705	0.9675	0.9490	0.9134
	30	**0.9928**	0.9545	0.9832	0.9784

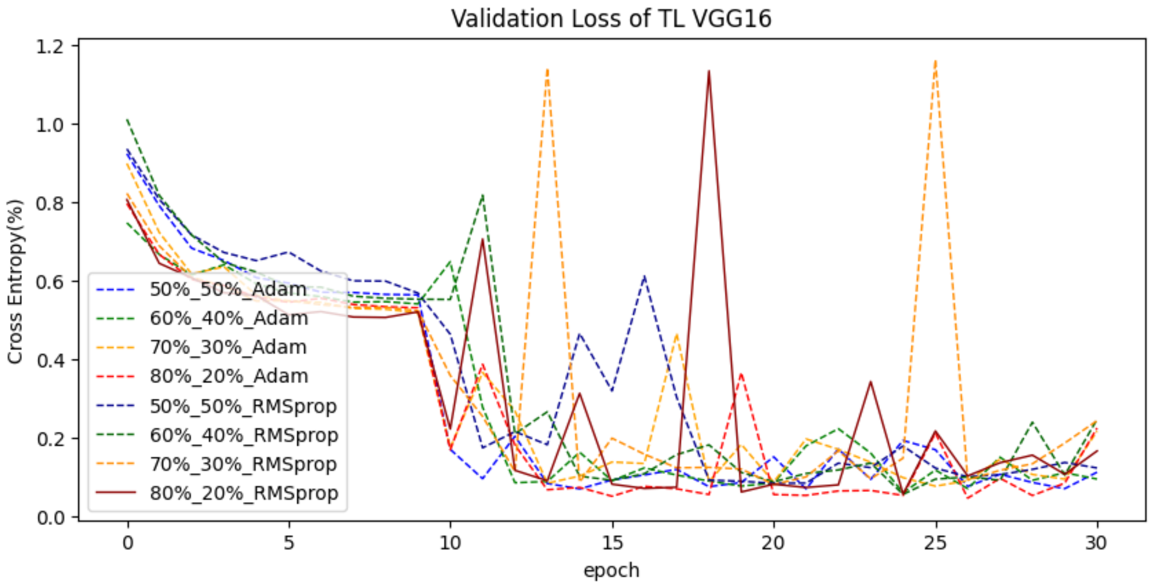
TABLE 5.2: Comparison of Cross Entropy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio

		VGG16_Adam		VGG16_RMSProp	
Split	**Epoch**	**Training loss**	**Validation loss**	**Training loss**	**Validation loss**
50%50%	1	1.2066	**0.9219**	1.3911	**0.9342**
	10	0.5649	0.5626	0.5718	0.5672
	15	0.0803	0.0688	0.1920	0.4646
	30	**0.0299**	**0.1113**	0.1104	**0.1220**
60%-40%	1	1.1090	0.7461	1.2487	1.0099
	10	0.5608	0.5396	0.5703	0.5517
	15	0.0623	0.1620	0.1736	0.1010
	30	0.0175	0.0938	0.0598	0.2457
70%-30%	1	1.1924	0.8972	1.0918	0.8211
	10	0.5550	0.5172	0.5590	0.5222

	15	0.0846	0.1014	0.1777	0.0859
	30	0.0241	0.2154	0.0663	0.2424
80%-20%	1	1.2455	0.7949	1.1238	0.8050
	10	0.5548	0.5297	0.5546	0.5194
	15	0.0786	0.0725	0.1370	0.3120
	30	0.0211	0.2238	0.0545	0.1653



(a)



(b)

FIGURE 5.1: Validation Accuracy(a) and Loss(b) comparison for proposed model VGG16 over all split ratios. Ex. 50_Adam represents the split ratio is 50%-50% and Adam optimizer used during fine tuning

The graphical comparison of training and validation accuracies of the proposed model for Adam and RMSprop optimizer over all split ratio is represented in Fig. 5.1. Fig 5.1(a) shows that the proposed model for both the optimizer during fine tuning increasing the accuracy during training period over train dataset. the maximum training accuracy achieved is 99.28% for 80%-20% split ratio for Adam optimizer. And the validation accuracy is improving for 80%-20% split ratio to 97.84% for RMSprop optimizer. The validation accuracy for 50%-50% split ratio for Adam optimizer is shown by dark line in the fig. 5.1(b). the graph conclude that the model is performed equally with minimum difference for all split ratio. From fig 5.2 (a) and (b), the cross entropy loss is reduced between 0.1 to 0.2 for train and validation dataset. the performance of model over 50%-50% split ratio for adam optimizer gives reduction in cross entropy loss during training and evaluation phase compared to other split ratio.

5.2.2 Result Analysis based on Performance parameters

Table 5.3 shows the evaluation parameter comparison for different train-test split of ND-IIITD dataset and optimizer used during fine-tuning of the proposed TL VGG16 model. To be precise, the RMSprop optimizer gives almost 100% accuracy in classifying the real samples of face images. The Adam gives slightly lower performs as compared to RMSprop which is 99.83% (max). However, RMSprop is not performing well to classify the fake samples of face images as compared to Adam. The maximum accuracy is noted 96.34% and 99.83% to classify fake and real images respectively, for 50%-50% split ratio and Adam optimizer, which is shown by bold letters in table 5.3.

TABLE 5.3: Performance parameter comparison for VGG16 models over different split ratio

Split ratio (Optimizer@fine-tuning)	**Image**	**Precision**	**Recall**	**F1-score**	**Accuracy**
80%-20%(Adam)	Fake	1	0.7749	0.8732	0.887
	Real	0.8149	1	0.898	
80%-20%(RMSprop)	Fake	0.9951	0.8874	0.9382	0.9413
	Real	0.8976	0.9956	0.9441	
70%-30%(Adam)	Fake	0.9968	0.9125	0.9528	0.9547
	Real	0.9189	0.9971	0.9564	
70%-30%(RMSprop)	Fake	1	0.8542	0.9214	0.9269
	Real	0.8721	1	0.9317	
60%-40%(Adam)	Fake	0.9976	0.9134	0.9537	0.9555

	Real	0.9198	0.9978	0.9572	
60%-40%(RMSprop)	Fake	1	0.8961	0.9452	0.9479
	Real	0.9055	1	0.9504	
50%-50%(Adam)	**Fake**	**0.9982**	**0.9634**	**0.9805**	**0.9808**
	Real	**0.9645**	**0.9983**	**0.9811**	
50%-50%(RMSprop)	Fake	1	0.9564	0.9777	0.9782
	Real	0.9581	1	0.9786	

5.2.3 Result Analysis based on training phase parameters for TL ResNet50 Model

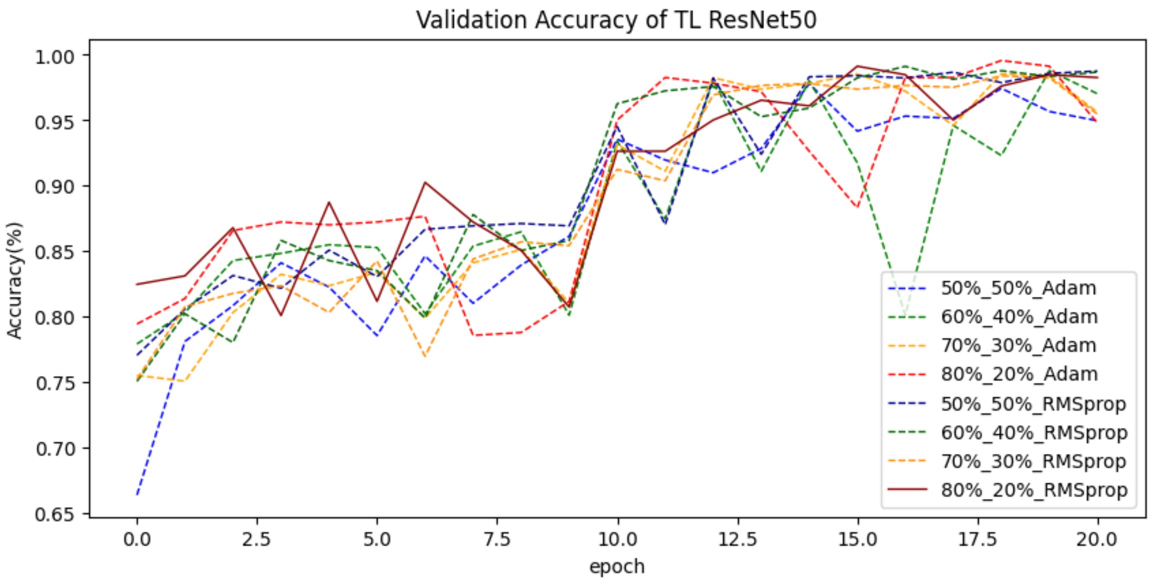
The performance of two distinct models, ResNet50_ADAM and ResNet50_RMSPROP, at various split ratios and epochs is shown in table 5.4. Notably, the training accuracy consistently reaches a high level of about 98% for all split ratios and both optimizers (ADAM and RMSPROP), while the validation accuracy nearly resembles this value. This consistency in high accuracy levels across many split ratios shows that the models do not appear to be overfitting. The matched training and validation accuracies show that the models have good generalisation capabilities, displaying stable and consistent performance across various training and validation datasets. The entropy loss is reduced over 20 epochs shows the model is learning the key features and not showing overfitting, as reflected in table 5.5. The accuracy and loss comparison of different split and both optimizers are shown in Fig. 5.3 and 5.4 for better understanding of performance of model during training and evaluation phase.

TABLE 5.4: Comparison of Accuracy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. ResNet50_xxx: xxx represents the optimizer used during fine-tuning

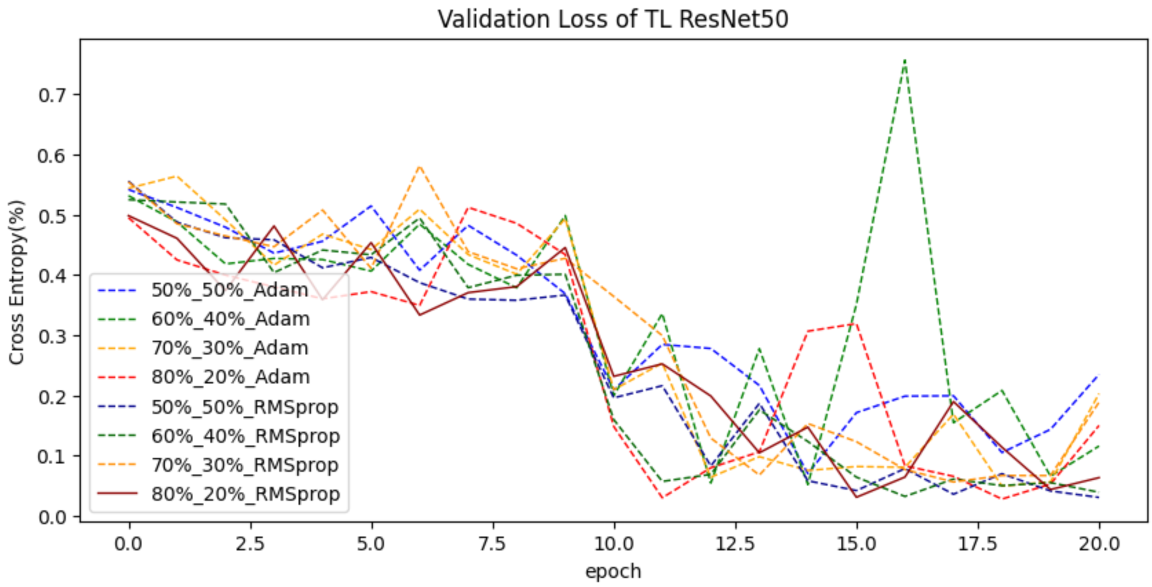
		ResNet50_Adam		**ResNet50_RMSprop**	
Split	Epoch	Training accuracy	Validation accuracy	Training accuracy	Validation accuracy
50%-50%	1	0.6220	0.6637	0.6277	0.7705
	10	0.8474	0.8614	0.8487	0.8694
	15	0.9797	0.9779	0.9599	0.9832
	20	0.9876	0.9497	0.9832	0.9876
60%-40%	1	0.6312	0.7791	0.6326	0.7505
	10	0.8475	0.8011	0.8397	0.8582
	15	0.9775	0.9802	0.9632	0.9593
	20	0.9941	0.9703	0.9878	0.9868
70%-30%	1	0.6181	0.7551	0.6666	0.7522
	10	0.8564	0.8105	0.8495	0.8542
	15	0.9890	0.9781	0.9685	0.9781
	20	0.9969	0.9563	0.9893	0.9548
80%-20%	1	0.6427	0.7944	0.6598	0.8247
	10	0.8400	0.8117	0.8471	0.8074
	15	0.9887	0.9264	0.9776	0.9610
	20	0.9724	0.9481	0.9934	0.9827

TABLE 5.5: Comparison of loss of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. ResNet50_xxx: xxx represents the optimizer used during fine-tuning

		ResNet50_Adam		ResNet50_RMSprop	
Split	Epoch	Training loss	Validation loss	Training loss	Validation loss
50%-50%	1	0.6260	0.5416	0.6394	0.5553
	10	0.3253	0.3688	0.3358	0.3664
	15	0.0540	0.0701	0.0986	0.0582
	20	0.0359	0.2347	0.0642	0.0306
60%-40%	1	0.6237	0.5313	0.6172	0.5246
	10	0.3401	0.4983	0.3363	0.4009
	15	0.0650	0.0519	0.1023	0.1234
	20	0.0183	0.1159	0.0345	0.0392
70%-30%	1	0.6440	0.5436	0.5846	0.5543
	10	0.3243	0.4933	0.3298	0.4274
	15	0.0319	0.0754	0.0887	0.1532
	20	0.0087	0.2029	0.0428	0.1887
80%-20%	1	0.6213	0.4948	0.5930	0.4981
	10	0.3367	0.4336	0.3283	0.4452
	15	0.0324	0.3064	0.0824	0.1473
	20	0.0832	0.1500	0.0216	0.0634



(a)



(b)

FIGURE 5.2: Validation Accuracy(a) and Loss(b) comparison for proposed ResNet50 model over all split ratios

5.2.4 Result Analysis based on performance parameter

The table 5.6 shows the comparative analysis of the accuracy of the proposed model for both optimizers and train-test split ratio. Comparing the performance of the ResNet50 model with the RMSprop optimizer, we observe that for an 80%-20% split ratio, the accuracy achieved is 98.26%. In this configuration, the recall for fake images is 96.97%, and the recall for real images is 99.56%. However, when using a 50%-50% split ratio, the accuracy improves to 98.52%. In this scenario, the recall for fake images also increases to 97.39%, while the recall for real images remains high at 99.65%. it is evident that the precision parameter for fake images consistently outperforms that of real images. In comparison to other split ratios, the 50%-50% split ratio yields notably high F1 scores, with a remarkable 98.50 for fake images and an even higher 98.53 for real images.

TABLE 5.6: Comparison of loss of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. ResNet50_xxx: xxx represents the optimizer used during fine-tuning

Split Ratio	Optimizer	Images	Precision	Recall	F1-score	Accuracy
80%-20%	Adam	Fake	1	0.8485	0.918	0.9239
		Real	0.8674	1	0.929	
80%-20%	RMSprop	Fake	0.9956	0.9697	0.9825	**0.9826**
		Real	0.9702	0.9956	0.9828	
70%-30%	Adam	Fake	0.9967	0.8746	0.9317	0.9357
		Real	0.8877	0.9971	0.9392	

70%-30%	RMSprop	Fake	0.9968	0.898	0.9448	0.9474
		Real	0.9067	0.9971	0.9497	
60%-40%	Adam	Fake	0.9976	0.9026	0.9477	0.9501
		Real	0.9107	0.9978	0.9523	
60%-40%	RMSprop	Fake	0.9955	0.9481	0.9712	0.9718
		Real	0.9502	0.9957	0.9724	
50%-50%	Adam	Fake	0.9977	0.8624	0.9252	0.9302
		Real	0.8785	0.9983	0.9345	
50%-50%	RMSprop	Fake	0.9964	0.9739	0.9850	**0.9852**
		Real	0.9744	0.9965	0.9853	

5.2.5 Comparative Discussion of proposed TL VGG16 and TL ResNet50

Both the models are trained over the same hyper parameters, except no. of iteration or epoch. The TL ResNet50 model achieves training accuracy ~85% over epoch 10 i.e during initial training. Whereas, TL VGG16 model achieves ~70% maximum training accuracy over epoch 10. ResNet50 performs better over epoch 20 achieving maximum training accuracy with improvement in entropy loss. Both ResNet50 and VGG16 perform exceptionally well in terms of precision, indicating their ability to classify fake images accuratelyResNet50 consistently achieves very high precision scores across all percentage splits and optimizers, with scores mostly above 0.995. The choice of optimizer, whether Adam or RMSprop, does not significantly affect precision scores in this context. The specific percentage split can influence precision scores. Overall, the models maintain high precision across different data distributions. Across all percentage splits except 70%-30% and for both optimizers, TL ResNet50 generally exhibits higher precision scores for real images compared to TL VGG16, as per Fig 5.5. ResNet50's precision scores are often above 0.90, indicating strong positive class classification ability. VGG16 also demonstrates good precision, but its scores are generally lower than those of ResNet50. As depicted in fig 5.5(b), the real images are classifying with maximum 96.34% accuracy(recall) and 99.83% for classifying fake images for TL vgg16 over 50%-50% ratio when Adam optimizer is used. Whereas, ResNet50 model, when RMSprop optimizer is used, give nearly equal recall parameters for real and fake samples over 50%-50% and 80%-20% split ratio. As per fig 5.6, Higher accuracy scores of 98.52% are typically achieved when there is a more balanced distribution of fake and real images in the dataset (e.g., 50%-50% split) and nearly same (98.26%) for 80%-20% split ratio. The choice of optimizer influences accuracy scores, with Adam often achieving slightly higher accuracy scores for VGG16 compared to RMSprop. Whereas, ResNet50 outperforms with RMSprop optimizer compared to Adam, as seen from fig 5.7.

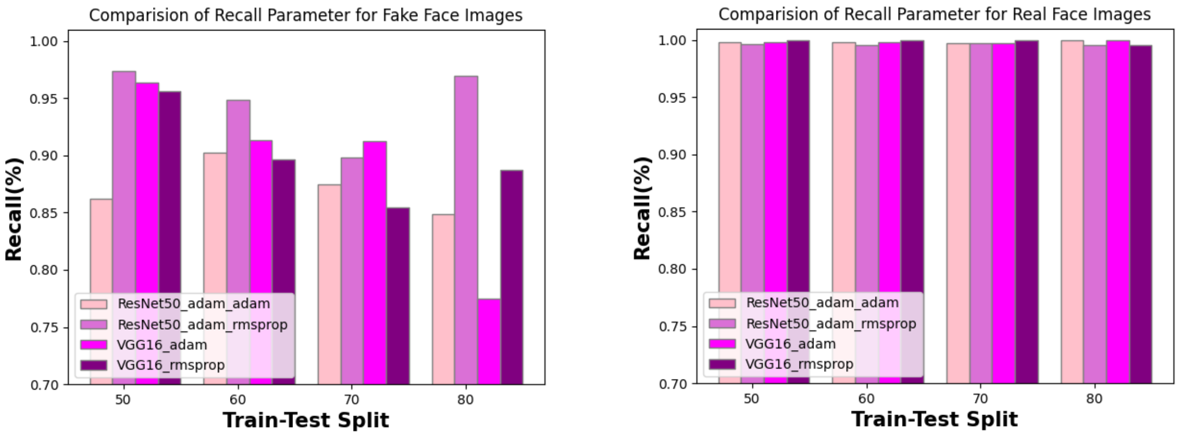


FIGURE 5.3: Comparison of Precision parameter for proposed model over all split ratio

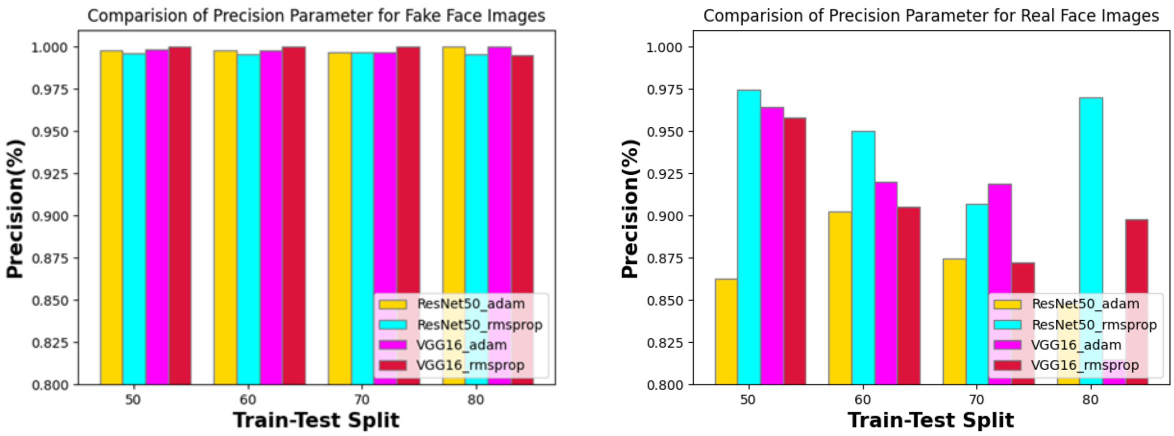


FIGURE 5.4: Comparison of Recall parameter for proposed model over all split ratio

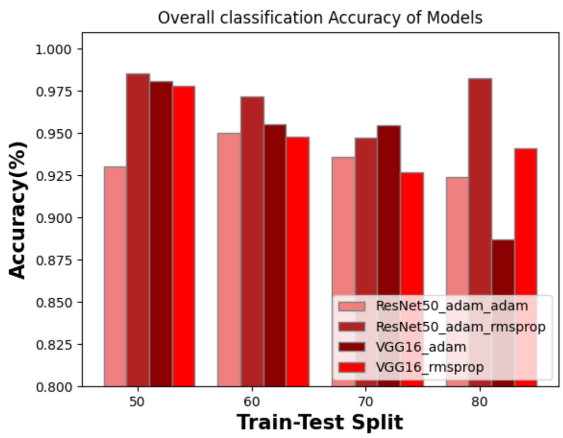


FIGURE 5.5: Overall Accuracy Comparison for Real & Fake samples

5.3 **Dataset -2**

5.3.1 Result analysis based on training phase parameters for TL VGG16

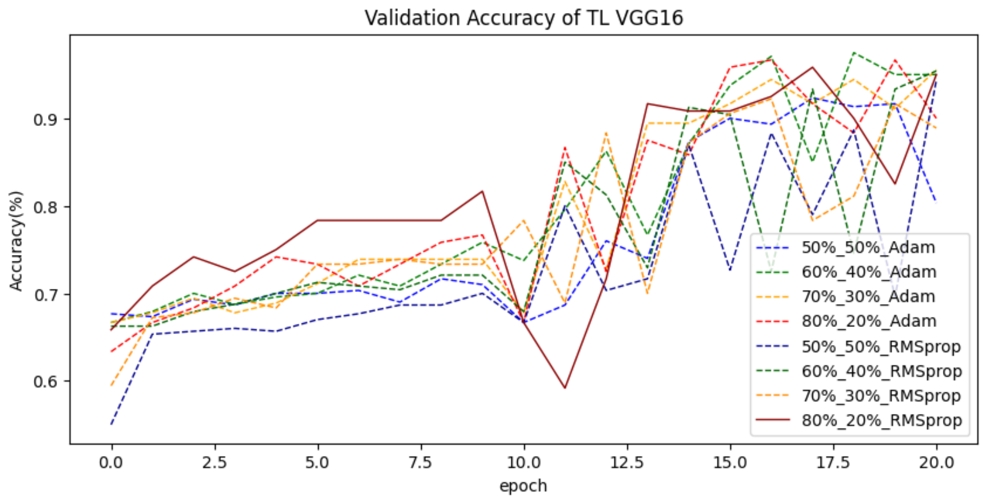
As per table 5.7 and 5.8, The VGG16 model with the Adam optimizer converges faster in terms of training accuracy and training loss but tends to overfit more quickly compared to the RMSprop optimizer. The RMSprop optimizer appears to exhibit slightly better generalization, as indicated by narrower gaps between training and validation accuracy and loss. The maximum training and validation accuracy achieved are 96.79% and 95.56% for the split ratio 70%-30% for Adam optimizer. The min loss achieved are 0.0913 and 0.1019 for the same configuration. 50%-50% split ratio showing underfitting when Adam optimizer is used for the proposed TL vgg16 model as compared to other split ratios and optimizers.

TABLE 5.7: Comparison of Accuracy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. VGG16_xxx: xxx represents the optimizer used during fine-tuning

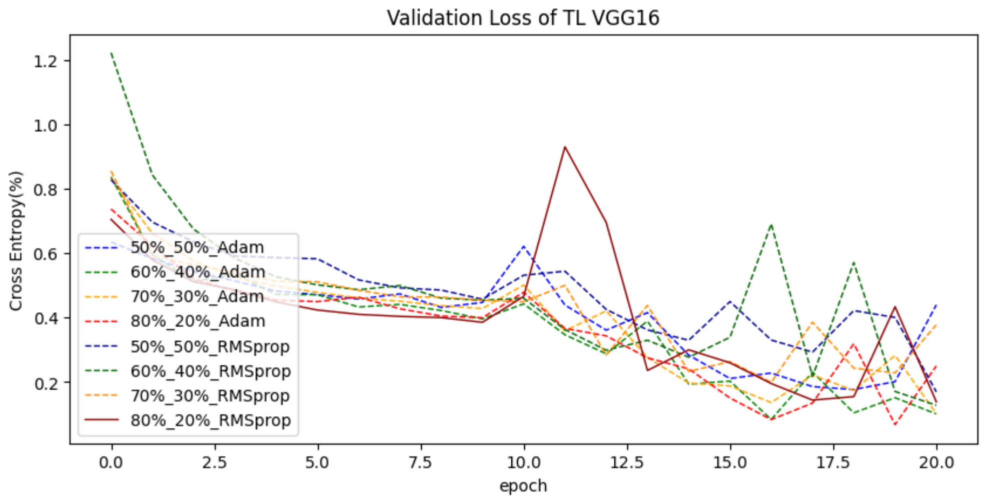
		VGG16_Adam		VGG16_RMSprop	
Split	**Epoch**	**Train_ACC**	**Validation_ACC**	**Train_ACC**	**Validation_ACC**
50%-50%	1	0.5967	0.6767	0.5550	0.5500
	5	0.7133	0.7000	0.7067	0.6567
	10	0.7583	0.7100	0.7600	0.7000
	15	0.8533	0.8733	0.8050	0.8700
	20	0.9483	0.8033	0.9000	0.9433
60%-40%	1	0.6236	0.6667	0.6000	0.6625
	5	0.7625	0.6958	0.7208	0.7000
	10	0.7903	0.7583	0.7764	0.7208
	15	0.8792	0.8708	0.8181	0.9125
	20	0.9542	0.9500	0.9000	0.9542
70%-30%	1	0.5488	0.6667	0.6345	0.5944
	5	0.7262	0.6889	0.7357	0.6833
	10	0.7821	0.7389	0.7845	0.7333
	15	0.8881	0.8944	0.8357	0.8722
	20	**0.9679**	**0.9556**	0.9321	0.8889
80%-20%	1	0.5437	0.6333	0.5885	0.6583
	5	0.7125	0.7417	0.7385	0.7500
	10	0.7656	0.7667	0.7917	0.8167
	15	0.8927	0.8583	0.8396	0.9083
	20	0.9667	0.9000	0.9240	0.9500

TABLE 5.8: Comparison of Accuracy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. VGG16_xxx: xxx represents the optimizer used during fine-tuning

		VGG16_Adam		VGG16_RMSprop	
Split	**Epoch**	**Train loss**	**Validation loss**	**Train loss**	**Validation loss**
50%-50%	1	1.0213	0.6348	1.3120	0.8279
	5	0.5960	0.4827	0.6267	0.5865
	10	0.4522	0.4473	0.4618	0.4573
	15	0.2838	0.2838	0.3294	0.3300
	20	**0.1445**	**0.4397**	**0.2014**	**0.1700**
60%-40%	1	0.9565	0.8369	1.1491	1.2222
	5	0.5252	0.4722	0.5688	0.5266
	10	0.4245	0.3963	0.4214	0.4541
	15	0.2266	0.1939	0.3328	0.2774
	20	**0.1153**	**0.1013**	**0.2309**	**0.1283**
70%-30%	1	1.1348	0.8315	1.0069	0.8553
	5	0.5401	0.4981	0.5537	0.5129
	10	0.4244	0.4284	0.4452	0.4541
	15	0.2266	0.1974	0.3312	0.2333
	20	**0.0913**	**0.1019**	**0.1688**	**0.3780**
80%-20%	1	1.2225	0.7371	1.0559	0.7050
	5	0.5744	0.4544	0.5267	0.4489
	10	0.4170	0.3999	0.3995	0.3857
	15	0.2237	0.2413	0.3081	0.3003
	20	**0.0700**	**0.2503**	**0.1920**	**0.1389**



(a)

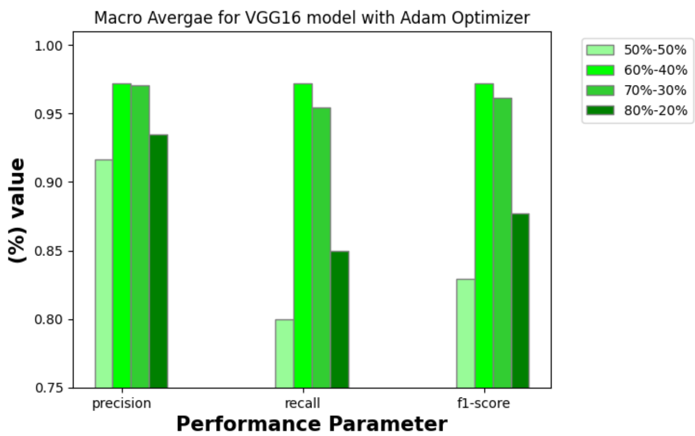


(b)

FIGURE 5.6: Comparison of Validation Accuracy(a) and Loss(b) of all split ratio for VGG16 Model

5.3.2 Result analysis based on performance parameters for TL VGG16 Model

The dataset 2 is imbalanced dataset, where the samples of fake images are higher than the real face images. Hence, the parameters are compared based on macro average. As per the comparison shown in Fig. 6, For the VGG16 model with the Adam optimizer performs better over RMSprop optimizer. Among these splits, the 60% split stands out with the



highest precision, recall, and F1-score of 0.9719, indicating excellent model performance on this specific data split.

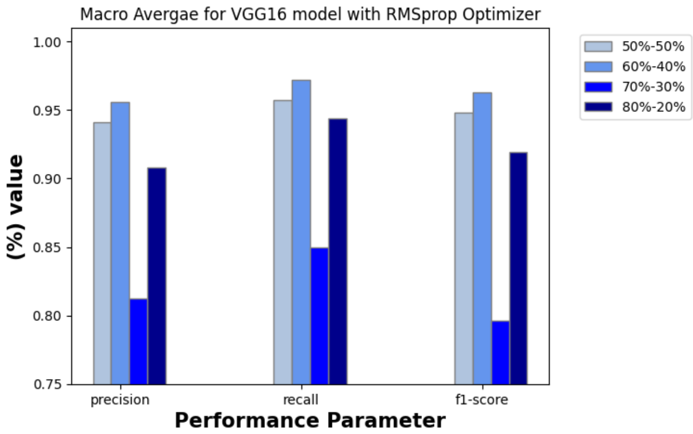


FIGURE 5.7: Comparison of Macro Average w.r.t Adam and RMSprop Optimizer for TL VGG16

5.3.3 Result analysis based on training phase parameters for TL ResNet50 Model

Across all data split ratios, the ResNet50 model consistently demonstrates strong performance in terms of both training and validation accuracy for both the Adam and RMSprop optimizers, as seen from table 5.9. The validation accuracy tends to closely track the training accuracy, indicating good generalization of the model. For the 60%-40% split, the model exhibits notably high validation accuracy values, surpassing 90% for both optimizers at various epochs. In the 80%-20% split, the ResNet50 model with the Adam optimizer achieves exceptionally high training accuracy, reaching 99.06%, although its validation accuracy is lower compared to other splits. The validation loss behavior varies for both optimizer and for different split ratios, as shown in Table 5.10. For some epochs and split ratios, the Adam optimizer outperforms RMSprop in terms of validation loss, while in others, RMSprop achieves lower validation loss. For split ratio 80%-20%, the initial training and validation loss were noted 0.5649 and 0.4640, which is reduced to 0.0599 and 0.1047 over epoch 20.

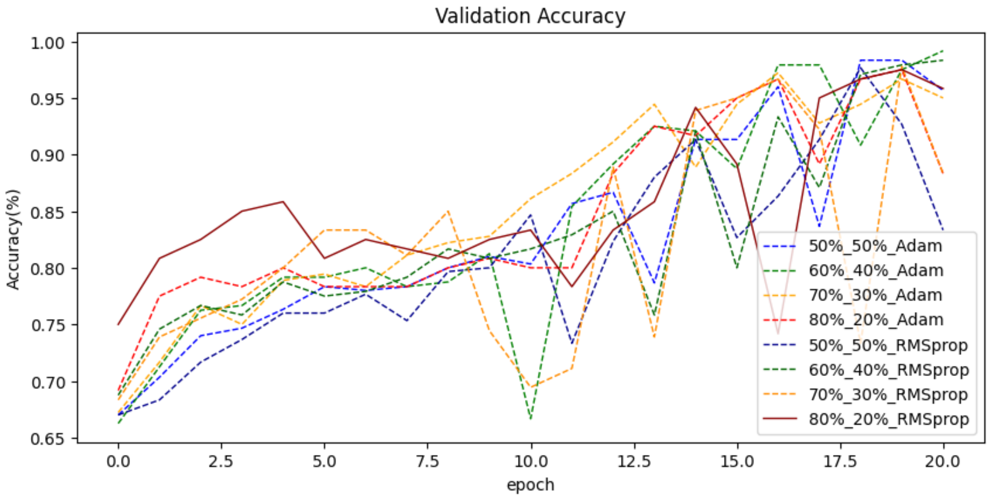
TABLE 5.9: Comparison of Accuracy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. ResNet50_xxx: xxx represents the optimizer used during fine-tuning

		ResNet50_Adam		ResNet50_RMSprop	
Split	**Epoch**	**Train_ACC**	**Validation_ACC**	**Train_ACC**	**Validation_ACC**
50%-50%	1	0.6933	0.6700	0.6867	0.6700

	5	0.8000	0.7633	0.8083	0.7600
	10	0.8517	0.8100	0.8617	0.8000
	15	0.9100	0.9133	0.8567	0.9133
	20	0.9783	0.9567	0.9383	0.8333
60%-40%	1	0.7028	0.6625	0.6764	0.6875
	5	0.8153	0.7917	0.8042	0.7875
	10	0.8514	0.8125	0.8611	0.8083
	15	0.9458	0.9208	0.8431	0.9208
	20	0.9889	0.9917	0.9556	0.9833
70%-30%	1	0.6631	0.6722	0.6988	0.6833
	5	0.8095	0.7889	0.8262	0.8000
	10	0.8405	0.8278	0.8702	0.7444
	15	0.9369	0.8889	0.8679	0.9389
	20	0.9512	0.9500	0.9500	0.8833
80%-20%	1	0.6458	0.6917	0.7073	0.7500
	5	0.8052	0.8000	0.8344	0.8583
	10	0.8667	0.8083	0.8771	0.8250
	15	0.9208	0.9167	0.8708	0.9417
	20	0.9906	0.8833	0.9760	0.9583

TABLE 5.10: Comparison of Cross Entropy of Training and validation Dataset for the proposed model w.r.t different optimizers and train-test split ratio during initial training of model Adam optimizer is used for all split ratio. ResNet50_xxx: xxx represents the optimizer used during fine-tuning

		ResNett50_Adam		**ResNet50_RMSprop**	
Split	Epoch	Train loss	Validation loss	Train loss	Validation loss
50%-50%	1	0.5913	0.5517	0.5999	0.5124
	5	0.3743	0.4131	0.3514	0.4082
	10	0.2934	0.3423	0.2918	0.3436
	15	0.2117	0.1934	0.2944	0.2120
	20	0.0600	0.0849	0.1892	0.3874
60%-40%	1	0.5813	0.5292	0.6172	0.5299
	5	0.3452	0.3953	0.3732	0.3895
	10	0.2947	0.3343	0.2781	0.3423
	15	0.1168	0.1271	0.3209	0.2911
	20	0.0257	0.0415	0.1467	0.0619
70%-30%	1	0.6614	0.5225	0.5477	0.4988
	5	0.3643	0.4081	0.3378	0.4008
	10	0.2969	0.3525	0.2746	0.3644
	15	0.1467	0.1795	0.2960	0.2415
	20	0.1354	0.1546	0.1457	0.1932
80%-20%	1	0.6646	0.5111	0.5649	0.4640
	5	0.3640	0.4516	0.3250	0.3577
	10	0.2730	0.3529	0.2710	0.3331
	15	0.1673	0.1657	0.2871	0.1388
	20	0.0197	0.4098	**0.0599**	**0.1047**



(a)

(b)

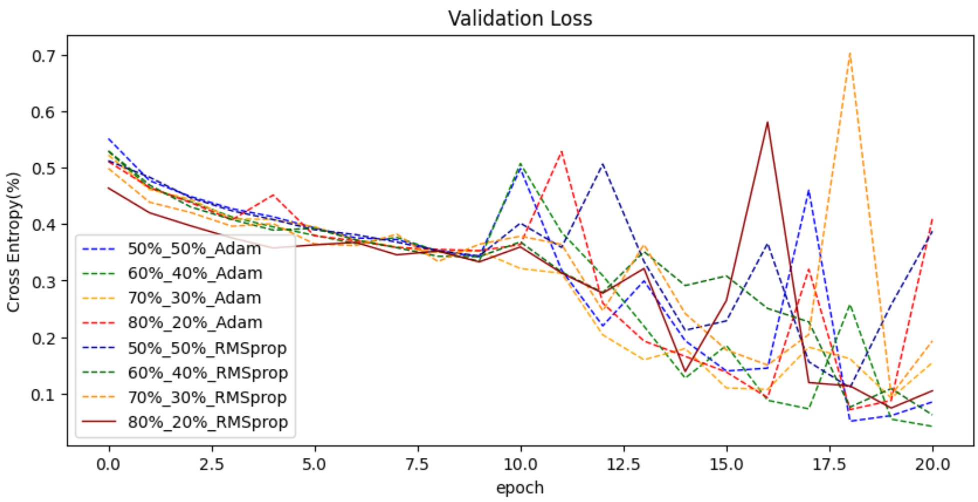


FIGURE 5.8: Comparison of Validation accuracy(a) and Validation Loss(b) of all split ratio for ResNet50 Model

5.3.4 Result analysis based on performance parameters for ResNet50 model

As shown in fig 5.9, macro average is calculated from the performance parameters of CM (confusion Matrix) for all four split ratios and both the optimizers used for TL ResNet50. Precision, recall and f1-score is achieved maximum 97.85%, 98.44%, and 98.14% for ResNet50 model with Adam optimizer for 60%-40% split over all train-test split ratio. Precision, recall and f1-score is achieved a maximum 98.78%, 99.38%, and 99.07% for ResNet50 model for 80%-20% split ratio with RMSprop optimizer. Hence, based on precise comparison of all the parameters ResNet50 models gives better performance when RMSprop optimizer is used during fine tuning. Moreover, the analysis reveals that the model gives maximum classification accuracy in terms of macro average for 80%-20% train-test split ratio.

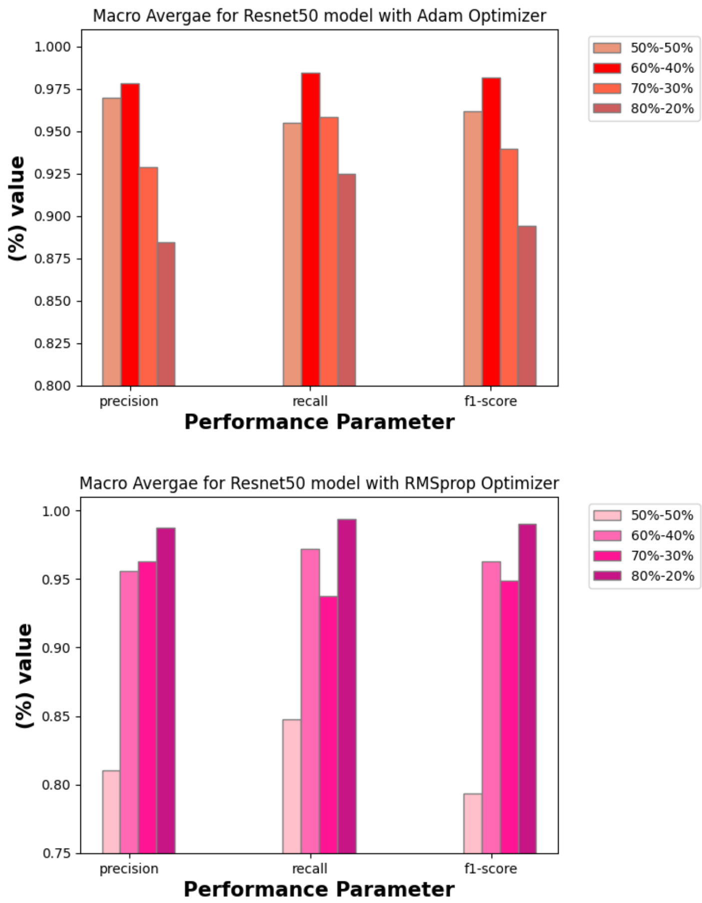
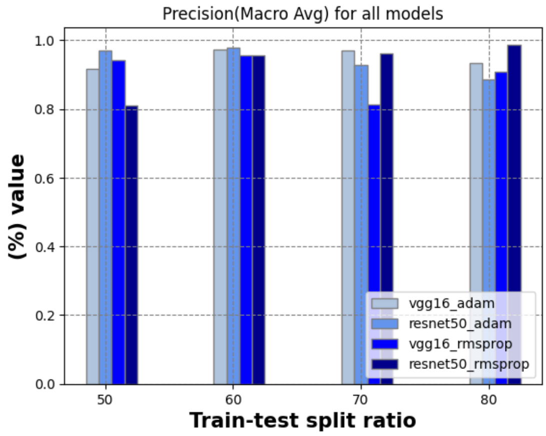


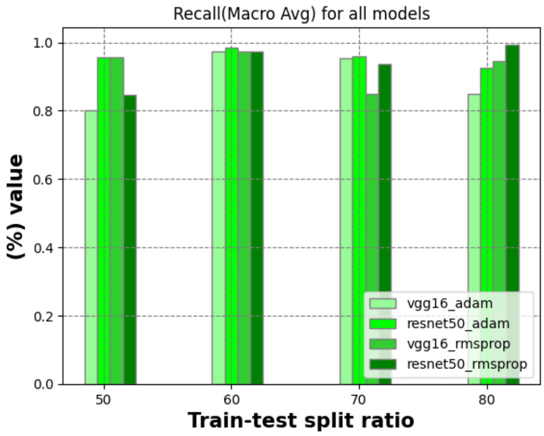
FIGURE 5.9: Comparison of Macro Average w.r.t Adam and RMSprop Optimizer for TL ResNet50 model

5.3.5 Comparative Discussion of proposed TL VGG16 and TL ResNet50

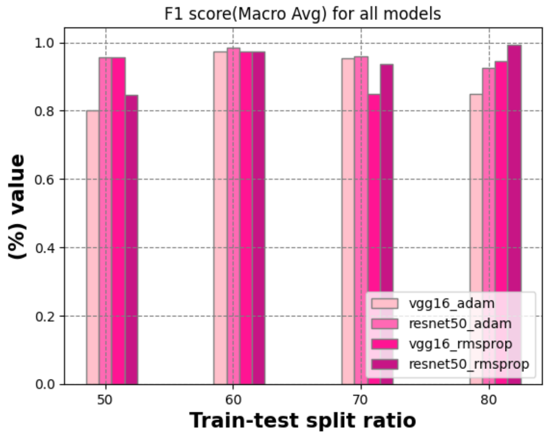
Figure 5.10 shows the comparison of performance parameters calculated based on macro average for both proposed model. The comparative analysis conveys that over all split ratio, the resnet50 model performs maximum with 98.78% precision, 99.38% recall and 99.07% f1_score when RMSprop optimizer is used, as depicted in Fig. 5.10 (a), (b), (c) respectively. As per the fig 5.11, for Vgg16, as the split ratio increases from 50% to 80%, the accuracy decreases from 0.8667 to 0.9 with Adam optimizer, and from 0.9467 to 0.925 with RMSprop. For Resnet50, with the Adam optimizer, the accuracy remains relatively stable across different split ratios (around 0.97), while with RMSprop, it increases from 0.7967 to 0.9917 as the split ratio goes from 50% to 80, which is depicted in fig 5.11.



(a)



(b)



(c)

FIGURE 5.10: Comparison of Macro Average (Precision, Recall and F1-score w.r.t Adam and RMSprop Optimizer for TL ResNet50 model

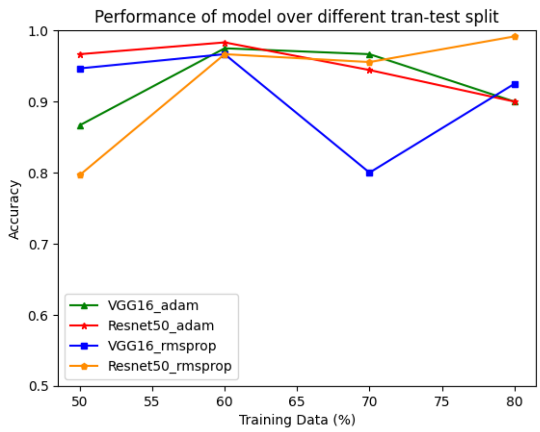


FIGURE 5.11: Comparison of performance of TL VGG16 and TL ResNet50 w.r.t Adam and RMSprop Optimizer and over all data partitioning

5.3.6 Comparative with Existed Research

The proposed model achieved highest overall classification accuracy of 98.52%, real and retouched image classification accuracy of 99.65% and 97.39% respectively for TL ResNet50 with Adam optimizer used during initial training and RMSprop Optimizer used during fine-tuning for the train-test split ratio of 50%-50%. Moreover, the proposed model shows improvement of 16.62% and 11.42% [9] in classifications of retouching for the same dataset, as shown in Table 8. As compared to reference[18], recent paper, the proposed work improved the classification accuracy by 18.52% when the model is trained and evaluated on whole image rather than the face patches. It is worth noting that in Table 5.11 and 5.12, the detection accuracies for the existing techniques are taken directly from reference [9][10][18]because of similar dataset settings. In context to Dataset 2, the previous work utilized a 5-fold cross-validation approach with an 80%-20% split ratio based on the dataset for analysis. As far as our knowledge extends, no subsequent research has been conducted on this dataset. In contrast to the prior work, our model, employing Transfer Learning with a ResNet50 architecture using Adam during initial training and RMSprop optimizer during fine-tuning, exhibits a significant improvement of 5.14% in overall classification accuracy. Furthermore, given the dataset's inherent class imbalance, our approach achieves an impressive 99.17% macro average accuracy. These results clearly demonstrate the superiority of our proposed model over the existing work. These findings lead to the conclusion that the proposed model exhibits superior performance in discerning genuine from retouched images when compared to state-of-the-art models in terms of requirements of computation time and resources.

TABLE 5.11: Comparison of proposed TL model with other state-of-the-art for Dataset 1

Train-Test Split	DL algorithm	Overall	Real	Retouched
50%-50%	Perceptual Metric[4]	48.80%	32.70%	71.90%
	Supervised DBM	81.90%	74.30%	90.90%
	Unsupervised DBM[4]	87.10%	81.10%	93.90%
	Residual CNN[18]	90.00%	93.30%	86.30%
	Proposed TL ResNet50 model	**98.52%**	**99.65%**	**97.39%**

TABLE 5.12: Comparison of proposed TL model with other state-of-the-art for Dataset 2

Reference Paper	Method	Overall Accuracy
Bharti el. Al.[10]	VGG Face+SVM	79.30%
	Supervised DBM	84.20%
	Subclass Supervised Sparse Autoencoder	94.30%
Proposed TL ResNet50 model	RMSprop Optimizer	**99.44%**

5.4 Summary

In this chapter, two transfer learning models were employed, each with specific and identical weights. The chapter focuses on assessing the generalization and performance capabilities of these models, utilizing two distinct facial datasets. The research also examines the robustness of the models through an analysis of their performance with two different optimizers employed during the fine-tuning process. Notably, the study demonstrates that varying models exhibit diverse performance outcomes for a specific train-test split ratio, which is verified in this section. our experiments demonstrate that the proposed Transfer Learning (TL) model, ResNet50, surpasses TL VGG16 architecture in the task of detecting facial retouching. Notably, the model shows enhanced classification efficacy when RMSprop is employed during fine-tuning, compared to Adam Optimizer. Furthermore, ResNet50 exhibits superior generalization, as it achieves the highest retouching classification accuracy for both Dataset 1 and Dataset 2 when an 80%-20% train-test split ratio is considered, making it a strong candidate for real-world applications, like photos uploaded on social media