

Name: Fnu Sowrabh

SAT5115 – Final project progress report 2

1. AI Model(s) Used in Research Project (5 points)

Answer:

A) Models Used: VGG19, InceptionV3, ResNet50

Ensemble Models: VGG19 + InceptionV3, VGG19 + InceptionV3 + ResNet50

B) Transfer Learning: The convolutional base is followed by:

GlobalAveragePooling2D

Dense layers (128 units + Dropout)

Final SoftMax layer for 5-class classification (diabetic retinopathy levels).

Input Size: 224x224x3 for all models.

Optimizers: Adam with learning rate of 1e-4.

Loss Function: Categorical Cross entropy.

C) VGG19: Known for simplicity and effectiveness in medical imaging classification tasks.

InceptionV3: Efficient at capturing multi-scale features via factorized convolutions.

ResNet50: Overcomes vanishing gradient issues, effective for deeper learning and high classification accuracy.

Ensemble Models: Capture complementary strengths of individual models. Improve robustness and generalization, important in real-world medical diagnosis.

These models are well-documented and widely used in healthcare image classification research, making them ideal for benchmarking and ensuring replicability.

1. Performance Metrics Analysis (10 points)

Answer:

A}

	Model	Accuracy	ROC_AUC_Macro	F1_Score	Precision	Recall
0	VGG19	0.748977	0.893550	0.421207	0.615075	0.442997
1	InceptionV3	0.718963	0.880727	0.489765	0.517680	0.478900
2	ResNet50	0.780355	0.904865	0.595193	0.653965	0.574915
3	Ensemble (VGG+Incep)	0.758527	0.903390	0.482716	0.597173	0.478127
4	Ensemble (All 3)	0.777626	0.916162	0.556068	0.666825	0.535287

B) Performance Metrics Analysis (10 points)

Accuracy: Overall correct predictions, good general indicator but limited for imbalanced classes.

ROC AUC (macro): Measures model's ability to distinguish between classes. It is essential in medical imaging, especially for early diagnosis.

F1 Score: Harmonic mean of precision & recall, crucial for imbalanced medical data to balance false positives/negatives.

Precision: How many predicted positives were correct, especially important to avoid over-diagnosis.

Recall (Sensitivity): How many actual positives were correctly predicted, vital to minimize missed diagnoses.

2. Project Status Summary (10 points)

Answer:

Several studies report 5-class classification accuracy of 89–99% on the APTOS dataset using advanced techniques. Kaggle’s APTOS 2019 leaderboard shows top models typically achieve 85–90% accuracy with tuning and ensembles. Our best model, ResNet50, achieved 78.0% accuracy, 90.5% ROC AUC, and 0.595 F1 score.

These results are strong, given we used only single-model training for 25 epochs with basic augmentation. Further improvements can be achieved with advanced preprocessing, ensembling, and extended training.

PROJECT ON TRACK-

Steps remaining:

1. Will try to improve the accuracy by making changes in image pre-processing.
2. Completion of report.