

Skin Cancer Detection Model

Dateset

Dataset used: Kaggle - Skin Cancer Binary Classification Dataset

Evaluation Metrics

Optimal threshold: 0.7452

=== Final Evaluation ===

precision recall f1-score support

Non-Cancer 0.84 0.72 0.77 50

Cancer 0.75 0.86 0.80 50

Accuracy 0.79 100

Macro avg 0.80 0.79 0.79 100

Weighted avg 0.80 0.79 0.79 100

Confusion Matrix:

[[36 14]

[7 43]]Total Misclassified: 21/100 (21.00%)

False Positives: 14

False Negatives: 7

Accuracy: 0.8

Workflow

Skin Cancer Detection Using Deep Learning (MobileNetV2 + Cross-Validation)

1. Dataset Preparation

- Cancer and Non-Cancer images divided into Training and Testing folders (50 each).
- Images loaded using OpenCV, resized to 224x224, and converted to RGB.

2. Label Assignment

- Assigned labels: 1 = Cancer, 0 = Non-Cancer.
- Combined into x_train_full, y_train_full and x_test, y_test.

3. 5-Fold Stratified Cross-Validation

- StratifiedKFold preserves class distribution.
- Each fold includes:
 - Training/Validation split
 - Image augmentation using ImageDataGenerator
 - MobileNetV2 model training with layers:
 - GlobalAveragePooling

- Dense(128) + ReLU
- Dropout(0.5)
- Dense(1) + Sigmoid

4. Callbacks & Training

- Used EarlyStopping, ReduceLROnPlateau, ModelCheckpoint.
- Best model per fold saved based on validation recall.

5. Ensemble Learning

- Loaded all 5 models.
- Predictions averaged on full train and test sets.

5. Threshold Tuning

- Precision-Recall curve used to find optimal threshold (based on F1-score).

6. Evaluation

- Classification report: Precision, Recall, F1-score.
- Confusion matrix with:
 - False Positives
 - False Negatives
 - Accuracy

7. Final Model Saving

- Averaged ensemble outputs from 5 folds.
- Saved as skin_cancer_model.h5

Outcome:

- Improved generalization through cross-validation.
- Balanced performance via threshold tuning.

Skin Cancer Detection using MobileNetV2 with Cross-Validation & Ensemble Learning

