INTERNSHIP REPORT

on

Histopathological Image Analysis Through Deep Learning-Based Image Classification Techniques

Submitted by

YASHARTH KESARWANI

in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology
in
Computer Science and Engineering

Under the supervision of

Dr. Joohi Chauhan
Assistant Professor
Department of Computer Science and Engineering
MNNIT Allahabad, Prayagraj – 211004

Internship Duration: 2nd June – 13th July, 2025

Department of Computer Science and Engineering
Motilal Nehru National Institute of Technology, Allahabad

July 2025

The full block is represented as:

$$\mathcal{F}(\mathbf{x}) = \text{ReLU}(\text{BN}(W_2 \cdot \text{ReLU}(\text{BN}(W_1 \cdot \mathbf{x}))))$$

and the final output becomes:

$$\mathbf{y} = \mathcal{F}(\mathbf{x}) + \mathbf{x}$$

This residual addition is then optionally passed through another ReLU activation.

Network Architecture Variants

Different ResNet architectures differ primarily in depth and the type of residual blocks used:

Architecture	Block Type	Depth	Description
ResNet-18	Basic Block	18	Two 3×3 convolutions per block
ResNet-34	Basic Block	34	Deeper version of ResNet-18
ResNet-50	Bottleneck Block	50	1×1 → 3×3 → 1×1 convolutions
ResNet-101	Bottleneck Block	101	Deeper variant using same block as ResNet-50
ResNet-152	Bottleneck Block	152	Extremely deep architecture with over 60M params

The **bottleneck block** improves computational efficiency by using 1×1 convolutions to reduce and then restore dimensionality around a 3×3 convolution:

$$\mathcal{F}(\mathbf{x}) = W_3 \cdot \text{ReLU}(\text{BN}(W_2 \cdot \text{ReLU}(\text{BN}(W_1 \cdot \mathbf{x}))))$$

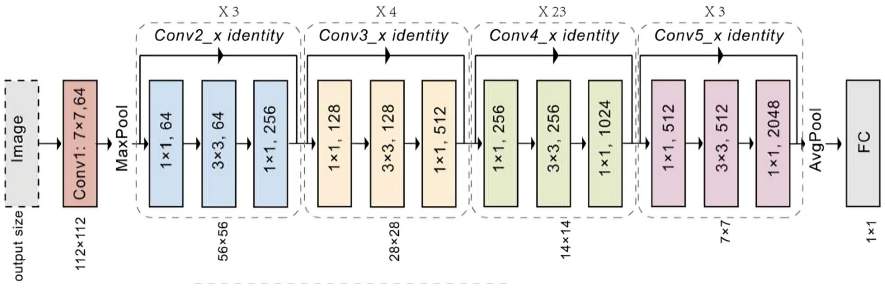

Figure 2: Illustration of a Residual Network (ResNet) architecture, highlighting residual blocks with skip connections that allow the network to learn identity mappings and improve gradient flow in deep convolutional models.

Benefits of ResNet

- **Improved gradient flow**: Skip connections help gradients bypass multiple layers, mitigating vanishing gradients.
- **Identity mapping**: Enables the network to preserve important features when no further transformation is required.
- **Better generalization**: Residual learning regularizes the model implicitly.

Building Blocks

EfficientNet is composed of **MBConv blocks** (Mobile Inverted Bottleneck Convolution), which are adapted from MobileNetV2. Each MBConv block includes:

- A pointwise 1×1 convolution for dimensionality expansion
- A depthwise separable 3×3 or 5×5 convolution
- A squeeze-and-excitation (SE) module to model inter-channel dependencies
- A final 1×1 pointwise convolution to project back to the desired dimensions

The activation function used is **Swish**, defined as:

$$\text{Swish}(x) = x \cdot \sigma(x)$$

where $\sigma(x)$ is the sigmoid function. Swish provides smoother gradients and improves convergence over ReLU in deeper networks.

Architecture Overview

The EfficientNet-B0 architecture contains the following key stages:

- **Stem**: One 3×3 convolution layer with stride 2
- **MBConv Blocks**: Stacked across several stages with increasing number of channels and decreasing spatial resolution
- **Head**: Global average pooling followed by dropout and a fully connected dense layer for classification

Each version from B1 to B7 increases model size, input resolution, and computational cost while maintaining the architectural skeleton.

Model	Parameters	Input Resolution	Top-1 Accuracy
EfficientNet-B0	$\tilde{5}$.3M	224×224	$\tilde{7}$7.1%
EfficientNet-B3	$\tilde{1}$2M	300×300	$\tilde{8}$1.6%
EfficientNet-B4	$\tilde{1}$9M	380×380	$\tilde{8}$2.9%

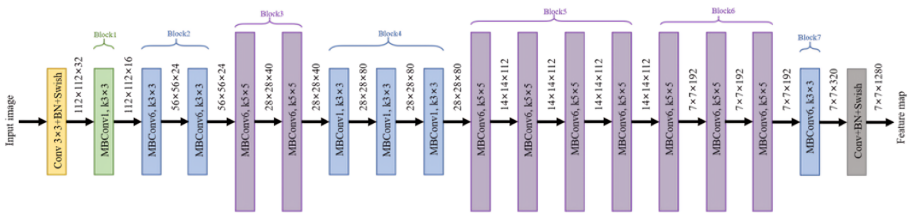

Figure 3: EfficientNet architecture showcasing compound scaling across depth, width, and resolution, along with its modular MBConv blocks and squeeze-and-excitation mechanisms.

Evaluation Metrics

The trained model was evaluated on the validation and test sets using the following metrics:

- **Accuracy**
- **Precision**
- **Recall**
- **F1-score**

Results

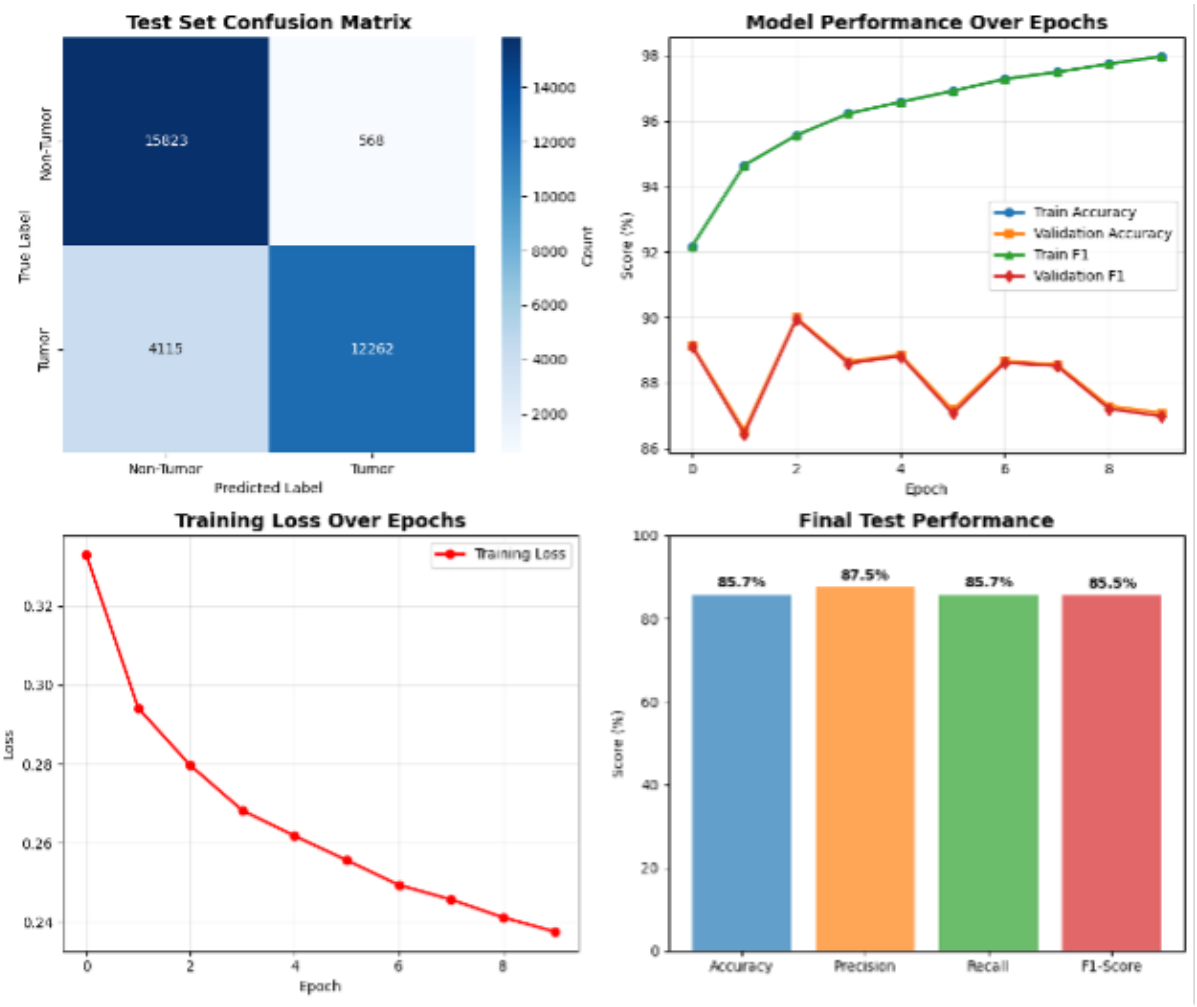

Figure 4: Evaluation Results on PCam Dataset: Confusion Matrix, Training Loss, Validation Metrics, and Final Test Performance

Validation Performance:

- Accuracy: 89.99%
- Precision: 90.65%

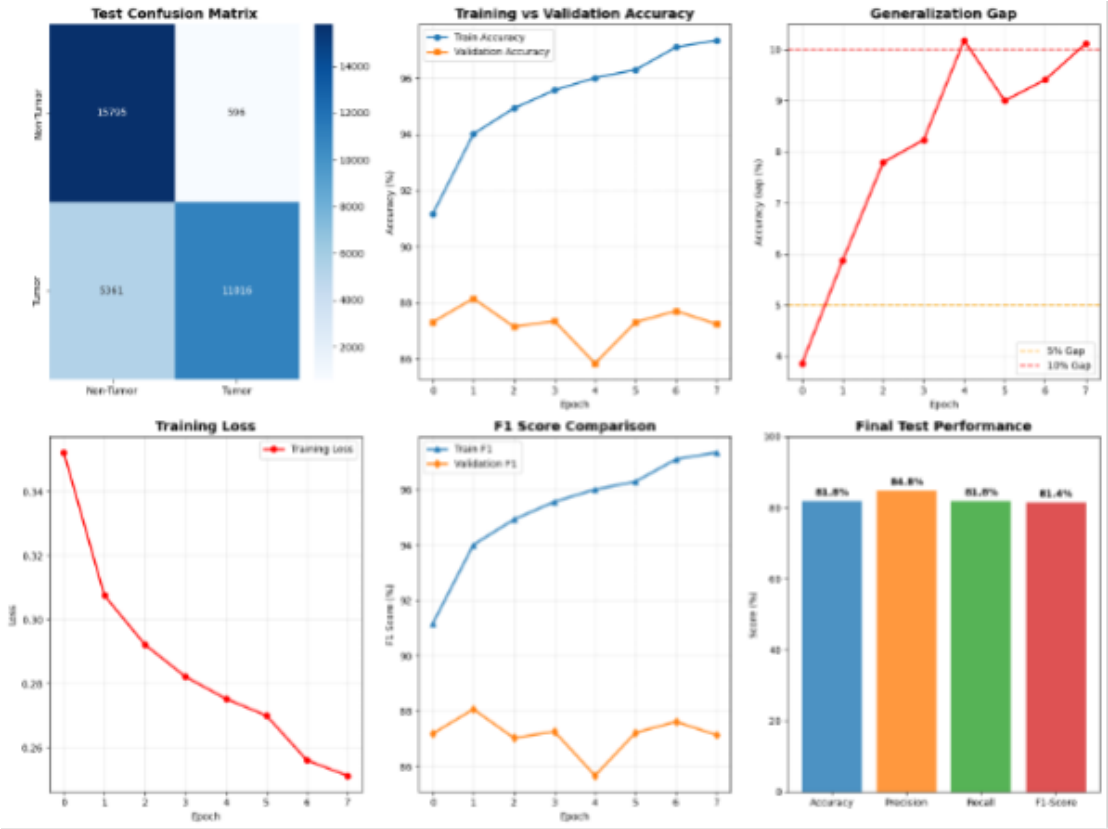

Figure 5: (Top-left) Confusion Matrix on Test Set, (Top-middle) Accuracy over Epochs,(Top-right)Generalization-Gap, (Bottom-left) Training Loss over Epochs, (Bottom-middle) F1-Score Comparison, (Bottom-right) Final Test Performance.

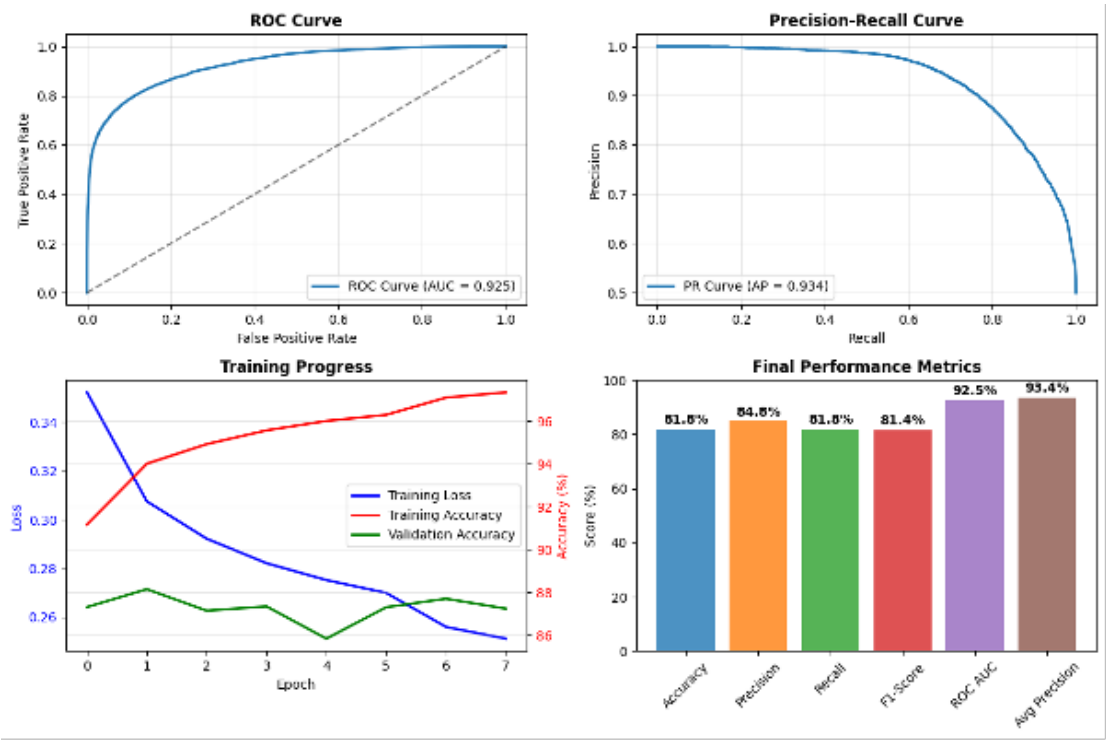

Figure 6: (Top-left) ROC Curve, (Top-right) Precision-Recall Curve, (Bottom-left) Training Progress, (Bottom-right) Final Performance Metrics.

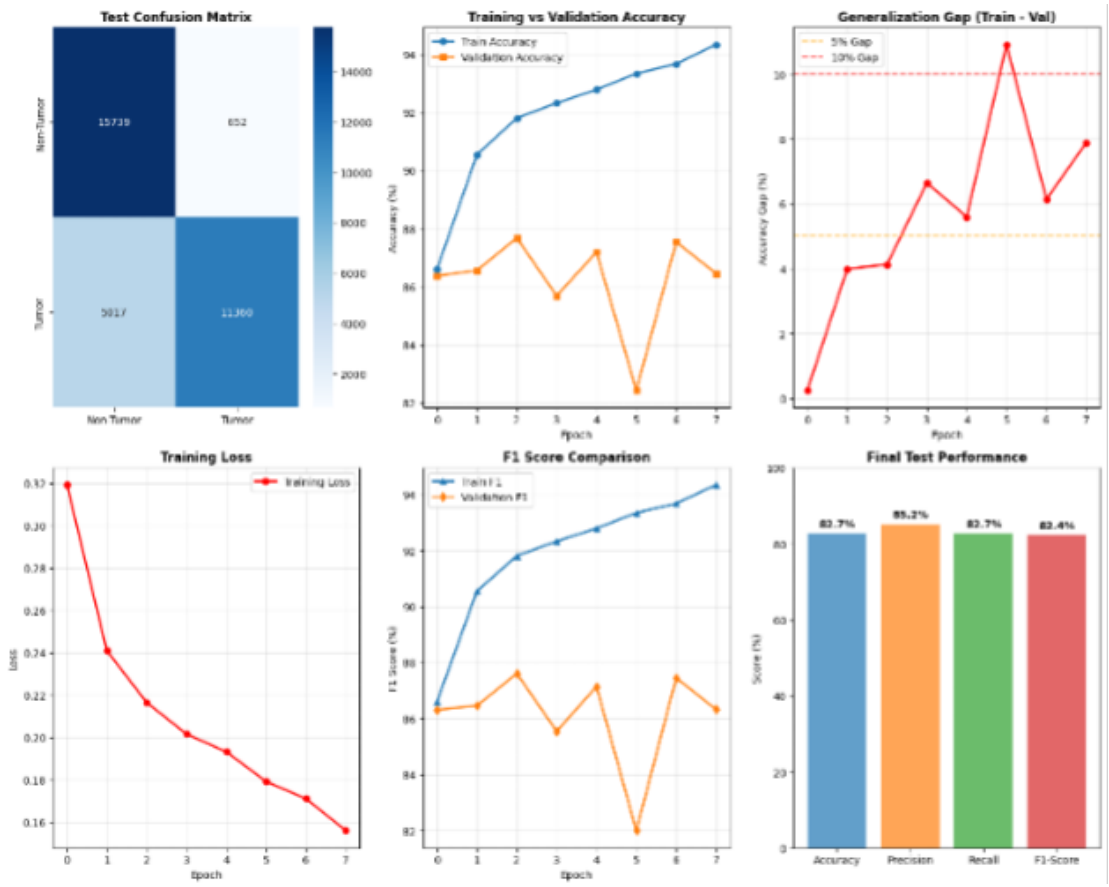

Figure 7: (Top-left) Confusion Matrix on Test Set, (Top-middle) Accuracy over Epochs, (Top-right) Generalization Gap, (Bottom-left) Training Loss over Epochs, (Bottom-middle) F1-Score Comparison, (Bottom-right) Final Test Performance.

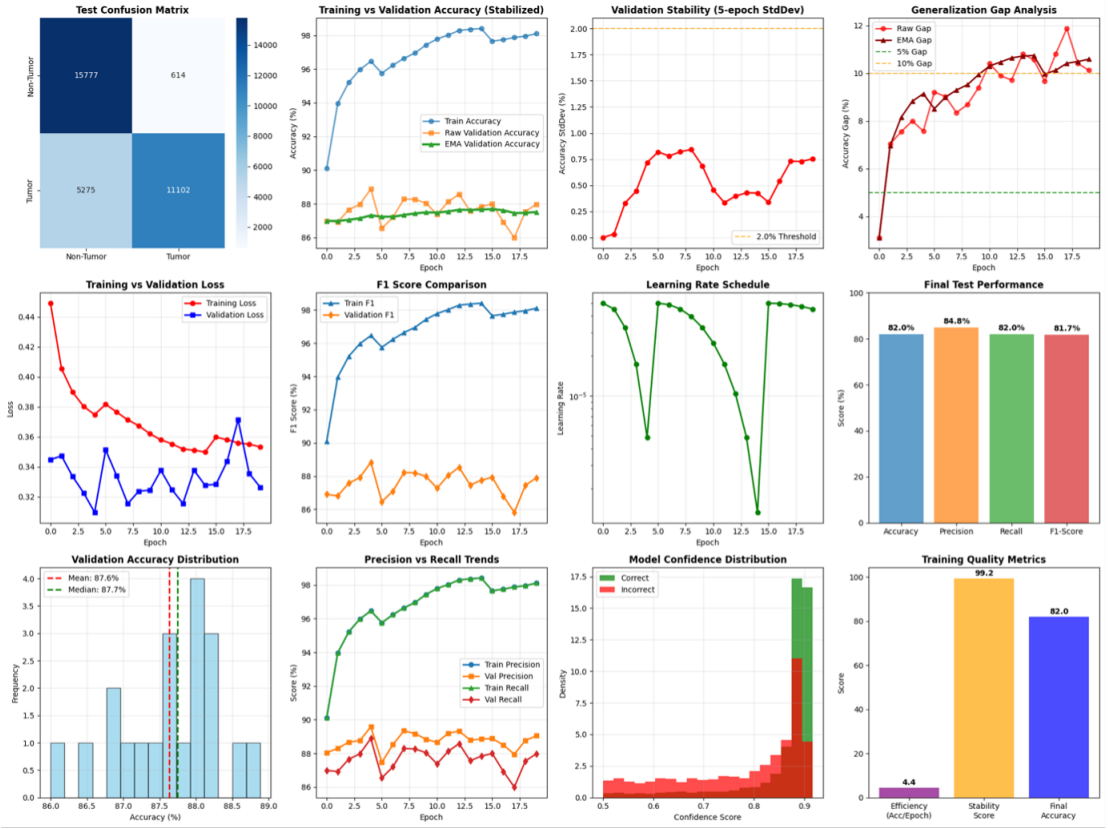

Figure 8: (Top-left) Test Confusion Matrix, (Top-middle-left) Training vs Validation Accuracy (Stabilized), (Top-middle-right) Validation Stability (5-epoch StdDev), (Top-right) Generalization-Gap Analysis, (Middle-left) Training vs Validation Loss, (Middle-middle-left) F1 Score Comparison (Middle-middle-right) Learning Rate Scheduler Curve , (Middle-right) Final Test performance, (Bottom-left) Validation Accuracy Distribution, (Bottom-middle-left) Precision vs Recall Trends,(Bottom-middle-right) Model Confidence Distribution, (Bottom-right) Model Confidence Distribution and Training Quality Metrics.

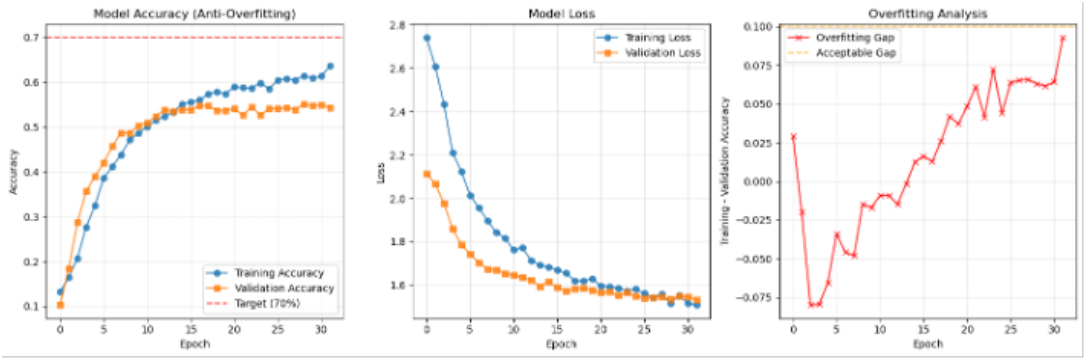

Figure 9: (Left) Training and Validation Accuracy Progression over Epochs, (Middle) Training and Validation Loss over Epochs, (Right) Overfitting Anaysis.

5.6 Implementation 6: Histopathologic Cancer Detection Using EfficientNet-B4 on BRACS Dataset

This section presents a comprehensive experimental framework for histopathologic cancer detection using an ultra-optimized EfficientNet-B4 architecture applied to the BRACS dataset. The study utilized advanced optimization techniques, including aggressive fine-tuning, enhanced data augmentation, and refined regularization strategies.

Dataset Configuration

The BRACS dataset consists of 4,516 histopathologic images categorized into seven diagnostic classes: 0_N (Normal), 1_PB (Papillary Benign), 2_UDH (Usual Ductal Hyperplasia), 3_FEA (Flat Epithelial Atypia), 4_ADH (Atypical Ductal Hyperplasia), 5_DCIS (Ductal Carcinoma In Situ), and 6_IC (Invasive Carcinoma). The class distribution is as follows:

- 0_N: 484 samples (10.7%)
- 1_PB: 834 samples (18.5%)
- 2_UDH: 515 samples (11.4%)
- 3_FEA: 756 samples (16.7%)
- 4_ADH: 506 samples (11.2%)
- 5_DCIS: 779 samples (17.2%)
- 6_IC: 642 samples (14.2%)

A class imbalance ratio of 1.72 was observed. Patient-level stratified splitting was used to partition the dataset into:

- Training: 4,163 images (92.2%)
- Validation: 290 images (6.4%)
- Test: 63 images (1.4%)

All images were resized to 512×512 pixels using LANCZOS resampling.

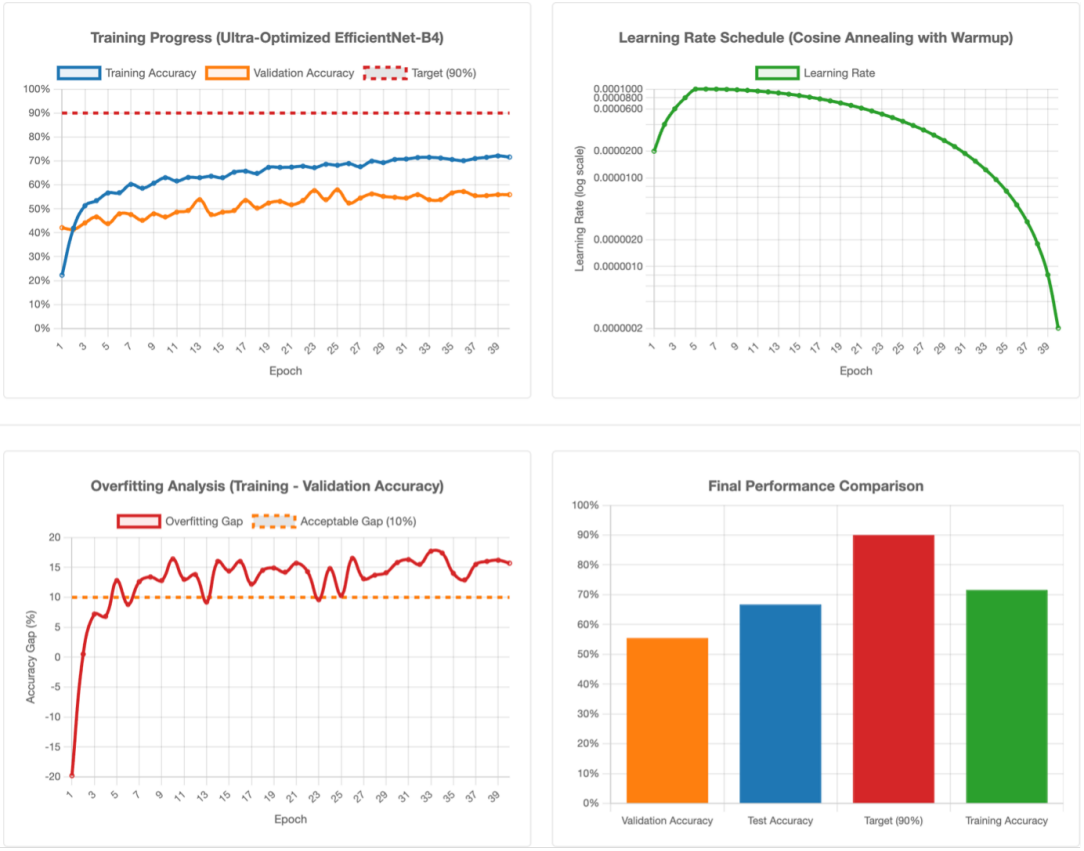

Figure 10: (Top-left) Training Progress (Training vs Validation Accuracy), (Top-right) Learning Rate Schedule (Cosine Annealing with Warmup), (Bottom-left) Overfitting Analysis (Training - Validation Accuracy), (Bottom-right) Final Performance Comparison

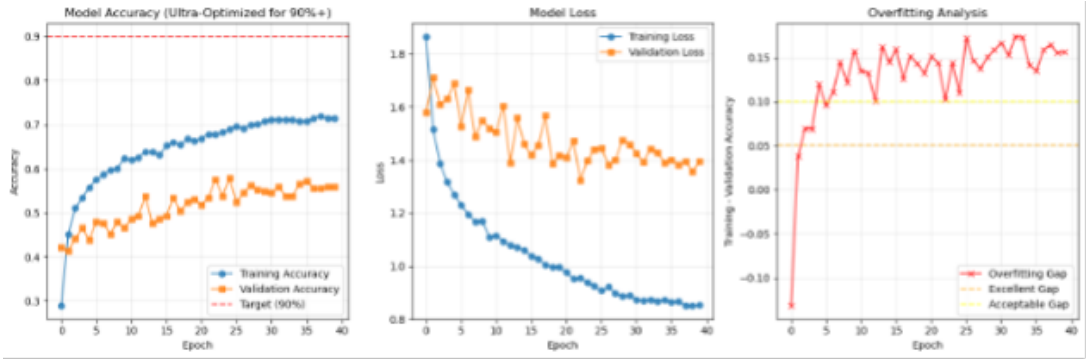

Figure 11: (Left) Training and Validation Accuracy Progression over Epochs, (Middle) Training and Validation Loss over Epochs, (Right) Overfitting Anaysis.

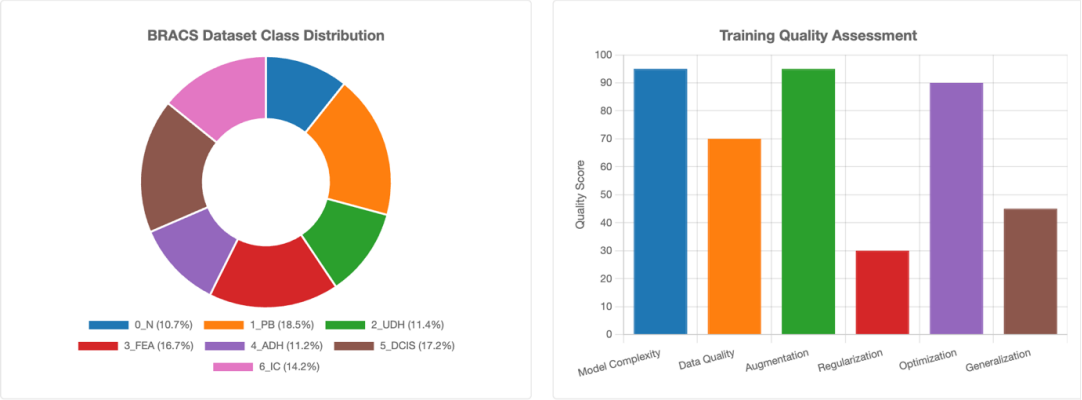

Figure 12: (Left) BRACS Dataset Class Distribution (0_{N}: 10.7%, 1_{PB}: 18.5%, 2_{UDH}: 14.4%, 3_{FEA}: 16.7%, 4_{ADH}: 11.2%, 5_{DCIS}: 17.2%, 6_{IC}: 14.2%), (Right) Training Quality Assessment (Model Complexity, Data Quality, Augmentation, Regularization, Optimization, Generalization).

5.7 Implementation 7 : EfficientNet-B3 for BRACS Histopathologic Cancer Detection

This section presents a comprehensive experimental framework for histopathologic cancer detection using an EfficientNet-B3 architecture applied to the BRACS dataset. The study implemented advanced optimization techniques, enhanced data augmentation, and refined regularization strategies, targeting over 70% validation accuracy.

Dataset Configuration

The BRACS dataset comprises 4,516 histopathologic images labeled across seven diagnostic categories: 0_N (Normal), 1_PB (Papillary Benign), 2_UDH (Usual Ductal Hyperplasia), 3_FEA (Flat Epithelial Atypia), 4_ADH (Atypical Ductal Hyperplasia), 5_DCIS (Ductal Carcinoma In Situ), and 6_IC (Invasive Carcinoma).

Class distribution ranged from 484 to 834 samples per class (10.7% to 18.5%), with a moderate imbalance ratio of 1.72. Patient-level stratified splitting was employed to avoid data leakage. The dataset was partitioned as:

- **Training Samples:** 3,854 images (85.3%)
- **Validation Samples:** 477 images (10.6%)
- **Test Samples:** 185 images (4.1%)

All images were resized to $320 \times 320 \times 3$ before model input.

Data Preprocessing and Advanced Augmentation

Advanced Preprocessing Pipeline:

- Resizing: 320×320 using LANCZOS resampling
- Normalization: Pixel values normalized to float32 range

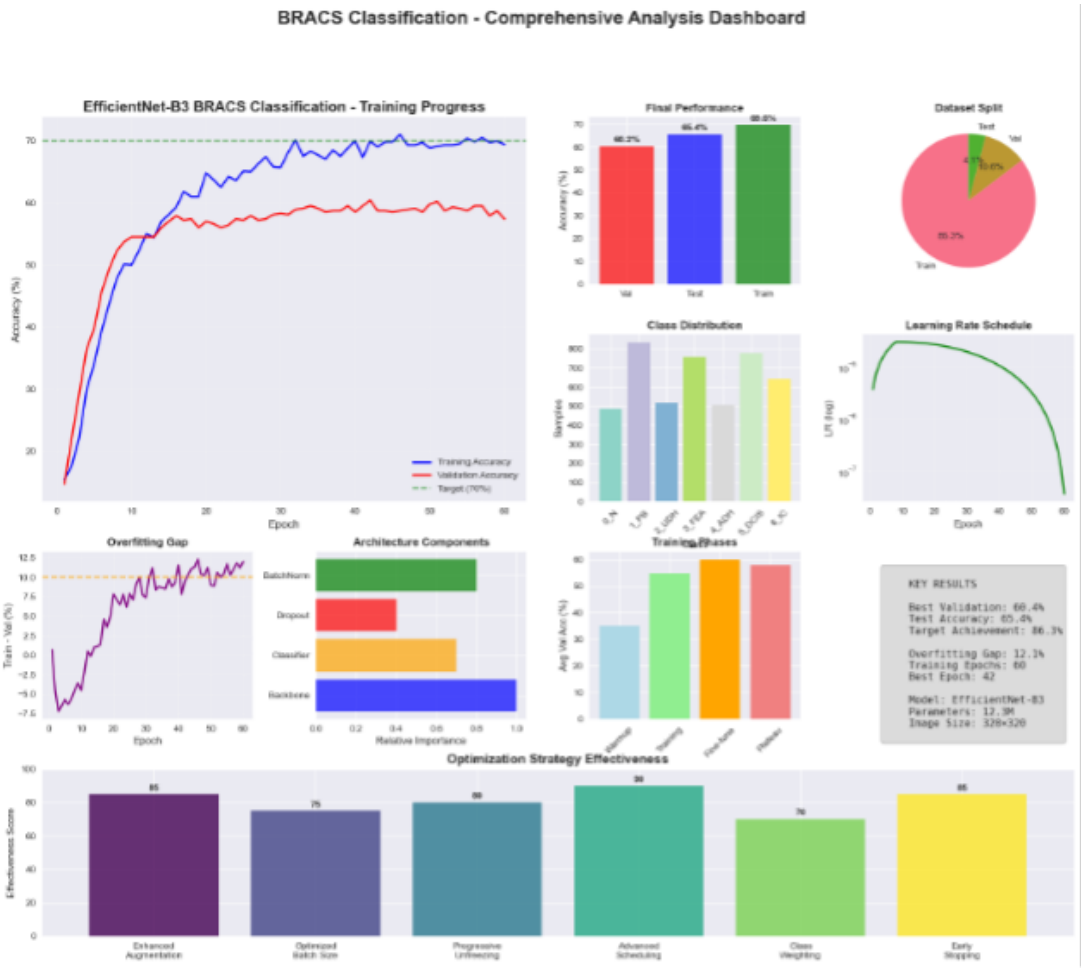

Figure 13: (Left) Training Progress (Performance Accuracy for Validation, Test, and Train Dataset Split), (Middle) Learning Rate Schedule and Epoch Overfit Gap, (Right) Architecture Components (BatchNorm, Dropout, Classifier, Backbone) and Class Distribution. Optimization Strategy Effectiveness: Enhanced Augmentation, Optimized Batch Size, Progressive Unfreezing, Advanced Scheduling, Class Weighting. Key Results: Target Achievement, Model Parameters, Image Size, Chart/Graph Visualization, Data Insight, Report Summary, Score Improvement Technique, Efficiency and Effectiveness, Strategy Enhancement, Progress Monitoring, Evaluation Metric, Result Comparison, Trend Analysis, Data Distribution, Performance Metrics across Training, Testing, and Validation Phases, Target Accuracy, Relative Importance, Early Stopping.

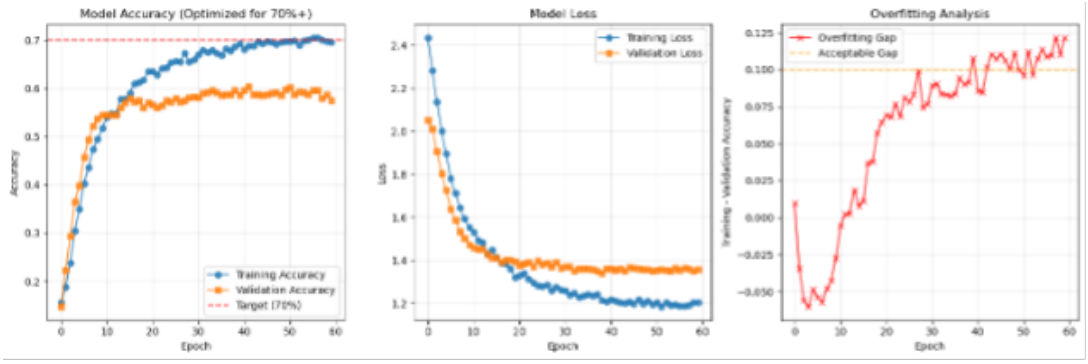

Figure 14: (Left) Training and Validation Accuracy Progression over Epochs, (Middle) Training and Validation Loss over Epochs, (Right) Overfitting Anaysis.

5.8 Implementation 8 : EfficientNet-B3 for BRACS Histopathologic Cancer De- tection

This section presents a comprehensive experimental framework for histopathologic cancer detection, employing a refined EfficientNet-B3 architecture applied to the BRACS dataset. This iteration prioritized training stability and correctness, aiming for a validation accuracy of 70% or higher.

Dataset Configuration

The BRACS dataset consists of 4,516 histopathologic images, categorized into seven diagnostic classes: 0_N (Normal), 1_PB (Papillary Benign), 2_UDH (Usual Ductal Hyperplasia), 3_FEA (Flat Epithelial Atypia), 4_ADH (Atypical Ductal Hyperplasia), 5_DCIS (Ductal Carcinoma In Situ), and 6_IC (Invasive Carcinoma).

Class distributions range from 484 to 834 images (10.7%–18.5%), yielding a moderate class imbalance (ratio 1.72). The dataset was split as follows:

- **Training Samples:** 3,854 images (85.3%)
- **Validation Samples:** 477 images (10.6%)
- **Test Samples:** 185 images (4.1%)

All input images were resized to $300 \times 300 \times 3$ before model ingestion.

Data Preprocessing and Conservative Augmentation

Basic Preprocessing:

- Resizing: All patches resized to 300×300
- Tensor Conversion: HWC $\rightarrow$ CHW format
- Pixel Normalization: Scaled to $[0, 1]$ by dividing by 255
- EfficientNet-Compatible Normalization: ImageNet mean $= [0.485, 0.456, 0.406]$, std $= [0.229, 0.224, 0.225]$

Conservative Augmentation Strategy (50% Application Rate):

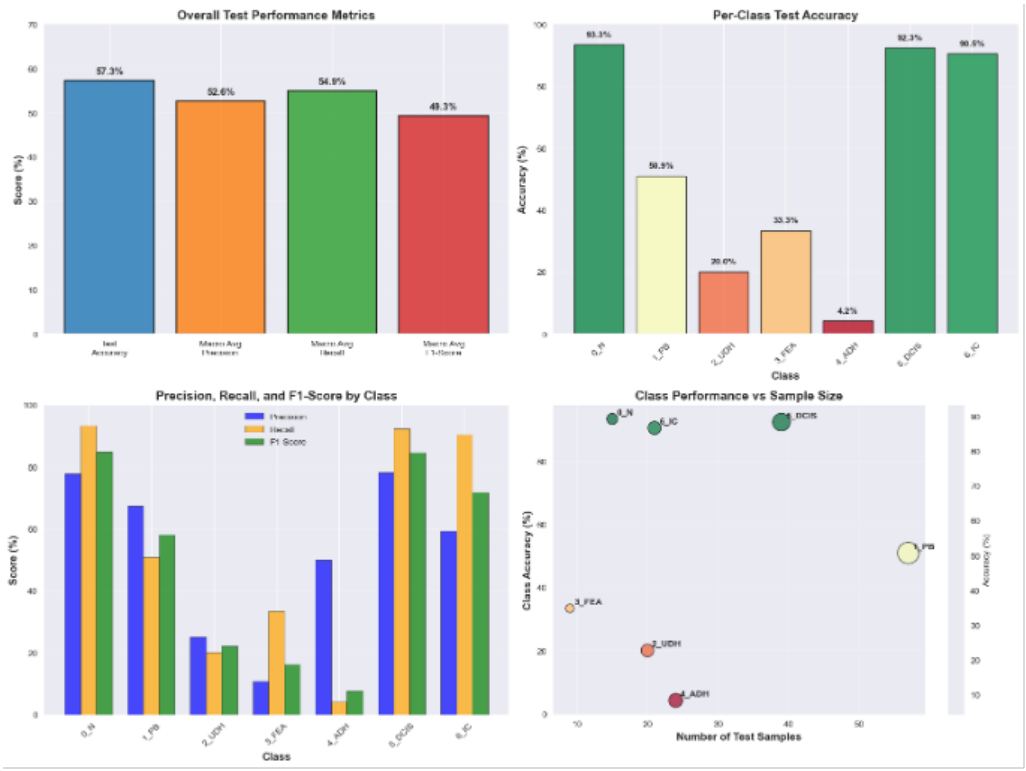

Figure 15: (Top-left) Overall Test Performance Metrics , (Top-right) Per-Class Test Accuracy (0_{N}, 1_{PB}, 2_{UDH}, 3_{FEA}, 4_{ADH}, 5_{DCIS}, 6_{IC}), (Bottom-left) Precision, Recall, and F1-Score by Class, (Bottom-right) Class Performance vs Sample Size.

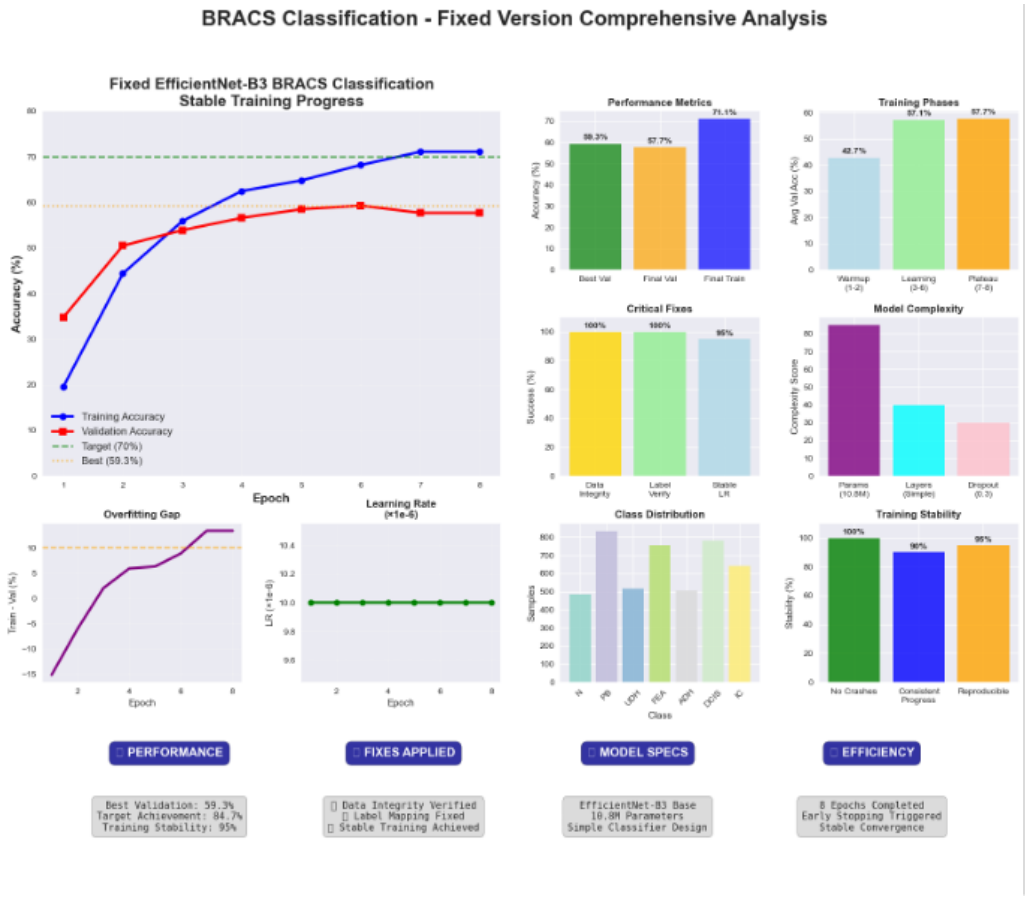

Figure 16: (Left) Stable Progress (Training vs Validation Accuracy), (Right) Performance Metrics with Fixes (Complexity, Learning Rate, Overfitting, Epochs, Integrity, Mapping, Efficiency, Targets, Layers, Dropout, Parameters, Early Stopping, Convergence, Reproducibility).

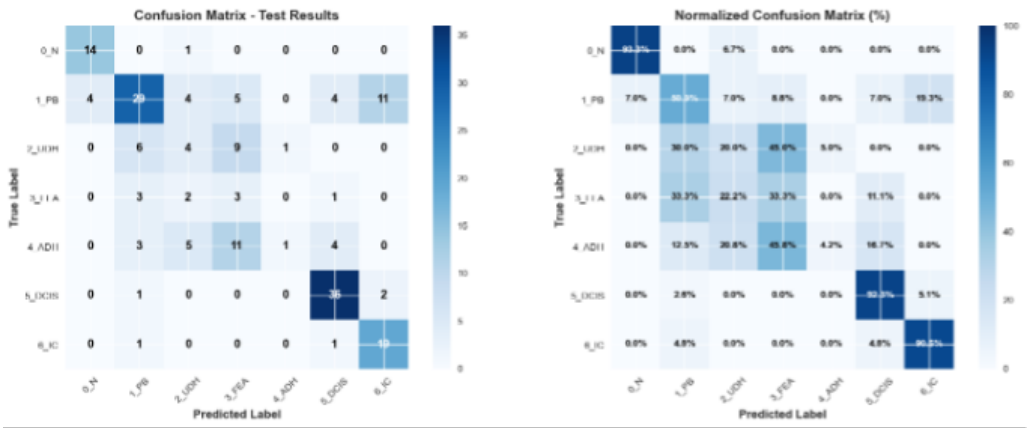

Figure 17: (Left) Confusion Matrix, (Right) Performance Metrics (Normalized Accuracy, Precision, Recall, F1-Score, Error Rate).

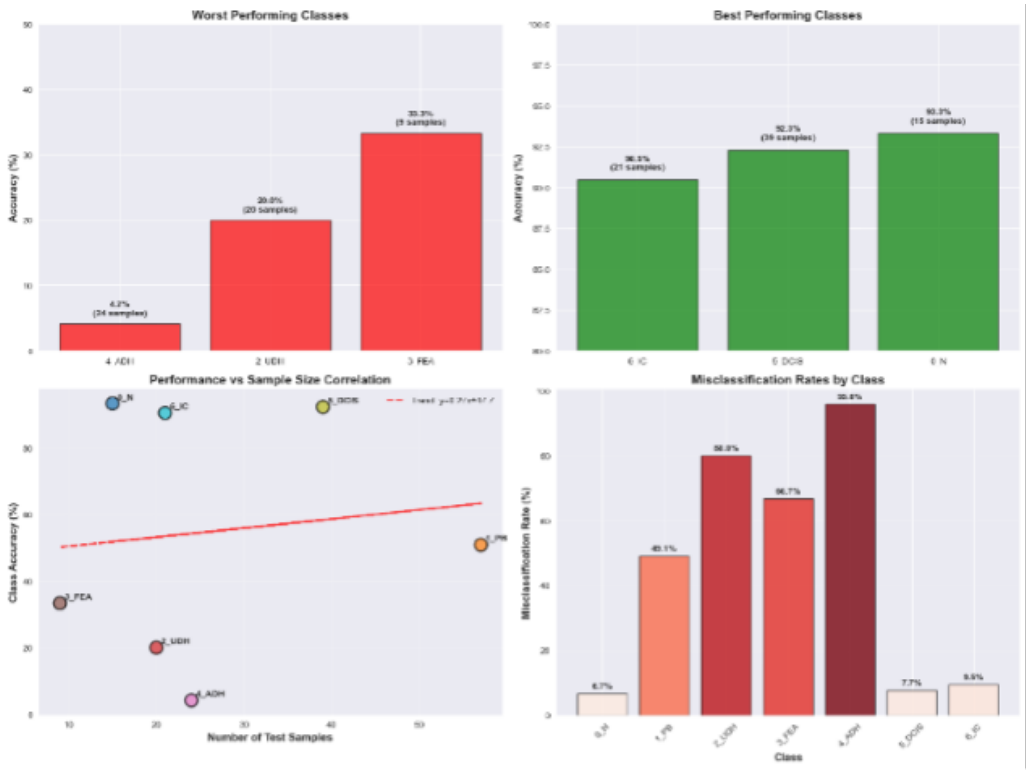

Figure 18: (Top-left) Worst Performing Classes, (Top-right) Best Performing Classes, (Bottom-left) Performance vs Sample Size Correlation, (Bottom-right) Misclassification Rates by Class.

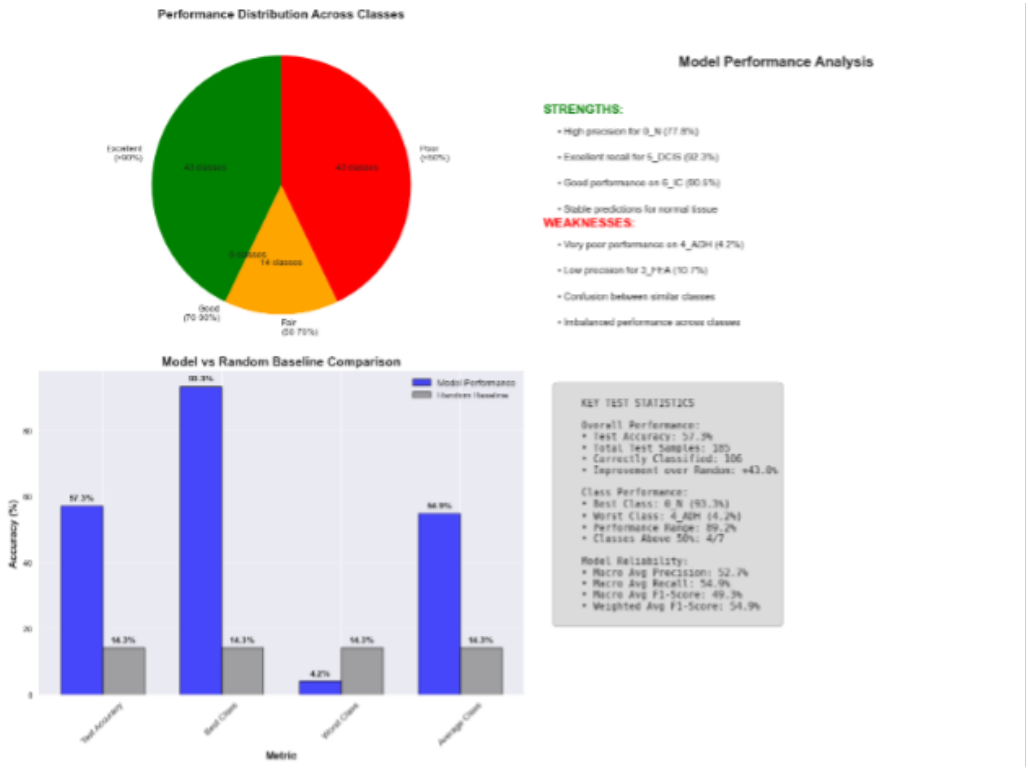

Figure 19: (Left) Performance Distribution Across Classes, (Middle) Model vs Random Baseline Comparison, (Right) Model Performance Analysis (Strengths, Weaknesses, Key Test Statistics).

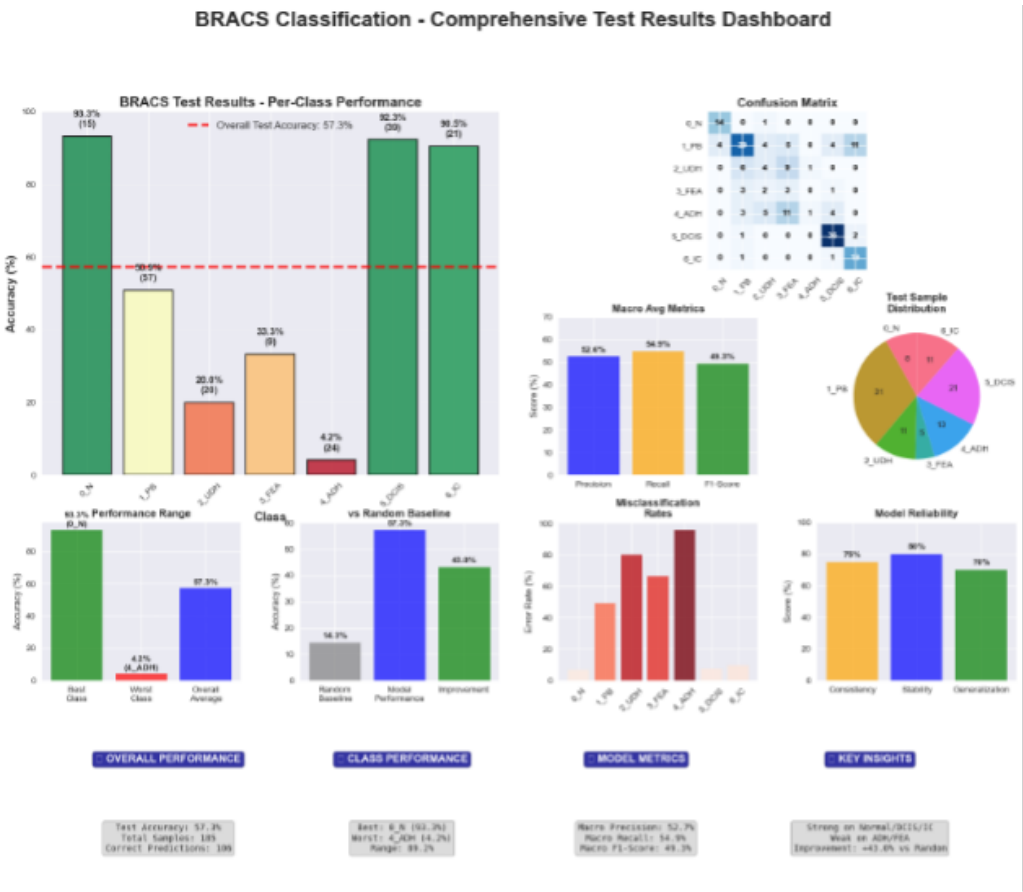

Figure 20: (Top-left) BRACS Test Results - Per-Class Performance, (Top-right) Confusion Matrix, (Middle-left) % Performance Range, (Middle-right) Macro Avg Metrics, (Bottom-left) Misclassification Rates, (Bottom-right) Model Reliability.