SKIN CANCER CLASSIFICATION: ORIGINAL VS. MODIFIED APPROACH

AMAN ALLAH ALTALAY 2021028

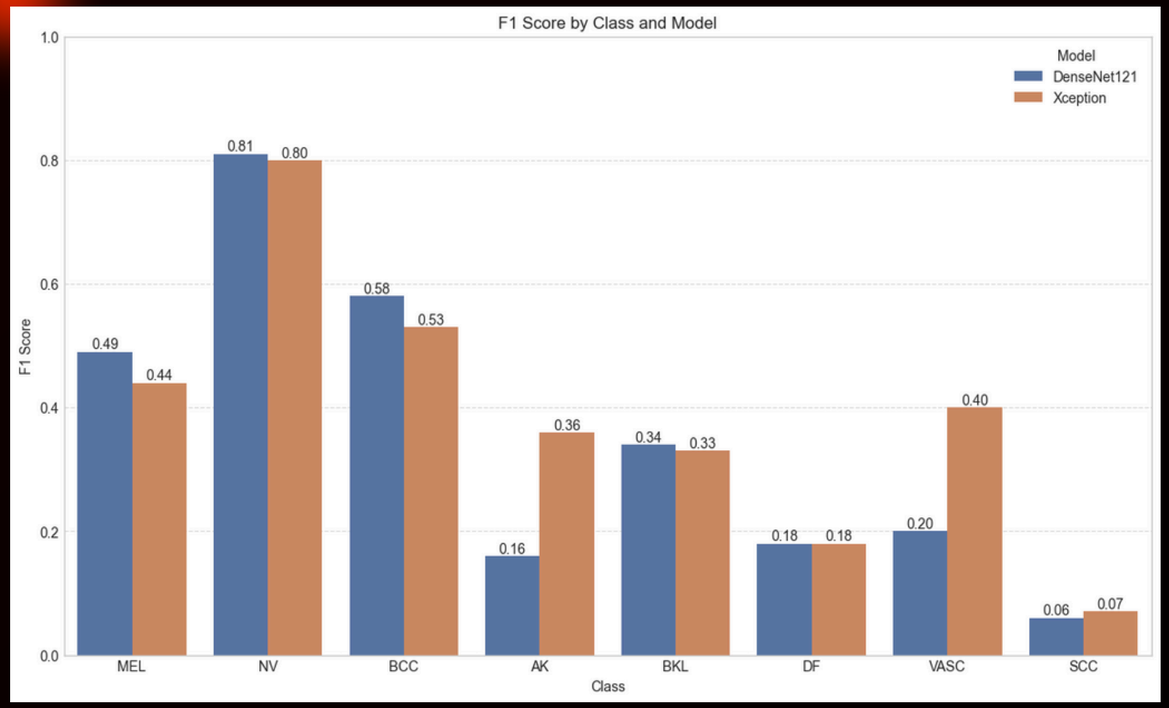
HALID KERDI 2105551

PROJECT OVERVIEW

- Goal: Compare our implementation of DenseNet121 and Xception with research paper models
- Dataset : ISIC Dataset with dermoscopic images (8 skin lesion classes)
- Analysis
- Research Paper: "Classifying Melanoma in ISIC Dermoscopic Images Using Efficient Convolutional Neural Networks and Deep Transfer Learning"

RESEARCH PAPER REFERENCE

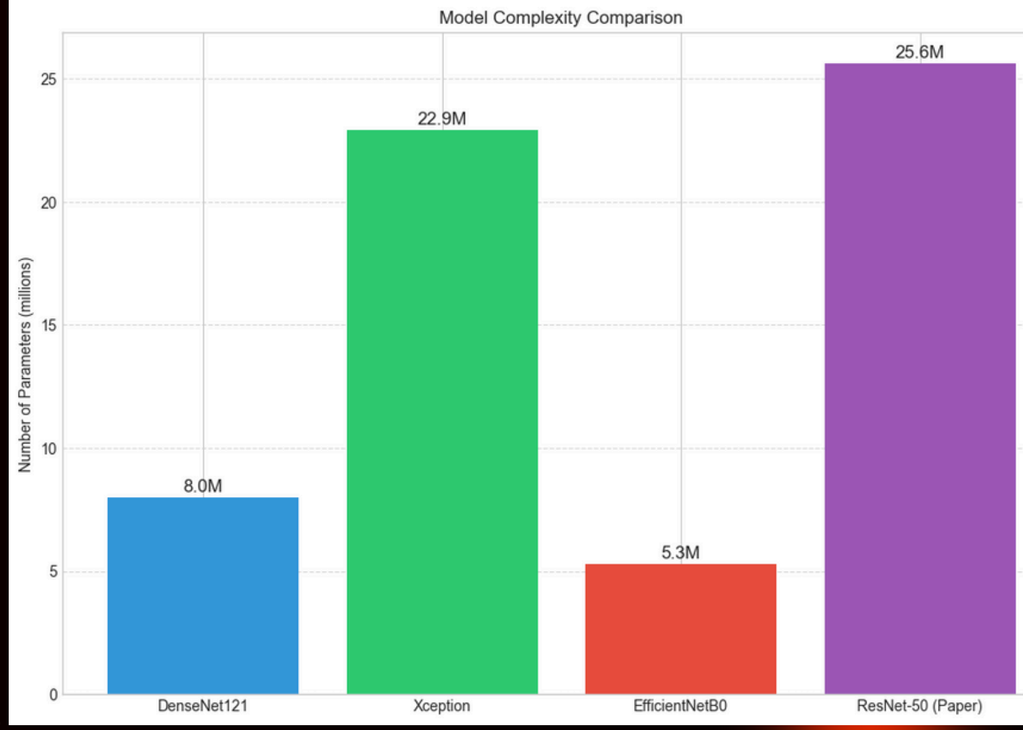
- Paper Focus : Binary classification (melanoma vs. nevus)
- Top Performing Models :
  - ResNet-50: 92.98% accuracy, 0.927 AUC, 95.08% recall, 92.46% precision, 93.75% F1
  - EfficientNetB0: 91.20% accuracy, 0.908 AUC, 94.20% recall, 90.33% precision, 92.22% F1
  - AlexNet: 92.56% accuracy, 0.923 AUC, 94.77% recall, 92.04% precision, 93.38% F1
  - VGG-19: 91.79% accuracy, 0.911 AUC, 97.54% recall, 88.76% precision, 92.94% F1



METHODOLOGY

Key Differences from Paper:

- Our project tackled 8-class classification, while the paper focused on binary classification (melanoma vs. nevus).
- We used a 20% subset of the ISIC 2020 dataset due to hardware limitations.
- All models were trained locally on an Apple M2 MacBook for 3 epochs, while the paper used extended training on high-end GPUs.
- We added EfficientNetB0 for a direct performance comparison with the paper.
- Our evaluation emphasized F1-score, confusion matrices, and validation vs. test accuracy gaps.



Models Implemented:

- DenseNet121 (Halid): 8M parameters, 224x224 input, custom classifier head. Performed well across classes.
- Xception (Aman): 22.9M parameters, 224x224 input, deep dense layers. Strong on common classes.
- EfficientNetB0 (Joint): 5.3M parameters, 224x224 input, lightweight and stable across classes.

All models were trained using the Adam optimizer with categorical cross-entropy loss and early stopping.

TRAINING PROCESS COMPARISON

Training Process Comparison

DenseNet121:

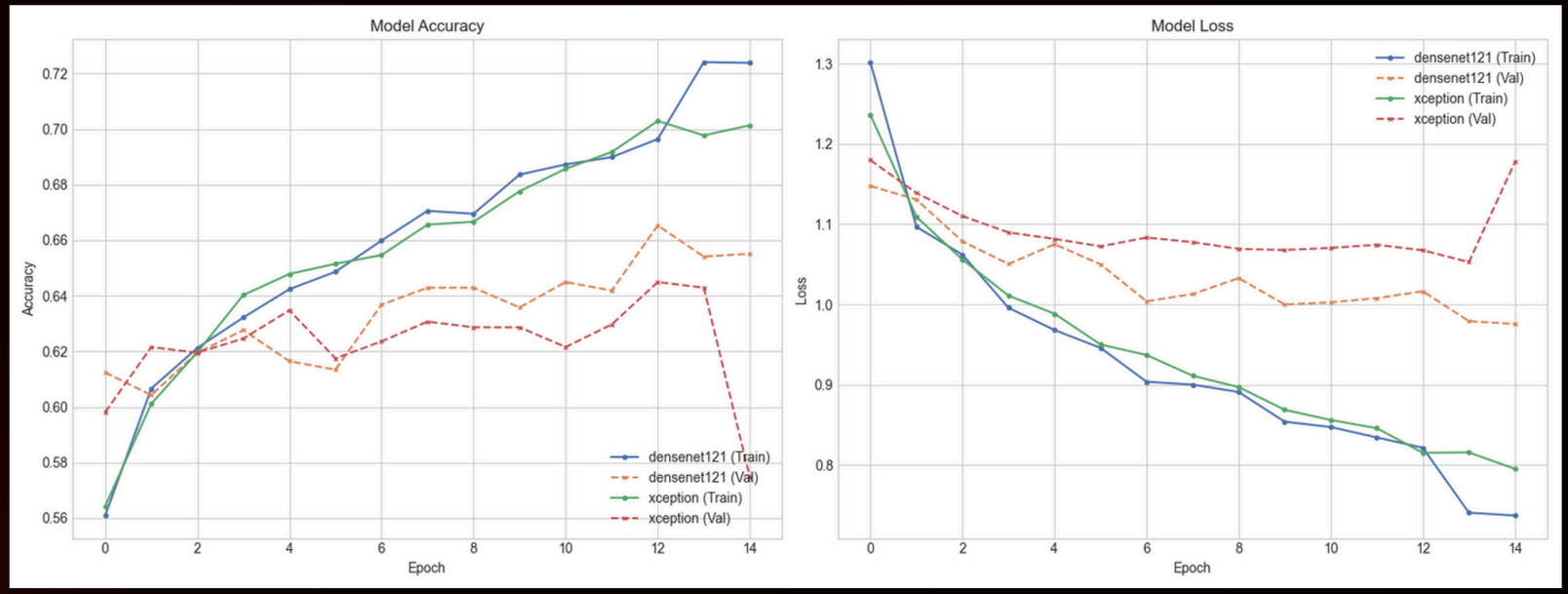
- Validation accuracy: 65.51%
- Training progression showed gradual improvement across epochs
- Final model achieved best results on NV class (F1: 0.81)

Xception:

- Validation accuracy: 64.29%
- Similar training pattern to DenseNet121
- Strong performance on NV (0.80) and BCC (0.53) classes

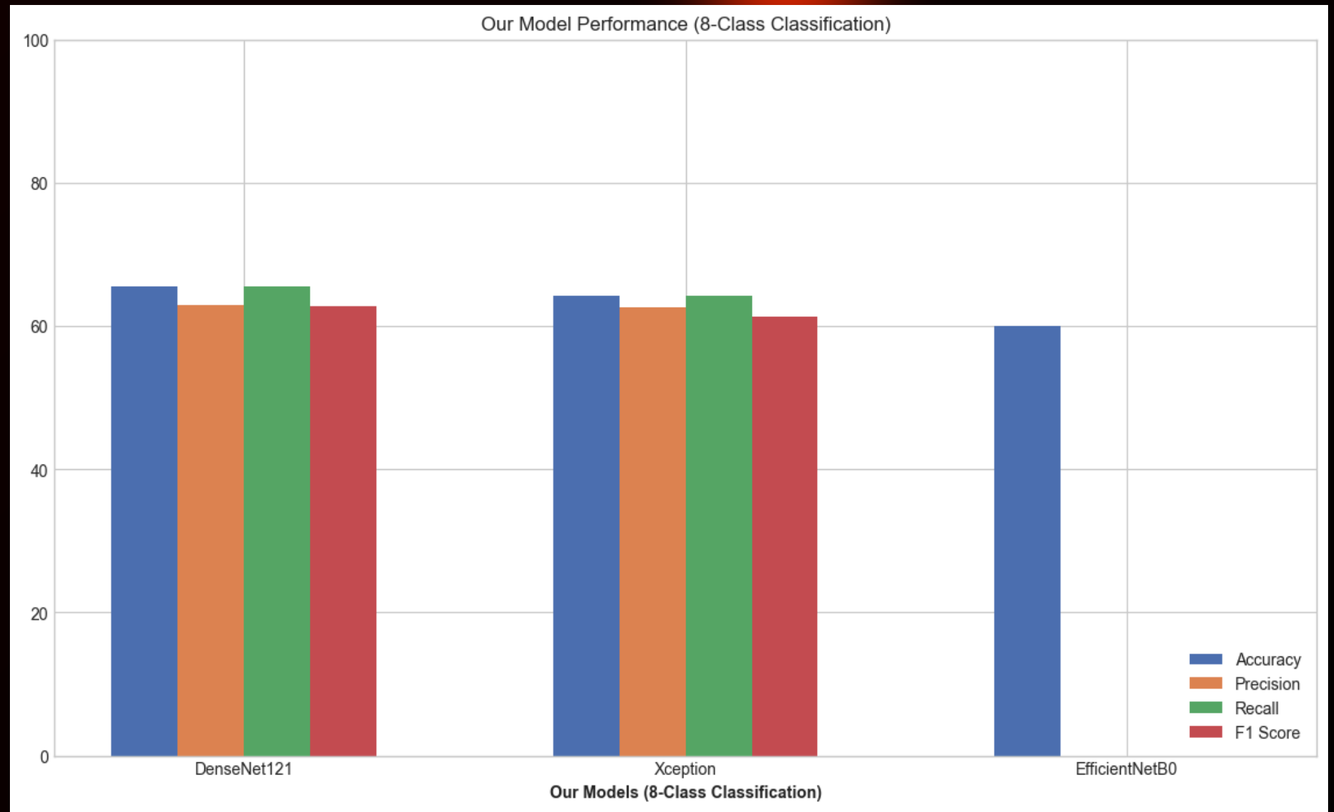
EfficientNetB0:

- Run locally with limited epochs
- Reached ~60% validation accuracy
- Showed promising performance despite limited training time



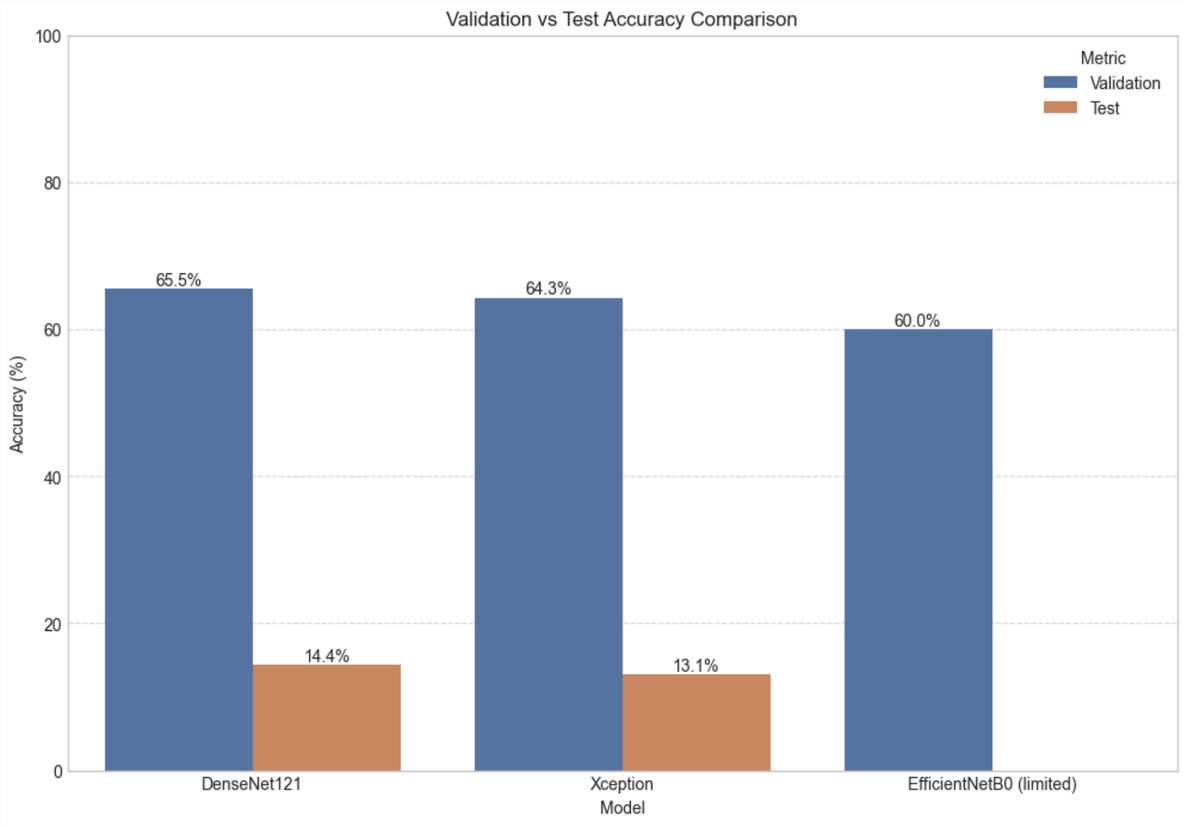
PERFORMANCE METRICS VS. PAPER

- Our Models (8-class classification):
  - DenseNet121: Accuracy: 65.51%, Precision: 62.89%, Recall: 65.51%, F1: 62.83%
  - Xception: Accuracy: 64.29%, Precision: 62.54%, Recall: 64.29%, F1: 61.37%
  - EfficientNetB0: ~60% accuracy (run locally on laptop with limited training epochs)
- Paper Models (binary classification) :
  - ResNet-50: 92.98% accuracy, F1: 93.75%
  - EfficientNetB0: 91.20% accuracy, F1: 92.22%



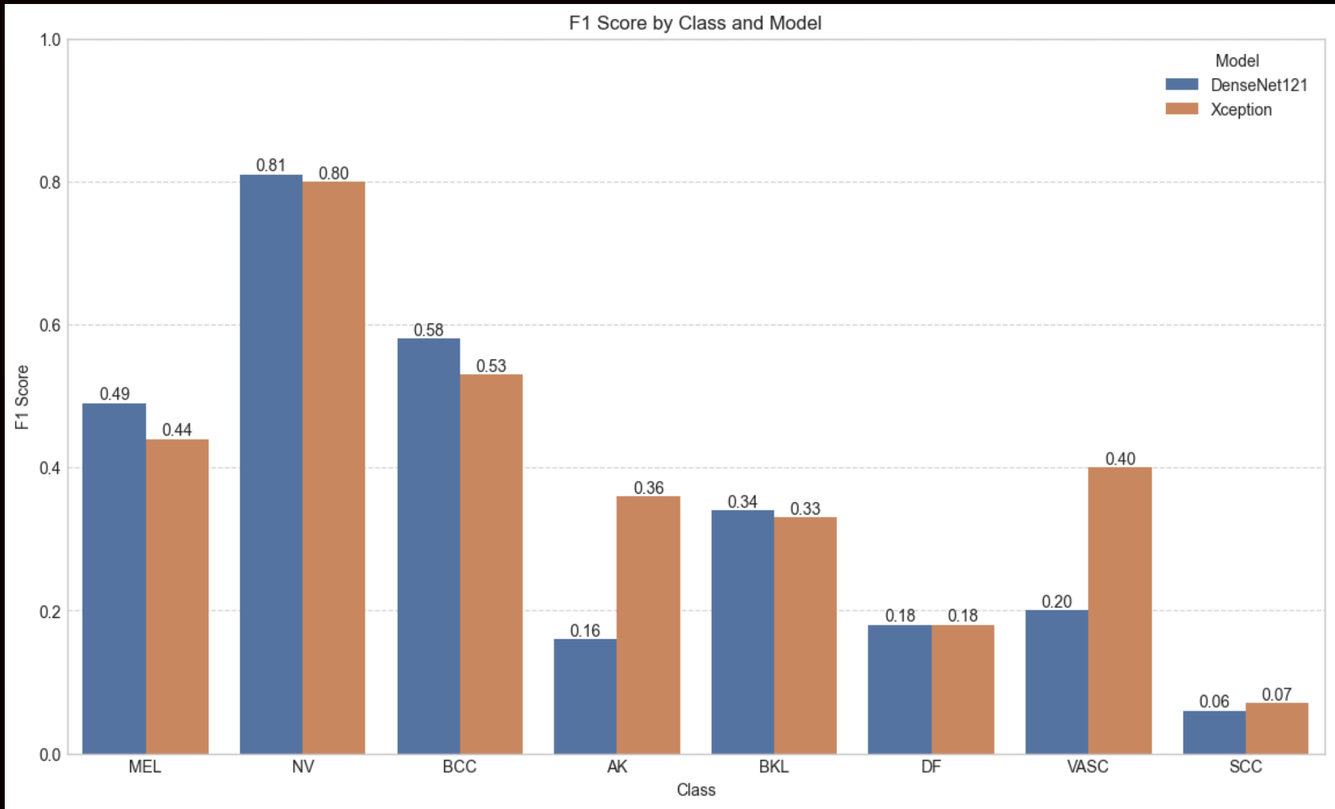
VALIDATION VS TEST PERFORMANCE

- DenseNet121: Training: 65.51% | Test: 14.4%
  - Xception: Training: 64.29% | Test: 13.11%
  - EfficientNetB0: Training: ~60%
- Significant Test-Train Gap: Especially notable in Xception (64.29% vs 13.11%) and DnseNet



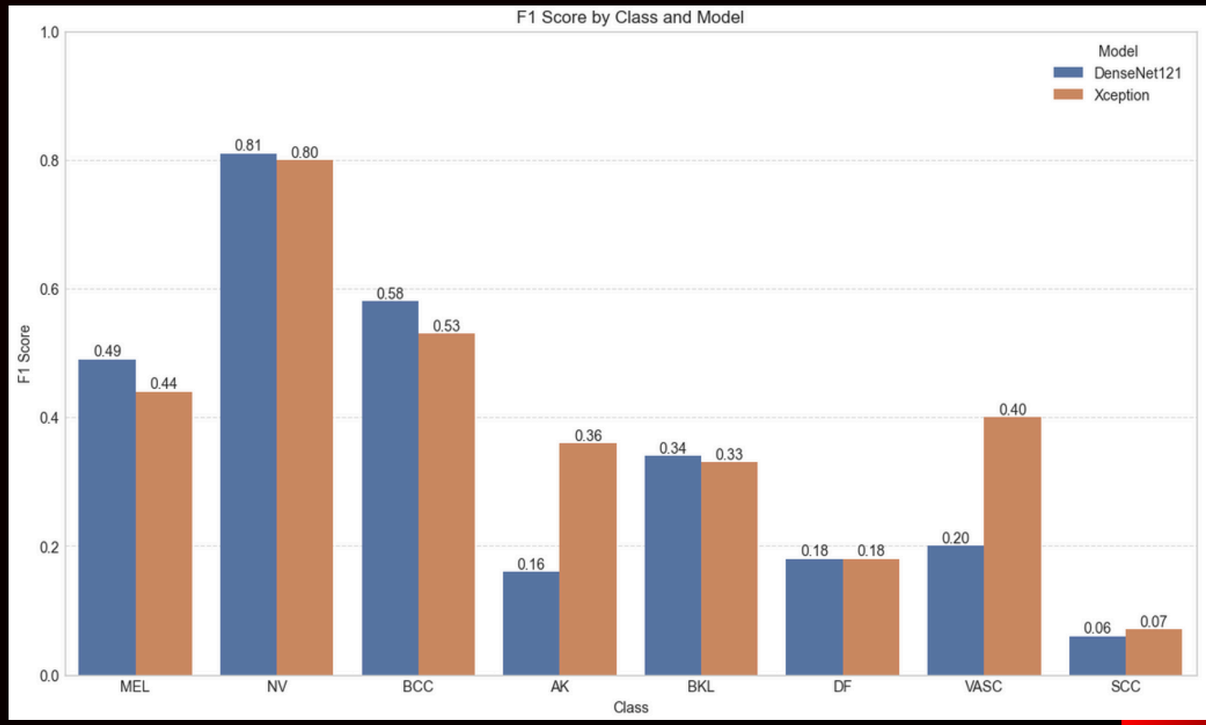
DENSENET121 ANALYSIS

- Class Performance (F1 Scores):
  - Best: NV (0.81), BCC (0.58), MEL (0.49)
  - Worst: AK (0.16), SCC (0.06)
- Test Prediction Distribution*:
  - DF: 62.2%, BKL: 14.3%, NV: 11.6%, BCC: 9.6%, Others: <1% each
- Comparison to Paper: Lower performance due to 8-class complexity vs. binary



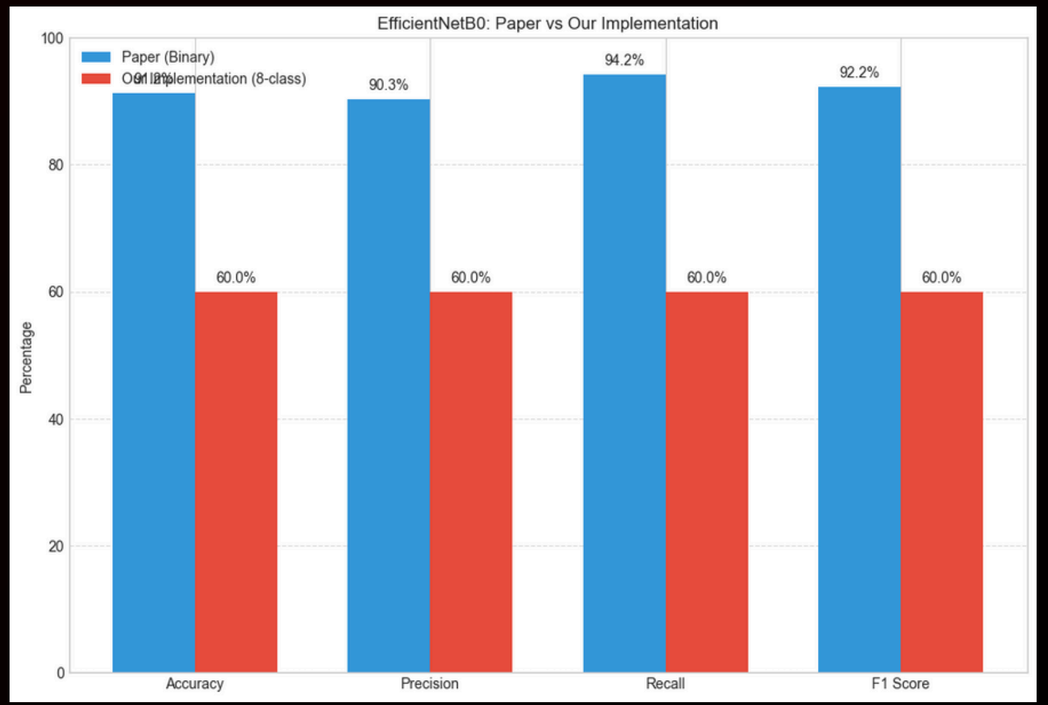
XCEPTION ANALYSIS

- Class Performance (F1 Scores):
  - Best: NV (0.80), BCC (0.53), MEL (0.44)
  - Worst: DF (0.18), SCC (0.07)
- Test Performance:
  - Test Accuracy: 13.11%, Precision: 12.46%, Recall: 13.11%,
  - Significant performance drop in test environment
- Class-specific test F1-scores: BCC (0.18), DF (0.19), MEL (0.14), low performance across all classes



EFFICIENTNETB0 ANALYSIS

- Direct Comparison with Paper:
  - Paper (Binary): 91.20% accuracy, 0.908 AUC, 94.20% recall, 90.33% precision, 92.22% F1
  - Our Implementation (8-class): ~60% validation accuracy (run locally on laptop)
- EfficientNetB0 Local Training Results:
  - Initial training: Epoch 0: 54.53% accuracy, validation: 60.53%
  - Epoch 2: 61.55% accuracy, validation: 60.94%
  - Limited epochs due to local laptop resource constraints
  - Results from efficientnet_b0_training_log.csv



CONCLUSIONS

- Performance Analysis:
  - Expected decrease in accuracy from binary to 8-class classification
  - DenseNet121 (65.51%) slightly outperformed Xception (64.29%) in training
  - Significant validation-test gap indicates generalization challenges
  - EfficientNetB0 showed promising initial results at ~60% accuracy even with limited training
- Key Learnings:
  - 8-class classification is significantly more challenging than binary
  - Models perform well on common classes (NV) but struggle with rare conditions
  - Model architecture choice impacts both performance and computational requirements

THANK YOU

FOR LISTENING