Advanced Artificial Intelligence Project

Fruit Recognition

USING DEEP LEARNING MODELS

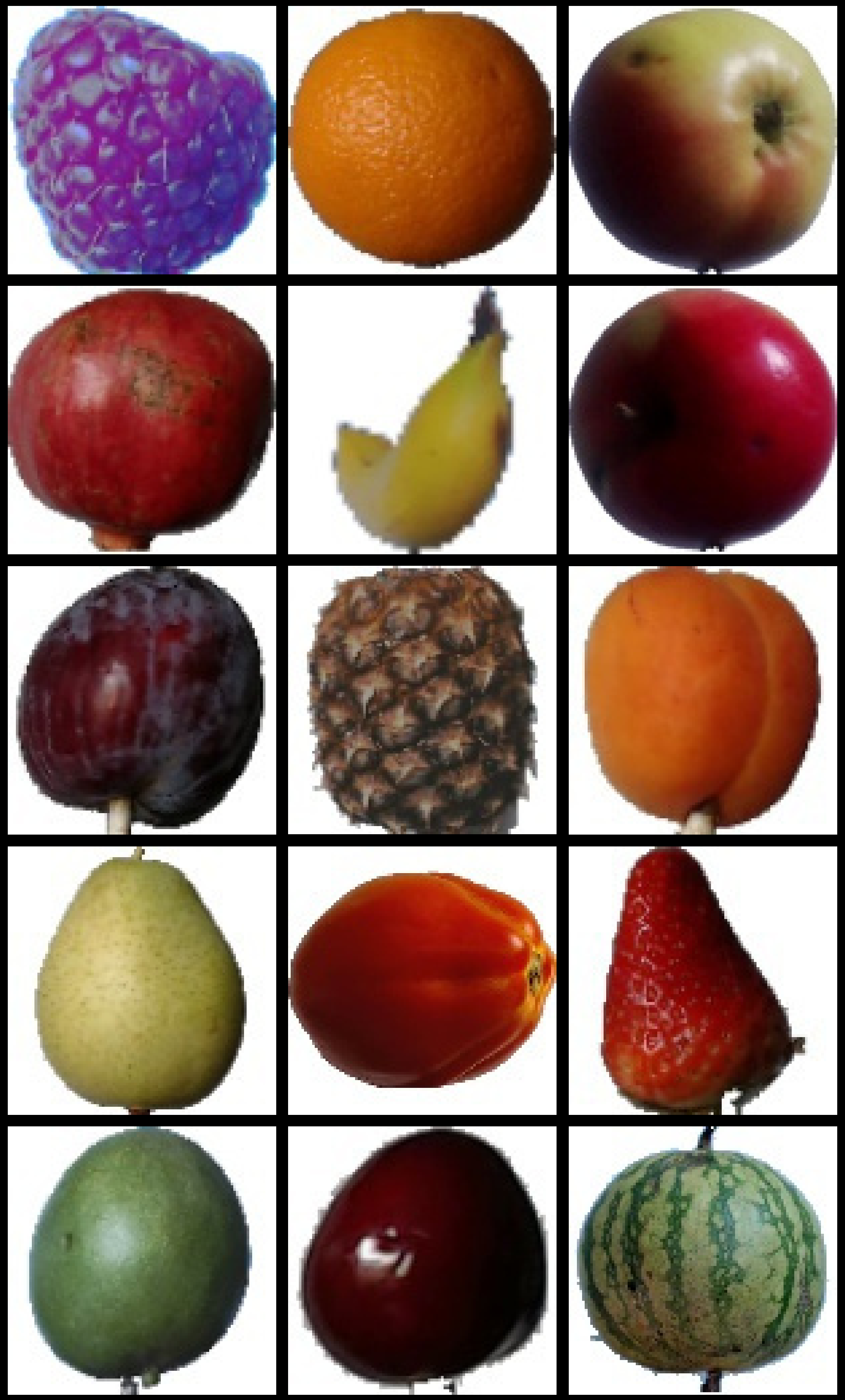
Jana Ghoniem – 2022/08912

DATASET

The dataset used was the fruits-360 by Mihai Oltean which consists of 137,104 images of 201 fruits, vegetables, nuts and seeds.

The dataset was reduced to 53,345 images belonging to 25 fruit classes for the training set, and 17,809 images for the test set.

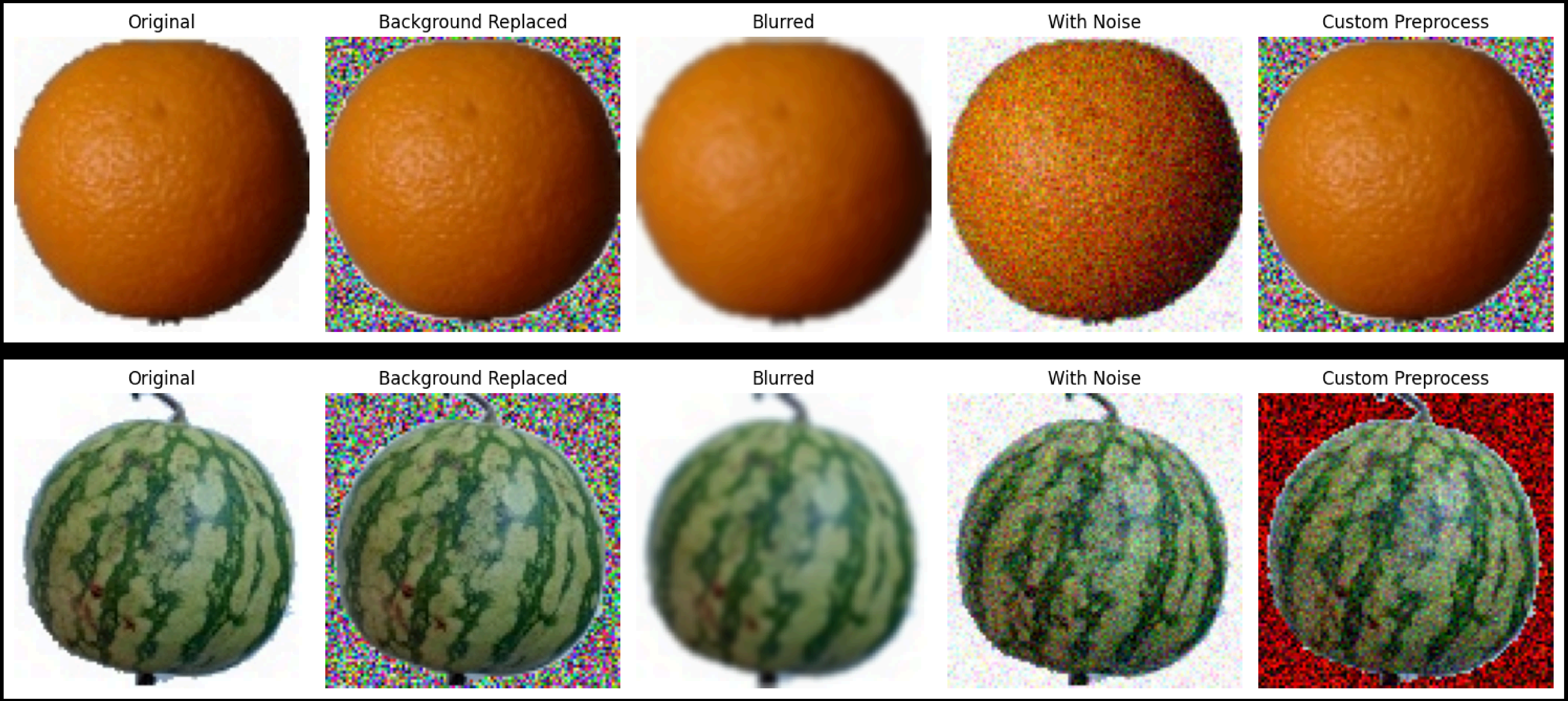
The issue with the original dataset is that all class images contain a white-only background with a single fruit in the foreground, leading to the models' inability to generalize when it comes to noisy backgrounds or images with multiple fruits.



Data Preprocessing & Augmentation

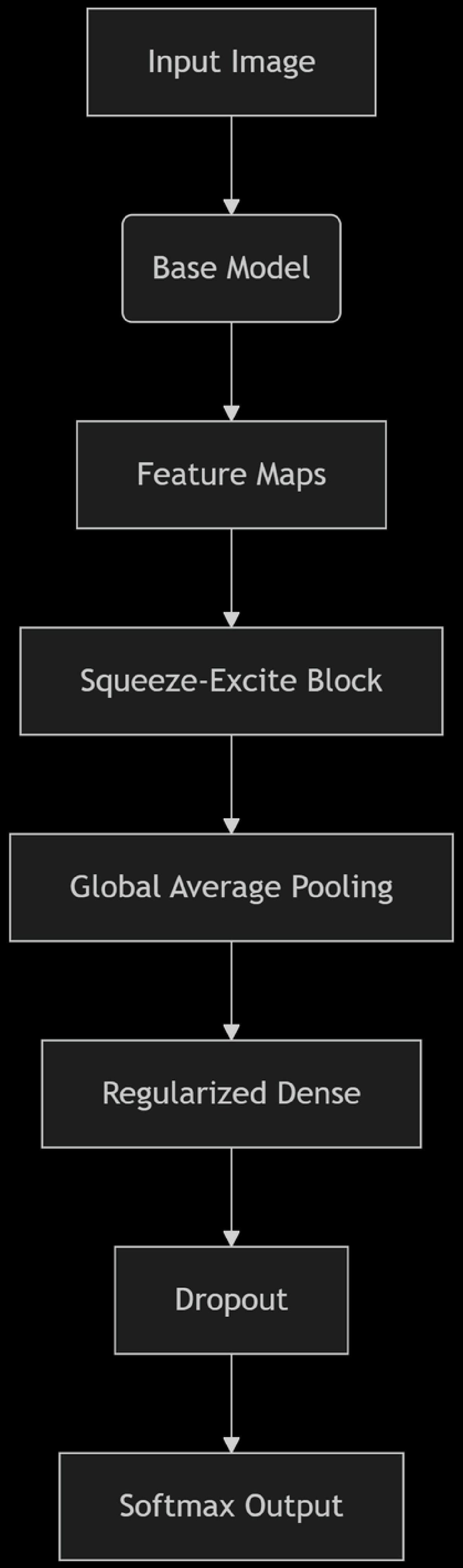
Preprocessing was performed in a customized fashion to simulate real-world variability. Each image underwent probabilistic augmentations such as: random background replacement (solid color or noise), Gaussian blur, and random noise addition. These augmentations were applied using OpenCV and NumPy, and then each model-specific preprocessing function (e.g., preprocess_input) was applied. The intention behind this approach is to train models that can generalize better to non-ideal environments, such as poor lighting or cluttered backgrounds.

DATA PREPROCESSING & AUGMENTATION



Model Architecture

Three CNN architectures were selected as backbones: MobileNetV2, EfficientNetB0, and ResNet50. These were initialized with pretrained ImageNet weights and fine-tuned by adding a custom classifier head. This head consists of a 2D convolutional layer, a Squeeze-and-Excitation (SE) Block for attention, dropout for regularization, and a dense output layer with softmax activation for multi-class classification. The SE Block plays a crucial role by adaptively recalibrating channel-wise feature responses, allowing the network to emphasize important information.


Input Image
Base Model
Feature Maps
Squeeze-Excite Block
Global Average Pooling
Regularized Dense
Dropout
Softmax Output

Model Training

Each model was trained using the Adam optimizer with a learning rate schedule controlled by ReduceLROnPlateau. Early stopping was applied to prevent overfitting, and model checkpoints were used to save the best-performing weights. Training was run for up to 15 epochs with a batch size of 32. The loss function was categorical cross-entropy, suitable for multi-class classification.

Class Imbalance

Unfortunately, though the dataset was extremely comprehensive, showcasing images of fruits from every angle with different illuminations, class “Apple” had thousands of images more than any other class, leading to class imbalance, and bias towards it. Overall performance was great, with almost all classes reaching a precision metric of 100%. The results of the training and evaluation on the unbalanced dataset are in the following slides.

MODEL EVALUATION (UNBALANCED)

Evaluation for MobileNetV2

	precision	recall	f1-score	support
apple	0.89	0.94	0.92	5348
apricot	0.62	0.70	0.66	164
banana	1.00	1.00	1.00	561
blackberrie	0.99	1.00	0.99	600
blueberry	0.95	1.00	0.97	154
cantaloupe	1.00	0.99	1.00	328
cherry	0.99	0.78	0.88	1373
dates	1.00	1.00	1.00	166
fig	0.97	0.97	0.97	234
grape	1.00	0.88	0.93	818
guava	0.96	0.98	0.97	166
kiwi	0.95	0.99	0.97	156
mango	0.64	0.70	0.67	308
orange	1.00	1.00	1.00	160
papaya	0.92	0.87	0.89	164
passion_fruit	1.00	0.43	0.60	166
peach	0.82	0.90	0.86	574
pear	0.85	0.96	0.90	1761
pineapple	1.00	1.00	1.00	329
plum	0.96	0.96	0.96	597
...				
accuracy			0.93	17809
macro avg	0.94	0.91	0.92	17809
weighted avg	0.93	0.93	0.93	17809

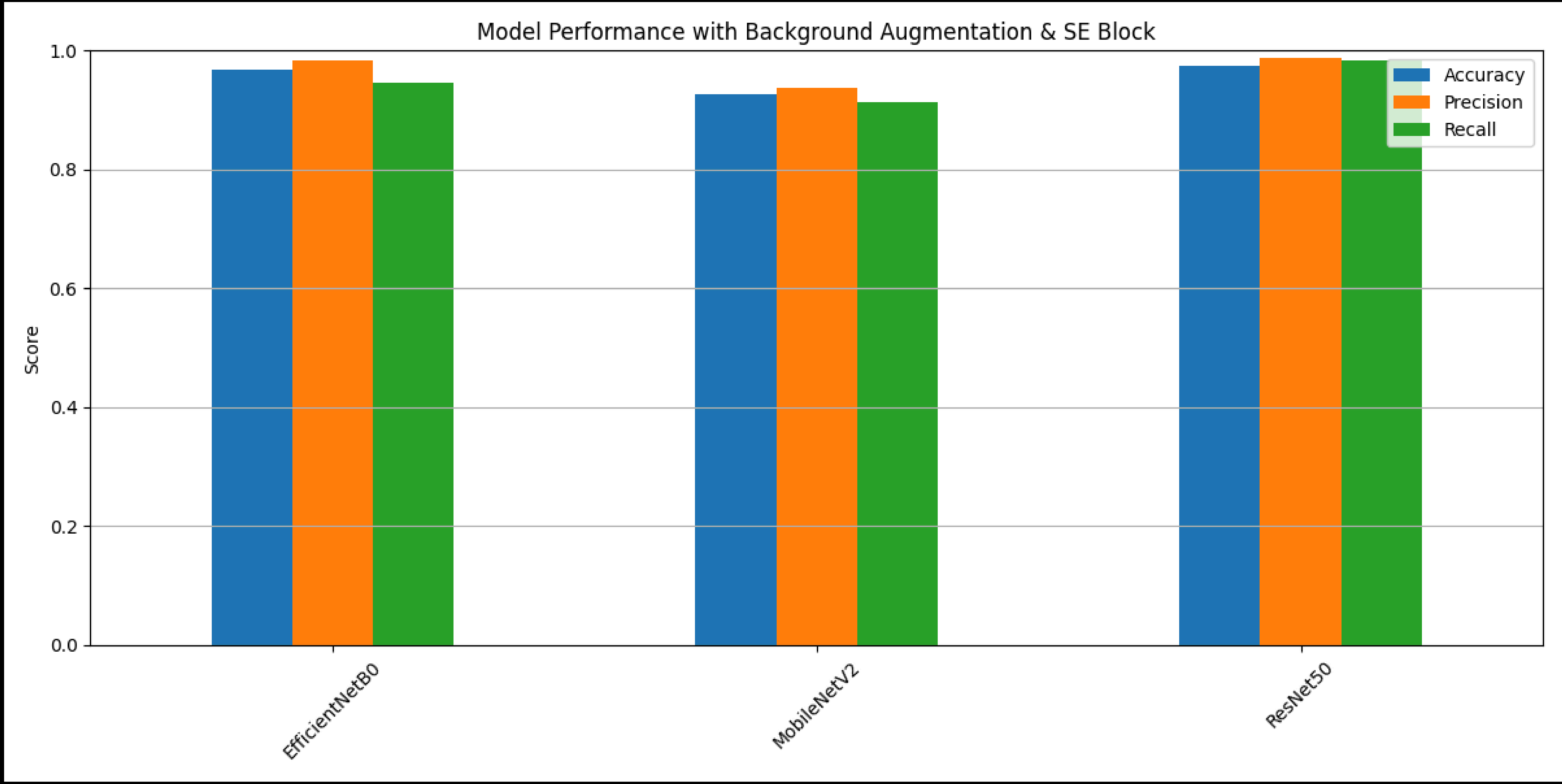
Evaluation for ResNet50

	precision	recall	f1-score	support
apple	1.00	0.96	0.98	5348
apricot	1.00	1.00	1.00	164
banana	1.00	1.00	1.00	561
blackberrie	1.00	1.00	1.00	600
blueberry	1.00	1.00	1.00	154
cantaloupe	1.00	1.00	1.00	328
cherry	0.94	1.00	0.97	1373
dates	1.00	1.00	1.00	166
fig	1.00	1.00	1.00	234
grape	1.00	0.99	1.00	818
guava	1.00	1.00	1.00	166
kiwi	1.00	1.00	1.00	156
mango	0.99	1.00	1.00	308
orange	1.00	1.00	1.00	160
papaya	1.00	0.90	0.95	164
passion_fruit	1.00	0.89	0.94	166
peach	0.96	0.86	0.91	574
pear	0.85	1.00	0.92	1761
pineapple	1.00	1.00	1.00	329
plum	0.96	1.00	0.98	597
pomegranate	1.00	1.00	1.00	164
raspberry	1.00	1.00	1.00	166
strawberry	1.00	1.00	1.00	410
tomato	1.00	0.97	0.98	2785
watermelon	1.00	1.00	1.00	157
accuracy			0.97	17809
macro avg	0.99	0.98	0.98	17809
weighted avg	0.98	0.97	0.98	17809

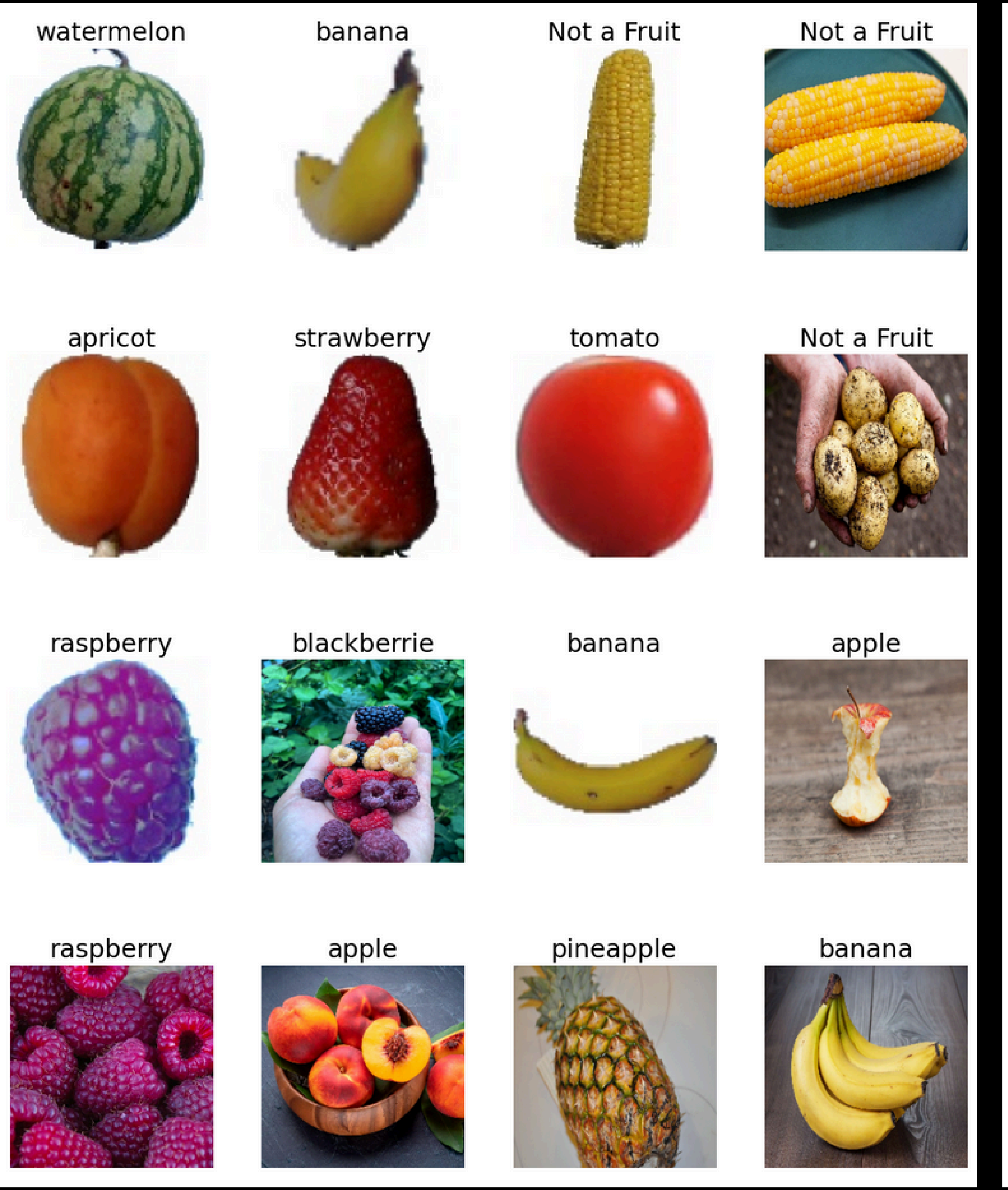
Evaluation for EfficientNetB0

	precision	recall	f1-score	support
apple	0.93	1.00	0.96	5348
apricot	1.00	0.90	0.95	164
banana	1.00	1.00	1.00	561
blackberrie	1.00	1.00	1.00	600
blueberry	1.00	1.00	1.00	154
cantaloupe	1.00	0.94	0.97	328
cherry	1.00	0.89	0.94	1373
dates	0.98	1.00	0.99	166
fig	1.00	1.00	1.00	234
grape	1.00	1.00	1.00	818
guava	1.00	1.00	1.00	166
kiwi	0.90	1.00	0.95	156
mango	1.00	0.96	0.98	308
orange	1.00	1.00	1.00	160
papaya	0.99	1.00	0.99	164
passion_fruit	1.00	0.30	0.46	166
peach	1.00	0.81	0.90	574
pear	1.00	0.92	0.96	1761
pineapple	1.00	1.00	1.00	329
plum	0.83	0.99	0.90	597
...				
accuracy			0.97	17809
macro avg	0.98	0.95	0.96	17809
weighted avg	0.97	0.97	0.97	17809

MODEL COMPARISON (UNBALANCED)



EFFICIENTNETB0 TESTING (UNBALANCED)



Class Balancing

In an attempt to cure the class imbalance, the majority classes were pruned of irrelevant or redundant images, and class weights were assigned to balance the data during the training process, enforcing penalties on the misclassification of the minority classes, forcing the model to focus on learning underrepresented patterns. Extra images were also added to the minority classes, in order to give the model the opportunity to differentiate between the classes. The results of the training and evaluation on the new balanced dataset are in the following slides.

MODEL EVALUATION (BALANCED)

Evaluation for MobileNetV2

	precision	recall	f1-score	support
0	0.91	0.71	0.80	3307
1	0.57	0.95	0.72	164
2	0.97	1.00	0.98	561
3	0.99	1.00	1.00	600
4	0.78	1.00	0.88	154
5	1.00	0.99	0.99	328
6	0.81	0.81	0.81	1373
7	1.00	0.90	0.95	166
8	0.72	0.99	0.83	234
9	0.92	0.97	0.94	818
10	0.92	1.00	0.96	166
11	0.74	1.00	0.85	156
12	0.54	0.83	0.65	308
13	0.65	1.00	0.78	160
14	0.40	0.93	0.56	164
15	0.78	0.87	0.83	574
16	0.87	0.86	0.87	1761
17	0.98	1.00	0.99	329
18	0.92	0.85	0.88	597
19	0.52	1.00	0.68	164
20	1.00	1.00	1.00	166
21	0.97	1.00	0.99	410
22	0.98	0.86	0.92	2632
23	0.88	1.00	0.93	157
accuracy			0.86	15449
macro avg	0.83	0.94	0.87	15449
weighted avg	0.89	0.86	0.87	15449

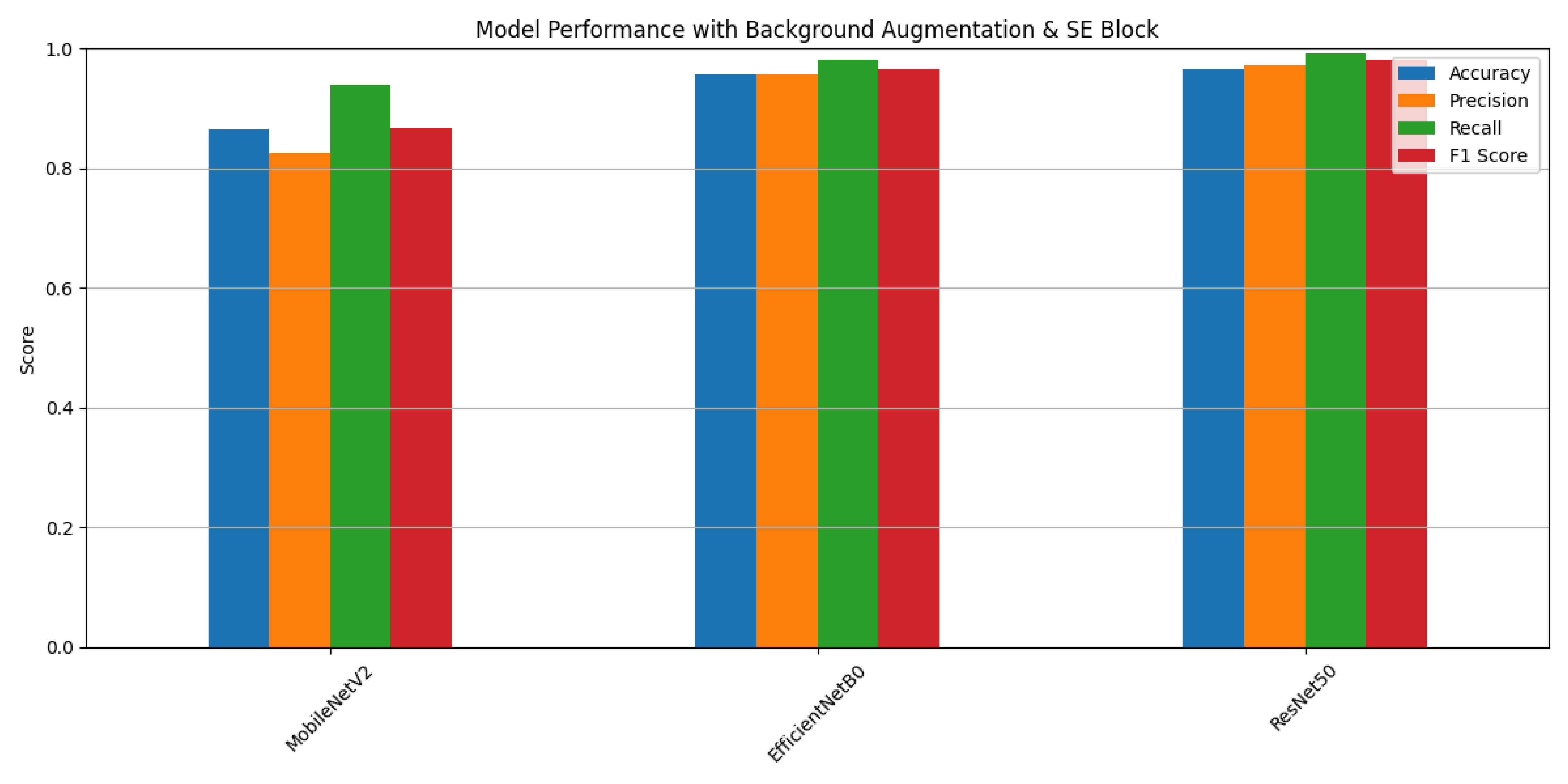
Evaluation for ResNet50

	precision	recall	f1-score	support
0	0.99	0.90	0.95	3307
1	1.00	1.00	1.00	164
2	1.00	1.00	1.00	561
3	1.00	1.00	1.00	600
4	1.00	1.00	1.00	154
5	1.00	1.00	1.00	328
6	0.88	1.00	0.93	1373
7	1.00	1.00	1.00	166
8	1.00	1.00	1.00	234
9	0.96	1.00	0.98	818
10	1.00	1.00	1.00	166
11	1.00	1.00	1.00	156
12	0.80	1.00	0.89	308
13	1.00	1.00	1.00	160
14	1.00	0.98	0.99	164
15	0.83	1.00	0.91	574
16	0.98	0.98	0.98	1761
17	1.00	1.00	1.00	329
18	0.99	1.00	1.00	597
19	0.91	1.00	0.95	164
20	1.00	1.00	1.00	166
21	1.00	1.00	1.00	410
22	1.00	0.94	0.97	2632
23	1.00	1.00	1.00	157
accuracy			0.97	15449
macro avg	0.97	0.99	0.98	15449
weighted avg	0.97	0.97	0.97	15449

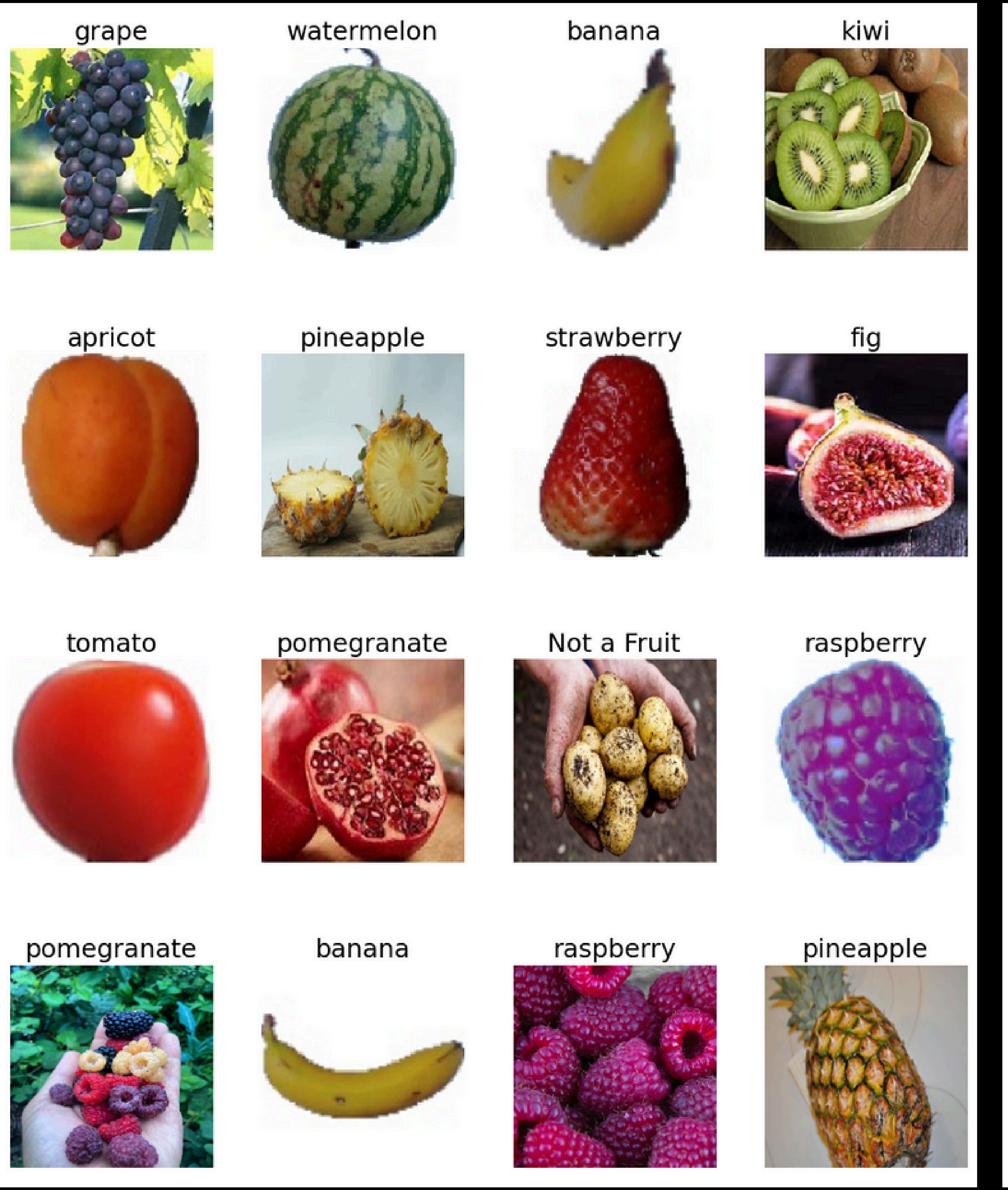
Evaluation for EfficientNetB0

	precision	recall	f1-score	support
0	0.91	0.93	0.92	3307
1	1.00	1.00	1.00	164
2	1.00	1.00	1.00	561
3	1.00	1.00	1.00	600
4	1.00	1.00	1.00	154
5	1.00	0.95	0.98	328
6	0.96	0.90	0.93	1373
7	0.99	1.00	1.00	166
8	0.97	1.00	0.99	234
9	1.00	1.00	1.00	818
10	1.00	1.00	1.00	166
11	0.93	1.00	0.96	156
12	0.92	1.00	0.96	308
13	1.00	1.00	1.00	160
14	0.76	1.00	0.86	164
15	0.87	0.93	0.89	574
16	1.00	0.87	0.93	1761
17	1.00	1.00	1.00	329
18	0.94	0.99	0.96	597
19	0.72	1.00	0.83	164
20	1.00	0.99	1.00	166
21	1.00	1.00	1.00	410
22	0.99	1.00	0.99	2632
23	1.00	1.00	1.00	157
accuracy			0.96	15449
macro avg	0.96	0.98	0.97	15449
weighted avg	0.96	0.96	0.96	15449

MODEL COMPARISON (BALANCED)



EFFICIENTNETB0 TESTING (BALANCED)



Conclusion

This project demonstrates that a combination of transfer learning, custom image augmentation, and attention mechanisms can yield high-performance models for fruit classification. EfficientNetB0, when combined with SE Block and background augmentation, achieved the highest scores across all metrics. These enhancements enabled the model to generalize better to challenging test samples. The results validate the importance of both data quality and architectural innovation in building robust deep learning systems.