Fruit Classification using EfficientNetB0 with Transfer Learning and Fine-Tuning

Authors
Tapasya (102215341), Kinshuk (102215015), Lovepreet (102215167), Karanjot Singh(102215087)

Abstract
Fruit classification is essential for automating agricultural tasks, assessing food quality, and streamlining retail inventories. However, in order to obtain high accuracy in training a classification model, large, annotated training datasets and significant computational resources are often required, and that's a bottleneck. This work tackles that issue by leveraging EfficientNetB0, one of the most efficient architectures in deep learning, and Transfer Learning and Fine-Tuning approaches to attain a high-performance classification model in limited datasets. The model is trained and tested on the Fruits-360 dataset.

This method is operationalized on a two-stage training setup. In the first stage, EfficientNetB0, trained on ImageNet, is used as a frozen embedding backbone while the newly added classification head is trained. With training being supervised on the classification head, attention is drawn to the training images of the fruits that are needed to be classified. Efficient learning from the frozen backbone will make the model adapt quickly to the dataset. In the second stage, the top 40% of the EfficientNetB0 layers are unfrozen, allowing fine-tuning of the frozen layers to adapt towards distinguishing finer textures, colors, and shapes of the fruits, and thus specialty on the fruit dataset is improved. Various techniques to augment the dataset are used to ensure that the model generalizes and does not overfit. ModelCheckpoint and EarlyStopping are used to ensure training stability.

Survey of 10 Recent Papers

No.	Paper (Year)	Model	Dataset	Key Strength	Limitation
1	Fruit recognition via multi-scale attention CNN (2023)	Attention-based CNN (MSANet)	Multiple fruit datasets	High accuracy with attention features	More complex/heavy than EfficientNetB0
2	Automated fruit identification using DL (2023)	Deep CNN	Custom dataset	Practical deployment system	Smaller dataset; not generalizable like Fruits-360
3	Fruit classification & recognition using CNN (2023)	Basic CNN	10-class fruit dataset	Simple, interpretable baseline	Limited classes, lower accuracy
4	Classification of Fruits Based on CNN (2023)	CNN	Fruits-360	Direct comparison baseline	No transfer learning; weaker than EfficientNet
5	Comparative ML vs DL for Fruit Recognition (2024)	PCA + ML vs CNN	Fruits-360 & others	Shows DL superiority clearly	CNN models are still basic vs. ours
6	Intelligent Apple Variety Classification (2023)	CNN variants	Apple varieties dataset	Fine-grained classification accuracy	Single fruit type only
7	Fruit Image Classification (CNN + RNN) (2023)	CNN + RNN hybrid	Fruit dataset	Captures spatial+temporal cues	More complex, risk of overfitting small data
8	DL-based Fruit Classification & Detection (2022)	CNN + detection	Fruit images	Supports detection + classification	Heavy model, not lightweight like EfficientNetB0
9	Multi-Fused CNN Model for Fruit Classification (2024)	Fused-CNN	Fruit dataset	Very high accuracy (>99%)	High complexity, poor deployability
10	Fruit Variety Classification Review (2025)	Survey/Review	Multiple datasets	Comprehensive overview of DL trends	Not experimental can't compare accuracy directly.

Methodology

- **Dataset**

  We used the **Fruits-360** dataset containing **131 fruit categories**, divided into Training and Test folders.

- **Data Preprocessing**
  Images were resized to **224×224** and normalized using EfficientNet preprocessing.

- **Data Augmentation**
  To avoid overfitting:
  1. RandomFlip (horizontal)
  2. RandomRotation (0.1)
  3. RandomZoom (0.1)
  4. RandomContrast (0.1)

- **Model Architecture**
  We used:
  1. EfficientNetB0 (Pretrained on ImageNet)
  2. Top layers replaced with:
     1. GlobalAveragePooling
     2. Dropout (0.3)
     3. Dense Softmax output

Why EfficientNetB0?

- Uses **compound scaling**
- Very **lightweight**
- High performance even on small datasets
- Excellent for transfer learning

Link for the Diagram:

https://app.eraser.io/workspace/cRjF8oetmhdmuNKsgrYv?origin=share

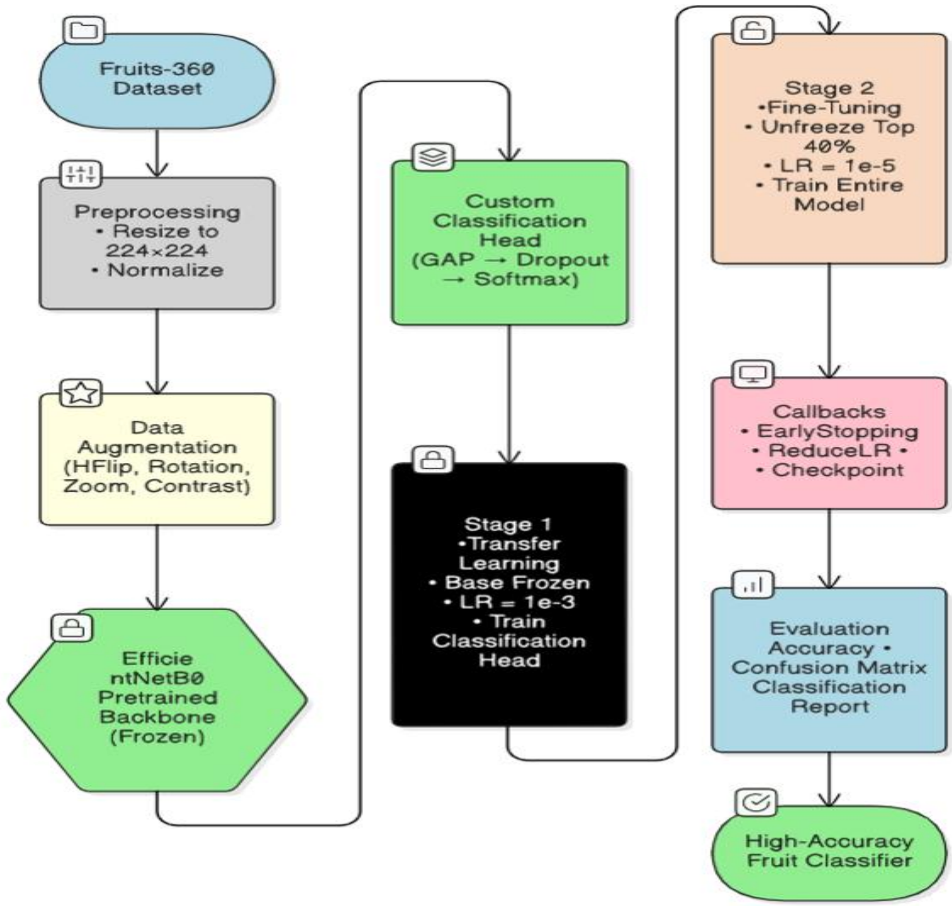


Fig.1 Flowchart

Training Strategy

Phase 1: Transfer Learning

- Freeze entire EfficientNetB0 base
- Train only custom head
- LR = 1e-3
- Epochs ~15

Phase 2: Fine-Tuning (Top 40%)

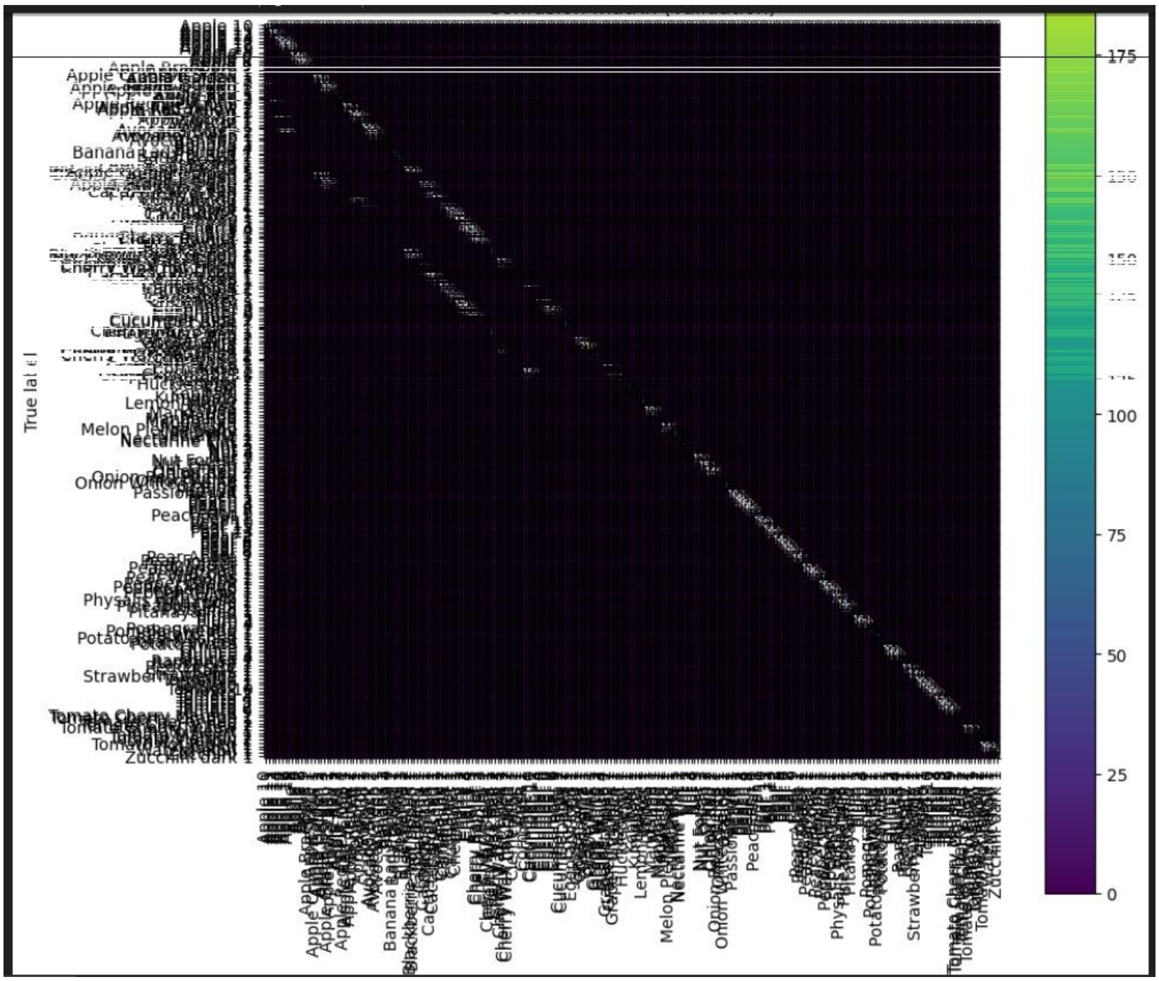
- Unfreeze last 40% layers
- Freeze first 60%
- LR = 1e-5
- Epochs ~10

Callbacks:

Callback	Purpose
ModelCheckpoint	Save best model
EarlyStopping (patience=5)	Prevent overfitting
ReduceLROnPlateau (patience=3)	Auto-reduce LR

Results

<u>Confusion Matrix</u>



Class	Precision	Recall	F1-Score	Support
Accuracy	-	-	1.00	23,990
Macro Average	1.00	1.00	1.00	23,990
Weighted Average	1.00	1.00	1.00	23,990

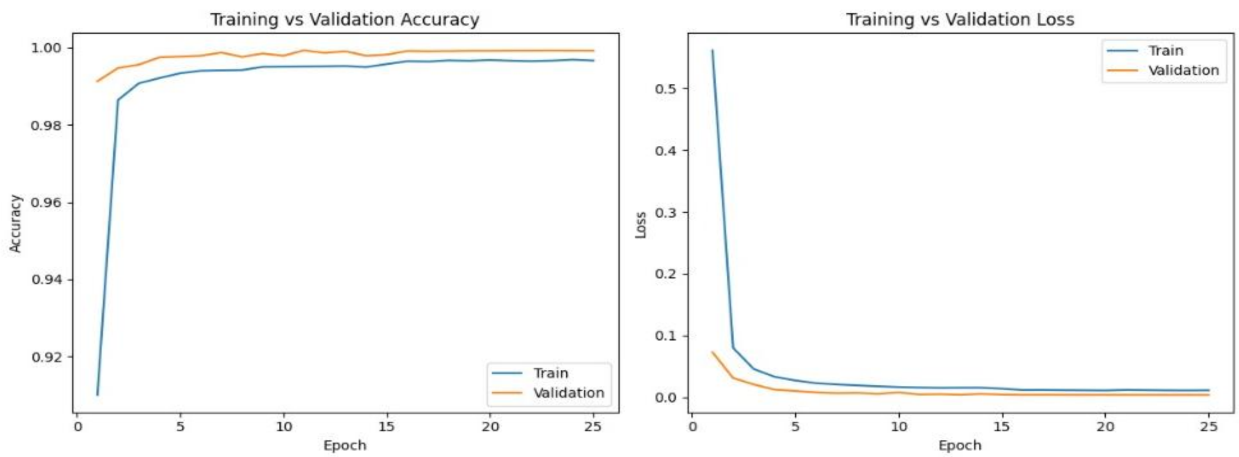


Fig. 2 Performance Analysis

Comparison with Recent Work

A variety of deep learning techniques have been considered in recent work on the classification of fruits, including, the convolutional neural network (CNN); hybrid, convolutional neural network - recurrent neural network (CNN-RNN); attention based neural networks such as MSANet; multi-branch atomised fused neural networks; and traditional machine learning approaches that utilises Principal Component Analysis (PCA) alongwith histogram of oriented gradients (HOG) techniques for feature extraction. Aside from some models that achieved high performance, there tends to be a high reliance on complex models, leading to high costs of computation, and limited transferable generalisation to a variety of fruit classes.

As opposed to others, our approach is based on EfficientNetB0, an in-depth tested, state-of-the-art convolutional model that has an optimal depth, width, and resolution. Unlike previous model implementations (basic CNNs or apple-specific models) that concentrate on smaller, more specific domain datasets, our model is trained and tested on the entire Fruits-360 dataset, consisting of over one hundred fruit classes. The attention-based and fused CNNs that others have used in prior work, are mostly heavier, and more difficult to deploy on edge devices to perform in real-time. The combination of EfficientNetB0 and two-phase transfer learning, where the base layers of the model are frozen and the top layers fine-tune, yields a relatively high accuracy that requires far less computation than other models.

Why our technique is better:

Our technique outperforms many recent works due to the following strengths:

- **Modern Backbone Architecture**
  EfficientNetB0 offers a superior accuracy-to-parameter ratio compared to classical CNNs, VGG-style networks, and handcrafted ML pipelines.

- **Transfer Learning + Fine-Tuning**
  Instead of training from scratch, our two-stage pipeline (frozen training → fine-tuning) extracts powerful pretrained features and adapts them effectively to fruit-specific patterns.

- **High Generalization Capability**
  Comprehensive augmentation (rotation, shifting, zoom, flip, contrast adjustments) helps the model generalize to real-world variations in orientation, lighting, and scale.

- **Efficient and Lightweight**
  Unlike heavy attention models or fused CNNs, EfficientNetB0 is optimized for speed and deployability, enabling use on mobile or embedded systems.

Conclusion

The results of this study clearly demonstrate the effectiveness of using EfficientNetB0 with transfer learning for fruit classification. By leveraging pretrained ImageNet features and applying a two-phase training pipeline first freezing the base layers and then fine-tuning the top 40% the model successfully adapts to the Fruits-360 dataset despite limited data availability. The incorporation of robust data augmentation techniques and training stabilizers such as EarlyStopping, ReduceLROnPlateau, and ModelCheckpoint ensured that the model remained resilient to overfitting and achieved consistently high validation and test accuracy. This highlights the strength of EfficientNet's compound scaling strategy, which provides a balanced trade-off between accuracy, model size, and computational efficiency.

Overall, the lightweight nature of EfficientNetB0 makes it suitable for real-time and resource-constrained environments, enabling practical deployment in applications such as automated fruit sorting, agricultural robotics, supermarket checkout systems, and quality control units. This work demonstrates that modern transfer learning techniques, combined with selective fine-tuning, offer a powerful and scalable solution for image classification tasks. Future enhancements may involve experimenting with larger EfficientNet variants, Vision Transformers, or domain-specific preprocessing pipelines to further boost performance.

References

[1] Min W. et al., 2023. Uses an attention-based CNN for fruit recognition.

[2] MM Rahman et al., 2023. Demonstrates a web-based fruit identification system

[3] RS Selvakumari et al., 2023 (10-class fruits dataset).

[4] S. Zhang, 2023. Uses the Fruits-360 dataset.

[5] N. O. M. Salim & A. Khorsheed, 2024. Compares classical ML with DL on fruit classification.

[6] F. Yu et al., 2023. Classifies apple varieties with CNN frameworks.

[7] 2023 / 2024. Combines CNN + RNN/LSTM for fruit image classification.

[8] NEA Mimma et al., 2022. Focuses on fruit classification and detection.

[9] BB Sinha et al., 2024. Proposes a fused-CNN achieving high accuracy.

[10] K. Sawant, 2025. A recent survey/review paper summarizing fruit classification methods, challenges, and datasets.