

RESEARCH ARTICLE

Multi-Fruit Classification and Grading Using a Same-Domain Transfer Learning Approach

LAMA A. ALDAKHIL¹ AND AESHAH A. ALMUTAIRI¹

Department of Information Technology, College of Computer, Qassim University, Buraydah 52571, Saudi Arabia

Corresponding author: Aeshah A. Almutairi (Aeshah.Almutairi@qu.edu.sa)

Researchers would like to thank the Deanship of Scientific Research, Qassim University, for funding publication of this project.

ABSTRACT The simultaneous classification and grading of fruits are essential yet underexplored facets of computer vision in agricultural automation. This study proposes the application of same-domain transfer learning using the EfficientNetV2 architecture to facilitate multi-fruit classification and grading. Our dual-model framework initially employs EfficientNetV2 to distinguish between six fruit types—bananas, apples, oranges, pomegranates, limes, and guavas—within the FruitNet dataset. Subsequently, the learned parameters are transferred to a second model, which focuses on grading the quality of the fruits. To address the class imbalance in the dataset, we incorporate a combination of AugMix, CutMix, and MixUp, significantly improving model generalization. Our experiments demonstrate robust performance, with classification and grading achieving an average test accuracy of 99%. These findings affirm the utility of same-domain transfer learning in enhancing grading accuracy using knowledge gained from classification tasks. The study shows promising potential for integrating this approach into machine vision systems to advance agricultural automation. Moving forward, this approach could be scaled to address broader cultivation challenges through the continued development of fine-grained visual analysis capabilities. The code is available on GitHub: MFCCG

INDEX TERMS Convolutional neural networks, deep learning, fruits, fruit classification, fruit quality grading, transfer learning.

I. INTRODUCTION

Fruits are indispensable for human nutrition, as they provide a rich source of vitamins and minerals. Traditionally, fruit quality assessment relies heavily on visual inspection to categorize the fruit type, freshness, and quality. Such assessments, performed by skilled personnel, involve evaluating multiple attributes, including color, shape, size, maturity, and cleanliness, to ensure that fruits meet market standards.

However, this practice has evolved significantly with technological advancements. Owing to the labor-intensive nature of manual inspection and increasing societal standards and expectations, the agriculture industry turned to computer vision [1], [2].

The introduction of machine vision systems in agriculture, specifically for automating tasks such as fruit sorting and

grading, has improved the production speed, efficiency, and accuracy of quality evaluation while simultaneously reducing overall production costs [1], [2], [3], [4]. These systems have been widely adopted in the agricultural industry for various tasks in pre-harvesting, harvesting, and post-harvesting fruits and vegetables [5].

At their core, these automation systems utilize machine learning algorithms that digitally interpret images to determine fruit characteristics and quality through accumulated experience, akin to human perception [3]. Deep learning (DL), an advanced subset of machine learning, now stands at the forefront of methods utilized for fruit classification and quality assessment, among other related tasks.

In recent years, convolutional neural network (CNN) methods, as one of the most popular image classification DL architectures, have been extensively researched for fruit classification and quality grading [2], [3], [4], [6]. Nevertheless, much of the research over the past decade has

The associate editor coordinating the review of this manuscript and approving it for publication was Jolanta Mizera-Pietraszko¹.

been confined to studying individual fruit types, [7], [8], which constrains the practicality of such methods in diverse agricultural and food industry settings where multiple fruit varieties are processed concurrently.

To address this gap, our study introduces a versatile method capable of classifying and grading multiple fruit types. We employed a transfer learning approach based on the EfficientNetV2 architecture [9], which operates in two stages. Initially, we trained a model to categorize the various fruits. Subsequently, we leveraged transfer learning to refine another model that utilizes the learned weights from the initial phase to grade fruit quality. This dual-stage approach presents an integrated solution for the automated classification and grading of multiple fruit types, aiming to enhance the scalability and utility of machine vision in agriculture.

A. MOTIVATION

Research in multi-fruit classification and grading has rarely addressed the simultaneous handling of both tasks for multiple fruit types. Some studies have proposed automated systems for multi-fruit classification without grading fruit [10], while others have focused on evaluating the grade of various fruits without preceding classification [11]. This study hypothesizes the effective use of same-domain transfer learning by training a CNN model to classify multiple fruits and then employing the learned knowledge to inform a subsequent grading model based on quality. This sequential dual-task approach is anticipated to streamline the classification and grading of fruits, enhancing the accuracy and efficiency of these processes. The proposed method also has the potential to be seamlessly integrated into existing machine vision systems, thereby improving their overall functionality.

B. CONTRIBUTIONS

The proposed work contributes to the current research domain in the following ways:

- 1) An innovative application of same-domain transfer learning is introduced, utilizing a DL approach to transfer knowledge from fruit type classification to fruit quality assessment.
- 2) The methodology tests the efficiency of the EfficientNetV2 DL architecture, emphasizing training speed and reduced computational costs [9].
- 3) It evaluates the performance of the FruitNet [12] dataset in fruit classification tasks.
- 4) Advanced data augmentation techniques, specifically CutMix [13], and MixUp [14] are utilized to improve the model's robustness and generalization in classifying fruit types and determining fruit quality.

Collectively, these contributions aim to enhance the capabilities of machine vision systems in accurately classifying and grading multiple fruit varieties, pushing towards a more integrated and efficient approach to agricultural automation.

C. RESEARCH ORGANIZATION

The remainder of this study is structured as follows: Section II reviews the relevant literature in the field. Section III describes the methodology adopted for this study. Section IV discusses the experimental results and provides a detailed analysis. Section V concludes the study with a summary of the findings and their implications for future research.

II. RELATED WORK

The classification of fruits based on their types, quality, freshness, and other visual attributes using DL approaches is a popular research area. Various DL models have been experimented with for different classification tasks of specific fruit types [7], [8], while others considered multiple fruits [10], [11], [15], [16]. Among the studies focused on one type of fruit is Raikar et al. [7], which proposed an automated CNN method for classifying and grading okra by size. Their study utilized a dataset comprising 3,200 samples, divided into four size classes (small, medium, large, and extra-large) for grading purposes. The experimental findings revealed that ResNet50 achieved an accuracy of 99.17%, outperforming AlexNet and GoogLeNet.

Iqbal and Hakim [8] conducted a study focusing on mangoes. They developed an automated system based on CNNs to grade images of classified mango cultivars. Two manually collected datasets were used: one for classification, containing images from eight mango cultivars, and another for grading, differentiated into three classes based on quality and defects. The results indicated that InceptionV3 surpassed VGG16 and ResNet152 in performance, achieving an accuracy of 99% for classification and 96% for grading.

Albarrak et al. [17] concentrated on classifying date fruits, proposing a CNN-based method utilizing the MobileNetV2 architecture chosen for its efficiency and mobile application compatibility. They fine-tuned the baseline network by modifying the last layer to include five layers: average pooling, flatten, dense, dropout, and softmax. This method achieved an accuracy of 99% in classifying eight different date fruit types, illustrating its superiority over other popular architectures with the specific dataset used.

Two studies among those reviewed focused on apples and bananas. Ismail and Malik [18] developed a cost-effective computer vision system based on DL to grade apples and bananas by defects and ripeness. The system used a Raspberry Pi module with a camera and a user interface for real-time fruit inspection. Image processing techniques such as Gaussian filtering and histogram equalization were applied to enhance image quality. During training and testing, mean shift clustering and watershed segmentation were employed for image segmentation. To compensate for the limited banana dataset, domain transfer learning with pre-trained weights from the apple dataset was utilized. EfficientNet outperformed other tested DL architectures, achieving average accuracies of 99.2% for apples and 98.6% for bananas, respectively.

Similarly, Knott et al. [16] proposed a different approach for the same task, utilizing an off-the-shelf, self-supervised vision transformer known as Self-Distillation with No Labels Vision Transformer (DINO ViT). The experimental outcomes indicated that the proposed model's accuracies were on par with CNN models, confirming its effective performance in domains with limited data.

Other studies adopted a broader scope, including a variety of fruits. Hossain et al. [10] suggested a DL approach for the automated classification of 25 varieties of fruits and vegetables using a lightweight CNN model with six layers and a fine-tuned VGG16 model. Two datasets were employed: The Supermarket Produce dataset and a self-collected dataset. The VGG16 model achieved an impressive accuracy of 99.75% on the Supermarket Produce dataset and 96.75% on the self-collected dataset.

Adopting an alternative approach, Nasir et al. [15] aimed to classify multiple fruits and plant diseases by introducing an automated classification system leveraging 5G and cloud technologies. The system combined a hybrid model for feature extraction, utilizing features from a pre-trained VGG16 and a Pyramid Histogram of Oriented Gradient (PHOG) technique with the minimum redundancy maximum relevance (mRMR) method. The best classification results were achieved using Cubic Support Vector Machines (SVM), with accuracies of 99.6% on the Fruits-360 dataset and 98.16% on the Plant Village Disease dataset.

Dhiman et al. [11] also focused on multi-fruit assessment, presenting an automated system for quality assessment using recurrent neural networks (RNNs). The study included 200 high-quality fruit images from the FIDS30 dataset and 200 self-collected images of poor-quality fruits, encompassing nine types of fruits. The system employed image pre-processing techniques such as Contrast-Limited Adaptive Histogram Equalization (CLAHE) and canny edge detection for segmentation. Principal Component Analysis (PCA) was used for feature extraction, and RNNs were used for classification. The results demonstrated that the RNN achieved an average accuracy of 98.47%, proving its capability to assess fruit quality accurately.

The studies reviewed highlight the significant potential of DL approaches in a wide range of fruit classification tasks. A summary of the related works is presented in Table 1. Our study aims to build upon the existing multi-fruit classification and grading research by proposing a same-domain transfer learning framework.

III. METHODS AND MATERIALS

A. DATASET PREPARATION

We employed the publicly accessible FruitNet dataset for model training and evaluation, as presented in [12], which can be accessed at this link [19]. This dataset comprises over 14,700 high-resolution images across six fruit classes: apples, bananas, pomegranates, guavas, oranges, and limes. Each class is categorized into good, bad, and mixed quality,

resulting in a structured dataset suited to a multi-task learning approach. The images within the dataset were captured using a mobile phone equipped with a high-resolution camera under various backgrounds and lighting conditions. All images in the dataset have dimensions of 256×256 pixels.

We reorganized the images into type-specific folders to tailor the dataset to our needs, with each fruit divided into quality-based sub-categories. We partitioned the data into training (70%), validation (20%), and testing sets (10%) using Python scripts.

B. DATA PRE-PROCESSING AND AUGMENTATION

In machine learning applications, the performance of models is heavily dependent on the quality and configuration of the dataset. In the domain of multi-fruit classification and grading, this necessitates rigorous pre-processing and augmentation of the data to ensure a high level of model accuracy and stability.

The initial analysis of the class distribution within the training dataset, as depicted in Fig. 1, indicated an over-representation of the pomegranate class. This is attributed to prior oversampling techniques employed by the dataset creators. A two-step pre-processing strategy was adopted to counteract this and diminish the risk of model overfitting.

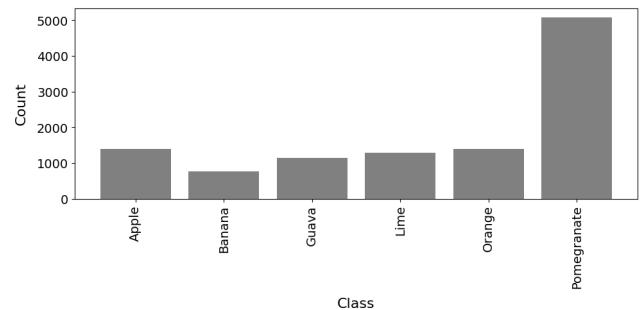


FIGURE 1. FruitNet class distribution.

The first step of this strategy involved augmenting under-represented classes using the AugMix [20] pipeline, which was selected for its effectiveness in introducing variability and enhancing model robustness against ambiguous data inputs. The AugMix method chains and mixes up simple augmentation operations by layering them stochastically to produce highly diverse synthetic samples. This method is used to balance the class distribution and promote consistency in model performance, implementing the Jensen-Shannon divergence to maintain the model's output reliability across augmented and original data.

The second step involved applying several augmentations to normalize the training data. This included resizing all images to 224×224 pixels to comply with the baseline model input size, implementing standard affine transformations such as random flips and rotations up to 45 degrees, and adding Gaussian blur. To optimize the data for training, a sequence of transformations was performed: image tensor conversion, type casting to 8-bit unsigned integers, and then

TABLE 1. Summary of related work.

Ref	Year	Fruit types	Dataset	Method	Accuracy	Contribution
[10]	2019	25 varieties of fruits and vegetables	Supermarket Produce dataset and a self-collected dataset	VGG16	99.75% using the Supermarket Produce dataset and 96.75% using the self-collected dataset.	Fruit and vegetable classification
[7]	2020	Okra	self-collected dataset	ResNet50, AlexNet, GoogLeNet	99.17% using ResNet50	Classifying and grading okra based on size
[15]	2020	131 varieties of fruits and vegetables	Fruits-360 and Plant Village Disease	VGG16	99.6% using the Fruits-360 dataset and 98.16% using the Plant Village Disease dataset.	Fruit type and disease classification
[18]	2020	Apples and bananas	CASC IFW dataset for apples and the banana ripeness dataset	ResNet, DenseNet, EfficientNet, NASNET, MobileNetV2	99.2% for the apple dataset and 98.6% for the banana dataset	Grading apples and bananas according to defects and ripeness
[11]	2021	Apples, bananas, guavas, pomegranates, tomatoes, grapes, mangoes, pears, and oranges	FIDS30 dataset and self-collected dataset	RNN	An average accuracy of 98.47%	Multi-fruit quality assessment
[8]	2022	Mango cultivars	Self-collected datasets	InceptionV3, VGG16, ResNet152	99% for classification and 96% for grading using the InceptionV3 model	Grading and classifying mango cultivars
[17]	2022	Dates	Self-collected dataset	MobileNetV2	99%	Classifying date fruit types
[16]	2023	Apples and bananas	CASC IFW dataset for apples and the banana ripeness dataset	DINO ViT	94.1% for the banana dataset and 95% for the apple dataset	Grading apples and bananas according to defects and ripeness

to 32-bit floating points, concluding with normalization based on a set mean and standard deviation.

Two data augmentations are randomly applied in the model training pipeline: the CutMix and the MixUp methods. Both of these methods aim to enhance the robustness and generalization of the proposed approach.

- 1) CutMix [13]: CutMix is an augmentation technique that involves cutting a patch from one image and pasting it onto another training image (x). At the same time, the ground truth labels (y) are also mixed in proportion to the area of the patches. Given two random samples (x_i, y_i) and (x_j, y_j) and a binary mask M , where regions from (x_j) are cut and pasted onto (x_i), the CutMix augmentation creates a synthetic sample (x', y') as follows:

$$x' = M \odot x_i + (1 - M) \odot x_j$$

$$y' = \lambda y_i + (1 - \lambda) y_j$$

- 2) MixUp [14]: MixUp is a simple augmentation technique that constructs synthetic training samples by

linearly interpolating between a pair of samples (x) and their labels (y) from a training set. Given two random samples (x_i, y_i) and (x_j, y_j), and mixing ratio $\lambda \in [0, 1]$, the MixUp augmentation creates a synthetic sample (x', y') as follows:

$$x' = \lambda x_i + (1 - \lambda) x_j$$

$$y' = \lambda y_i + (1 - \lambda) y_j$$

A separate pre-processing pipeline was used for the validation and test sets. This pipeline applied only the necessary augmentations to load the data, avoiding additional variations. Thus, it ensured that the data closely reflected real-world conditions without artificial alteration.

C. OVERVIEW OF THE EFFICIENTNETV2 ARCHITECTURE

The advanced EfficientNetV2 architecture [9] was selected for its superior performance in CNNs. EfficientNetV2 introduces compound scaling, which uniformly scales network width, depth, and image resolution, enhancing the model's ability to differentiate fine-grained features.

EfficientNetV2 is characterized by its:

- 1) Progressive learning capability that dynamically adjusts regularization during training.
- 2) Usage of Fused-MBConv blocks in early layers for faster optimization.
- 3) A novel scaling rule to optimize resource allocation by limiting input resolution.

By incorporating the characteristics mentioned, EfficientNetV2 achieves state-of-the-art performance over its predecessor, EfficientNet. For this reason, this network architecture was chosen as the primary model for conducting the experimentations in this study.

This architecture outperforms its predecessor, EfficientNet, particularly in resource-constrained environments, making it well-suited for the experimental framework of this study. The architecture comprises a combination of mobile inverted bottleneck convolution (MBConv) blocks and Fused-MBConv blocks, as shown in Fig. 2. The MBConv block, initially introduced in [21], optimizes neural network efficiency via an inverted bottleneck design, expanding and compressing channels to balance expressiveness and computational cost. The Fused-MBConv block, proposed in [22], streamlines this process further by merging layers for enhanced performance. The structures of the MBConv and Fused-MBConv blocks are displayed in Fig. 3. Consequently, EfficientNetV2 leverages these blocks to increase training speed without compromising accuracy [9].

D. TRANSFER LEARNING

Transfer learning is inspired by the human brain's innate ability to apply knowledge across related problems. When machine learning models encounter new data that differs somewhat from prior training, transfer learning can help bridge those gaps. By using what was learned from a source domain, models can more efficiently learn patterns in a target domain with less labeled data [23]. The process of transferring knowledge from task to task is illustrated in Fig. 4.

This study implements transfer learning through the use of pre-trained EfficientNetV2 models. Initially trained on the expansive ImageNet dataset, these models possess pre-learned weights that allow them to recognize a wide array of visual features from natural images.

Moreover, a homogeneous transfer learning strategy was employed, linking fruit classification to fruit quality assessment tasks. Despite the inherent variability due to factors like ripeness, the commonality of processing fruit images allows for transferring learned features from one task domain *classification* to another *quality assessment*.

Transfer learning significantly reduces the necessity for large labeled datasets by utilizing pre-existing knowledge, which is advantageous in machine learning [23]. This method reflects continual human learning, where new knowledge is assimilated and applied in conjunction with prior experiences.

E. EXPERIMENTAL SETUP

The experimental setup was conducted on Google Colaboratory using a Colab Pro subscription. This cloud-based service provides an interactive Jupyter Notebook environment, with computational resources including a T4 GPU, 12.7 GB of RAM, and 166.8 GB of disk space. The dataset was integrated through Google Drive for seamless access and management. The Python libraries utilized throughout the research are enumerated in Table 2, with the training hyperparameters detailed in Table 3.

To evaluate the proposed model, the following evaluation metrics were used:

- 1) Accuracy: This metric measures the proportion of correctly classified instances from all instances in the dataset.
- 2) Precision: This metric is the ratio of true positives to the model's total number of positive predictions.
- 3) Recall: This metric is the ratio of true positives to the total number of positive cases.
- 4) F1 score: This metric is the harmonic mean of precision and recall.

Equations (1), (2), (3) and (4) calculate accuracy, precision, recall, and F1-score, respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1 score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

TP= True Positives, TN= True Negatives, FP= False Positives, and FN= False Negatives.

F. PROPOSED FRAMEWORK

This study introduces a robust framework leveraging transfer learning for multi-fruit classification and grading tasks. The framework integrates the previously discussed methodologies, ensuring a systematic approach to these interrelated tasks. Fig. 5 illustrates the workflow of the proposed framework.

Initially, the FruitNet dataset undergoes several preparative and preprocessing steps, including dataset directory reorganization, partitioning, class distribution analysis, and balancing through AugMix augmentation. The preprocessed images are then fed into the subsequent transfer learning models.

Employing the EfficientNetV2-S as a base, the framework utilizes two sequential models: the first for fruit classification and the second for fruit quality grading, with the latter inheriting the weights from the former. The training pipeline integrates CutMix and MixUp augmentations, which are applied randomly to image batches, to improve feature robustness.

The proposed dual-task framework is distinctive for its application of same-domain transfer learning, which is

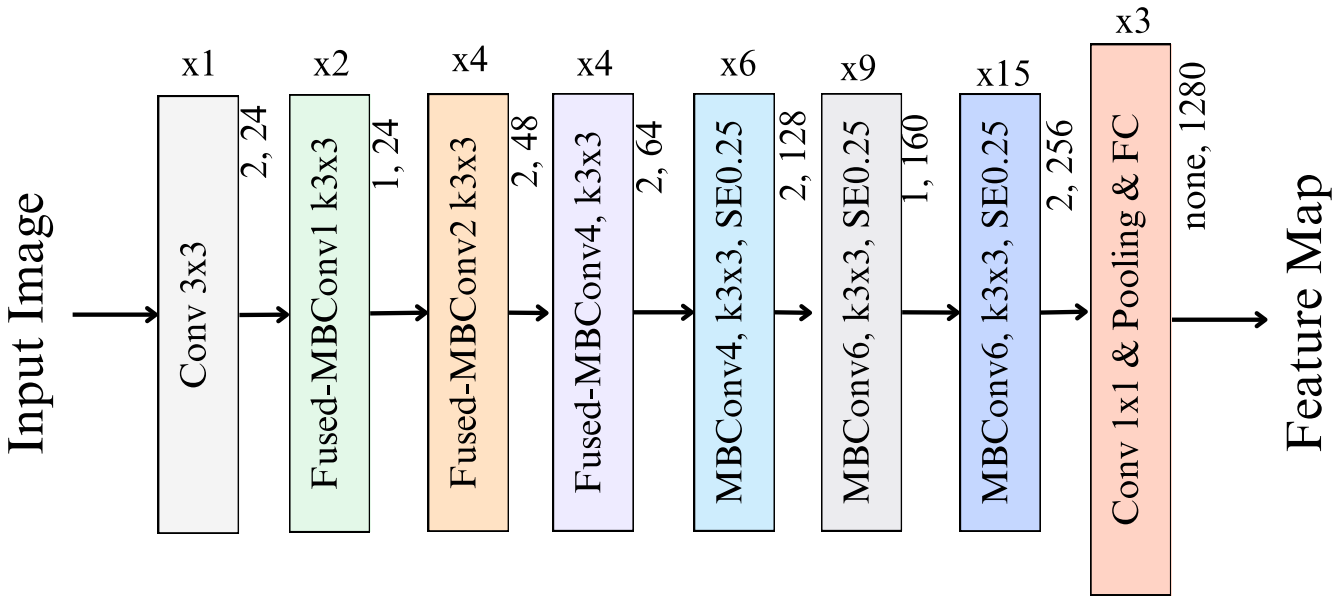


FIGURE 2. Architecture of EfficientNetV2 [9].

TABLE 2. Python libraries.

Library Name	Functions Used
os	Accessing files and making directories.
Time	Handling time-related tasks within code.
Torch	Creating the dataset classes and data loaders, loading the optimizer, criterion, and scheduler for model training, and logging model progress using tensorboard.
Torchvision	Loading data augmentations using V2 transforms and loading a pre-trained EfficientNetV2-S model.
PIL	Opening and manipulating images in the dataset.
Collections	Count the frequency of each class in the dataset and find the class with the maximum count.
Matplotlib	Plotting images and confusion matrix
Datetime	Creating directory names based on real dates and the time for logging model runs.
Numpy	Converting data types and performing mathematical operations.
Sklearn	Computing evaluation metrics.
Seaborn	Visualizing confusion matrix.
Streamlit	Creating the inference application.
gdown	Download files from Google Drive

TABLE 3. Training hyperparameters.

Hyperparameter	Value	Description
Learning Rate	0.01	The step size at each iteration while moving toward a minimum of a loss function.
Weight Decay	1e-2	Regularization technique applied to the weights of a model to reduce overfitting.
Batch Size	64	The number of samples loaded per training iteration.
Number of Epochs	30	The number of passes through the whole training dataset.
Loss Function	CrossEntropy-Loss	A function that measures the performance of a classification model whose output is a probability value between 0 and 1.
Optimizer	AdamW	A variant of Adam optimizer with more accurate weight decay regularization.
Learning Rate Scheduler	ReduceLROn-Pplateau	A scheduler that dynamically reduces the learning rate when learning stops improving according to a specified metric.

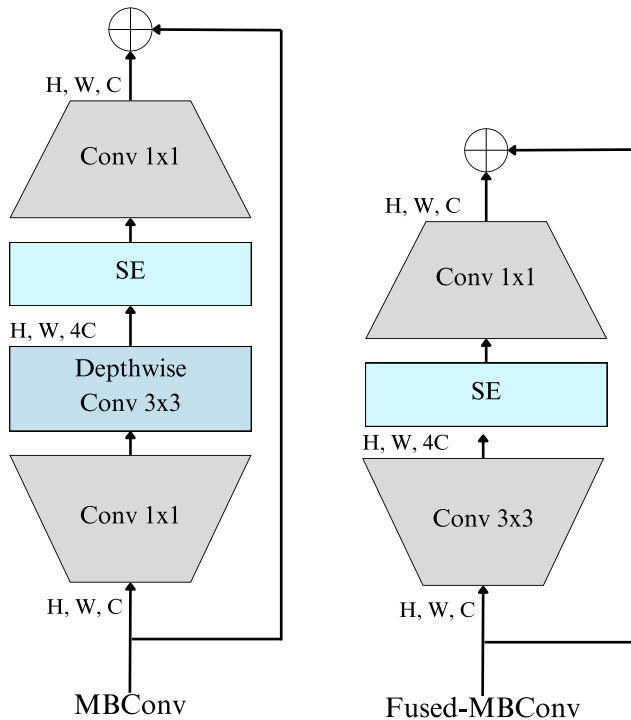


FIGURE 3. Structures of MBConv and Fused-MBConv [9], [21], [22].

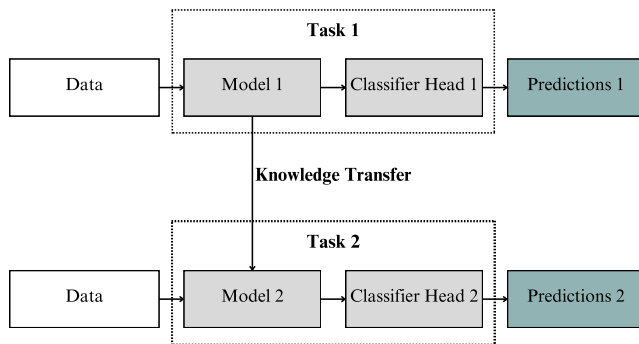


FIGURE 4. Knowledge transfer in transfer learning from one task to another.

enhanced using CutMix and MixUp. This approach is designed to improve fruit type classification and quality assessment accuracy.

The framework finalizes with the deployment of a Streamlit application, which provides real-time predictions for fruit type and quality, demonstrating the practical application of the research findings.

In essence, the framework delineated herein offers a practical and comprehensive solution for classifying and grading various fruit types. It showcases the successful implementation of transfer learning of two related tasks.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the outcomes of the proposed framework, highlighting its effectiveness and the impact of the

TABLE 4. Classification report for the fruit type model on the test set.

Fruit	Preci-sion	Re-call	F-1 Score	Sup-port
Apple	0.99	0.97	0.98	195
Banana	0.99	1.00	1.00	103
Guava	0.99	1.00	0.99	162
Lime	1.00	0.99	0.99	189
Orange	0.99	1.00	0.99	196
Pomegranate	1.00	1.00	1.00	727
Accuracy			0.99	1572
Macro Average	0.99	0.99	0.99	1572
Weighted Average	0.99	0.99	0.99	1572

selected data augmentation techniques. It also discusses these results and introduces a Streamlit application designed for real-world inference on fruit images.

A. VISUALIZING DATA AUGMENTATIONS

To illustrate the impact of our data preprocessing, we provide visual comparisons of the images before and after augmentation. The model's preprocessing pipeline incorporates standard image transformations, while the training pipeline is enhanced with the CutMix and MixUp techniques to introduce additional variability and complexity to the training data. Fig. 6 showcases these augmentations and their effects on the training images.

B. QUANTITATIVE RESULTS

Applying the proposed framework to the FruitNet dataset has yielded high-accuracy outcomes for classifying and grading multiple fruit types. The models underwent training on an augmented dataset comprising 30,450 samples, expanded from the original 14,700 samples through on-the-fly over-sampling. Fig. 7 illustrates the cross-entropy loss trends for training and validation datasets throughout the training epochs. Validation accuracy, which remained consistently high in the final epochs of training, is captured in Fig. 8. The testing phase saw the fruit type model attain an impressive average accuracy of 99.49%, with the fruit quality model closely following at 99.42%. A comprehensive breakdown of performance metrics, including precision, recall, F1-score, and macro and weighted averages, is presented for the fruit type model in Table 4 and the fruit quality model in Table 5. The confusion matrices, representing correct and incorrect predictions, are provided in Fig. 9 for both models.

C. DISCUSSION OF RESULTS

The performance metrics presented previously affirm the effectiveness of the proposed models in classifying fruit type and quality. Both models displayed a consistent decrease in training and validation losses, indicative of successful learning without overfitting —attributable to the effective use of data augmentation techniques.

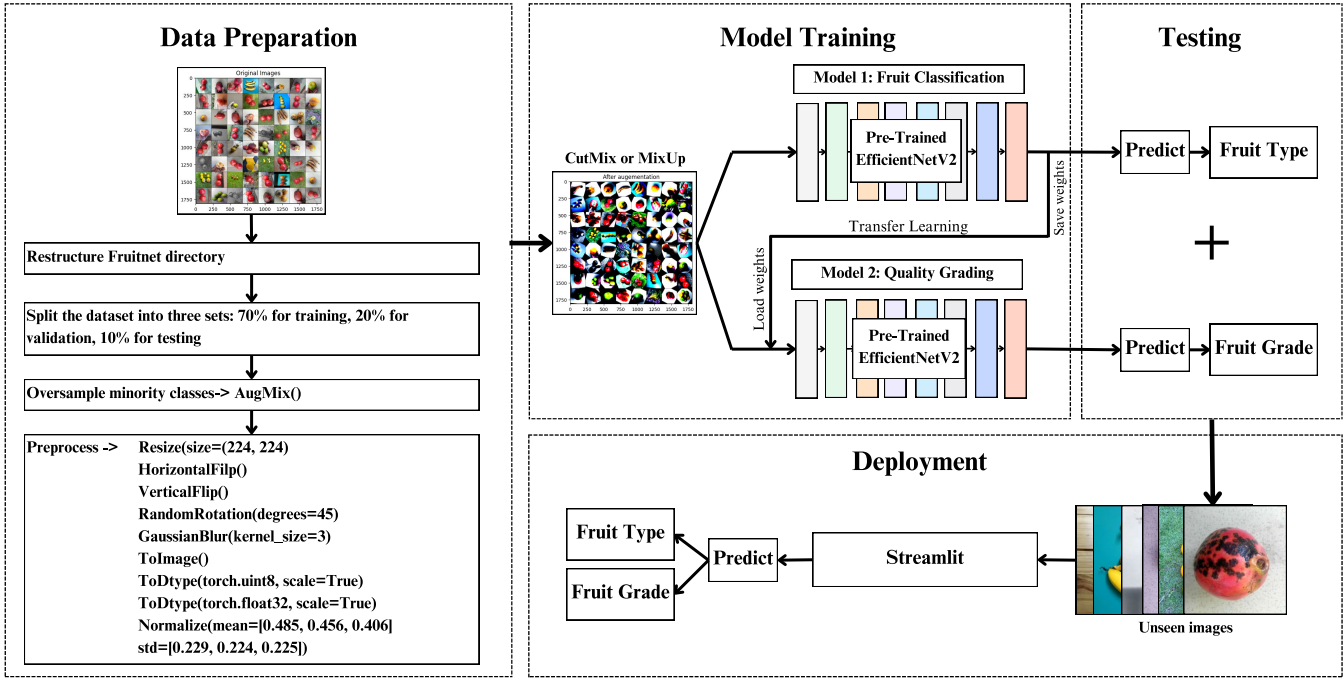


FIGURE 5. Proposed framework for multi-fruit classification and grading.

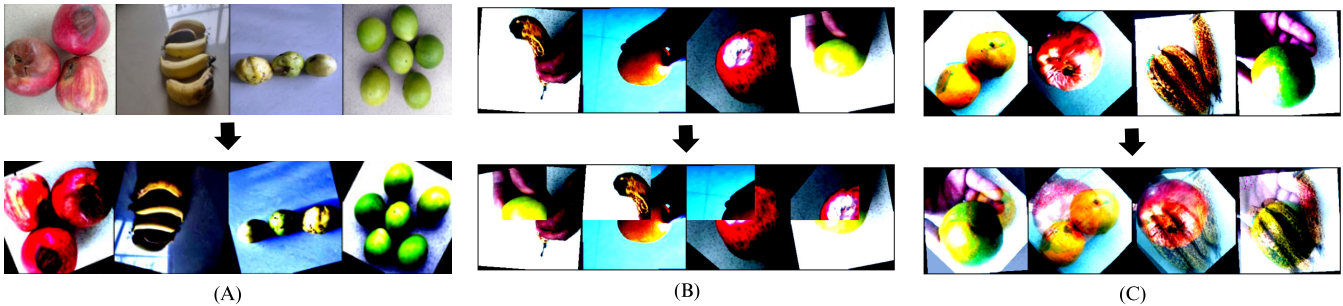


FIGURE 6. Examples of data augmentations applied for model training. Before and after traditional transformations are shown in (A). Before and after CutMix is shown in (B). Before and after MixUp is shown in (C).

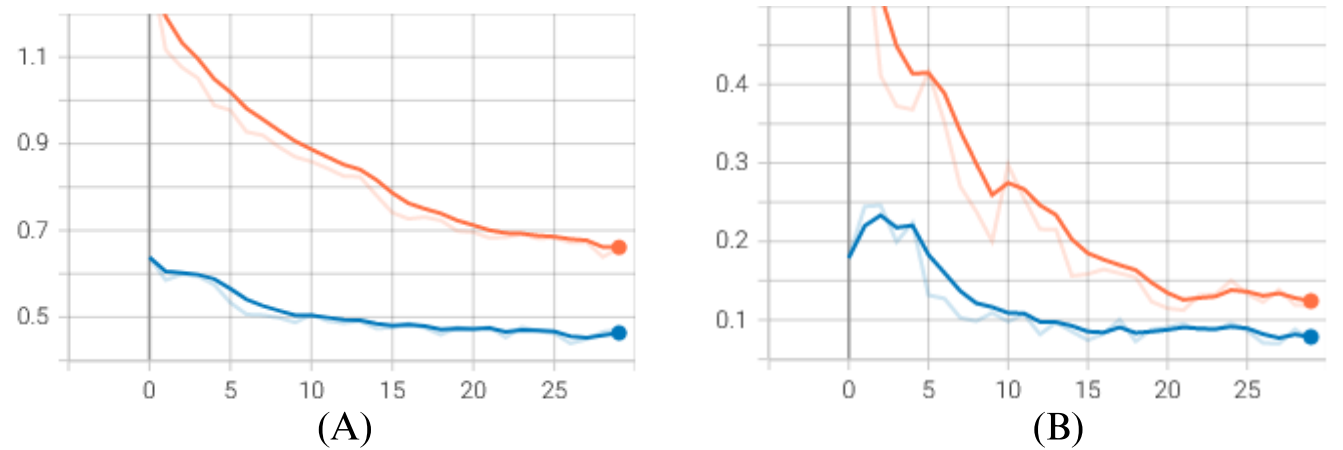


FIGURE 7. Cross-entropy loss for training and validation sets over the training epochs. Subfigure (A) presents the training loss, while Subfigure (B) details the validation loss, with the orange and blue lines representing the fruit type and quality models, respectively.

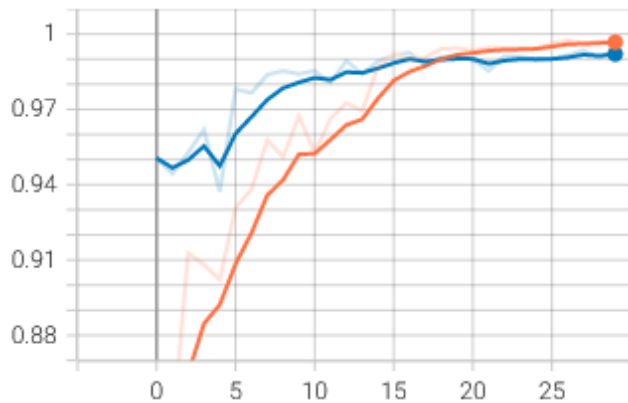


FIGURE 8. Validation accuracy trends for the fruit type and quality models during training. The orange and blue lines correspond to the fruit type and quality models, respectively.

TABLE 5. Classification report for the fruit quality model on the test set.

Fruit	Precision	Recall	F-1 Score	Support
Bad Quality	0.99	0.99	0.99	552
Good Quality	1.00	0.99	1.00	908
Mixed Quality	0.97	0.99	0.98	112
Accuracy			0.99	1572
Macro Average	0.99	0.99	0.99	1572
Weighted Average	0.99	0.99	0.99	1572

The validation loss served as a pivotal metric for adaptive learning rate adjustments, employing the ReduceLROnPlateau scheduler. This adaptive approach refined the learning process, as evidenced by the ascending accuracy trends in Fig. 8. The first model exhibited a steady increase in validation accuracy, exceeding 99% after 20 epochs. On the other hand, the second model—leveraging the first's learned features—demonstrated superior initial performance, achieving a validation accuracy of 95.1% in the first epoch and culminating at 99.3% by the 30th epoch, as displayed in Fig. 7 and Fig. 8.

Both models achieved high accuracy when testing a set of 1,575 unseen samples from the FruitNet dataset. The fruit type model attained 99% in both weighted and macro average accuracies across various fruit classes, as detailed in Table 4. Similarly, the fruit quality model reached 99% for weighted and macro average accuracies within the three quality categories, as Table 5 illustrates.

Analyzing the confusion matrices for both models in Fig. 9 (A) and (B), a high degree of predictive accuracy is evident. The fruit type model primarily confused apples with guavas and pomegranates and limes with oranges, likely due to their similar color and shape characteristics. For the fruit quality model, most misclassifications were within the good quality class, with some mislabeling as bad or mixed

TABLE 6. Accuracy comparison with previous multi-fruit classification studies.

Ref	Method	Accuracy
[10]	VGG16	96.75%-99.75%
[15]	VGG16	99.6%
[11]	RNN	98.47%
Proposed approach	EfficientNetV2	99.49%

quality, potentially due to shadow misinterpretation as a quality defect.

In conclusion, the models exhibited strong classification performance, yet there is potential for enhancement, especially in distinguishing between similar fruit varieties and refining quality assessment precision.

Compared to earlier work in multi-fruit classification, our method shows competitive or enhanced accuracy rates, as presented in Table 6.

D. WITH VS. WITHOUT SAME-DOMAIN TRANSFER LEARNING

We evaluated the impact of same-domain transfer learning by comparing its performance to a baseline model trained without this technique. The application of transfer learning led to a significant improvement in the model's learning dynamics. Specifically, the training and validation cross-entropy losses decreased consistently and reached lower final values when transfer learning was implemented, as depicted in Fig. 10. The initial validation accuracy was strikingly higher at 95.1% with transfer learning, compared to only 55% without it, as shown in Fig. 11. Moreover, the model with transfer learning achieved an average test set accuracy of 99.42%, which exceeds that of the non-transfer learning model by 1.02%.

The confusion matrix in Fig. 12 indicates a greater frequency of incorrect predictions across all classes when transfer learning was not used. Specifically, there were increases in misclassification by two, nine, and two instances for the bad, good, and mixed-quality classes, respectively. These results demonstrate the substantial advantages of employing same-domain transfer learning in enhancing the fruit quality model's accuracy and learning efficiency.

E. IMPLEMENTATION OF RESULTS

To showcase the practical application of our models for fruit classification and quality grading, we developed a Streamlit [24] application for inference. This application allows users to upload images and receive predictions of fruit type and quality, along with the confidence score for each prediction. Designed to test the trained models' performance in real-world scenarios, the application provides an intuitive platform for non-expert users. The user interface and a demonstration of the inference process are illustrated

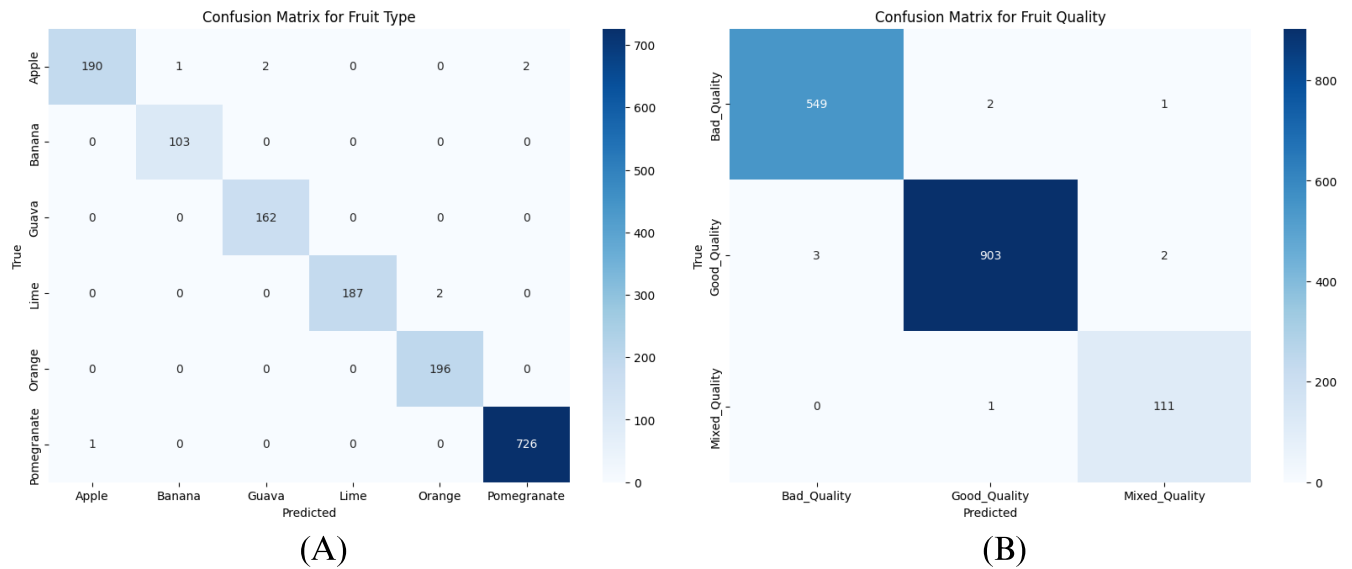


FIGURE 9. Confusion matrix results on the test set. The fruit type model confusion matrix is shown in (A). The fruit quality confusion matrix is shown in (B).

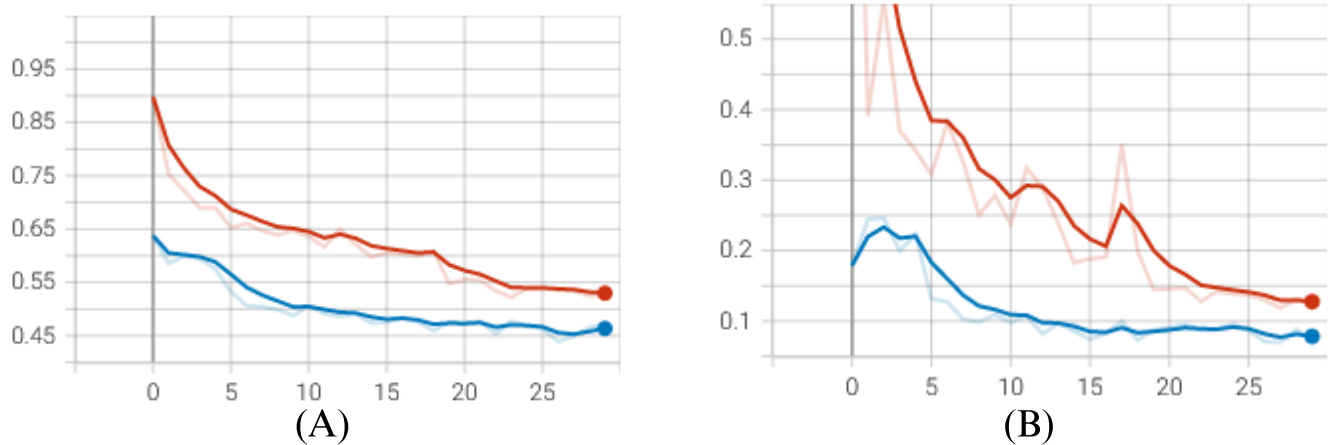


FIGURE 10. Training and validation loss trends with and without same-domain transfer learning. (A) Training loss comparison. (B) Validation loss comparison. The red and blue lines represent models without and with same-domain transfer learning, respectively.

in Fig. 13. The application can be accessed online at this link [25]

V. CONCLUSION AND FUTURE WORK

In this study, we have developed an advanced transfer learning framework utilizing the EfficientNetV2 architecture to classify and grade six fruit varieties from the FruitNet dataset, including bananas, apples, oranges, pomegranates, limes, and guavas. Our methodology incorporates a two-stage model training process: the first stage utilizes a pre-trained EfficientNetV2 for fruit classification, and the second stage applies same-domain transfer learning to grade fruit quality. We addressed the challenge of class imbalance in FruitNet by employing AugMix for oversampling and a suite of data augmentation techniques to enrich the training set diversity. Our models have demonstrated commendable accuracy,

reaching 99% on the test set, indicative of their robustness and potential for practical deployment in agricultural and food inspection applications.

Looking to the future, this work sets a strong precedent for creating more versatile and comprehensive visual classification systems. Further enhancements could expand the model's utility in precision agriculture, particularly in automating quality assessment processes.

A. LIMITATIONS

Despite the success of our proposed framework, we recognize its' limitations:

- 1) The models' generalization has been validated exclusively on the FruitNet dataset; real-world performance may vary due to unrepresented conditions in the training data.

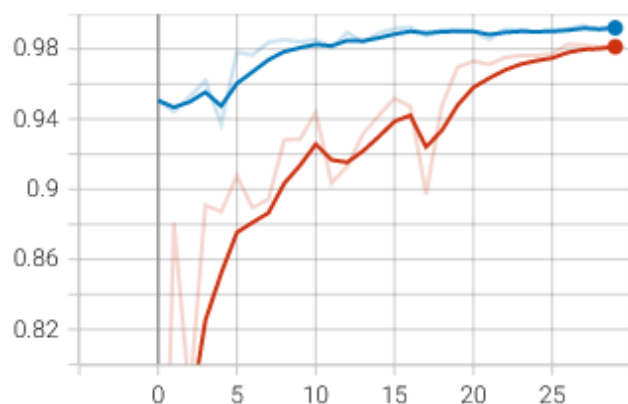


FIGURE 11. Validation accuracy trends during training. The red and blue lines represent models without and with same-domain transfer learning, respectively.

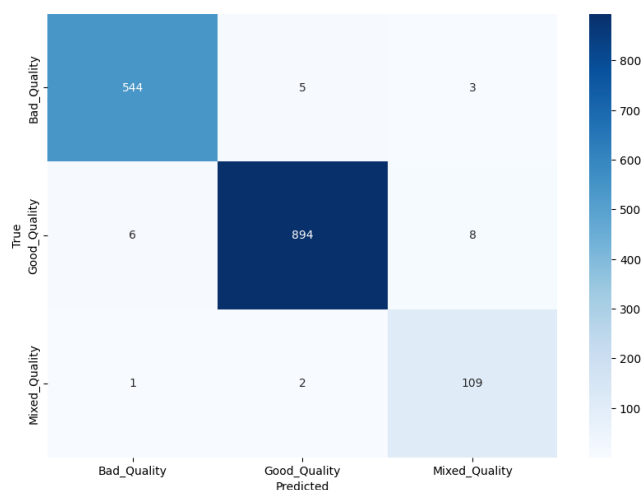


FIGURE 12. Confusion matrix for the fruit quality model test set without same-domain transfer learning.

- 2) Dependence on a singular CNN architecture, Efficient-NetV2, may limit the exploration of potentially more effective architectures.
- 3) The classification and grading process currently includes six fruit types; however, there is the potential to incorporate many more.

B. FUTURE RESEARCH

Further research in this domain should take into account the following considerations:

- 1) While the current research tackled the classification problem with two separate models, future studies should explore multi-task learning frameworks where a single model can simultaneously learn multiple task-specific features.
- 2) Compiling a more comprehensive and diverse dataset, potentially by combining existing datasets or curating new images, to improve model robustness.
- 3) Assessing the proposed model's integration within an automated machine vision system for real-time fruit recognition and grading.



Multi-Fruit Classification and Grading

Choose an image...



Drag and drop file here

Limit 200MB per file • PNG, JPG, JPEG

Browse files



IMG_8155.JPG 15.3KB



Uploaded Image.

Classifying...

Predicted fruit type: Apple (probability: 99.98%)

Predicted fruit quality: Bad Quality (probability: 85.39%)

FIGURE 13. Streamlit app user interface showcasing an example inference.

- 4) Refining fruit segmentation methods to enhance model focus on relevant features and minimize background noise.
- 5) For images that feature multiple fruits with mixed qualities, the model could be developed to classify each fruit individually as either "Bad Quality" or "Good Quality." Alternatively, a percentage could be assigned to represent the overall quality of a batch of fruits.

ACKNOWLEDGMENT

The authors acknowledge the use of ChatGPT-4 [26] accessed via <https://poe.com> for assistance with the coding aspects of this research. Specifically, It was utilized for:

- Providing coding templates.
- Correcting errors.

REFERENCES

- [1] C.-J. Du and D.-W. Sun, "Learning techniques used in computer vision for food quality evaluation: A review," *J. Food Eng.*, vol. 72, no. 1, pp. 39–55, Jan. 2006.
- [2] B. Dhiman, Y. Kumar, and M. Kumar, "Fruit quality evaluation using machine learning techniques: Review, motivation and future perspectives," *Multimedia Tools Appl.*, vol. 81, no. 12, pp. 16255–16277, May 2022.
- [3] A. Bhargava and A. Bansal, "Fruits and vegetables quality evaluation using computer vision: A review," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 33, no. 3, pp. 243–257, Mar. 2021.
- [4] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, "A review of convolutional neural network applied to fruit image processing," *Appl. Sci.*, vol. 10, no. 10, p. 3443, May 2020.
- [5] V. Meshram, K. Patil, V. Meshram, D. Hanchate, and S. D. Ramkteke, "Machine learning in agriculture domain: A state-of-art survey," *Artif. Intell. Life Sci.*, vol. 1, Dec. 2021, Art. no. 100010.
- [6] L. Zhou, C. Zhang, F. Liu, Z. Qiu, and Y. He, "Application of deep learning in food: A review," *Comprehensive Rev. Food Sci. Food Saf.*, vol. 18, no. 6, pp. 1793–1811, Nov. 2019.
- [7] M. M. Raikar, S. M. Meena, C. Kuchanur, S. Girraddi, and P. Benagi, "Classification and grading of Okra-ladies finger using deep learning," *Proc. Comput. Sci.*, vol. 171, pp. 2380–2389, Jan. 2020.
- [8] H. M. R. Iqbal and A. Hakim, "Classification and grading of harvested mangoes using convolutional neural network," *Int. J. Fruit Sci.*, vol. 22, no. 1, pp. 95–109, Dec. 2022.
- [9] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in *Proc. 38th Int. Conf. Mach. Learn.*, M. Meila and T. Zhang, Eds. PMLR, Jul. 2021, pp. 10096–10106. [Online]. Available: <https://proceedings.mlr.press/v139/tan21a.html>
- [10] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, "Automatic fruit classification using deep learning for industrial applications," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1027–1034, Feb. 2019.
- [11] B. Dhiman, Y. Kumar, and Y.-C. Hu, "A general purpose multi-fruit system for assessing the quality of fruits with the application of recurrent neural network," *Soft Comput.*, vol. 25, no. 14, pp. 9255–9272, Jul. 2021.
- [12] V. Meshram and K. Patil, "FruitNet: Indian fruits image dataset with quality for machine learning applications," *Data Brief*, vol. 40, Feb. 2022, Art. no. 107686.
- [13] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, "CutMix: Regularization strategy to train strong classifiers with localizable features," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6022–6031.
- [14] H. Zhang, M. Cissé, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, Apr./May 2017. [Online]. Available: <https://openreview.net/forum?id=r1Ddp1-Rb>
- [15] I. M. Nasir, A. Bibi, J. H. Shah, M. A. Khan, M. Sharif, K. Iqbal, Y. Nam, and S. Kadry, "Deep learning-based classification of fruit diseases: An application for precision agriculture," *Comput., Mater. Continua*, vol. 66, no. 2, pp. 1949–1962, 2021.
- [16] M. Knott, F. Perez-Cruz, and T. Defraeye, "Facilitated machine learning for image-based fruit quality assessment," *J. Food Eng.*, vol. 345, May 2023, Art. no. 111401.
- [17] K. Albarrak, Y. Gulzar, Y. Hamid, A. Mehmood, and A. B. Soomro, "A deep learning-based model for date fruit classification," *Sustainability*, vol. 14, no. 10, p. 6339, May 2022.
- [18] N. Ismail and O. A. Malik, "Real-time visual inspection system for grading fruits using computer vision and deep learning techniques," *Inf. Process. Agricult.*, vol. 9, no. 1, pp. 24–37, Mar. 2022.
- [19] K. Patil and V. Meshram, "FruitNet: Indian fruits dataset with quality (good, bad & mixed quality)," *Mendeley Data*, Mar. 2022. [Online]. Available: <https://data.mendeley.com/datasets/b6fftwbr2v/3>
- [20] D. Hendrycks, N. Mu, E. D. Cubuk, B. Zoph, J. Gilmer, and B. Lakshminarayanan, "AugMix: A simple data processing method to improve robustness and uncertainty," 2019, *arXiv:1912.02781*.
- [21] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520.
- [22] S. Gupta and B. Akin, "Accelerator-aware neural network design using AutoML," 2020, *arXiv:2003.02838*.
- [23] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [24] Streamlit. (Jan. 2024). *Streamlit: A Faster Way to Build and Share Data Apps*. [Online]. Available: <https://streamlit.io/>
- [25] (Jan. 2024). *MFCG Streamlit Application*. [Online]. Available: <https://5by5amzhzuhbjsqhsvj3s5.streamlit.app/>
- [26] OpenAI. (Jan. 2024). *GPT-4: Scaling up Deep Learning*. [Online]. Available: <https://openai.com/research/gpt-4>

LAMA A. ALDAKHIL received the B.Sc. degree in computer science from Qassim University, Saudi Arabia, in 2020, where she is currently pursuing the M.Sc. degree in information technology. She is an Information Technology Specialist with Onaizah Colleges. Her research interests include artificial intelligence, machine learning applications in image recognition, and the ethical implications of emerging technologies.

AESHAH A. ALMUTAIRI received the Ph.D. degree in computer science from Durham University, U.K. She is currently an Assistant Professor with the Computer College, Qassim University, Saudi Arabia. Her research interests include computer graphics, in particular, she aims to employ machine learning techniques for better recognition and identification.

• • •