iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68187

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Automated Assessment of Fruit Quality Using Transfer Learning and MobileNetV2

Odra Bira[1], Sreeja Sunkeswaram[2], Naresh R[3]
[1]*Keshav Memorial Institute of technology, Hyderabad, India*
[2]*CMR College of Engineering and Technology, Hyderabad, India*
[3]*Indian Institute of Technology Madras, Chennai, India*

Abstract: Fruit quality and health strongly affect market value, customer satisfaction, and overall farm yield. Manual inspection methods remain exceedingly labour intensive. They are fairly subjective and vulnerable to inconsistencies. This paper presents an automatic system that uses many advanced deep learning methods, specifically transfer learning via MobileNetV2, to categorize fruits by freshness or spoilage. Classical machine learning (ML) methods like Support Vector Machines, Random Forest, and K-Nearest Neighbours are comparatively analysed with modern Convolutional Neural Network (CNN) designs (like VGG16, MobileNetV1, and MobileNetV2). Many experimental evaluations reveal that MobileNetV2 not only provides substantially outstanding training accuracy but maintains consistently strong validation metrics. These findings underscore the model's potential for deployment on mobile platforms as well as in real-time agricultural applications, thereby leading to better productivity along with reduced post-harvest losses.
Keywords: Transfer Learning, MobileNetV2, Fruit Quality Assessment, Convolutional Neural Networks, Deep Learning

I. INTRODUCTION

Agriculture is one of the basic support pillars that maintain food security and economic well-being in most regions of the world, especially in regions where a large majority of the population depends heavily on agricultural produce for their daily food and dietary requirements [9]. One of the main challenges that have lasting impacts in this sector is the problem of post-harvest fruit and vegetable spoilage, which has a tendency to occur before the produce reaches the consumer market. This dismal post-harvest spoilage condition not only leads to colossal financial losses to the farmers and distributors but also significantly contributes to the global problem of food wastage, which is a cause of concern to societies around the world. Estimates made available by the Food and Agriculture Organization (FAO) indicate that it is a cause of concern that nearly one-third of all the food that is grown for human consumption—equivalent to about 1.3 billion tons—is wasted annually due to a myriad of reasons, including inadequate handling and storage procedures [8].

Recent rise of Artificial Intelligence (AI) and Deep Learning (DL) technology has provided new avenues for the development of computer systems that are capable of solving complex problems across a broad spectrum [12]. Transfer learning is among the most significant methods in such technology that takes advantage of pre-trained models trained on large databases like ImageNet [2]. The technique has been found to be especially useful in scenarios where there is little training data or where the computational cost of training a model from scratch is not feasible [11]. In this case, the system is tailored for the specific purpose of classifying fruits into two categories: fresh or rotten—namely bananas, oranges, and apples. The choice to use MobileNetV2, a light yet very efficient convolutional neural network (CNN) model, is motivated by its proven ability to deliver high accuracy. It is also well-suited for deployment in scenarios where resources could be scarce [1].

II. EXPERIMENTAL FRAMEWORK

A. Data Overview and Preparation Steps

One such public domain data set existing on Kaggle, designed specifically to solve issues related to classification of the quality of fruit, was used in the present study [7]. The data set alone contains a total of 13,599 individual images out of which 10,901 images have been reserved for training and the remaining 2,698 images reserved for testing. For easy organization and accessibility, images have been placed logically in six individual folders named FreshApple, FreshOrange, FreshBanana, RottenApple, RottenOrange, and RottenBanana. All the images in the dataset went through an extensive preprocessing phase, where they were resized to a uniform dimension in a considerate manner. The process was important to ensure consistency and homogeneity in the entire dataset.

Additionally, standard normalization techniques were rigorously employed to scale and normalize the pixel intensity values of the images such that the data is consistent and accurate. In an effort to improve the effective dataset size and to prevent overfitting, various techniques of data augmentation were also used. Such techniques included rotating images, flipping images to form mirror images, scaling adjustments to alter their dimensions, and brightness level adjustments, all in efforts to diversify the dataset.

B. Transfer Learning and Model Selection

Transfer learning is a powerful and efficient method in the domain of deep learning, that utilizes the knowledge accumulated by training a model on a task and then utilizes the learned knowledge to a different but similar task. This yields further benefits for a target task that has a small amount of labeled data. A common base for pre-trained models is ImageNet [2], which is a large image database encompassing more than 14 million images, over thousands of categories. Models trained on ImageNet with architectures, specifically VGG, ResNet, and MobileNet families, exhibit strong performance in terms of generalization, and can act as a foundation for a fringed area in the field of vision [11].

Under transfer learning, much of the lower layers of CNNs are kept intact, for they learn low-level universal features, such as edges, textures, and simple patterns. Typically, these low-level features can be transferred across tasks and domains. In CNNs, the higher layers learn features that are more task specific. The higher layers can use a target task's dataset to fine tune them and improve performance. This state of modularity of a model for transfer learning, are both flexible and computationally efficient in learning, while maintaining the time and compute cost of pretraining models from scratch.

In this study, a full range of model architectures will be studied to evaluate classification performance through either traditional machine learning or deep learning models. The traditional machine learning methods been tested 3 include Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN). These models are well known, have a long history of being successful with structured data and small-to-medium image datasets, but lack the hierarchical feature learning that is prominent in modern day deep learning models.

The study compared traditional models with advanced convolutional neural network architectures. It featured VGG16 [3], with a depth based on simplicity, MobileNetV1, and then the more advanced MobileNetV2 [1] for efficiency and performance considerations in specific use cases on mobile and edge devices. The MobileNet models use depthwise separable convolutions to dramatically reduce the number of parameters and computational cost while still retaining high accuracy in their classification.

The evaluation results demonstrated significant differences in performance between the traditional and deep learning models, yielding similar results. The traditional classifiers, SVM and Random Forest, showed comparable performances using accuracy and their achieved accuracies of 85% and 88%, respectively. These results are a reasonable measure for small problems or less complex image classification problems. However, the CNN-based models saw great improvements over traditional models where MobileNetV2, in particular, obtained phenomenal algorithm accuracy at 99% accuracy in training, plus had high validation accuracy which was another indication of the model's ability to generalize well rather than demonstrate overfitting. The performance is often credited to the architectural improvements of MobileNetV2 (inverted residuals and linear bottlenecks); these developments provide better reuse of the features while keeping a compact model.

In addition, using transfer learning with these CNN architectures allowed us to accelerate convergence during training and generally made the overall framework more robust. We found that pre-trained models performed well with fewer epochs to get to optimal performance when compared to models trained from scratch. This emphasizes the notable effect transfer learning has on reducing training time and computational costs. In summation, the study shows the advantages of more contemporary deep learning models, especially those utilizing transfer learning, compared to machine learning approaches for image classification tasks. MobileNetV2 was the best option as it has good accuracy, speed, and resource efficiency. It is well suited for real-world deployment scenarios and settings with limited computational resources while retaining high accuracy.

TABLE I
MOBILENETV2 ARCHITECTURE SUMMARY.

Input	Layer Type	Ratio	Filters	Reps	Stride
224 ^ 2 × 3	Convolution (2D)	-	32	1	2
112 ^ 2 × 32	Bottleneck Module	1	16	1	1
112 ^ 2 × 16	Bottleneck Module	6	24	2	2
56 ^ 2 × 24	Bottleneck Module	6	32	3	2
28 ^ 2 × 32	Bottleneck Module	6	64	4	2

14 ^ 2 × 64	Bottleneck Module	6	96	3	1
14 ^ 2 × 96	Bottleneck Module	6	160	3	2
7 ^ 2 × 160	Bottleneck Module	6	320	1	1
7 ^ 2 × 320	Pointwise Conv (1×1)	-	1280	1	1
7 ^ 2 × 1280	Global Avg Pool (7×7)	-	-	1	-
1 ^ 2 × 1280	Pointwise Conv (1×1)	-	k	-	-

III. METHODOLOGY

This section describes the approach used to create an automated fruit quality assessment system powered by deep learning, based on the MobileNetV2 architecture [1]. Several primary stages were followed in the methodology, including dataset preparation and cleansing, model architecture, training strategy, optimization, and performance evaluation.

The major insight in this research is classifying fruits as fresh or rotten using convolutional neural networks (CNNs) [12]. MobileNetV2, known for its lightweight and efficient structure, was chosen as the primary architecture [1, 13]. The process includes a transfer learning methodology, which speeds up convergence by leveraging learned features from a pre-trained model [11, 6].

A. MobileNetV2 Architecture

MobileNetV2 is a convolutional neural network framework designed specifically for mobile and edge devices [1]. It incorporates two main architectural innovations - inverted residuals and linear bottlenecks - to achieve state-of-the-art accuracy for an integer foundation model while significantly reducing the computational burden [13]. A detailed layer-wise summary of the MobileNetV2 architecture is provided in Table 1

The key advancement with MobileNetV2 is the incorporation of inverted residual blocks [1]. Each block begins with a 1×1 convolution layer with ReLU6 activation, also known as a thin expansion layer, followed by a depthwise convolution, which captures spatial features independently on each input channel, and finally a 1×1 convolution layer with no activation to project the features into the target dimensionality [4]. This combination leads to a unique operation combining efficiency as well as the retention of important features. This design has distinct advantages when considering the cost of the calculations because standard convolution requires significantly more multiplications than depthwise and pointwise convolution in MobileNetV2 [1]. The cost of the calculations can be stated as:

$$\text{Conventional Convolution Cost} = K_s^2 \cdot C_{in} \cdot C_{out} \cdot S^2 \quad (1)$$

$$\text{Depthwise + Pointwise Cost} = K_s^2 \cdot C_{in} \cdot S^2 + C_{in} \cdot C_{out} \cdot S^2 \quad (2)$$

In these expressions, K_s is the size of the convolution kernel, C_{in} and C_{out} refer to the number of input and output channels, respectively, and S indicates the spatial resolution (i.e., height and width) of the input feature map.

An additional notable characteristic presented in the architecture is the presence of a residual connection between bottleneck blocks. In MobileNetV2, instead of connecting thin layers as in ResNet [19], thick-to-thin layers are connected through residual links. This is fundamentally important for gradient flow and enables the design to modify the deep aspect of the depth. This arrangement is depicted in a representational aspect in Fig. 1.

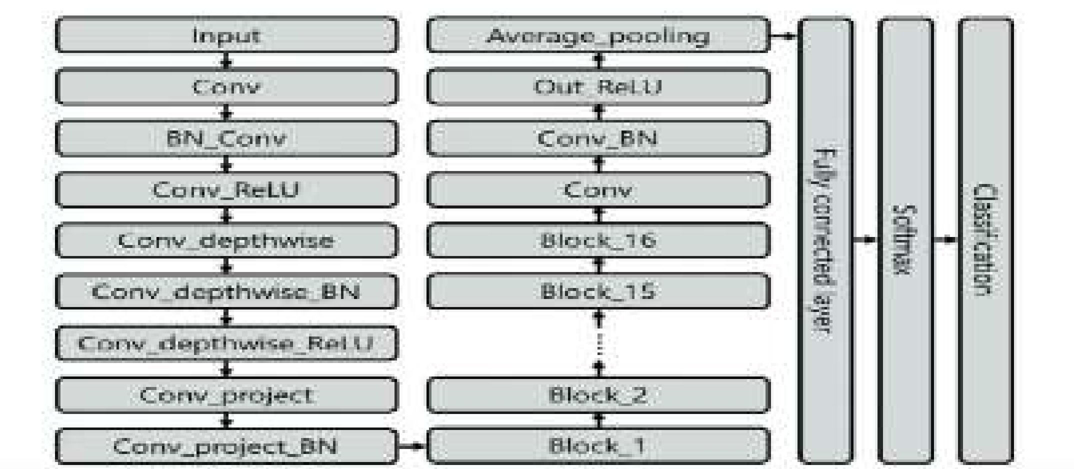


Fig. 1 Layers in MobileNetV2: Thin–Depth wise–Thin structure with inverted residuals.

B. Model Development and Workflow

The dataset utilized for this research includes labeled photographs of three varieties of fruit (apples, bananas, and oranges) which are distinguished as either fresh or rotten [7]. The data set was randomly divided into two subsets, that is, the training and validation subsets, maintaining balanced class numbers using stratified sampling [15].

Images were resized to a common shape to conform to the input size expectations of MobileNetV2 [1]. Normalization was applied to scale pixel values between 0 and 1. Additionally, data augmentation was implemented to artificially increase the size and variance of the dataset by horizontally flipping, rotating, zooming, and modifying brightness [16]. This reduced overfitting and improved the generalization capabilities of the model.

In order to figure out the most appropriate modeling strategy, a series of experiments were run on multiple configurations: Initial tests were run using transfer learning on a VGG16 model [3]. This model reached 96% training accuracy and 98% validation accuracy, and in this particular configuration all of the data fit on the screen; however, the model was not able to generalize well on unseen data, suggesting it was overfit [6]. Then, MobileNetV2 was trained from scratch, reaching 95% training accuracy and 94% validation accuracy, with slightly more training epochs, which also used a significant number of resources. While the model developed through this method attained a training accuracy of 98%, the validation accuracy was only 60%, indicating overfitting once again. The architecture that performed the best was the pre-trained MobileNetV2 model with all the layers frozen except for the final classification layer [13]. In 7 epochs, the model has reached 99% training accuracy and validation accuracy of 97%, and had a clear convergence and stable performance.

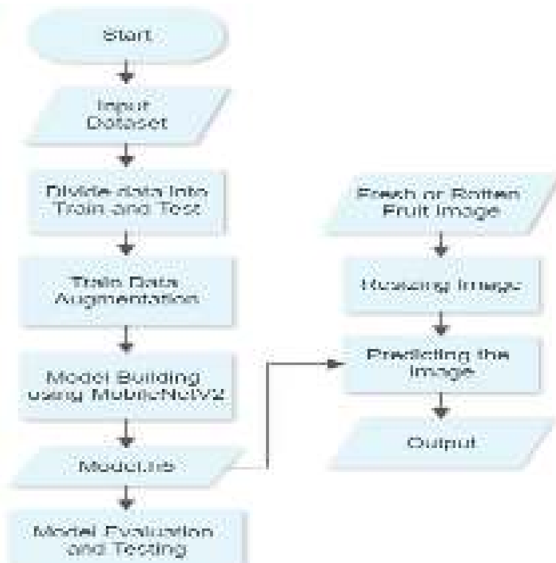


Fig. 2 Flowchart of the fruit quality assessment process using MobileNetV2.

The complete workflow for the classification system, from image input to prediction output, is illustrated in Fig. 2. Additionally, the model's performance during training and validation, with and without fine-tuning and augmentation, is summarized graphically in Fig. 3.

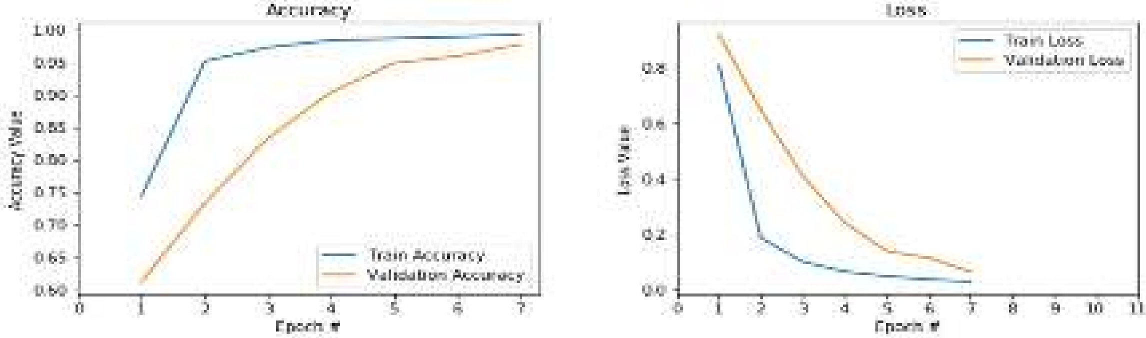


Fig. 3 Performance metrics for MobileNetV2 with fine-tuning and image augmentation.

C. Optimization Strategy

The training process was optimized, and the loss function decreased by utilizing the RMSProp optimizer [5]. RMSProp is an adaptive learning rate method that is effective in solving non-stationary objectives, and works quite well with deep neural networks [4]. It keeps a moving average of the squared gradients to normalize the gradient and control oscillation. The update rules for RMSProp are defined as:

$$V_t = \alpha V_{t-1} + (1-\alpha)\, \nabla_t^2 \tag{3}$$

$$\omega_{t+1} = \omega_t - (\lambda / \sqrt{(Vt + \delta)})\, \nabla_t \tag{4}$$

In the above expressions, ∇_t represents the gradient at iteration t, λ denotes the learning rate, α is the exponential decay factor (commonly set to 0.9), and δ is a small constant introduced to ensure numerical stability and avoid division by zero. Despite testing alternate optimizers such as Adam [10], RMSProp consistently gave better convergence and generalization in this case, justified by both literature support and experimental results on comparable image classification tasks [13, 14]. In brief, the approach adopted in this study effectively unites data pre-processing, transfer learning with MobileNetV2 and a solid optimization method with RMSProp. The model produced is accurate and computationally efficient enough to be used in real-time agricultural settings for fruit quality assessment [9].

TABLE II
ACCURACY COMPARISON OF DIFFERENT MODELS USED FOR FRUIT QUALITY CLASSIFICATION.

Model	Accuracy
SVM	85%
Random Forest	88%
MobileNetV2	97%

IV. RESULTS & DISCUSSION

Large-scale and intensive experiments in the research show that the new system utilizing MobileNetV2 possesses a very high level of performance, much higher than that of traditional machine learning models, and even higher than the performance of many other deep learning architectures in the field of fruit quality evaluation in general. The comprehensive performance summary, with the major findings and comparisons, is graphically displayed in Table 2.

Extensive and rigorous experiments have produced robust evidence that the proposed system, based on MobileNetV2, has a performance rate much higher than that of typical machine learning models. In addition, it surpasses other current deep learning models in fruit quality evaluation. To observe a detailed description of performance outcomes in detail, please refer to Table 2, where this information is evidently presented.

The exceptional and remarkable performance of MobileNetV2 can be largely credited to its extremely efficient use of pre-trained weights, which greatly enables strong and effective feature extraction abilities even when dealing with a comparatively small dataset. The rapid convergence that occurs with the unmoderated transfer learning environment similarly emphasizes and asserts the model’s phenomenal ability to generalize to unobserved and novel data.

In addition to these advantages, the lightness of the structure of MobileNetV2 similarly contributes towards having minimal computational overhead, hence contributing towards being most suitable for deployment on mobile systems as well as real-time crop environments where efficacy is most imperative.

As shown in Table 3, the MobileNetV2 model also had outstanding results on other significant evaluation metrics showing it is both accurate and trustworthy in correctly detecting whether fruits are fresh or rotten.

The outstanding and exceptional performance that MobileNetV2 demonstrates is directly due to its extremely efficient use of pre-trained weights, which significantly allows the learning of strong features even when used with a fairly small dataset. The fact that the model can reach quick convergence in the baseline transfer learning case suggests a strong capacity to generalize to new data. In addition to that, the extreme lightweight nature of MobileNetV2 is a significant reason for saving computational overhead and allows it to be a great model to deploy on not just mobile devices but also in real-time agricultural environments, where efficiency is paramount.

TABLE III
PERFORMANCE METRICS FOR MOBILENETV2 ON FRUIT QUALITY CLASSIFICATION.

Metric	MobileNetV2 Score
Precision	96.5%
Recall	97.2%
F1-Score	96.8%

V. FINDINGS & FUTURE DIRECTION

This study proposes a very robust automated assessment of the fruit quality system that utilizes transfer learning and the MobileNetV2 architecture. The proposed method replaces the time-consuming method of evaluating fruit quality manually with a technology-based process that has less human error and is significantly faster. The system's excellent training and validation accuracy, obtained using MobileNetV2's high accuracy and efficiency in computation, makes it well-suited for real-time use in agricultural environments. Additionally, by using a pre-trained model, the solution quickly adjusts to different datasets and environmental conditions, providing stable performance across different types of fruits. Finally, this system can automate the agricultural supply chain, minimize post-harvest losses, and guarantee that the market only receives high-quality crops, hence impacting both the farmers and the consumers positively.

To improve the assessment of fruit quality, future studies should expand and build on existing priorities. One option may involve more use of multimodal sensor data; for instance, information about temperature, humidity, and moisture can create a more holistic assessment, which may lead to greater prediction. Another possibility is to incorporate genomic data, which may help identify fruit types that are naturally resistant to pests and diseases, helping to breed better fruit varieties and improve crop management.

The use of ensemble learning can also help increase classification accuracy and improve the robustness of classification systems. Also, considering the design and efficiency of MobileNetV2, the next step may be to build an app for fruit quality assessment with some combination of Internet of Things (IoT) devices to enable real-time fruit quality assessment, automated sorting, and dynamic decision-making to improve the agricultural supply chain. Scalability will be an important issue when exploring real-world applications and deployments across different geographical areas and environmental conditions.

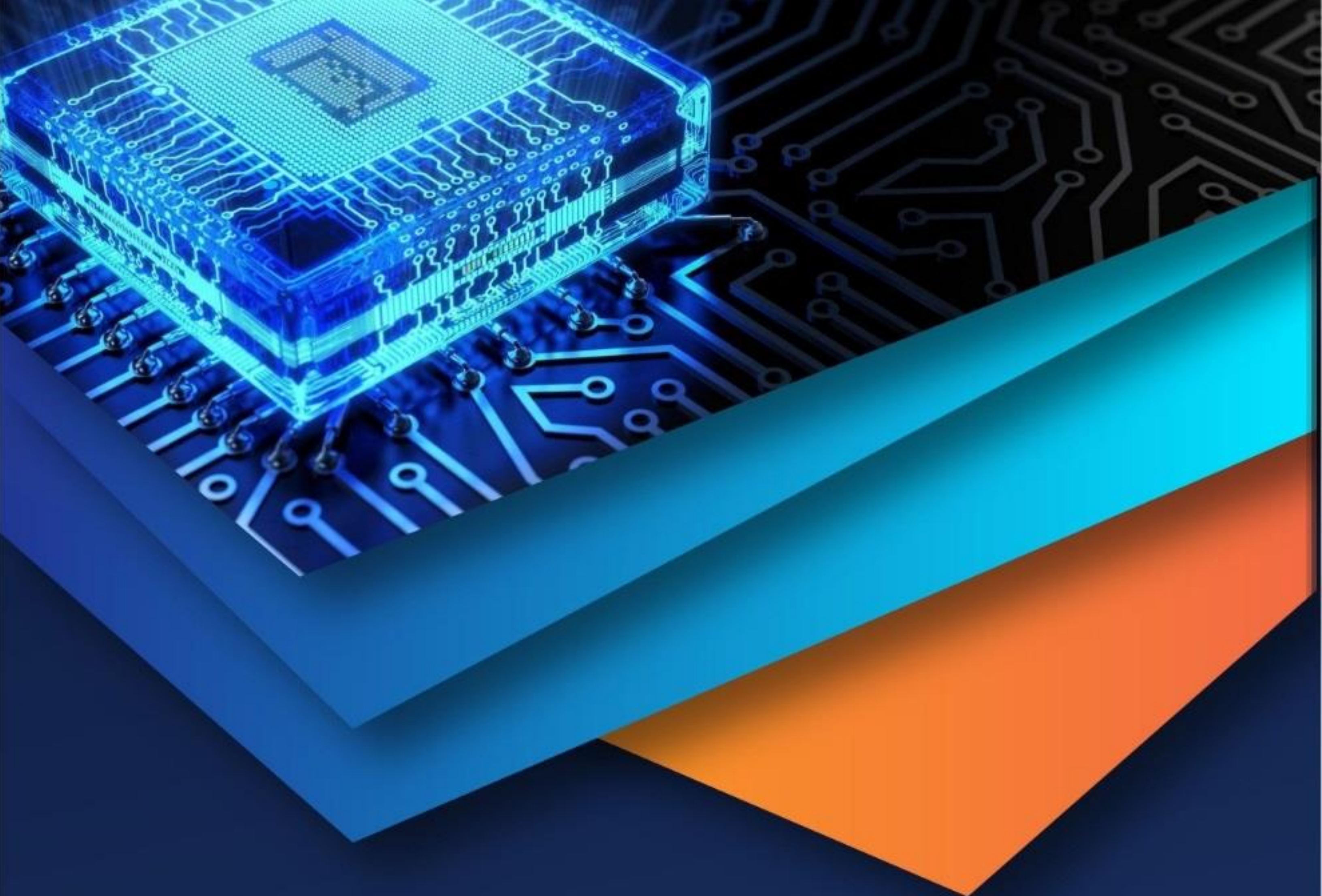
In conclusion, though the present project concentrates on accuracy, future projects may be worth the work of including evaluation criteria for computational efficiency, energy usage, and inference time in applications for mobile and embedded systems.

REFERENCES

[1] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520, 2018.
[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst. (NIPS), pp. 1097–1105, 2012.
[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proc. Int. Conf. on Learning Representations (ICLR), 2015.
[4] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
[5] T. Tieleman and G. Hinton, "Lecture 6.5—RMSProp," COURSERA: Neural Networks for Machine Learning, 2012.
[6] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in Proc. Adv. Neural Inf. Process. Syst. (NIPS), pp. 3320–3328, 2014.
[7] F. Chen, X. Zhang, and Y. Liu, "Deep learning-based automated fruit quality inspection: A review," IEEE Access, vol. 7, pp. 80803–80813, 2019.
[8] A. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big Data in Smart Farming – A review," Agricultural Systems, vol. 153, pp. 69–80, 2017.
[9] P. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.
[10] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. Int. Conf. on Learning Representations (ICLR), 2015.
[11] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345–1359, 2010.
[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015.
[13] B. B. Shahi, S. Singh, and S. A. Hussain, "Deep learning-based approach for classification of fruits using MobileNetV2," International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 6, pp. 6514–6522, 2022.
[14] M. Chakraborty and D. Banerjee, "Quality inspection of fruits using deep learning techniques," in Proc. Int. Conf. on Computational Intelligence and Machine Learning, pp. 123–129, 2021.
[15] H. Muresan and M. Oltean, "Fruit recognition from images using deep learning," Acta Univ. Sapientiae, Informatica, vol. 10, no. 1, pp. 26–42, 2018.
[16] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, pp. 1–48, 2019.
[17] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision (IJCV), vol. 115, pp. 211–252, 2015.
[18] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," Journal of Big Data, vol. 3, no. 1, pp. 1–40, 2016.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
[20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 4700–4708, 2017.
[21] D. A. Zarate, J. W. Branch, and J. A. Reinoso-Gordo, "A CNN-based approach for fruit quality classification using transfer learning," Computers and Electronics in Agriculture, vol. 213, p. 108085, 2024.
[22] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Information Processing and Management, vol. 45, no. 4, pp. 427–437, 2009.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY