International Journal of Innovative Research in Engineering
*Volume 6, Issue 5 (September - October 2025), PP: 13-18.*
https://www.doi.org/10.59256/ijire.20250605003
*www.theijire.com*
**ISSN No: 2582-8746**

# Fruit and Vegetable Image Recognition Using Convolutional Neural Networks

**Shaik Mushraf[1], Dr. Mohd Rafi Ahmed[2]**
[1] *Student, MCA, Deccan College of Engineering and Technology, Hyderabad, Telangana, India.*
[2]*Associate Professor, MCA, Deccan College of Engineering and Technology, Hyderabad, Telangana, India.*

**How to cite this paper:**
Shaik Mushraf[1], Dr. Mohd Rafi Ahmed[2] ,"Fruit and Vegetable Image Recognition Using Convolutional Neural Networks", IJIRE-V6I5-13-18.



**Abstract:** *The classification and identification of fruits and vegetables through image recognition is a significant challenge in modern agriculture, food processing, and retail automation. Manual sorting methods are often error-prone, time-consuming, and inefficient for large-scale operations, necessitating the development of automated, scalable solutions. This study introduces a deep learning–based framework for fruit and vegetable recognition using Convolutional Neural Networks (CNNs). A sequential CNN model was designed and trained on a publicly available dataset containing thousands of labeled fruit and vegetable images. Preprocessing techniques such as resizing, normalization, and data augmentation (rotation, flipping, zooming, and shearing) were applied to enhance generalization and mitigate overfitting. The model was implemented using TensorFlow and Keras, trained with categorical cross-entropy loss, and evaluated using accuracy, precision, recall, and confusion matrix analysis. Results indicate that the CNN achieved high classification accuracy, demonstrating its effectiveness in distinguishing between visually similar categories of produce. The framework shows strong potential for integration into commercial retail systems, automated inventory management, and agricultural inspection workflows. Furthermore, this work lays the foundation for future enhancements, including expansion to additional produce categories and real-time mobile or web-based deployment, thereby contributing to intelligent, AI-driven solutions for food quality control and supply chain optimization.*

**Key Words:** *Fruit and Vegetable Recognition; Convolutional Neural Network (CNN); Deep Learning; Image Classification; Data Augmentation; TensorFlow; Keras; Agricultural Automation; Retail Technology; Computer Vision.*

## I.INTRODUCTION

Fruits and vegetables form a vital component of human nutrition and global food supply chains, making their accurate classification and quality control a priority in agriculture, retail, and logistics. Traditional methods of sorting and identifying produce are primarily manual, depending on human workers' visual inspection and experience. While effective in small-scale operations, these approaches are inherently limited by subjectivity, inconsistency, and inefficiency. As the demand for food safety and large-scale distribution continues to grow, there is a critical need for automated, reliable, and scalable solutions.

Advancements in artificial intelligence (AI) and deep learning have revolutionized image recognition and classification tasks across diverse fields. Convolutional Neural Networks (CNNs), in particular, have demonstrated remarkable performance in extracting hierarchical image features, enabling robust classification of complex visual patterns. By leveraging such architectures, it has become feasible to design automated systems capable of distinguishing between different types of fruits and vegetables, even under varying lighting, orientation, and background conditions.

This project aims to design and implement a CNN-based deep learning framework that can automatically classify images of fruits and vegetables. The system integrates essential stages of dataset preprocessing, feature extraction, model training, and performance evaluation. Publicly available datasets containing labeled fruit and vegetable images were utilized, ensuring diversity in terms of size, shape, and color variations. Preprocessing steps such as resizing, normalization, and augmentation were employed to enhance the robustness and generalization ability of the model.

The proposed system addresses the limitations of existing manual and rule-based approaches by offering scalability, consistency, and high accuracy. Furthermore, it establishes the foundation for potential deployment in agricultural inspection systems, smart retail platforms, and automated inventory management workflows. By harnessing deep learning, this research contributes to the growing adoption of AI in agriculture and highlights the role of intelligent automation in ensuring efficiency, food quality, and supply chain optimization.

## II.MATERIAL AND METHODS

The development of an automated fruit and vegetable recognition framework using Convolutional Neural Networks (CNNs) required a structured methodology that ensured robustness, reproducibility, and scalability. This section outlines the study design, dataset acquisition, preprocessing steps, exploratory data analysis (EDA), model development, training strategy, evaluation techniques, and system deployment approach. Each phase was carried out systematically, ensuring that insights from earlier stages informed the refinement of subsequent stages.

### Study Design

The research followed a supervised deep learning pipeline for multi-class image classification of fruits and vegetables. Unlike traditional approaches that rely on manual inspection or barcode-based identification, this study employed CNNs to extract discriminative features directly from image data. The pipeline consisted of:

1. Dataset collection from publicly available repositories.
2. Preprocessing of raw images to standardize size, scale, and quality.
3. Data augmentation to improve generalization and prevent overfitting.
4. Model development using a sequential CNN architecture.
5. Training and validation using appropriate hyperparameters.
6. Evaluation through quantitative metrics such as accuracy, precision, and recall.
7. Deployment considerations for real-world use cases in agriculture and retail.

This structured design ensured that the system was not only accurate in controlled environments but also scalable and adaptable for practical applications.

### Data Acquisition

The dataset was sourced from a publicly available **Kaggle fruit and vegetable image dataset**, which provides structured folders for training, validation, and testing. The dataset comprised thousands of labeled images representing multiple categories of fruits and vegetables. Images varied in orientation, size, lighting, and background, ensuring diversity for robust model training.

- **Training Set (70%)** – Used to learn feature representations.
- **Validation Set (15%)** – Used to tune hyperparameters and implement early stopping.
- **Test Set (15%)** – Used for unbiased performance evaluation.

This distribution aligned with best practices in machine learning, providing sufficient samples for both learning and generalization assessment.

### Data Preprocessing

Image preprocessing was a critical step to enhance the dataset's quality and make it suitable for CNN-based training. The following techniques were applied:

1. **Resizing and Normalization**
   - All images were resized to **150 × 150 pixels** to maintain consistency.
   - Pixel intensity values were normalized to a [0,1] range, ensuring uniform scaling and stable training.

2. **Data Augmentation**
   - To artificially expand the dataset and increase robustness, augmentation techniques were applied:
     - Random rotations
     - Horizontal and vertical flipping
     - Zooming and shearing
     - Brightness adjustments
   - These transformations ensured that the model could generalize to unseen variations in fruit and vegetable appearances.

3. **Class Balance Adjustment**
   - Augmentation was also applied selectively to underrepresented categories, mitigating potential class imbalance.

Through these steps, the dataset became standardized, balanced, and representative of real-world conditions.

### Exploratory Data Analysis (EDA)

Before model development, exploratory analysis was performed to understand dataset characteristics.

- **Category Distribution**: Histograms revealed the number of samples per class, highlighting imbalance in certain fruit categories.
- **Color and Texture Features**: Analysis confirmed that fruits and vegetables exhibited distinctive color and surface patterns, suitable for CNN-based classification.
- **Sample Visualization**: Random samples were visualized to assess variability in background and lighting, reaffirming the need for augmentation.

EDA validated that the dataset contained sufficient discriminative features for CNN-based learning.

**Model Development**
A **Sequential CNN model** was designed using TensorFlow and Keras. The architecture included:

- Multiple **convolutional layers** for feature extraction.
- **Pooling layers** to reduce spatial dimensions and computation.
- **Dropout layers** to minimize overfitting.
- **Fully connected dense layers** for classification. Final **softmax output layer** for multi-class predictions.

This architecture was selected because of its proven effectiveness in image classification tasks while maintaining computational efficiency.

**Training Strategy and Hyperparameter Tuning**
Training was conducted on GPU-enabled hardware for faster convergence. Key parameters included:
- **Optimizer**: Adam optimizer with learning rate = 0.001.
- **Loss Function**: Categorical Cross-Entropy.
- **Batch Size**: 32.
- **Epochs**: 50, with early stopping after no improvement for 5 consecutive epochs.
- **Learning Rate Scheduling**: Reduce-on-plateau strategy for adaptive learning.

Hyperparameter tuning was performed using grid search to optimize dropout rates, learning rate, and batch size.

**Evaluation Metrics**
To ensure a holistic evaluation of the model, the following metrics were applied:
- **Accuracy** – Percentage of correctly classified samples.
- **Precision** – Proportion of predicted positives that are true positives.
- **Recall (Sensitivity)** – Proportion of actual positives correctly classified.
- **F1-Score** – Harmonic mean of precision and recall.
- **Confusion Matrix** – Breakdown of predictions across categories.

These metrics provided insights into both overall performance and class-specific strengths and weaknesses.

**System Deployment**
The trained model was integrated into a deployment-ready framework. A lightweight interface was designed using **Streamlit**, allowing users to upload fruit or vegetable images and obtain real-time predictions.

**Features of the deployment framework included:**
- **Prediction Results** – Output probabilities for each category.
- **Visualization Tools** – Graphs for training/validation accuracy and confusion matrix results.
- **Scalability** – Compatibility with web or mobile applications for agricultural or retail usage.

This ensured that the research could move beyond theoretical analysis and serve practical, real-world applications.

## III.RESULT

**A. Data Preprocessing Outcomes**
Preprocessing significantly improved dataset uniformity and quality, making it suitable for deep learning training.
**1. Resizing & Normalization –** All images were standardized to 150 × 150 pixels, ensuring input compatibility across the model. Normalizing pixel values to the range [0,1] stabilized training and prevented dominance of higher-intensity values.
**2. Noise Reduction** – Although most images were clean, preprocessing filters minimized background variation, ensuring consistency across categories.
**3. Data Augmentation** – Techniques such as rotation, horizontal/vertical flipping, zooming, and shearing expanded the dataset size by nearly threefold, improving model robustness against orientation, lighting, and positional variations.
**4. Class Balance** – Augmentation of underrepresented categories ensured better class balance, reducing the risk of bias toward majority classes.These steps created a high-quality, balanced dataset, which directly contributed to improved generalization during training.

**B. Model Training and Performance**
The CNN model was trained using TensorFlow and Keras on a GPU-enabled system. Training converged within ~40 epochs, with early stopping applied to prevent overfitting.
**Key performance metrics on the test dataset:**
- Accuracy: 95.8%
- Precision: 0.94
- Recall: 0.93
- F1-Score: 0.935

The high accuracy and F1-score indicate that the CNN was effective in learning discriminative features across multiple fruit and vegetable categories.

**C. Visualization and Graphical Analysis**

**Several graphical outputs were generated to evaluate model performance:**

**1. Training vs. Validation Accuracy** – Accuracy improved steadily across epochs, with minimal divergence between training and validation curves.

**2. Loss Curves –** Both training and validation loss decreased consistently, plateauing near convergence.

**3. Confusion Matrix –** The confusion matrix revealed strong classification performance across most categories. Misclassifications occurred mainly between visually similar items.

**4. ROC Curves (Class-wise) –** ROC curves indicated high discriminative ability across categories, with most classes achieving an AUC > 0.95.

**D. Error Analysis**

Despite strong performance, some misclassifications were observed:

- **False Positives –** Cases where one fruit was incorrectly labeled as another with similar shape or color.
- **False Negatives –** Images captured under poor lighting or with partial occlusion led to missed detections.
- **Complex Backgrounds –** Certain test samples contained cluttered or non-uniform backgrounds, which confused the model.

These errors suggest that further augmentation and background filtering techniques could enhance accuracy.

**E. Feature Importance and Interpretability**

To enhance transparency, Grad-CAM (Gradient-weighted Class Activation Mapping) was applied:

- For apples, the model highlighted round contours and red surface textures.
- For leafy vegetables like spinach, the model focused on leaf venation and green intensity patterns.
- Misclassifications often revealed attention shifts to background areas rather than the fruit/vegetable itself.

This interpretability confirmed that the CNN relied on biologically meaningful features rather than arbitrary patterns.

**F. Comparative Results**

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Logistic Regression | 82.3% | 0.79 | 0.76 | 0.775 |
| VGG16 (Transfer Learning) | 92.7% | 0.90 | 0.89 | 0.895 |
| ResNet50 (Transfer Learning) | 94.5% | 0.93 | 0.92 | 0.925 |
| Proposed CNN Model | 95.8% | 0.94 | 0.93 | 0.935 |

**G. Summary of Results**

1.Preprocessing improved dataset quality and corrected class imbalance.
2. The CNN achieved 95.8% accuracy, demonstrating robust classification capability.
3. Visualization confirmed effective convergence with minimal overfitting.
4. Error analysis revealed challenges with visually similar items and cluttered backgrounds.
5. Grad-CAM analysis verified that the CNN learned meaningful fruit/vegetable features.
6. Comparative analysis showed superior performance of the CNN over classical ML models and competitive results with transfer learning approaches.

## IV.DISCUSSION

**A. Comparative Insights**

The proposed CNN-based system for fruit and vegetable recognition demonstrates clear advantages over traditional manual classification methods and earlier computational approaches. Manual inspection is often subjective, labor-intensive, and inconsistent, while conventional rule-based systems lack adaptability to real-world variability. In contrast, the CNN model achieved an accuracy of 95.8%, significantly outperforming classical machine learning models such as Logistic Regression, which recorded 82.3%. Even when compared with deeper transfer learning models such as VGG16 and ResNet50, the proposed architecture showed competitive or superior results, balancing computational efficiency with accuracy. These findings validate the suitability of CNNs for agricultural and retail automation applications.

**B. Strengths of the Proposed System**

The strength of the system lies not only in its accuracy but also in its robustness and adaptability. Through preprocessing and augmentation, the model was trained to recognize produce under varying lighting conditions, orientations, and backgrounds. The use of Grad-CAM visualization enhanced interpretability, allowing users to understand the features contributing to model decisions. This transparency fosters trust in the system when integrated into agricultural inspection or retail environments. Furthermore, the relatively lightweight design ensures that the framework can be scaled for both industrial-level deployment and lightweight mobile applications, making it versatile for diverse use cases.

### C. Limitations of the Study

Despite strong results, the study presents certain limitations. The dataset, although augmented, remains constrained to specific classes of fruits and vegetables, which may not fully represent global diversity. Some misclassifications occurred in visually similar categories such as apples versus tomatoes, or cucumbers versus zucchinis, indicating the need for more refined feature extraction. Additionally, performance in cluttered or complex backgrounds was less reliable, suggesting the potential benefit of background segmentation techniques. Another limitation is computational dependency, as GPU support significantly reduced training time, which may pose a challenge in low-resource environments.

### D. Implications for Practice

The implications of this system for agriculture and retail are substantial. Automated recognition can streamline supply chain management, quality control, and retail checkout systems, reducing human labor and error rates. In agricultural settings, the system can support farmers by enabling rapid and objective classification of harvested produce, improving both quality assurance and market readiness. In retail environments, the model can be integrated into self-checkout systems, allowing customers to scan fruits and vegetables without barcodes. Furthermore, the transparent decision-making provided by Grad-CAM visualizations ensures that stakeholders maintain confidence in AI-assisted classification.

### E. Future Directions

Future research should focus on expanding the dataset to include a broader variety of fruits and vegetables from diverse regions, captured under real-world conditions. Incorporating advanced architectures such as ResNet101 or Vision Transformers could improve feature extraction and classification accuracy. Additionally, combining CNNs with object detection algorithms like YOLO or Faster R-CNN would allow not only classification but also localization of produce within complex backgrounds. Deployment on mobile and edge devices should be further optimized to enhance accessibility in rural and resource-constrained areas. Finally, integrating the system with supply chain management platforms could provide end-to-end automation from farm to retail.

### F. Summary of Discussion

In summary, the discussion highlights the effectiveness of CNNs in fruit and vegetable recognition, their strengths in accuracy, robustness, and interpretability, as well as the limitations concerning dataset diversity and background complexity. The system offers significant implications for agriculture and retail, enabling automation and improving efficiency. With further research and refinement, particularly in expanding datasets and enhancing real-time deployment, this approach can become a vital contribution to AI-driven food quality control and retail automation.

## V.CONCLUSION

This research has demonstrated the successful application of Convolutional Neural Networks (CNNs) for automated fruit and vegetable image recognition, addressing the limitations of traditional manual and rule-based classification methods. By employing systematic preprocessing, including resizing, normalization, and augmentation, coupled with a carefully designed CNN architecture, the system achieved a high accuracy of **95.8%** on the test dataset. These results validate the capability of CNN-based models to accurately capture discriminative features such as shape, color, and texture for reliable produce classification.

The study's findings underscore the transformative potential of deep learning in agriculture and retail automation. The proposed framework offers significant improvements in efficiency, scalability, and accuracy compared to manual sorting and classical machine learning approaches. The interpretability of model predictions using Grad-CAM further enhances transparency and trust, making the system practical for real-world integration.

However, the research also revealed certain limitations, particularly in handling visually similar categories and complex background environments. These limitations highlight opportunities for refinement, such as incorporating more diverse datasets, employing advanced architectures, and integrating background segmentation techniques.

Looking forward, the proposed system can be extended to include additional fruit and vegetable categories, real-time detection capabilities, and deployment on mobile or edge devices for practical use in both agricultural fields and retail environments. Such advancements will not only optimize food quality control and supply chain processes but also contribute to the broader adoption of intelligent AI-driven systems in agriculture and commerce.

In conclusion, this work establishes a strong foundation for future research in food recognition systems, bridging the gap between AI innovation and real-world agricultural and retail challenges. By advancing automated classification technologies, it paves the way toward sustainable, efficient, and scalable solutions that align with the growing global demand for intelligent food management.

## References


1. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
2. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1251–1258.
3. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018.
4. A. Mureșan and M. Oltean, "Fruit Recognition from Images Using Deep Learning," *Acta Universitatis Sapientiae, Informatica*, vol. 10, no. 1, pp. 26–42, 2018.
5. S. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, Apr. 2018.
6. H. Hasan, R. Abdul Rahman, and M. H. Marhaban, "Classification of fruits using convolutional neural network," in *Proc. IEEE Int. Conf. on Control System, Computing and Engineering (ICCSCE)*, 2018, pp. 242–247.
7. M. R. Khan et al., "Smart Farming Using Image Processing and Machine Learning," *IEEE Access*, vol. 7, pp. 170645–170661, 2019.
8. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
9. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv:1409.1556*, 2014.
10. A. 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, Jul. 2019.