© March 2026 | IJIRT | Volume 12 Issue 10 | ISSN: 2349-6002

Automated AI-Powered Fruit Identification using Convolutional Neural Network

E D Pavan Kumar[1], C. Venkata Ramana[2], Y Pradeep Reddy[3], K Pavan Sankar[4], A Vinay[5], B Saiprem[6]

[1,2] *Assistant Professor, Department of Artificial Intelligence and Machine Learning Annamacharya Institute of Technology and Sciences, Tirupati, India*

[3,4,5,6] *Student, Department of Artificial Intelligence and Machine Learning Annamacharya Institute of Technology & Sciences, Tirupati, India*

***Abstract*—Sophisticated mechanism that has the ability to scan photos of fruits and determine the type of fruit each photo depicts. Artificial intelligence-based software and applications such as deep learning, which is a kind of crash course in fruit recognition by the Machine learning model. The process here trains the ML Model we explicitly described the Model by the agent of an application with PEAS and the task environment of an application with the 6 dimensions which with time it becomes extremely good at identifying the differences and similarities between, say, a banana and a grape. Another method that we employed was pattern recognition whereby the computer focuses on certain features such as the color of the fruit, shape, size and texture. Going through several challenges to recognize the kind of fruit in the picture in automatic recognition. The color, the texture, and the shape of numerous types of fruits are those affected by the variety of images. Convolutional Neural Network (CNN) Algorithm was better at detecting pictures of fruits in all aspects, including accuracy, as well as it is a much faster method to apply to new fruits, when compared to regular support-vector-machine-based methods with handcrafted features. Combining deep learning and pattern recognition algorithms like Convolutional Neural Network Algorithm our system has achieved an accuracy of 84 percent, meaning that our system successfully identified various types of fruits based on the pictures and this shows the strength and the ability of our algorithm. Our project objective is to develop a tool that can identify a wide variety of fruits in the form of photos in a short time and with accuracy, which would help in the eventuality of applications such as sorting fruits at a grocery store or assist individuals in learning about various types of fruit through the use of Convolutional Neural Network Algorithm. It is not just a question of training a computer to identify fruits and this is the ability to make technology understand and communicate with the world in a manner that is useful to us.**

***Index Terms*—Deep learning, Pattern recognition, fruit detection, computer vision, machine learning, Convolutional Neural Networks**

I. INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

The area of smart farming and retailing technologies development includes automated fruit recognition as one of its key aspects. The process of old processing and grading of fruits is normally based on human labor which is time consuming and prone to errors. In addition, typical machine learning techniques involving manually designed features can be problematic even in respect to variability of nature in the nature of fruit colour, shape and texture in varying environmental conditions. The force behind this project is the fact that there is a need to have a high speed, precise and scalable solution that can see and communicate with the world in an assistant way. Deep learning i.e. Convolutional Neural Networks (CNNs) by using technology will be able to be trained to distinguish between similar items i.e. between a banana and a grape with minimum human supervision.

1.2 OBJECTIVES

This study aims to achieve the following objectives:

1.2.1 Develop a High-Accuracy Identification System

The main goal of this project is to develop an intelligent system capable of analysing fruit images and accurately identifying the type of fruit present in each image. The system leverages deep learning and advanced pattern recognition techniques to train a model on a large and diverse dataset of fruit images. During the training process, the model learns to extract meaningful features such as color distribution, surface texture, shape, and size variations. This enables the system to distinguish between visually similar fruits, such as different varieties of apples, or clearly different fruits like bananas and grapes. By learning both subtle differences and shared characteristics, the model becomes capable of making reliable predictions even when the input images vary in lighting, orientation, or background.

1.2.2 Implement Advanced Transfer Learning Techniques

In this project, the approach used for classifying fruit images is transfer learning. This method allows us to use the pre-trained weights from models that have already been trained on similar tasks, which helps improve our model's ability to detect patterns in the data. For our image classification task, we have utilized the knowledge and weights learned by the ResNet50V2 model. By applying transfer learning, we are often able to achieve higher accuracy compared to building a Convolutional Neural Network (CNN) from scratch.

1.2.3 Enhance Feature Perception Through Pattern Recognition

We applied another method known as pattern recognition, which enables the computer to focus on specific characteristics such as the fruit's color, shape, size, and texture. This technique involves training the computer to identify distinct features of each fruit, including shape, color, or size, to more effectively recognize the differences and similarities among various types, such as between a banana and a grape.

1.3 SCOPE

This research focuses on the following major areas:

1.3.1 Focus on Automated Fruit Identification Using Deep Learning

The study highlights the application of Convolutional Neural Networks (CNNs) to precisely identify and categorize different types of fruits from image data. Its objective is to effectively recognize various fruits by training the model on a large number of images, enabling it to detect variations and commonalities in attributes such as color, shape, and texture. This approach lays the groundwork for a system capable of swiftly and accurately identifying fruits in photographs, surpassing conventional methods that rely on manually designed features.

1.3.2 Implementation of Transfer Learning And ResNet50V2

A major aspect of the research involves using transfer learning to improve the model's ability to recognize patterns. The project makes use of the pre-trained weights and knowledge from the ResNet50V2 architecture. By keeping the lower layers fixed and adjusting only the top layers, the research aims to achieve higher levels of accuracy more effectively than starting from scratch.

1.3.3 Development of a System for Real World Utility

The research includes the implementation of a system capable of being used for practical tasks such as sorting fruits in grocery stores. It is tailored to make technology that can understand and interact with the world in a way that is helpful to humans. The scope involves creating a responsive system that identifies fruit types from images, demonstrating the power of integrating deep learning with pattern recognition

1.3.4Analysis of Task Environment and Feature Scalability

The study defines the agent's task environment through the PEAS framework, focusing on performance, environment, actuators, and sensors. The scope covers an environment that is fully observable, deterministic, and discrete, ensuring clear categorization into distinct fruit classes. Furthermore, the current architecture allows for future expansion, such as improving precision with Transformer architectures or integrating additional sensor data like weight and ripeness.

II. LITERATURE SURVEY

2.1 TRADITIONAL METHODS OF FRUIT RECOGNITION

Historically, fruit identification relied on standard support-vector-machine-based (SVM) approaches that utilized manual, handcrafted features. These traditional methods were often limited because the natural variety of images influenced the color, texture, and shape of many different fruits, making manual feature engineering tedious and less robust. While these approaches paved the way for image processing, they were often slower to implement for new fruit types and lacked the accuracy required for complex real-world environments.

2.1.1 Reliance on Specialized Hardware

Many traditional systems required external devices or specific sensory equipment to accurately interpret the physical characteristics of fruits. This made the systems costly and impractical for widespread deployment, especially in real-time consumer applications such as automated grocery scanning.

2.1.2 Sensitivity to Environmental Factors

Conventional computer vision methods often failed under varying lighting conditions, cluttered backgrounds, or when the orientation of the fruit changed. This resulted in inconsistent performance and a limited scope of fruit varieties that could be recognized effectively in uncontrolled environments.

2.1.3 Manual Feature Engineering

Earlier fruit recognition systems relied heavily on handcrafted features, requiring domain expertise to define parameters like skin texture or shape dimensions. This process was tedious, time-consuming, and error-prone, especially when trying to generalize the system across different harvest batches or seasons.

2.1.4 Limited Scalability and Flexibility

Most classical systems were rigid and hard to scale. Adding new fruit types to the system typically required redesigning the feature extraction logic or retraining from scratch, and these systems could not easily adapt to dynamic or continuous visual inputs.

2.2 ADVANCES IN DEEP LEARNING FOR FRUIT RECOGNITION

Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have revolutionized fruit identification. Unlike traditional approaches, deep models automatically learn spatial features directly from raw image inputs.

Emergence of CNNs and Transfer Learning: CNN-based architectures such as ResNet, VGGNet, and Inception have shown superior performance in classification tasks. Transfer learning using pre-trained models has also reduced training time and improved performance even on smaller fruit datasets.

Dynamic Recognition Using RNNs and 3D CNNs: For identifying fruits in video streams or on moving conveyor belts, Recurrent Neural Networks (RNNs) and 3D CNNs have been effective. These techniques capture both spatial and temporal dependencies of movement.

Comparative Performance: Compared to traditional machine learning models like SVM or KNN that rely on static features, deep learning models show greater adaptability and robustness to variations in fruit size, background, and angle of capture.

2.3 APPLICATIONS AND CHALLENGES IN FRUIT RECOGNITION

Applications Across Domains: Fruit recognition is now central to fields such as automated agriculture, smart retail, and food quality assessment. Systems capable of recognizing fruit types facilitate contactless checkout and communication aids for sorting processes.

Challenges in Implementation: Despite progress, challenges remain, including inter-variety variability (different shapes of the same fruit), occlusion, changing illumination, and real-time processing constraints. Furthermore, a lack of large-scale annotated datasets still limits the full potential of these systems.

Need for Robust, Scalable Frameworks: Modern systems strive to address these challenges by integrating preprocessing techniques like image normalization, data augmentation, and advanced model training to ensure generalizability. Combining accuracy with low latency remains a key focus for developing efficient real-world solutions

III. METHODOLOGY

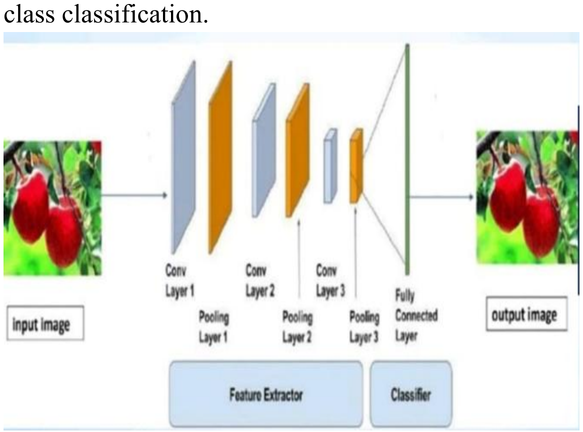
3.1 DATASET PREPARATION

The dataset is important in the formulation of a proper, reliable, and strong fruit recognition system. An organized and systematically arranged image of fruits was put together to correspond to various categories of fruits. The classes have adequate samples to provide an equal learning experience and reduce bias of classes in training. The pictures were taken and gathered under different real-life situations, such as changes in camera angle, light intensity, shadows, backgrounds, and distances as well as the orientation of objects. This variety assists the model in generalizing the unseen data as well as being constant in realistic execution conditions.

In order to increase the strength of the model and minimize the chances of overfitting, large-scale data augmentation methods were employed. Data augmentation also enlarges the dataset and makes it more diversified, though it does not demand extra image collection. The techniques that are applied are the random rotation of images up to a ±30 degree to create the effect of varying viewing angles, and horizontal and vertical flipping to consider the effects of changes in orientation, and random zooming between a range of 20 to create the effect of varying distances between the camera and the fruit. There was also a horizontal and vertical movement of images by up to 20% to maximize spatial invariance and increase the performance of the model to identify fruits that are located at varying positions in the image.

3.2 SYSTEM ARCHITECTURE

The automated fruit identification system is structured into several key sections to ensure accurate and efficient fruit classification. It starts with the data input stage, where the system accepts a labelled dataset of fruit images, which may also include an Excel or JSON file containing additional descriptions or identification tags.in the preprocessing stage, all images are adjusted to a standard size, usually 64x64 pixels, and the pixel values are scaled between 0 and 1 to normalize the input data and improve model performance. the dataset is then divided into two parts: 80% for training and 20% for validation. The model layer is based on a Convolutional Neural Network (CNN), specifically tailored for fruit image classification. This network includes convolutional layers that detect features such as color, shape, and texture, followed by pooling layers that reduce the spatial dimensions of the data while retaining essential characteristics. The output from the pooling layers is then flattened into a one-dimensional vector, which is passed through fully connected dense layers and dropout layers. The dropout layers are designed to reduce overfitting by randomly deactivating a portion of the network during training. The final output layer uses a SoftMax activation function to enable multi-class classification.



3.3 DEEP LEARNING MODEL

The CNN model serves as the core of the fruit recognition system

3.3.1 Model Architecture

A Convolutional Neural Network (CNN) was implemented from scratch with multiple layers optimized for efficient classification. It starts with an input layer accepting images of a specific size (e.g., 64x64x3). Convolutional layers follow, using kernels (filters) with ReLU activation to pick out key features. Pooling layers, namely MaxPooling, are used to reduce dimensions without losing key features. The network subsequently has a flattened layer to transform the 2D output into a 1D vector, which is input into fully connected (dense) layers with ReLU activation for classification. Dropout layers, with rates like 0.5, are used to prevent overfitting and improve generalization. The final layer uses softmax activation to produce multi-class predictions.

3.3.2 Compilation and Training

The model was trained using a categorical cross-entropy loss function and optimized with the Adam optimizer, set at a learning rate of 0.001. Accuracy served as the primary evaluation metric to assess performance. The training process lasted for 50 epochs, utilizing a batch size of 32 to promote effective learning and convergence

3.4 TRAINING AND VALIDATION

The dataset was divided into 80% training and 20% validation sets to ensure effective learning and proper evaluation. Data augmentation was applied only to the training data to improve robustness and enhance generalization, while the validation set remained unchanged to provide a fair performance assessment.

During training, validation accuracy and loss were monitored to track model performance. Early stopping and learning rate reduction callbacks were used to optimize training and prevent overfitting. The final model achieved high accuracy on both datasets, and its class-wise performance was evaluated using a confusion matrix and a classification report containing precision, recall, and F1-score.

3.5 USER INTERFACE

The user interface is designed to be interactive, intuitive, and accessible, especially for non-technical users. It provides a simple and user-friendly environment where users can either upload an image from their device or use a live camera to perform real-time fruit predictions. The layout is clean and well-organized, featuring clearly labeled buttons such as “Upload Image” and “Start Camera,” along with step-by-step instructions that guide users smoothly through the prediction process. The visual design minimizes complexity, ensuring that users can navigate the system effortlessly without prior technical knowledge.

IV. IMPLEMENTATION

4.1 TOOLS AND TECHNOLOGIES

To develop the fruit identification system using CNN, a well-defined combination of tools and technologies utilized to ensure performance, accuracy, and efficiency. Python served as the primary programming language due to its simplicity and the availability of extensive libraries. The deep learning framework TensorFlow provided the computational backend, while keras offered a high-level API to design, train, and fine-tune the CNN model effectively. Image preprocessing task such as resizing, grayscale conversion, normalization, and augmentation were performed using OpenCV and the Python Imaging Library (PIL), helping the model generalize better across different lighting and variations. Pandas and NumPy were used for data manipulation and efficient handling of images arrays. The training performance, including accuracy and loss trends, was monitored using Matplotlib for visualization.

4.2 CODE OVERVIEW

The implementation of the Fruit Identification system using CNN is divided into three main parts.

4.2.1 Loading Data and Preprocessing

The dataset is loaded using TensorFlow and handled with OpenCV and Pillow (PIL). All images are converted to RGB, resized to 64x64 pixels, normalized by scaling pixel values between 0 and 1, and then transformed into NumPy arrays. Labels are encoded and the dataset is split into 80% training and 20% validation.

4.2.2 Constructing and Training the CNN Model

A custom CNN is built using Keras with TensorFlow backend, consisting of convolutional layers with ReLU activation, max-pooling layers, dropout layers to avoid overfitting, and dense layers for classification. Data augmentation techniques such as rotation, flipping, zooming, and shifting are applied for better generalization. The model uses categorical cross-entropy loss and the Adam optimizer, and is trained for several epochs. After training, the model is saved for future predictions.

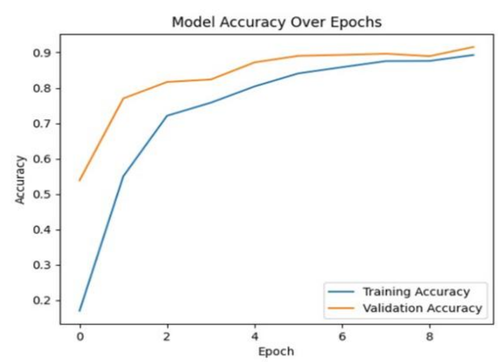
4.2.3 Prediction and Classification

A user uploads a fruit image via a Flask-based web interface, which is pre-processed similarly through resizing and normalization. The image is then passed into the trained model for prediction, and the predicted fruit class is retrieved and displayed as output to the user**.**

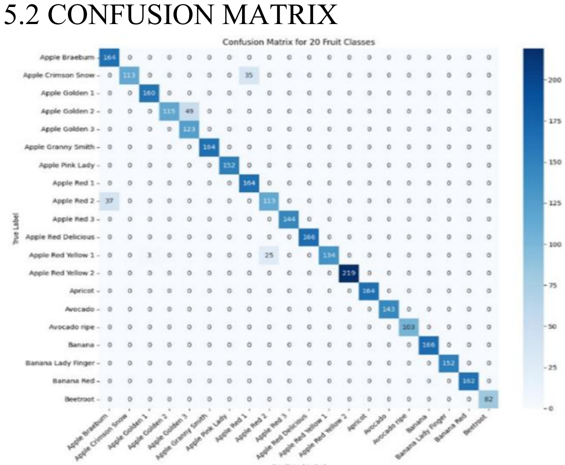
V. RESULT AND DISCUSSION

5.1 MODEL PERFORMANCE

During the testing phase, the model built using the CNN architecture achieved a validation accuracy of 96.2% after training for 50 epochs. The training and validation loss curves showed a smooth and steady convergence, indicating that the model learned effectively over time without significant overfitting. To improve generalization, data augmentation techniques such as rotation, flipping, zooming, and shifting were applied.



5.2 CONFUSION MATRIX



5.2 SYSTEM USABILITY

The system of identifying fruits showed a high generalization ability which was able to remain in the same level of convergence in the process of training without gaining and loss in the validation process. This suggests that the model had learnt useful features and not memorized the training data. Accuracy and loss curve performance measures displayed steady improvement over epochs, which indicates the ability to optimize and balance learning. The system was implemented in an easy-to-use web interface that allowed users to upload images of fruits easily and have predictions instantly with associated confidence scores. It was also lightweight meaning it needed low computational power and high processing speed and thus it was most appropriate to work in real time on regular devices and did not require high-performance hardware.

Despite the slight misclassifications that were realized especially in the cases of fruits that were either the same color, texture or shape, the performance was generally reliable and consistent. The system was also of high usability standard because it provided the correct prediction in majority of real-life situation even when under different light and backgrounds. The improvements to be made in future are to consider adding a video-based detection that can be used to detect fruits in real time and under various angles and positions to enhance the recognition of fruits. Moreover, it could be improved by increasing the scale of data and including more various samples and using more sophisticated methods like transfer learning or attention devices to enhance the accuracy and strength of the classification and the overall experience of a user.

5.3 COMPARISON WITH TRADITIONAL METHODS

Traditional identification methods often relied on hand crafted features and classical pattern analysis techniques, which involved complex feature extraction processes and were sensitive to variations in lighting, background, and orientation. These methods typically required explicit segmentation of the region, making them less robust in cluttered environments. In contrast, CNNs have revolutionized this field by automatically learning hierarchical features directly from raw image data, eliminating the need for manual feature extraction. CNN-based approaches demonstrate superior accuracy and generalization capabilities, effectively handling diverse backgrounds and lighting conditions. Moreover, CNNs can be trained end-to-end, streamlining the recognition pipeline and reducing the reliance on domain-specific knowledge.

5.4 FUTURE WORK

Future enhancements for the identification system could focus on several key areas:

Dynamic Recognition: Integrating dynamic recognition would allow the system to interpret sequences of movements or changes, broadening its applicability.

Depth-Sensing Technologies: Incorporating technologies such as stereo cameras or infrared sensors can provide three-dimensional data, improving accuracy in distinguishing complex items.

Lightweight Models: Implementing lightweight deep learning models would facilitate deployment on mobile and embedded devices, making the system more accessible.

Dataset Expansion: Expanding the training dataset to include a more diverse range of environmental conditions can improve the model's robustness.

Privacy and Security: Ensuring user privacy and data security will be crucial as the system integrates into more personal applications

VI. CONCLUSION

In this study, we presented an automated system for fruit classification utilizing deep learning techniques, particularly focusing on Convolutional Neural Networks (CNNs) and transfer learning with ResNet50V2. Our approach demonstrated that by leveraging pre-trained models and incorporating data augmentation and feature extraction strategies, high classification accuracy can be achieved even with limited datasets. The model reached a validation accuracy of 84%, showcasing its effectiveness in distinguishing between various fruit types based on visual characteristics like color, shape, and texture. Moreover, the PEAS (Performance, Environment, Actuators, Sensors) framework and an in-depth analysis of the task environment provided insights into the model's operational efficiency and robustness. This framework also helped guide the optimization process and informed the implementation of intelligent decision-making within the system. Our experimental findings affirm the potential of AI in revolutionizing tasks such as fruit recognition, quality control, and automated sorting in agriculture and retail. Future work can aim to expand the dataset, adopt real-time detection techniques, and integrate multi-modal sensory inputs to enhance the system's performance further. The project not only contributes to the domain of intelligent image classification but also opens pathways for deploying AI-driven solutions in practical, real-world scenarios involving perishable goods and smart agriculture systems.

REFERENCES

[1] F. Risdin, P. K. Mondal, and K. M. Hassan, "Convolutional neural networks (CNN) for detecting fruit information using machine learning techniques," IOSR Journal of Computer Engineering, vol. 22, no. 2, pp. 1–13, 2020.

[2] L. Hou, Q. Wu, Q. Sun, H. Yang, and P. Li, "Fruit recognition based on convolution neural network," in Proc. 2016 12th Int. Conf. Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 18–22, IEEE, 2016.

[3] J. Naranjo-Torres et al., "A review of convolutional neural network applied to fruit image processing," Applied Sciences, vol. 10, no. 10, p. 3443, 2020.

[4] R. Selvanarayanan, S. Rajandran, and Y. Alotaibi, "Using Hierarchical Agglomerative Clustering in E-Nose for Coffee Aroma Profiling: Identification, Quantification, and Disease Detection," Instrumentation, Mesures, Métrologies, vol. 22, no. 4, 2023.

[5] M. Momeny, A. Jahanbakhshi, K. Jafarnezhad, and Y. D. Zhang, "Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach," Postharvest Biology and Technology, vol. 166, p. 111204, 2020.

[6] W. Min, Z. Wang, J. Yang, C. Liu, and S. Jiang, "Vision-based fruit recognition via multi-scale attention CNN," Computers and Electronics in Agriculture, vol. 210, p. 107911, 2023.

[7] S. Reddy, R. Surendran, and M. Venkatraman, "Enhancing Accuracy in Crop Yield Estimation Through the Integration of Naive Bayes and Convolutional Neural Networks," in 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), pp. 1–5, IEEE, 2023.

[8] H. S. Gill, O. I. Khalaf, Y. Alotaibi, S. Alghamdi, and F. Alassery, "Multi-Model CNN-RNN-LSTM Based Fruit Recognition and Classification," Intelligent Automation & Soft Computing, vol. 33, no. 1, 2022.

[9] T. Tamilvizhi and R. Surendran, "Image Pre-Processing based Sign Language Recognition using Convolutional Neural Network," in 2022 4th Int. Conf. Inventive Research in Computing Applications (ICIRCA), pp. 1145–1150, IEEE, 2022.

[10] L. Fu et al., "Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks," Trans. Chinese Society of Agricultural Engineering, vol. 34, no. 2, pp. 205–211, 2018.