



ANDROID-BASED CLASSIFICATION MODEL ON MANGO FRUIT DATASET USING DEEP LEARNING

¹Dr. K.K.Tripathi, ²Santosh Bhosale, ³Pritam Dhumal, ⁴Pradip Kayande

¹Assistant Professor, ^{2,3,4}UG Student

^{1,2,3,4}Dept. of Computer Engineering,

^{1,2,3,4}Shivajirao S. Jondhale College of Engineering, Dombivli East, Mumbai University.

Abstract: This project introduces an Android-based mobile application aimed at fulfilling the crucial need for precise mango variety identification. Understanding the significance of mango varieties in both agriculture and consumer preferences, the app employs Deep learning algorithms such as Convolutional Neural Networks (CNN) to revolutionize the accuracy and efficiency of this process. Traditional methods, primarily reliant on manual assessments and basic image recognition, suffer from limitations such as time consumption and reduced accuracy. This project overcomes these challenges by harnessing CNN technology, facilitating automated and accurate mango variety identification. The CNN model, trained on an extensive dataset of mango varieties, evaluates visual attributes from input images, resulting in a significant enhancement in accuracy. Users can capture a mango image for quick and reliable variety identification. Additionally, real-time access to a comprehensive mango variety database enriches the user experience, providing valuable information on taste, appearance, and cultivation practices for each recognized variety. In conclusion, the Android-based Mango Variety Identification Using CNN app offers an efficient and highly accurate solution, benefiting farmers, vendors, and consumers alike, facilitating informed decisions, and fostering appreciation for this beloved fruit.

IndexTerms - CNN, Mango Classification, Deep Learning.

1. Introduction

There exists a vast array of mango species worldwide, numbering in the thousands. While seasoned taxonomy experts can readily distinguish between these species, the task proves challenging for most individuals. Typically, people seek assistance from specialists, consult mango guidebooks, or scour the internet for relevant information through keyword searches to identify mango varieties or their distinguishing characteristics. One effective approach to this identification conundrum involves the classification of mango images, leveraging the widespread use of digital cameras and mobile phones. However, despite the potential utility and significance of mango classification, several hurdles impede its realization. Unlike classifying more straightforward objects like cars or desks, mango classification presents unique difficulties due to the considerable similarity between species and the considerable variation within each species, often compounded by environmental factors such as weather conditions. Moreover, distinguishing between mango varieties can be challenging, especially when differences hinge on subtle variations in aspect ratio and angle. Thus, the adoption of Convolutional Neural Networks (CNNs) emerges as a promising solution for accurately classifying mangoes.

2. Advantages

ADVANTAGES AND DISADVANTAGES

1. **Accuracy:** CNN algorithms typically deliver high levels of accuracy in classification tasks, ensuring reliable results in mango classification.
2. **Automation:** CNN-based mango classification systems automate the sorting process, reducing the need for manual labor and increasing efficiency.
3. **Speed:** CNNs can process large volumes of mango images rapidly, allowing for quick sorting and grading, which is crucial for industries with high production demands.
4. **Scalability:** CNN-based classification systems are scalable and adaptable to different production scales, making them suitable for both small-scale farms and large commercial operations.
5. **Consistency:** By standardizing the classification criteria, CNNs ensure consistency in mango grading, reducing variability and improving overall product quality.

2.1 Disadvantages

1. **Data Dependency:** CNN performance heavily relies on the quality and quantity of training data. Insufficient or biased training data may lead to inaccurate classification results.
2. **Complexity:** Developing and fine-tuning CNN models for mango classification requires expertise in machine learning and computational resources, which can be challenging and costly.
3. **Sensitivity to Variations:** CNNs may struggle with classifying mangoes with subtle variations in appearance or shape, leading to misclassification errors.
4. **Hardware Requirements:** Implementing CNN-based classification systems may necessitate powerful hardware infrastructure, particularly for real-time applications or processing large datasets.
5. **Interpretability:** CNN models are often considered as "black-box" systems, making it difficult to interpret the underlying decision-making process, which can be a limitation in certain contexts, such as regulatory compliance or quality assurance.

3. LITERATURE SURVEY

In a comparative analysis of various research papers on fruit classification methodologies, the following observations were made: In "A Novel Classification Model of Date Fruit Dataset Using Deep Transfer Learning" by Alsirhani et al. (2023), the utilization of deep transfer learning algorithms like DenseNet201 facilitates the classification of date fruits, though the accuracy remains below 97%. Similarly, "Date fruit classification and sorting system using Artificial Intelligence" by Vikraman et al. (2022) employs different transfer learning algorithms to achieve good accuracy in date fruit classification, albeit with some instances of lower accuracy. Meanwhile, "Fruit Image Classification using Deep Learning" by Gill and Khehra (publication year not specified) utilizes a hybrid classifier (CNN, RNN & LSTM) to maintain accuracy levels between 90%-95%, albeit with relatively lower accuracy. In "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers" by Sakib et al. (2022), the CNN algorithm achieves a remarkable 99% accuracy for some fruits but falls short in accuracy for others. Lastly, "Identification and Classification of Mango Fruits Using Image Processing" by Sahu and Dewangan (2017) relies on a clustering algorithm (k-means) for mango fruit classification, providing accurate and consistent results, although occasionally leading to incorrect identifications.

Summary of Literature Survey:

1. "A Novel Classification Model of Date Fruit Dataset Using Deep Transfer Learning" by Amjad Alsirhani, Muhammad Hameed Siddiqi, Ayman Mohamed Mostafa, Mohamed Ezz, and Alshimaa Abdelraof Mahmoud (2023)
 - Methodology: Classification of date fruits utilizing deep transfer learning algorithms like DenseNet201.
 - Advantages: Easily identifies date fruits.
 - Disadvantage: Accuracy is below 97%.
2. "Date fruit classification and sorting system using Artificial Intelligence" by Bindu Puthentharayil Vikraman, Vanitha Mahadevan, Dua Al Hashmi, Azhar Al Harrasi (2022)
 - Methodology: Classify date fruits using different transfer learning algorithms sorted in accuracy.
 - Advantages: Different transfer learning algorithms yield good accuracy.
 - Disadvantage: Less accuracy.
3. "Fruit Image Classification using Deep Learning" by Harmandeep Singh Gill, Baljit Singh Khehra
 - Methodology: Classification of fruits using a hybrid classifier (CNN, RNN & LSTM).
 - Advantages: Maintains accuracy level between 90%-95%.

- 0 Disadvantage: Relatively low accuracy.
- 4. Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers" by Shadman Sakib, Zahidun Ashrafi, Md. Abu Bakr Sidique (2022)
 - 0 Methodology: Classification of fruits using the CNN algorithm.
 - 0 Advantages: Achieved 99% accuracy for some fruits; identifies various fruits.
 - 0 Disadvantage: Lacks accuracy for some fruits.
- 5. Identification and Classification of Mango Fruits Using Image Processing" by Dameshwari Sahu and Chitesh Dewangan (2017)
 - 0 Methodology: Classification of mango fruits based on the clustering algorithm k-means.
 - 0 Advantages: Provides accurate and consistent results.
 - 0 Disadvantage: Sometimes leads to incorrect identification.

4. EXISTING SYSTEM

The existing systems for mango fruit classification, which predominantly rely on traditional image processing techniques and some early machine learning methods, exhibit several limitations that underscore the need for a more advanced approach using deep learning algorithms like CNN.

Traditional image processing has limitations in effectively handling the variability in mango fruit appearances. Handcrafted feature extraction methods struggle to accurately capture the nuances of mangoes, resulting in suboptimal classification performance. Existing systems cannot also adapt to changes in lighting conditions and background clutter, which is a critical limitation in agricultural settings where mango classification occurs in diverse environments.



5. PROPOSED SYSTEM

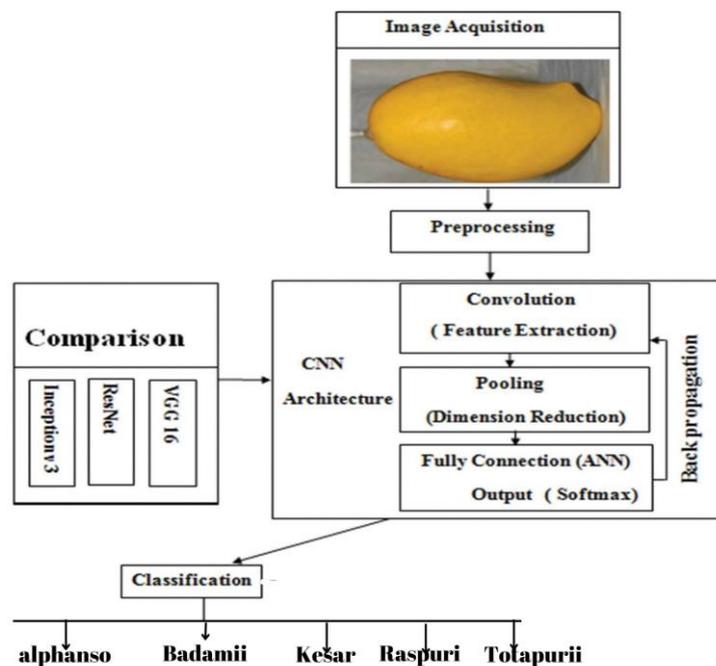


The proposed system employs Convolutional Neural Networks (CNN), a specialized deep learning algorithm tailored for image processing and recognition. Unlike other classification models, CNNs streamline preprocessing by autonomously learning hierarchical feature representations directly from raw input images.

A novel approach is introduced to revolutionize mango variety identification through an Android-based mobile application leveraging CNN as its fundamental algorithm. By integrating CNN, the system ensures precise identification of mango varieties from images, granting users immediate access to an extensive mango variety database. The subsequent sections detail the algorithm, analysis process, and framework of this pioneering solution.

A diverse dataset comprising mango images depicting various varieties is meticulously collected, refined, and annotated. Utilizing deep learning frameworks such as TensorFlow, the CNN model undergoes training on this dataset. Through multiple iterations, the model is fine-tuned to enhance its proficiency in recognizing mango varieties. Subsequently, rigorous testing and validation are conducted using an independent dataset to evaluate the model's accuracy and generalization. Finally, a mango variety database is constructed and seamlessly integrated into the application.

5.1 Architecture:



5.2 Framework

Graphical User Interface: Our Android-powered mobile app offers a user-friendly interface enabling users to snap images of mangoes and obtain diverse identification outcomes.

Preprocessing: Images captured undergo preprocessing to refine quality and eliminate distortions, ensuring precision in analysis.

CNN Analysis: Post preprocessing, images are fed into our Convolutional Neural Network (CNN) model. This model evaluates visual traits such as color, form, and texture to discern the type of mango..

Output Generation: Once identified, the system generates a report showcasing the recognized mango variety along with relevant details.

5.3 Algorithm

Step 1: Start.

Step 2: Prompt the user to upload an image of a mango. Step 3: Check if an image was uploaded.

Step 4: If no image was uploaded, return to step 2 to prompt the user to upload an image again. Step 5: If an image was uploaded, proceed with the following steps:

Step 5.1: Preprocess the uploaded image (resize, normalize, etc.).

Step 5.2: Apply a pre-trained Convolutional Neural Network (CNN) model specifically trained for mango classification to the pre-processed image.

Step 5.3: Check if the mango classification was successful.

Step 5.4: If a mango variety is identified, retrieve information about the mango variety from a pre-defined database. Step 5.5: Display the mango variety information to the user.

Step 5.6: If no mango variety is identified, display "Mango variety could not be identified". Step 6: End.

6. OBJECTIVE

Our system leverages cutting-edge deep learning techniques, powered by TensorFlow, To identify the variety of mangoes using images. By leveraging image recognition technology, the application seeks to offer users a seamless platform for distinguishing between different mango varieties based on their distinct visual attributes. This initiative intends to empower farmers, consumers, and

mango enthusiasts by providing them with a convenient tool to enhance mango-related activities, including identification, and consumer education, and promoting appreciation for the rich diversity of Indian mangoes.

7. APPLICATION

- a. **Agricultural Assistance:** Farmers and agricultural experts can use the app to quickly identify different mango varieties in orchards. This can help in monitoring and managing crops more effectively, as well as in planning harvest and distribution schedules.
- b. **Consumer Education:** Consumers shopping for mangoes can use the app to learn about different varieties and their characteristics, helping them make more informed purchasing decisions based on flavor preferences and other factors.
- c. **Food Industry Quality Control:** Food processing and manufacturing companies can utilize the app to ensure the quality and consistency of mangoes used in their products. It can assist in sorting mangoes based on variety, size, and ripeness, improving the overall quality of processed mango products.
- d. **Research and Conservation:** Researchers and conservationists studying mango biodiversity can use the app to quickly identify and catalog different mango varieties in the wild or in collections. This can aid in conservation efforts and genetic studies aimed at preserving and enhancing mango diversity.
- e. **Educational Tool:** The app can serve as an educational tool for students and enthusiasts interested in botany, agriculture, or machine learning. Users can learn about different mango varieties and gain insights into the application of deep learning algorithms for image classification.
- f. **Tourism and Culinary Exploration:** In regions known for mango cultivation, tourists and food enthusiasts can use the app to explore and discover local mango varieties. It can enhance culinary experiences by helping users identify unique mango flavors and dishes made with specific mango varieties.
- g. **Export and Import Industry:** Exporters and importers of mangoes can use the app to verify the authenticity and classification of mango shipments, ensuring compliance with trade regulations and standards.

8. RESEARCH METHODOLOGY

Once the Convolutional Neural Network (CNN) is trained to recognize a specific feature, it can identify that feature regardless of its position within the image. This section offers a succinct overview of CNN's functioning. Initially, the image undergoes segmentation into overlapping smaller image tiles by employing a sliding window approach. Each of these tiles is then fed into a small neural network for processing. The outcomes from processing each tile are organized in a grid structure, resembling the original image layout, thus forming a matrix containing pertinent feature information. This convolution process, depicted in Fig. 1, is followed by down-sampling, often performed through the max pooling algorithm, which selects the maximum value from each sub-region of a rectangular area, effectively reducing dimensionality. Eventually, the reduced array is converted into a feature vector and inputted into a fully connected neural network. Each feature undergoes a voting mechanism, and the one with the highest votes and lowest error rate is classified as the specific mango class. In practical scenarios, CNN is trained on a substantial dataset, and the trained CNN weights are utilized either for initialization or as a fixed feature extractor for the targeted task.

The methodology section delineates the plan and approach employed in the study. This encompasses the study's scope, sample selection, data sources, variables under study, and the analytical framework. The details are expounded upon as follows:

Working Of CNN

Convolutional Neural Networks (CNNs) are specifically designed to detect visual patterns directly from pixel images with minimal preprocessing. This architectural design constitutes a unique form of artificial neural networks, primarily composed of convolutional layers and pooling layers, both of which are relatively straightforward components.

1. Data Collection: The process involves gathering a diverse dataset of mango images, encompassing different varieties, ripeness stages, and environmental conditions. This is achieved through the utilization of online repositories, databases, and image collections. Ensuring proper labeling and categorization of the dataset is essential.

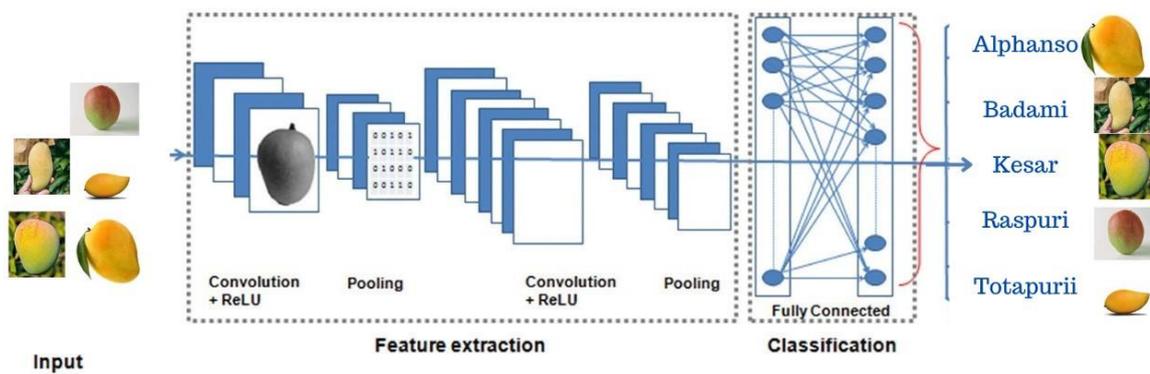
2. Preprocessing: Techniques such as resizing, normalization, and augmentation are applied to enhance model performance. Data cleaning is also conducted to eliminate noise or irrelevant images.

3. Model Development: This stage involves selecting an appropriate CNN architecture tailored for mango dataset identification. The CNN model is then trained on the prepared dataset using frameworks like TensorFlow or PyTorch. Fine-tuning is performed to enhance accuracy and generalization.

4. Mobile App Development: An Android-based mobile application is developed for real-time mango image identification. The trained CNN model is integrated into the mobile app, accompanied by the design of an intuitive user interface for seamless interaction. Features such as image capture, processing, and result display are implemented.

5. Model Evaluation: The performance of the trained model is evaluated on a held-out test set that was not used during training. Various evaluation metrics, including accuracy, precision, recall, and F1-score, are computed to gauge the model's effectiveness in distinguishing between real and AI-generated images.

6. Performance Analysis: The evaluation metrics and visualizations such as confusion matrices and ROC curves are analyzed to gain insights into the model's performance. This analysis aids in identifying areas for improvement and fine-tuning the model's hyperparameters to achieve better accuracy and robustness in mango variety identification.



Three-layer CNN architecture consisting of convolutional + rectified linear unit (ReLU), pooling, and classification layers.

Categories of CNN layers:

- 1. Input Layer:** The input layer receives raw images with a width and height of 32 each, and a depth of 3 representing RGB channels.
- 2. Convolution Layer:** In this layer, the dot product is computed between filters and patches of the input image. For example, if there are 8 filters in a layer, the output will be of dimensions $32*32*8$.
- 3. Activation Function Layer:** This layer applies activation functions such as sigmoid or ReLU to introduce non-linearity into the network.
- 4. Pooling Layer:** Intermittently interleaved in convolutional networks, the pooling layer reduces the size of the volume, thereby conserving memory and preventing overfitting.
- 5. Fully-Connected Layer:** Taking the output of the preceding layer as input, this layer generates a 1D array whose size corresponds to the number of classes, facilitating classification.

9. IMPLEMENTATION

It involves data collection and training , testing.

Data Collection:

The Mango Varieties Dataset is a collection of images representing five different types of mangoes: Alphonso, Badami, Kesar, Raspuri, and Totapuri. Each variety is a distinct class in the dataset. This dataset is curated for the purpose of mango variety classification tasks.

Key Features:

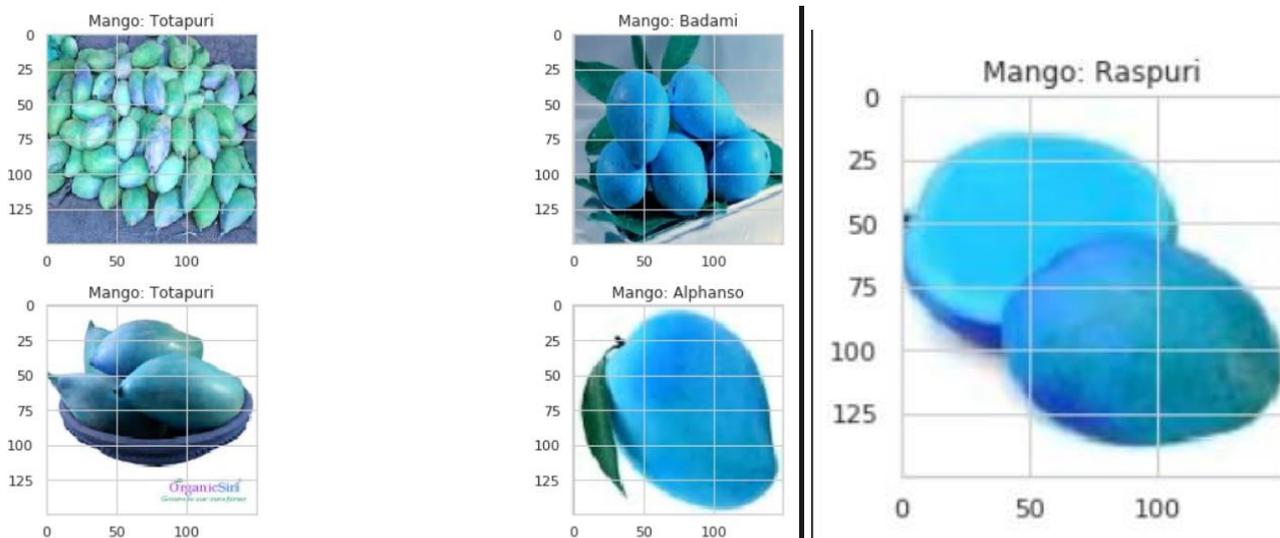
Class Labels: The dataset comprises five classes, each corresponding to a specific mango variety.

Image Count: The dataset contains a total of 250 images, with 50 images per mango variety class.

Image Diversity: Images in the dataset capture various aspects of mangoes, including their shape, color, texture, and size.

image Quality: High-quality images are included in the dataset, ensuring clarity and detail necessary for accurate classification.

Balanced Distribution: The dataset maintains a balanced distribution of images across all classes, with an equal number of samples per class.



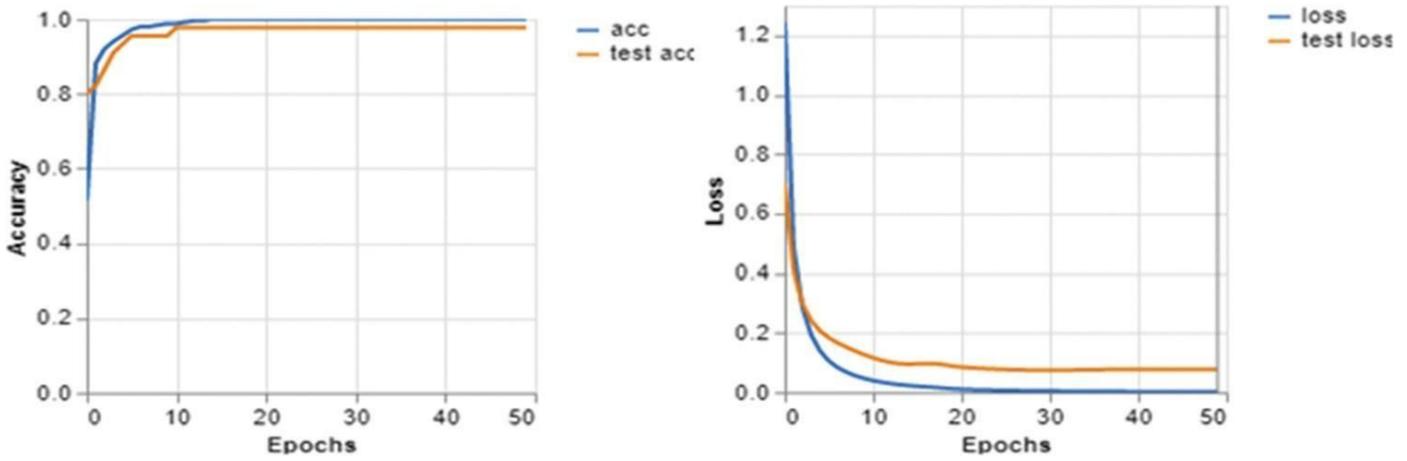
Class Id	Class Type	Number of Training Images	Number of Testing Images
1	Alphonso	150	100
2	Badami	150	100
3	Kesar	150	100
4	Raspuri	150	100
5	Totapuri	150	100

Training & Testing:

The CNN mango classification algorithm progresses by first training the model with gathered mango images and adjusting its parameters to minimize errors in classification through backpropagation. Validation using a separate dataset confirms the model's accuracy and fine-tunes its parameters if necessary. Once validated, the trained CNN model stands ready for mango classification tasks. When presented with a new mango image, the algorithm uses the trained model to predict the depicted mango variety. Data augmentation, transfer learning, and model ensembling are employed to bolster the model's robustness and generalization capabilities. Through this process, the algorithm reliably and accurately identifies mango varieties from images, contributing to efficient agricultural practices and food industry operations.



Test accuracy and test loss achieved at 50 epoch with 32 batch size for Mango cultivar classification.



Test accuracy and test loss achieved for Mango Classification.

10. RESULTS



11. Conclusion

In our final year project, we endeavored to pioneer an innovative mango fruit classification system using Convolutional Neural Networks (CNNs), aiming to address the urgent need for automation in the mango industry and provide a dependable method for categorizing mangoes based on their physical characteristics. By conducting an extensive literature review, we observed a significant shift toward deep learning techniques, particularly CNNs, in image-based mango classification approaches, marking a substantial advancement over traditional machine learning and image processing methods. Our project successfully developed a user-friendly software system capable of real-time mango classification, showcasing the potential of deep learning in practical agricultural settings and achieving remarkable accuracy and adaptability in the classification process. As agriculture grapples with escalating demand and supply chain intricacies, our project represents a stride toward optimizing the mango industry, facilitating enhanced quality control, efficient sorting, and informed decision-making for farmers, distributors, and consumers alike. The integration of deep learning technology in agriculture holds promise for revolutionizing the classification and management of agricultural products, signaling a future where advanced technologies bring benefits to both industry stakeholders and consumers.

REFERENCES

- [1] Al-Yahyai R., and Khan M.M. (2015) Date Palm Status and Perspective in Oman. In: Al-KhariJ., Jain S., Johnson D. (eds) Date Palm Genetic Resources and Utilization. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9707-8_6
- [2] AbdulRazak, N.A. (2010). Economics Of Date Palm Agriculture in The Sultanate Of Oman, Current Situation And Future Prospects. *Acta Hortic.* 882, 137-146 DOI: 10.17660/ActaHortic.2010.882.15
- [3] H. Altaheri, M. Alsulaiman and G. Muhammad, "Date Fruit Classification for RoboticHarvesting in a Natural Environment Using Deep Learning," in *IEEE Access*, vol. 7, pp. 117115-117133, 2019, doi: 10.1109/ACCESS.2019.2936536.
- [4] Al-Yahyai R, Manickavasagan A. An overview of date palm production. In: Manickavasagan, Essa MM, Sukumar E, editors. *Dates: production, processing, food, and medicinal values*. New York: CRC Press; 2012, p. 3-12.
- [5] Aisha Hakami and Mohammed Arif (2019), "Automatic Inspection of the external quality of the date fruit", *ScienceDirect,Procedia Computer Science*, 2019.
- [6] Rashid Al-Yahyai, and Sulaiman Al-Khanjari (2008), "Biodiversity of Date palm in the Sultanate of Oman", *African Journal of Agricultural Research*, Vol. 3 (6), pp. 389-395, June 2008.
- [7] May Z, and Amaran M.H, (2011, "Automated Oil Palm Fruit Grading System using Artificial Intelligence", *IJENS*, Vol: 11 No: 03.
- [8] Abi sen Adnan et al.(2020), "A classification Algorithm for date fruits", *Proceedings of the 14th INDIACom; INDIACom-2020; IEEE Conference ID: 49435, DOI:10.23919/INDIACom49435.2020.9083706*.