2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC) | 979-8-3503-7651-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICEC62498.2024.10808883

Agricultural Practices: Deep Learning for Vegetable Classification

Ruchitha Jillelamudi *, Sai suchitra Tarimella , Indhu Kotha, Niharika Peruri, Sujatha Kamepalli, Anjana Chekuri

School of Computing, VFSTR Deemed to be University, Guntur, India

**ruchithaj458@gmail.com*

***Abstract*— The agricultural industry is undergoing significant advancements, yet traditional methods for classifying vegetables by attributes such as size, color, texture, and freshness remain a challenge for farmers and vendors. This paper explores the application of deep learning models to automate the classification and recognition of vegetables, aiming to enhance efficiency and innovation in agricultural practices. Deep learning-powered automated systems can handle large volumes of produce, optimizing processes like sorting, grading, and quality inspection, thereby accelerating these tasks. Additionally, such systems can improve self-checkout mechanisms in retail by accurately identifying and pricing vegetables, enhancing customer convenience and minimizing errors. Utilizing a specialized dataset of vegetable images, this study employs various deep learning models, incorporating preprocessing techniques like data augmentation and normalization. The effectiveness of these models, particularly a Convolutional Neural Network (CNN) and VGG16, is rigorously tested. The results demonstrate that the CNN model achieves superior accuracy at 92.69 percentage, compared to 87.37 percentage for VGG16, highlighting the potential of deep learning models in revolutionizing vegetable recognition and classification.**

Keywords—Vegetable Recognition and Classification; CNN; VGG-16; Agriculture; Smart Farming and Agriculture.

I. INTRODUCTION

The agricultural sector plays a vital role in global food production but faces significant challenges in efficiently classifying and assessing the quality of vegetables [1]. Traditional methods, which depend on manual inspection, are labor-intensive, inconsistent, and prone to errors. As the demand for high-quality produce and streamlined supply chain management increases, innovative solutions that can enhance accuracy and efficiency are urgently needed [2].

Recent advancements in artificial intelligence (AI) and machine learning (ML) [3] have shown great promise in tackling various agricultural challenges. Among these advancements, deep learning, a branch of machine learning, has proven to be particularly effective for image recognition tasks. Convolutional Neural Networks (CNNs) [4], a type of deep learning model, have achieved notable success in accurately identifying and classifying intricate patterns within images.

This research explores the use of deep learning models to automate the recognition and classification of vegetables. By employing a specialized dataset of vegetable images and leveraging advanced deep learning techniques, the study aims to assess the accuracy and efficiency improvements in vegetable sorting, grading, and quality inspection. Specifically, the study compares the performance of the CNN model with the VGG16 [5] model to determine the most effective approach for practical application in agricultural environments.

The implications of this study extend beyond agricultural fields to retail settings. Automated systems utilizing deep learning can significantly improve self-checkout processes in grocery stores by accurately identifying and pricing vegetables, thereby enhancing customer convenience and reducing errors. The integration of advanced AI technologies [6] into everyday agricultural and retail operations represents a significant advancement towards modernizing the sector, ensuring consistent quality control, and boosting overall operational efficiency.

This research provides a thorough analysis of the effectiveness of deep learning models in vegetable recognition, highlighting their potential to transform traditional agricultural practices and contribute to the development of more advanced and efficient agricultural solutions.

A. Structure of the paper

The introduction gives information about how the vegetable is predicted according to the class label, what kind of deep learning algorithms were utilized, and the comprehension of the deep learning algorithms The methodology section describes how the algorithms were employed and how they are beneficial in the prediction of vegetables. The results part gives the outcomes of the study, while the discussion section contextualizes the findings. Finally, the conclusion summarizes the main findings of our investigation.

B. Objectives

- To examine the deep learning models used for predicting vegetables

- To create deep learning models capable of recognizing vegetables based on different classes with high accuracy.
- To assess the deep learning models and how accurately they are identifying the images based on features.
- Examine the efficiency of various image pre-processing approaches.
- To predict whether the given image is predicted or not.
- Evaluate the impact of the dataset on the capabilities and solidity of the deep learning models.

II. RELATED WORK

Using texture components, in [1] suggested a comprehensive method for effectively identifying fruits and vegetables. The texture component of the color image is created by calculating the total and contrast of the intensity values of the adjacent pixels. The author suggests using a texture component obtained from a color image, which could result in erroneous detection, rather than taking into account the texture and color parts separately. [7] presented a fruit recognition system that makes use of the features of color, shape, and size. use the k closest neighbor technique to combine three features to identify which robust and increases the method's accuracy. Some fruits are misclassified because they share the same color and shape when color and shape are the only factors taken into account. Fruits can be identified 90% of the time. [8] suggested using texture, shape, and color components to categorize dates. Descriptors are used to calculate texture, form, and color feature binary patterns. Support vector machines (SVM) are used for classification, and by combining features, date fruit categorization is achieved with 98% accuracy. Using the Inception v3 model, [9] suggested grade disease detection and fruit recognition. The inception model uses a convolution neural network (CNN) with an accuracy of 99% to identify the grade and degree of sickness in fruits. This method works effectively and accurately for fruit detection. Fruit classification using a multiclass support vector machine and computer vision was suggested by [10]. Fruits are categorized using a combination of criteria that increase accuracy and reduce misclassification. This approach contrasts various support vector machines (SVM) and determines which SVM is most accurate at categorizing fruits with a 95% accuracy rate. [11] suggested employing neural networks to recognize texture and color components for both intra-class and inter-class fruit recognition. In this experiment, just four fruits out of 270 photos from a tiny dataset were used for testing and training, which could result in false positives. Fruit varieties and subtypes can be identified with 90% accuracy and satisfactory performance. A method for identifying different types of fruits and vegetables based on their color and texture qualities is proposed by [5] Color and texture features are fused using a matching score fusion approach, and then a NN classifier is employed to achieve fruit and vegetable detection. As we can see, most fruit classification techniques are conventional and outdated. Fruit classification heavily relies on linear classifiers and KNN classifiers, whose scarcity restricts technique development and accuracy.

III. METHODOLOGY

The research begins by identifying the challenges inherent in traditional vegetable classification methods and establishing the objectives for improving accuracy and efficiency through deep learning models. A diverse dataset of vegetable images is collected and preprocessed, including data augmentation and normalization, to enhance the robustness and performance of the models. Next, suitable deep learning models are selected, focusing on Convolutional Neural Networks (CNNs) and VGG16, and these models are trained using the preprocessed dataset with optimization techniques.

The trained models are then evaluated using a separate validation dataset to measure their performance in terms of accuracy and other relevant metrics. The performance of the CNN and VGG16 models is compared, highlighting the strengths and weaknesses of each. Practical implications of the findings are analyzed, considering how automated systems powered by these models can improve efficiency in agricultural and retail settings, such as automated sorting, grading, and self-checkout systems in grocery stores. Fig.1 shows flow of the work.

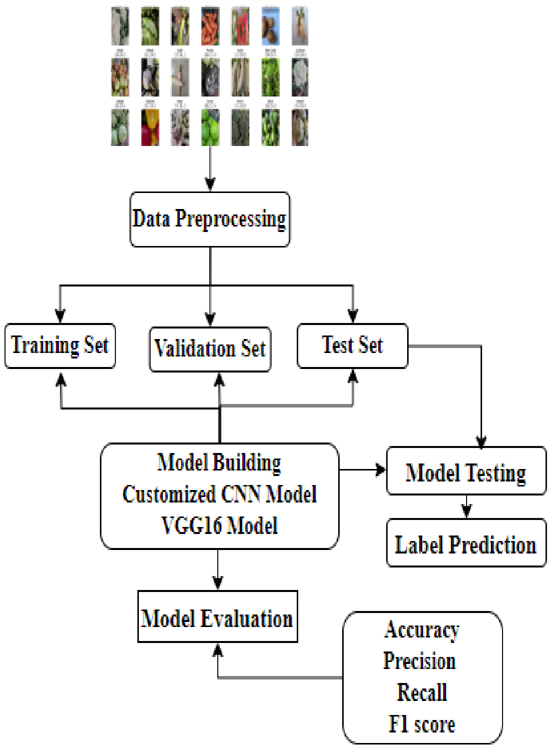

Fig. 1. Flow of the work

A. Convolution Neural Network

The architecture of CNN has the initial convolution layer (conv2d) employing 15 filters, followed by a subsequent layer featuring 64 filters.

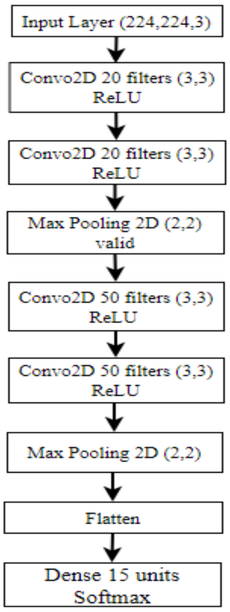

Fig. 2. CNN Architecture

Along with the conv2d, Max-pooling also serves as the down sample of feature maps. Next, we have an additional two convolution layers, with the subsequent layer featuring 64 filters. followed by an additional max-pooling layer. Again, there will be two convolution layers with 128 filters each, followed by a maximum pooling layer. The architectural configuration is completed with three fully connected dense layers, featuring 512, 256, and 16 units, respectively, ultimately culminating in an output layer incorporating sigmoid activation. Fig.2 shows the CNN architecture for real life application vegetable classification using convolution layers, max pooling and dense layers.

The Table-I shows different layers, parameters considered in the CNN model implemented.

TABLE I. LAYERS IN CNN MODEL

Layer	Output shape	Parameters
conv2d 4	(None, 222, 222, 20)	560
conv2d_5	(None, 220, 220, 20)	3620
max_pooling2d_2	(None, 110, 110, 20)	0
conv2d_6	(None, 108, 108, 50)	9050
conv2d_7	(None, 106, 106, 50)	22550
max_pooling2d_3	(None, 53, 53, 50)	0
flatten_1	(None, 140450)	0
dense_1	(None, 15)	2106765

B. VGG 16

The VGG-16 model, which has 16 layers total 13 convolutional layers and three fully connected is distinguished by its depth[12]. Dimensions of the input: (224, 224, 3). Convolutional layers (same padding, 3x3 filters, 64 filters).

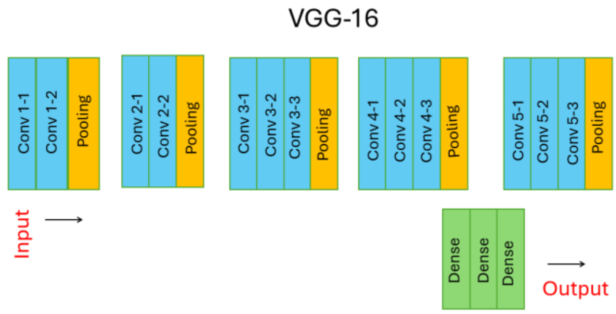

Fig. 3. VGG 16 Architecture

Two successive convolutional layers with a filter size of 3×3 and 64 filters apiece. The spatial dimensions are maintained by applying the same padding. Layer max-pooling with a 2×2 pool and a stride of 2. Two successive convolutional layers with a filter size of 3×3 and 128 filters apiece. Layer max-pooling with a 2×2 pool and a stride of 2. Two convolutional layers with 256 filters each, separated by a 3x3 filter size. There are two sets of three successive convolutional layers, each with 512 filters and a 3x3 filter size. Layer max-pooling with a 2×2 pool and a stride of 2. Convolutional layers two more after the preceding stack. Dimensions of the filter: 3x3. Flatten the 7x7x512 feature map output to create a 25088-size vector. Three ReLU activated layers that are entirely connected. First layer: 4096 output size and 25088 input size.4096 input and 4096 output sizes make up the second tier. The third layer has an output size of 1000 and an input size of 4096, which matches the 1000 classes in the ILSVRC challenge. The third fully linked layer's output is subjected to softmax activation in order to facilitate classification.Fig.3 shows VGG 16 architecture for an application.

IV. RESULTS AND INTERPRETATION

In this research study, we employed multiple deep-learning algorithms for predicting vegetables based on classes. Several deep learning models were built and including CNN, VGG 16. Our goal is to determine whether models successfully predict vegetables.

A. Dataset Description

Vegetable Dataset consists of 21000 images with 15 different class labels such as cauliflower, Tomato, Cabbage, Carrot, Potato, Radish, Pumpkin, Bitter_Gourd, Papaya, Brocoli, Brinjal, Cucumber, bean, capsicum. Dataset consists of 15 folders with 15 different class labels.

Each Class label contains more than 1000 images. Using different parameters like color, weight, taste and size. Each image size is 224 pixels and it is standard size. Fig.4 represents 15 different class labels and presents single image in each class.

Fig. 4. Class Labels of dataset

The vegetable image dataset consists of 21000 more images with 15 different class labels, such as cauliflower, tomato, cabbage, carrot, potato, radish, pumpkin, bitter green, capsicum, papaya, broccoli, brinjal, cucumber, and bean. Figure 5 shows bar graph representation of 15 different class labels.

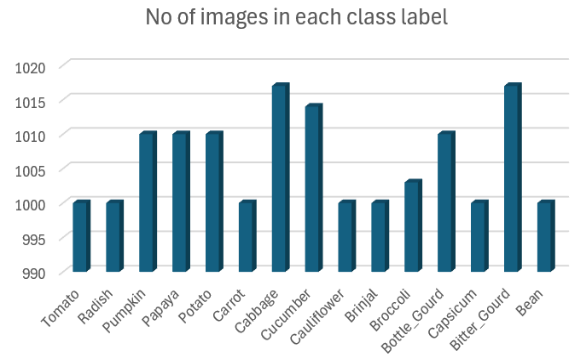

Fig. 5. Class Labels Representation

Table II lists 15 different class labels and their images count in each folder. More than 21,0000 images are present in whole dataset.

TABLE II. THE DATASET CONTAINS CLASS-WISE SAMPLES

Class Name	Total Count	Class Name	Total Count
Tomato	1000	Cauliflower	1000
Radish	1000	Brinjal	1000
Pumpkin	1010	Broccoli	1003
Papaya	1010	Bottle_Gourd	1010
Potato	1010	Capsicum	1000
Carrot	1000	Bitter_Gourd	1017
Cabbage	1017	Bean	1000
Cucumber	1014	Total	21,000

Table III represents splitting percentage is 70% for training and 15% for validation and 15% for testing.

TABLE III. TRAIN AND TEST SPLIT

Class Name	Training	Testing	Validation
Images	15,000	3,000	3,000

The following bar graph represents training, testing and validation splits in vegetable dataset.

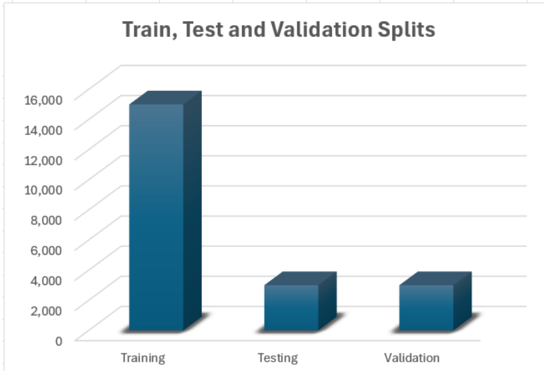

Fig. 6. Train,Test and Validation Splits of dataset

Figure 7 represents the class distribution of train data. Here we use bar graph for representation of train data. X-axis consists of different types of class labels and Y-axis consists of number of images.

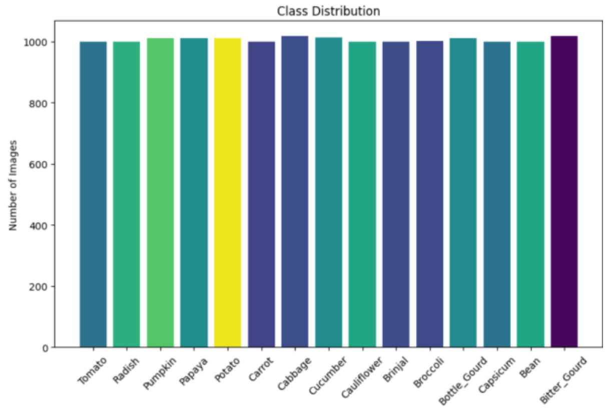

Fig. 7. Class Distribution of train data

B. *Model Performance*

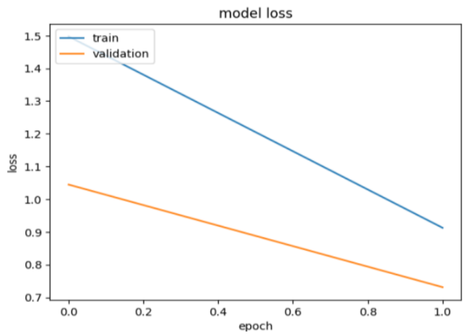

Fig. 8. Model Loss over Epochs

Figure 8 shows the model loss over epochs for CNN Model. Model loss is based on train and validation sets. Train loss is more compared to validation loss.

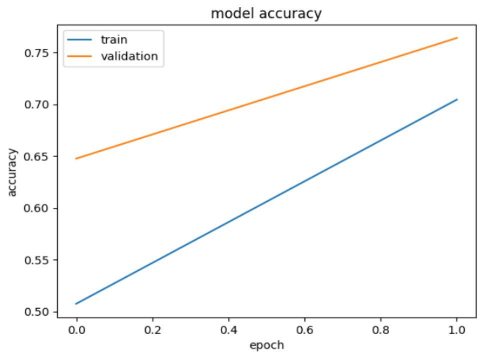

Fig. 9. Model Accuracy over Epochs

Figure 9 shows the Model Accuracy over Epochs for the CNN model. Model Accuracy between training and validation splits. In training, when epochs get increased then MSE also more compared to validation sets.

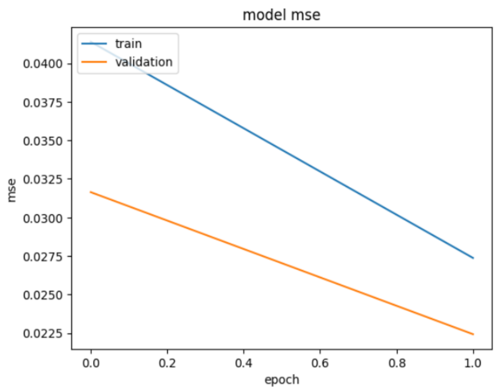

Fig. 10. Model MSE over Epochs

Figure 10 shows the Model MSE over Epochs for the CNN model.

By showing the proportion of accurate and inaccurate instances based on the model prediction, the confusion matrix is used to assess a model's performance. This Confusion matrix provide prediction results. This matrix provides true class labels and predicted class labels. The Confusion matrix consists of four components such as:

- True Positives (TP)
- True Negatives (TN)
- False Positives (FP)
- False Negatives (FN)

Figure 11 shows the confusion matrix for the CNN model with 15 different class labels as true and predicted.

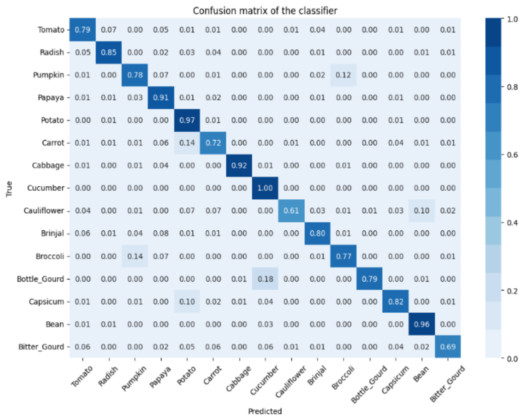

Fig. 11. Confusion Matrix of true and predicted labels

Table IV shows the various performance metrics for CNN model. Performance metrics such as accuracy, precision, recall and f1-score.

TABLE IV. CNN PERFORMANCE METRICS

Accuracy	**Precision**	**Recall**	**F1-Score**
92.69	86	93	89

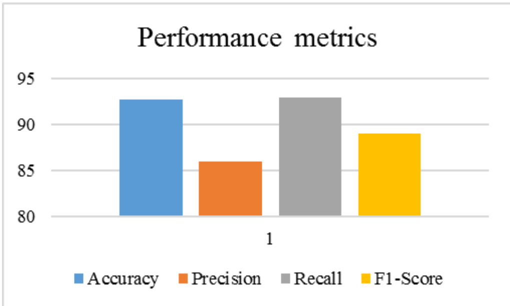

Fig. 12. CNN Performance Metrics

Figure 12 shows bar graph which represents different performance metrics for CNN model.

TABLE V. VGG-16 PERFORMANCE METRICS

Accuracy	**Precision**	**Recall**	**F1-Score**
87.37	83	86	80

Table V shows the various performance metrics for the VGG-16 model.

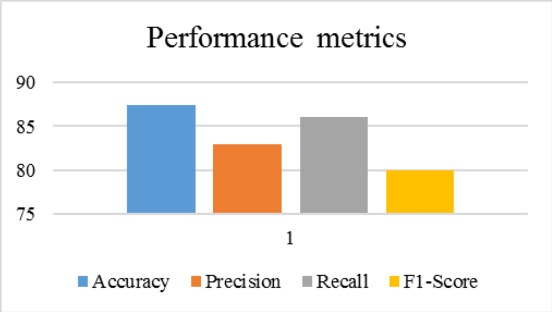

Fig. 13. VGG-16 Performance Metrics

Figure 13 shows bar graph which represents performance metrics of VGG-16 model. Metrics are accuracy, precision, recall and f1-score.

Table VI displays the total accuracy of the models. The values were obtained by deploying various deep-learning models on the dataset.

TABLE VI. ACCURACY OF THE VARIOUS MODELS

Model	Accuracy
CNN	92.69
VGG 16	87.37

Figure 14 shows graph which represents overall accuracy of two different models such as CNN, VGG-16 models.

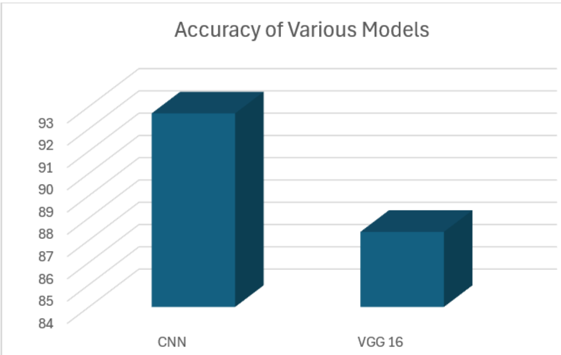

Fig. 14. Accuracy of Various Models

The CNN Model achieves an accuracy of 92.69 Finally, the results show that CNN achieved the maximum accuracy on the dataset bagging 92.69%, while the other models VGG 16 had an accuracy of 87.37%.Based on these findings, we may prefer CNN to recognize vegetable images.

V. CONCLUSION

Automated systems based on deep learning can revolutionize traditional agricultural practices, providing a robust solution for the rapid and accurate classification of vegetables. This not only enhances operational efficiency but also ensures higher consistency in quality control processes. Furthermore, integrating such systems into self-checkout mechanisms in grocery stores can significantly improve customer experience by reducing checkout times and minimizing pricing errors. This research underscores the transformative potential of deep learning models in automating vegetable recognition and classification within the agricultural sector. By leveraging a specialized dataset of vegetable images and employing advanced deep learning techniques, we demonstrated significant improvements in the accuracy and efficiency of sorting, grading, and quality assessment tasks. The Convolutional Neural Network (CNN) model achieved an impressive accuracy of 92.69%, surpassing the VGG16 model, which recorded an accuracy of 87.37%. These findings illustrate the superior capability of CNNs in handling complex image recognition tasks pertinent to agriculture. Future research should focus on expanding the models' recognition capabilities to a wider variety of vegetables and improving their performance under diverse conditions. Integrating other deep learning architectures could further enhance robustness.

REFERENCES

[1] L. Zhu, Z. Li, C. Li, J. Wu, and J. Yue, "High performance vegetable classification from images based on AlexNet deep learning model," *Int. J. Agric. Biol. Eng.*, vol. 11, no. 4, pp. 190–196, doi: 10.25165/j.ijabe.20181104.2690, 2018.

[2] S. R. Dubey and A. S. Jalal, "Robust approach for fruit and vegetable classification," *Procedia Eng.*, vol. 38, pp. 3449–3453, doi: 10.1016/j.proeng.2012.06.398, 2012.

[3] K. Sujatha and B. Srinivasa Rao, "Recent Applications of Machine Learning : A Survey," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 6, pp. 263–267, doi: Retrieval Number: F10510486C219 /19©BEIESP, 2019.

[4] K. Sujatha, K. V. K. Kishore, and B. S. Rao, "Performance of Machine Learning Algorithms in Red Wine Quality Classification based on Chemical Compositions," *Indian J. Ecol.*, vol. 47, pp. 176–180, 2020.

[5] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 8595–8598. doi: 10.1109/ICASSP.2013.6639343, 2013.

[6] K. Sujatha, K. Gayatri, M. S. Yadav, N. C. Sekhara Rao, and B. S. Rao, "Customized Deep CNN for Foliar Disease Prediction Based on Features Extracted from Apple Tree Leaves Images," in *IEEE International Interdisciplinary Humanitarian Conference for Sustainability (IIHC-2022)*, pp. 193–197. doi: 10.1109/iihc55949.2022.10060555, *November 18th & 19th 2022*, 2023.

[7] W. C. Seng and S. H. Mirisaee, "A new method for fruits recognition system," in *Proceedings of the 2009 International Conference on Electrical Engineering and Informatics, ICEEI 2009*, vol. 1, no. June, pp. 130–134. doi: 10.1109/ICEEI.2009.5254804, 2009.

[8] B. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2012.

[9] M. Nikhitha, S. Roopa Sri, and B. Uma Maheswari, "Fruit Recognition and Grade of Disease Detection using Inception V3 Model," in *Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019*, pp. 1040–1043. doi: 10.1109/ICECA.2019.8822095, 2019.

[10] Y. Zhang and L. Wu, "Classification of fruits using computer vision and a multiclass support vector machine," *Sensors (Switzerland)*, vol. 12, no. 9, pp. 12489–12505, doi: 10.3390/s120912489, 2012.

[11] Y. Zhang, S. Wang, G. Ji, and P. Phillips, "Fruit classification using computer vision and feedforward neural network," *J. Food Eng.*, vol. 143, pp. 167–177, doi: 10.1016/j.jfoodeng.2014.07.001, 2014.

[12] N. Abhinav and K. Sujatha, "Content-Based Movie Recommendation System Using cosine similarity measure," in *AIP Conf. Proc. 2901, 060035 (2023)*, vol. 2901, pp. 1–10. [Online]. Available: https://doi.org/10.1063/5.0178819, 2023.