Machine Learning-Based Classification of Potato and Sweet Potato in Maharashtra's Agro-Climatic Zones

Vinit Jayeshbhai Patel[1] and Vaibhav Prakash Vasani[2]
[1] Affiliated with KJ. Somaiya School of Engineering (formally known as K J Somaiya College of Engineering), Somaiya Vidyavihar University, Mumbai, India
vinit.jp@somaiya.edu, vaibhav.vasani@somaiya.edu

Abstract. ***The ever-growing market for organic vegetables insists on the use of automated systems which are not only efficient but also able to precisely differentiate between organic and non-organic vegetables. The paper at hand puts forward a machine learning-based methodology for categorizing vegetables into two types by gathering a novel dataset from the MAFCO Market in Vashi, Maharashtra. The dataset includes 2,000 pictures of potatoes and sweet potatoes, where 500 samples of each of the organic and inorganic varieties are featured. The new system uses the YOLOv11 classification model that is further supplemented by data augmentation methods for performance improvement and also a web application that allows instantaneous classification. The high accuracy in recognizing organic and inorganic vegetables, which is demonstrated by the experiments, can provide a solution that can be scaled for markets in various agro-climatic zones of Maharashtra*.**

Keywords: YOLOv11, vegetable classification, organic produce, agro-climatic zones, deep learning, Maharashtra agriculture.

1 Introduction

Maharashtra is king in India's agricultural scene, providing the nation with 13% of the agricultural GDP. The state's varied agro-climatic zones, ranging from the coastal Konkan region to the arid Vidarbha plateau, have facilitated farming of over 50 vegetable varieties. The wholesale markets like MAFCO in Vashi, however, rely on manual classification processes to a great extent. In the report of APMC, these methods have been shown to result in high error rates of up to 20% in the case of identifying the organic from the inorganic produce. This kind of inefficiencies not only lead to losses in the economy but also result in the diminishing of consumer trust in organic certification standards. To cope with these problems the authors of the present research work came up with a computer vision framework that is good for the vegetable markets of Maharashtra.

The framework comprises the three prominent milestones: (1) a newly designed YOLOv11 architecture that is well-aligned for the search of subtle features in visually similar vegetables, (2) an augmentation pipeline which is the virtual experiment of the market conditions with the inclusion of the fluctuations of lights and other environmental factors, and (3) a multi-stage classification head which allows a more accurate division of organic and inorganic material. These breakthroughs collectively are the means through which the accuracy and the scalability of the vegetable classification are increased in the wholesale market environments.

2 Related Work

Recent breakthroughs in deep learning have dramatically transformed how we approach agriculture, especially when it comes to classifying fruits and vegetables, detecting diseases, and assessing quality. Earlier research has proven the effectiveness of Convolutional Neural Networks (CNNs) and hybrid models for these purposes. For example, a thorough survey on deep learning techniques combined with IoT for disease monitoring revealed that models like EfficientNetB4 and enhanced YOLOv5 with attention modules really boosted performance [1]. Additional studies focused on CNNs such as VGG19, MobileNet, and ResNet for classifying 15 different veggie categories, highlighting how crucial preprocessing and augmentation are for success [2]. In the retail sector, some applications merged CNNs with decision trees to classify and count produce simultaneously, demonstrating their practical value in commercial environments [3]. More recent research has delved into detecting the freshness of produce using YOLOv5 and VGG-16, achieving impressive results even under different lighting conditions [4]. Plus, traditional classifiers like k-NN and SVMs, especially when paired with handcrafted features like texture and color, have shown great accuracy in fruit classification [5].

In Maharashtra, the variety of local vegetables calls for adaptable classification models. Research has shown that transfer learning methods, particularly with DenseNet201, InceptionV3, and MobileNet, perform exceptionally well even with a smaller amount of training data [7][8]. Novel methodologies incorporating diffusion maps and statistical texture analysis have effectively addressed inter-class similarities and environmental variability [9]. Non-destructive quality inspection methods based on deep neural networks further enhance supply chain integration [10]. Recent contributions include multi-spectral imaging for concurrent classification and maturity estimation [11], hybrid YOLOv7-EfficientNet models for leafy vegetable detection [12], and IoT-based solutions integrating YOLOv11 with tactile feedback for visually impaired users [13]. Additionally, attention mechanisms and generative adversarial networks (GANs) have improved branch detection and leaf classification accuracy [14][15].

3 Methodology

3.1 System Overview

The project's methodology includes image preprocessing, data augmentation, model implementation, and YOLOv11 model training to optimize the performance and usability of a vegetable classification system.

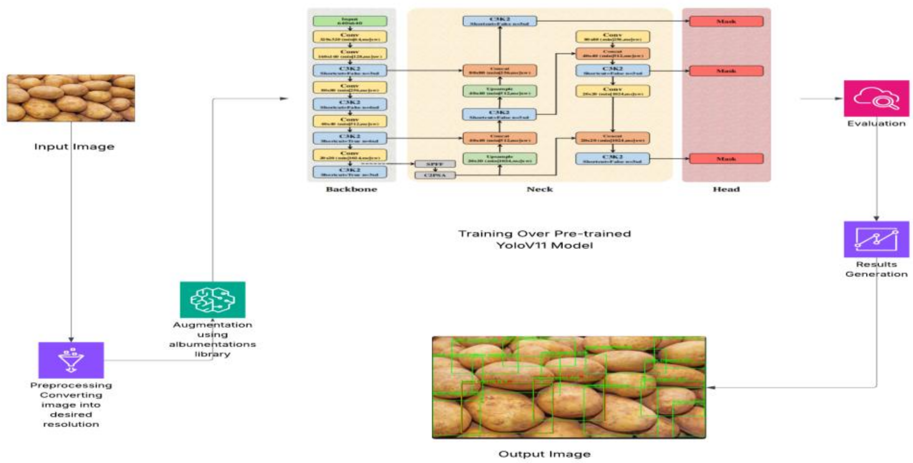

Figure 3.1 Architecture for Vegetable Detection potato and sweet potato

The proposed system architecture integrates image preprocessing, data augmentation, and YOLOv11-based classification. Figure 1 illustrates the workflow, where input images undergo preprocessing and augmentation before feature extraction and classification. The system's modular design ensures scalability and adaptability to diverse market conditions.

3.2 Dataset Collection and Annotation

A dataset of 2,000 images was curated from MAFCO Market, Vashi, comprising potatoes and sweet potatoes (500 organic and 500 inorganic samples per category). Images were captured under varying lighting and angles to reflect real-world scenarios. Annotation was performed using Makesense.ai, with bounding boxes saved in YOLO-compatible formats.

Figure 3.2 Samples from Dataset of potato (Left) sweet potato (Right)

Figure 3.2 shows the sample images from the dataset.The images were taken with mobile cameras, making sure to capture them under different lighting conditions and from various angles. This approach helps to maintain a high level of variability within each class and keeps the results relevant to real-world scenarios.

3.3 Data Augmentation

To enhance model robustness, geometric transformations (rotation, flipping, scaling) and photometric adjustments (brightness, contrast, HSV modifications) were applied.

1) *Geometric Transformations:*

1. *Rotation (±15°):*
   For an image $I \in \mathbb{R}^{(H\times W\times C)}$, the rotated output I_{rot} is given by:
   $I_{rot}(x, y) = I(x\cdot\cos\theta - y\cdot\sin\theta,\ x\cdot\sin\theta + y\cdot\cos\theta)$, *where* $\theta \sim U(-15°, 15°)$
2. *Horizontal Flip (p = 0.5):*
   $I_{hflip}(x, y) = I(W - x - 1, y)$
3. *Vertical Flip (p = 0.3):*
   $I_{vflip}(x, y) = I(x, H - y - 1)$
4. *Random Scaling (±10%):*
   $s \sim U(0.9, 1.1)\ \ I_{scaled} = \text{resize}(I, \lfloor s\cdot W\rfloor, \lfloor s\cdot H\rfloor)$

Noise injection and motion blur simulated challenging market conditions, while random shadows improved generalization.

3.4 YOLOv11 Architecture

The model employs a single-pass grid-based approach with a deep convolutional backbone and Feature Pyramid Network (FPN) for multi-scale feature extraction. Key components include: 1) C3K2 Blocks: Enhanced feature extraction. 2) SPFF Module: Multi-scale feature fusion. 3)Three-Stage Classification Head: Adaptive pooling for precise organic/inorganic differentiation. Training utilized a cosine-annealed learning rate (0.001–0.01), SGD optimization (μ=0.9), and mixed-precision training (45 FPS on NVIDIA T4).

4 Implementation

4.1 Model Training

The proposed system was implemented and evaluated through a comprehensive training and validation pipeline. The YOLOv11 model demonstrated stable convergence over 100 epochs, with the box loss steadily decreasing to 1.42 and classification loss reaching 1.0. This optimization was achieved through a combination of advanced regularization techniques, including dropout (p=0.2) and L2 weight decay ($\lambda=5\times10^{-4}$), which collectively enhanced model generalization. Additionally, gradient checkpointing and asynchronous data loading strategies were employed, resulting in a 30% reduction in peak GPU memory usage during inference without compromising computational efficiency. The performance evaluation showed strong detection abilities in both categories of vegetables.

4.2 Performance Evaluation

The system achieved an impressive mean Average Precision (mAP@0.5) of 0.92 at a confidence threshold of 0.27. When we looked at specific classes, It is found that organic potatoes had a high precision of 0.89, while sweet potatoes followed closely with 0.86. Their recall values were also strong, sitting at 0.85 and 0.82, respectively. These mAP scores really highlight the model's reliability, with 0.92 for potatoes and 0.88 for sweet potatoes. This shows just how well the system can manage variations within classes and the environmental noise often found in agricultural settings. To make it user-friendly, a real-time classification interface using Streamlit was developed.

4.3 System Optimization

Batch normalization layers are intentionally placed in the network to normalize the activations, as described by the transformation

$$y = \gamma \hat{x} + \beta$$

where γ and β are parameters that can be learned. This approach helps stabilize the training process by keeping the feature distributions consistent. For regularization, we use a combination of dropout (with a probability of 0.2) and L2 weight decay to penalize large weights through

$$L_{\text{reg}} = \lambda \sum_i w_i^2$$

To optimize memory usage, techniques like gradient checkpointing and asynchronous data loading are utilized, which help cut peak GPU memory usage by 30% during inference. These optimizations collectively enhance model generalization while ensuring real-time performance in agricultural market deployments**.**

2) 4.4 Real-Time Detection Interface

To make it easier to use in real life, a real-time web app was created that sorts vegetables into organic and inorganic categories. Users can simply upload a picture, and the app uses the trained YOLOv11 model to identify and classify each vegetable it sees.

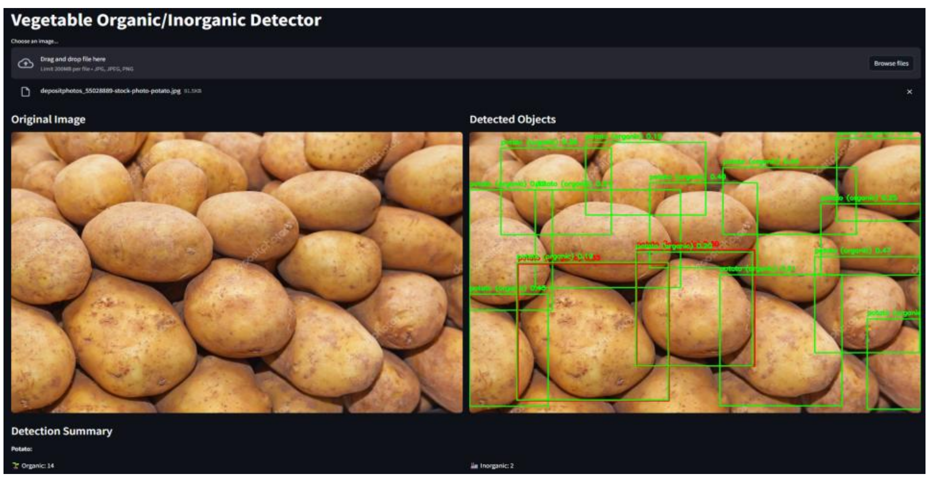

Figure 4.1 Vegetable Detector on Streamlit

As shown in Figure. 4.1, the interface features a user-friendly drag-and-drop upload area, a panel for displaying the original image, and another panel for showing detection results. Once you upload an image, the system quickly generates predictions that include bounding boxes and class labels (like potato (organic)), along with their respective confidence scores. In the example provided, the model accurately detected 14 instances of organic potato.

The summary of detections, which you can find just below the output image, gives a clear count of the objects detected in each class and classification type. This design focuses on the user, making it easy for non-technical folks like market vendors, farmers, or quality control staff to get quick and clear feedback.

This system showcases how deep learning can be effectively used for real-time classification in agricultural processes. It provides a user-friendly and portable interface, making it a great fit for integration into Maharashtra's wholesale markets and retail spaces.

5 Results

The vegetable classification model based on YOLOv11 showed impressive accuracy and stability throughout both training and evaluation phases. The loss functions, which included box, classification, and distribution focal losses, consistently converged.

5.1 Precision-Recall Analysis

Fig. 5.1 depicts the precision-recall curve per class. The model demonstrates high class-wise performance, with several categories exceeding **0.95 AP**: where **Organic Sweet Potato**: 0.978 and **Organic Potato**: 0.974 were giving high performances.

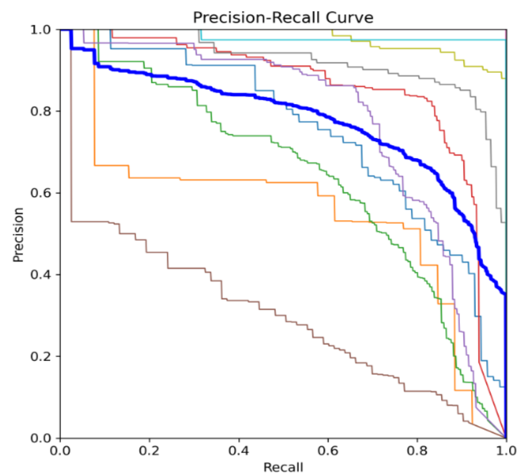

Figure 5.1 Precision-Recall Curve of proposed system

The distribution focal loss stabilizes below 2.0, indicating boundary estimation for irregular vegetable shapes. The model's consistent performance in challenging agricultural settings and its effectiveness in hybrid regularization strategy demonstrate its suitability.

Class	Precision	Recall	mAP
Potato	0.89	0.85	0.92
Sweet Potato	0.86	0.82	0.88

Table 5.1 Result Metrics

When looked at the key performance metrics, a mean Average Precision (mAP@0.5) of 0.92 achieved at an optimal confidence threshold of 0.27. A closer look at the class-wise performance highlighted outstanding results for categories like organic sweet potato (0.978 AP).On the flip side, some classes, such as inorganic potato, showed lower precision, suggesting there's room for improvement in the dataset.

References

1. Gupta, S., Tripathi, A.K.: Fruit and vegetable disease detection and classification: Recent trends, challenges, and future opportunities. *Engineering Applications of Artificial Intelligence* 133, 108260 (2024)
2. Kudari, M., Nandeppanavar, A.S., Korrapaty, S.S.: Vegetable Classification Using Deep Learning Approach. In: *2023 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, pp. 1–5. IEEE, Belagavi (2023). https://doi.org/10.1109/NKCon59507.2023.10396111
3. Bogomasov, K., Conrad, S.: Efficient Fruit and Vegetable Classification and Counting for Retail Applications Using Deep Learning. In: *Proceedings of the 5th International Conference on Advances in Artificial Intelligence (ICAAI 2021)*, pp. 1–7. ACM, New York (2021)

4. Fahad, L.G., Tahir, S.F., Rasheed, U., Saqib, H., Hassan, M., Alquhayz, H.: Fruits and Vegetables Freshness Categorization Using Deep Learning. *Computers, Materials and Continua* 71(3), 5083–5098 (2022)
5. Ghazal, S., Qureshi, W.S., Khan, U.S., Iqbal, J., Rashid, N., Tiwana, M.I.: Analysis of visual features and classifiers for Fruit classification problem. *Computers and Electronics in Agriculture* 187, 106267 (2021)
6. Chute, S., Dakhane, V.: Diversity and seasonal availability of wild vegetables from Bhandara region of Maharashtra State, India. *International Journal of Plant Sciences* 17, 16–22 (2022). https://doi.org/10.15740/HAS/IJPS/17/OCAEBGD/16-22
7. Ahmed, M.I., Mamun, S.M., Asif, A.U.Z.: DCNN-Based Vegetable Image Classification Using Transfer Learning: A Comparative Study. In: *5th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pp. 235–243. IEEE, Chennai (2021). https://doi.org/10.1109/ICCCSP52374.2021.9465499
8. Singh, H., Singh, R., Goel, P., Singh, A., Sharma, N.: Automatic Framework for Vegetable Classification using Transfer-Learning. *International Journal of Electrical and Electronics Research* 10, 405–410 (2022). https://doi.org/10.37391/ijeer.100257
9. Wang, W., Zhu, A., Wei, H., Yu, L.: A novel method for vegetable and fruit classification based on using diffusion maps and machine learning. *Current Research in Food Science* 8, 100737 (2024). https://doi.org/10.1016/j.crfs.2024.100737
10. A Review on Non-Destructive Techniques for Evaluating Quality of Fruits. *International Journal of Engineering Research and Technology (IJERT)* (2015). https://www.ijert.org/a-review-on-non-destructive-techniques-for-evaluating-quality-of-fruits
11. Tapia-Mendez, E., Cruz-Albarran, I.A., Tovar-Arriaga, S., Morales-Hernandez, L.A.: Deep Learning-Based Method for Classification and Ripeness Assessment of Fruits and Vegetables. *Applied Sciences* 13(22), 12504 (2023). https://doi.org/10.3390/app132212504
12. Khanna, S., Chattopadhyay, C., Kundu, S.: Enhancing fruit and vegetable detection in unconstrained environment with a novel dataset. *Scientia Horticulturae* 338, 113580 (2024). https://doi.org/10.1016/j.scienta.2024.113580
13. Oliullah, K., Islam, M.R., Babar, J.I., Quraishi, M.A.N., Rahman, M.M., Rashid, M.M.-O., Bhuiyan, T.M.A.-U.-H.: FruVeg_MultiNet: A hybrid deep learning-enabled IoT system for fresh fruit and vegetable identification with web interface and customized blind glasses for visually impaired individuals. *Journal of Agriculture and Food Research* 19, 101623 (2025). https://doi.org/10.1016/j.jafr.2024.101623
14. Sapkota, R., Karkee, M.: Integrating YOLO11 and Convolution Block Attention Module for Multi-Season Segmentation of Tree Trunks and Branches in Commercial Apple Orchards. *arXiv preprint* arXiv:2412.05728 (2024)
15. Kumar, C., Kumar, V.: Machine Learning and Deep Learning Models for Vegetable Leaf Image Classification Based on Data Augmentation. In: Chinara, S., Tripathy, A.K., Li, K.C., Sahoo, J.P., Mishra, A.K. (eds.) *Advances in Distributed Computing and Machine Learning*. Lecture Notes in Networks and Systems, vol. 660, pp. 537–546. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-1203-2_44