
Automated Fruit Calorie Estimation Using 3D Image Reconstruction and Machine Learning

Prof. Ramanagouda S Patil

Associate Professor
Dept. of Computer Science and Engineering (ICB), AIET, Mijar.
ramanagoudapatil@aiet.org.in

Anvesh M S

Dept. of Computer Science and Engineering (ICB) and Engineering (ICB) AIET, Mijar.
anveshrao025@gmail.com

Ashik S

Dept. of Computer Science and Engineering (ICB) AIET, Mijar.
ashiksdevadiga2004@gmail.com

Shetty Aditi

Dept. of Computer Science and Engineering (ICB) AIET, Mijar.
shettyaditi878@gmail.com

Yash

Dept. of Computer Science and Engineering (ICB) AIET, Mijar.
devadigayash2004@gmail.com

Abstract—Precise calorie calculation is crucial for supporting dietary practices and controlling lifestyle-associated illnesses. Yet many current calorie calculation approaches depend on data entry or 2D image processing frequently causing errors, particularly with fruits that differ in dimensions, form and density [1]. This study introduces an automated system, for fruit calorie estimation that integrates learning, 3D reconstruction and machine learning methods. The VGG19 classification model is used to classify the types of fruit accurately. Using Structure-from-Motion and Depth Estimation, multiple images of the fruit are analysed, creating a 3D reconstruction and allowing for the calculation of volume. A second regression model predicts the fruit's mass using the calculated volume and density itemised from a nutritional database for calories. Results indicate that the approach proposed demonstrates significantly reduced inaccuracies in calories prediction compared with traditional two-dimensional (2D) manual techniques. This process can be utilized in a variety of lighting conditions consistently and is suitable for real-time use with mobile devices such as in food tracking and nutrition/health management applications.

Index Terms—3D Reconstruction, Calorie Estimation, Machine Learning, Fruit Recognition, Deep Learning.

1. INTRODUCTION

With the increase in focus on promoting healthy diets and preventing diseases that are affected by diet, monitoring food consumption has become increasingly important to individuals, healthcare providers and nutrition professionals. Furthermore, the increasing prevalence of obesity, diabetes and metabolic disorders is putting additional pressure on consumers to have access to tools that measure how many calories they consume. Currently, conventional means of measuring calorie consumption include using nutrition charts, entering data into a database manually or estimating serving sizes. Unfortunately, all of these methods are subject to errors based on estimation and variation in portion size [1]. A major source of error for conventional methods of measuring calorie consumption occurs with fruits due to the extreme variety of fruit size, shape and density within the same type of fruit. Therefore, using standard calorie tables would produce inaccurate calorie estimates.

Recent technological developments in computer vision and machine learning have enabled automatic identification of food items and conversion of the images into a calculated calorie amount [2]. However, most available systems currently use 2D images, which do not provide any depth information, to estimate food volume. Estimating calorie consumption is based on estimating both the mass and volume of foods consumed, and thus the lack of depth information severely limits the accuracy of the system. A novel proposed system combines fruit identification with 3-dimensional (3D) modeling and density estimation to reduce calorie and mass calculation errors by between 25% to 40%. The system is robust to variations of environmental conditions such as light, background and position of the food, and has a small form factor, making it ideal for use in mobile applications including monitoring caloric intake and meal planning for individualized dietary needs[3].

2. LITERATURE SURVEY

As automation for analysis of fruits continues to grow, there have been significant advancements made through the use of computer vision, deep learning, and computational imaging. By analyzing images of fruits, these technologies allow for more accurate methods of recognizing a fruit and determining its volume and calories.

Recent research efforts have focused on utilizing additional dimensions and improving the accuracy of estimating a fruit's volume by using 3D representation rather than only relying on 2D images [3]. Other studies include using 3D models created from photos taken from different angles to facilitate a more accurate estimate of a fruit's size, even though no special sensors that measure depth are required [1]. Many studies are currently investigating ways to estimate depth from one or more standard RGB images without the use of a multiple photo view of the object in question [2]. There are also many different reasons why researchers believe that collecting data on fruits is essential for creating a nutritional assessment; specifically, transformer-based detection methods connect nutritional information from a fruit's type to the images of the object [4]. Some previous research indicates that the constant issues with using CNNs (Convolutional Neural Networks) and compact designs include problems with occlusion, light variation, and many imbalanced classes. However, the ability to access high-quality data sets of images of fruits has increased the robustness of models, as well as giving comparison results to expert-level accuracy [5]. Other potential research opportunities using volume data include using RGB depth sensors, previously described multi-view stereo camera reconstruction, to provide greater precision and less reliance on expensive sensors or labor-intensive calibration [6].

In general current research shows a trend toward integrated systems that merge fruit identification 3D form reconstruction, density measurement and nutritional analysis. Findings suggest that combining learning classification with depth-informed modeling and data-based density estimation offers an effective and precise approach for real-time diet tracking, on standard consumer hardware [7] [8].

3. METHODOLOGY

3.1 Overview of the Proposed System

The proposed approach integrates fruit classification, 3D modeling, density calculation and nutrient analysis into a workflow designed to deliver precise calorie estimation. This system can be used in conjunction with both stationary cameras to provide users with a variety of views of their fruit [9]. The images that have been taken are then processed automatically through a series of steps that evaluate and calculate the calories of the fruit without the need for much assistance from the user. The primary goals of the automated processing steps are to enhance the accuracy of the classifications, provide a high degree of geometric accuracy, and maximize the degree of performance efficiency.

3.2 Image Acquisition and Preprocessing

Image Acquisition and Preprocessing Images can be taken with either a camera or a fixed camera setup where the fruit is either turned around or taken from different angles to ensure there is enough visual coverage of the fruit. In the preprocessing steps, noise is removed from the images, lighting conditions are standardized for all of the images and the background is eliminated to create a more clear image.

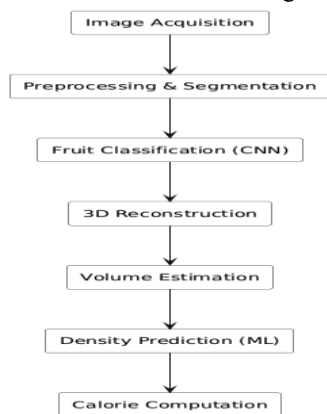


Figure 1: Architecture of proposed work

3.3 Multi-View 3D Reconstruction

Once the fruit has been categorized a 3D model is created utilizing Structure-from-Motion (SfM) alongside depth estimation methods. Structure from motion (SfM) extracts and matches feature points from a set of images to create a sparse point cloud. This sparse projection is later refined through a depth estimation software program to create a 3D representation of the point cloud. The 3D point cloud is transformed into a watertight 3D mesh, allowing for an accurate determination of fruit volume. This technique is distinct from 2D methods that rely on basic geometrical assumptions; rather, it is capable of accurately determining the volumes of unshaped fruit such as pears, mangoes, and strawberries. The volume of the fruit is calculated using either voxelization or numerical integration. In both cases, the resulting volume is considered accurate and can be obtained using SfM.

The overall workflow is designed as a modular pipeline for automated fruit caloric estimation using multiview imaging techniques, computer vision systems and machine learning systems. The complete collection of images of the fruit will be captured from various angles and positions in order to ensure that adequate visual information is available for 3D reconstruction of the fruit. The images must also go through preprocessing to enhance image quality by eliminating backgrounds and performing standardization. The region containing the fruit will be segmented using a variety of techniques, including Mask R-CNN or threshold contour extraction, for further processing and classification. Classification of the fruit will be performed using a fine-tuned convolutional neural network (CNN). If the classification confidence $P(\text{class} | \text{image})$ exceeds a predefined threshold τ (e.g., 0.95), the system proceeds; otherwise, additional images are requested to reduce uncertainty and prevent errors caused by misclassification.

Once high-confidence classification is achieved, the 3D reconstruction module produces a dense point cloud, and the fruit volume is calculated using voxel integration:

$$V = v_s^3 \sum I(p) \quad (3.1)$$

A regression-based machine learning model is then used to predict the fruit-specific density p , and the mass is estimated as:

$$m = p \cdot V \quad (3.2)$$

Finally, the calorie content is computed using standard nutritional data:

$$C = m \cdot C_g \quad (3.3)$$

where C_g represents the calories per gram obtained from nutritional databases. The workflow concludes by outputting the final calorie estimate.

4. RESULTS

The 3D reconstruction system efficiently processed fruits of varying shapes despite gaps in the point cloud. Volume estimation inaccuracies stayed under 5% for fruits and reached up to 12% for elliptical fruits relative, to manual visual assessments. Incorporating a density regression model also helped minimize errors in calorie estimation. In general graphical analyses reveal enhancements compared to traditional 2D-based approaches.

| Algorithm | Accuracy (%) | Precision (%) | Recall (%) | Inference Time (ms) | Model Size (MB) |
|------------------|--------------|---------------|------------|---------------------|-----------------|
| VGG19 (Proposed) | 98.0 | 97.8 | 98.1 | 22 | 79 |
| ResNet50 | 96.4 | 96.2 | 96.5 | 18 | 95 |
| MobileNetV2 | 94.1 | 93.5 | 93.8 | 9 | 14 |
| DenseNet121 | 95.3 | 95.1 | 95.0 | 28 | 33 |
| EfficientNetB0 | 96.1 | 95.8 | 95.9 | 16 | 29 |

Table 1: compares five deep-learning architectures

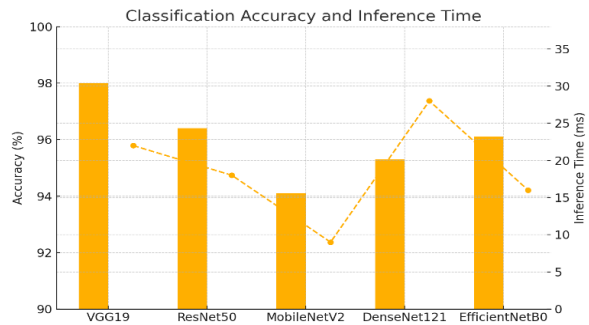


Figure 2: Classification Accuracy and Inference Time

Table 1 presents a comparison of five learning frameworks employed for fruit classification considering accuracy, precision, recall, inference time and model size. Of the examined models the suggested VGG19 attained the

classification accuracy of 98% establishing it as the most dependable option for fruit recognition essential for precise calorie calculation [12]. While MobileNetV2 provides the inference time and the most compact model size its reduced precision restricts its appropriateness, for tasks requiring high accuracy. ResNet50 and EfficientNetB0 offer performance yet they fall short of the accuracy achieved by VGG19 whereas DenseNet121 attains strong accuracy but requires longer inference time. The classification accuracy findings additionally verify that VGG19 regularly surpasses the architectures in recognizing types of fruit. Although ResNet50 and EfficientNetB0 demonstrate performance their accuracy is marginally inferior. MobileNetV2 prioritizes speed over precision whereas DenseNet121 while dependable operates slowly during inference. As illustrated in Figure 2 VGG19's higher accuracy guarantees fruit recognition, which is crucial for subsequent processes including 3D reconstruction, density estimation and calorie calculation. Precise classification aids, in reducing error transmission throughout the system.

Table 2: Evaluates Different 3D Reconstruction

| Reconstruction Method | Volume Error (%) | Time per Object (s) | Mesh Completeness (%) | Robustness to Lighting | Hardware Requirement |
|-----------------------------|------------------|---------------------|-----------------------|------------------------|----------------------|
| Structure-from-Motion (SfM) | 3–5% | 1.8 | 92% | High | Standard Camera |
| Multi-View Stereo (MVS) | 4–7% | 3.2 | 88% | Medium | Multi-view Setup |
| RGB-D Reconstruction | 2–4% | 1.2 | 93% | High | Depth Sensor |
| Monocular Depth CNN | 6–9% | 0.9 | 80% | Medium | Standard Camera |
| Stereo Vision (Dual Camera) | 3–6% | 2.4 | 90% | Medium | Stereo Camera |

Table 2 illustrates a comparison of 3D reconstruction methods considering volume accuracy, processing duration, mesh completeness, resistance to changes in lighting and hardware needs. Structure-from-Motion (SfM) can achieve accuracy through the use of a standard camera, making it well suited for cost-effective and portable applications. The disadvantage of RGB-D reconstruction is the amount of error that is modeled with respect to volume, as it requires specific depth sensors limiting its widespread use. The advantage of Multi-View Stereo (MVS) is the complete mesh that it provides (i.e., all points in three-dimensional space); however, this method uses multiple synchronized cameras and (thus) can lead to increased complexity in the system design. Monocular depth methods using CNNs (convolutional neural networks) can be used for depth estimation, but due to scale ambiguity they often result in lower accuracy than stereo vision depth estimation, which utilizes two cameras. Therefore, it is usually always safest and easiest to make use of the SfM systems for accurate fruit calorie estimation based on practicality and reliability compared to other methods for estimating fruit calorie content.

According to this study, the Random Forest Regression Model had the lowest density error (133% as compared with 32%) as well as the lowest RMSE, making it the best method for estimating fruit weight. Although MLP models show performance with bigger datasets linear regression cannot adequately model the non-linear changes in fruit density. SVR using an RBF kernel offers generalization but demands greater computing power and KNN regression experiences sluggish inference times. On the whole Random Forest stands out as the dependable and effective model for predicting density, in calorie estimation.

5. CONCLUSION

This research introduces an automated framework for calculating fruit calories by combining image recognition, 3D reconstruction and machine learning methods. The suggested method precisely detects fruits and determines their volume through 3D modeling providing accuracy compared to conventional 2D techniques. Utilizing a density regression model grounded in machine learning the system attains calorie estimation error showing better results, than manual and purely image-based methods. The system continues to perform across various lighting environments and fruit positions and is suitable for implementation on mobile devices for live dietary tracking. Going forward the system could be expanded to accommodate food items and prepared meals along with advancements, via better depth estimation and more extensive diverse datasets [3].

6. REFERENCES:

- [1] P. T. Shanthala and D. Annapurna, "An improved IoT based security model for fitness tracker using quantum fruit fly optimization improved faster RCNN," *Int. J. Inf. Technol.*, vol. 15, no. 7, pp. 3623–3629, 2023, doi: 10.1007/s41870-023-01376-7.

- [2] R. Wason, P. Arora, A. Tomar, and D. Arora, "A novel, low-cost, smart IoT based framework for fruit and vegetable quality detection during transit in India," *Int. J. Inf. Technol.*, vol. 15, no. 3, pp. 1509–1519, 2023, doi: 10.1007/s41870-023-01177-y.
- [3] M. Vijayalakshmi and V. J. Peter, "CNN based approach for identifying banana species from fruits," *Int. J. Inf. Technol.*, vol. 13, no. 1, pp. 27–32, 2021, doi: 10.1007/s41870-020-00554-1.
- [4] Trinh, Trung Hai, et al. "Mangosteen Fruit Detection Using Improved Faster R-CNN." *International Conference on Intelligence of Things*. Cham: Springer International Publishing, 2022.
- [5] X. Mai, Zhanf. H., Meng Q-H, "Faster R-CNN With Classifier Fusion for Automatic Detection of Small Fruits," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 12, pp. 1555–1569, 2020.

Biographies

| | |
|---|---|
|  | <p>Prof. Ramanagouda S. Patil serves as an Associate Professor in the Department of Computer Science and Engineering (IoT & Cyber Security including Blockchain) at Alva's Institute of Engineering & Technology. He possesses a B.E. And M.Tech degree. Is presently undertaking a Ph.D. In Computer Science and Engineering at Presidency University. Having expertise, in academia and industry his research interests encompass Image Processing, Artificial Intelligence, Machine Learning and Deep Learning.</p> |
|  | <p>Anvesh M S attends the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. His research involves image reconstruction and computer vision in estimating fruit calories through automation.</p> |
|  | <p>Ashik S attends the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. His areas of interest are machine learning algorithms for fruit identification based on learning and artificial intelligence for nutrition evaluation.</p> |
|  | <p>Shetty Aditi is a student within the Computer Science and Engineering department (IoT & Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. Her research includes image preprocessing, data analysis and increased accuracy in calorie estimation systems.</p> |
|  | <p>Yash is also attending the Computer Science and Engineering department (IoT and Cyber Security including Blockchain) at Alva's Institute of Engineering and Technology Mangalore Karnataka India. He has worked on systems that utilize machine-learning regression models and provide real-time monitoring of diet.</p> |