2024 IEEE 9th International Conference for Convergence in Technology (I2CT)
Pune, India. Apr 5-7, 2024

Food Image Recognition and Volume Estimation: A Comprehensive Study for Dietary Assessment

Druwil Jain
School of Computer Engineering
MIT Academy of Engineering
Alandi, Dist. Pune, India
druwil.jain@mitaoe.ac.in

Aditi Uphade
School of Computer Engineering
MIT Academy of Engineering
Alandi, Dist. Pune, India
aditi.uphade@mitaoe.ac.in

Sakshi Ugale
School of Computer Engineering
MIT Academy of Engineering
Alandi, Dist. Pune, India
sakshi.ugale@mitaoe.ac.in

Soham Kumbhar
School of Computer Engineering
MIT Academy of Engineering
Alandi, Dist. Pune, India
soham.kumbhar@mitaoe.ac.in

Vaishali C. Wangikar
School of Computer Engineering
MIT Academy of Engineering
Alandi, Dist. Pune, India
vcwangikar@mitaoe.ac.in

2024 IEEE 9th International Conference for Convergence in Technology (I2CT) | 979-8-3503-9447-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/I2CT61223.2024.10543812

***Abstract*—The paper discusses the importance of health and wellness in today's tech-dependent society and introduces the model named as "Food Diet Recaller App - FDRA". This app employs AI and computer vision to help users track and manage their diets effectively. It features food recognition, quantity estimation, and provides detailed nutritional information. The research utilizes techniques like image recognition and volume estimation, to accurately measure food intake. It also incorporates the InceptionV3 model for image classification an impressive accuracy of approximately 92% is achieved. The Ninja API offers nutritional information for classified food items. With the reference of aruco marker users capture images of their food, and the app calculates volume and density to determine nutritional content. The food volume estimation model is specifically trained for linear-shaped food items, characterized by square or rectangular surfaces. Additionally, the development of a dataset and regression models is mentioned to improve accuracy in estimating food mass. The paper also highlights the popularity of diet recall apps and their role in helping users maintain balanced nutrition.**

***Index Terms*—Health and Wellness, Food Diet Recall App, Food Recognition, Nutritional Information, Volume Estimation, Density Calculation, InceptionV3 Model**

I. Introduction

In today's health-conscious era, prioritizing a balanced diet and closely monitoring food consumption has become integral to personal well-being. Balanced Diet have surged the popularity, emerging as indispensable tools for individuals striving to achieve diverse health goals, whether it's managing weight, addressing chronic conditions like diabetes, or simply adopting a healthier lifestyle. Modern dietary assessment methods, including food image recognition and volume estimation, are revolutionizing nutritional monitoring, especially in a world of rising lifestyle diseases. The research serve as digital gateways to a more informed and intentional approach to nutrition, providing a means to meticulously record and analyze one's dietary habits. Food image recognition uses advanced computer vision techniques to identify and classify food items in images, deep learning algorithms to recognize dishes, ingredients, and portion sizes. The technology is crucial for diet consciousness, as it helps maintain a healthy lifestyle, make informed decisions about nutrition, and manage health conditions like diabetes and hypertension. It help individuals to monitor their food intake, promoting diet consciousness.Tracking food intake enhances awareness, accountability, and decision-making in diets. It helps analyze nutritional content, ensuring balanced protein, carbohydrate, fat, vitamins, and minerals for overall health.

The FDRA model leverages advanced technologies like ArUco markers and computer vision to enhance the user experience. [1] The FDRA uses ArUco markers to accurately estimate food volume, enhancing dietary tracking precision. It offers personalized meal plans, recipe suggestions, and monitoring of physical activity and water consumption. Additionally, the incorporation of image processing techniques, such as a record de-duplication algorithm, further refines the FDRA model, ensuring accurate and streamlined data for a comprehensive analysis of dietary habits. [2] This innovative tool revolutionizes dietary choices, allowing users to easily capture and analyze meals, gain insights into nutritional content, portion sizes, and trends, and empowering individuals to take control of their health and well-being.

II. Literature review

The study "Using distance estimation and deep learning to simplify calibration in food calorie measurement" introduces a novel approach to streamline food calorie measurement by fusing deep learning and distance estimation. Traditional methods involve cumbersome manual calibration prone to errors. [3] The method leverages deep neural networks and computer vision to accurately gauge the camera-to-food distance, a pivotal factor in calorie calculation. Through training on a comprehensive dataset of food images with known calorie values, system learns to estimate calorie content based on visual cues and distance data. Extensive trials across various food types, sizes, and calorie contents confirm this approach's accuracy, surpassing traditional calibration methods. This user-friendly method eliminates time-consuming manual steps, democratizing calorie measurement for individuals, dietitians, and food industry professionals. This study presents an innovative calibration technique for precise food calorie quantification, enabling informed dietary choices.

This research proposes an innovative approach for "Object Detection and Distance Measurement Using AI," integrating advanced models like YOLO or Faster R-CNN with AI-based distance estimation [4]. Trained on a large dataset, the system identifies objects in real-time, incorporating distance estimation for tasks like spatial mapping. Results show improved performance over traditional methods, with reduced sensor complexity and human calibration requirements. By simplifying calibration, the solution enhances accessibility, utility, accuracy, and efficiency, offering potential applications in surveillance technology.

The study presents an innovative method for robust object size measurement in challenging environments using computer vision and image processing. Overcoming issues like lighting variations and occlusions, the approach captures high-resolution images, preprocesses for noise removal, extracts object features, and estimates size through geometrical concepts and scale methods. The technique ensures precision by incorporating reference items or calibration markers, proving its utility in diverse fields such as robotics, construction, and manufacturing [5].

This study presents a novel method for real-time object detection and size measurement, overcoming previous limitations through current object detection algorithms and efficient measurement techniques [6]. The system operates in real-time, utilizing deep learning-based models like Faster R-CNN or YOLO for object identification in images or video streams. After segmentation and processing, the system employs feature extraction and geometric aspects for precise size estimation. To ensure real-time operation on embedded devices, optimization techniques such as model compression and hardware acceleration are incorporated. The study validates the system's accuracy through diverse datasets and real-world situations, emphasizing its applicability in robotics, autonomous vehicles, and surveillance systems. The paper explores potential applications across various domains.

This study proposes an innovative method for "A Novel Measurement Method for the Object Size and Shape Based on Image Recognition." The approach addresses challenges in detecting the characteristics from digital images and aims to provide a reliable solution applicable in manufacturing, quality control, and robotics. The method employs advanced image recognition algorithms to locate and distinguish objects in photographs. [7] Key features such as length, width, area, perimeter, and curvature are extracted using shape analysis techniques to determine the object's size and shape. The method incorporates techniques for handling occlusions, noise, and variations in lighting to guarantee measurement accuracy. Most importantly, the method offers flexibility by making it possible to modify measurement settings in line with specific application needs. With so many applications in industries requiring accurate item analysis, this study provides a workable approach to precisely characterise objects through photo recognition.

The paper "MUSEFood: Multi-sensor-based Food Volume Estimation on Smartphones" introduces an innovative system that utilizes smartphone sensors like the accelerometer and magnetometer to estimate food serving volumes. [8] Previous research on food volume estimation primarily relies on computationally intensive computer vision algorithms or specialized equipment like 3D scanners. Smartphone sensors have gained traction in health and fitness applications, such as monitoring physical activity and mental health screening. [7] Using smartphone sensors to estimate food volume has been investigated in studies. These sensors include the camera and accelerometer for monitoring food displacement and the microphone and accelerometer for sound analysis. The MUSEFood system stands out by integrating multiple smartphone sensors to provide more precise and reliable food portion size estimates, representing a significant advancement.

The research study presents a novel method for identifying and measuring the amount of food consumed while using a mobile device through "Recognition and Volume Estimation of Food Intake using a Mobile Device". The approach uses machine learning algorithms to recognise the object and computer vision techniques to photograph meals. [9] Using geometric concepts and photo analysis, the system determines the amount of food consumed. The approach provides a useful and precise tool for nutritional monitoring with food consumption tracking. Based on trial results, the system's accuracy in classifying a wide range of food items and predicting their quantities open the way for future applications in nutrition tracking, weight control, and personalised healthcare.

Thie study aims to develop a wearable sensor system for precise food quantity measurement by estimating the dimensions of plates and bowls. Using embedded sensors, the system captures images of the meal alongside the dishware and employs computer vision techniques to calculate the dimensions of the kitchenware. [10] Using the approach that analyses food size and shape in relation to the utensil, method determines plate and bowl sizes from gathered photographs. The accuracy of the method in estimating dimensions has been validates different standard-sized dishes by testing. The technology support dietary treatments and research in clinics, promote healthier eating habits, and make dietary monitoring easier. wearable sensors are a useful tool for precisely determining meal portion amounts.

The study introduces ”NutriTrack an Android-based food recognition app” designed to enhance nutritional awareness. Utilizing computer vision algorithms, the app identifies and classifies various food items based on pictures taken with a smartphone camera. NutriTrack employs a deep learning method, specifically a CNN model, trained on a vast database of food images. The CNN model identifies the food components with accuracy when users take pictures of their meals. Following that, NutriTrack offers detailed nutritional data, including calories, macronutrients and micronutrients. Users can track the nutrients they consume each day and record their meals. Tests confirmed the software identified and categorised food items. NutriTrack is an easy-to-use way to track nutrition, assisting people in choosing healthy foods.[11] The software provides real-time nutritional data using smartphone cameras and deep learning algorithms. This helps users make informed effective dietary decisions. According to the study, NutriTrack should be used in initiatives that support dietary therapies, weight loss, and healthy eating.

The Smart Chef project introduces an innovative approach to cooking with its personalized recipe recommendation system. Motivated by the demand for unique experiences and the growing interest in cooking, Smart Chef employs web mining and web scraping techniques to enhance its recipe suggestion engine.[12] The system uses web mining to collect food data from websites, food blogs, and online cooking groups. The process generates a vast library of recipes. By getting specific ingredient data, including names and measurements, from recipe websites, web scraping enhances. The collected data isgrouped for easy retrieval. Smart Chef evaluates user-supplied items by an ingredient-based algorithm that consider user preferences, compatibility, popularity, and online culinary knowledge. Research transforms the cooking experience by offering options based on their preferred ingredients and availability.

The work presents an innovative method for food volume estimation using virtual reality (VR) technology from a single image capture. [13] Accurate food volume assessment is crucial for various applications, including dietary evaluation, portion control, and meal planning. Leveraging the immersive capabilities of VR, this approach constructs a 3D virtual food model, calculating its volume based on user interactions. The paper describes the creation of a virtual reality that lets users use hand controls and VR headsets to interact with digital food models. In order to simulate portioning and rearrangement for getting depth and volume information, users can manipulate, slice, and study virtual food objects. The accuracy in calculating food volume was proven through testing across a variety of foods and portion sizes. The possible uses of VR system provides understanding of meal sizes and more interesting user experience for nutritional assessment, nutrition education, and customised meal planning are covered in the study. Accuracy could be increased by occlusions and intricate food structures through better modelling and machine learning methods. Overall, VR technology has potential to improve nutritional assessment and management through precision and efficiency.

The research introduce a method for precise 3D localization of circular objects within 2D images, with a potential application in estimating food volume. Accurate food volume measurement is vital for various health-related purposes, including nutrition assessment and portion control. The method leverages 3D localization techniques to identify and size circular food features in 2D images, enabling more accurate volume estimations. [14] The method makes volume calculations easier by determining the radii and 3D spatial coordinates of circular food features using a combination of geometric and statistical techniques. The accuracy is demonstrated using both actual and synthetic food photographs, under different conditions. The approach enables to compute meal volume effectively, which helps with portion management and dietary assessment. This 3D localization method offers precise and effective nutritional evaluation and quantity control, with potential applications in personalised meal planning and nutrition research.

The research ”Dimensional Analysis of Objects in a 2D Image. Advancements in Quantitative Image Processing” introduces innovative techniques for accurately measuring object dimensions in 2D images.[15] It signifies dimensional analysis includes computer vision.The study offers a framework incorporating automatic and human techniques, for improving accuracy using feature extraction, edge detection algorithms, and machine learning. The focus is on calibration techniques to lower measurement errors.

The research is a well planned and has solution that combines cutting-edge technologies and systematic methodologies to ensure seamless functioning and accuracy in food identification. The system development involves several key phases, including creating a wide and well-labeled collection of Indian food items, classifying food accurately using the InceptionV3 model, and accurately estimating meal volume using computer vision techniques and ArUco markers.The system increases the accuracy of dietary tracking by precisely estimating the volume of food by utilising ArUco markers. By employing the sophisticated approaches, the programme demonstrates a high accuracy and reliability in tackling the challenging difficulties of identifying and measuring ingredients from Indian cuisine. The model’s systematic approach to dataset curation and model training demonstrates its commitment to robustness and efficacy. The model has the potential to completely transform the way people manage their diets. Empower users with vital nutritional insights and portion control for better health and well-being.

III. Methodology

The methodology employed in developing the Food Diet Recall App is a multi-step process designed to ensure accurate tracking and management of users dietary intake. The initial phase involves the creation of a diverse Indian food dataset, comprising high-quality images of various dishes, each labeled with their respective class names. This dataset forms the basis for training the image classification model, utilizing the powerful InceptionV3 architecture Fig. 1. [16]

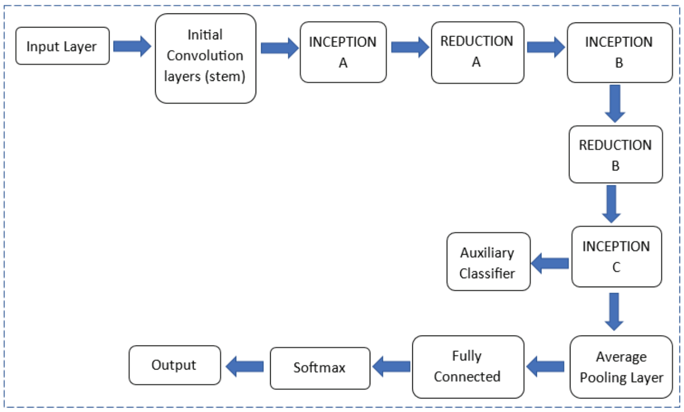

Fig. 1. InceptionV3 Architecture

The initial image is received by an input layer in Inception-v3, a powerful neural network for image classification. It uses 2D convolutions of varying filter sizes in the initial convolution layers to extract essential features. The Inception modules are at the heart of the architecture, consisting of parallel branches with various convolutions for tasks such as dimension reduction and spatial information capture. Reduction blocks reduce spatial dimensions even further, while auxiliary classifiers deal with the vanishing gradient problem. A pooling layer on average reduces dimensions, preparing data for fully connected layers. These layers process features in preparation for final classification, and the softmax layer generates classification probabilities, with the output layer providing the final prediction. This sturdy architecture performs admirably in image classification tasks. [17][18]

The app utilizes deep learning algorithms to enhance accuracy in food item classification, ensuring more precise data for comprehensive analysis and dietary recommendations. Once trained, this model enables the app to classify food items captured by users, providing valuable input for further analysis. To estimate food volume, a combination of computer vision techniques and ArUco markers are employed. Users are instructed to capture top and side view images of the food, and with the aid of known-size Aruco markers shown in fig. 2, the app extracts dimensions to calculate volume accurately.

Density calculation is crucial for converting volume to mass, and users are prompted to provide the specific density value for the Dhokla shown in fig. 3. This information, combined with the known volume, allows the app to calculate the mass of the consumed food accurately. After calculating the volume, the Nutritional API output will be recalculated. The results, including classified food name, nutritional content from the Ninja API, and calculated quantity consumed, are then presented to the user.

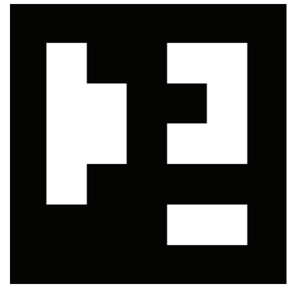

Fig. 2. Sample image of Aruco Marker

Name	Dhokla
Calories	254.4
Serving size (g)	100
Total fat (g)	12.9
Saturated Fat (g)	4.4
Protein (g)	9.8
Sodium (mg)	535
Potassium (mg)	167
Cholesterol (mg)	5
Total carbohydrates (g)	27.5
Fiber (g)	4.5
Sugar (g)	9.8

Fig. 3. Sample output of API Ninja's Nutritional API

To enhance user experience, the implementation utilizes the Streamlit framework, a Python library simplifying the creation of interactive, data-driven web applications. This ensures the app is intuitive and visually appealing, facilitating easy navigation and interaction. The app's deployment involves hosting it on a server or cloud platform, ensuring accessibility from various devices, such as smartphones, tablets, or computers, providing flexibility and convenience to users.Overall, this comprehensive methodology ensures precise tracking and management of dietary intake for users of the FDRA.

A. *FDRA Architecture*

As shown in fig. 4 and fig. 5, the Food Diet Recall App is designed with a structured approach, ensuring seamless functionality and accuracy in food recognition. The initial phase involves the creation of an Indian food dataset, a crucial step for training the image classification model. This dataset encompasses a diverse array of high-quality images, each meticulously labeled with their corresponding class names, and organized for optimal training efficiency.

Once the dataset is curated, the InceptionV3 model is employed for food classification, with its parameters fine-tuned through a rigorous training process using the dataset.

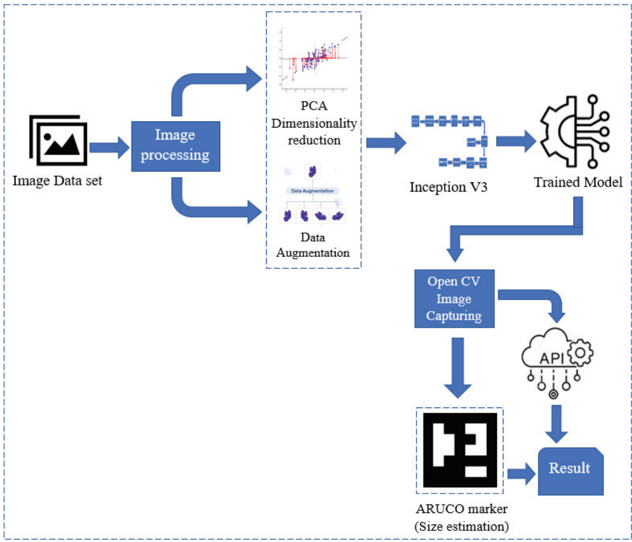

Fig. 4. FDRA Architecture Diagram

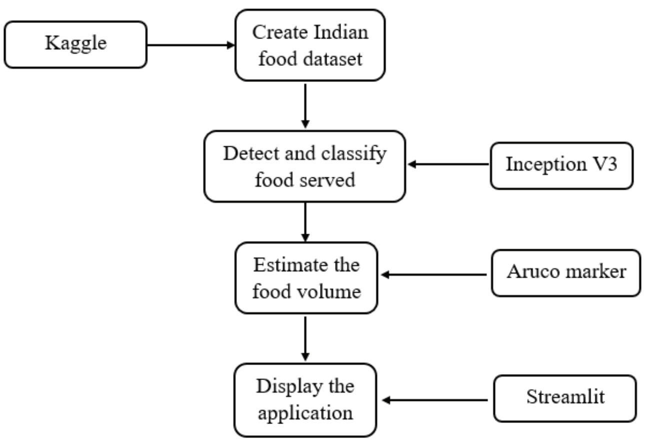

Fig. 5. FDRA Data Flow Diagram

This allows the model to make precise predictions when presented with images of Indian dishes captured by users.

To estimate food volume, computer vision techniques in conjunction with ArUco markers are utilized. Users capture both top and side view images of the food, which are then processed using OpenCV to extract accurate dimensions. By introducing a known-size ArUco marker alongside the food, the app calculates its volume by multiplying the three determined dimensions.

The final phase encompasses the deployment of the Food Diet Recall App, entailing the hosting of the application on a server or cloud platform, ensuring easy accessibility for users. This comprehensive architecture ensures the app's effectiveness in recognizing and quantifying Indian food items, offering a valuable tool for dietary monitoring.

B. *Calculating the Volume of Food*

After obtaining the dimensions from the top view and side view of the food with reference to the ArUco marker placed near the food, we can determine the mass of the food item using the formula:

$$\text{Mass} = \text{length (top view)} * \text{breadth (top view)} * \text{height (side view)} \tag{1}$$

After calculating the mass of the food item, we will take the density input from the user manually or user has to select from predefined density value and apply it for using the following equation for calculating the volume of the food.

$$\text{Volume} = \frac{\text{Mass}}{\text{Density}} \tag{2}$$

IV. Results and Analysis

A. *Food item color detection*

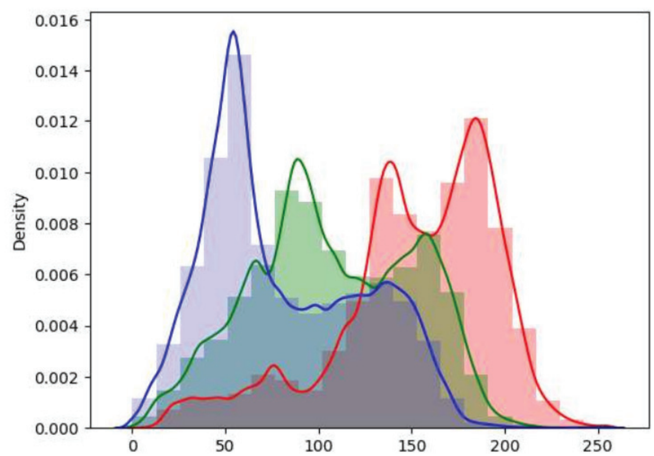

Fig. 6. Plot for visualizing pixel intensities for RGB in color space

The provided graph in fig. 6 uses the seaborn library to create three overlaid plots, each representing the distribution of pixel intensities in the Red, Green, and Blue channels of an image to check whether the rgb colors are normalized for the food dataset. This information can be valuable for several reasons:

- Color Balance: It helps in assessing whether the RGB values are balanced. An image with balanced color channels will have a relatively even distribution of pixel intensities across the channels.
- Color-Based Features: Food items often have distinctive colors. For example, a ripe banana is typically yellow, while a tomato is red. By examining the color distribution, a classifier can potentially learn to recognize foods based on their color characteristics.
- Color Normalization: If the images in the dataset have not been properly color normalized, it might introduce biases that could affect the classifier's performance. For instance, if one channel is consistently brighter or darker

across the dataset, the classifier might become overly sensitive to that channel.
- Color Variability: Food items can vary in color due to factors like lighting conditions, camera settings, and natural variations. Understanding the range of color intensities present in the dataset can help the classifier account for this variability.

If the RGB values in colour space are not normalised, histogram equalisation should be used to standardise the contrast and brightness levels across images. This can aid in the creation of consistent colour representations.

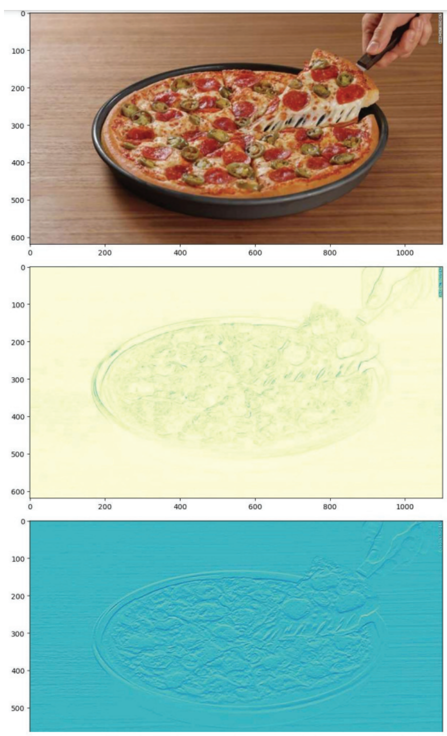

Fig. 7. Edge Detection

In fig. 7, the first subplot displays an image, the second subplot shows the Sobel edge detection applied to the blue channel, and the third subplot shows the horizontal Sobel edge detection applied to the green channel. Sobel filter is effective in highlighting edges or boundaries in an image. Edges represent areas where there is a sudden change in intensity, which often corresponds to object boundaries or significant features in an image. It demonstrates the effectiveness of the Sobel filter in detecting edges or boundaries in the image. The visualization allows for a qualitative assessment of the effects of different parameters or variations in the image processing steps. To check the variations in lighting conditions or color biases on the food image dataset. Other edge detection approaches, such as canny edge detection, are utilized when precise edge localization, noise reduction, and well-connected edges are required, such as in computer vision applications such as object recognition and image segmentation.[19] However, when computing efficiency is a requirement and a simpler edge detection is acceptable. It is appropriate for real-time applications or circumstances requiring rapid, sobel edge detection.

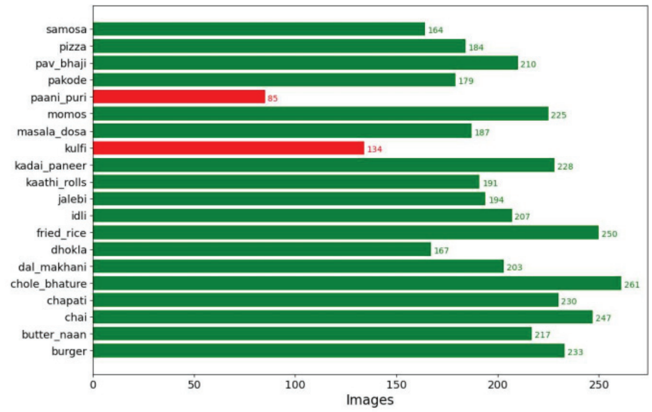

Fig. 8. Selection of food item for model training

Fig. 8 visually represents image distribution across various food categories. Green bars indicate categories with over 150 images, while red bars represent those with fewer than 150 images. Notably, during training, there was a shortage of images for kulfi and pani puri. To address this, data augmentation techniques were applied, such as random rotation, image resizing, and shearing. These techniques expanded the kulfi and pani puri datasets, overcoming the initial image scarcity and creating a more diverse representation. Data augmentation not only solved the image shortage but also improved the model's generalization. The augmented dataset allowed the model to better handle variations in scale, orientation, and other factors, resulting in a more robust classifier.

TABLE I
COMPARATIVE ANALYSIS

Food	Measured Serving (g)	Observed Serving (g)
Dhokla	81	79.6
Pizza (Whole)	198	3404.9
Pizza (Slice)	53.7	248.2

This model was tested on a variety of food items, including pizza and dhokla, chosen for their distinct shapes, sizes, and textures. Dhokla is known for its simple texture, whereas pizza is known for its varied and not always simple composition. We manually measured the food items with a weighing machine before running the model. The model yields accurate results for plain and edgy surfaces, but it struggles to yield accurate results for food items with irregular and varied shapes. Table 1 compares the results of the measured serving of the food item calculated with a food weighing machine and the observed serving calculated by the food model. The results for dhokla are quite similar, whereas the results for food items with irregular shapes differ more.

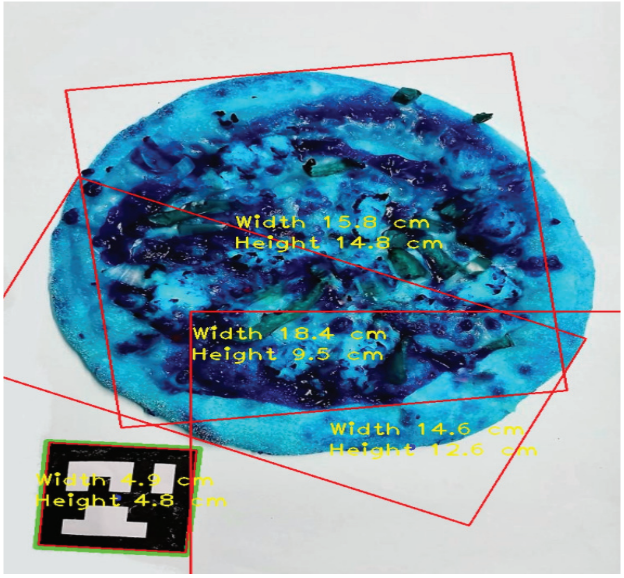

Fig. 9. Volume Estimation of Pizza

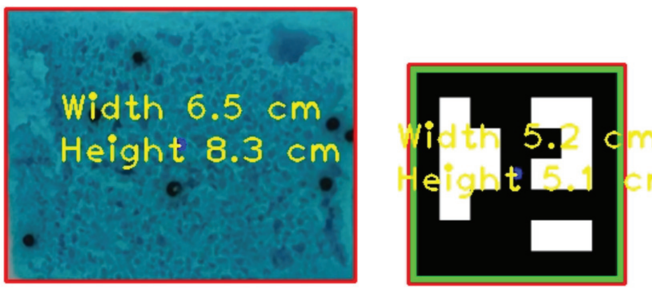

Fig. 10. Volume Estimation of Dhokla

The food volume estimation model yields accurate results when assessing the volume of food items with plain and edgy surfaces. However, its precision diminishes when applied to food items with irregular surfaces or non-linear shapes. As observed in Fig. 9 and Fig. 10, the model performs correctly and accurately identifies the edges for the dhokla food item. Conversely, when conducting volume analysis for pizza, it struggles to identify the edges accurately. Due to this, the performance is compromised when dealing with food items that have irregular surfaces or non-linear shapes, such as pizza, burger, etc, where the model struggles to identify edges correctly. This suggests a limitation in the model's ability to handle diverse and complex food shapes, and further refinement or adaptation may be necessary to improve its accuracy across a broader range of food items.

Fig. 11 and 12 illustrate the experimental analysis conducted on food items having well defined linear and edgy geometric structure, exemplified by Dhokla and samosa, across varying weights. The comparison involves assessing the calorie content determined by standard norms against the calorie predictions generated by the Food Detection and Recognition Algorithm-FDRA. The results are striking in that they show a high degree of consistency, with an accuracy rate of almost 92%. This observation leads to the compelling conclusion that FDRA exhibits notable results when applied to food items characterized by distinct edges. The algorithm's capacity to accurately predict calorie content underscores its reliability and potential utility in nutritional assessments. This suggests that FDRA holds promise as a valuable tool for quantifying and analyzing the nutritional composition of food items, particularly those with well-defined geometric features.

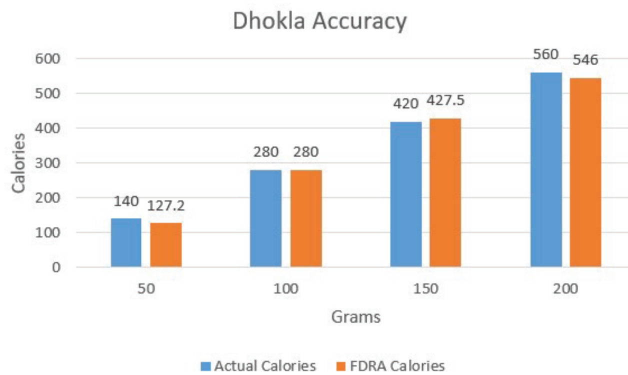

Fig. 11. Accuracy analysis for linear food item (with edges): Dhokla

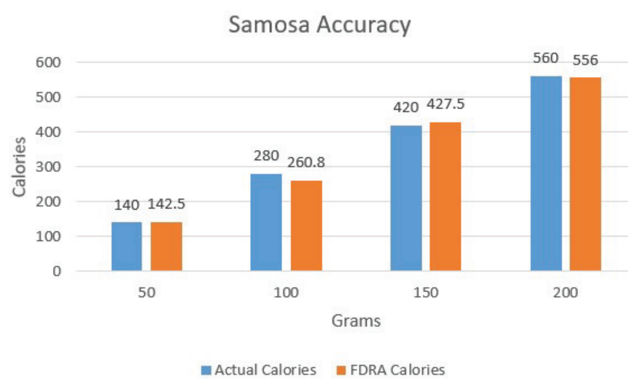

Fig. 12. Accuracy analysis for linear food item (with edges): Samosa

V. Conclusion

The Food Diet Recall App represents a meticulously designed solution that integrates state-of-the-art technologies and systematic approaches to ensure seamless operation and precision in food identification and volume estimation. The model leverages InceptionV3 for food item identification and aruco markers for food volume estimation.

During food identification with the InceptionV3 model, an impressive accuracy of approximately 92% is achieved. However, when estimating food volume using aruco markers, the model excels with higher accuracy for food items exhibiting linear and regular shapes. Unfortunately, its performance declines when dealing with food items of irregular and non-linear shapes. This is primarily attributed to the challenge of inaccurate edge detection during the food volume estimation phase.

The food volume estimation model is specifically trained for linear-shaped food items, characterized by square or rectangular surfaces. In instances where irregularly shaped food items are presented, the model successfully identifies edges but struggles to accurately delineate them on the food item. Notably, in Fig. 9, numerous squares are erroneously drawn over the pizza. The model tends to select the square with the highest height and width for calculating food volume, leading to potential inaccuracies. Additionally, the use of squares impacts surface area calculations, particularly for circular food items, as they possess distinct surface areas compared to squares.

The food estimation model also encounters difficulties in accurately estimating the volume of food items with small surfaces, like the sides of a pizza. This limitation contributes to the generation of inaccurate results. To improve the model's accuracy, it is imperative to incorporate edge detection capabilities for irregularly shaped food items. By enabling the model to detect and account for the edges of such items, it can more effectively scale and analyze the volumes of diverse food shapes, including those with small and intricate surfaces. Overall, the food estimation model performs quite well for square and rectangular surfaced objects.

VI. Future Scope

The current research has accurately estimated the mass of food items, with rectangular and square food items. With image classification, object detection, and volume estimation algorithms the mass detection and calorie calculations are quite comparable to standard methods.

In the future, we hope to expand our research to handle various shaped food items, including irregular shapes. To achieve this, point cloud data and mesh grid generation approaches can be used. Point clouds are collections of data points in a three-dimensional coordinate system that can be obtained using methods such as 3D scanning or depth sensing. We can reconstruct the 3D model of the food item by processing the point cloud data, allowing for more accurate analysis and volume estimation for irregular shapes.These 3D models can provide additional information about the shape, volume, and dimensions of the food. We can improve the accuracy of object detection, volume estimation, and nutritional content prediction for a wider range of food items by combining information from 3D models with our existing algorithms.

References

[1] ArUco Marker-ArUco markers and boards detection for robust camera pose estimation

[2] Vaishali Wangikar, Sachin Deshmukh, and Sunil Bhirud. 2016. Study and Implementation of Record De-duplication Algorithms. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies (ICTCS '16). Association for Computing Machinery, New York, NY, USA, Article 8, 1–6. https://doi.org/10.1145/2905055.2905063.

[3] P. Kuhad, A. Yassine and S. Shimohammadi, "Using distance estimation and deep learning to simplify calibration in food calorie measurement," 2015 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Shenzhen, China, 2015, pp. 1-6, doi: 10.1109/CIVEMSA.2015.7158594.

[4] M. M. Faisal et al., "Object Detection and Distance Measurement Using AI," 2021 14th International Conference on Developments in eSystems Engineering (DeSE), Sharjah, United Arab Emirates, 2021, pp. 559-565, doi: 10.1109/DeSE54285.2021.9719469.

[5] A. F. Said, "Robust and Accurate Objects Measurement in Real-World Based on Camera System," 2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 2017, pp. 1-5, doi: 10.1109/AIPR.2017.8457954.

[6] N. A. OTHMAN, M. U. SALUR, M. KARAKOSE and I. AYDIN, "An Embedded Real-Time Object Detection and Measurement of its Size," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-4, doi: 10.1109/IDAP.2018.8620812.

[7] J. He, Z. Tao, J. Zhang, Y. Tan, L. Xie and Y. Xiong, "A Novel Measurement Method for the Object Size and Shape Based on Image Recognition," 2021 International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS), Xi'an, China, 2021, pp. 610-613, doi: 10.1109/ICITBS53129.2021.00155.

[8] M. Puri, Zhiwei Zhu, Q. Yu, A. Divakaran and H. Sawhney, "Recognition and volume estimation of food intake using a mobile device," 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 2009, pp. 1-8, doi: 10.1109/WACV.2009.5403087.

[9] M. Puri, Zhiwei Zhu, Q. Yu, A. Divakaran and H. Sawhney, "Recognition and volume estimation of food intake using a mobile device," 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 2009, pp. 1-8, doi: 10.1109/WACV.2009.5403087.

[10] V. B. Raju, D. Hossain and E. Sazonov, "Estimation of Plate and Bowl Dimensions for Food Portion Size Assessment in a Wearable Sensor System," in IEEE Sensors Journal, vol. 23, no. 5, pp. 5391-5400, 1 March1, 2023, doi: 10.1109/JSEN.2023.3235956.

[11] A. B. Ocay, J. M. Fernandez and T. D. Palaoag, "NutriTrack: Android-based food recognition app for nutrition awareness," 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2017, pp. 2099-2104, doi: 10.1109/CompComm.2017.8322907.

[12] S. Chaudhari, R. Aparna, V. G. Tekkur, G. L. Pavan and S. R. Karki, "Ingredient/Recipe Algorithm using Web Mining and Web Scraping for Smart Chef," 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2020, pp. 1-4, doi: 10.1109/CONECCT50063.2020.9198450.

[13] Zhengnan Zhang, Yongquan Yang, Yaofeng Yue, J. D. Fernstrom, Wenyan Jia and Mingui Sun, "Food volume estimation from a single image using virtual reality technology," 2011 IEEE 37th Annual Northeast Bioengineering Conference (NEBEC), Troy, NY, 2011, pp. 1-2, doi: 10.1109/NEBC.2011.5778625.

[14] W. Jia, Y. Yue, J. D. Fernstrom, Z. Zhang, Y. Yang and M. Sun, "3D localization of circular feature in 2D image and application to food volume estimation," 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 2012, pp. 4545-4548, doi: 10.1109/EMBC.2012.6346978.

[15] A. S. Parihar, M. Gupta, V. Sikka and G. Kaur, "Dimensional analysis of objects in a 2D image," 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Delhi, India, 2017, pp. 1-7, doi: 10.1109/ICCCNT.2017.8203937.

[16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.

[17] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi , Google Inc.Rethinking the Inception Architecture for Computer Vision .2015,Thirty-First AAAI Conference on Artificial Intelligence.

[18] Dr. Domonkos Varga, "Multi-Pooled Inception Features for No-Reference Image Quality Assessment." Department of Networked Systems and Services, Budapest University of Technology and Economics, Budapest H-1111, Hungary,23 March 2020.

[19] Chikmurge, Diptee & Harnale, Shilpa. (2018). Feature Extraction of DICOM Images Using Canny Edge Detection Algorithm. 10.1007/978-981-10-5520-1-18.