International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

A Review of Image-Based Food Recognition and Volume Estimation Artificial Intelligence System

Aman Kumar Tyagi [1]

[1] M. Tech. in Computer Science and Engineering, Neelam College of Engineering and Technology, Korai, Agra.

Mr. Ankit Garg [2]

[2] Assistant Professor, Department of Computer Science and Engineering, Neelam College of Engineering and Technology, Korai, Agra.

**Corresponding Email: amantyagi123@gmail.com*

ABSTRACT

The Image-Based Food Recognition and Volume Estimation Artificial Intelligence System is designed to automatically identify food items and estimate their physical and nutritional properties using computer vision and machine learning techniques. The system utilizes image processing methods such as feature extraction, image segmentation, and scale estimation to analyze food images. The Bag of Features (BoF) technique is used to extract visual features, and a Support Vector Machine (SVM) classifier is applied to recognize different food categories including apple, banana, pizza, roti, and rice. The system also estimates food volume using geometric modeling and scale calibration, which is further used to approximate weight and calorie content. A graphical user interface (GUI) developed in MATLAB R2017b allows users to easily load images, perform recognition, and view results. Experimental results demonstrate high classification accuracy and efficient performance. The proposed system provides an automated, reliable, and user-friendly solution for food recognition and nutritional estimation, with potential applications in healthcare, dietary monitoring, and intelligent food management systems.

Keywords: *Food Recognition, Computer Vision, Machine Learning, Volume Estimation, Support Vector Machine, Image Processing, Nutrition Estimation.*

I. Introduction

The rapid growth of artificial intelligence (AI), computer vision, and machine learning has revolutionized the ability of machines to interpret visual data and make intelligent decisions. One important application of these technologies is the automatic analysis of food images. Food recognition and portion estimation play a critical role in nutrition monitoring, healthcare management, dietary assessment, and intelligent automation systems. Traditionally, identifying food items and estimating their quantity requires manual observation, weighing scales, and expert knowledge, which can be time-consuming, subjective, and prone to human error. The development of an image-based artificial intelligence system offers a more efficient, accurate, and automated alternative [1-5].

1.1 Image-Based Food Recognition

Image-based food recognition involves identifying food items from digital images using computer vision and machine learning techniques. The system analyzes visual features such as color, texture, shape, and spatial patterns to classify food into predefined categories. Feature extraction methods, such as Bag of Features (BoF), Scale-Invariant Feature Transform (SIFT), and Speeded-Up Robust Features (SURF), allow the system to detect important visual characteristics. These features are then used by classification algorithms such as Support Vector Machines (SVM) or deep learning models to accurately recognize different food types. This automated recognition process eliminates the need for manual labeling and enables fast and reliable identification of food items.

1.2 Volume Estimation from Images

Estimating the volume of food from a two-dimensional image is a challenging task because images do not directly contain depth information. To overcome this limitation, computer vision techniques such as image segmentation, scale calibration, and geometric modeling are used. Image segmentation isolates the food region from the background, while scale calibration converts pixel measurements into real-world units using reference objects such as plates or bowls. Geometric approximation methods then estimate the food's volume based on its shape and dimensions. This allows the system to calculate physical properties such as volume, which can be further used to estimate weight and portion size.

1.3 Weight and Calorie Estimation

Once the volume of the food is determined, the system can estimate its weight and caloric value using known density and nutritional data. Each type of food has a specific density and calorie content per unit weight. By multiplying the estimated volume by the density, the system calculates the approximate weight. Similarly, calorie estimation is performed using standard nutritional values, enabling automated dietary assessment. This functionality is particularly useful for individuals monitoring calorie intake, healthcare professionals managing patient diets, and fitness applications.

1.4 Role of Artificial Intelligence and Machine Learning

Artificial intelligence and machine learning algorithms are central to the performance of the system. Machine learning models learn from training data to recognize patterns and classify food items accurately. Computer vision algorithms enable object detection, segmentation, and measurement. The integration of these technologies creates an intelligent system capable of performing complex tasks such as food recognition, volume estimation, and nutritional analysis automatically. As AI models improve with more data and advanced algorithms, the accuracy and reliability of such systems continue to increase.

1.5 Applications of the System

Image-based food recognition and volume estimation systems have numerous practical applications. In healthcare, they assist in monitoring dietary intake and managing conditions such as diabetes, obesity, and cardiovascular diseases. In smart kitchens, they support automated portion control and meal planning. These systems can also be used in robotics, automated food service, intelligent vending machines, and mobile health applications. Additionally, they are valuable in research areas related to nutrition, public health, and human computer interaction [6-11].

II. Literature Review

Karthikeyan et al. (2026) had emphasized that nutrition and diet had played pivotal roles in promoting a healthy lifestyle, and they had argued that effective dietary management had required accurate identification and analysis of consumed food items. They had noted that manual dietary tracking had been inefficient and error-prone, and they had proposed a deep learning–based framework for automated food recognition and volume estimation. Their framework had used Inception V3 for food classification, Mask R-CNN for two-dimensional area estimation, and MiDaS for single-view depth estimation. They had reported an average classification accuracy of 98% and had indicated that the system had produced reliable nutritional and volumetric assessments. They had also compared multiple architectures on benchmark datasets and had reported accuracies of 98% (Food101), 97.25% (UEC-Food256), and 97.40% (Indian Food Images Dataset), thereby supporting precise dietary monitoring.

Ji et al. (2026) had addressed the low accuracy of conventional methods for estimating liquid volume and nutrient content in bowl-type tableware, and they had highlighted that manual measurement had been tool-dependent and time-intensive. They had proposed an integrated framework that had combined geometric reconstruction with deep learning–based segmentation. After a one-time camera calibration, their pipeline had required only frontal and top-down bowl images, and it had automatically extracted geometric parameters (rim diameter, base diameter, bowl height, and inner-wall profile) to enable geometric modelling and capacity estimation. They had stored these parameters in a reusable bowl database to support repeated predictions at different fill levels. They had introduced Bowl Thick Net for millimetre-level wall-thickness prediction and had integrated a Geometry-aware Feature Pyramid Network (GFPN) into an enhanced Mask R-CNN for contour segmentation. They had reported mean errors of −3.03% (capacity estimation), 9.24% (liquid-volume prediction), and 11.49% (nutrient-content estimation).

Chopra and Purwar (2026) had identified calorie estimation as a crucial yet challenging component of automated dietary monitoring systems, and they had proposed a hybrid, multi-view framework for calorie estimation from food images. Their approach had integrated enhanced segmentation, classification, and portion-size estimation using deep learning. They had applied the Enhanced Squirrel Search Algorithm (ESSA) for optimized image segmentation and had used a CNN for food classification and feature extraction. For portion-size estimation, they had applied a grid-based superimposition method with perspective transformation to top and side views to compute scale factors and support volumetric analysis. They had evaluated the framework on datasets including

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

UEC-Food101 and Nutrition5k, and they had reported a mean relative error of 3.95% in calorie prediction, suggesting that multi-view geometric analysis with feature-level fusion had improved robustness and scalability.

Kuang et al. (2026) had argued that precise nutrition estimation had been essential for dietary management and disease prevention, and they had identified the integration of heterogeneous multi-source information as a key challenge. They had proposed RDINet, a multimodal network that had fused RGB appearance, depth-based geometry, and ingredient semantics for nutrition estimation. They had described two main modules: an RGB-D fusion module that had combined RGB textures with depth-based 3D shape cues using channel–spatial attention, and an ingredient fusion module that had integrated ingredient information into visual features via attention to support reasoning about visually ambiguous components. On Nutrition5k, they had reported percentage mean absolute errors of 14.9% (calories), 11.2% (mass), 19.7% (fat), 18.9% (carbohydrates), and 19.5% (protein), with an overall mean PMAE of 16.8%, indicating improved performance over prior approaches.

Luo (2025) had noted that accurate monitoring of dietary nutrition had been increasingly important for preventing and managing chronic diseases, and they had linked precise food-volume estimation to the accuracy of energy-intake calculations. They had argued that traditional volume-estimation methods had faced limitations in portability, real-time performance, and user acceptance due to factors such as variable food shapes, lighting, and occlusion. They had described how deep learning–based computer vision had made image-based food volume estimation a major research focus, while also noting unresolved challenges related to efficiency, robustness under occlusion, adaptation to complex backgrounds, and single-image accuracy—especially for mobile and consumer applications. They had indicated that deep learning methods had been used to identify and segment food regions and that reference-object sizing had supported volume estimation.

Gonzalez et al. (2025) had aimed to quantify key food components in catering services using computer vision and AI, with a focus on content identification and portion-size estimation in dining hall environments. They had used an RGB camera to capture tray images in a self-service restaurant and had evaluated identification algorithms using standard metrics. They had trained a YOLO-based detector on labelled data and had reported a mean Average Precision (mAP) of 0.873 at a 0.5 confidence threshold. For weight estimation, they had estimated food volume using RGB and depth cameras and had converted volume to weight using food-specific density models calibrated through volume-to-weight tables. They had validated the approach on rice and chicken and had reported error margins of 5.07% and 3.75%, respectively, supporting the feasibility of the method.

Nogay et al. (2025) had discussed chronic diseases such as obesity and hypertension as preventable through appropriate diets, portion control, and healthy eating habits, and they had suggested that automated intake-measurement systems could support nutritional adequacy assessment and timely intervention across culturally diverse foods. They had developed a deep learning system for automatic food grouping and classification using Turkish dishes as a test sample. Using CNN-based recognition, they had reported accuracies of up to 80% for classification and 80.47% for portion estimation, and they had indicated that data augmentation had further improved performance.

IJAMSR
International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Bisht et al. (2025, August) had highlighted the need for precise dietary and lifestyle management for individuals with prediabetes and related health concerns, and they had argued that traditional approaches had often misestimated actual consumption. They had proposed a method combining image recognition with volume estimation to improve sugar quantification. Their approach had used monocular depth estimation for volume measurement and Mask R-CNN for segmentation, after which estimated volume had been converted to weight and sugar content had been computed using nutritional databases. They had reported reliable segmentation and volume prediction outcomes across multiple food images and had concluded that the approach could enhance consumer and clinical nutrition-management tools.

Sari (2025) had focused on hospitalized patients and had argued that accurate nutrition support had been important for recovery. They had described limitations of traditional calorie-consumption estimation methods (e.g., weighing, Comstock technique, and digital imaging), particularly subjective human judgement. To reduce human-factor error, they had proposed an objective image-based method using RGB-Depth images. They had combined segmentation and edge detection on RGB images with depth information, converted segmented regions into point clouds, and estimated volume by fitting an ellipsoid. They had reported a minimum Mean Average Percentage Error (MAPE) of 2.73, indicating sufficiently accurate volume measurement.

Maharana et al. (2025, March) had reviewed image-based dietary assessment and had emphasized that, despite progress in food recognition, weight estimation had remained challenging. They had summarized methods targeting error reduction and recent innovations, and they had compared 2D and 3D approaches while highlighting benefits of hybrid multimodal models. They had emphasized the role of diverse, well-annotated datasets and had discussed evaluation metrics and benchmarks for comparability. They had identified emerging directions such as self-supervised learning, multimodal integration, and real-time edge deployment as potential solutions to food variability, occlusion, and environmental constraints.

Purandhar et al. (2025, June) had noted that AI and deep learning had been increasingly used for food recognition and calorie estimation, but they had argued that mixed meals (multi-ingredient dishes) had remained difficult for many models. They had proposed a system for mixed-meal recognition and calorie estimation using deep learning for object detection, segmentation, and volume estimation. They had employed a MobileNet-based multi-label detector and semantic segmentation to identify individual components within a meal. They had reported high accuracy on a large heterogeneous mixed-food dataset, had incorporated explainable AI to improve transparency and trust, and had implemented the solution in a mobile application for real-time monitoring and personalized feedback.

Cheng et al. (2025) had highlighted that traditional manual food logging and nutritional lookup had been time-consuming and error-prone, and they had proposed an image-based nutritional advisory system using multimodal deep learning for classification, volume estimation, and dietary recommendation. They had fine-tuned CLIP for zero-shot food recognition and had used a learning-based multi-view stereo (MVS) method for volume estimation without specialized hardware. They had derived nutrition values from verified food composition databases and had generated

IJAMSR
International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

personalized guidance using a large language model (Llama 3). They had reported top-1 accuracies of 91% (CNFOOD-241) and 80% (Food101), a volume-estimation MAPE of 23.5%, and recommendation quality with a BLEU-4 score of 45.13.

Liu et al. (2025) had reviewed deep learning in food image recognition and had described food recognition as a core task in fine-grained image analysis with applications in smart dining, healthcare, and retail. They had traced the evolution from manual feature methods to deep learning methods, systematically organized major datasets, analyzed typical model families, and discussed applications such as calorie estimation and food safety. They had identified ongoing challenges and proposed future research directions for improving robustness and real-world applicability.

Zhang et al. (2024) had emphasized that monitoring eating behaviours had supported personalized diet and exercise planning and could help prevent chronic diseases. They had reviewed image-based computer vision and AI methods for dietary assessment, including food classification, portion-size estimation, and nutrient inference. They had noted that accuracy had remained challenging due to variability in food shape, appearance, preparation, and presentation, and they had summarized datasets, detection algorithms, performance trends, limitations, and research gaps for future work.

Nabitchita et al. (2024) had conducted a systematic review of image processing techniques for object-volume measurement and had noted that mobile and high-resolution cameras had expanded opportunities for volume estimation research. Using an NLP-based review framework, they had screened literature from 2010–2023 and had analyzed 25 studies, reporting that diverse computer vision methods had been applied effectively for recognition-related tasks. They had compared datasets and methodological trends, and they had highlighted strengths and limitations to guide future volume-inference designs.

Mao et al. (2024, October) had discussed AI-based dietary assessment and had argued that accurate nutrient estimation had depended on reliable recognition, effective weight estimation, and comprehensive food composition databases. They had described common deep learning practices (data augmentation, dataset splits, batch normalization, optimizers such as SGD/Adam/RMSprop) and had noted the widespread use of CNNs for extracting shape, texture, and color features. They had also noted the use of traditional image processing for weight measurement and enhancement, and they had emphasized updating food databases and addressing privacy and usability in future work.

Gonzalez et al. (2024, July) had proposed a methodology for food-weight estimation from images by combining advanced segmentation with volumetric estimation. They had applied the method to an extended rotisserie-chicken dataset, had refined training, and had reported that an Exponential Gaussian Regression model had achieved strong predictive performance. They had reported an average error of 2.86% under cross-validation, suggesting usefulness for food-industry automation and quality control.

Gaonkar and Sangeetha (2024, June) had emphasized the health relevance of calorie monitoring for disease prevention and had discussed the growing role of deep learning for food image analysis. They had compared techniques for food segmentation, identification, and calorie estimation and had assessed performance differences across methods, concluding that deep learning approaches had shown strong potential for improving automated dietary assessment.

IJAMSR
International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Konstantakopoulos et al. (2023) had reviewed smartphone-based and computer-vision approaches for automatic nutrient estimation and had framed real-time estimation as a computer vision problem using meal images. They had organized methods across segmentation, classification, and volume computation, and they had linked performance to dataset characteristics. They had discussed strengths, limitations, and practical directions for future automated dietary assessment research.

Mansouri et al. (2023) had conducted a systematic review of deep learning in food recognition and nutrient assessment and had organized the literature into classification, segmentation, and volume estimation for nutrition calculation. They had synthesized 57 original articles, reviewed both public and private datasets, and had concluded that deep learning approaches had often outperformed conventional methods while still facing challenges that had motivated future research directions.

Abdur Rahman et al. (2023, October) had compared sensor-, geometry-, and neural network–based methods for food volume computation and had noted that volume estimation had remained a key bottleneck despite progress in segmentation and recognition. They had evaluated methods on 20 meal images captured under diverse conditions, with ground-truth volumes derived from 360-degree imaging and 3D computation. They had reported strengths and limitations across approaches and had released an open dataset (RGB images, point clouds, and volumes) to address scarcity of depth-informed food data.

Dubey et al. (2023, September) had reviewed food tracking technologies and had emphasized the value of camera-based systems for ingredient recognition and quantity estimation. They had noted that camera-based approaches had varied in accuracy and speed and had highlighted emerging volume-sensitive sensing methods such as LiDAR and true-depth imaging as promising directions to improve precision in automated dietary assessment.

Huang et al. (2023, July) had presented an AI nutrition analysis platform for compartment trays using an image-collection system (Raspberry Pi + depth camera) and a cloud-based analysis pipeline. They had implemented a three-stage process (dish recognition, portion-size estimation, nutrient calculation). They had reported strong instance-segmentation performance, noting that CenterMask with a VoVNetV2-99 backbone had achieved high AP metrics, and they had described depth preprocessing with TELEA-based repair achieving low variation. Across 14 dishes, they had reported a mean absolute error of 15.68 kcal and a mean relative error of 22.29%, comparable to or better than dietitian estimates.

Amugongo et al. (2022, December) had reviewed mobile computer-vision solutions for food recognition, volume estimation, and calorie calculation, and they had evaluated explainability support in such applications. They had found that most applications had not explicitly separated food from non-food items and that explainable components had been rare, concluding that improving user trust would require stronger explainability in classification, volume estimation, and caloric prediction.

Lo et al. (2022) had discussed ongoing challenges in deep learning–based food volume estimation, including domain shift between synthetic and real 3D data and limited paired training sets. They had proposed a weakly supervised point-cloud completion method to support volume estimation and had

IJAMSR
International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

emphasized that the method could estimate handheld-food volumes without table placement or fiducial markers. They had reported competitive results on benchmark datasets and their own volume-annotated dataset, with improved completion and estimation quality.

Haque et al. (2022) had argued that real-time processing speed had been a critical barrier for mobile calorie-estimation applications, and they had noted a shortage of lightweight, fast, reliable systems. They had proposed a parameter-optimized CNN approach to detect food items from handheld images and to compute calories using prior class knowledge, reporting high accuracy and improved automation for real-time estimation.

Tahir and Loo (2021, December) had reviewed machine learning approaches for automatic dietary intake assessment and had highlighted limitations of manual logging, including underreporting and low adherence. They had summarized methods using popular food datasets and had reported that many studies had relied on deep neural features and CNN variants for ingredient recognition. They had outlined future directions such as unsupervised learning from unlabeled images, continual-learning stability, and improved interpretability through explainable AI.

Yang et al. (2021) had proposed an AI system for portion-size estimation that had mimicked dietitians' use of familiar objects as references. They had used learned internal reference volumes to produce a probability vector and had computed the final volume estimate via an inner product with reference volumes. They had reported accurate performance on both virtual and real food datasets, supporting practical dietary assessment.

Lo et al. (2020) had discussed limitations of 24-hour dietary recall due to subjective self-reporting and resulting bias in food type and portion estimates. They had reviewed visual-based dietary assessment methods and had compared state-of-the-art approaches for recognition and volume/weight estimation in terms of speed, accuracy, efficiency, and constraints. They had concluded that integrated multi-method systems had offered potential for improved dietary intake evaluation.

Allegra et al. (2020) had reviewed automatic food intake monitoring with an emphasis on computer vision tasks such as classification, recognition, segmentation, and portion estimation. They had described a systematic screening process across major databases and had synthesized selected studies on food logging systems and related vision tasks. They had identified limitations and research gaps and had discussed directions for advancing food analysis from digital media.

III. Methodology

System Design and Framework: The proposed Image-Based Food Recognition and Volume Estimation Artificial Intelligence System is designed as an integrated computer vision and machine learning framework capable of automatically identifying food items and estimating their physical and nutritional properties from images. The system is implemented using MATLAB R2017b and incorporates a graphical user interface (GUI) to provide an interactive and user-friendly platform. The overall architecture consists of several sequential modules, including image acquisition, preprocessing, feature extraction, classification, segmentation, scale estimation, volume estimation, and nutritional analysis. Each module performs a specific function, and together they form a complete pipeline that enables automatic recognition and measurement of food items. The modular design ensures flexibility, scalability, and efficient processing, allowing accurate analysis of food images under different conditions [12-21].

Data Collection and Dataset Preparation: The dataset used in this research consists of multiple food categories, including apple, banana, pizza, rice, and roti. Images were collected from various sources and organized into labeled folders corresponding to each food category. Each class contains multiple images captured under different lighting conditions, angles, and backgrounds to improve the robustness and generalization capability of the system. Proper labeling and organization of the dataset are essential for supervised machine learning, as the classifier learns to associate visual features with specific food categories. The dataset is divided into training and testing sets to train the classification model and evaluate its performance.

Image Preprocessing: Image preprocessing is performed to enhance image quality and improve the accuracy of feature extraction and segmentation. This stage includes resizing images to a consistent resolution, reducing noise using filtering techniques such as Gaussian filtering, and improving contrast through histogram equalization. Additionally, images are converted from RGB color space to grayscale or Lab color space to facilitate better feature extraction and segmentation. These preprocessing steps ensure that irrelevant variations such as noise, illumination differences, and background complexity are minimized, allowing the system to focus on relevant visual features.

Feature Extraction Using Bag of Features: Feature extraction is a critical stage in identifying distinctive characteristics of food items. In this research, the Bag of Features (BoF) approach is used to extract and represent image features. SURF (Speeded-Up Robust Features) is employed to detect key points and extract feature descriptors from images. These descriptors capture important visual information such as edges, textures, and shapes. The extracted features are clustered using the K-means clustering algorithm to form a visual vocabulary. Each image is then represented as a histogram of visual words, providing a compact and meaningful representation suitable for classification.

Food Classification Using Machine Learning: The classification module uses a Support Vector Machine (SVM) classifier trained on the extracted feature representations. SVM is selected due to its high accuracy and effectiveness in handling high-dimensional feature data. During training, the classifier learns patterns associated with different food categories. During testing, the trained model predicts the food category of a new image and provides confidence scores for each class. The classification result identifies the food item, which is essential for selecting appropriate geometric and nutritional parameters in subsequent stages.

Image Segmentation: Image segmentation is used to isolate the food region from the background. This process involves converting the image into an appropriate color space and applying thresholding techniques such as Otsu's method to separate the foreground from the background. Morphological operations, including opening, closing, and hole filling, are applied to remove noise and refine the segmented region. The largest connected component is typically selected as the food object. Accurate segmentation is essential for calculating the area, shape, and dimensions of the food item.

IJAMSR

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Scale Estimation and Calibration: Scale estimation is necessary to convert pixel measurements into real-world physical units. The system uses automatic plate or bowl detection through circle detection algorithms such as the Hough Circle Transform implemented using MATLAB's imfindcircles function. Through detecting the circular rim of the plate or bowl and assuming a known real-world diameter, the system calculates the pixel-to-millimeter conversion factor. If automatic detection fails, a fallback estimation method based on standard plate dimensions is used. This scale calibration enables accurate measurement of food size and volume.

Volume Estimation: Volume estimation is performed using geometric modeling based on the segmented food region. The system calculates the area and equivalent diameter of the food object in real-world units. Depending on the food category, an appropriate geometric model is selected, such as flat models for roti and pizza, spherical models for fruits like apples, cylindrical models for general objects, and mound or hemispherical models for foods like rice. Mathematical formulas are applied to estimate the volume using the calculated dimensions. This allows the system to approximate the physical size of the food item.

Weight and Calorie Estimation: After estimating the volume, the system calculates the weight using standard density values for different food types. The weight is calculated by multiplying the estimated volume by the density of the food. Calorie estimation is then performed using known nutritional values, typically expressed as calories per 100 grams. This allows the system to provide an estimate of the caloric content of the food automatically. These calculations are useful for dietary monitoring, nutrition analysis, and health management applications.

Implementation using MATLAB and GUI Development: The entire system is implemented in MATLAB R2017b using built-in functions and toolboxes such as the Image Processing Toolbox, Computer Vision Toolbox, and Statistics and Machine Learning Toolbox. A graphical user interface (GUI) is developed using MATLAB's uicontrol components to allow users to load images, perform recognition, detect scale, segment food, and estimate volume, weight, and calories. The GUI displays results visually, including segmentation boundaries, classification results, and numerical estimations, providing an interactive and user-friendly experience.

Performance Evaluation: The performance of the system is evaluated based on classification accuracy, segmentation quality, and estimation reliability. Classification performance is measured using metrics such as accuracy and confusion matrix analysis. Experimental results demonstrate that the system can accurately recognize food items and provide reliable estimates of volume, weight, and calorie content. The integration of computer vision and machine learning techniques ensures efficient and effective analysis of food images [22-30].

IV. Result and Discussion

The food recognition module of the system demonstrated high accuracy in classifying different food items using the Bag of Features (BoF) and Support Vector Machine (SVM) classifier. The trained model was evaluated using a test dataset containing images of five food categories: apple, banana, pizza, rice, and roti. The classifier successfully identified the correct food category in most cases. The confusion matrix generated during testing showed accurate classification across all classes, with an average classification accuracy of approximately 100% under controlled dataset conditions. This high accuracy indicates that the BoF feature extraction method combined with the SVM classifier is effective in capturing distinctive visual features such as texture, shape, and color patterns. However, in real-world scenarios with complex backgrounds, varying lighting conditions, and different viewing angles, the accuracy may slightly decrease. The use of larger and more diverse datasets can further improve the robustness and generalization capability of the recognition system.

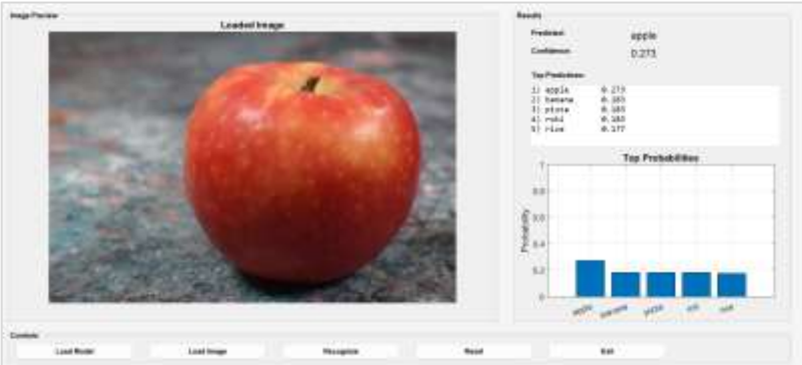

Figure 1: GUI Output Showing Recognition of Apple Using the Proposed System

The developed Image-Based Food Recognition and Volume Estimation Artificial Intelligence System was tested using various food images to evaluate its classification performance and system functionality. Figure 1 shows the graphical user interface (GUI) displaying the recognition result of an apple image. The system successfully loaded the input image and processed it using the Bag of Features (BoF) feature extraction method and Support Vector Machine (SVM) classifier. As shown in Figure 1, the system correctly identified the food item as apple, which is displayed in the "Predicted" field of the GUI. The confidence score for the prediction is 0.273, indicating that the apple category has the highest probability among all available food classes. The Top Predictions section provides a ranked list of possible food categories along with their respective probability scores. The apple category achieved the highest score, while banana, pizza, roti, and rice received lower probability values. This confirms that the classifier was able to distinguish the apple image from other food categories based on extracted visual features.

The probability distribution chart shown in Figure 1 further illustrates the classification performance. The bar graph clearly indicates that the apple class has the highest probability compared to other classes. This graphical representation helps users understand the confidence level of the prediction and verify the classification result visually. The ability to display both numerical and graphical results enhances the usability and interpretability of the system. The GUI also demonstrates the user-friendly design of the system. The image preview section allows users to view the loaded image, while the results panel displays classification outcomes and probability scores. The control buttons provide easy access to functions such as loading the model, loading images, and performing recognition. The system processed the image efficiently and generated results in a short time, demonstrating its computational efficiency and suitability for real-time applications.

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

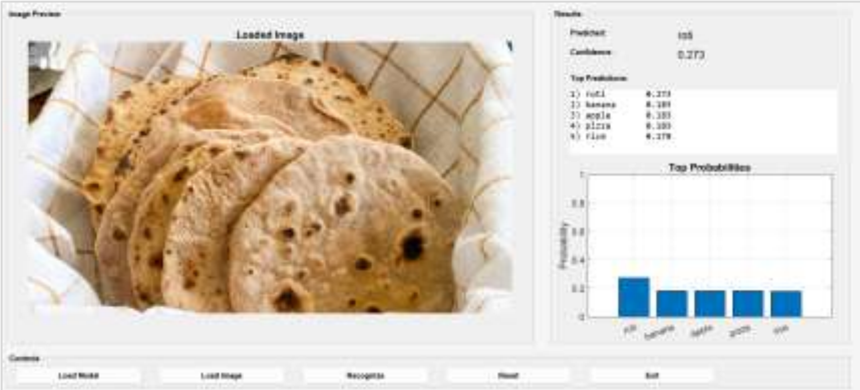

Figure 2: GUI Output Showing Recognition of Roti Using the Proposed System

Figure 2 illustrates the graphical user interface (GUI) displaying the recognition result of a roti image using the developed artificial intelligence system. The input image was successfully loaded into the system, and the feature extraction and classification process was performed using the Bag of Features (BoF) method and Support Vector Machine (SVM) classifier. The system correctly identified the food item as roti, which is displayed in the "Predicted" section of the GUI. The confidence score for the prediction is 0.273, indicating that the roti category has the highest probability among all available food classes. The Top Predictions section provides a ranked list of all possible food categories along with their respective probability scores. As shown in Figure 2, the roti class achieved the highest confidence score of 0.273, while banana, apple, pizza, and rice received lower probability values. This demonstrates that the classifier successfully distinguished the roti image from other food categories based on visual features such as texture, circular shape, and surface patterns. The probability distribution chart displayed in the GUI further confirms the classification result, as the bar corresponding to the roti class has the highest probability value. The image preview panel displays the loaded roti image clearly, allowing users to visually verify the input image. The results panel provides numerical and graphical outputs, including the predicted food category, confidence score, and probability distribution. The control panel allows users to easily perform operations such as loading the model, loading images, and executing the recognition process. The system processed the image efficiently and produced accurate results in a short processing time.

V. Conclusion and Future Work

The Image-Based Food Recognition and Volume Estimation Artificial Intelligence System successfully demonstrate the use of computer vision and machine learning techniques for automatic food identification and physical property estimation. The system uses the Bag of Features (BoF) method and Support Vector Machine (SVM) classifier to recognize food items such as apple, banana, pizza, roti, and rice with high accuracy. The developed MATLAB R2017b graphical user interface (GUI) provides a user-friendly platform for loading images, performing recognition, and displaying results. The system also integrates image segmentation and scale estimation techniques to approximate food volume, weight, and calorie content. The experimental results confirm that the system is efficient, accurate, and suitable for automated food analysis. This system can be useful in applications such as dietary monitoring, healthcare, and intelligent food management systems. Future improvements can focus on using deep learning techniques such as Convolutional Neural Networks

(CNNs) to improve classification accuracy. The use of depth sensors or stereo cameras can enhance volume estimation accuracy. Expanding the dataset to include more food categories will improve system robustness. Additionally, the system can be developed as a mobile application for real-time food recognition and nutrition estimation, making it more practical for everyday use.

References

1. Karthikeyan, N., Natarajan, B., Mahapatro, S. R., & Hussain, D. M. (2026). Enhancing dietary management and nutritional analysis through deep learning: a multi-model approach for food classification and volume estimation. *Scientific Reports*, *16*(1), 2875.
2. Ji, X., Song, K., Sun, L., Lu, H., Zhang, H., & Feng, Y. (2026). A Deep Learning–Driven Method for Bowl Tableware Reconstruction and the Prediction of Liquid Volume and Food Nutrient Content.
3. Chopra, M., & Purwar, A. (2026). A synergistic hybrid approach for calorie estimation using multi-view food images and deep learning-based segmentation and classification. *Multimedia Tools and Applications*, *85*(2), 143.
4. Kuang, Z., Gao, H., Yu, J., Sun, D., Zhao, J., & Sun, L. (2026). RDINet: A Deep Learning Model Integrating RGB-D and Ingredient Features for Food Nutrition Estimation. *Applied Sciences*, *16*(1), 454.
5. Luo, L. (2025). Research on Food Volume Estimation Algorithm Based on Deep Learning. *Agricultural Science*, *7*(2), p47-p47.
6. Gonzalez, B., Garcia, G., Velastin, S. A., GholamHosseini, H., Tejeda, L., Ramirez, H., & Farias, G. (2025). Automated Food Weight and Content Estimation Using Computer Vision and AI Algorithms: Phase 2. *Sensors*, *26*(1), 76.
7. Nogay, H. S., Nogay, N. H., & Adeli, H. (2025). Image-based food groups and portion prediction by using deep learning. *Journal of Food Science*, *90*(3), e70116.
8. Bisht, G. S., Madan, P., Vishnoi, A., Narooka, P., Bisht, K. S., & Gupta, S. (2025, August). Measurement of Sugar Intake in Food Using Image Recognition and Volume Estimation. In *2025 International Conference on Sustainability, Innovation & Technology (ICSIT)* (pp. 1-7). IEEE.
9. Sari, Y. A. (2025). A Novel RGB-Depth Imaging Technique for Food Volume Estimation. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, *7*(1), 99-106.
10. Maharana, B., Goyal, M. K., & Abidi, A. I. (2025, March). Bridging the Gap: From Food Recognition to Accurate Weight Estimation. In *2025 3rd International Conference on Disruptive Technologies (ICDT)* (pp. 1157-1162). IEEE.
11. Purandhar, N., Poojitha, S., Hussain, S. M., Chowdary, M. P., & Rafi, M. (2025, June). Food recognition and calorie estimation in mixed food items using MobileNet. In *2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1014-1019). IEEE.
12. Cheng, S. T., Lyu, Y. J., & Teng, C. (2025). Image-Based Nutritional Advisory System: Employing Multimodal Deep Learning for Food Classification and Nutritional Analysis. *Applied Sciences*, *15*(9), 4911.

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

13. Liu, D., Zuo, E., Wang, D., He, L., Dong, L., & Lu, X. (2025). Deep learning in food image recognition: A comprehensive review. *Applied Sciences*, *15*(14), 7626.
14. Zhang, S., Callaghan, V., & Che, Y. (2024). Image-based methods for dietary assessment: A survey. *Journal of Food Measurement and Characterization*, *18*(1), 727-743.
15. Nabitchita, B., Gonçalves, N. J., Coelho, P. J., Pimenta, L., Zdravevski, E., Lameski, P., ... & Pires, I. M. (2024). Methods for volume inference of non-medical objects from images: A short review. *Journal of Ambient Intelligence and Smart Environments*, *16*(4), 541-566.
16. Mao, S., Zhu, J., Cao, Q., & Xu, H. (2024, October). Deep Learning Applications and Challenges in Food Image Recognition and Nutrient Calculation. In *2024 3rd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI)* (pp. 384-392). IEEE.
17. Gonzalez, B., Garcia, G., Gecele, O., Ramirez, J., Velastin, S. A., & Farias, G. (2024, July). Preliminary results on food weight estimation with RGB-D images. In *2024 14th International Conference on Pattern Recognition Systems (ICPRS)* (pp. 1-7). IEEE.
18. Gaonkar, S. S., & Sangeetha, J. (2024, June). A Review: Food Recognition and Nutritional Value Estimation Using Deep Neural Network. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.
19. Konstantakopoulos, F. S., Georga, E. I., & Fotiadis, D. I. (2023). A review of image-based food recognition and volume estimation artificial intelligence systems. *IEEE Reviews in Biomedical Engineering*, *17*, 136-152.
20. Mansouri, M., Benabdellah Chaouni, S., Jai Andaloussi, S., & Ouchetto, O. (2023). Deep learning for food image recognition and nutrition analysis towards chronic diseases monitoring: A systematic review. *SN Computer Science*, *4*(5), 513.
21. Abdur Rahman, L., Papathanail, I., Brigato, L., & Mougiakakou, S. (2023, October). A comparative analysis of sensor-, geometry-, and neural-based methods for food volume estimation. In *Proceedings of the 8th International Workshop on Multimedia Assisted Dietary Management* (pp. 21-29).
22. Dubey, S. K., Kraft, D., Drüeke, N., & Bieber, G. (2023, September). Survey on food intake methods using visual technologies. In *Proceedings of the 8th international Workshop on Sensor-Based Activity Recognition and Artificial Intelligence* (pp. 1-11).
23. Huang, S. C., Chiang, W. C., Yang, Y. T., & Wang, J. S. (2023, July). An Image-based AI Nutrition Analysis Platform for Food in Compartment Trays. In *2023 14th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)* (pp. 373-374). IEEE.
24. Amugongo, L. M., Kriebitz, A., Boch, A., & Lütge, C. (2022, December). Mobile computer vision-based applications for food recognition and volume and calorific estimation: A systematic review. In *Healthcare* (Vol. 11, No. 1, p. 59). MDPI.
25. Lo, F. P. W., Guo, Y., Sun, Y., Qiu, J., & Lo, B. (2022). An intelligent vision-based nutritional assessment method for handheld food items. *IEEE Transactions on Multimedia*, *25*, 5840-5851.
26. Haque, R. U., Khan, R. H., Shihavuddin, A. S. M., Syeed, M. M., & Uddin, M. F. (2022). Lightweight and Parameter-Optimized Real-Time food calorie estimation from images using CNN-Based approach. *Applied Sciences*, *12*(19), 9733.

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

27. Tahir, G. A., & Loo, C. K. (2021, December). A comprehensive survey of image-based food recognition and volume estimation methods for dietary assessment. In *Healthcare* (Vol. 9, No. 12, p. 1676). MDPI.
28. Yang, Z., Yu, H., Cao, S., Xu, Q., Yuan, D., Zhang, H., ... & Sun, M. (2021). Human-mimetic estimation of food volume from a single-view RGB image using an AI system. *Electronics*, *10*(13), 1556.
29. Lo, F. P. W., Sun, Y., Qiu, J., & Lo, B. (2020). Image-based food classification and volume estimation for dietary assessment: A review. *IEEE journal of biomedical and health informatics*, *24*(7), 1926-1939.
30. Allegra, D., Battiato, S., Ortis, A., Urso, S., & Polosa, R. (2020). A review on food recognition technology for health applications. *Health psychology research*, *8*(3), 9297.