Object Detection and Segmentation in Indian Flat Bread Chapati Using AI Models

Sharada Y. Desai
Department of Electronics & Telecommunication, Faculty of Electronics & Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, India
sharada.desai@vit.edu (corresponding author)

Sharda L. Kore
Department of Electronics & Telecommunication, Faculty of Electronics & Telecommunication Engineering, Bharati Vidyapeeth's College of Engineering for Women, Pune, India
sharda.kore@bharatividyapeeth.edu

Received: 12 May 2025 | Revised: 10 June 2025 | Accepted: 21 June 2025

 DOI: https://doi.org/10.48084/etasr.12088

ABSTRACT

The present research evaluates the implementation of the YOLOv11n model, which stands as one of the best industry-level object detection algorithms, and the application of the SAM2 model, an advanced segmentation algorithm for the digital identification of Indian flatbread, commonly known as chapati. The research demonstrates how the enhanced features and real-time processing of the YOLOv11n system detect and classify chapati images under multiple environmental conditions. The main goal exploits SAM2's advanced ability to divide intricate shapes and textures to separate the chapati pictures from multiple backgrounds and lighting variations. Researchers used high-resolution chapati images under different lighting circumstances with diverse backgrounds for the training and validation process. The identification accuracy of chapatis using YOLOv11n reached notable levels, which established both its effectiveness and speedy operation. The SAM2 performance evaluation method incorporates metrics related to segmentation precision combined with system execution speed and its ability to handle different conditions. This study demonstrates how YOLOv11 object identification and SAM2 segmentation are combined to simplify the industrial food processing operations and automate the food production systems within the food industry framework.

Keywords- Indian flat bread; chapati; YOLO11n; SAM2; object detection; object segmentation

I. INTRODUCTION

Digital images and videos are essential for the food industry development, as they help perform quality checks, safety assessments, and analyze the nutritional values. Indian traditional flatbread known as Chapati maintains a significant cultural weight and culinary position. The chapati preparation process includes specific cooking techniques that shape both its texture and puffing properties, as well as its appearance, which are essential for determining its quality and taste. Computer vision represents a promising method for analyzing food quality, as the interest in automated quality control continues to rise [1].

Being advanced computer vision models, YOLOv11n and SAM2 strengthen these procedures through a precision-oriented delivery of efficient operation and scalable application [2, 3]. Their essential features are.

Quality Control: Food quality measurements based on texture and color as well as size and shape can be evaluated instantly through digital image analytics. Given its ability to detect objects, YOLOv11n effectively identifies various kinds of defects on chapatis including irregular cooking, physical abnormalities, and burn scarring. The food quality assessment capabilities of SAM2, such as its ability to process individual food components through its segmentation model, outperform previous models. SAM2 divides chapati surfaces to examine problematic areas, which help maintain strict product standards. Rapid detection of contaminants, foreign objects, and food anomalies remains a top priority in food safety, requiring advanced digital analysis. Through its quick identification capabilities, YOLOv11n finds unsafe production line irregularities while omitting the need for manual checks. The food inspection system gets enhanced detail through SAM2's segmented analysis, which can investigate specific areas of food image or video contaminants. The joint operation between these systems reduces the health risks as it enhances the compliance with food safety requirements. Nutritional analysis is made possible through the accurate identification and segmentation of food items, followed by the measurement of their components. SAM2 sends detailed portion segments to YOLOv11n for understanding the food item classification, which enables calorie estimation, macro/micronutrient ratio calculation, or ingredient quantity assessment. The segmentation capabilities of SAM2 enabled studying nutritional components and making recipe improvements by breaking down the ingredients presented in cooking videos of chapatis.

A. Why YOLOv11n and SAM2?

The food technology space finds these models advantageous because they offer distinctive advantages.

- YOLOv11n provides exceptional speed and real-time detection, making it suitable for continuous monitoring in production lines.
- SAM2 offers unparalleled segmentation capabilities, allowing for a detailed and granular analysis essential in research and development [4].

Together, these models enable smarter, automated workflows that streamline operations, enhance safety, and enrich nutritional data. Their application in food analysis shows how Artificial Intelligence (AI)-driven solutions are revolutionizing the food industry. Real-time object detection and segmentation operate with the following restrictions in the current methodologies:

Traditional methods of digital image object detection and segmentation depend on pre-constructed features as well as traditional machine learning algorithms, which normally face difficulties with complex visual data and lack the ability to capture high-level semantics. Convolutional Neural Networks (CNNs) represent a crucial development in deep learning that enables the automatic extraction of complex hierarchical features from data, which fosters better precision and resilience. Deep learning techniques have revolutionized object detection and medical imaging, while a successful shift from classical to deep learning approaches has been observed. Due to its real-time and fast detection capabilities YOLOv11n performs optimally for sustaining an ongoing observation in production lines. The segmentation functions of SAM2 enable performing a detailed analysis, which is significant during the research and development activities. These models create collaborative systems, which offer better workflow automation that benefits production improvement, safety enhancement, and data enrichment. Food analysis proves how AI solutions transform the entire food industry sector.

However, the detection methods available for digital image objects and the segmentation through traditional algorithms struggle with the complex data input, as they require hand-built features with traditional machine learning algorithms. Deep learning strategies, including CNNs, have enhanced the object recognition capabilities by extracting multidimensional features directly from input data, which boosts the system accuracy and performance robustness. The conversion from classical to deep learning approaches allows for the efficient detection of objects across multiple domains and medical image analysis.

Traditional methods rely on manually designed features and show limited capability in detecting advanced image patterns and semantic details [4, 5]. These approaches demonstrate limited accuracy and sturdy performance when dealing with variations in object orientation, size, and occlusions [6]. The commonly used thresholding techniques, along with K-means segmentation, achieve lower performance metrics than deep learning methods, according to findings of [7].

Deep learning algorithms, including U-Net, SegNet, and FCN-8, deliver superior results in segmentation tasks because they provide better accuracy, precision, and recall statistics than regular techniques [6, 7]. These methods are effectively applied to medical imaging, autonomous vehicles, surveillance systems, and numerous other disciplines [5, 8]. However, they require substantial data resources and powerful computational capabilities. Scientific efforts continue to investigate matters related to fairness and transparency combined with the accountability problems in deep learning models [5].

B. Research Objectives

The present study establishes an advanced computer vision system composed of YOLOv11n and SAM2 models to investigate the cooking quality and appearance of Indian flatbread chapati. Its objectives are:

- To employ the YOLOv11n system for real-time quality assessment of chapatis through its capability to detect puffing features, surface characteristics, and color consistency for consistent product standards.
- To implement the SAM2 model for picture segmentation that focuses on finding stray spots, poorly cooked areas, and texture that appears inconsistent.
- To conduct a taste correlation analysis to investigate visual attributes, such as color and texture combinations, for predicting sensory results using picture-based information.
- To enable the system to monitor chapati images within operating environments allowing for immediate quality feedback.
- To develop an effective model containing a broad dataset with chapati pictures that show changes in the lighting quality, cooking states, and background variations for model learning and testing.
- To investigate the feasibility of the permanent implementation of this integrated framework in large scale food production plants to recreate manufacturing facilities with both automated systems and cut down manual quality examination tasks.

The current research combines traditional sensory assessment strategies with state-of-the-art AI technologies for the development of quality evaluation.

II. OVERVIEW OF EXISTING APPROACHES TO FOOD IMAGE ANALYSIS

Food image analysis encompasses various approaches aimed at recognizing and assessing food items through image processing. The evolution of this domain occurred since deep learning and self-supervised learning have emerged. The key approaches currently employed in food image analysis are:

A. Self-Supervised Learning

This approach utilizes vast quantity of unlabeled food images to learn visual representations. Six self-supervised models were evaluated on the Food-101 data set, highlighting their strengths and weaknesses in food image analysis. Future work aims to enhance performance through improved model designs [9].

B. Deep Learning Techniques

Food recognition has evolved through deep learning, matching or even surpassing traditional methods in accuracy [10]. The three main approaches consist of scratch-built model design, transfer learning, and platform-based solutions, each with its own benefits and drawbacks [11]. The widespread adoption of CNNs makes it possible to discard hand-crafted features.

C. Vision-Based Dietary Assessment

First-order dietary assessment depends on three main components: food image analysis, volume estimation, and nutrient derivation. End-to-end deep learning algorithms show promise for directly detecting the nutrition values in food images [12].

TABLE I. SUMMARY OF EXISTING APPROACHES TO FOOD IMAGE ANALYSIS

Sr. No	Study	Findings
1	[9]	Existing approaches to food image analysis primarily rely on deep learning methods that require human annotations. However, these methods are impractical for real-world applications due to the abundance of unlabeled images, prompting a focus on self-supervised learning techniques.
2	[13]	Existing approaches to food image analysis include use of imaging methods, like TEM and x-ray microscopy, combined with software tools, such as Dragonfly, for segmentation and visualization, enabling the characterization of microstructure properties related to food formulation and optimization,
3	[14]	This study evaluates Digital Image Analysis (DIA) techniques by discussing image acquisition and analyzing features while reviewing the processing methods. Digital processing techniques allow fast assessment and non-harmful inspection that detect toxins from pesticides and perform quality checks and moisture measurement in food throughout the supply chain.
4	[15]	Various methods used in food image analysis incorporate the 2D and 3D imaging systems, stereological procedures for shape measurements coupled with linear discriminant analysis for classification, the enhancement of details, and measurement of structural elements, including volume fraction and object size.
5	[16]	Existing food image analysis methods include time-consuming manual classification and inconsistencies, whereas contemporary computer vision techniques, such as CNN and YOLOv4, perform rapid quality checks and segmentation based on the dimension, shape, and color characteristics.
6	[17]	In this study, MobileNetV2 and EfficientNetV2 were used to extract features, with a focus on essential leaf attributes, such as color, texture, and shape. With the help of the combined method of feature extraction, the model proved to be more efficient.
7	[18]	The findings of this study demonstrate that YOLOv5 has a higher accuracy and shorter inference time compared to other models. This evaluation is importance as it establishes YOLOv5 as the best model for real-time cashew categorization.
8	[19]	According to this study, the existing methods for food image analysis consist of manual model development, transfer learning implementation, and platform-based solution deployment. The strength of each method differs. Platforms achieve the highest deployment speed although their efficiency performance remains unknown.
9	[20]	Traditional approaches to food image analysis include classical features with neural networks, SVMs, HMMs, and contemporary features based on CNNs with deep learning. The drawbacks of these approaches include the requirement for large dataset collections to improve the accuracy levels.
10	[21]	According to this study, existing approaches to food image analysis include preprocessing techniques, like scaling, noise reduction, feature extraction using CNNs, and segmentation methods, such as color-based, texture-based, and graph-based segmentation, all aimed at enhancing ingredient recognition and classification accuracy.
11	[22]	This study provides an overview of the current food image analysis methods, highlighting the development of image-based tools for accurate dietary intake measurement. It emphasizes the need for further research to address the existing challenges in this field.

III. METHODOLOGY

Figure 1 shows the installation of a computer vision setup for the chapati cooking research. The computer vision setup used in this study includes the following hardware components: a Raspberry Pi 5 (8GB) serving as a high-speed storage processor, a Sony IMX708 NoIR 12 MP Auto Focus camera, an MLX90640 IR Array Thermal Imaging Camera with a resolution of 32×24 pixels, a 1TB PCIe NVMe SSD disk for data storage, and an ESP32S3-based LED light intensity controller to ensure uniform lighting conditions across the dataset. The software stack comprises Python 3.9, libcamera2 Python library for image capture, SAM for image segmentation, the YOLOv11n model for object detection, and matplotlib to convert temperature arrays into raw image captures. The following steps are performed to implement the object detection and segmentation in chapati using AI models:

A. Data Acquisition

Following are the steps performed during the data collection phase of the study.

- Capturing of high-resolution images and videos of chapatis at various cooking stages.
- Collection data under diverse conditions, such as different lighting setups and texture variations.
- Inclusion of images that showcase common defects, like burnt areas, uneven puffing, and surface irregularities.
- Organizing the dataset into training, validation, and testing subsets.

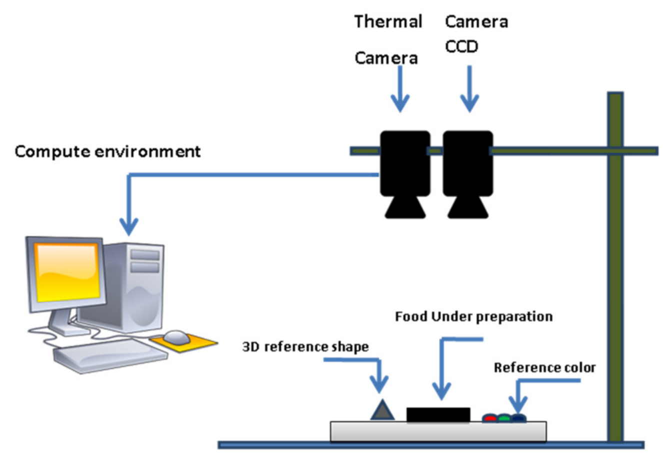


Fig. 1. Computer Vision setup of chapati cooking.

B. Creating Chapati Data

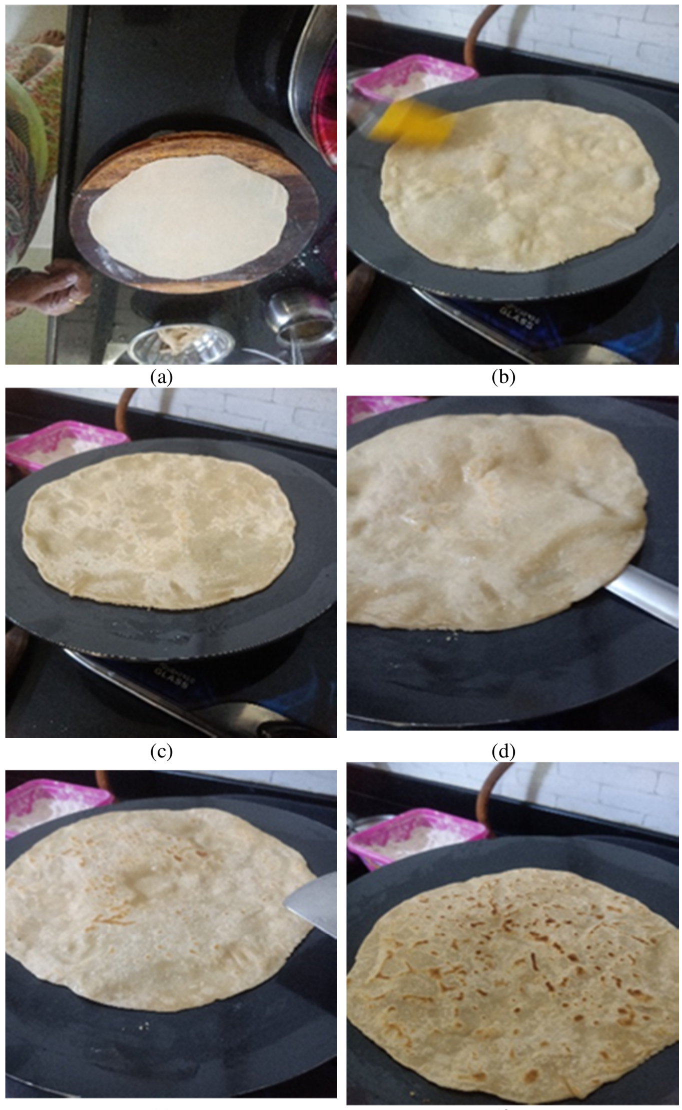
A custom dataset was made for identifying and segmenting chapatis in computer vision. All appropriate ethical guidelines for the data collection were followed, and the dataset was constructed under suitable and controlled settings. All images in the dataset were taken with standard light and the same color settings for a higher model precision. No identifiable information or invasive methods were used while collecting the images. Also, no publicly available datasets were utilized.

Lokwan whole wheat flour grinded from the domestic grinding meal was used to make the dough of the chapati. The dough was rested for approximately half an hour before being used to make chapatis. The experiments were conducted utilizing chapati images captured through the computer vision setup shown in Figure 1. Images from the initial cooking stage to the final fully cooked chapati were captured and stored. One epoch of complete chapati making may consist of a minimum of five stages, as depicted in Figures 2(a)-(f). The study used approximately 2,500 chapati images (500 chapatis × 5 stages each) to train the model. All chapati images are captured at a home kitchen on different days. These self-made data aim to help with the academic research on using computer vision to spot and separate the chapatis in an image.

C. Dataset Preprocessing

After images were captured, data processing was performed, involving data augmentation and annotation. Augmentation techniques, such as rotation, scaling, noise addition, and brightness adjustments, were applied to simulate real-world variability. Regarding annotation, images were manually labeled with bounding boxes to be used with the YOLOv11n model. Segmentation masks were also used in the SAM2 model to define the regions of interestYOLOv11n Model Setup.

The YOLOv11n model was configured for real-time detection tasks. The model was trained to detect chapati-specific features, such as puffing, texture uniformity, and cooking irregularities. Hyperparameter tuning tests were conducted for the learning rate and batch size to optimize both the detection speed and accuracy levels.



(a) (b) (c) (d) (e) (f)

Fig. 2. Stages of chapati preparation: (a) raw chapatti, (b) cooking stage 1, (c) flip side 1, (d) flip side 2, (e) flip side 3, (f) fully cooked chapati.

D. SAM2 Model Integration

SAM2 was deployed for the segmentation tasks by being trained to segment the chapati images and isolate critical regions, such as burnt spots or unevenly cooked surfaces. The model was fine-tuned for precision segmentation to ensure adaptability across diverse chapati shapes and textures.

E. Training and Validation

The trained YOLO11vn and SAM2 models receive data from the prepared dataset through the training and validation steps. The performance of the YOLOv11n and SAM2 models must be assessed using key metrics, namely accuracy, recall, and Intersection over Union (IoU).

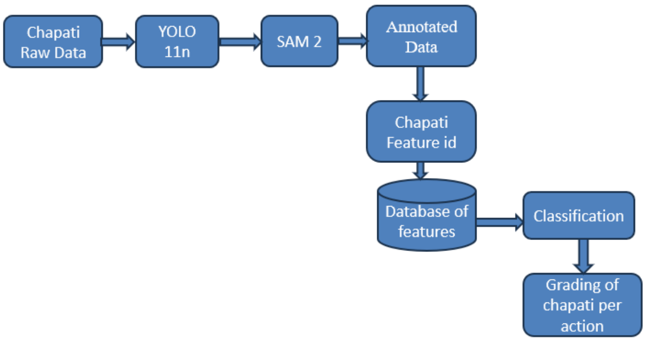


Fig. 3. AI-based model architecture for chapati classification.

F. *Multi-Model Workflow Integration*

YOLOv11n and SAM2 were integrated into a unified pipeline by implementing sequential detection (via YOLOv11n), followed by segmentation (via SAM2) for detailed analysis. The integrated system was tested on real-time video inputs to assess the dynamic performance.

G. *Quality and Taste Correlation*

The segmented regions and detected features were analyzed to quantify the visual indicators of taste and quality. Puffing patterns and uniform texture were correlated with sensory parameters, like softness and elasticity, and finally, predictions were evaluated against human taste for validation.

IV. RESULTS AND DISCUSSION

The proposed model runs its training and testing processes on Linux 64-bit operating system with Intel Xeon, 32GB memory system. Figure 4 shows the YOLOv11n and SAM2 integrated output images of various stages of chapati cooking, where Yolo11n is used to identify the chapati object and SAM2 to conduct the chapati segmentation. The results of YOLOv11N integrated with the SAM2 models are listed in Table I.

TABLE I. RESULTS OF YOLO11N AND SAM2 MODELS

Class Label	Precision	Recall	F1-Score
Chapati	0.975	1	0. .924

The performance assessment of YOLOv11n and SAM2 model output relies on three evaluation metrics including F1 score, precision, and recall.

A high F1 score across all values signifies a consistent performance, which is considered ideal. The traditional F-measure, or balanced F1 score, is calculated as the harmonic mean of precision and recall:

$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2\,\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{2\text{TP}}{2\text{TP+FP+FN}} \qquad (1)$$

Precision is calculated as $TP/(TP+FP)$ and recall as $TP/(TP+FN)$, where TP is True Positive, FN is Flase Negative, and FP is the False Positive. The F1 score balances these two measures to evaluate the model performance.

Fig. 4. Yolo11n and SAM2 integrated output images of various stages of the chapati making process.

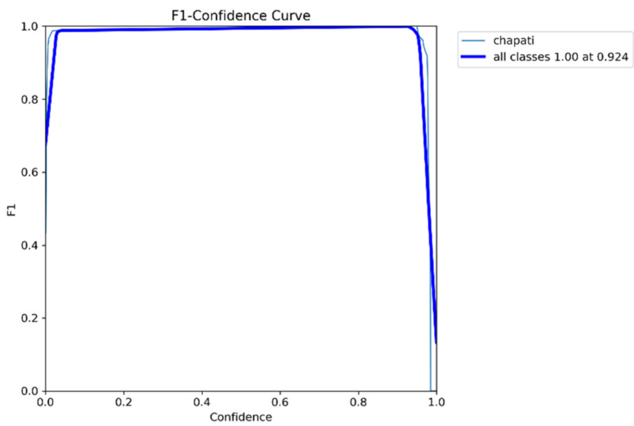


Fig. 5. F1- score curve for chapati.

Proper selection, scaling, and extraction of features directly affect model performance and can result in either improved or reduced precision and recall.

The Precision-Recall (P-R) curve serves as a primary assessment tool for model classification, especially in the case of imbalanced data. A model which achieves high values of precision and recall across all thresholds generates a P-R curve that strongly follows the top-right boundary because of its excellent performance abilities. The P-R curve serves as an evaluation tool during the food image analysis to assess the model performance when identifying specific products and flaws (for instance, quality analysis of chapatis) without producing excessive FP or FN result. The following metrics are employed to monitor and evaluate the model's training and validation performance:

- Box Loss (train/box_loss and val/box_loss): The model uses this metric to quantify the errors in predicting the box coordinates for the detected objects. A downward pattern in error, as displayed in Figure 8, indicates that the model becomes better at localizing objects.
- Classification Loss (train/cls_loss and val/cls_loss): This metric measures the errors in classifying objects into the

correct categories. A steady decrease, as illustrated in Figure 8, implies improved prediction accuracy for the object classes.

- Distribution Focal Loss (train/dfl_loss and val/dfl_loss): This metric focuses on refining the localization of the bounding boxes. Its reduction, as portrayed in Figure 8, implies enhanced precision in determining the object boundaries.
- Precision (metrics/precision): The precision metric shows the proportion of the correct positive predictions. A high precision value, as observed in Figure 9, suggests that the model makes fewer FP predictions.
- Recall (metrics/recall): Recall represents the proportion of actual positives correctly identified. A high recall means fewer FN, as seen in Figure 9 .

Mean Average Precision (mAP) (metrics/mAP50 and metrics/mAP50-95): This metric evaluates the overall model performance in detecting objects across different thresholds of IoU. Higher values, as evidenced in Figure 9, indicate a better detection accuracy.

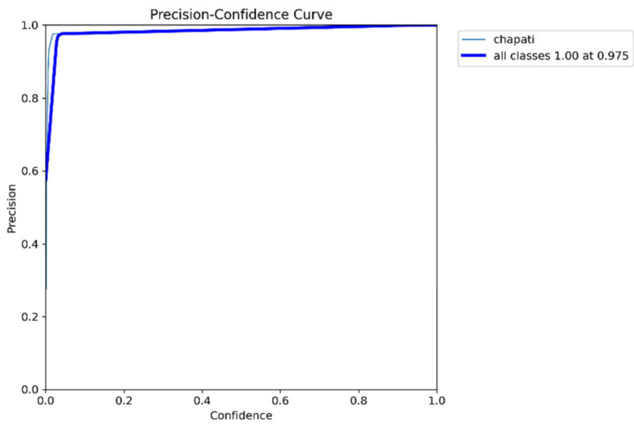


Fig. 6. Precision curve for chapati.

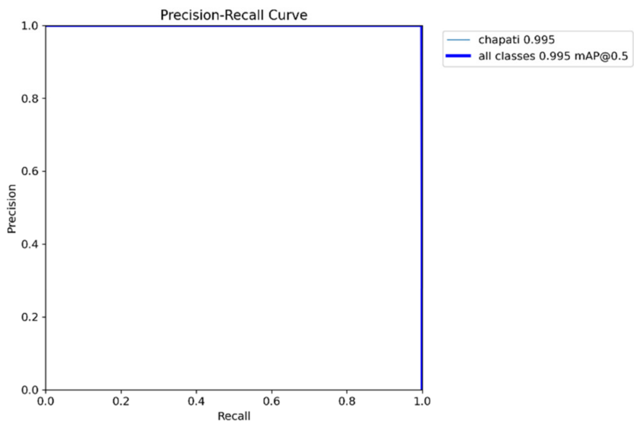


Fig. 7. P-R curve for chapati.

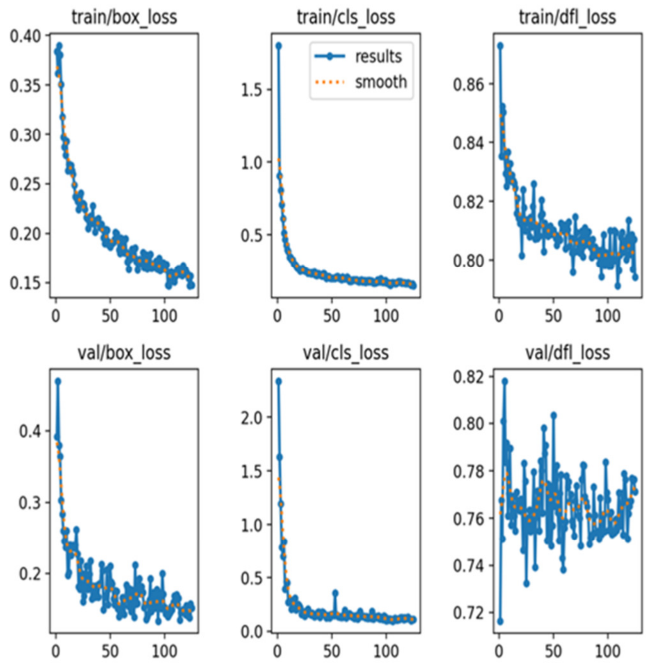


Fig. 8. Training curve of YOLO11n model.

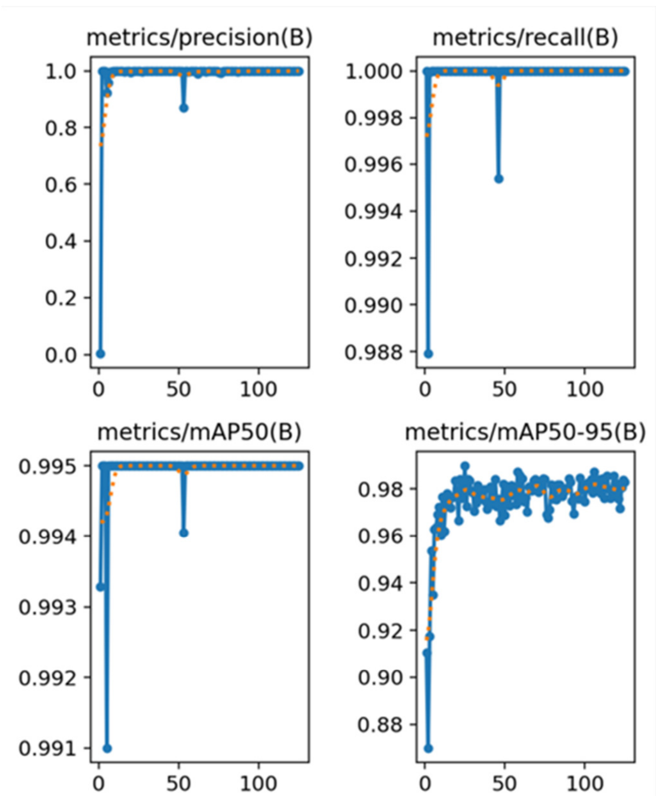


Fig. 9. Training curve of YOLO11n.

V. CONCLUSION

The present study demonstrates the research process undertaken to identify the chapati objects and segment the chapati images using computer vision techniques. The experimental results confirm that using YOLOv11n for instant food detection and SAM2 for precise image segmentation significantly enhances the chapati identification. One important finding is that both YOLOv11n and SAM2 perform optimally when identical lighting and accurate color calibration is used while taking chapati pictures. However, thermal readings are not required for detecting and segmenting the pan and chapati, since the chapati is always at the same temperature during cooking. For the sake of performance, expressing temperature data in a numerical form proves to be effective. The findings of this study can be employed for:

- The utilization of automated robots and software to check the quality of chapatis in large-scale manufacturing environments.
- Studying regional cultural variations in chapatis.
- Employing self-supervised models to boost the accuracy in recognizing food objects with minimum labeled data.

With the help of this study's findings, the chapati analysis can be performed with Artificial Intelligence, which supports both the automation in food factories and the increase in real-time image analysis capabilities. Future studies can combine vision with deep learning to create Artificial Intelligence tools that quantify both taste and quality.

ACKNOWLEDGMENT

The authors would like to thank Mr. Yogesh Desai for the technical guidance and moral support throughout the research work.

DATA AVAILABILITY

The dataset used in this study can be made available upon request for research and validation purposes. Interested individuals may contact the authors, subject to the signing of a data-sharing agreement.

REFERENCES

[1] S. Y. Desai and S. L. Kore, "The Current State of Art-Indian Unleavened Flat Bread Cooking," in *IOT with Smart Systems*, vol. 720, J. Choudrie, P. N. Mahalle, T. Perumal, and A. Joshi, Eds., Singapore: Springer Nature Singapore, 2023, pp. 1–9.

[2] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 12, Dec. 2024, Art. no. 336, https://doi.org/10.3390/computers13120336.

[3] X. Lan *et al.*, "FoodSAM: Any Food Segmentation," *IEEE Transactions on Multimedia*, vol. 27, pp. 2795–2808, 2025, https://doi.org/10.1109/TMM.2023.3330047.

[4] F. Neha, D. Bhati, D. K. Shukla, and M. Amiruzzaman, "From classical techniques to convolution-based models: A review of object detection algorithms," *arXiv*, 2024, https://doi.org/10.48550/ARXIV.2412.05252.

[5] S. Tasnim and W. Qi, "Progress in Object Detection: An In-Depth Analysis of Methods and Use Cases," *European Journal of Electrical Engineering and Computer Science*, vol. 7, no. 4, pp. 39–45, Jul. 2023, https://doi.org/10.24018/ejece.2023.7.4.537.

[6] S. M. Khaniabadi, H. Ibrahim, I. A. Huqqani, F. M. Khaniabadi, H. A. M. Sakim, and S. S. Teoh, "Comparative Review on Traditional and Deep Learning Methods for Medical Image Segmentation," in *2023 IEEE 14th Control and System Graduate Research Colloquium (ICSGRC)*, Shah Alam, Malaysia, Aug. 2023, pp. 45–50, https://doi.org/10.1109/ICSGRC57744.2023.10215402.

[7] A. Plaksyvyi, M. Skublewska-Paszkowska, and P. Powroznik, "A Comparative Analysis of Image Segmentation Using Classical and Deep Learning Approach," *Advances in Science and Technology Research Journal*, vol. 17, no. 6, pp. 127–139, Dec. 2023, https://doi.org/10.12913/22998624/172771.

[8] Z. Liu *et al.*, "A novel acoustic emission signal segmentation network for bearing fault fingerprint feature extraction under varying speed conditions," *Engineering Applications of Artificial Intelligence*, vol. 126, Nov. 2023, Art. no. 106819, https://doi.org/10.1016/j.engappai.2023.106819.

[9] A. W. Peng, J. He, and F. Zhu, "Self-supervised visual representation learning on food images," *Electronic Imaging*, vol. 35, no. 7, pp. 269-1-269–6, Jan. 2023, https://doi.org/10.2352/EI.2023.35.7.IMAGE-269.

[10] C. Kiourt, G. Pavlidis, and S. Markantonatou, "Deep learning approaches in food recognition," in *Machine Learning Paradigms: Advances in Deep Learning-based Technological Applications*, G. A. Tsihrintzis and L. C. Jain, Eds., Cham, Switzerland, 2020, pp. 83–108.

[11] S. Arora, G. Chaware, D. Chinchankar, E. Dixit, and S. Jain, "Survey of Different Approaches Used for Food Recognition," in *Information and Communication Technology for Competitive Strategies*, vol. 40, S. Fong, S. Akashe, and P. N. Mahalle, Eds. Singapore: Springer Singapore, 2019, pp. 551–560.

[12] W. Wang, W. Min, T. Li, X. Dong, H. Li, and S. Jiang, "A review on vision-based analysis for automatic dietary assessment," *Trends in Food Science & Technology*, vol. 122, pp. 223–237, Apr. 2022, https://doi.org/10.1016/j.tifs.2022.02.017.

[13] N. Piche, E. Jabason, M. Marsh, and A. H. Vollmer, "Survey of Image Analysis Methods Applied to Consumer Foods," *Microscopy and Microanalysis*, vol. 24, no. S1, pp. 1208–1209, Aug. 2018, https://doi.org/10.1017/S1431927618006529.

[14] Y. E. Nagaty, "Digital Image Analysis (DIA) in Food Technology," *Alexandria Journal of Food Science and Technology*, vol. 22, no. 2, pp. 29–34, Jan. 2025, https://doi.org/10.21608/ajfs.2025.340104.1062.

[15] J. C. Russ, "Image Analysis of Foods," *Journal of Food Science*, vol. 80, no. 9, Sep. 2015, https://doi.org/10.1111/1750-3841.12987.

[16] R. Karpe, P. Patil, P. Shahane, H. Patki, S. Vispute, and R. Kannan, "A Review of Fruits Image Analysis Using Computer Vision and Deep Learning Techniques," in *Soft Computing for Security Applications*, vol. 1449, G. Ranganathan, Y. El Allioui, and S. Piramuthu, Eds., Singapore: Springer Nature Singapore, 2023, pp. 707–723.

[17] S. Khandelwal, A. Raut, H. Vyawahare, D. Theng, and S. Dhande, "Optimizing Performance in Mango Plant Leaf Disease Classification through Advanced Machine Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18476–18480, Dec. 2024, https://doi.org/10.48084/etasr.8220.

[18] S. N. Karnam, V. S. Vaddagallaiah, P. K. Rangnaik, A. Kumar, C. Kumar, and B. M. Vishwanath, "Precised Cashew Classification Using Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 17414–17421, Oct. 2024, https://doi.org/10.48084/etasr.8052.

[19] Y. Zhang *et al.*, "Deep learning in food category recognition," *Information Fusion*, vol. 98, Oct. 2023, Art. no. 101859, https://doi.org/10.1016/j.inffus.2023.101859.

[20] Q. L. Tran, G. H. Lam, Q. N. Le, T. H. Tran, and T. H. Do, "A Comparison of Several Approaches for Image Recognition used in Food Recommendation System," in *2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, Purwokerto, Indonesia, Jul. 2021, pp. 284–289, https://doi.org/10.1109/COMNETSAT53002.2021.9530793.

[21] M. Marinis, E. Georgakoudis, E. Vrochidou, and G. A. Papakostas, "Visual Recognition of Food Ingredients: A Systematic Review," in *Artificial Intelligence Annual Volume 2024*, G. A. Papakostas, M.

Antonio Aceves-Fernández, and M. Emin Aydin, Eds., IntechOpen, 2023.

[22] Y. Wang *et al.*, "Food image analysis: The big data problem you can eat!," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2016, pp. 1263–1267, https://doi.org/10.1109/ACSSC.2016.7869576.

AUTHORS PROFILE

Sharada Y. Desai, Orcid ID- https://orcid.org/0009-0009-4400-9502, Scopus ID: 58623925700, Web of Science (WoS) Researcher- MVW-8817-2025, received the M.S. degree in Electronics Engineering from the Bharati Vidyapeeth University, Pune, India. Sharada Desai. Since 2008, she serves as an assistant professor at E&TC Department of the Vishwakarma Institute of Technology, Pune, India. Machine learning, artificial intelligence, and IoT constitute the main research areas she focuses on. As part of her Ph.D., she devotes her studies to artificial intelligence and machine learning at Savitribai Phule Pune University while conducting research at the VIIT research center.

Dr. Sharda L. Kore, https://orcid.org/0000-0003-1757-0424, Scopus ID-0000-0003-1757-0424, Web of Science (WoS) Researcher ID- ADU-7649-2022: Dr. Sharada Kore received her PhD degree in E & TC Engineering. She holds an Associate Professor position at the Department of E & TC Engineering within Bharati Vidyapeeth's College of Engineering for Women, Pune, India. Her academic track record includes 32 publications, three book chapters, and two patents.