CSC8628 Report - Food Identification and Segmentation

Vaishnavi Jeurkar

Abstract

The primary objective of this research is to implement an efficient and accurate approach to food recognition, an integral part of an automated food analysis system. This work presents an image segmentation pipeline tailored for food item extraction from images. Leveraging computer vision techniques and algorithms, the system employs a multi-step approach integrating the GrabCut and Mean Shift algorithms for precise delineation of food regions. The workflow involves preprocessing steps to enhance image quality and adaptively segment regions of interest. Furthermore, post-processing techniques refine the segmentation output. Evaluation metrics, particularly the mean Intersection over Union (mIoU), are utilized to quantitatively assess the segmentation accuracy against manually annotated ground truth. Visual comparisons between the segmented output and ground truth images facilitate qualitative evaluation. The system's efficiency and accuracy lay a foundation for various applications in automated food analysis, dietary assessment, and nutritional analysis.

Introduction

Identifying and segmenting food items on a plate poses a significant challenge in computer vision and image processing. Automating the recognition of different food types within an image is crucial for numerous applications, including dietary assessment, food analysis, and nutritional tracking. The complexity arises from the diverse appearance, shapes, textures, and occlusions of various food items on plates, making their accurate isolation and identification a non-trivial task.

Traditional methods rely on classical segmentation algorithms to address this problem. In recent research [1], a novel approach was explored, integrating thermal and RGB imaging to segment food items of similar appearance but different temperatures. K-means clustering was attempted, but its reliance on a predetermined number of clusters poses limitations as the quantity of food items varied. Similarly, another study [2] combined k-means clustering and watershed segmentation for image segmentation and edge detection. However, the effectiveness of these techniques in accurately identifying and delineating individual food items varies due to the inherent variability in food appearances, lighting conditions, and occlusions. Various automatic segmentation methods have been devised for fruit image analysis to facilitate fruit defect detection. Among these, a novel hybrid algorithm based on a split and merge approach is proposed in [3]. It implements k-means clustering along with graph-based algorithms. In their work, Anthimopoulos et al. [4] devised a methodology focused on segmenting and recognizing multiple food items in images, particularly aimed at carbohydrate counting purposes. Their approach involved several key steps: initially, pyramidal mean shift filtering coupled with a region growing method was employed to effectively segment the plate from the various food items in the image.

The objective of this research is to develop an intelligent system that effectively addresses these challenges by employing classical segmentation algorithms to segment and isolate food items on a plate. The dataset utilized in this assignment corresponds to the work by X. Wu et al., titled 'A Large-Scale Benchmark for Food Image Segmentation,'[5] which consists of 2135 food images. By leveraging techniques such as GrabCut and Mean Shift, the system aims to accurately identify distinct food entities within an image, overcoming issues related to varying appearances, overlapping items, complex backgrounds, and the need to predefine the number of clusters.

The ultimate goal is to create a robust and accurate food recognition system capable of segmenting different food items accurately, enabling applications in automated dietary assessment, food portion estimation, and nutritional analysis. Addressing these challenges is critical for advancing computer vision systems' capabilities in understanding and processing food-related images accurately.

Methodology

1. Algorithm.

The python code implements a multi-step image segmentation and evaluation process applied to food-related images. The code utilizes libraries like OpenCV and NumPy for image processing and matplotlib for visualization. The aim is to segment food items from images and evaluate the segmentation quality using the mean Intersection over Union (mIoU) metric.

The workflow begins with preprocessing the input image, by applying gamma correction for light adjustment and normalization. It then employs the GrabCut algorithm to isolate food items by iteratively refining a segmented mask. Subsequently, the Mean Shift algorithm is used to further segment the regions obtained from GrabCut, enhancing the segmentation results.

The post-processing step involves morphological operations to refine the segmentation. Finally, the algorithm calculates the mIoU metric, comparing the processed segmentation with the ground truth images, where food items are manually annotated.

To assess the segmentation quality, the code iterates over specific images, performs the segmentation pipeline, calculates the mIoU score for each image by comparing the segmented output with its corresponding ground truth, and displays the segmented output alongside the ground truth for visual comparison.

This implementation is designed for food segmentation in images, offering a systematic pipeline for evaluation and comparison of segmentation results against manually annotated ground truth, essential for assessing segmentation accuracy in computer vision tasks. The code's structure and methodology provide a comprehensive framework for image segmentation evaluation in food-related domains, essential for research in automated food recognition and analysis.

2. Libraries/Functions.

A. cv2 (OpenCV): OpenCV is a popular computer vision library used for image and video processing. Functions from OpenCV are employed for image manipulation, segmentation algorithms (GrabCut, Mean Shift), morphology operations, and reading/writing images.

B. numpy: NumPy is a fundamental package for numerical computing in Python. It provides support for arrays and matrices, which are used extensively for mathematical operations and array manipulation in this code.

C. matplotlib.pyplot: Matplotlib is a plotting library in Python. The pyplot module from Matplotlib is used for generating visualizations such as displaying images and creating subplots.

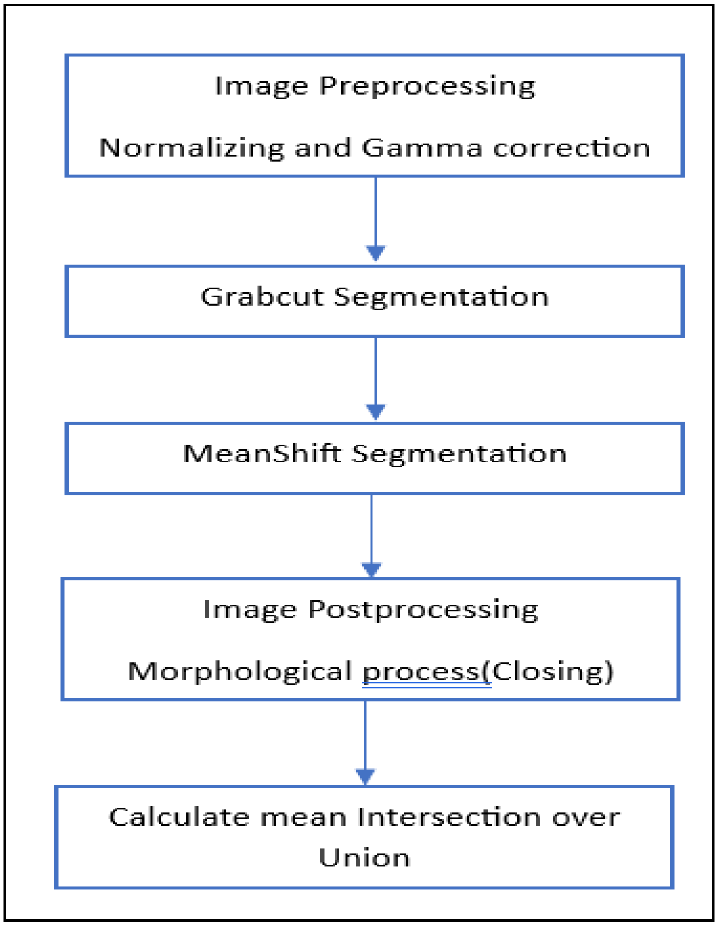
D. cv2.grabCut(img, mask, rect, bgd_model, fgd_model, iterCount, mode=cv2.GC_EVAL): Performs the GrabCut algorithm for image segmentation. It iteratively refines a segmented mask based on the provided rectangle.

E. cv2.pyrMeanShiftFiltering(image, sp, sr): Applies the Mean Shift filter for image segmentation. It performs clustering to group similar pixels.

F. cv2.cvtColor(ms_clustering, cv2.COLOR_RGB2GRAY): Converts the color image to grayscale, often used for further image processing.

G. cv2.morphologyEx(segmented_img, cv2.MORPH_CLOSE, kernel): Applies morphological operations to the image. This code uses a closing operation with a specified kernel to refine the segmentation.

3. Flowchart.



Results and Discussion

The algorithm underwent extensive testing on a dataset comprising 2135 food images, showcasing its performance in segmenting various food items. An average Intersection over Union (IoU) score of 0.57 was achieved, indicating a moderate level of accuracy in delineating food items from the background. This score, ranging between 0 and 1, quantifies the overlap between segmented regions and ground truth annotations, serving as a key performance metric for segmentation algorithms.

However, the algorithm encountered challenges when dealing with specific scenarios. Images containing multiple plates or combinations of a plate and a cup posed difficulties in accurate segmentation. The complexity increased when attempting to segment such multifaceted arrangements, affecting the algorithm's precision in these instances.

Despite these challenges, the algorithm exhibited noteworthy strengths. It demonstrated a robust capability in effectively segmenting contents from single plates, displaying consistent performance regardless of variations in plate color, food texture, or shapes of the items. This resilience suggests the algorithm's adaptability and reliability in scenarios where a single plate is the dominant subject.

Notably, variations in lighting conditions does not affect the algorithm. Images with diverse lighting setups exhibited no decrease in segmentation accuracy, highlighting the algorithm's insensitivity to lighting variations.

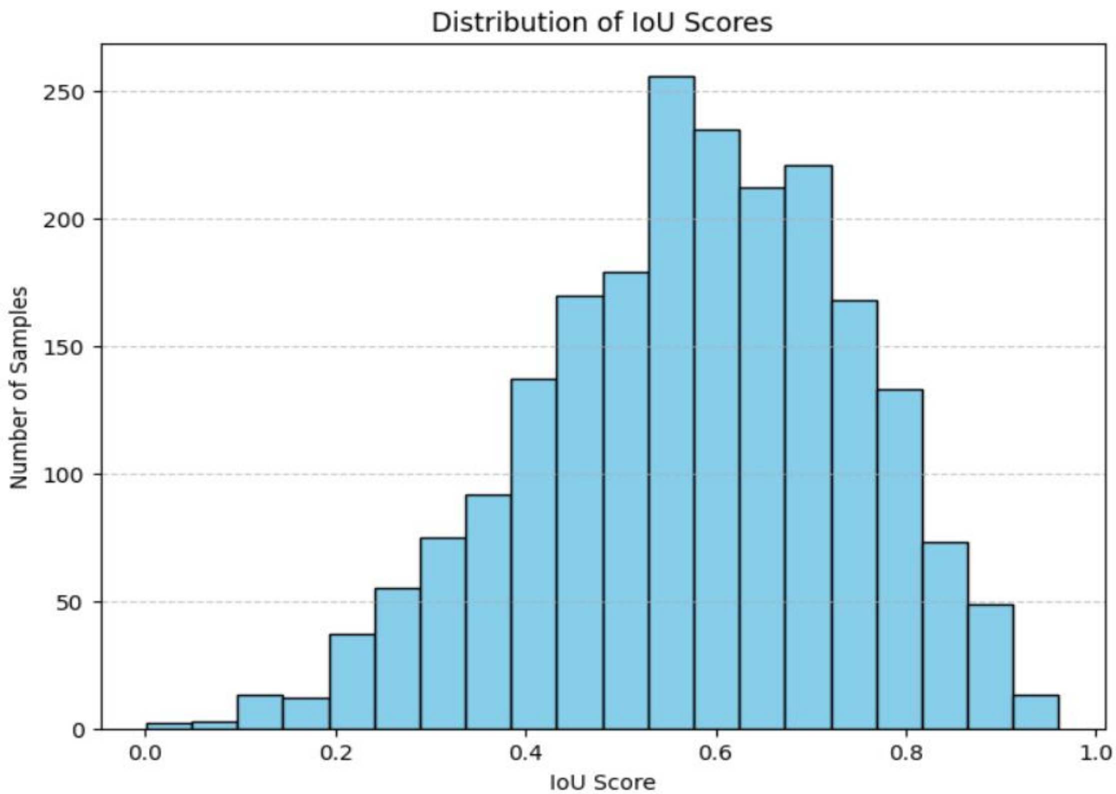


Fig. 1. Histogram of mIoU scores distribution

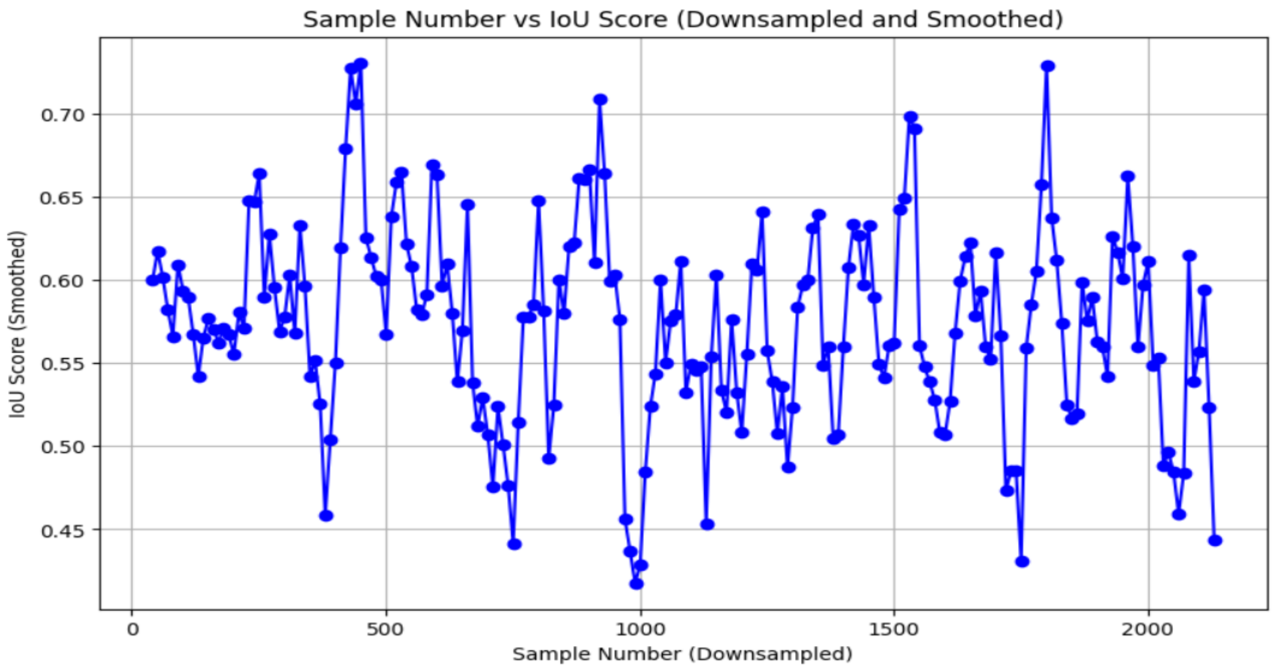


Fig. 2. Line chart to show the mIoU score for each sample image.

To visualize the distribution of the IoU scores across the dataset, Figure 1 presents a histogram depicting the frequency distribution of the obtained scores. Additionally, Figure 2 provides a line chart illustrating the smoothed and downsampled distribution of the 2135 segmented images' IoU scores. This chart aims to offer a clearer representation of the IoU distribution for better comprehension.

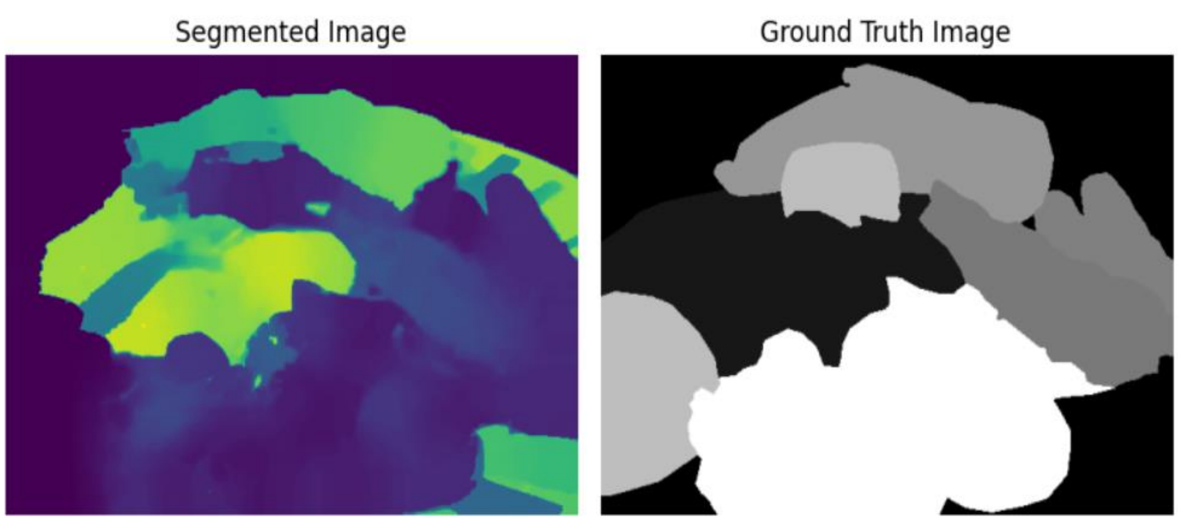


Fig. 3. Example 1 of segmented and ground truth images mIoU = 0.77.

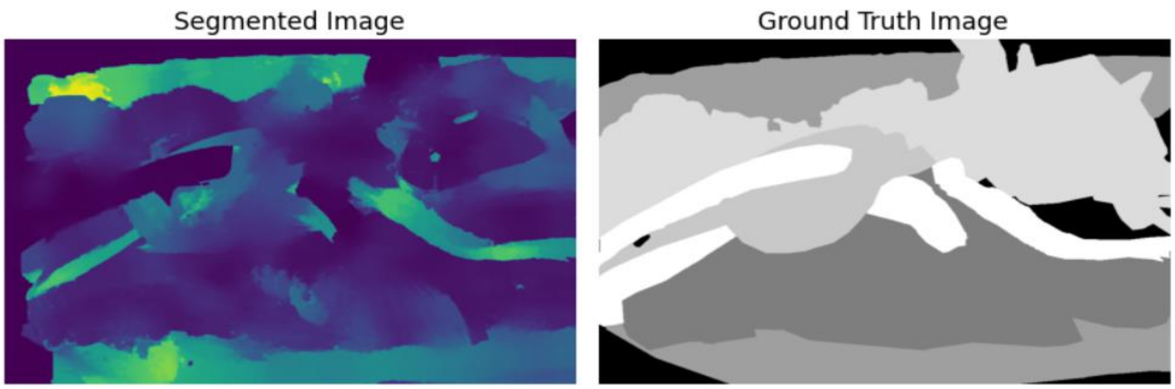


Fig. 4. Example 2 of segmented and ground truth images mIoU = 0.91.

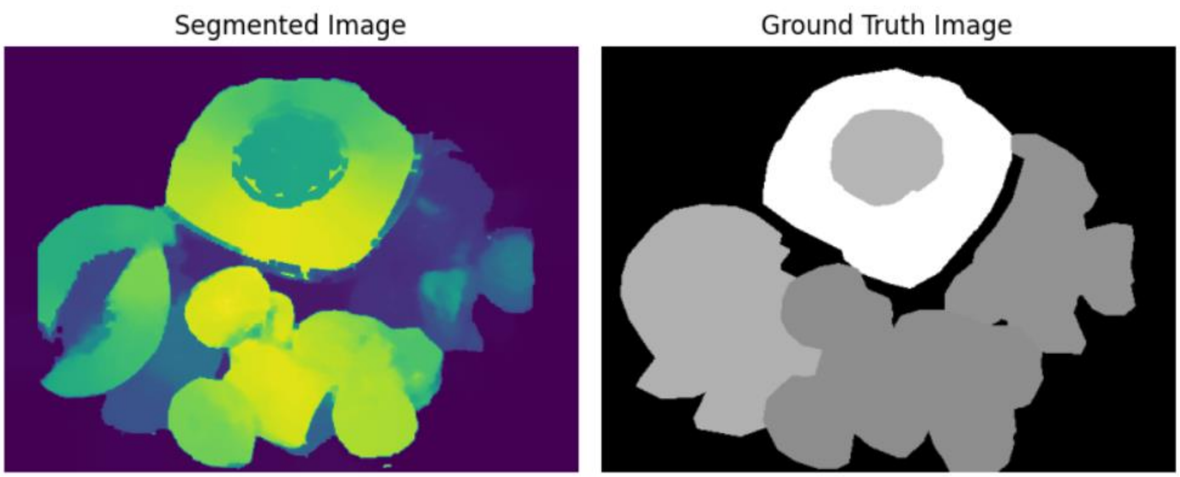


Fig. 5. Example 3 of segmented and ground truth images mIoU = 0.78.

Figures 3, 4, and 5 present visual comparisons between the algorithm's segmented images and their corresponding ground truth images. These examples showcase the algorithm's performance through specific outputs, allowing for a qualitative assessment of segmentation accuracy.

Conclusion

In conclusion, this project embarked on the development and evaluation of an image segmentation algorithm tailored for food item segmentation within a dataset comprising 2135 diverse food images. The algorithm showcased commendable performance metrics, notably yielding an average Intersection over Union (IoU) score of 0.57 on a scale of 0 to 1, signifying moderate accuracy in segmenting food items. Notably, the algorithm excelled in consistent seperation of contents from individual plates, irrespective of plate color, food texture, or item shapes. However, challenges emerged in scenarios involving multiple plates or combinations of plates and cups, highlighting areas for algorithm refinement.

A significant departure from traditionally employed techniques was the utilization of mean shift clustering over the conventionally used k-means algorithm. This strategic shift negated the necessity of pre-defining the number of clusters, thereby enhancing the algorithm's adaptability and generalizability across various food compositions. This pivotal change allowed for a more flexible segmentation approach, accommodating diverse arrangements without imposing rigid cluster constraints, thus contributing to the algorithm's robustness.

While the algorithm presents promising potential in segmenting single plate food arrangements, addressing challenges related to multi-object scenarios and computation cost remains a focal point for future enhancements. Further refinement and robustness in handling complex scenes are crucial for elevating the algorithm's accuracy and versatility in real-world applications.

References

1. Raju VB, Imtiaz MH, Sazonov E. Food Image Segmentation Using Multi-Modal Imaging Sensors with Color and Thermal Data. Sensors (Basel). 2023 Jan 4;23(2):560. doi: 10.3390/s23020560. PMID: 36679357; PMCID: PMC9860575.

2. Nassir Salman, Computer Science Department, Zarqa Private University of Jordan: Image Segmentation Based on K-Means clustering. and Edge Detection Techniques, 2006.

3. Pham, V.H., Lee, B.R. An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm. *Vietnam J Comput Sci* **2**, 25–33 (2015). https://doi.org/10.1007/s40595-014-0028-3

4. MariosAnthimopoulos, Joachim Dehais, Peter Diem and StavroulaMougiakakou: "Segmentation and Recognition of MultiFood Meal Images for Carbohydrate Counting". 2013 IEEE.

5. Xiongwei Wu, Xin Fu, Ying Liu, Ee-Peng Lim, Steven C.H. Hoi, and Qianru Sun. 2021. A Large-Scale Benchmark for Food Image Segmentation. In Proceedings of the 29th ACM International Conference on Multimedia (MM '21). Association for Computing Machinery, New York, NY, USA, 506–515. https://doi.org/10.1145/3474085.3475201