CSC8628: Image Informatics Coursework - Multi-Food Image Segmentation

Sandra Mishale Niño Arbelaez

I. Introduction

Food image segmentation is being used in different areas of research such as nutrition intake, carbohydrate measurements, and food volume estimation [1]. This is a challenging task in computer vision due to the “large diversity in food appearances and the imbalanced distribution of ingredients” [2]. Moreover, lighting conditions increase the difficulty of the segmentation process [3].

The dataset provided has a variety of images with different sizes and food types, food from different countries, and food with multiple ingredients and consistencies. Some images contain different non-food objects in the background, such as plates, forks, glasses, hands, and tables which make the segmentation task more difficult.

Consequently, different techniques were searched to approach these problems. In [3], image thresholding is applied to each compartment of the tray box which outputs a binary image and in [4], two phases of thresholding are applied to distinguish the foreground and background of food images on a white plate. However, our dataset is not suitable for image thresholding because we don't have a common single background and single-food images, but many different backgrounds and multi-food images. Moreover, the use of the K-means algorithm is applied in [5] where the number of clusters (K) is known for MR images to do the segmentation of the regions of the head, in this case, 4. Nevertheless, we don’t know the number of clusters for our images because of the variety of food items. Furthermore, in [6], a combination of pyramidal mean shift with watershed segmentation is applied to detect unique number of segments.

In this way, the core of the proposed method is the pyramidal mean shift algorithm which doesn’t require a number of clusters beforehand as K-means and is an “iterative algorithm for feature space cluster analysis” [7] as suggested in [7] and [6].

II. Proposed method

The proposed algorithm is divided into the following steps: image normalisation, gamma correction, pyramidal mean shift, morphological opening, and edge detection (Figure 1). The algorithm uses methods from OpenCV and Scikit-image.

Image preprocessing

We normalised the image to be in the range of [0, 255]. Then, we calculated the average intensity of the image. If this number is close to 255, we consider the image might be too light, therefore, for averages of intensities greater than 200, we applied a gamma correction with a gamma of 1.5 which results in a curve that darkens the image and makes some light correction. Otherwise, for all images, we made them lighter by applying a gamma of 0.5 to make certain features more visible.

Segmentation

The gamma-corrected image is used in the pyramidal mean shift to identify clusters in the images without specifying the number beforehand which iteratively replaces the locally weighted mean [8]. Moreover, due to the variety of image resolutions, using a pyramid makes the algorithm advantageous by starting with the original image and then reducing the resolutions to compute the algorithm, in our case, the maximum level of the pyramid is 2. For this, after a lot of experiments, the best parameters were sp = 15 and sr=40 which controls the spatial and colour radius, respectively.

Postprocessing

After segmentation, we applied morphological opening which is an erosion followed by a dilation which helps to smooth the contours and remove thin protrusions [9, pg 702]. For this operation, we applied a kernel of 5x5. Afterwards, we converted the image to grayscale to be compared with ground truth images and performed edge detection using the Sobel operator which computes the gradient of the image and helps to suppress noise [9, pg 781].

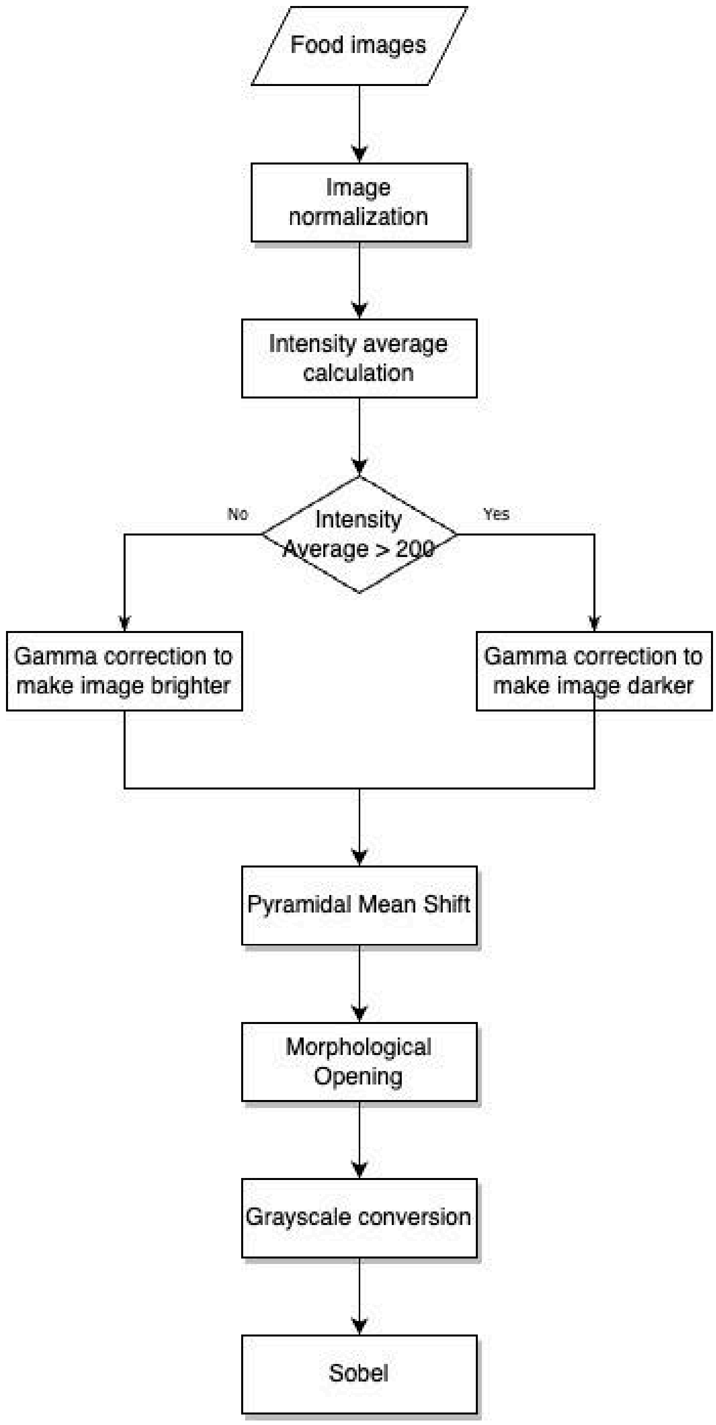

Figure 1. Proposed algorithm flowchart

III. Results, Key Findings and Discussion

After running the whole dataset of 2,135 images, the Mean Intersection Over Union was 0.5348, which means that the proposed algorithm has a large room for improvement.

Nevertheless, the proposed algorithm did very well in images where the food items are predominant in the image or zoomed in, without non-food items in the background, and the average intensity is in the range of 100 to 180, which means that the image is not too light but either too dark, indicating normal lighting conditions. For instance, the Intersection Over Union (IoU) for Figure 2 is 0.9319. This image has the food items exposed and the segmentation was performed correctly. We see that Mean Shift could detect the clusters of tomatoes, breads, meat, and vegetables. Also, we can notice a similar segmentation performance in Figure 3.

The algorithm produced a decent metric in some of these situations. In Figure 4, the IoU decreased when there is an increased texture in the food items and a region of the image is occupied by some non-food item, like in Figure 5, a white plate has more presence in the image. Additionally, in Figure 6, the algorithm started to struggle when the food ingredients overlapped. Probably an addition of the watershed algorithm could improve the segmentation of items which are overlapped with each other [6].

The introduced algorithm struggled in images with non-food objects in the background which made a lot of noise in the segmentation task as shown in Figure 7. This was because we didn't do recognition of non-food items to remove and clean the image with the correct target. Moreover, it didn’t perform very well when the food was far from the view (zoom out), therefore, it mixed the segmentation with the noise of the background such as in Figure 8. Additionally, in Figure 9, we can notice that when the food item was hidden or in a container such as a cup, bottle, glass or bowl, the algorithm also segmented the object where the food was contained.

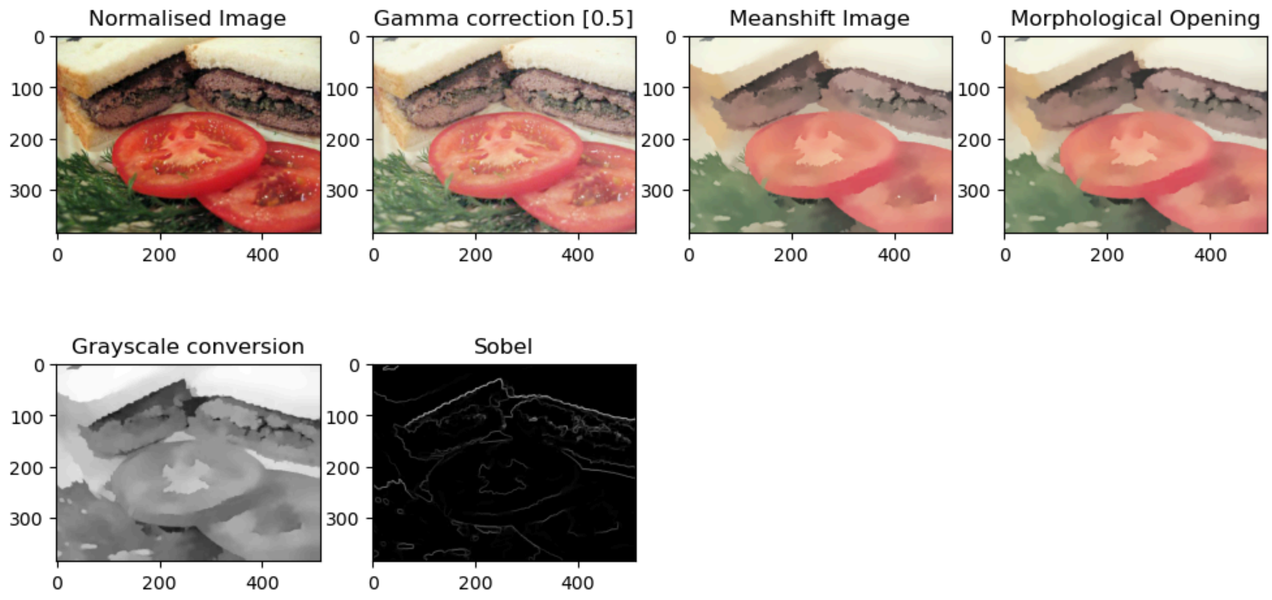

Figure 2. Sample image with IoU of 0.9319. Gamma correction of 0.5.

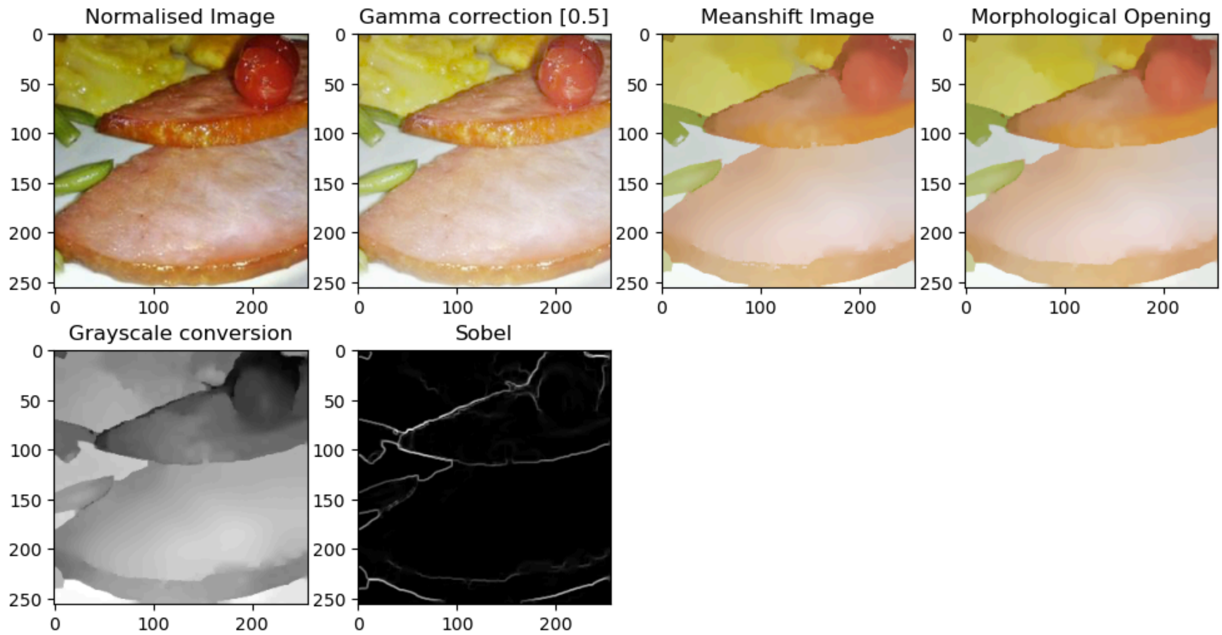

Figure 3. Sample image with IoU of 0.904. Gamma correction of 0.5

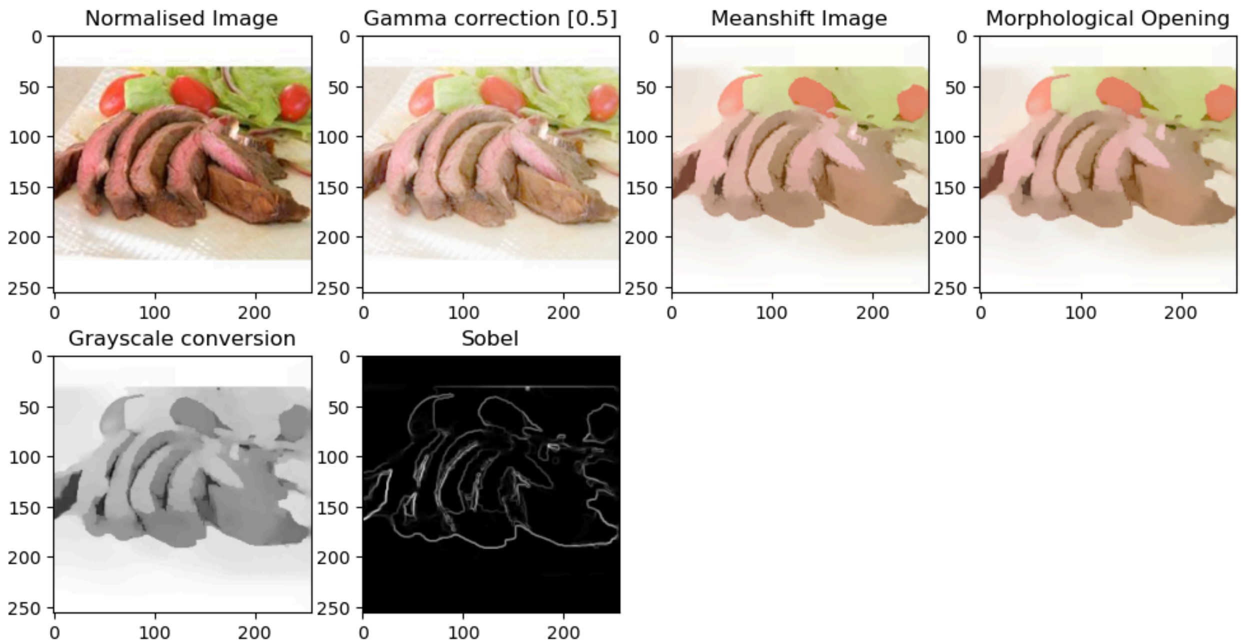

Figure 4. Sample image with IoU of 0.572.

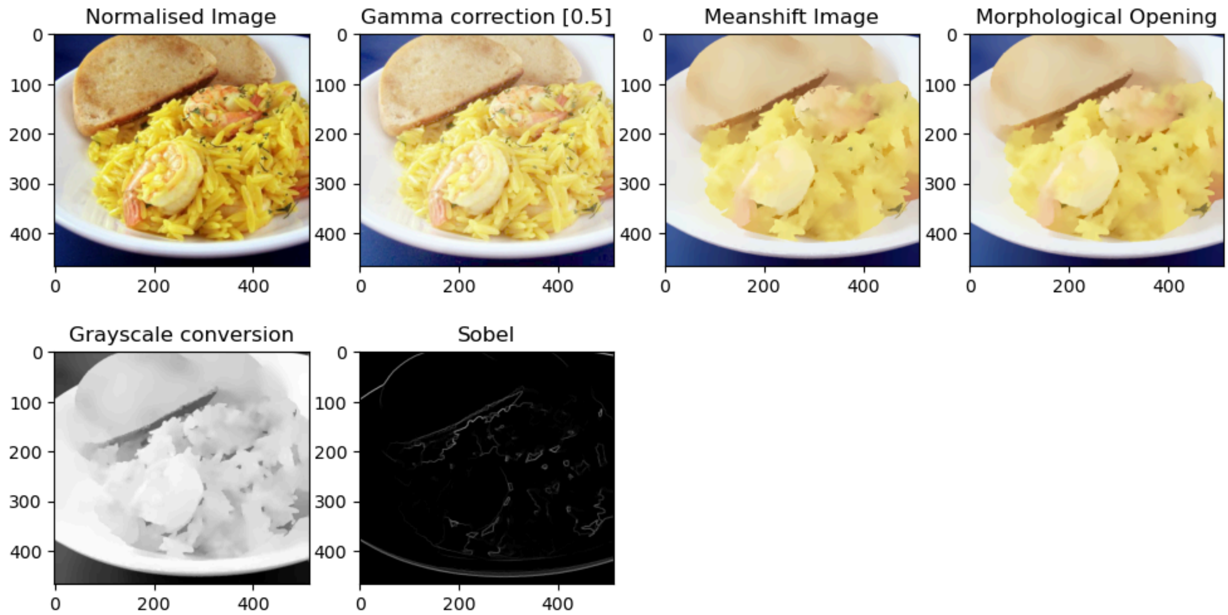

Figure 5. Sample image with IoU of 0.684.

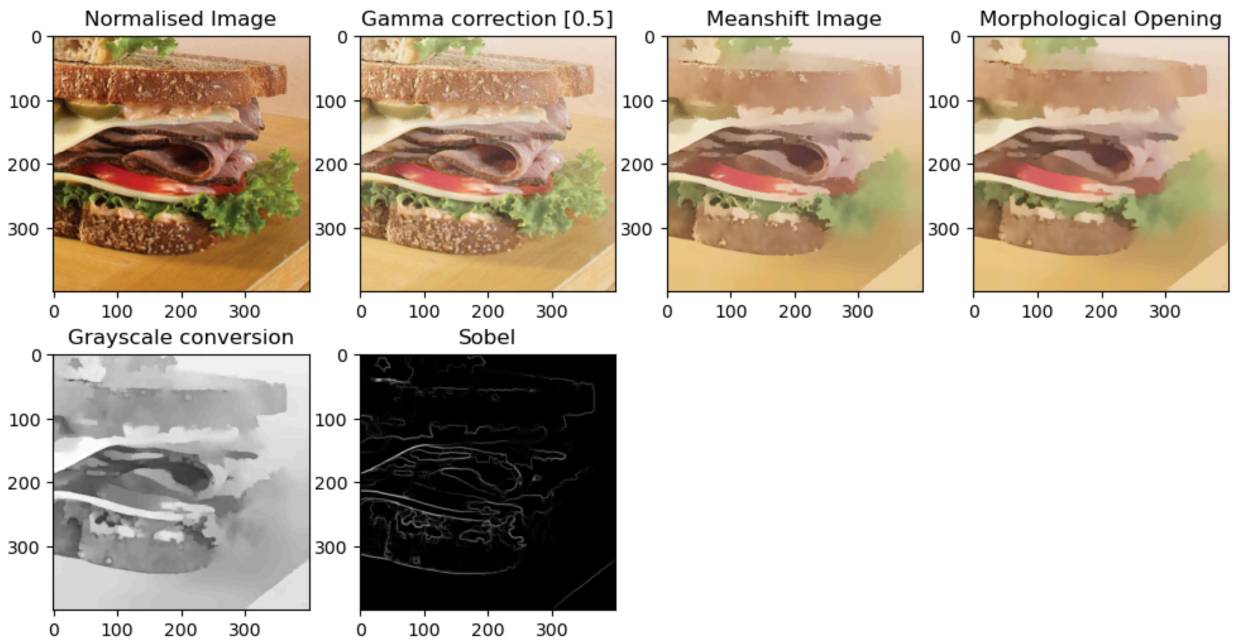

Figure 6. Sample image with IoU of 0.697

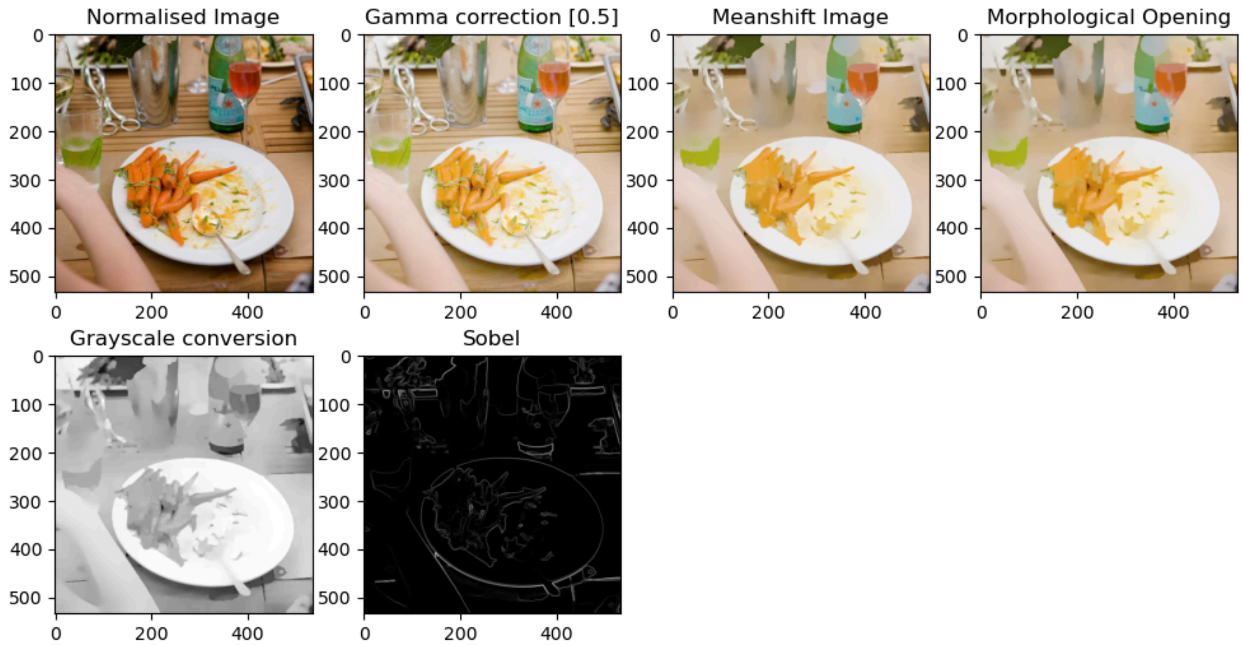

Figure 7. Sample image with IoU of 0.121.

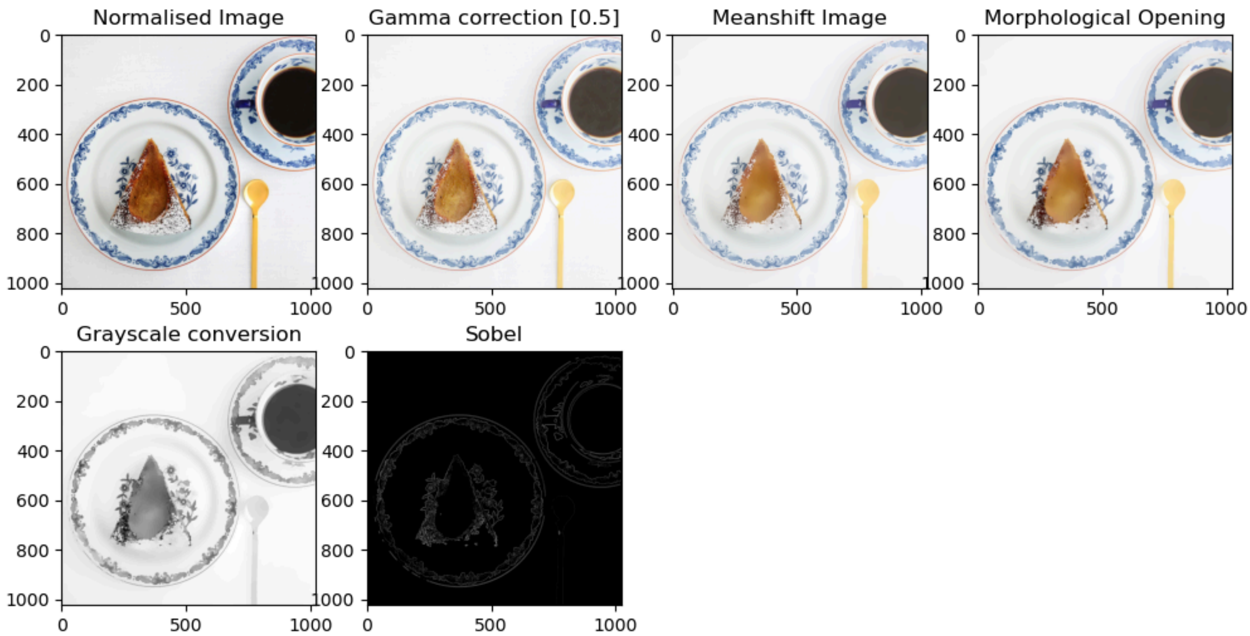

Figure 8. Sample image with IoU of 0.139

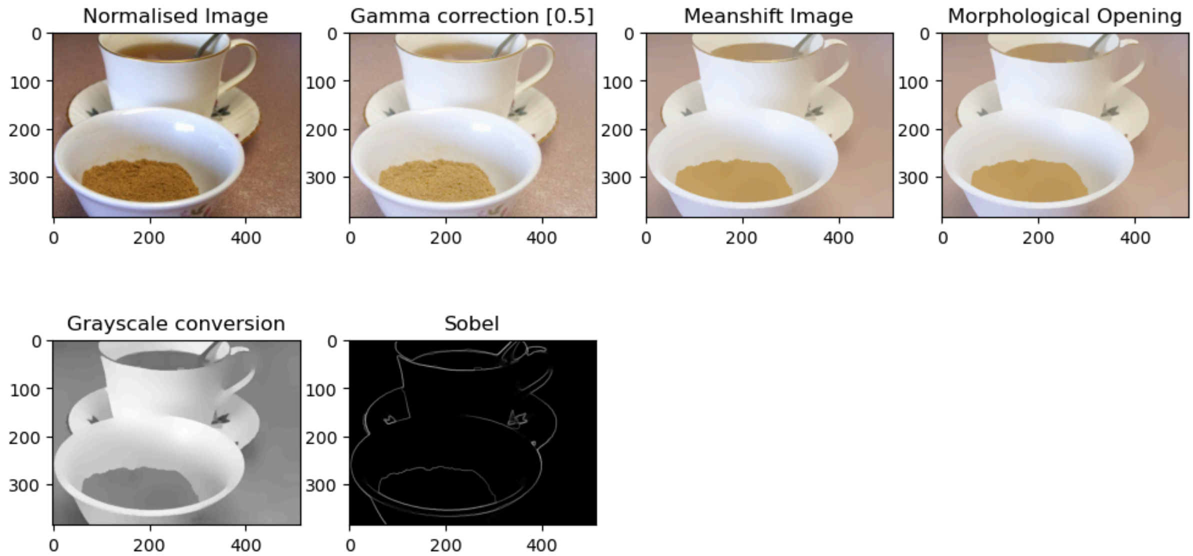

Figure 9. Sample image with IoU of 0.106

IV. Conclusions

This coursework highlights the challenges of multi-food image segmentation and proposes an algorithm with classical methods. To address this limitation, the core of the algorithm was the use of pyramidal mean shift with gamma correction as a preprocessing step and, as postprocessing steps, included morphological opening and edge detection with the Sobel operator.

The output of the algorithm was indifferent to the image resolutions, but it was more affected by the food's placement, the background noise with non-food items, and the average intensities. Consequently, the algorithm performed very well when the food items occupied the majority of the image and the average intensity was in normal lighting conditions but struggled when there were many non-food objects in the background, when the food items had an increased texture and overlapped with each other.

In summary, the algorithm introduced in this project reaches a mIoU of 0.5348 which is a decent metric, but we are aware that it has room for improvement that can be achieved with more preprocessing work, noise reduction, non-food items detection and the use of Convolutional Neural Networks.

V. References

[1] Burkpalli, Vishwanath & Patil, Priyadarshini C & Professor, Asst. (2016). A REVIEW ON SEGMENTATION TECHNIQUES FOR FOOD IMAGES. 250.

[2] Wu, X., Fu, X., Liu, Y., Lim, E., Hoi, S. C., & Sun, Q. (2021). A Large-Scale Benchmark for Food Image Segmentation. *ArXiv*. /abs/2105.05409

[3] Y. A. Sari, J. M. Maligan and A. F. Prakoso, "Improving the Elementary Leftover Food Estimation Algorithm by Using Clustering Image Segmentation in Nutrition Intake Problem," *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, Surabaya, Indonesia, 2020, pp. 435-439, doi: 10.1109/CENIM51130.2020.9298005.

[4] Yuita Arum Sari, Vriza Wahyu Saputra, Andini Agustina, Yudi Arimba Wani, and Yusuf Gladiensyah Bihanda. 2021. Comparison of image thresholding and clustering segmentation methods for understanding nutritional content of food images. *In Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology (SIET '20).* Association for Computing Machinery, New York, NY, USA, 124–129. https://doi.org/10.1145/3427423.3427441

[5] H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh and W. L. Nowinski, "Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm," *2006 IEEE Southwest Symposium on Image Analysis and Interpretation*, Denver, CO, USA, 2006, pp. 61-65, doi: 10.1109/SSIAI.2006.1633722.

[6] A. D. Anantha Padmanabha Reddy and P. Sriramya, "Segmentation of Food Items Using Watershed Algorithm and Predicting the Country of Food Items," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 7, no. 6S4, April 2019, ISSN: 2277-3878

[7] M. Anthimopoulos, J. Dehais, P. Diem and S. Mougiakakou, "Segmentation and recognition of multi-food meal images for carbohydrate counting," *13th IEEE International Conference on BioInformatics and BioEngineering*, Chania, Greece, 2013, pp. 1-4, doi: 10.1109/BIBE.2013.6701608.

[8] Richard Szeliski 1958-, *Computer vision algorithms and applications*. London ; New York: Springer, 2011.

[9] R. C. Gonzalez, *Digital image processing : global edition*, 4th ed. Harlow, United Kingdom : Pearson Education Canada, 2017, p. 1 online resource (1022 pages).

[10] V. Kumar, T. Lal, P. Dhuliya and D. Pant, "A study and comparison of different image segmentation algorithms," *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall),* Bareilly, India, 2016, pp. 1-6, doi: 10.1109/ICACCAF.2016.7749007.

[11] A. L. Siridhara, K. V. Manikanta, D. Yadav, P. Varun and J. Saragada, "Defect Detection in Fruits and Vegetables using K Means Segmentation and

Otsu’s Thresholding," *2023 International Conference on Networking and Communications (ICNWC)*, Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICNWC57852.2023.10127559.