# Carbohydrate Estimation using Computer Vision

Ganesh Rajasekharan

## ABSTRACT

One of the main challenges faced by individuals with type 1 diabetes is the estimation of carbohydrates (CHO) in non-packaged food. Estimating carbohydrates is crucial since it is necessary to administer the correct dose of insulin to maintain blood glucose levels post meal. In this project, an effort to identify images of food on a plate using computer vision techniques is carried out. Carbohydrate content is estimated by detecting the food type and area on the plate using Computer Vision. The information obtained about the food is compared against the nutritional information from the United States Food and Agriculture Departments(USDA) database. The amount of carbohydrates is estimated by finding the surface area after segmenting the food items from the plate.

## 1. INTRODUCTION

Bolus insulin is the type of dose of insulin which is taken before a meal to prevent hyperglycemia. Such a type of insulin usually needs to act rapidly since the glucose levels are elevated when carbohydrates in the meal are digested. One of the difficulties faced by individuals with Diabetes type 1 is to correctly estimate the amount of carbohydrates in their food intake in order to administer the correct bolus dose.

The aim of this project is to create an application which can be used by the diabetics to capture pictures of their meal to help with CHO estimation. The captured image would be run against an image processing pipeline which is described in detail in Section 7. This section describes the problems and challenges involved in implementing a Computer Vision based CHO estimation model. The first problem in identifying the food items on the table includes identifying the plate. A circular plate is used to serve the meals in the estimation purpose image. The shape of the plate in the image taken is leveraged to find an ellipse which fits around the boundary of the plate. The shape of the plate becomes elliptical due to the orientation of the camera in taking the picture.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

In this project, an edge detection technique such as Canny's edge detection algorithm is used to find the edges of the plate and then it is segmented out of the plate based on color. The food types are clustered by using Mean Shift clustering[2] based on texture, gradient and color features. Once the image is identified the features from the clustered image is extracted using a feature extraction algorithm to identify the food class. The extracted features such as the color would be used to identify the class of the food. Post identification of the food type based on features, the calorie information mainly of carbohydrates is verified against the information obtained from the United States Food and Agriculture Department's database.

## 2. RELATED WORK

There have been several attempts carried out to perform carbohydrate estimation in the past based on Computer Vision approaches. One of the first approaches towards tackling this problem was called the DiaWear System [8]. This model involved taking the picture of the food and using a reference object placed next to it to aid in the estimation of carbohydrates. Although the model was successful in identifying four fast food types based on a reference table it did not address the problem of calculating the volume of the food on the plane. Furthermore, there was no adaptive threshold calculation for dealing with images of low contrast. Another paper by Chen et al, [1] involved capturing a panoramic video of the food and comparing it with the known 7 classes of food types stored in a database. This was carried out on the assumption that the meals were always served in similar size.

Many of the proposed systems make use of only one image of the image and this limited the estimation of volume of food. The DietCam application proposed by Kong et al [4], involves capturing three images of the meal in order to aid in obesity prevention. This model also involves sparsely creating the 3D model of the food based on matching a database and involves the detection of only known objects. The Food Intake Visual and Voice Recognizer [7] application by Puri et al. uses 3 images as input and also asks the user to identify the food type by speech. Segmentation was carried based on classification of the local areas in the image.

A technique based on finding the volume by taking the two images of the food was proposed by Pouladzadeh et al [6], which involves taking one images from the side of the food and another form the top of the food. This was used to estimate the volume in the food by multiplying the height by the surface area of the food calculated from the top of

the image. This method obviously introduced significant error in estimation since food has to be at a uniform height throughout the plate to get accurate results.

## 3. ETHICAL CONSIDERATIONS

Since the bolus insulin is administered based on the carbohydrate calculation, it is imperative to achieve an estimation accuracy between the mean error range of 6 Âś 8 grams per dish which is acceptable. Any system with a medical application should always consider the situations when the estimation accuracy can suffer due to the varying size of the plates and overlapping food items. Also since the bolus insulin is fast acting, it would rapidly affect the blood glucose levels therefore any such estimation system has to be thoroughly tested for all possible scenarios.

An incorrectly dosed type-1 diabetic individual can undergo severe hypoglycemia. This can lead to several ill effects on the individuals health and can cause accidents due to suffering from hypoglycemia while driving. Although several systems have been proposed, none of them are currently approved by the FDA for medical purposes. There are several problems associated with such systems which requires further improvement in nutrient detection techniques.

## 4. BUSINESS CASE

We can clearly see that an accurate system to estimate carbohydrates would be beneficial. Especially in the case of diabetics, the estimation is crucial for their blood glucose level management to stay healthy and to avoid hyperglycemia. Also, many people are unaware or simply do not possess the skills required to estimate the carbohydrates in food manually. Also the success of such a commercial system would require intensive FDA trials for approval for commercial use as an aid for Diabetes. Another area where the stakes are not very high is the market of weight loss and healthy nutrition. People are more focused than ever on keeping track of carbohydrates in their meals especially while trying to stay healthy or to loose weight. Many people often face the dilemma of not knowing the calorie contents of food served in restaurant meals. This application could be a potential problem solver for such situations. Also the class of food items can be increased to classify more food which are not complex in their constituent ingredients.

## 5. DATA FOR IDENTIFICATION

The data for estimation would be created with a meal plate of the same size for different food items as discussed earlier. The image used for the demonstration contains food items such as pasta, breaded food and green salads for identification. The estimate of a single plate of items not overlapping each other would then carried out using the main function in Matlab. As a baseline the carbohydrate values would be manually calculated using the FDA database and compared against the results from the classification. The images would be scaled down to perform faster segmentation and edge detection.

## 6. NOISE REMOVAL TECHNIQUES

The food items can be segmented based on mean shift clustering algorithm [2]. This technique would preserve the strong color edges while smoothing out the fine grain texture. The outcome is desirable since the color pixels of the same food segments would have the same color compared to other food segment. The method would grow and cluster the pixels of the same food together resulting in the initial segmentation for the food items on the plate. This means that the food items does have to be physically separated but should not overlap each other. Overlapping would lead to the area estimation of the food items favoring the top item.

The figure(a) shows the Original image of the food. Figure(b) shows the resultant image after identifying the boundary of the image. The effect of segmentation can be noticed in figure(c) in which the plate pixels are removed by applying a mask to the original image. Median Shift Clustering on the image is carried and we can see that it has clustered the food based on texture and color in figure (d).

The food is segmented from the plate by converting the image to the Hue, Saturation and Value color space. A mask is created for the colors of food and is applied to the image as mentioned in [5]. The plate is removed by applying the mask to the original image to segment contents based on the color of the food. This would improve the accuracy of the next clustering and estimation techniques in finding the area of the food items. All the other pixels are set to zero and the original image is segmented to contain only the food constituents in order to identify them.

## 7. FEATURES AND ALGORITHM USED

This section summarises the algorithm used to detect the CHO content from the food. The following steps are carried out in the image processing pipeline.

1. The image is converted to binary with a threshold and Canny edge detection is performed.

2. The best fitting elliptical shape of the plate boundary is computed [9].

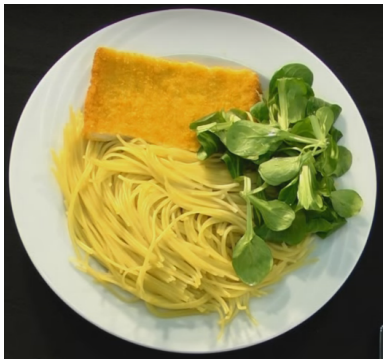
3. The area of the plate in pixels is calculated using minor and major axis of the boundary.

4. A mask is created for the plate by converting to HSV and applied on the original image.

5. Then the food is clustered from the plate using Median Shift Clustering

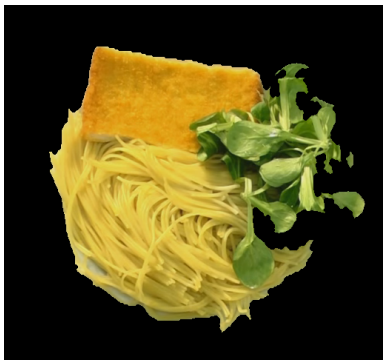
6. Finally the areas of the individual food items are calculated using pixel counts and finally total CHO is estimated.

The same plate is used for the estimation purpose, therefore the area of the plate is assumed to be known beforehand. In order to estimate the size in pixels, we need to find the pixels of the plate in the image. This would help in dealing with the scale space problem. Area enclosed by the plate in terms of pixels would be the area of the ellipse calculated in Figure(b). The area of the elliptical plate is calculated by multiplying the lengths of minor and major axes with the value of pi. Color and texture features are used for the food recognition step which consists of two steps. The first step is to cluster the food items based on the color and texture. Therefore Mean Shift Clustering Algorithm is used to cluster based on the pixel intensity and texture features of the food image.



(a) Original Image

(b) Elliptical boundary detected



(c) Food segmented from plate

(d) After Mean Shift Clustering

**Figure 1: Title for both**

## 8. SOLUTION FOR PROVIDING ESTIMATION

The challenging part of this project is to carry out texture and color based segmentation based on the food types. This was solved by using the Mean Shift Clustering technique which uses texture, color and gradient features for clustering. Also a region growing technique was tried out to grow the regions and segment the food cluster but the method did not provide satisfactory results. Another problem is to accurately estimate the area of the food present on the plate. The ellipse fitting around the plate helps to find the area in pixels which can be compared against the already known area of the plate. This method results in solving the scale issue of the picture due to the distance of the camera from the plate. Also the regionprops utility of Matlab was tried to get the properties of the segmented food clusters to obtain the area but the method did not yield accurate results due to the shape of food present in the plate like salads.

The approach of scanning bar codes would be suitable for estimating nutritional information of packaged items. This feature would also be included in a mobile application for future work. The project would be extended to create a mobile application to detect carbohydrates in packaged as well as non packaged food on a plate in the future. The use of mobile can also leverage the mobile phone sensors such as gyroscopes to provide better orientation to assist in image capture from the application. Also, I believe that for packaged food items once the image logo has be segmented using an appropriate edge detection algorithm (Sobel or Canny) the characters in the logo can be extracted using the Google Tesseract Optical Character Recognition Engine [10].

The application was executed against a set of images containing food items such as pasta, salads and breaded items. The CHO information for the foods and their corresponding nutritional information was found out from the USDA database. The data sets of images would be downloaded from the Internet consisting of various commonly available packaged items with logos. We can measure the success of this approach by testing an image and cross verifying it with the information available on the USDA database. Since the images are available before processing, we can establish the baseline by running the application on an item whose nutritional values are known beforehand.

## 9. FUTURE WORK

The Canny edge detection would have to implement some adaptive threshold setting for edge detection to improve the accuracy of the detection for low contrast images. The method for adaptive threshold computation would be carried out in future work. It would be done by calculating the average gradient magnitude of the food image to produce higher detection rates for images of poor contrast. Also the training data set needs to be improved to use a classification technique such as Support Vector Machine to assign the food to support many food class labels. Additionally, there are scenarios where the food can be placed on different plate sizes therefore the current model would have to be extended to accommodate such changes. In the future, I would like to extend the work carried out in this project to detect CHO in meals and to estimate the volume by placing a reference object placed next to the food such as thumb size or a reference card as mentioned by Anthimopoulos, Marios et al [3].

## 10. REFERENCES


[1] R. M. S. B. Chen N, Lee YY. Toward dietary assessment via mobile phone video cameras. *AMIA Annual Symposium Proceedings. AMIA*, pages 106–110, November 2008.

[2] M. P. Comaniciu D. Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Analysis Machine Intelligence.*, pages 603–619, November 2002.

[3] A. M. D. J. S. S. et al. Computer vision-based carbohydrate estimation for type 1 patients with diabetes using smartphones. *Journal of Diabetes Science and Technology*, 9(3):507–513, November. http://doi.org/10.1177/1932296815580159, YEAR = 2015.

[4] T. J. Kong F. Dietcam: Automatic dietary assessment with mobile camera phones. *Pervasive Mobile Computing 2012*, pages 147–163, November 2012.

[5] Mathworks. SimplecolordetectioâĂŃnbyhue. https://www.mathworks.com/matlabcentral/fileexchange/28512-simplecolordetectionbyhue--MONTH = "August", YEAR = 2010.

[6] A. T. Pouladzadeh P, Shirmohammadi S. Intelligent svm based food intake measurement system. *Computational Intelligence and Virtual Environments for Measurement Systems and Applications. IEEE*, pages 87–92, November 2013.

[7] Y. Q. D. A. S. H. Puri M, Zhu Z. Recognition and volume estimation of food intake using a mobile device. *Workshop on Applications of Computer Vision. 2009*, pages 1–8, November 2009.

[8] S. D. Shroff G, Smailagic A. Wearable contextaware food recognition for calorie monitoring. *International Symposium on Wearable Computers. IEEE*, 12(3):119–120, November 2008.

[9] M. Simonovsky. Ellipse detection using 1d hough transform. https://www.mathworks.com/matlabcentral/fileexchange/33970-ellipse-detection-using-1d-hough-transform MONTH = "November", YEAR = 2011.

[10] R. Smith. Tesseract open source ocr engine. https://github.com/tesseract-ocr/tesseract, 2013.