An Efficient approach for Estimation and Calculation of Calories and Nutrition in Food Image

Reshma V P[1], Karthik L[2]
[1]ECE Dept., Karavali Institute of Technology, Mangalore, Email:reshu610@gmail.com
[2]Asst. Professor, ECE Dept., Karavali Institute of Technology, Mangalore, Email:karthik.l465@gmail.com

Abstract. Each person should daily consume a certain amount of calories. If the amount of calories expenditure is increased, it will lead to weight gain and results in obesity. This paper proposed "efficient method for estimation and calculation of calories and nutrition in food image". The goal of the proposed measurement system is to record food intake, measure the size of food portions, measure calories and nutritional facts, compared with the existing clinical method. The system will start analyzing the image by using several steps, which are image processing, pre-processing, shape recognition, image segmentation. The process involves Gabor filter to measure local texture properties and SVM algorithm for data calculation (training data and testing data). The result of the system is acceptable and is an improvement over manual calorie calculation technique. The estimation and calculation of the food volume and amount of calories in the image is an essential step in the system. The proposed system is implemented in MATLAB.

Keywords: Calorie measurement, Image segmentation, Obesity.

I.INTRODUCTION

OBESITY causes more serious problems in adults. The imbalance between the amount of food intake and energy consumed by the individual's results in obesity which leads to people suffer serious health conditions like hypertension, heart attack, type II diabetes, high cholesterol, breast and colon cancers, and breathing disorders. Hence a balancing food intake must be followed to reduce body weight in healthy manner. In this paper, we propose a personal software instrument to measure calorie and nutrient intake which is differ from all other existing obesity treatment techniques require the patient to record all food intakes per day to compare the food intake with consumed energy. The system will start analyzing the image by using several steps, which are image processing, pre-processing, shape recognition, image segmentation. The process involves Gabor filter to measure local texture properties and SVM algorithm for data calculation (training data and testing data). The proposed system involves image processing and image segmentation to identify food portions and measures the volume of each food portions. The nutritional facts of each portion of food are obtained by calculating the total mass of each portion from its measured volume and matching it over available nutritional facts tables. While a preliminary description of this work has been presented in [1], here it is extend by proposing a more accurate measurement method for estimating food portion volume, which also works for food portions with an irregular shape, and by evaluating this approach with more food items. More importantly, the segmentation features are enriched by involving texture as well as color, shape, and size of the objects. The results show reasonable accuracy in the estimation of nutritional values of food types for which this system has been trained. Color and texture are the fundamental characters of natural images, and play an important role in visual perception. Color has been used in identifying objects for many years. Texture is one of the most active topics in machine intelligence and pattern analysis. Recently, different features of color and texture are combined together to measure food nutrition more accurately [2].

The proposed system has the following contributions.
1) This system is currently the only one that not only explains and discusses uncertainties in image-based food Calorie measurement, but also measures and presents actual uncertainty results using food images and its application scenario. This puts the system properly in the context of instrumentation and measurement research, and leads to more meaningful results for food recognition systems.
2) It gives clear idea about food image processing, Image segmentation, texture properties, classification, identification, and calorie measurement system that uses several images, but also under different conditions. The system uses variety of food such as solid, mixed or non mixed food.
3) The proposed system has following features color, texture, size, and shape, whereas most of the existing methods in this area, such as [3], use only color and shape.
4) In the proposed system for texture segmentation of food images a Gabor filter is used. To an input image a bank of Gabor filters with different wavelength and desired orientations are applied. With the same size of the input image, hence the outcome of each of these Gabor filters is a 2-D array. Therefore the sum of all the elements in an array is a number which represents the spatial frequency and matching orientation of the input image. In our system, six orientations are used as Gabor parameter.

Paper organization

The proposed paper is organized as follows. Section II covers related work in this area. Section III presents our system design, which is followed by Section IV, where our food portion volume measurement technique is proposed. Section V covers the performance evaluation of our proposed method, while Section VI analyzes the proposed work. Finally, Section VII concludes this paper.

II. RELATED WORK

For measuring daily food's dietary information there are several proposed method, which is typical of current clinical approaches, is the 24-h dietary recall [4]. The idea of this method is the listing of the daily food intake using a special format for a period of 24 h. This method requires a trained interviewer, such as a dietician, to ask the respondent to remember in details all the food and drinks she/he has consumed during a period of time in the recent past (often the previous 24 h). The 24 h requires only short-term memory, and if the recall is unannounced, the diet is not changed. In addition, the interview is relatively brief (20–30 min), and the subject burden is less in comparison with other food recording methods [5]. However, it is not always easy for a person to remember the actual contents as well as the amount of the food intake. In addition, to see an expert every 24 h is difficult and in many cases not feasible. The great majorities of existing clinical methods are similar to this, and typically require food records to be obtained for three to seven days, with seven days being the gold standard [1]. The problem with this manual approach is obvious: people not remembering exactly what they ate, forgetting to take note, and needing to see an expert dietician on a very frequent basis so the dietician can guess how much calories and nutrient the person has taken. To alleviate the shortcomings of these clinical methods, researchers have been trying to come up with improved techniques. To measure the amount of calorie some of the techniques requires the person to take a picture of the food before eating it so that the picture can be processed offline or manually or automatically. For example, the work in [6] proposes a method that uses a calibration card as a reference; this card should be placed next to the food when capturing the image, so that food dimensions are known. Since, this card must always be present in the photo when the user wants to use the system. The drawback is that the system will not work without this card, which means that in the case of absence or misplacement of the card. The proposed measurement system also uses a photo of the food, but uses the patient's thumb for calibration, which solves the problem of carrying cards or special trays. More specifically, an image of the thumb is captured and stored with its measurements in the first usage time (first time calibration). This unique method will lead to relatively accurate results without the difficulties of other methods. Food images will then be taken with the user's thumb placed next to the dish, so which allows to measures easier real life size portions. Finally we apply image processing and classification techniques to find the food portions, their volume, and their nutritional facts.

III. PROPOSED SYSTEM

The overall design of this system and its blocks are shown in Fig. 1. As the figure shows, at the early stage, images are taken by the user with a camera or mobile device followed by a preprocessing step. Then, at the segmentation step to extract various segments of the food portion each image will be analyzed.

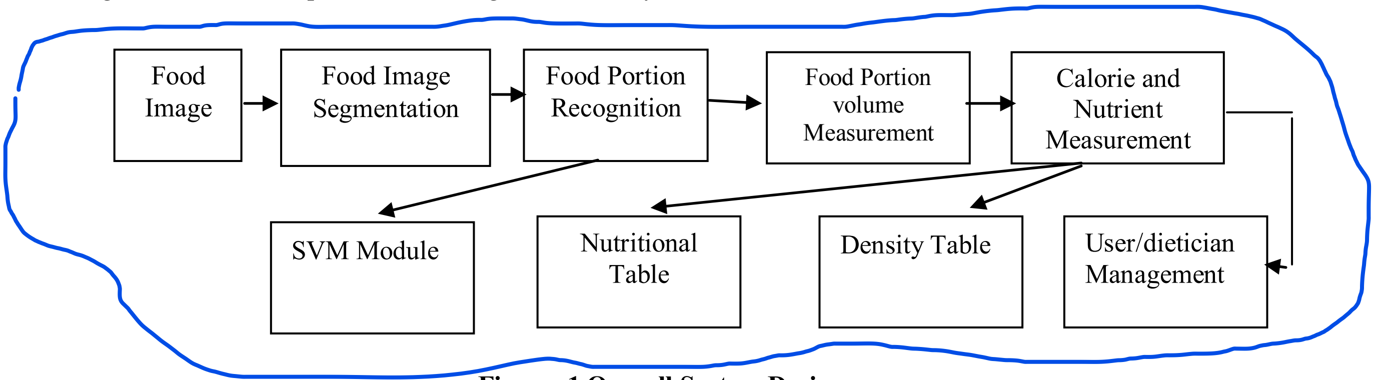

Figure. 1 Overall System Design

It is known that without having a good image segmentation mechanism, it is not possible to process the image appropriately. That is why color and texture segmentation tools are used. For each detected food portion, a feature extraction process has to be performed. In this step, various food features including size, shape, color, and texture will be extracted. The extracted features will be sent to the classification step where, using the support vector machine (SVM) scheme, the food portion will be identified. Finally, by estimating the area of the food portion and using some nutritional tables, the calorie value of the food will be extracted.

There is a one-time calibration process for the thumb, which is used as a size reference to measure the real-life size of food portions in the picture. The concept of using the thumb for calibration, as well as its implementation and evaluation is reported in [7] and [8] , respectively, and so we do not repeat them here.

The user captures two photos of the food: one from above and one from the side; the side photo is needed to measure depth, to have a more accurate volume measurement, as will be explained in Section IV. The system uses image segmentation on the photo taken from the top and uses contours to isolate various food portions. The detailed design, implementation, and evaluation of this image processing and segmentation component were described in [8].To measure local texture properties in the frequency domain a Gabor filter is used. Also in the proposed system Gabor filter-bank used [9]. It is highly suitable for the purpose where the texture features are obtained by subjecting each image to a Gabor filtering operation in a window around each pixel. We can then estimate the mean and the standard deviation of the energy of the filtered image.

Figure. 2 show the overall sequence of steps in this system.

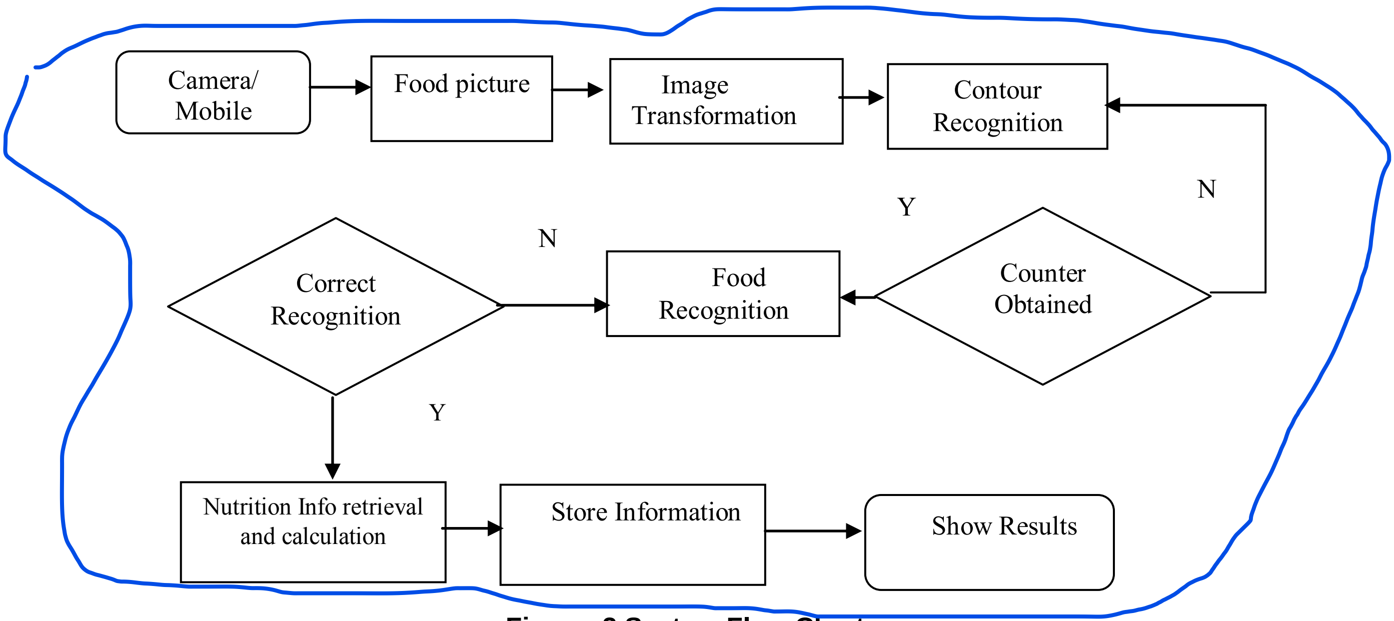

Figure. 2 System Flow Chart

The size of the block is proportional to the size of the segment. A Gabor impulse response in the spatial domain consists of a sinusoidal plane wave of some orientation and frequency, modulated by a 2-D Gaussian envelope. It is given by:

$$h(x, y) = -\exp\tfrac{1}{2}\{x^2/\sigma_x^2 + y^2/\sigma_y^2\} \cos(2\pi U_x + \varphi) \qquad (1)$$

Where *Ux* and *φ* are the frequency and phase of the sinusoidal plane wave along the *z*-axis (i.e., the 0° orientation), and *σx* and *σy* are the space constants of the Gaussian envelope along the *z*- and *y*-axis, respectively. A Gabor filter-bank consists of Gabor filters with Gaussian kernel function of several sizes modulated by sinusoidal plane waves of different orientations from the same Gabor-root filter, as defined in (1), it can be represented as:

$$g_{m,n}(x, y) = a^{-m} h(x', y') \qquad a>1 \qquad (2)$$

Where
$x' = x \cos\theta + y \sin\theta$
$y' = -x \sin\theta + y \cos\theta$
$\theta = n\pi/k$ (k = total orientation, n = 0, 1… k-1),
m = 0, 1... S-1).
Give an image I_E *(r, c)* of size *HxW*, the discrete Gabor filtered output is given by a 2-D convolution:

$$I_g(r,c) = \sum_{s,t} I_E(r-s, c-t) gm, n(st) \qquad (3)$$

As a result of this convolution, the energy of the filtered image is obtained and then the mean and standard deviation are estimated and used as features. We used the following parameters: five scales (*S* = 5) and six orientations (*K* = 6). In proposed system, we used Gabor filter for texture segmentation. In the implementation phase, each image is divided into 4×4 blocks, and each block is convolved with Gabor filter. Six orientations and five scales Gabor filters are used, and the mean and variance of the Gabor sizes are calculated for each block. In this project, using Gabor filter, we can identify five different textures and their identities as soft, rough, smooth, porous, and wavy. As the figure below shows, are used as classification inputs and the results will be the input of the SVM phase. For each feature, several categories are engaged, as shown in Figure. 3.

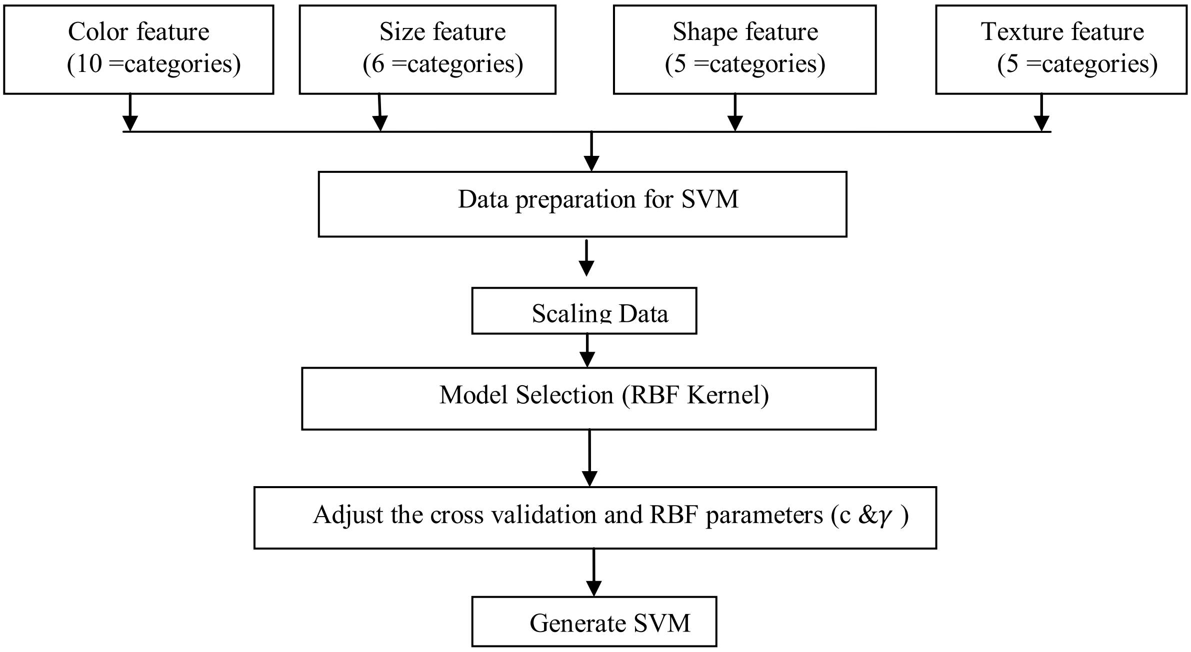

Figure 3. SVM algorithm

Some examples of various food types and their segmented portions are shown in Figure. 4.

Once the Image segmentation process completed next process is the Pattern Recognition which involves identifying the food items. The proposed system used SVM method for data classification, which is a well known method. A data classification usually involves training and testing data, which consist of some data occurrences. Each occurrence in the training set contains one class label and several features. The main aim of the SVM method is to produce a system, which predicts target value of data occurrences in the testing set, which are given only by their attributes. A Radial basis function (RBF) kernel is also used in proposed system which maps samples into a higher dimensional space in a nonlinear manner. In the proposed system, the feature vectors of SVM contain five texture, five color, three shape, and five size features.

The training vectors of SVM will be the feature vectors of each food items which was extracted during image segmentation process. For increasing the accuracy, after the SVM module has determined each food portion type, the system can optionally interact with the user to verify the kind of food portions. The system accuracy will increase. The system measures the volumes of each food portions and converts into mass, using available density tables. Finally to measure the overall calorie and nutrients in the food uses the mass and nutritional tables.

Figure 4. Segmentation of dishes into food portions.

IV. PROPOSED MEASUREMENT METHOD

A. Required Accuracy of the Measurement System

Before discussing any technical issues, it is important to understand what level of accuracy is expected from our system. To answer this question, we must first see what level of accuracy existing clinical methods have in their measurement of food's nutritional facts. There are two things to consider. First, if we put a plate of food in front of an expert dietician, she/he cannot give an accurate measurement of its nutritional facts by simply looking at it or even examining it manually, because it is impossible to know the exact contents of the dish, such as if this dish contains salt, and if so how much, or contains oil, and if so what type (olive, corn, animal-based, etc.), and how much, and so on. In addition, some food portions can be obstructed, for example, a piece of meat could be deep inside a soup, making it invisible to the dietician. Therefore, we can see already that high accuracy of calorie measurement is not possible in real life. Second, when we add this to what happens in existing clinical methods, in which the dietician goes over a list of food items recorded by the patient without necessarily even seeing the actual food or its picture, and without knowing size of portions, it becomes clear that accuracy is decreased even more. This is very important, because it directly affects the objectives of the system. The goal of this measurement system is to record food intake, measure the size of food portions, and measure nutritional facts, compared with the existing clinical methods. Hence it is very important to understand that high accuracy is not possible when dealing with food pictures only.

B. Measurement Unit: Calorie Definition and Nutritional Tables

Calorie is a typical measuring unit, which is defined as the amount of heat energy needed to raise the temperature of one gram of water by 1° [10]. This unit is commonly used to measure the overall amount of energy in any food portion that consists of the main food components of carbohydrate, protein, and fat. In addition to gram units, calorie units are also adopted in developing nutritional facts tables. Each person should take a certain amount of calories daily. If this amount is increased, it will lead to gain weight. Table I shows a small sample of a typical nutritional facts table.

C. *Food Portion Volume Measurement*

To measure the size of the food inside the dish, two pictures must be taken: one from the top and one from the side, with the user's thumb placed beside the dish when taking the picture from the top. The picture from the side can be used to see how deep the food goes, and is needed for measuring the food portions' volumes.

The system, which already has the dimensions of the user's thumb, can then use this information to measure the actual area of each food portion from the top picture, and can multiply this area by the depth (from the side picture) to estimate the volume of food. Let us see this in more details in the following paragraphs. To calculate the surface area for a food portion, a grid of squares is superimposed onto the image segment so that each square contains an equal number of pixels and, therefore, equal area. Fig.5 shows an example with an actual food portion. The reason for using a grid is twofold. When compared with other methods the grid will more easily match with irregular shapes, which is important for food images because most of the food portions will be irregular. Naturally, there will be some estimation error, but this error can be reduced by making the grid finer. If the grid is made finer, measurements become more accurate but will take longer time, and if the grid is made coarser, measurements become less accurate but the response time will be faster.

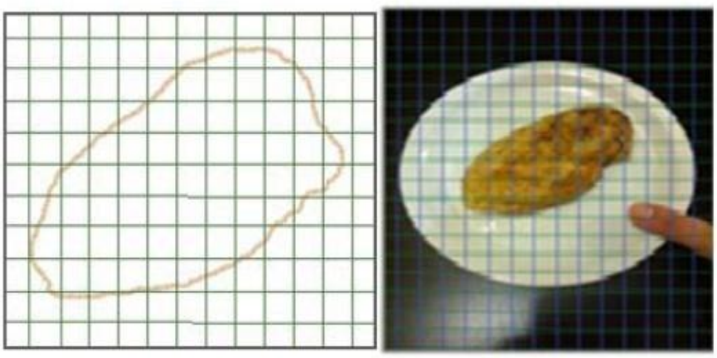

Figure 5. Methodology for food portion area measurement

The food portion of the total area (TA) is calculated as the sum of the sub areas (Ti) for each square (i) in the grid, as shown in equation (4)

$$TA=\sum_{i=1}^{n} Ti \qquad (4)$$

Where n is the total number of squares in the food portions area. After that, and using the photo from the side view, the system will extract the depth of the food, d, to calculate the food portion's volume, V, using the following:

$$V=TA \times d \qquad (5)$$

For better accuracy, if some food portions happen to be regular shapes such as square, circle, triangle, and so on, we can use geometric formulas to calculate their area, instead of using a grid. This requires an additional module that can recognize regular shapes.

D. Calorie and Nutrition Measurement

The volume measurement method described above is really just an interim step to measure the mass of the food portion. Once we have the mass we can easily calculate the amount of calories and other nutrition, as described next. It is known that the nutritional facts database is an important component for a useful and successful food recognition system [11].

Table I: Sample Of A Typical Nutritional Table.

Food Name	Measure	Weight (Grams)	Energy
Apple with Skin	1	140	80
Potato boil no skin	1	135	116
Cucumber	1	100	30
Tomatoes raw	1	123	30
Egg	1	150	17
Orange	1	110	62
cake	1	100	250

The data of nutritional values of foods are stored in these tables and are available from national and international health organizations. These tables, similar to the one shown in Table I, help us to calculate the amount of calories quickly and without reference to the Internet or an expert. At this point, we have the measurement for the volume of each food portion, and we can use the following general mathematical equation to calculate their mass:

$$M=\rho V \qquad (6)$$

Where M is the mass of the food portion and ρ is its density.

Food density can also be obtained from readily available tables. To extract the density of each food portion, the system needs to know the type of the food, which is done by the SVM-based food recognition module. An example of the information that is fed into the SVM module is shown in Fig. 6 right column. The SVM module uses this information and recognizes the type of food for each portion [12].

In addition, as mentioned earlier, at this stage, the system can ask the user to verify whether the food type recognized by the SVM module is correct. If not, the user can then enter the correct type. Now the system can calculate the mass by having the type of food. Therefore, the amount of calorie and nutrition of each food portion can be derived using nutritional tables, such as Table I, and based on the following:

$$\text{Calorie in the photo}=\frac{\textit{Calorie}\ \text{From Table} \times \text{mass in the Photo}}{\text{Mass from table}} \qquad (7)$$

E. Partially Eaten Food

It is possible that a user does not finish the entire food captured in the first picture that was taken before eating the food. This paper proposed a simple technique to increase measurement accuracy in such cases. If a user does not finish a meal, she/he should take another top picture of what is left of the meal. All of the above process can then be repeated on this new picture to calculate the amount of calorie and nutrient in the remaining food. The actual value of in-take is then adjusted by deducting the values of the remaining food.

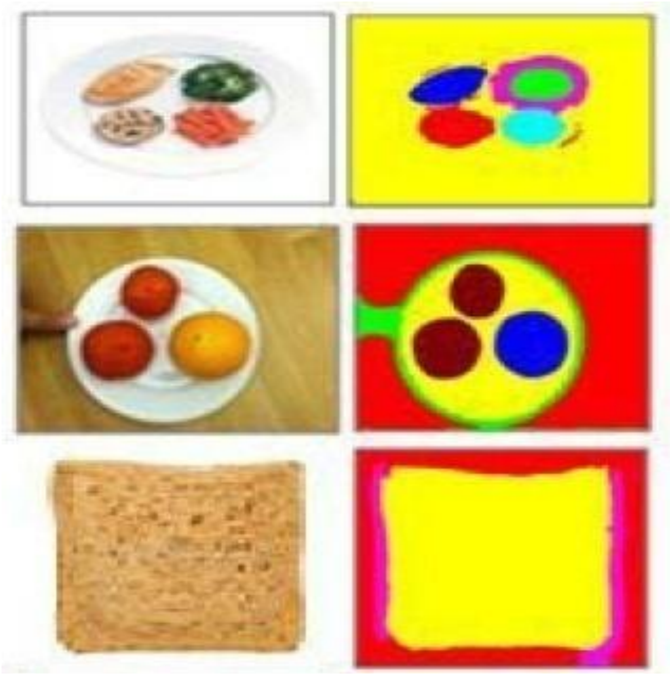

Figure 6. Before (left) and (right) color analysis and contour detection Right column is fed into SVM

V. PERFORMANCE EVALUATION

A. *Evaluation Strategy*

The system is implemented as a software prototype, where it is successfully segmented the food images and identified food portions using their contour inside of the dish [8]. Each portion is extracted one-by-one and analyzed them using the methods described in this paper. For the SVM part, we used around several different images for our method, which means a set of more than 300 images for each food portions, Approximately 150 for training set and then another 150 images as a testing set.

In the experiment, the color, texture, size, and shape properties of the food images were extracted after preprocessing, as shown in the examples of Fig. 6. Then observed the recognition result with features separately, which were color, texture, size, and shape, respectively.

B. Evaluation of the Recognition Systems

The system has low accuracy results for each separate feature, whereas involving joint combination of all features works well with an accuracy of approximately 93.22%. The system performance has examined using tenfold cross validation technique, and the accuracy of results is acceptable. Since in tenfold cross-validation, the input data are divided into 10 different groups, in each
iteration we have to test the method on the group of images, meaning that the results are for a group of images, not only for one single image. Compared with the tenfold cross method with the previous model in which we have tested the system using only one image in each step and the result is the accuracy of finding one food portion.

C Evaluation of the Area Measurement Technique

The proposed evaluation of the area measurement technique on a variety of simple food was done. (Not liquid like soup, curry, and so on). The area of each food portion is measured twice: once by hand from the image, and once using our proposed method. The experimental results, some of which are shown in Table II, show that the area measurement method achieves a reasonable error of about 10% in the worst case, and less than 1% in the best case.

Table II: Area Measurement Experiment Results

Food Type	Error Percentage
Bread	0.63%
Cake	2.30%
Cookies	0.50%
Omelet	10.5%

D. System Accuracy

To evaluate the accuracy of the proposed method, two different simulation scenarios are performed. In the first one, the proposed method is applied on several food portions, and their type and volume are extracted. Using the type and volume of each food portion, its mass is extracted using a density table [13]. Using the extracted mass, the calorie of each food portion is derived using Table I. In the second scenario, the real food portion is actually weighted and its real calorie is extracted using the tables. Finally, we have compared the extracted calories from these two scenarios. Some of the results are shown in Table III. The accuracy of the proposed system in non mixed food is approximately around 87.16%. as shown in Table III.

E. Uncertainty Measurements

When dealing with repeated measurements, there are three important statistical quantities: average (or mean), standard deviation, and standard error. In this system, the following parameters may have effects on the results: illumination, camera angle, and the camera itself. Illumination is one of the important parameters, which affect the system outcome because illumination directly affects the segmentation algorithm, which in turn affects the rest of the algorithms. To consider, pictures are taken with the same plate placed in three different locations with different illuminations. This strategy was repeated for all of the images in our database. The second effective parameter is the angle of photography; three different angles are chosen, which are approximately 30°, 90°, and 150° from the plate of food for all pictures. This means that for each plate in three different locations we have also gotten three more pictures from different angles. Finally, the camera itself will have an effect on the results in terms of its lens, hardware, and software. We discussed above that three different illuminations are selected for the plates, each illumination combined with three different angles, and each angle taken with three different cameras. This means that we have 27 images for each plate of food in various conditions. This gives a good opportunity to measure uncertainties.

Table III: Accuracy 0f Proposed Method in Comparison With Real Values.

Food Portions	Weight (Grams)	Calculated calorie	Real Calorie	Absolute Accuracy (%)
Cake	100	275	250	90
Egg	150	15	17	88
Apple	200	100	114	87
Tomato	150	23	30	76
Cucumber	100	27.5	30	91
Orange	160	98	90	91
Average Accuracy				87.16

VI. ANALYSIS

There are still some limitations with our system, as follows.

1) This method still has problems in detecting some mixed foods. In the current version of the proposed method, the segmentation step often fails to properly detect various food portions in mixed foods. In addition, illumination of food portions in a mixed food may be changed as they get mixed, making it harder to extract different food portions. Furthermore, the size of food portions indifferent mixed food are not similar, hence the method fails to segment food portions properly. To solve this problem, the segmentation mechanism can be improved, to better support mixed food as well, with the following plan for the future work.

a) To apply and test other methods such as graph cut segmentation. Having a more accurate segmentation method helps us to extract more reliable features for recognition phase.
b) To train the system with more mixed foods, to expand the operation range of the system.
c) To increase the accuracy of segmentation, also to increase the range of each feature; e.g., expanding the range of color or texture features.

2) To achieve higher accuracy following steps are followed:
a) Using more and improved segmentations methods for accurate estimation of the area of each food portion.
b) White background plates with smooth texture are used for simulations.

VII. CONCLUSION

In this paper, the proposed system identifies the food items in an image using image processing and segmentation, food classification using SVM, food portion volume measurement, and calorie measurement based on food portion mass and nutritional tables. The result of the system is acceptable and is an improvement over manual calorie calculation technique. The estimation and calculation of the food volume and amount of calories in the image is an essential step in the system. The proposed system is implemented in MATLAB.

REFERENCES

[1] R. Almaghrabi, G. Villalobos, P. Pouladzadeh, and S. Shirmohammadi, "A novel method for measuring nutrition intake based on food image," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Graz, Austria, May 2012, pp. 366–370.
[2] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 124–128, Jan. 1998.
[3] S. Qin, Y. Wu, and Y. Deng, "An automatic food recognition algorithm with both," *Proc. SPIE*, vol. 7489, pp. 1–8, Jul. 2009.
[4] M. Livingstone, P. Robson, and J. Wallace, "Issues in dietary intake assessment of children and adolescents," *Brit. J. Nutrition*, vol. 92, pp. S213–S222, Oct. 2004.
[5] (2011, Oct.). *Obesity: A Research Journal* [Online]. Available: http://www.nature.com/oby/journal/v10/n11s/ full / oby2002192a.html
[6] S. Mingui, L. Qiang, K. Schmidt, Y. Lei, Y. Ning, J. D. Fernstrom, *et al.*, "Determination of food portion size by image processing," in *Proc. 30th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2008, pp. 871–874.
[7] K. E. Barrett, S. Boitano, S. M. Barman, and H. L. Brooks, "Digestion, absorption & nutritional principles," in *Ganong's Review of Medical Physiology*. New York, NY, USA: McGraw-Hill, 2009, ch. 27.
[8] G. Villalobos, R. Almaghrabi, P. Pouladzadeh, and S. Shirmohammadi, "An image processing approach for calorie intake measurement," in *Proc. IEEE Symp. Med. Meas. Appl.*, Budapest, Hungary, May 2012, pp. 1–5.
[9] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognit.*, vol. 24, no. 12, pp. 1167–1186, 1991.
[10] G. Villalobos, R. Almaghrabi, B. Hariri, and S. Shirmohammadi, "A personal assistive system for nutrient intake monitoring," in *Proc. ACM Workshop Ubiquitous Meta User Inter.*, Scottsdale, AZ, USA, 2011, pp. 17–22.
[11] W. Wu and J. Yang, "Fast food recognition from videos of eating for calorie estimation," in *Proc. Int. Conf. Multimedia Expo*, Jul. 2009, pp. 1210–1213.
[12] P. Pouladzadeh, G. Villalobos, R. Almaghrabi, and S. Shirmohammadi, "A novel SVM based food recognition method for calorie measurement applications," in *Proc. IEEE Int. Workshop Interact. Ambient Intell. Multimedia Environ.*, Melbourne, Australia, Jul. 2012, pp. 495–498.
[13] K. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181–201, Mar. 2001.