

Real Time South Indian Food Recognition

N. Rajesh Kumar and M. Vignesh Babu

Assistant Professor,

ECE, SACS MAVMM Engineering College,

Madurai, Tamil Nadu, India.

Abstract— In this paper, we propose an automatic food recognition from various food images. The main theme of our project is to propose a novel methodology with CIELAB conversion, Mean shift filtering, Region growing and SVM classification for automatic food recognition. CIELAB conversion is used to convert the food images from RGB to L*A*B. Mean shift filtering is used to find the set of neighbouring pixels in the food image. Region growing process is used for segmentation of the region iteratively by comparing all unallocated neighbouring pixels to the region. Images are trained using SVM classifier with the features extracted. From the trained features, the SVM classifier detects the food images more accurately. The accuracy of food detection in our proposed method was obtained as 93.25%.

Index Terms— SVM classifier, image processing, food recognition, region growing.

I. INTRODUCTION

There has been a clear cut increase in the health consciousness of the global urban community in the previous few decades. Given the rising number of cases of health problems attributed to obesity and diabetes reported every year, people (including elderly, blind or semi-blind or dementia patients) are forced to record, recognize, and estimate calories in their meals.

Consequently, there is a rise in the market potential for such fitness applications products which cater to the demand of logging and tracking the amount of calories consumed. Food recognition system is essential for diabetic patients to analyze the calorie value before consumption. A region growing algorithm is used to grow a seed pixel over an entire area of the corresponding food region. The complementary information from learned color histograms and Local Binary Patterns provide an informative description that lead to high classification rates.

After finishing the iterations, the starting pixel gets the final mean color. The distance was chosen instead of the Euclidean in order to reduce the significance of lightness and eliminate the shadow effect.

II. RELATED WORKS

The segmentation of easily separated food items in a white plate, adaptive thresholding and connected component analysis have been used. Food items were segmented by matching the Scale Invariant Feature Transform (SIFT) points of the image with a reference image dataset. Multi-class classification was proposed for segmentation via the generation of saliency maps.

Fractal information has also been investigated for texture description, Xu et. al. [1] proposed the usage of the

multifractal spectrum as an extension of the multifractal analysis [2, 3]. It provides an efficient framework combining global spatial invariance and local robust measurements by capturing essential structure of textures with low dimension. Their methods achieved bi-Lipschitz transform invariance, i.e. invariance against translation, rotation, perspective transformation, and texture warping on regular textures. Varma et al. [4] proposed locally invariant fractal features for statistical texture description, which are also based on multifractal analysis.

Food images were segmented using normalized cuts based on intensity and color. Average color values and Gabor features were used. Color neighborhood and maximum response features have been used within a bag-of-features model. Classification accuracy was less.

III. PROPOSED SYSTEM

In our proposed system, we use Support Vector Machine and mean-shift algorithm for the food segmentation. We adopt a pyramidal approach which reduces substantially the computational cost and together with the proposed LAB-based color distance achieves both accuracy and efficiency. The information from learned color histograms and Local Binary Patterns provide an informative description that lead to high recognition rates.

A. Food Recognition

Food recognition is performed for each of the window images within the given bounding boxes. Firstly, image features are extracted from each window, secondly feature vectors are built based on the pre-computed codebook, and finally we evaluate the feature vectors with the trained linear SVMs on 50 food categories. The top five candidates which have the top five SVM scores over 50 categories and all the bounding boxes are shown on the screen as food item candidates.

B. Work modules

Module 1: Acquiring the image

Initially we need to produce a digital image. This can be done using camera.

Module 2: Preprocessing and segmentation

Module 3: Feature extraction

Module 4: Classification of food image

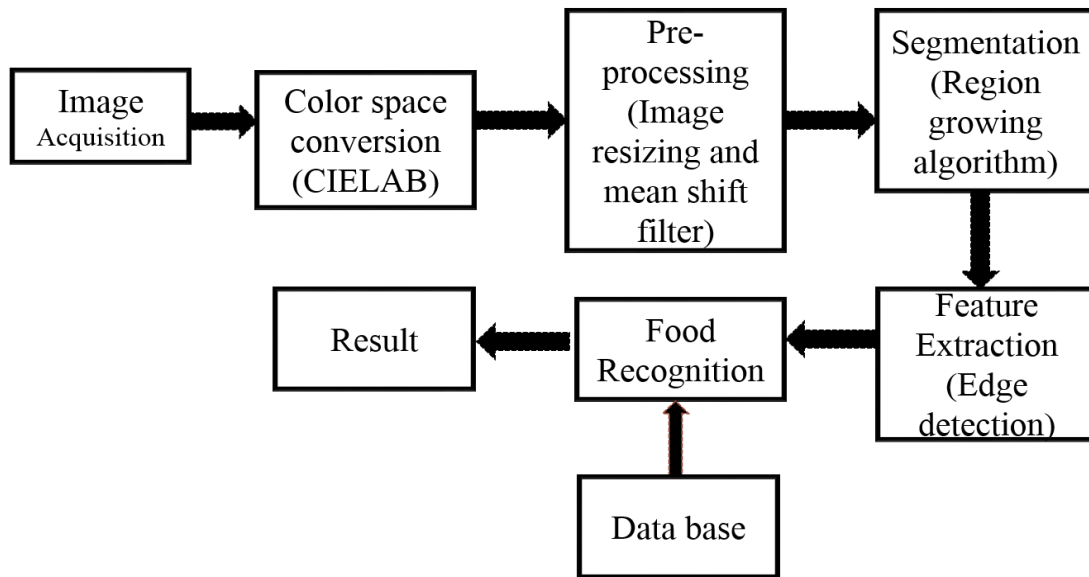


Fig. 1. Block diagram of proposed food recognition method.

C. Image acquisition

The South Indian food images are taken using Smart Phone and transferred via the USB cable.



Fig. 2. Image acquisition.

D. Color space conversion

In this phase, the conversion of RGB color values to YCbCr color space is carried out.

- Y- Luminance
- Cb- Chroma:blue
- Cr- Chroma:red

Digital Y'CbCr (8 bits per sample) is derived from analog R'G'B' as follows:

$$Y' = 16 + (65.481 \cdot R' + 128.553 \cdot G' + 24.966 \cdot B')$$

$$C_B = 128 + (-37.797 \cdot R' - 74.203 \cdot G' + 112.0 \cdot B')$$

$$C_R = 128 + (112.0 \cdot R' - 93.786 \cdot G' - 18.214 \cdot B')$$

Fig. 3. Color space conversion.

E. CIELAB Conversion

The input image is converted to the perceptually uniform CIELAB color space. The Euclidean distance between two colors in CIELAB is considered more representative of their visual color difference.

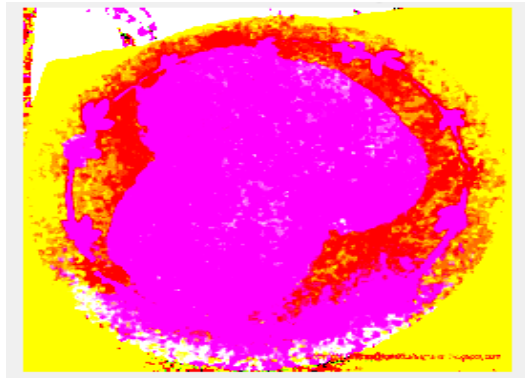


Fig. 4. CIELAB color space conversion.

IV. PROCESS OF FOOD RECOGNITION

A. Pre-processing

The aim of preprocessing (Fig. 5) is an improvement of image data that suppresses unwanted distortions or enhances



Fig. 5. (a) Original image, (b) Resized image.

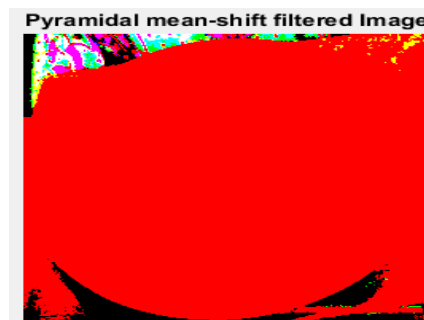


Fig. 6. Mean shift filtered image.

B. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. In our paper, we use Region Growing Algorithm for the Segmentation process.

C. Region Growing Algorithm

The region is iteratively by comparing all unallocated neighboring pixels to the region. The difference between a pixel's intensity value and the region's mean, is used as a

some image features important for further processing. They are,

- Image Resizing
- Pyramidal mean shift filtering

Image resizing

Image resizing is necessary when we need to increase or decrease the total number of pixels.

Mean Shift Filtering

Mean shift filtering is a data clustering algorithm. For each pixel of an image, the set of neighboring pixels is determined. The new spatial center and new color mean value are calculated. These calculated mean values will serve as the new centre for next iteration. The result of mean shift filtered image is shown in Fig. 6.

measure of similarity. The pixel with the smallest difference measured this way is allocated to the region. The code segments a region based on the value of the pixel selected (the seed) and to which thresholding region it belongs. Based on the region growing algorithm, we consider four neighboring pixels. Region growing algorithm is used for segmentation.

- means that the segmentation must be complete; that is, every pixel must be in a region.
- Requires that points in a region must be connected in some predefined sense.
- Indicates that the regions must be disjoint.

- Deals with the properties that must be satisfied by the pixels in a segmented region.

EG: $P(R_i)=TRUE$ if all pixels in R_i have the same grayscale.

- Indicates that region R_i and R_j are different in the sense of predefined P .

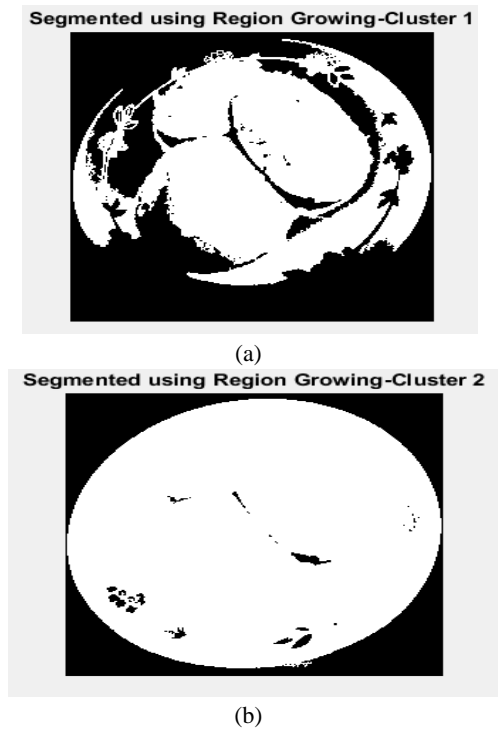


Fig. 7. Cluster formation using Region growing – (a) Cluster 1, (b) Cluster 2.

D. Feature Extraction

Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved.

In this paper, we use the feature extraction of LBP from the Edge detected images using Canny operator.

E. Edge Detection

Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Edge detection is used for image segmentation and data extraction in areas such as image processing. Edge detection is used to identify the edges in an image.

Algorithm for Canny edge detection

```
I = imread ('database/idly.jpg');
BW1 = edge (I,'canny');
Fig.;
imshowpair (BW1,'montage')
title ('Canny Filter');
```

Canny edge detection is a technique to extract useful structural information from the different vision object.

1. Apply the filter to smooth the image in order to remove the noise.
2. Find the intensity gradients of the image.
3. Apply the non-maximum suppression to get rid of spurious response to edge detection.
4. Apply double threshold to determine potential edges.
5. Track edge by hysteresis.

F. SVM Classification

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges (Fig. 8). However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well. SVM classification consists of two phases:

- » Training
- » Testing

SVM works really well with clear margin of separation. It is effective in high dimensional spaces. It is effective in cases where number of dimensions is greater than the number of samples. It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

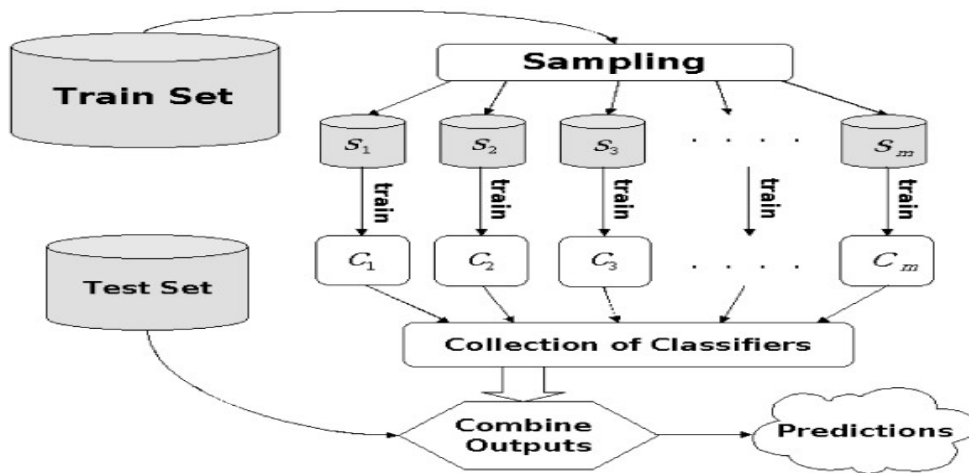


Fig. 8. Classification steps in SVM.

G. KNN Classifier

It is defined as K-Nearest Neighbors – Classification. K-nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970’s as a non-parametric technique.

Algorithm

A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor.

$$kNN = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Where, p and q are Euclidian vectors.

V. RESULTS AND DISCUSSION

The proposed food detection methodology uses region growing algorithm for segmentation and SVM classifier for accurate classification. The SVM is trained with features extracted from the source or training image and based on the features the classifier detects the test image. The proposed method is analysed using certain parameters and tabulated in Table 1. The same is compared with other conventional methodologies in terms of classification accuracy and detection time in Table 2.



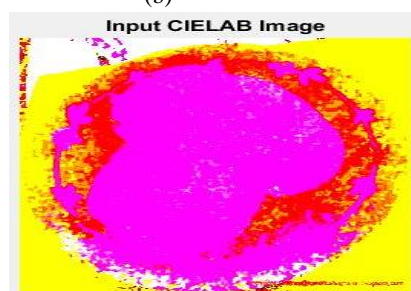
(a)



(b)



(c)



(d)

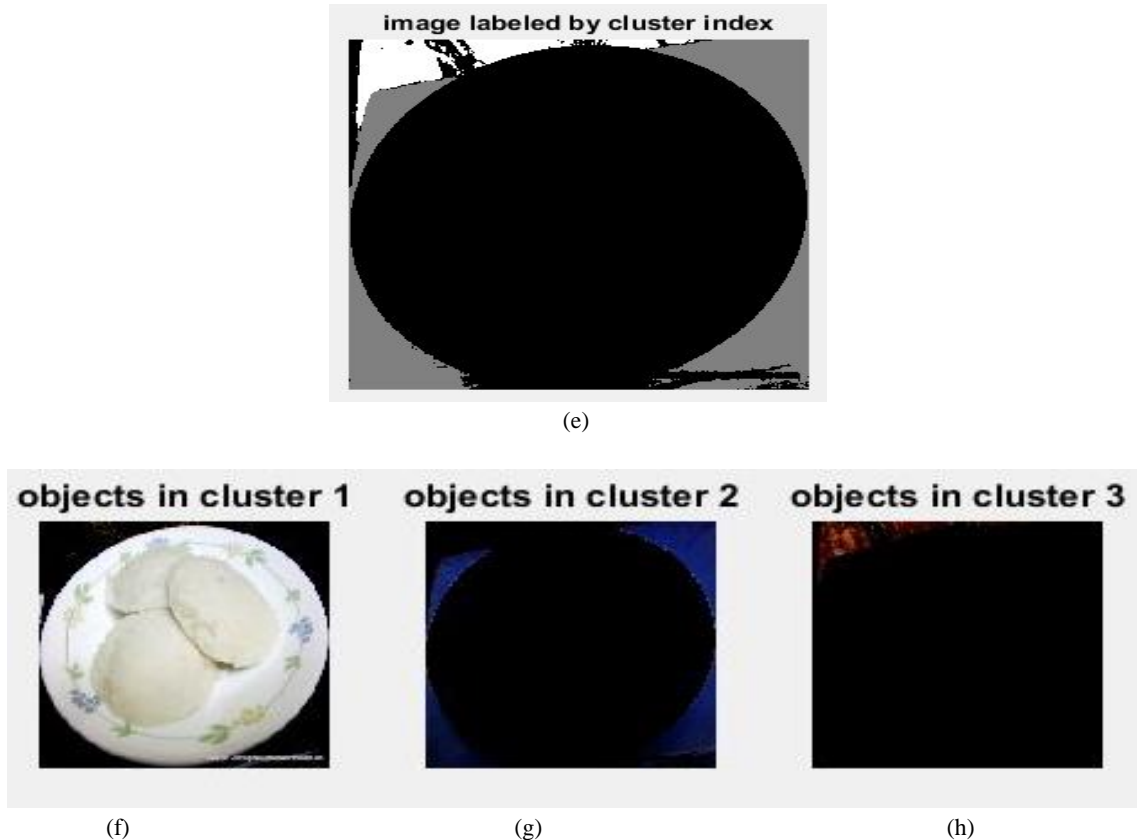


Fig. 9. (a) Source image, (b) Resized image, (c) Mean shift filtered image, (d) CIELAB converted image, (e) Image labeled by cluster index, (f)-(h) Cluster images.

Table 1. Performance Analysis of proposed method

PARAMETERS	SIMULATION RESULTS
SVM classification (%)	93.25%
Detection Time(sec)	5.02sec

Table 2. Performance comparison of proposed method

Author	Classification Rate (%)	Detection Time (Sec)
Paritosh Pandey et al.	86%	10
S.Zhang et al.	82%	12
A.Myers et al.	80%	13

VI. CONCLUSION

In this paper, a novel system for the automatic segmentation and recognition of multi-food images towards Diabetes patient’s health and calorie management has been proposed. The food images have been segmented by region growing algorithm and classified using SVM classifier. The provided results prove the effectiveness of the methods despite the challenges of the problem.

REFERENCES

[1] Y. Xu, X. Yang, H. Ling, and H. Ji, “A New Texture Descriptor Using Multifractal Analysis in Multi-orientation Wavelet Pyramid,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, 2010.

[2] J. Vehel, P. Mignot, and J. Merriot, “Multifractals, Texture, and Image Analysis,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 661- 664, 1992.

[3] K. Falconer, Fractal Geometry: Mathematical Foundations and Applications, Wiley, England, 1990.

[4] M. Varma, and R. Garg, “Locally invariant fractal features for statistical texture classification,” Proceedings of the International Conference on Computer Vision, vol. 1, pp.1-8, 2007.

[5] Eatly, “Eat smart (snap a photo of your meal and get health ratings),” 2017. [Online]. Available: <https://itunes.apple.com/us/app/eatlyeat-smart-snap-photo/id661113749>.

[6] A. Myers et al., “Im2calories: Towards an automated mobile vision food diary,” in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 1233–1241.

- [7] S. Sasano, X. H. Han, and Y. W. Chen, "Food recognition by combined bags of color features and texture features," in Proc. 9th Int. Congr. Image Signal Process., Bio Med. Eng. Informat., Oct. 2016, pp. 815–819.
- [8] C. Pham and T. N. T. Thanh, "Fresh food recognition using feature fusion," in Proc. Int. Conf. Adv. Technol. Commun., Oct. 2014, pp. 298–302.
- [9] N. Martinel, C. Piciarelli, and C. Micheloni, "A supervised extreme learning committee for food recognition," *Comput. Vis. Image Understanding*, vol. 148, pp. 67–86, 2016.
- [10] H. Hoashi, T. Joutou, and K. Yanai, "Image recognition of 85 food categories by feature fusion," in Proc. IEEE Int. Symp. Multimedia, Dec. 2010, pp. 296–301.
- [11] T. Joutou and K. Yanai, "A food recognition system with multiple kernel learning," in Proc. IEEE Int. Conf. Image Process., Nov. 2009, pp. 285–288.
- [12] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [13] S. Liu, D. He, and X. Liang, "An improved hybrid model for automatic salient region detection," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 207–210, Apr. 2012.