International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 4 (2015) pp. 9167-9174

http://www.ripublication.com

Large Scale Learning For Food Image Classification

E. Nagapavani
Sathyabama University,e.n.pavani14@gmail.com
CH. Manvitha
Sathyabama University,ch.manvi@gmail.com
Ms. J. Lydia Jeba
Assistant Professor, Faculty of Computing,Sathyabama University,lydiajeba@gmail.com

Abstract

Many large databases are available with huge number of images. A proper classification of images is in need to organize the images in an efficient manner. For that we need a classifier system which analyses the image in a correct way and identifies the properties of it accurately. The Bag of Features model analyses the image in an efficient manner and classifies the image into the class to which it belongs. It converts the image into HSV color space modeland then extracts the key points from it. Later we need to apply descriptors on it. Then on extracting features it classifies the image. RGB values of the key points should be calculated and maintained in a matrix to compare with test image values. The accuracy of the proposed system increases due to this comparison.

Key Terms: Image Classification, Features Extraction, Spatial Relationship, Quantization.

Nomenclature:

X x co-ordinate of key point
Y y co-ordinate of key point

Introduction

Large scale training dataset contains huge number of images to classify. In that case everyone's choice is to be arranged, that is related separate datasets in the database itself. The problem comes during the time of classification of image. In order to classify accurately our testing image should be very similar to the image in the database. In that case the result is accurate and the image can be identified correctly. Many people facing the same problem because even the lighting conditions, angle,

clarity should be similar in the testing image to that of training image. A solution was provided to this by specifying that each and every transformation of the image should be placed in the database.But problem arises in the maintenance of such a large database.

Many applications are available at present to solve this but with no accuracy parameter. Even diabetes patients need the image analysis to check their carbohydrate intake amount. A large number of approaches were available for image classification with only a few parameters similar between the training and testing images like spatial information, pixel information, shape, geometry etc. Several techniques were proposed like ANN, SVM, decision tree, fuzzy measure etc. All these approaches and techniques classifies the images but with less accuracy.

A famous Bag of Words model is currently in use for image classification. But it loses the accuracy due to the method of classification it follows. It also uses document classification that does not consider the grammar and words order ,but considers how frequently the words are occurring in the image to be classified. Hence it faces the problem in accuracy due to improper matching of images. Bag of Features model solved the issues of Bag of Words. It analyses each feature of the image and matches the testing image very accurately. Here image is classified based on RGB values on the key points extracted from the image and descriptors can be applied and features can be extracted, thus the accurate classification is possible in this model. The present study uses matlab high level computing language for image classification. It solves technical issues faster than other available languages like c, c ++. One more advantage in matlab is that it gives an interactive application developing environment. This is faster because it contains many tools, inbuilt math functions etc. It analyses the images and helps to build the applications. A visual dataset which this study contains existing 5000 images and added 8 more classes of images. Nearly 2500 images can be added in these 8 classes. Each image should be analysed.

Till now only few particular classes or categories of images are taken for classification and identification purpose. Bag of features model has two phases . One is training and the other one is testing phase. Training is to collect the all possible comparative features of the images to be classified and cluster them into classes. Hence a k- means clustering algorithm is required in this scenario. In k- means clustering similar images are kept under one category using some similarity measurement. Cluster is nothing but grouping of images which are closely related in features. During testing the image features should be extracted and these are compared to the training set image features. Thus accurate identification of image is possible and is easy to handle a database which contains number of datasets of different classes and categories.

For example, we can say a database which contains a dataset of cancer patient food images and diabetes patient food images etc. In any type of similar image even with some variations, the image gets classified into appropriate class with great accuracy. In Bag of Words, frequency was taken as a measure. In the proposed work features are taken as measure and the space relationship plays a major role.

Related Work

Bag of words was used before to classify the documents based on the count of occurrence vector. Later it is being introduced in image classification by treating image features as words. For that image should be treated as word document. Local patches should be extracted from the image and keep it in a bag. Later descriptor is applied to convert patches into numerical vectors. But one defect in this method is that, it does not follow any order of occurrence. Finally code word should be produced. In document classification it does not even follow grammar and order of words. Diabetes patients faces many problems to count amount of CHO intake in their food. It is a better solution if there exists an application to analyse the image and generate its properties. For this accurate results should be needed.

Even if small variations are present in the image ,it does not identify the object in the image. The variations may include illumination, scale etc. This causes noise in the image. If the noise exists, it cannot identify the object in the image. Hence there should be a database with the images of the same object with all possible variations. This causes overhead of database size. In the existing database 11 classes of images are present. Totally 5000 images are available in those 11 classes.

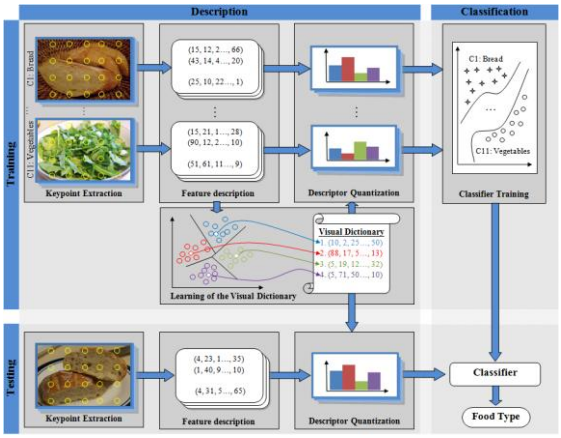
Experimental Procedure:

The current method contains 2 different stages 1) Training and 2) Testing.

In both the stages 3 methods are common i) key point extraction ii) feature description iii) descriptor quantization. In this, the proposed block diagram is taken from [1] with few changes using SIFT key point extraction method. Here SIFT is preferred of all key point extraction algorithms because of its invariant scale extraction. Mainly to increase the proposed system accuracy the SIFT method is being selected for recognition of images according to the features extracted. SIFT is better because it extracts the key points which are strong enough to provide the image features, even if there are some changes in noise, illumination, scale.

Three parameters are considered in each key point a) radius b) centre coordinates c) angle. It detects the key points and compose a circle over it. Key points are explored at numerous positions and angles. Hence it is invariant to any rescaling, transition of the image.A training set of images should be taken and the features of each image should be extracted by using SIFT algorithm. The images along with the details should be maintained in a database. Here part of training phase gets completed.

In testing phase, the test image should be taken and the features should be extracted by using the SIFT descriptor. Then the features of the test image should be compared with the features in the database. Maximum similarities should be compared from which, we can identify the class of the test image. For clustering of database, k-means algorithm is needed. With some similarity measurement, clusters should be formed with similar images. These are called classes. In the proposed system,more number of classes and our local food images can also be added to the database. Thus the image gets classified in Bag of Features model.



Block diagram

Implementation:

Dataset Collection

In this module there is a need to construct a huge database which contains images of all food types. The overall goal of the task is to collect the images of different food types as a training set along with the peculiarities. This causes amid testing stage to match the features of the image with the training set.

Key Point Extraction:

In SIFT, key points are extracted along the high contrast regions. It collects the key points in such a regions , because even if the scale changes or any other property of the image changes it could identify the picture legitimately. This strategy obliges less number of examples. SIFT class or function should be applied on HSV image to extract key points. It gives the details like key point descriptor, orientation, magnitude, location etc. As key points are minute a circle should be drawn over it.

Feature Description:

After the key points got extracted a descriptor should be applied on it. Matlab, vl2 toolbox gives many feature description functions . Points around the key point to extract more features can also be considered.

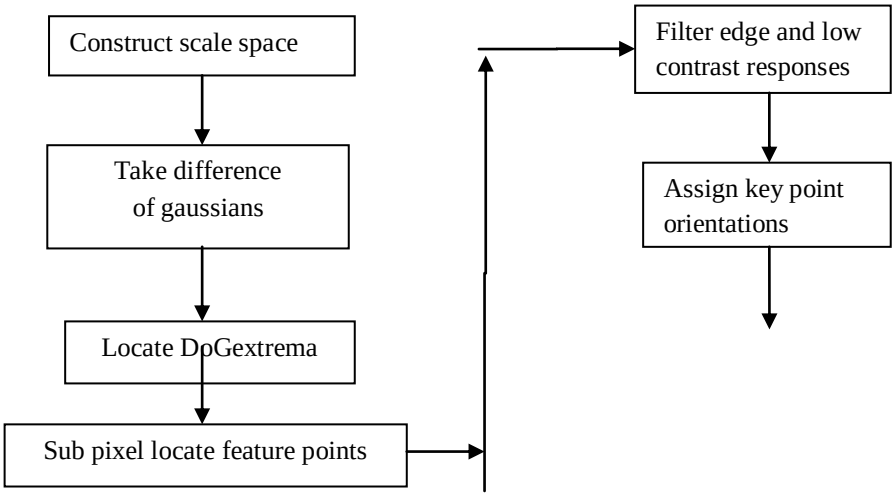
Descriptor Quantization:

In this method descriptor values have to be plotted for each h, s, v images. A histogram should be generated for each image. Histogram contains bins of values. Final result should be how many 1’s and 0’s in each part of image. These are global values and should be placed in a training set. Later it should be used for testing the image.

Technique Used

K-Means Clustering Algorithm:
A clustering algorithm finds the objects which are similar in some measure and group them in a class. This algorithm even finds the center of a group. It divides n observations into k clusters with some nearest mean. It divides the images into classes of nearest featured images.

Sift Algorithm:
This is the only method available for matching images even if they are differ in scale, rotation, illumination, view point etc. If everything is same it can be identified easily by using corner detectors.



Scale space:
$L(x,y,\sigma)=G(x,y,\sigma) * I(x,y)$

where $G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$

Difference of Gaussians:

$D(x,y,\sigma) = \left(G(x,y,k\sigma) - G(x,y,\sigma)\right) * I(x,y)$

$= L(x,y,k\sigma) - L(x,y,\sigma)$

- Scale space is constructed for ensuring the invariance of scale.
- Key points are located at difference of gaussians.
- Bad key points should be eliminated to increase the performance of the algorithm.
- Bad key points indicate low contrast regions.
- Orientation also should be calculated for each key point, so that it will be orientation invariant.

Theory:

To overcome all the problems in existing methodology a Bag of Features model is proposed. It considers spatial relationship instead of count of occurrence. A best descriptor called SIFT can be applied on key points to overcome the variations like scale, illumination etc. Even if the image has undergone any change, it can identify the object in the image accurately. The methods used in the proposed system are described below.

Results and Discussions

The training phase result would be the tested images along with their features. These should be maintained in a database. While in testing phase, each image should be tested and that image features should be checked with the database values.

In this way images can be classified very quickly and in an accurate manner using this model.

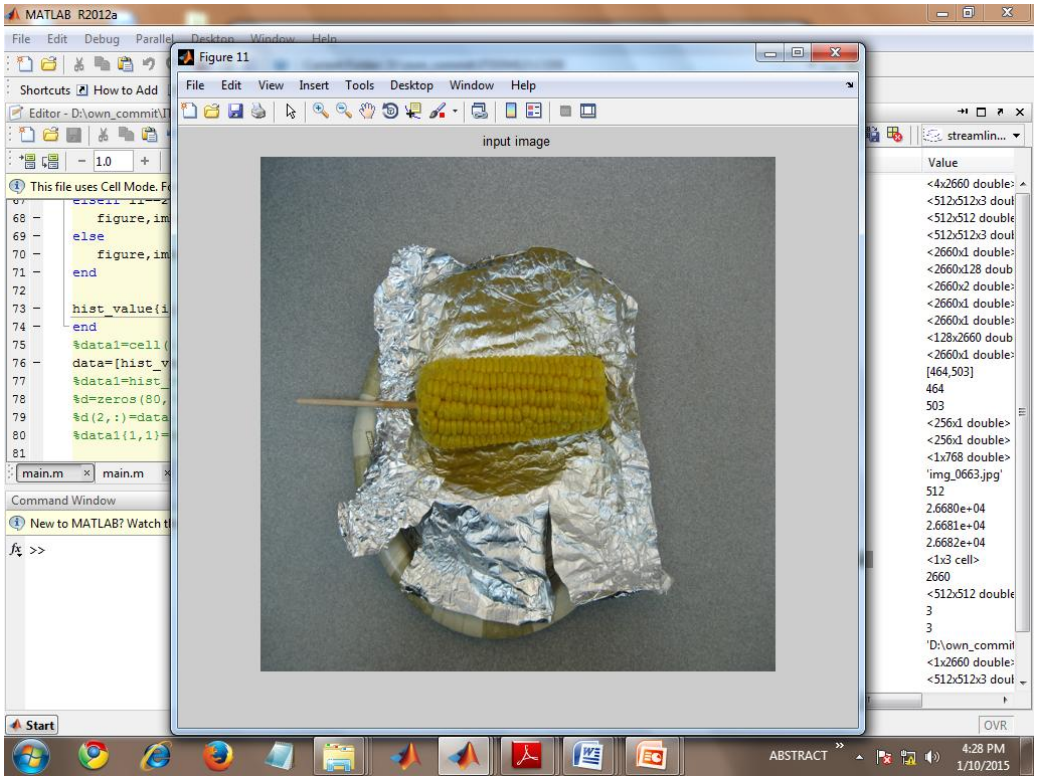


Figure 2: shows the input image taken

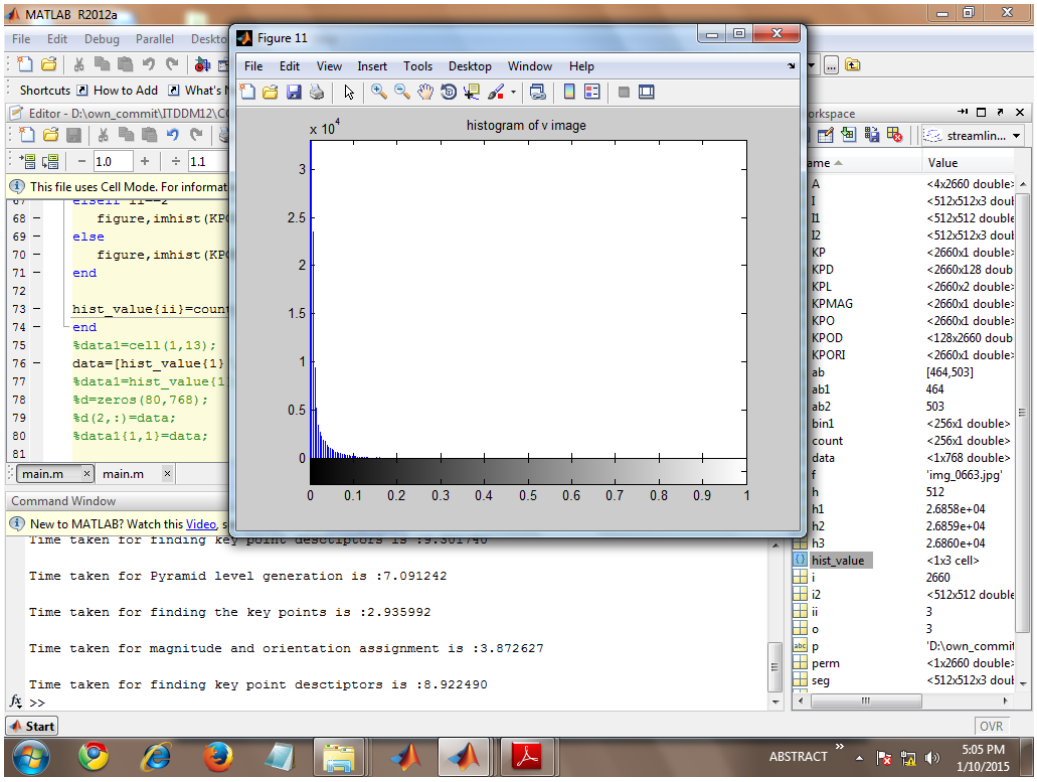


Figure 3: shows the output values of the image which are to be saved in a database

Conclusion

The proposed work gives more accuracy compare to existing one. It is easy to implement this in matlab. The descriptor is better than any other descriptors as it needs less samples and can identify the images even in many variations. This also helps the diabetes patients to analyse the CHO count on their own. This reduces overhead of database. Maximum similarities should be verified between the images, so that it can identify the object exactly.

References

[1] Marios M.; Gianola, Lauro; Scarnato, Luca; Diem, Peter; Mougiakakou, Stavroula G. (2014). *A Food Recognition System for Diabetic Patients based on an Optimized Bag of Features Model.* IEEE Journal of Biomedical and Health Informatics, 18(4), pp. 1261-2194. Institute of Electrical and Electronics Engineers 10.1109/JBHI.2014.2308928

[2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110.

[3] K. E. A. Van de Sande, T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 32, no. 9, pp. 1582–1596, Sep. 2010.

[4] T. Joutou and K. Yanai, "A food image recognition system with multiple kernel learning," in *16th IEEE International Conference on Image Processing (ICIP)*, pp. 285–288, 2009.

[5] M. Graff, T. Gross, S. Juth, J. Charlson, "How well are individuals on intensive insulin therapy counting carbohydrates?," *Diabetes Research and Clinical Practice*, vol. 50, sup. 1, pp. 238-239, 2000.