DIGITAL IMAGE PROCESSING
26 NOV 2018

PROJECT REPORT

-HIMANSH SHEORAN
-LOVISH NARANG
-AYUSH ANAND

PROBLEM STATEMENT:

Given pictures of large Indian meal spreads, identify items (label and segments).

MAIN GOALS:

Need to successfully segment out regions in an image consisting of Indian Foods using some sophisticated image segmentation techniques. Then train our model to so that it can recognise the label of any given image of an Indian meal spread.

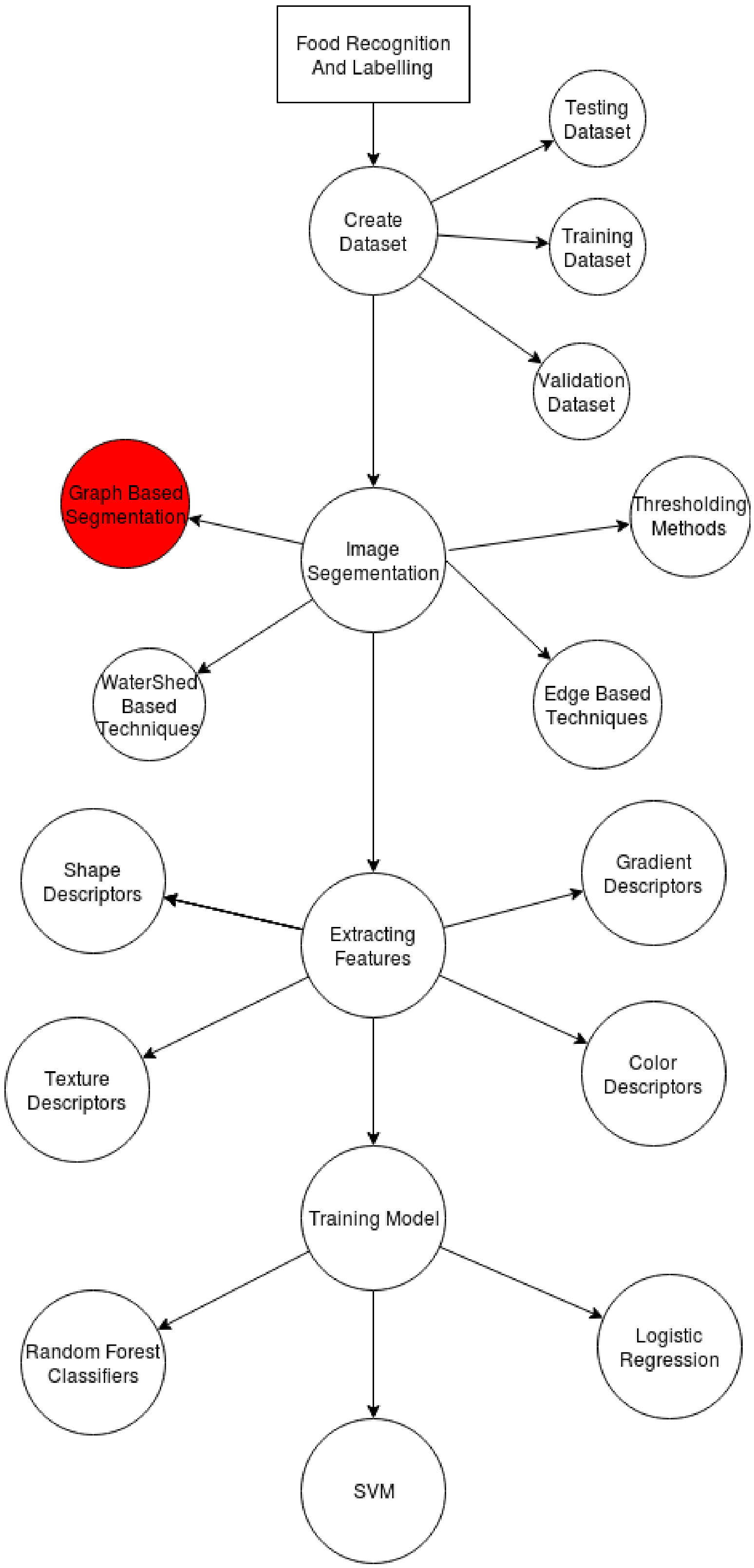
MOTIVATION:

Automatic food image recognition systems are alleviating the process of food-intake estimation and dietary assessment. However, due to the nature of food images, their recognition is a particularly challenging task.

As people are becoming increasingly aware of the importance of a healthy diet, the need for automatic food and drink recognition systems has arisen. Not only can such systems provide the automatic recognition of food and drink items, but they can also enable an estimation of their nutritional values, making them especially useful for dietary assessment and planning, which is applicable for patients with different dietary restrictions, as well as for healthy individuals by preventing nutrition-related conditions.

The problem of food and drink image detection and recognition is challenging due to the nature of food and drink items. Foods are typically deformable objects, which makes the process of defining their structure difficult. Furthermore, some food types can have a high intra-class (similar foods look very different) and low inter-class (different foods look very similar) variance, making the process of specifying the food type even more challenging. The issue with drink recognition is that there is only a limited amount of information that can be gained using images of drink items; an example of such information is the drink’s color, whether the drink is well-lit, and the drink’s density. All of these obstacles make food and drink image detection and recognition a particularly challenging computer vision problem.

PROJECT OVERVIEW:



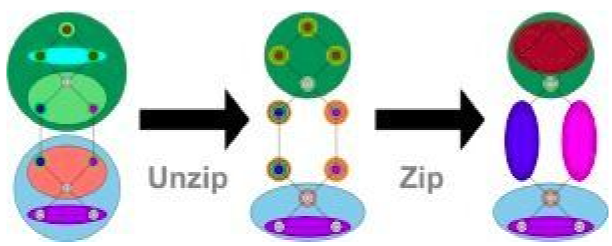
Food Recognition And Labelling
Create Dataset
Testing Dataset
Training Dataset
Validation Dataset
Image Segementation
Graph Based Segmentation
Thresholding Methods
WaterShed Based Techniques
Edge Based Techniques
Extracting Features
Shape Descriptors
Gradient Descriptors
Texture Descriptors
Color Descriptors
Training Model
Random Forest Classifiers
Logistic Regression
SVM

VARIOUS METHODS OF SOLVING SEGMENTATION PROBLEM

Image segmentation is the techniques are used to partition an image into meaningful parts have similar features and properties.The aim of segmentation is simplification i.e. representing an image into meaningful and easily analyzable way. Image segmentation is the first step in image analysis.

The main techniques to Image Segmentation are: Structural Segmentation Techniques Stochastic Segmentation Techniques Hybrid Techniques Different Methods which follow the given above techniques are Thresholding Method Edge Based Segmentation Method Region Based Segmentation Method Clustering Based Segmentation Method Watershed Based Methods Partial Differential Equation Based Segmentation Method

The type of segmentation done in this model is graph based segmentation.



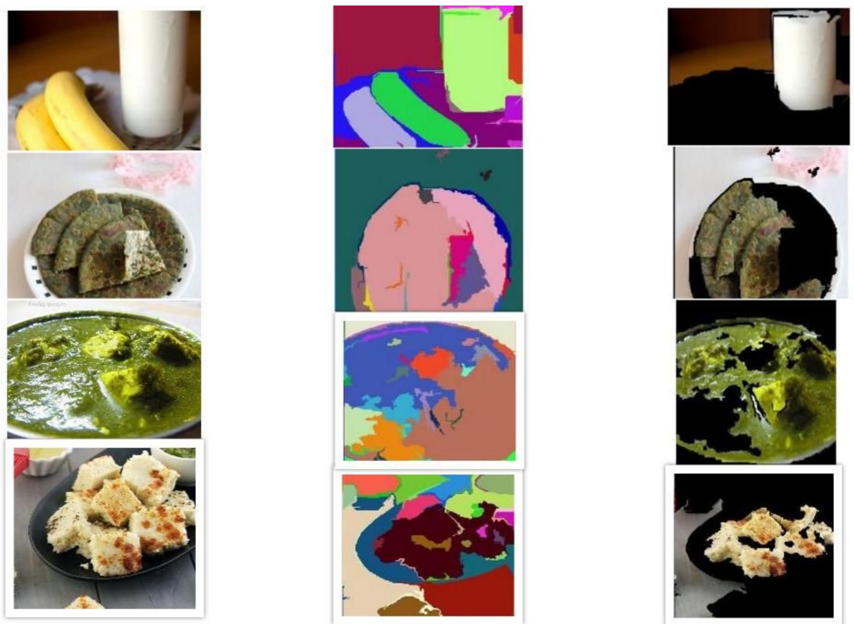
GRAPH BASED SEGMENTATION:

Graph partitioning methods are an effective tools for image segmentation since they model the impact of pixel neighborhoods on a given cluster of pixels or pixel, under the assumption of homogeneity in images. In these methods, the image is modeled as a weighted, undirected graph. Usually a pixel or a group of pixels are associated with nodes and edge weights define the (dis)similarity between the neighborhood pixels. The graph (image) is then partitioned according to a criterion designed to model "good" clusters. Each partition of the nodes (pixels) output from these algorithms are considered an object segment in the image. Some popular algorithms of this category are normalized cuts, random walker, minimum cut, isoperimetric partitioning, minimum spanning tree-based segmentation, and segmentation-based object categorization.

FEATURE EXTRACTION:

In machine learning, pattern recognition and in image processing, **feature extraction** starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is a dimensionality reduction process, where an initial set of raw variables is reduced to more manageable groups (features) for processing, while still accurately and completely describing the original data set

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.



Original Image Connected Components Segmented Image

DIFFERENT TYPES OF FEATURES:

Low-level

- Edge detection
- Corner detection
- Blob detection
- Ridge detection
- Scale-invariant feature transform

Shape based

- Thresholding
- Blob extraction
- Template matching
- Hough transform

Curvature

- Edge direction
- Changing intensity
- Autocorrelation.

Image motion

- Motion detection
- Area based
- Differential approach
- Optical flow

FEATURE EXTRACTION:

●HU MOMENTS (SHAPE DESCRIPTOR) :

In image processing, computer vision and related fields, an image moment is a certain
particular weighted average (moment) of the image pixels' intensities, or a function of
such moments, usually chosen to have some attractive property or interpretation.Image
moments are useful to describe objects after segmentation. Simple properties of the
image which are found via image moments include area (or total intensity), its centroid,
and information about its orientation.
Hu Moments are normally extracted from the silhouette or outline of an object in an
image. By describing the silhouette or outline of an object, we are able to extract a
shape feature vector (i.e. a list of numbers) to represent the shape of the object.
We can then compare two feature vectors using a similarity metric or distance function
to determine how “similar” the shapes are.

● **HARALICK (TEXTURE DESCRIPTOR):**

Haralick's texture features were calculated to describe the texture of the image.
The basis for these features is the gray-level co-occurrence matrix.
This matrix is square with dimension N g , where N g is the number of gray levels in
the image. Element [i,j] of the matrix is generated by counting the number of
times a pixel with value i is adjacent to a pixel with value j and then dividing the
entire matrix by the total number of such comparisons made.
Each entry is therefore considered to be the probability that a pixel with value i
will be found adjacent to a pixel of value j.

HISTOGRAMS (COLOR DESCRIPTOR):

Color can be considered a very crucial descriptor for our model. Each object has a
distinct color which is specific to all the instances belonging to one class.
Therefore it is very important to store the colors of the segmented out part of the
image as it can help the classifier to recognize an image by analyzing the color of
the largest component of the image.

HOG (GRADIENT DESCRIPTOR):

The histogram of oriented gradients (HOG) is a feature descriptor used in
computer vision and image processing for the purpose of object detection. The
technique counts occurrences of gradient orientation in localized portions of an
image. This method is similar to that of edge orientation histograms,
scale-invariant feature transform descriptors, and shape contexts, but differs in

that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

PROBLEMS FACED:

The four main problem that will arise in this project is to :

1) Find an algorithm which can successfully segment the required part which we require, out of the image so that our training can be made more efficient.
2) Implementing efficient feature descriptors which can extract out features from the segmented out part.

3)Training the model on the basis of the features extracted using different Multi-Class classifiers and finding out respective accuracies.

4) Prevent overfitting.

GLIMPSE OF DATASET:

- Aloo Gobhi
- Biryani
- Butter Paneer
- Chapati
- Dhokla
- Dosa
- Gulab Jamun
- Ladoo
- Lassi
- Palak Paneer

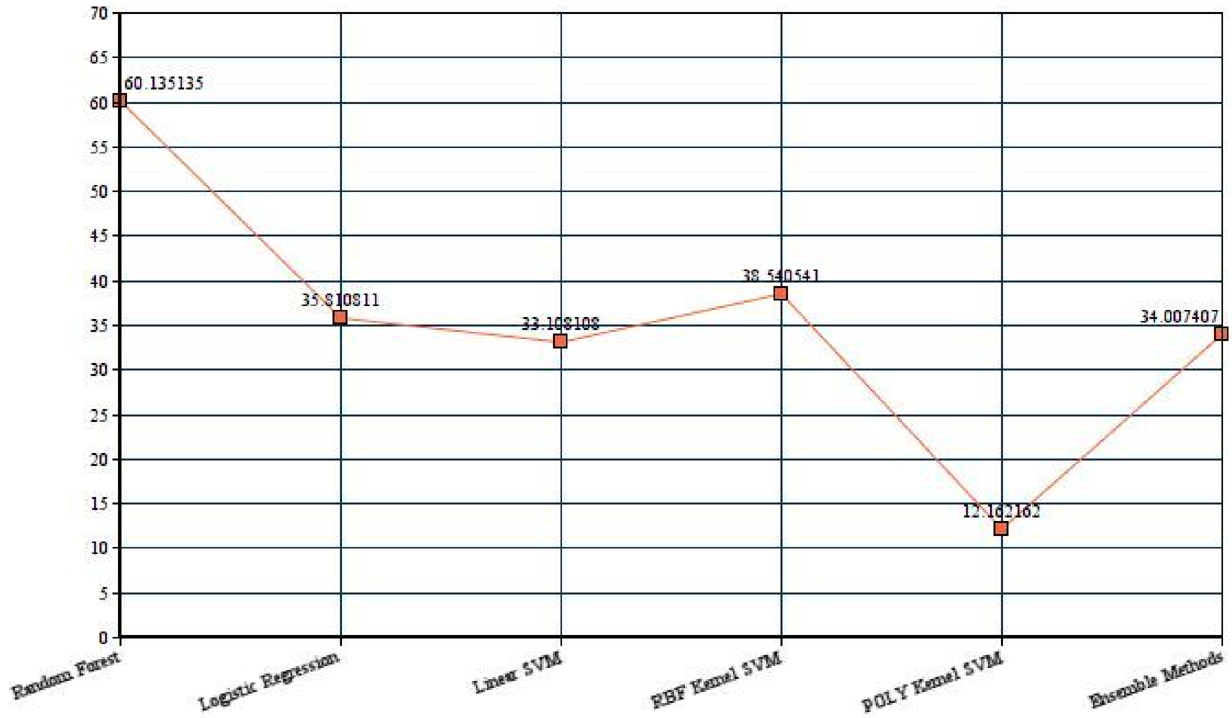
MAIN STEPS FOLLOWED:

Create Training And Testing Dataset of all 10 classes with 300 Images in Training
Set And 100 Images in Testing Images.
Performing Segmentation using a graph based technique.
Extracting the following features out of the segmented part:
○ Shape Descriptor -> HU Moments
○ Texture Descriptor -> Haralick
○ Color Descriptor -> Histograms
○ Gradient Descriptor -> HOG
Scaling and appending all features of an image together.
Training the model using different Multi-Class Classifiers and then testing it on
new samples to report respective accuracies
Tuning the model to predict better.

CLASSIFIERS USED:

The model trains using 5 different classifiers:

-Random Forest Classifier
-Linear SVM
-RBF Kernel SVM
-POLY Kernel SVm
-Logistic Regression
-Ensemble Methods(Logistic Regression,Decision Tree,Linear SVM)



FINAL RESULTS

EXAMPLES OF CORRECTLY CLASSIFIED OBJECTS:

EXAMPLES OF WRONGLY CLASSIFIED OBJECTS:

FUTURE GOALS:

1) Creating a more robust model capable of deciding among larger dataset with
2) greater number of classes Implementing neural network and test the model on it to compare with the classical machine learning algorithms
3) Implementing an even better type of feature descriptor more specific to a particular class so as to increase the accuracy of that particular class.
4) Implementing a bounding box object detector which correctly segments out the `required pixels out of an image.

GITHUB LINK https://github.com/sheoranhimansh/DIP-Project

TASK ASSIGNEMENT:

DATASET COLLECTION -> AYUSH ANAND
DATASET SEGMENTATION -> HIMANSH SHEORAN
DATASET FEATURE EXTRACTION -> LOVISH NARANG
SCALING UP FEATURES AND SETTING UP THE MODEL -> AYUSH ANAND
TRAINING THE DATA ON A CLASSIFIER -> LOVISH NARANG
IMPROVING THE MODEL ALONG WITH TUNING OF HYPERPARAMETERS -> HIMANSH SHEORAN
REPORT MAKING -> HIMANSH SHEORAN
TESTING DATASET -> AYUSH ANAND
PRESENTATION COMPLETION -> LOVISH NARANG

Each member has given almost equal time to the project.

ACKNOWLEDGEMENT:

1) https://github.com/stratospark/food-101-keras
2) https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwj31_Pm5_HeAhVEqI8KHToDATsQFjACegQICBAB&url=http%3A%2F%2Fbennycheung.github.io%2Fyolo-for-real-time-food-detection&usg=AOvVaw0FF3gtcJJYApFYWQKvGS3K
3) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5537777/
4) https://en.wikipedia.org/wiki/Image_segmentation#Graph_partitioning_methods
5) https://www.pyimagesearch.com/2017/09/18/real-time-object-detection-with-deep-learning-and-opencv/
6) https://github.com/igorrendulic/food_recognition_with_calorie_mama