**INDIAN FOOD IMAGE SEGMENTATION AND CLASSIFICATION**

**A PROJECT COMPONENT REPORT**

***Submitted by***

|  |  |
| --- | --- |
| **AISWERYAA R** | **(Reg. No. 201904007)** |
| **MAHASIVAPRIYA B** | **(Reg. No. 201904083)** |

***for the Theory Cum Project Component***

***of***

**19CS693 – DIGITAL IMAGING**

***during***

***VI Semester – 2021 – 2022***

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**MEPCO SCHLENK ENGINEERING COLLEGE, SIVAKASI**

**(An Autonomous Institution affiliated to Anna University Chennai)**

**April 2022**

**MEPCO SCHLENK ENGINEERING COLLEGE, SIVAKASI**

**(An Autonomous Institution affiliated to Anna University Chennai)**

**Department of Computer Science and Engineering**

**BONAFIDE CERTIFICATE**

Certified that this project component report titled **INDIAN FOOD IMAGE SEGMENTATION AND CLASSIFICATION** is the bonafide work of **R. AISWERYAA (Reg. No. 201904007),** and **B. MAHASIVAPRIYA (Reg. No. 201904083)** who carried out this work under my guidance for the Theory cum Project Component course **“19CS693 – DIGITAL IMAGING”** during the sixth semester**.**

|  |  |
| --- | --- |
| **Mr.B.LAKSHMANAN,** M.Tech.,(Ph.D).  Assistant Professor (Sr. Grade)  Course Instructor  Department of Computer Science & Engg, Mepco Schlenk Engineering College,  Sivakasi. | **Dr. J. RAJA SEKAR,** M.E., Ph.D.  Professor  Head of the Department  Department of Computer Science & Engg, Mepco Schlenk Engineering College,  Sivakasi. |

Submitted for viva-Voce Examination held at **MEPCO SCHLENK ENGINEERING COLLEGE (Autonomous), SIVAKASI** on **………/……/…. 20..........**

**Internal Examiner 1 Internal Examiner 2**

**ABSTRACT**

Food recognition plays an important role in food choice and intake, which is essential to the health and well-being of humans.  In this paper we provide an overview about food intake monitoring, by focusing on technical aspects and Image Processing works which solve the main involved tasks (i.e., classification, recognition, segmentation, etc.).

Previously, Users documented their intake using a food diary. However, many users now use smartphone applications to document their energy intake. The increase in smartphone usage has led to the increase of well-being applications that are able to facilitate food logging. The objective of our project is to categorize the detected food items fuzzy c-means algorithm. The input food images are segmented and classified using Google Net.

**ACKNOWLEDGEMENT**

First and foremost, we thank the **LORD ALMIGHTY** for his abundant blessings that is showered upon our past, present and future successful endeavors.

We extend our sincere gratitude to our college management and Principal **Dr. S. Arivazhagan M.E., Ph.D.,** for providing sufficient working environment such as systems and library facilities. We also thank him very much for providing us with adequate lab facilities, which enable us to complete our project.

We would like to extend our heartfelt gratitude to **Dr. J. Raja Sekar M.E., Ph.D.,** Professor and Head, Department of Computer Science and Engineering, Mepco Schlenk Engineering College for giving me the golden opportunity to undertake a project of this nature and for his most valuable guidance given at every phase of our work.

We would also like to extend our gratitude and sincere thanks to **Mr.B.Lakshmanan M.Tech., (Ph.D).,** Assistant Professor (Sr. Grade), Department of Computer Science and Engineering, Mepco Schlenk Engineering College for being our Project Mentor. He has put his valuable experience and expertise in directing, suggesting and supporting us throughout the Project to bring out the best.

Our sincere thanks to our revered **faculty members and lab technicians** for their help over this project work.

Last but not least, we extend our indebtedness towards out beloved family and our friends for their support which made the project a successful one.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO.** | **TITLE** | **PAGE  NO.** |
|  | **ABSTRACT** | **ii** |
|  | **LIST OF TABLES** | **vii** |
|  | **LIST OF FIGURES** | **viii** |
|  | **LIST OF ABBREVIATIONS** | **ix** |
| **1** | **INTRODUCTION** | **1** |
|  | 1.1 Motivation | **1** |
|  | 1.2 Objectives | **1** |
|  | 1.3 Outcomes | **1** |
|  | 1.4 Overview | **2** |

|  |  |  |
| --- | --- | --- |
| **2** | **LITERATURE SURVEY** | **3** |
|  | 2.1 Fine-Grained Food Classification Methods on  the UEC FOOD-100 Database | **3** |
|  | 2.2 Food and Ingredient Joint Learning for Fine  Grained recognition | **4** |
|  | 2.3 DeepFood: Food Image Analysis and  Dietary Assessment via Deep Model | **5** |
|  | 2.4 State Recognition of Food Images Using Deep  Features | **6** |
|  | 2.5 Multi-Scale Multi-View Deep Feature  Aggregation for Food Recognition | 7 |
|  | 2.6 Image-Based Food Classification and Volume  Estimation for Dietary Assessment: A Review | 8 |
|  | 2.7 DietCam: Multi-View Food Recognition Using  Multi-Kernel SVM | 9 |
|  | 2.8 Recognition of multiple food images detected by  candidate regions | 10 |
|  | 2.9 DeepFood: Automatic Multi-class classification  of Food Ingredients using Deep Learning | 11 |
|  | 2.10 Exploring Food Detection using CNNs | 12 |
|  | 2.11 Recognition and Classification of fast food images recognition and classification of fast food images | 13 |
| **3** | **SYSTEM STUDY** | 15 |

|  |  |  |
| --- | --- | --- |
| **4** | **SYSTEM DESIGN** | 19 |
|  | 4.1 Food Segmentation | 19 |
|  | 4.2 Food Classification | 20 |
|  | 4.2.1 GoogleNet | 20 |
| **5** | **SYSTEM IMPLEMENTATION** | **23** |
|  | **5.1 Pre-processing steps** | **23** |
|  | 5.1.1 Resized Image | 23 |
|  | 5.1.2 RGB to Gray Scale Image | 23 |
|  | 5.1.3 Image Histogram | 23 |
|  | 5.1.4 Gray to Black and White Image | 23 |
|  | **5.2 GLCM Feature Extraction** | **23** |
|  | 5.2.1 Contrast | 24 |
|  | 5.2.2 Correlation | 24 |
|  | 5.2.3 Energy | 24 |
|  | 5.2.4 Homogeneity | 24 |
|  | 5.2.5 Mean | 24 |
|  | 5.2.6 Standard Deviation | 24 |
|  | 5.2.7 Entropy | 24 |
|  | 5.2.8 IDM | 25 |
|  | 5.2.9 Kurtosis | 25 |
|  | 5.2.10 GLCM Algorithm Steps | 25 |
|  | **5.3 Food Segmentation** | **26** |
|  | 5.3.1 FCM Algorithm Steps | 26 |
|  | **5.4 Food Classification** | **27** |

|  |  |  |
| --- | --- | --- |
| **6** | **RESULTS AND DISCUSSION** | **28** |
|  | 6.1 RGB to Gray Scale Image | 30 |
|  | 6.2 Image histogram | 31 |
|  | 6.3 Gray to Black and White | 31 |
|  | 6.4 Food Image Segmentation using FCM | 32 |
|  | 6.5 Food Image Classification using GoogleNet | 33 |
| **7** | **CONCLUSION AND FUTURE ENHANCEMENTS** | **35** |

|  |  |  |
| --- | --- | --- |
| **APPENDIX –A** | **SYSTEM REQUIREMENTS** | **36** |
| **APPENDIX –B** | **SOURCE CODE** | **37** |
|  | **REFERENCES** | **44** |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **TABLE NO.** | **TABLE CAPTION** | **PAGE NO.** |
| 1. | Table 4.1 GoogleNet Architecture | 22 |
| 2. | Table 6.1 Food Category List | 29 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Caption** | **Page No.** |
| Fig 4.1 | System flow of the Entire System | **19** |
| Fig 4.2 | GoogleNet Architecture | **21** |
| Fig 6.1.1 | Re-sized image of fried rice | **30** |
| Fig 6.1.2 | Re-sized image of idli | **30** |
| Fig 6.2.1 | gray scale image of fried rice | **30** |
| Fig 6.2.2 | gray scale image of idli | **30** |
| Fig 6.3.1 | histogram of fried rice | **31** |
| Fig 6.3.2 | histogram of idli | **31** |
| Fig 6.4.1 | black and white image of fried rice | **31** |
| Fig 6.4.2 | black and white image of idli | **31** |
| Fig 6.5.1 | segmented image of fried rice | **32** |
| Fig 6.5.2 | segmented image of idli | **32** |
| Fig 6.6.1 | input image is identified as fried rice | **33** |
| Fig 6.6.2 | input image is identified as idli | **33** |
| Fig 6.6.3 | input image is identified as butternaan | **33** |
| Fig 6.6.4 | input image is identified as chai | **33** |

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| **ABBREVIATION** | **DESCRIPTION** |
| CNN | Convolutional Neural Network |
| FCM | Fuzzy C Means |
| GoogleNet | Google Network |
| GLCM | Gray Level Co-occurrence Matrix |
| HoG | Histogram of Oriented Gradients |
| IVM | Inverse Difference Moment |
| MSMVFA | Multi-Scale Multi-View Feature Aggregation |
| MKL | Multiple Kernel Learning |
| SVM | Support Vector Machines |

**CHAPTER 1**

**INTRODUCTION**

**1.1MOTIVATION**

Food is an important part of everybody life. It is important to maintain the health of a person to live a happy life. Good health can be maintained based on the diet chart they follow and the physical condition. If health is not maintained then it leads to other problems like Obesity which is increasing worldwide and it is of major concern since it causes serious health problems like heart disease, type -2 diabetes and some cancer. To follow the diet chart people should know what food to eat and what not to eat. They should also know about the calorie of each food items they consume to maintain a good health. With the help of the Computer Vision System, it is possible to classify the food items present in the image.

Recognition of dishes would not only help users effortlessly organize their extensive photo collections but we would also help online photo repositories make their content more accessible. It also helps patients estimate and track their daily calorie intake, outside of any constraining clinical environment.

Humans can identify the food items once they see it. But this is not the case with computers. They have to be trained in order to carry out the process. Machines can be trained using Google Network (GoogleNet). Features are the properties which are used to determine the object. Classification is done using Google Net. Though the process of classification, the food items are recognized, understood and are categorized unto their respective categories.

**1.2 OBJECTIVES**

* To segment the food images using fuzzy c-means algorithm.
* To classify the food images using GoogleNet.

**1.3 OUTCOMES**

* Food image is segmented from the input image.
* Segmented image is classified into their respective food categories.

**1.4 OVERVIEW**

Our project ‘Indian Food Image Segmentation and Classification’ detects the

object(food) using fuzzy c-means. GoogleNet classifier is used for Classification of the food images.

This chapter deals with the importance of GoogleNetwork for classifying the

food items in the images.

Chapter 2 deals with the existing techniques that prevails for segmentation and

classification.

Chapter 3 deals with the detailed description of the proposed system.

Chapter 4 deals with the methods and modules involved in the design of the

proposed system.

Chapter 5 deals with the implementation methods used in various modules of

the proposed system.

Chapter 6 deals with the results obtained from the proposed system and analysis

on the obtained results.

Chapter 7 deals with the conclusion of the work done and the future

enhancement for the system.

**CHAPTER 2**

**LITERATURE SURVEY**

This chapter deals with the literature survey of various existing techniques. This includes techniques used in those implementations and the drawbacks identified in those methodologies.

**2.1 FINE-GRAINED FOOD CLASSIFICATION METHODS ON THE UEC FOOD-100 DATABASE**

This work is proposed by Berker Arslan, Sefer Memi¸s, Elena Battini Sonmez , and Okan Zafer Batur in 2021, introducing the main databases of food items currently available, and reaching the state-of-the-art performance in the best-shot classification experiment of the UEC Food-100 database.This aims to contribute to automatic food recognition by presenting the most common algorithms used for food classification. That is, this article improves the current best-shot performance by 0.44 percentage points, and fixes it to 90.02%. Furthermore, with the best of our knowledge, this is the first article to introduce to the research community comparison of performances of the classification experiment on the UEC Food-100 database averaged over five-trails.

There are only a few number of food image databases, which have different characteristics such as the number of food categories, the total number of images, the type of cuisine, i.e., Western (French, Italian, Turkish, etc), Asian (Japanese, Chinese, Thai, etc), or the fast food (as it may be considered a world-wide type of food), quality and type of the images, i.e., single versus multifood images, and different contexts, i.e., the same dish may contain more than one food, or different comestibles are separated in several dishes on a tray. The diversity present in the available databases makes the comparison a difficult task. This analysis focuses on publicly available databases with more than 60 classes and 1000 images.

It overviews the main algorithms used for food classification, it details the databases of food items currently available and it presents the results of several deep learning algorithms, considering both the best-shot performance as well as the average over five trials.

**ADVANTAGE:**

The advantage of the module in the right is that the intermediate dimensions are

lower; that is, each of the 32 first level’s sub-block compresses the original signal from 256 to 4, each sub-block of the second level uses 3 × 3 filters and outputs a 4-D feature, and each of the 32 third level’s sub-block converts the input signal into a

256 output signals.

**LIMITATION :**

Despite the big importance of food classification systems, the current number of studies and improvements are still too limited. The main drawback is the lack of big and international databases, which are necessary to train the algorithms. Also, new methods can help to improve the performance on available databases.

**2.2 FOOD AND INGREDIENT JOINT LEARNING FOR FINE-GRAINED RECOGNITION**

The detailed classification that provides more specialized and professional attribute information of food, proposed by Chengxu Liu, Yuanzhi Liang, Yao Xue, Xueming Qian in 2020. It is the basic work to realize healthy diet recommendations and cooking instructions, nutrition intake management, and cafeteria self-checkout system. Chinese food lacks structured information, and ingredients composition is an important consideration.

The research on fine-grained food classification and recognition not only has high scientific value but also has high practical application value. Therefore, an excellent method of food and ingredients classification and recognition is very necessary. It aims to distinguish and recognize the categories with similar visual features within a general category. The effectiveness of the method is proved through the comparative experiments. The use of the method get the state-of-the-art performance of ingredients recognition and improved the Micro− F1, Macro− F1, and Accuracy of ingredients to 74.10%, 58.80%, and 34.29%.

**ADVANTAGE :**

Balance focal loss method takes full advantage of multi-label ingredients

information and improves the learning ability of ingredients by BFL.

**LIMITATION :**

The equilibrium effect is best when f p Equilibrium (t)/ f n Equilibrium (t) is approximately equal to five to ten. When the balance term is selected properly, the constraint term can improve the performance of ingredient recognition to a certain extent. On the contrary, the constraint term will further lead to worse performance if the equilibrium term is not appropriate.

**2.3 DEEPFOOD: FOOD IMAGE ANALYSIS AND DIETARY ASSESMENT VIA DEEP MODEL**

This work by Xue liu and chenxi huang (2020) proposes three types of most commonly used methods to manually access dietary intake including diet records, 24 hour recall and food frequency questinarie.For diet records, subjects need to record the food and beverage consumed over three consecutive days (two weekdays and one weekend day). Detailed instructions on how to record intake must be provided by trained staff and

the completed records need to be entered into a application such as Nutrition Data System for Research for analysis. By applying 24-hour recall, subjects are asked to report all food/meals consumed in the past 24 hours, which can be done via telephone call or face-to-face interview. The data from subjects are required to be collected and analysed, an interview for details will be conducted by trained staff Subjects using FFQ method are asked to report how frequently certain food and beverage items were consumed over a specific period of time (e.g., 1 year). Most FFQs are available in paper or electronic format listing general questions about everyday diet and cooking practice. Software programs are deployed to calculate nutrient intake by multiplying the reported frequency of each food by the amount of nutrient in each food item.

**ADVANTAGE** :

CNN achieved a very high accuracy of 93.8% on food and non-food item detection. The experimental results on food recognition showed that the proposed CNN solution outperformed all other baseline methods achieved an accuracy of 73.7% for 10 classes.

**LIMITATION :**

Though our proposed model do perform well on different types of datasets, there is still room for the improvement compared with some of the state-of-the-art models.

**2.4 STATE RECOGNITION OF FOOD IMAGES USING DEEP FEATURES**

To state and recognize the food images, Gianluigi Ciocca, Giovanni Micali, and Paolo Nnapoletano proposed this article state recognition of food images using deep features in 2020, a recent topic that is gaining a huge interest in the Computer Vision community. Recently, researchers presented a dataset of food images at different states where unfortunately no information regarding the food category was included.

Computer vision techniques can help to build systems to automatically locate and recognize diverse foods as well as to estimate the food quantity. For example, one may simply take the associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano, a picture of a plate of food using a smartphone and the whole process towards measuring the total calorie in the plate can be achieved by a visual understanding framework. The evaluation pipeline includes a feature extraction module and a classification module (one for each task) based on an SVM classifier with a radial basis function (RBF). The validation set of the dataset is used for the choice of the RBF parameters. Learned features are extracted from several CNNs trained using the dataset.

**ADVANTAGE** :

A picture of a plate of food using a smartphone and the whole process towards measuring the total calorie in the plate can be achieved by a visual understanding framework.

**LIMITATION :**

These features outperform hand-crafted features by a large margin. Moreover, there is no significant advantages in combining hand-crafted features with learned ones. On the overall, it seems that the state recognition problem is more approachable by the CNN-based features than the food classification one. Good results notwithstanding, which need to further investigate the robustness of machine learning methods to the variability of real world foods in images and videos in terms of illumination, scale, point of view, and cluttered scenes.

**2.5 MULTI-SCALE MULTI-VIEW DEEP FEATURE AGGREGATION FOR FOOD DETECTION**

Multi-Scale Multi-View Feature Aggregation (MSMVFA) scheme for food recognition was proposed byShuqiang Jiang , Weiqing Min, Linhu Liu, and Zhengdong Luo in 2020, which can conduct two-level fusion, namely multi-scale fusion for each type of features and multi view aggregation for various types of features with different granularity to produce more robust, discriminative and comprehensive fine-grained representation. Two key technologies are exploited in MSMVFA. First, aggregate high-level semantic features, mid-level attribute features and deep visual features into unified representation. Different types of features describe food images from different granularity. Therefore, the aggregated features can capture semantic information of food images with the greatest probability. Second, unlike general objects, food typically does not exhibit distinctive spatial patterns. In order to solve this, for each type of features, multi-scale patches based feature fusion is utilized to obtain more robust and discriminative representation. such multi-scale feature representation not only contains ones from discriminative image regions, but also is insensitive to geometrical deformation.

**ADVANTAGE** :

The advantage of MSMVF can be derived from twofold. First, MSMVF can obtain different types of deep features under different supervised signals. Category-supervised deep network can provide high-level semantic features while ingredient-supervised deep network can provide fine-grained attribute features.Second, MSMVF can explore discriminative image regions with different scales. Fusing these regional features from the coarse scale to the fine scale contain discriminative information with the greatest

probability.

**LIMITATION :**

Itdoes not provide perfect accuracy for some food categories. We can see that these food categories are very similar in the visual appearance and texture. Even the humans are not easy to distinguish among these food categories. Probably, it need to design more fine-grained visual feature learning methods to classify these food categories.

**2.6 IMAGE-BASED FOOD CLASSIFICATION AND VOLUME ESTIMATION FOR DIETARY ASSESSMENT**

A daily dietary assessment method named 24-hour dietary recall, proposed by Frank Po Wen Lo, Yingnan Sun, Jianing Qiu, and Benny Lo in 2020, has commonly been used in nutritional epidemiology studies to capture detailed information of the food eaten by the participants to help understand their dietary behaviour. This study provides an overview of computing algorithms, mathematical models and methodologies used in the field of image-based dietary assessment. It also provides a comprehensive comparison of the state of the art approaches in food recognition and volume/weight estimation in terms of their processing speed, model accuracy, efficiency and constraints. It will be followed by a discussion on deep learning method and its efficacy in dietary assessment.

Food recognition is a crucial part in the dietary assessment process. Only after

recognising the type of food can we further compute the calorie intake and analyse nutritional information. In food recognition, the characteristics of food can be greatly attributed to their surface colours, shapes and texture. Therefore, if a system tries to identify a particular food object, feature descriptors containing those underlying information should be extracted.

**ADVANTAGE** :

The advantage of using deep neural network for food volume estimation is that the scale of the monocular image can be learned from the global cues of the scene without

the need of camera calibration, which means reference objects with known dimension are not required

**LIMITATION :**

3D models cannot be reconstructed and the volume estimation will fail if the food surface does not have distinctive characteristics or texture. Another concern is that stereo-based approach requires users to capture multiple images from different viewing angles before and after eating, which in turn makes this approach very tedious and not suitable to be applied on wearable sensors for long-term health monitoring and data collection.

**2.7 DIETCAM: MULTIVIEW FOOD RECOGNITION USING A MULTIKERNAL SVM**

To address the variation of food appearances, an automatic food classification method. DietCam is proposed by Hongsheng He et.al.DietCam in 2015 consists of 2 major components:

1.Ingredient Detection

2.Food Classification

Food ingredients are detected through a combination of a deformable part-based model and a texture verification model. From the detected ingredients, food categories are classified using a multi-view kernel SVM. The Current part-based object recognition model is improved towards texture-oriented and location-flexible to detect food ingredients.

The goal of ingredient detection is to predict bounding boxes of each food ingredient in the image. A food ingredient detector is developed by combining part-based models and texture models. Part-based models are popular for rigid shaped objects detection and classification, taking into account the shape of each part and their geometric relations. The part-based models, however, cannot be directly applied to food ingredient detection because of the shape and texture variance in food appearances. Therefore

Texture filters are integrated into part-based detectors, where textures are classified in texture filters.

Semantic Texton Forest(STF), an image segmentation and classification technique that guarantees soft labels for each pixel based on their local texture properties is chosen to detect food textures. This is achieved through manually labelled sample images and building decision forest.

The distribution of food ingredients has different geometric patterns due to the intraclass deformation. Therefore, the deformation is modelled in high level using the distribution of the texture parts. The distribution of the texture parts is obtained by analysing the arrangement of features from the food image.

**ADVANTAGE :**

DietCam achieved the best performance in terms of precision and recall. The recognition precision of all the methods dropped as the the number of food items in the images increased.

**LIMITATION :**

DietCam presents promising performance as compared with commonly used food classification method. But it needs many computing resources for classification especially for the ingredient detection part.

**2.8 RECOGNITION OF MUTIPLE FOOD IMAGES BY DETECTING BY CANDITATE REGIONS**

This work by Yuji Matsuda et. al.(2012) proposed a 2-step method to recognize multiple food images by detecting candidate regions and classify them with various kind of features. Initially, the candidate regions are detected from the input image; feature vectors are generated and the image is classified.

Candidate regions are detected in the image through 4 methods, namely:

1. Whole Image- In this method, the region is detected by assuming that the image contains a single food item.
2. Deformable Part Model (DPM)- In DPM, an approach called sliding window is adopted to detect object regions.
3. Circle detector- It detect the region of dishes by extracting circular contours from the input image.
4. Region segmentation- It segments an image into several pieces of regions. JSEG algorithm, which detects an image by color space quantization and color class map is used for region segmentation

As a next step, the candidate region detected by the 4 methods are aggregated in the candidate set and are converted into bounding boxes each of circumscribes each detected region. Irrelevant regions are removed from the candidate set to reduce classification cost and noisy candidates. Various kinds of image features are extracted from the bounding boxes.

The extracted image features are described using following methods:

1. Bag-of-features of SIFT and CSIFT- A set of local image points are sampled and visual descriptors are extracted by the Scale Invariant Feature Transform (SIFT) descriptor on each point. CSIFT is also extracted, which is extracted from SIFT from a RGB color space.

2. Spatial Pyramid Representation- To account for the spatial information of the image, spatial pyramid representation is used, which divides the object regions by hierarical grids.

3. Histogram of Oriented Gradients (HoG)- HoG describes the local patterns in the image which is based on the gradient histogram.

4. Gabor texture feature- It represents texture patterns of local regions with several scales and orientations.

After the extraction of feature vectors from each candidate region, evaluation values of the candidate region regarding each of all the given categories are calculated using support vector machines (SVM) which are trained by Multiple Kernel Learning (MKL)

As results, a classification rate of 55.8% is achieved, which improves the baseline result in case of using only DPM by 14.3 points, for a multiple food image dataset.

**LIMITATION:**

Though the proposed model recognizes images with a single food item, it misidentifies images with multiple food items.

**2.9 DEEPFOOD: AUTOMATIC MULTI-CLASS CLASSIFICATION OF FOOD INGREDIENTS USING DEEP LEARNING**

Deep learning techniques have brought significant improvements in the classification of food images. This paper by Lili Pan et. al in 2017 proposed a new framework, called DeepFood which extracts rich and effective features from a dataset of food ingredient images using deep learning and improves the average accuracy of multi-class classification by applying advanced machine learning techniques.

DeepFood framework is evaluated on a multi-class dataset that includes 41 classes of food ingredients and 100 images for each class. First, a set of transfer learning algorithms based on Convolutional Neural Networks (CNNs) are leveraged for deep feature extraction. Then, a multi-class classification algorithm is exploited based on the performance of the classifiers on each deep feature set.

A two-level CNN feature extraction module is utilized in this paper. The raw images are given as input to a pre-trained CNN and then activation vectors are derived from its intermediate layers. CNN vectors are propagated into the upper layers and the extracted vectors are regarded as the image features. The CNN features are extracted from the last output layer of the pre-trained CNN.

Information Gain (IG) is used for the feature selection. IG is very effective and popular approach that evaluates features respectively for category prediction. Specifically, it measures the entropy or uncertainty in a feature set and selects the one with the highest information.

Sequential Minimal Optimization (SMO) algorithm is used for classification task. It is an improved algorithm for training Support Vector Machines (SVM) on classification tasks. In the training phase, multiple SMOs are trained to classify multi-class food ingredients using the training instances and different feature sets. In testing, all the trained SMO models are utilized to predict the label of each testing instance.

Experimental results illustrate the effectiveness of the DeepFood framework for multi-class classification of food ingredients. This model that integrates ResNet deep feature sets, Information Gain (IG) feature selection, and the SMO classifier has shown its supremacy for food ingredients recognition compared to several existing work in this area.

**LIMITATION**:

Though this framework provides results with good accuracy for smaller datasets, it is not suitable for larger datasets.

**2.10 EXPLORING FOOD DETECTION USING CNNS**

This work by Eduardo Aguilar et. al. (2018) has focused his efforts on several areas involved in the visual food analysis such as food detection, food recognition, food localization, and portion estimation. Food detection is explored using GoogleNet because it presents the best result in classification of object and in particular for food detection it also presents good results on multiple datasets with images acquired in different conditions.

In this work food detection model is proposed based on GoogleNet for feature extraction, PCA (Principal Component Analysis) for feature selection and SVM (Support Vector Machines) for classification.

The first step involves training the GoogleNet CNN model. The model is previously trained on ImageNet Dataset. Since GoogleNet is going to be utilized for binary classification (food/non-food), the number of classes in the output layer is changed from 1000 to 2. Then, the Googlenet is finetuned on the last two years until the accuracy of training set stops to increase. The finetuned model is used for feature extraction. A feature vector with 1024 dimensions of the images obtained.

PCA (Principal Component Analysis) is utilized in the next step to reduce the dimesnions of the feature vectors obtained in the previous step. PCA transforms the data to new quardinate system leaving the greatest variance of the images in the first axis. To apply PCA transformation, features are selector based on the kaiser Criterion, which consists of retaining those components with eigen values greater than 1.

The training of the Support Vector Machine (SVM) classifier is made by means of the GridSearchCV strategy on 3-folds using the feature vectors obtained from the previous step. Finally, the trained SVM classifier is used for binary classification (food/non-food).

The proposed model provides an overall accuracy of 97.41%, with the GoogleNet model providing high accuracy on the food detection problem.

**LIMITATION :**

The proposed model is not able to address larger datasets containing a much wider range of dishes and beverages.

**2.11 RECOGNITION AND CLASSIFICATION OF FASTFOOD**

**IMAGES**

This work is by Amatul Bushra Akhi in 2018. Obesity is conceding a great

problem in today’s life. The preeminent reason of obesity is consuming more calories that

we burn which can seriously undermin the quality of life. Researchers says, accurately

accessing diatary intake is an important factor to reduce this risk. To meet this exigency,

researchers have taken some approaches to measure the calorie of a food.

They use use Gabor and color features to represent food items. A multi-label

SVM classifier combined with multi class Adaboost algorithm is used to show that the

new technique can successfully improve the performance of original.

Deep learning gradually becomes a very powerful image recognition

technique, and CNN is the most popular deep learning architecture.

**ADVANTAGE** :

By using Support Vector Machine (SVM) classifier with a trained CNN to extract

and to classify fast food images of ten different classes and achieved accuracy 99.5.

**LIMITATION :**

A raw image contains of certain factors such as noise, climatic conditions, poor resolution and unwanted background for which it is not suitable enough to classification and identification.

**SUMMARY:**

The next chapter will provide a detailed study of the proposed system and explain

how these limitations can be overcome in the proposed system.

**CHAPTER 3**

**SYSTEM STUDY**

This chapter deals with the detailed description of the proposed system. In the last few years, recognition of food items has become a popular research topic due to the availability of a large number of images on the internet and because of the interest of people in social networks. Automatic understanding of food is an important research challenge. One of the first challenges in this field is the discrimination between images containing food versus the others. Other challenging task in food classification is to determine which food items are present in the picture. This becomes more complicated when there are more than one food item in the image.

Automatic food understanding is becoming more and more important to provide services for self-nutrition. User’s feeding habits have to be taken into account to try to combat obesity, which is dramatically increasing also in childhood. Having a daily record of actual intake of food has a crucial importance to provide the user with a personalized Diet and beneficial information on the food intake balance. More specifically, the procedure of measuring energy and nutrients intake requires the record of all food consumed by an individual, the identification of the portion size and the determination of the frequency with which each food is eaten.

Previously dietary management has been carried out through food logging. Food logging is an activity in which the users document their energy intake to monitor their diet. Standard approaches for food intake monitoring demand the user to self-report all the food he/she eats and to recognize and describe quantities. Other methods may include the use of an exercise log book to document physical activities and the duration. Previously, users documented their intake using a food diary. But these are time-consuming process. However, the increase in smartphone usage has led to the increase of well-being applications that are able to facilitate food logging. The novel methods include allowing the user to photograph the food items to determine their category. This can be achieved using computer vision methods.

Indian food image dataset is used for food classification. The dataset consists of testing, training and validation set. Each contain 11 major categories of food items. The number of images in each category varies. First the object is resized from the images and then it is converted into grayscale image, image histogram was plotted and then black and white conversion and then the food items are classified into their respective category.

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do image recognition and imageclassification. CNN image classification takes an input image, processes it and classifies under certain categories. CNN models take the input image and pass it through the series of layers and apply Softmax function to classify an object with probabilistic values based on the number of classes. CNN architecture consist of three layers namely

* Convolutional layer
* Pooling layer
* Fully Connected(FC) layer

**Convolutional Layer**

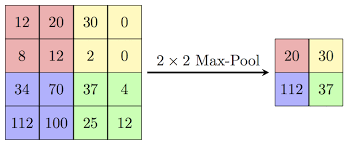
Convolutional layer is the first layer to extract features from the input image.

This layer preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes 2 inputs such as image matrix and filter. Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters.

**Pooling Layer**

The pooling layer takes small rectangular blocks from the convolution layer and

sub samples it to produce single output from the block. The most common approach used in pooling is maxpooling, that is that take the maximum of the block they are pooling.



**Fully Connected Layer**

The high level reasoning in the neural network is done via fully connected layer.

The feature map matrix obtained from the previous layer is flattened into a vector. In this layer, features are combined to create a model. Neurons in a fully connected layer have dull connections to all activations in the previous layer. Finally, activation function is used to classify the output.

The convolution and pooling layers are used to extract features and reduce the

number of parameters from the original images. The fully connected layer is used to

generate the output. The convolution and pooling layers are present alternatively and fully

connected layer is present at the end.

The performance of the neural network can be improved by

increasing in size. This can be done by increasing its depth (number of network layer) as well as its width (number of units in each layer).

**GoogleNet**

GoogleNet consists of 22 layers and 9 inception modules. The introduction Inception modules utilized the concept of using approximation of sparse structure with repeated dense components. The output of inception modules is the best among the modules with each inception. Each layer in GoogleNet is 2 or 3 layers deep. Hence the total number if layers in GoogleNet is 144. The last two layers are used for classification. Features are extracted from the 20th layer which is the Dropout layer. The size of the input is 224 x 224 x 3. Hence the cropped images are resized according to the size of the input layer and then fed into the network. The output of this nx1024 matrix where n is the number of images which are given as input to this layer, The obtained feature vector is stored for classification.

**Classification**

Food classification is done by means of GoogleNet.

**GoogleNet**

Another approach is to use GoogleNet for Classification itself. The training images are converted to the size equal to the size of the input layer in GoogleNet. Then the features from the fully connected layer are obtained and trained using the image to the data store. The number of classes in the output layer is adjusted as per the requirement. After making changes to those layers they are connected to the network and training is started, Then images from the evaluation dataset are given as input to this trained network and the predicted label is obtained.

This chapter gave an overview about the techniques and algorithms used in the project. The next chapter will give a detailed description about the design of the project.

**CHAPTER 4**

**SYSTEM DESIGN**

This chapter deals with the modules use in the project and explains the sequence in which data flows. The flowchart describing the entire flow of the proposed system is shown in the **Figure 4.1**

Image Classification

Food Image Segmentation

Food Portion Recognition

Food Image

GoogleNet

Fuzzy C Means

**Figure 4.1 System Flow of the Entire System**

The modules present in the proposed model are as follows:

1. Food Segmentation
2. Food Classification

**4.1 Food Segmentation (Fuzzy c-means Thresholding)**

FCM clustering is used to partition N objects into C classes. In our method, N is equal to the number of pixels in the image i.e., N=Nx x Ny and C=3 for 3-class FCM clustering. The FCM algorithm uses iterative optimization of an objective function based on a weighted similarity measure between the pixels in the image and each of the c-cluster centers.

The aim of FCM algorithm is to find an optimal fuzzy c-partition by evolving the fuzzy partition matrix U iteratively and computing the cluster centers. In order to achieve this, the algorithm tries to minimize the objective function Q ( by iteratively updating the cluster centers and the membership functions using the following equations. After performing FCM clustering, finally each pixel is assigned to the cluster for which its membership value is maximum. Based on the intensity distribution obtained using histogram of the image, the threshold value is calculated by taking mean of maximum of cluster 1 and minimum of cluster 2 or maximum of cluster 2 and minimum of cluster 3. This method of threshold selection takes into account the intensity distribution in the image. This choice helps in obtaining optimum threshold values for different images obtained under different conditions.

**4.2 Food Classification**

In order to identify the category of food item present in the image, it needs to be classified. Classification is used to analyze the numerical properties of various image features and Organizes data in to Categories. Generally, classification algorithms employ 2 phases of processing: training and testing.

In the initial training phase, characteristic properties of typical image features are

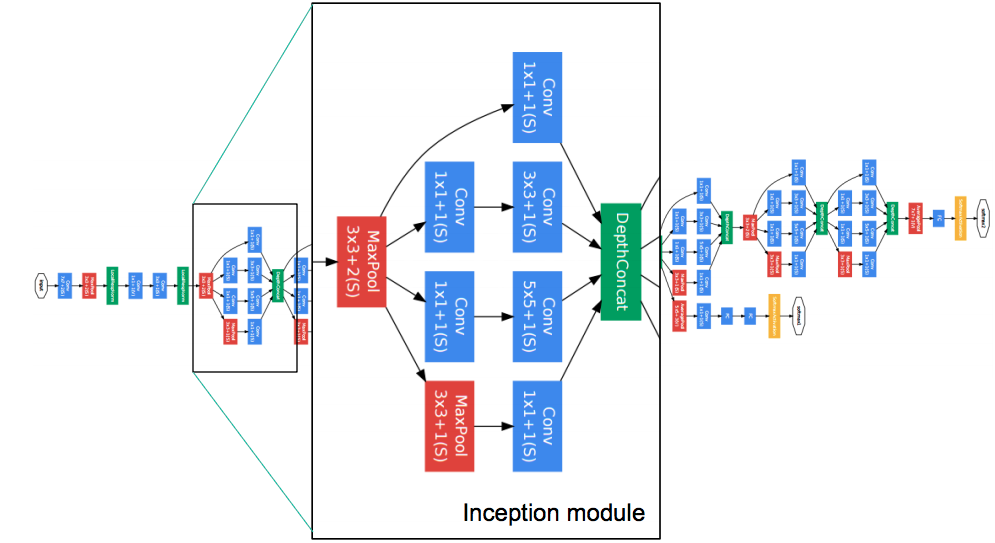
isolated and based on these, a unique description of each classification category that is training classes created. In the subsequent testing phase, these feature space partitions are used to classify the image features.

**4.2.1 Google Net**

Google net is also used for food image classification by training the network with food images. The size of the input layer of GoogleNet is 224x224x3. Therefore, the cropped image must be resized as suitable to be fed to the input layer of GoogleNet. As Indian Food Image Dataset has 20 categories of food items, the number of classes is changed from 1000 to 20 in the output layer of GoogleNet.

The training process is started by feeding the resized cropped images of dimension 224x224x3 to the input layer of GoogleNet. The trained model is used for Classification.The testing image is given as input to the trained GoogleNet and the predicted label is obtained. Table 4.2 describes GoogleNet Architecture.

This chapter has explained the design of the proposed system and how it works. The next chapter will provide details about how this design is implemented to build a classifier that classifies the food images.



**Fig 4.2 GoogleNet Architecture**

**Table 4.2.1 for GoogleNet Architecture**



**CHAPTER 5**

**SYSTEM IMPLEMENTATION**

This chapter deals with the models that are represent in the proposed system. Section 5.1 used the preprocessing steps. Section 5.2 deals with steps for GLCM feature extraction. Section 5.3 deals with food segmentation. Section 5.4 deals with food classification. Section 5.5 deals with accuracy estimation.

**5.1 Preprocessing steps**

Preprocessing steps such as, resized image, RGB to grayscale image, Image Histogram, gray to black and white image.

**5.1.1 Resized image**

Image interpolation occurs when you resize or distort your image from one pixel grid to another. Image resizing is necessary when you need to increase or decrease the total number of pixels, whereas remapping can occur when you are correcting for lens distortion or rotating an image.

**5.1.2 RGB to grayscale image**

To take the average of three colors. Since its an RGB image, so it means that you have add r with g with b and then divide it by 3 to get your desired grayscale image.

Its done in this way,

Grayscale = (R + G + B / 3)

**5.1.3 Image Histogram**

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution at a glance.

**5.1.4 Gray to black and white**

Binarization converts a grayscale image to a black/white image. This transformation is useful in detecting blobs and further reduces the computational complexity.

**5.2 GLCM feature extraction**

Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture feature**s**. The approach has been used in a number of applications, Third and higher order textures consider the relationships among three or more pixels.

A co-occurrence matrix measures the probability of appearance of pairs of pixel values located at a distance in the image. This algorithm is known as GLCM. The matrix defines the probability of joining two pixels, ( , ) that have values i and j with distance d and as an orientation angular.

**Contrast :**

The intensity variation of threshold and its nearest pixel is determined by contrast.

**Correlation :**

It is the analytical form of correlation. It is used to measure the relationship

between the threshold and nearest pixel.

**Energy :**

Homogeneity is calculated by the energy. It is also called as angular second moment or uniformity.

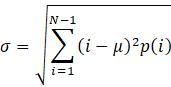
**Mean :**

Mean is defined as the average level of intensity of image.

 …(1)

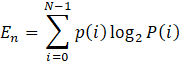
**Standard Deviation :**

The mean value of the pixels and their probability densities are used to measure the standard deviation.

 …(2)

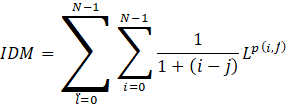
**Entropy :**

The uncertainty in the random variable is measured by the entropy. It depends on the probability density p(i).

 …(3)

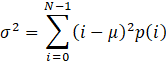
**Inverse Difference Moment (IDM) :**

The local homogeneity of an image is calculated by IDM.

 …(4)

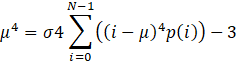
**Variance :**

The variation in the intensity is measured with the help of variance. It is also calculated by squaring the standard deviation.

 …(5)

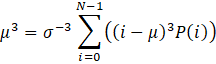
**Kurtosis :**

The histogram flatness is measured by kurtosis. It is the mathematical form kurtosis which depends on the standard deviation, mean and probability density.

 …(6)

**Skewness :**

Symmetry of an image is defined by the skewness. It is denoted by μ3

 ….(7)

**GLCM ALGORITHM STEPS:**

1. Image is given as the input.
2. Quantize the input image by mapping the intensity value of each pixel into specified discrete gray levels.
3. Define each element I ,j of the GLCM as the number of times two samples of intensities I and j occur in specified spatial relationship.
4. Make the GLCM symmetric by adding the matrix with its transpose.
5. Normalize the GLCM by dividing each element by the sum of all elements.
6. Calculate GLCM features such as inverse difference moment, sum average, entropy, variance, correlation.
7. Calculate total number of features extracted.

**5.3 Food segmentation**

FCM clustering is used to partition N objects into C classes. In our method, N is equal to the number of pixels in the image i.e., N=Nx x Ny and C=3 for 3-class FCM clustering. The FCM algorithm uses iterative optimization of an objective function based on a weighted similarity measure between the pixels in the image and each of the c-cluster centers.

FCM Algorithm steps :

Let X={x1, x2,…,xn} be the set of data points and V={v1,v2,v3,….,vn}be the set of centers.

1. Randomly select ‘c’ cluster centers.
2. Calculate the fuzzy membership ‘µij’ using:

1. Compute the fuzzy centers ‘vj’ using:

/( (uij)m ), Vj=1,2,….,c

1. Repeat step 2 and 3 until the minimum ‘J’ value is achieved or

|| U(k-1) – U (k) ||<β.

‘k’ is the iteration step.

‘β’ is the termination criterion between [0 , 1].

‘U=(µij)n\*c is the fuzzy membership matrix.

‘J’ is the objective function.

where,

’n’ is the number of data points.

‘vij’ represents the jth cluster center.

‘m’ is the fuzziness index m € [1,∞].

‘C’ represents the number of cluster center.

‘µij’ represents the membership of ith data to jth cluster center.

‘dij’ represents the Euclidean distance between ith data and jth cluster center.

**5.4 Food classification**

All the data containing the features of training images are collected and the label are added to it and stored. Classification is done by GoogleNet.

GoogleNet is used for food image classification by training the network with the food images. The cropped images are resized and fed to the input layer of the GoogleNet.The number of classes in the output layer is also modified accordingly.

//input:FVSTrain-feature vector of all training images,labels-labels for training image,FVSTest-feature vector for all testing image.s

//output:predLabels-predicted labels of the testing images

classifier()

{

trainNet=buildTrainingModel(net,FVSTrain,labels);

predLabels=findLabels(trainNet,FVSTest);

}

The features and labels from the training images are given as the input to the classifier model and it is trained and then the testing image is given to the model and the label is predicted.

This chapter has dealt on how the proposed system is implemented. The next chapter will provide illustration of the results provided by the system.

**CHAPTER 6**

**RESULTS AND DISCUSSIONS**

This chapter illustrates the result of the ‘Indian Food Image Segmentation and Classification’ that operates on features from the segmented image to classify the food images. Food images form Indian food image dataset were taken analysis.

**EXISTING SYSTEM:**

Automatic understanding of food is an important research challenge. One of the first challenges in this field is the discrimination between images containing food versus the others.

Dietary management has been carried out through food logging. Food logging is an activity in which the users document their energy intake to monitor their diet. Standard approaches for food intake monitoring demand the user to self-report all the food he/she eats and to recognize and describe quantities. Other methods may include the use of an exercise log book to document physical activities and the duration. Previously, users documented their intake using a food diary. But these are time-consuming process. However, the increase in smartphone usage has led to the increase of well-being applications that are able to facilitate food logging.

**PROPOSED SYSTEM:**

In the last few years, recognition of food items has become a popular research topic due to the availability of a large number of images on the internet and because of the interest of people in social networks.. The user just to take a picture of the food image then to recognize the image to detect the type of food portion and classify using GoogleNet we are performing segmentation, food portion recognition using fuzzy c-means and classification using GoogleNet.

**DATASET DESCRIPTION**

Indian Food Image dataset is used for food classification. The dataset consists of testing, training, and validation set. Each contains 20 major categories of food items.

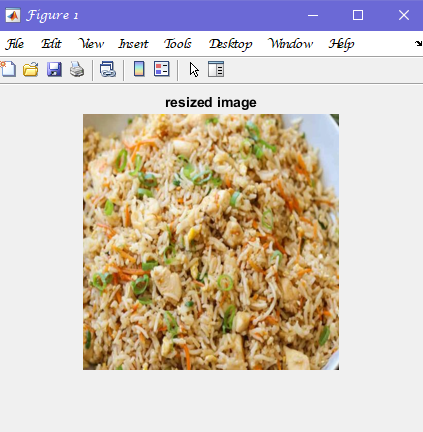
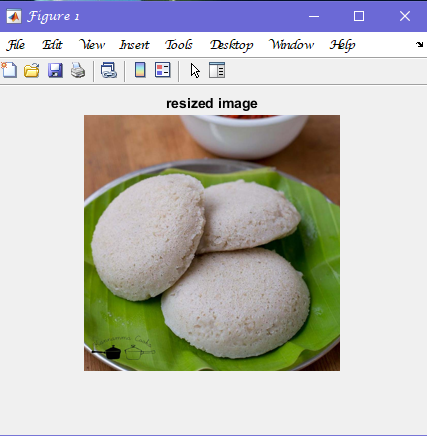
The images from Indian Food Image Dataset cover a wide range of food types in order to train a strong classifier that has the ability to classify varieties of foods. Many of the images contained in the food dataset were taken from real world environments, therefore the images contain the high color variation and some noise (unrelated food item) may be present.

**Table 6.1 Food category list**

|  |  |
| --- | --- |
| CATEGORY NUMBER | CATEGORY NAME |
| C0 | Burger |
| C1 | Butter naan |
| C2 | Chai |
| C3 | Chapathi |
| C4 | Chole\_bhature |
| C5 | Dal\_markhani |
| C6 | Dhokla |
| C7 | Fried\_rice |
| C8 | Idli |
| C9 | Jilabi |
| C10 | Kathi\_rolls |
| C11 | Kadai\_paneer |
| C12 | Kulfi |
| C13 | Masala\_dosa |
| C14 | Momos |
| C15 | Pani\_poori |
| C16 | Pakode |
| C17 | Pav\_bhajji |
| C18 | Pizza |
| C19 | Samosa |

**6.1 Resizing image**

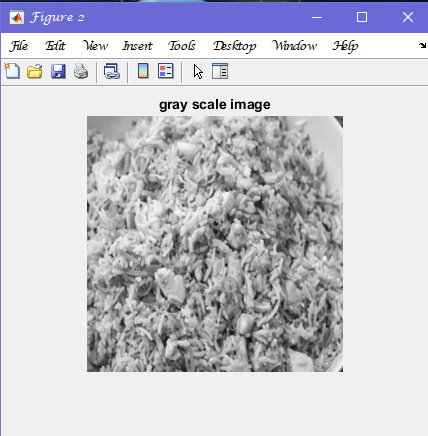
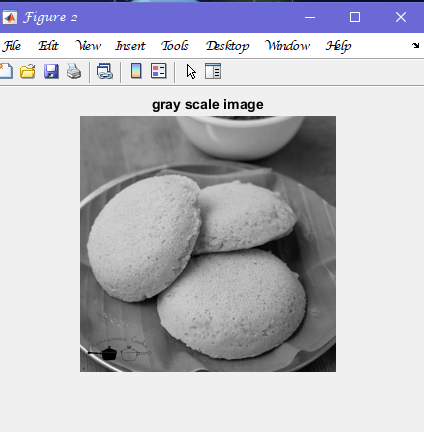
The food image from Indian Food Image Dataset are given as inputfor the process of Resizing. Figure 6.1.1, 6.1.2 describes resized image output for food items fried rice and idli.

**Fig 6.1.1** resized image of fried rice **Fig 6.1.2** resized image of idli

**6.2 RGB to Gray Scale Image**

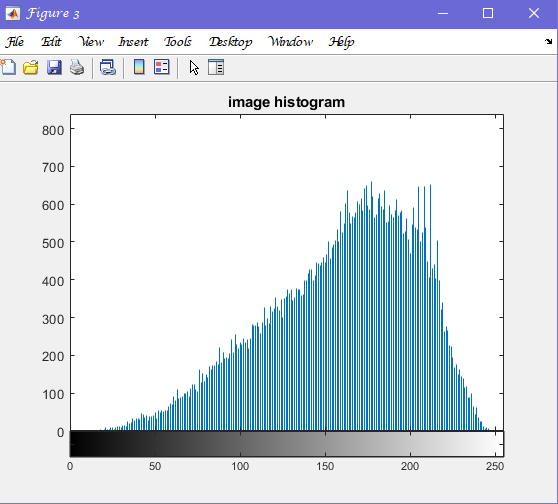
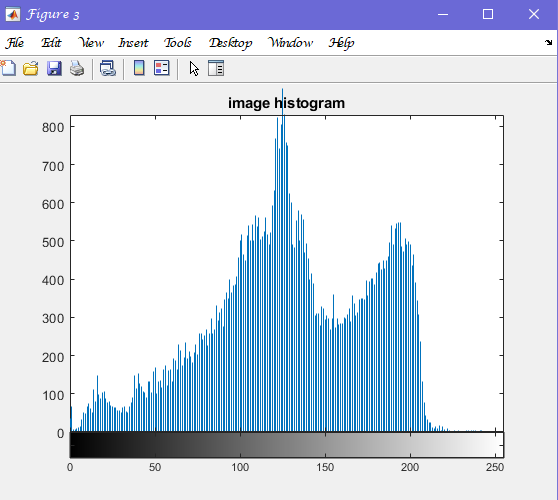
The food images from Indian food Image Dataset are given as input for the process of RGB to Gray Scale Image. Figure 6.2.1,6.2.2 describes the RGB to Gray Scale Image process.

**Fig 6.2.1** gray scale image of fried rice **Fig 6.2.2** gray scale image of idli

**6.3 Image Histogram**

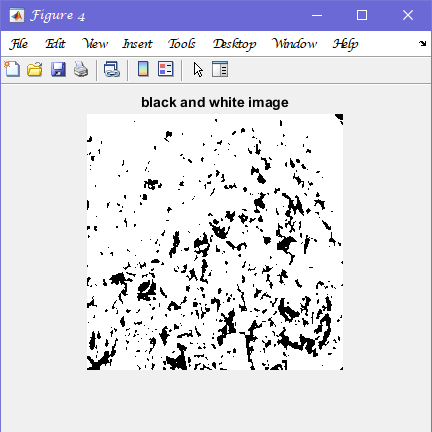
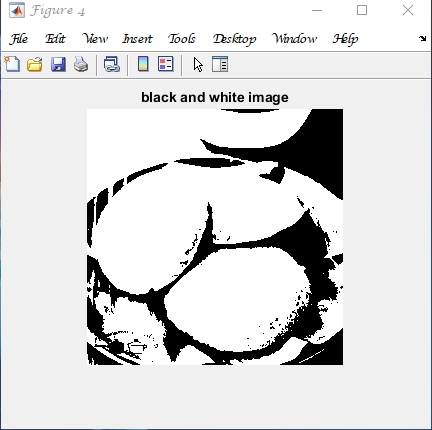
Preprocessed Gray Scale Image is given as input for the process of Image Histogram. The output is in the form of Graphical representation. Figure 6.3.1 and 6.3.2 describes the Image Histogram Process.

**Fig 6.3.1** histogram of fried rice **Fig 6.3.2** histogram of idli

**6.4 Gray to Black and White Image**

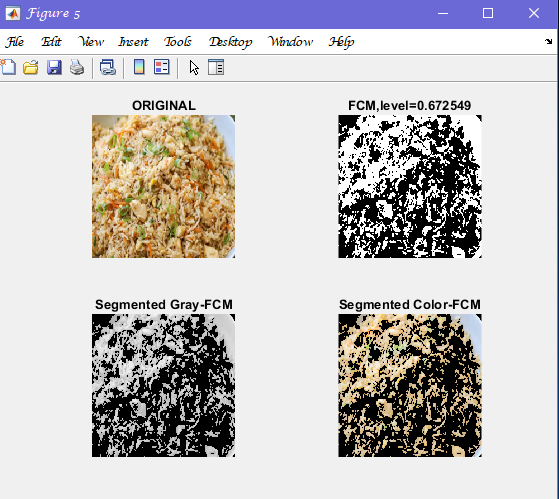
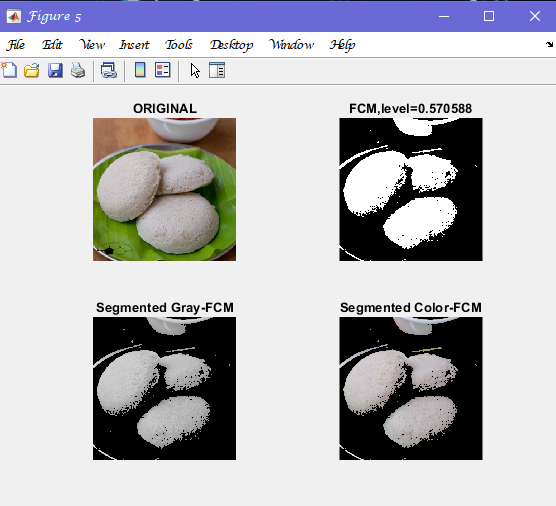
Food Image is given as input for the process of Gray to Black and White Image. Figure 6.4.1 and 6.4.2 describes the Gray to Black and White Image process.

**Fig 6.4.1** black and white image of fried rice **Fig 6.4.2** black and white image of idli

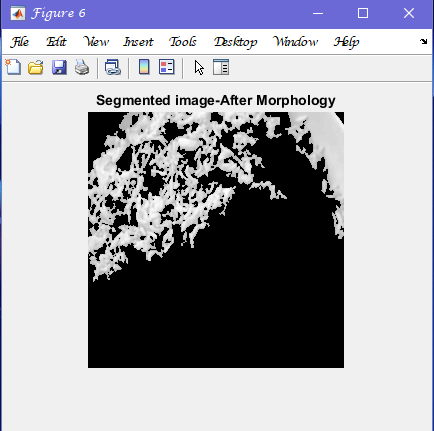
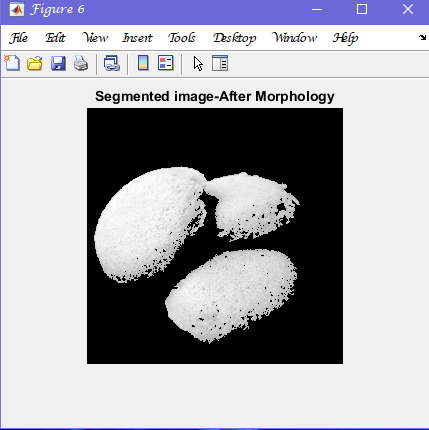
**6.5 Food Image Segmentation using FCM**

The input image is segmented using Fuzzy c-means algorithm. Food item is extracted from the original image. Figure 6.5.1,6.5.2 describes the Original image, level 1 FCM, Segmented Gray FCM, Segmented color FCM, Segmented image after morphology.

**Fig 6.5.1** segmented image of fried rice **Fig 6.5.2** segmented image of idli

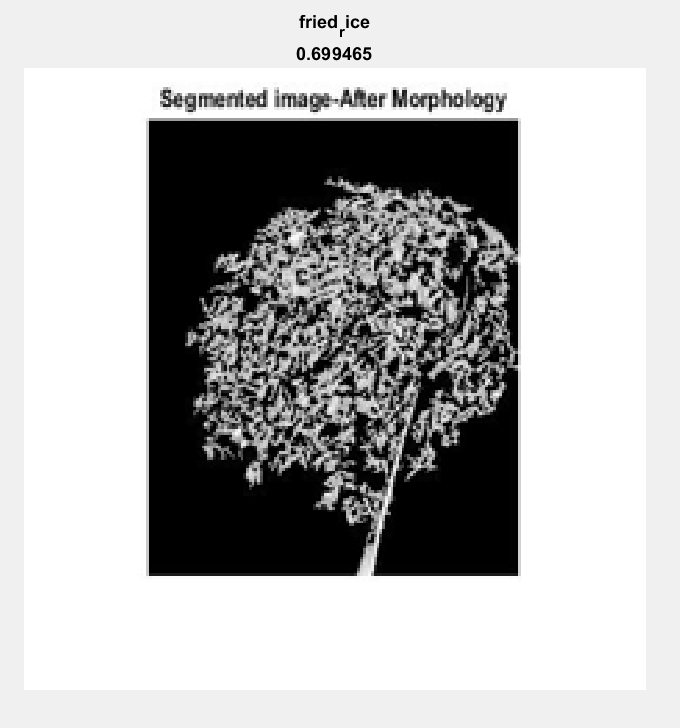
**AFTER MORPHOLOGY :**

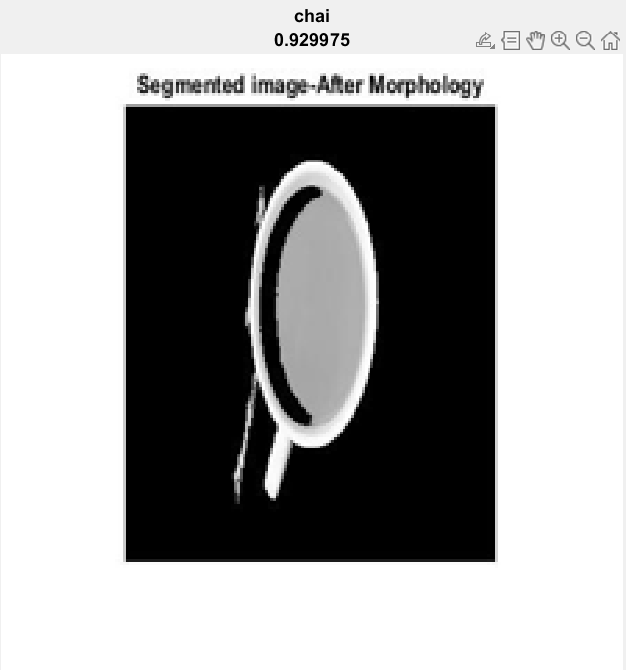
 

**Segmented image after morphology for fried rice and idli**

**6.6 Food Image Classification using GoogleNet**

After the Segmentation, food items are classified for training, the features along with the labels are given as input to supervised machine learning algorithms such as GoogleNet. These classifiers when trained with feature vector and label generates a model. The features of the image to be classified are given as input to the model to predict the class of the image. The output generated is a class label to which the test image belongs. **Figure 6.6.1** describes the image containing only one object then using the classifiers such as GoogleNet, the class labels are identified.

  
**Fig 6.6.1 image is identified as fried rice Fig 6.6.2 Image is identified as idli**

** **

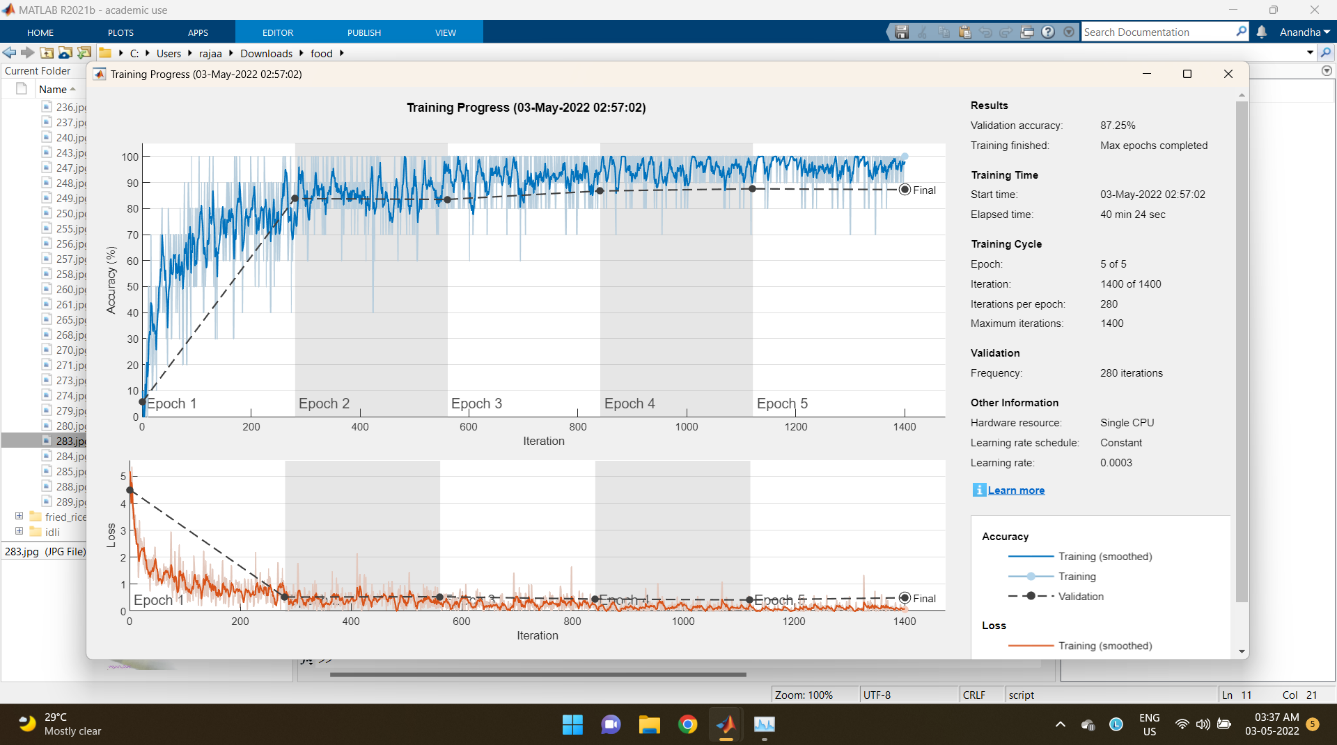
**Fig 6.6.3 image is identified as butternaan Fig 6.6.4 Image is identified as chai**

**6.7 PERFORMANCE ANALYSIS:**

Performance analysis is accomplished for 300 images from each category. For this purpose, the classifiers are trained with 300 training image and 300 testing image. **Fig 6.7.1** shows the accuracy (in %) results for the 20 food categories using GoogleNet.

The food categories co-c9 are classified by GoogleNet with total accuracy **87.25%**

**Fig 6.7.1** shows the results of accuracy and loss function .



**Fig 6.7.1 Performance Analysis**

Thus, the chapter has dealt with the results obtained from the proposed system. The performance of the system is also discussed. The next apter will discuss about the future enhancements related to the proposed system.

**CHAPTER 7**

**CONCLUSION AND FUTURE ENHANCEMENTS**

**Conclusion:**

The food image dataset is divided into training set and testing set. From the image the objects (food items) are segmented. Then the features are extracted from these images. To the extracted features labels are added. This process is done for training images. For testing, images are collected and then the features are extracted. Classifiers such as GoogleNet is used to train with the labels and features from the training set and a model is obtained. Using this model and features from the testing image food category can be predicted.

**Future Enhancement:**

The work can be further extended by finding the calorie content in the food item which has been classified then can be used for diet monitoring. The Pre-trained network which has been used for feature extraction, can be fine tuned to increase the accuracy. Food items can be classified as healthy and non-healthy.

**APPENDIX A**

**SPECIFICATIONS**

**A1.1 HARDWARE REQUIREMENTS**

RAM :8GB

PROCESSOR: Intel core i5

**A1.2 SOFTWARE REQUIREMENTS**

IDE: Matlab-(Version R2021b)

OS : Windows

**APPENDIX B**

**SOURCE CODE**

**CODE 1: FCM THRESHOLDING**

function [bw,level]=fcmthresh(IM,sw)

%FCMTHRESH Thresholding by 3-class fuzzy c-means clustering

if (nargin<1)

error('You must provide an image.');

elseif (nargin==1)

sw=0;

elseif (sw~=0 && sw~=1)

error('sw must be 0 or 1.');

end

data=reshape(IM,[],1);

[center,member]=fcm(data,3);

[center,cidx]=sort(center);

member=member';

member=member(:,cidx);

[maxmember,label]=max(member,[],2);

if sw==0

level=(max(data(label==1))+min(data(label==2)))/2;

else

level=(max(data(label==2))+min(data(label==3)))/2;

end

bw=im2bw(IM,level);

**CODE 2: Main**

clc;

clear;

close all;

X=uigetfile('\*.jpg;\*.tiff;\*.ppm;\*.pgm;\*.png','pick a jpge file');

I=imread(X);

im=imresize(I,[256,256]);

imshow(im);title('resized image');

I = rgb2gray(im);

figure;

imshow(I);title('gray scale image');

figure

imhist(I);title('image histogram');

figure

BW =im2bw(I,0.4);

imshow(BW);title('black and white image');

fim=mat2gray(im);

[bwfim1,level1]=fcmthresh(fim,1);

imgG=double(bwfim1).\*double(rgb2gray(im));

imgClr(:,:,1)=double (bwfim1).\*double(im(:,:,1));

imgClr(:,:,2)=double (bwfim1).\*double(im(:,:,2));

imgClr(:,:,3)=double (bwfim1).\*double(im(:,:,3));

figure;

subplot(2,2,1);

imshow(fim);title('ORIGINAL');

subplot(2,2,2);

imshow(bwfim1),title(sprintf('FCM,level=%f',level1));

subplot(2,2,3);

imshow(uint8 (imgG));title('Segmented Gray-FCM');

subplot(2,2,4);

imshow(uint8 (imgClr));title('Segmented Color-FCM');

%%FEATURES CALCULATION

%morphological operation

BWF=bwareaopen(bwfim1,3000);

img2=double(BWF).\*double(rgb2gray(im));

figure;

imshow(img2,[]);title('Segmented image-After Morphology')

G=img2;

%GLCM feature extraction

g=graycomatrix(G);

stats=graycoprops(g,'Contrast Correlation Energy Homogeneity');

Contrast=stats.Contrast;

Correlation=stats.Correlation;

Energy=stats.Energy;

Homogeneity=stats.Homogeneity;

Mean=mean2(G);

Standard\_Deviation=std2(G);

Entropy=entropy(G);

RMS=mean2(var(double(G)));

Variance=mean2(var(double(G)));

a=sum(double(G(:)));

smoothness=1-(1/(1+a));

Kurtosis=kurtosis(double(G(:)));

skewness=skewness(double(G(:)));

%INVERSE DIFFERENCE MOMENT

m=size(G,1);

n=size(G,2);

in\_diff=0;

for i= 1:m

for j=1:n

temp=G(i,j)./(1+(i-j).^2);

in\_diff=in\_diff+temp;

end

end

IDM=double(in\_diff);

%

GLCMfeat=[Contrast,Correlation,Energy,Homogeneity,Mean,Standard\_Deviation,Entropy,RMS,Variance,a,smoothness,Kurtosis,skewness];

mesg=sprintf('Successfully segmented');

msgbox(mesg)

function test\_network(net,image)

I = imread(image);

R = imresize(I,[224,224]);

[Label,Probability] = classify(net,R);

figure;

imshow(R);

title({char(Label),num2str(max(Probability),6)})

end

**CODE 3: Classify**

dirname = 'C:\Users\Elango\Documents\MATLAB\food';

Dataset = imageDatastore(dirname,...

'IncludeSubfolders',true,...

'LabelSource','foldernames');

[Training\_Dataset,Validation\_Dataset] = splitEachLabel(Dataset,0.7,'randomized');

Number\_of\_Classes = numel(categories(Training\_Dataset.Labels));

net = googlenet;

analyzeNetwork(net);

Input\_Layer\_Size = net.Layers(1).InputSize(1:2);

Resized\_Training\_Image= augmentedImageDatastore(Input\_Layer\_Size,Training\_Dataset,"ColorPreprocessing","gray2rgb");

Resized\_Validation\_Image=augmentedImageDatastore(Input\_Layer\_Size,Validation\_Dataset,"ColorPreprocessing","gray2rgb");

Feature\_Learner = net.Layers(142);

Output\_Classifier = net.Layers(144);

New\_Feature\_Learner = fullyConnectedLayer(Number\_of\_Classes,...

'Name','Food Feature Learner',...

'WeightLearnRateFactor',10,...

'BiasLearnRateFactor',10);

New\_Classifier\_Layer = classificationLayer('Name','Food Items Classifier');

Layer\_Graph = layerGraph(net);

New\_Layer\_Graph= replaceLayer(Layer\_Graph,Feature\_Learner.Name,New\_Feature\_Learner);

New\_Layer\_Graph = replaceLayer(New\_Layer\_Graph,Output\_Classifier.Name,New\_Classifier\_Layer);

analyzeNetwork(New\_Layer\_Graph);

Size\_of\_Minibatch = 5;

Validation\_Frequency = floor(numel(Resized\_Training\_Image.Files)/Size\_of\_Minibatch);

Training\_Options = trainingOptions('sgdm',...

'MiniBatchSize',Size\_of\_Minibatch,...

'MaxEpochs',20,...

'InitialLearnRate',3e-4,...

'Shuffle','every-epoch',...

'ValidationData',Resized\_Validation\_Image,...

'ValidationFrequency',Validation\_Frequency,...

'Verbose',false,...

'Plots','training-progress');

net2 = trainNetwork(Resized\_Training\_Image,New\_Layer\_Graph,Training\_Options);

YPred = classify(net2,Resized\_Validation\_Image);

YTest = Validation\_Dataset. Labels;

accuracy = sum(YPred == YTest)/numel(YTest)

save('classify\_20.mat',"net2");

**CODE 4: Testing**

function test\_network(net,image)

I = imread(image);

R = imresize(I,[224,224]);

[Label,Probability] = classify(net,R);

figure;

imshow(R);

title({char(Label),num2str(max(Probability),6)})

end

**REFERENCES**

1. Y. Matsuda, H. Hoashi, and K. Yanai, “Recognition of multiple-food images by detecting candidate regions,” in Proc. IEEE Int. Conf. Multimedia Expo., 2012, pp. 25–30.
2. Y. Kawano and K. Yanai, “Automatic expansion of a food image dataset leveraging existing categories with domain adaptation,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 3–17.
3. S. Memis, B. Arslan, O. Z. Batur, and E. B. Sonmez, “A comparative¨ study of deep learning methods on food classification problem,” in Proc. Innov. Intell. Syst. Appl. Conf., 2020, pp. 1–4.
4. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[5] K. Aizawa, Y. Maruyama, H. Li, and C. Morikawa, ``Food balance estimation

by using personal dietary tendencies in a multimedia food log,'' IEEE

Trans. Multimedia, vol. 15, no. 8, pp. 2176\_2185, Dec. 2013.

[6] D. Albanes, ``Total calories, body weight, and tumor incidence in mice,''

Cancer Res., vol. 47, no. 8, pp. 1887\_1892, 1987.

[7] S. Ao and C. X. Ling, ``Adapting new categories for food recognition with

deep representation,'' in Proc. IEEE Int. Conf. Data Mining Workshop

(ICDMW), Nov. 2015, pp. 1196\_1203.

[8] K. Ashton, ``That `Internet of Things' thing,'' RFID J., vol. 22, no. 7,

pp. 97\_114, 2009.

[9] S. Belongie, J. Malik, and J. Puzicha, ``Shape context: A new descriptor for

shape matching and object recognition,'' in Proc. Adv. Neural Inf. Process.

Syst., 2001, pp. 831\_837.