FOOD CLASSIFICATION AND CALORIE ESTIMATION USING COMPUTER VISION TECHNIQUES

Dhanalakshmi S[1], Harshitha S V[1], Mukeshwar varma D[1], and Mayuri P[2].

1. *Department of Computer Science Engineering, Sir M Visvesvaraya Institute of Technology, Bangalore, India,*
2. *Faculty of Department of Computer Science of Engineering, Sir M Visvesvaraya Institute of Technology, Bangalore, India.*

Abstract: In today's world, the important thing that matters is health. people have become more health-conscious and are careful about their diet. To consume quantified food every day, automatic recognition of the food image helps. As of now, there are no applications that can recognize food and estimate their calories automatically. Our proposed system not only detects varieties of fruits & vegetables but also provides per serving calories of each food detected in a single image. To achieve this, we will take the input of the food image from the user. This food item is detected with the help of the CNN algorithm. In the next step, we do image segmentation with the help of morphological functions of OpenCV. After Segmentation, the Volume of the food is calculated. After this, with the help of formulas calories of the food are calculated.

Keywords: Automatic recognition, Convolution Neural Network, Morphological functions, OpenCV.

I. INTRODUCTION

According to the NIH, obesity is the second leading cause of preventable death. A million people die yearly due to obesity. Nowadays it is very difficult for a person to track the calories consumed by them. The intake of calories plays a very vital role in one's healthy lifestyle. Earlier the users used to track their calorie intake with the help of charts. These methods are very hectic to follow and lead to an unquantified meal diet. Having a meal which is quantified helps you reduce that extra fat. We came up with an idea to help the people track the number of calories that it takes in with the help of simple images of the food that is captured by the user instantly. There are various apps out there available. In those apps, the user has to manually input all the data. Computer vision is also used to estimate the amount of calories present. There are a lot of algorithms available for object detection. Each method has its advantages and disadvantages. In our project, we are using the CNN algorithm.

II. RELATED WORK

S. Jasmine Minija et al. (2017), proposed a method where they pick a stable segmented region from multiple such segmented regions. After all images are segmented, the global and local features area is obtained via feature extraction based on texture, local neighborhood pixel and color. Once classification is performed, calorie value is obtained and an accuracy of 97% was obtained.

Hong Liang et al. (2018), proposed a method where the volume of food is estimated by 3-D reconstructions through depth camera or planar image sequence. Food calorie estimation is done using the Deep-learning approach. Datasets used are Microsoft COCO and Pascal VOC for image recognition. The calorimeter R2 and RMSE are about 0.95 and 43 and MSE is 32. Updating food ingredients information to meet the needs of different users. And gather larger training sets to improve accuracy and speed is aimed in the future.

Shaikh Mohd. Wasif et al. (2019), proposed a method where they have taken an image as input from the user. The image is then passed to faster R-CNN model. After image detection, they performed image segmentation. For image segmentation, they used grab cut algorithm. Segmentation of the image is then performed. For calorie estimation, they calculated the volume of the food item. The volume is then used to calculate the calories present in the food item. All the above modules are then integrated to make the software for calorie estimation using images of the food item. The accuracy obtained is 90%.

Yanchao Liang et al. (2017), proposed a deep learning method. In the dataset, both food volume and mass records are given. To estimate calories, it takes the top view and side view of the food. Every image has a calibration object which will be used to find out the scale factor of the image. Food(s) and calibration objects are detected by the object detection method called Faster R-CNN and each food counter is obtained by applying the Grab Cut algorithm. After that, they estimated each food's volume and calories.

S. Jasmine Minija et al. (2019), proposed an automatic method of food category recognition and calorie estimation using the BFC and IpCA-DBN. For automatic recognition, the features are extracted, for which the optimal segments are offered using the BFC. The features increase the robustness of the recognition, and the method consumes less time, which is applicable for assessing the diet in daily life. Bayesian Fuzzy Clustering (BFC)is the optimal clustering mechanism, and the main role of BFC is to generate the optimal segments from the food image. BFC is characterized using a set of unknown memberships, and prototypes. DBNs are the generative neural networks that consist of the multiple layers of Restricted Boltzmann Machines (RBM) with each layer holding the input and the hidden neurons and the hidden neurons constituting the output layer. Once the food category is recognized using the proposed IpCA-DBN, the calorie of the food is determined. The accuracy obtained is 96%.

Pallavi Kuhad et al. (2015) , proposed a method where the Category of tools is been considered, which uses image processing to recognize single and multiple mixed food objects, namely deep learning and SVM. Finger -based calorie measurement,

distance measurement methods are applied for calorie estimation. A block resize method is also been proposed that uses the measured distance values along with the recognized food object name to further estimated calories. The accuracy obtained is 97%.

W R Samemmanuel et al. (2018), proposed a method which introduces the dietary assessment system through the efficient identification of the calorie values of each food items from the food images. The dataset contains various food or the meal images consumed by the person every day. The images in the food database are preprocessed for better processing. The proposed CSW-WLIFC algorithm does segmentation on each food. After segmentation, The feature extraction will extract features like shape, texture from the image. The proposed WLM-NN classifier performs to train the neural network with the existing LM model and the Whale Optimization Algorithm (WOA), which also performs classification of the food items from the image. Food items obtained from the WLM-NN classifier. The accuracy obtained is 96%.

Hattarki Pooja et al. (2017), proposed a system which takes as input an image of a food item and outputs the number of calories in this food item. The image is passed through Mathworks Image Processing Toolbox. The toolbox extracts visual features from the food item. The compressed image is passed through a classifier that predicts the type of food item in the image. The compressed image is also passed to a regressor which estimates the size of the image. Lastly, the compressed image is passed to another regressor that predicts the number of calories in the food. The accuracy obtained is 93%.

Sulfayanti F. Situju et al. (2019), proposed a method where Multi-task CNN is used for food ingredient estimation. Dataset is obtained from the collection of a large number of food images. Two-stage transfer learning was been proposed to improve the estimation accuracy. The accuracy obtained is 92.6 %. In calorie estimation, the error in categories with a large mean and large intra-class variance is large.

Gozde Ozsert Yigit et al. (2018), proposed a method which uses Datasets Food11 and Food101. Deep convolution neural network structure is been trained from scratch by using different learning methods l- Stochastic gradient descent, Nesterov's accelerated gradient, and adaptive moment estimation. The tool used for pattern recognition problems is CNN, which is based on the cats' visual cortex method. The time taken for the training and testing process is very high. Adam and Nesterov's technique took almost 2.5 hours. Pre-trained food recognition structures with different classifiers are aimed in the future.

III. METHODOLOGY

A system is developed which detects the food from the given input. In addition to this, the system also helps to estimate the calorie intake of the food. The Proposed flow diagram is depicted in the figure 1.

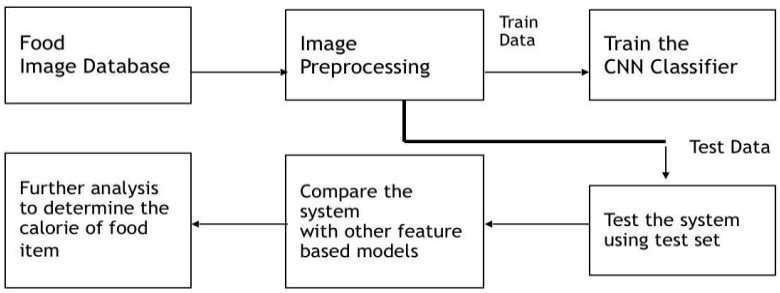

fig 1. Block Diagram

The steps involved in image processing are preprocessing and neural network training and from this, the trained model is obtained which will classify any supplied image based on the trained dataset.
The proposed methodology allows for automatic food detection and calorie estimation.

The system consists of four stages:
1. Image Acquisition and Preprocessing
2. Neural Network Training
3. Image segmentation
4. Calorie Estimation

I. Image Acquisition and Preprocessing:

Fruits 360 dataset [11] is used which contains 90483 images of 131 fruits and vegetables. For our project we consider 15 types of fruits and vegetables. The dimensions of the images are 100x100 which are resized to 224*224. After resizing the images are converted to 4D tensors with the shape of (1, 224, 224, 3) and can then be passed to the Convolutional Neural Network for learning

II. Neural Network Training:

The base of the model being made is deep learning and the technique used is convolution neural networks (CNN), the model is being developed from scratch. CNN also called ConvNet is a class of deep neural networks which is based on shared-weights architecture and translation invariance characteristics. In CNN, the network employs a mathematical operation called Convolution. Convolution is a specialized kind of linear operation. A CNN consists of an input and an output layer with the multiple hidden layers between input and output layer.

The model now takes the input array of dimension 4 and processes it further through the network. The main aim was to give perfect result with less errors so the number of filters were increased. First we started with 16 filters and kept padding as same as it may help not to lose the data and then the MaxPooling layer with pool_size=2 to reduce our data width wise. Then to go deep inside and classify, the number of filters were increased to 128 in following sequence 16,32,64,128 and droupout layer were used to reduce the chances of overfitting. The flattering layer is used to connect CNN layer to fully connected flattering layer. Then the hidden layers are connected having 'relu' activation function. The end layer of the model is dense layer of 15 nodes as we have 15 types of food which uses Softmax function' which gives the probabilities of the type of fruit in the Image. The Sequential CNN model is compiled with a Root Mean Square Propagation (RMS Prop) optimizer which uses moving average of squared gradients which normalizes itself. Also, the model is compiled with the 'categorial_crossentropy' loss.

```
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_9 (Conv2D)            (None, 224, 224, 16)      208
_________________________________________________________________
max_pooling2d_9 (MaxPooling2 (None, 112, 112, 16)      0
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 112, 112, 32)      2080
_________________________________________________________________
max_pooling2d_10 (MaxPooling (None, 56, 56, 32)        0
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 56, 56, 64)        8256
_________________________________________________________________
max_pooling2d_11 (MaxPooling (None, 28, 28, 64)        0
_________________________________________________________________
conv2d_12 (Conv2D)           (None, 28, 28, 128)       32896
_________________________________________________________________
max_pooling2d_12 (MaxPooling (None, 14, 14, 128)       0
_________________________________________________________________
dropout_5 (Dropout)          (None, 14, 14, 128)       0
_________________________________________________________________
flatten_3 (Flatten)          (None, 25088)             0
_________________________________________________________________
dense_5 (Dense)              (None, 500)               12544500
_________________________________________________________________
dropout_6 (Dropout)          (None, 500)               0
_________________________________________________________________
dense_6 (Dense)              (None, 15)                7515
=================================================================
Total params: 12,595,455
Trainable params: 12,595,455
Non-trainable params: 0
```

Fig 2.Model Summary

III. Image Segmentation:

Segmentation is the process of partitioning a digital image into multiple segments with the aim to simplify or change the representation of the image into something that is more meaningful and easier to analyze. An adaptive thresholding is used to convert an image consisting of gray scale pixels to just black and white scale pixels. Usually a pixel value of 0 represents white and the value 255 represents black with the numbers from 1 to 254 representing different gray levels. A Contour based segmentation is done by calculating the number of contours and by finding the biggest contour. The biggest contour corresponds to the plate and the food. An interface was used that allowed users to segment the food portion using a polygonal tool and the application also detected the one-centimeter square using Hue Saturation Value (HSV) color thresholds to measure food portion area. Once the image has been convert to HSV to remove plate and fruit pixels, we obtain fruit pixels and then it is converted to binary. Later we have implemented two fundamental morphological operations that is erosion and dilation. First erosion has been applied to remove the pixels from object boundaries and then dilation has been implemented to add the pixels to the object and the area of the food has been calculated. Finally, we then have image of just the fruit (in this case apple). For different foods different set of morphological operations need to be performed multiple times in order to get the food region. To calculate area we convert the pixels to area in cm square. We do this by dividing the area with the skin area and multiple it with a constant which is found by trail and error.

IV. Calorie Estimation:

Once area of the food has been found, the volume of the food has been calculated considering the different shape of the food. In our case we have used an apple. First Radius is obtained by dividing it with pi (3.14). Then volume of sphere is obtained using the formula

Volume of sphere = (4/3)*pi*radius*radius

Volume of other fruit shapes such as cylinder and ellipsoid can be found by using the appropriate volume formulae. An excel sheet is used which consists the density values of all the food. By using the density values and the calculated volume value, the mass of the fruit has been obtained. By considering the value of 100grams of food, the total calories of the food has been calculated.

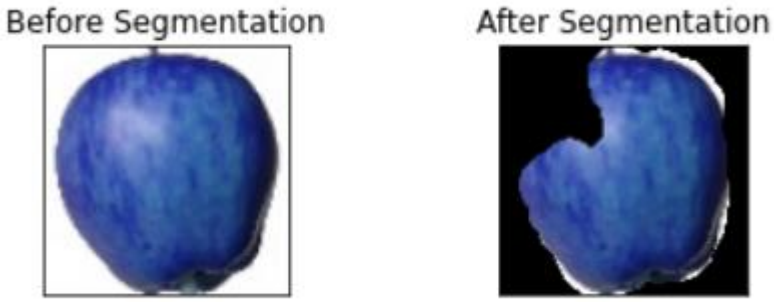

Fig 3. Food Segmentation

IV. RESULTS

The training and the validation images are run for 15 epochs. Of batch size of 50. Early stopping was used which monitors the validation loss. It stops the training of the network at 7th epoch as the validation accuracy didn't improve.

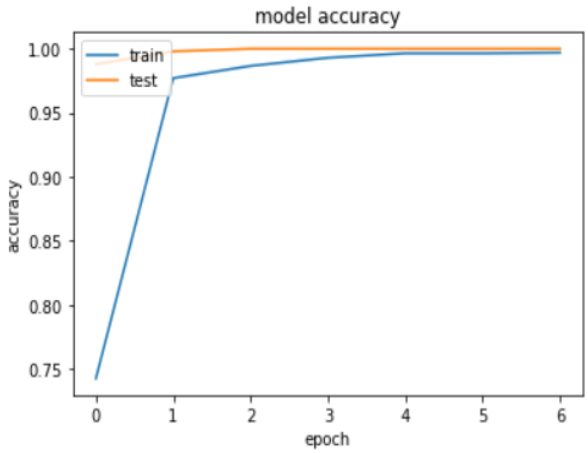

Chart 1 Accuracy of CNN model

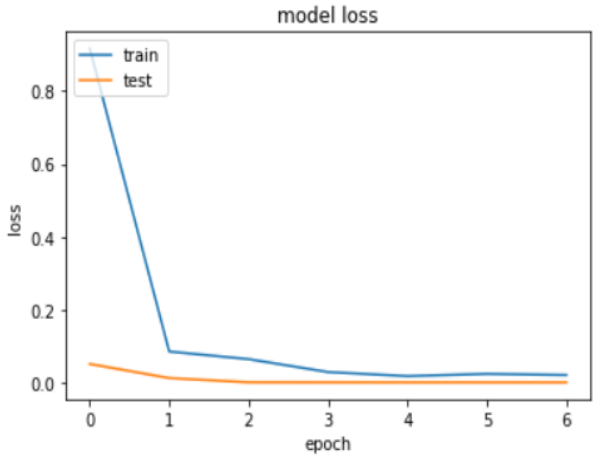

Chart 2 Loss of CNN model

Confusion Matrix is also obtained after training and testing model as accuracy alone is not enough to decide if the model is working as expected

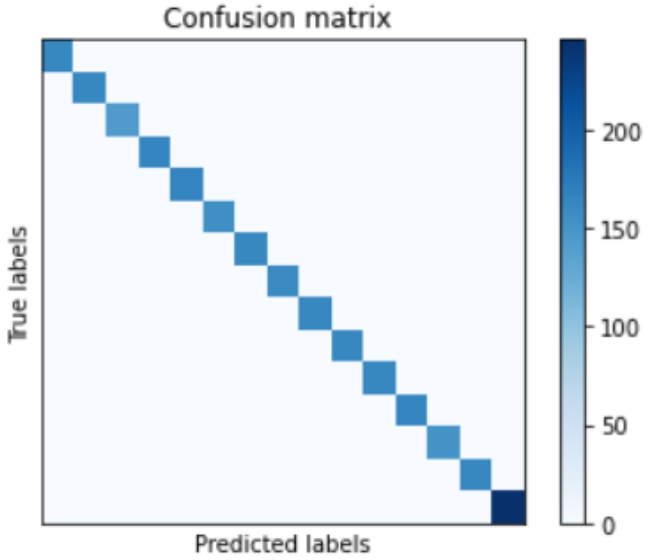

Fig 5.Confusion Matrix

The calories of the apple was found to be 69 calories which is about the average which is found in medium sized apples.

V. CONCLUSION

In our project, we have kept the scope of the project only for fruits and some vegetables. We cannot calculate the calories of fast food and other cooked food items. The reason for that is we are unable to identify the ingredients present in the food item just with the help of an image which contains food item. We need to know all the ingredients and the quantity of the same in the food. We can further improve our project by adding fast food and more cooked food. In the above project, we have calculated the calories of food item using machine learning and image processing. For that we have implemented three modules, First, we detect the object using CNN. Then we Segment the image using grabcut algorithm. After segmentation, we calculate the volume of the food item. In the final step, we calculate the calories of the food item. By the results of our project, it has been implemented successfully. The overall process of development of the model from starch was little though but due to this project we learnt many new techniques and cleared concepts regarding machine learning. The most interesting aspect was to pre-process the data and to do the same it requires large memory. The model takes an input of 224*224, and predicts the output accordingly. This can be used as a part for many application and automation systems.

VI. ACKNOWLEDGEMENTS

The Author is thankful to the Sir M Visvesvaraya Institute of Technology for providing computer lab facilities.

REFERENCES

[1] “Imperialist competitive algorithm-based deep belief network for food recognition and calorie estimation”, Springer-Verlag GmbH Germany, July 2019.
[2] “Food Constituent Estimation for Lifestyle Disease Prevention by Multi-Task CNN”,Taylor and Francis , 2019
[3] “Food calorie estimation using machine learning and image processing”, International journal publication, Feb 2019.
[4] “CEP- Calories estimation from food photos”, Taylor and Francis, June 2018
[5] “Fuzzy clustering and Whale-based neural network to food recognition and calorie estimation for daily dietary assessment” Indian Academy of Sciences, 2018
[6] “Comparison of convolutional neural network models for food image classification”, Taylor and Francis, Mar 2018
[7] “Neural network classifier and multiple hypothesis image segmentation for dietary assessment using calorie calculator”, Taylor and Francis, Aug 2017
[8] “Computer vision-based food calorie: Dataset, method, and experiment”, School of Information Science and Engineering, may 2017.
[9] “Food recognition and calorie estimation using Bag of surf and facial pyramid matching methods, 2017.
[10] “A deep Learning and Auto-Calibration Approach for Food Recognition And Calorie Estimation in Mobile e-health”, School of EE and CS, 2015
[11] Horea Muresan, Mihai Oltean, Fruit recognition from images using deep learning, Acta Univ. Sapientiae, Informatica Vol. 10, Issue 1, pp. 26-42, 2018.