INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

Food calorie estimation using machine learning and image processing

Shaikh Mohd. Wasif
mohd.wasif.s@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Swapnil Thakery
swapnil.thakery@gmail.com
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Amir Nagauri
aamir.n@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Sheetal Ignatius Pereira
sheetalpereira@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

ABSTRACT

In today's world a healthy lifestyle is a must for every individual and what they consume is of utmost importance in order to achieve the same. Our paper focuses on creating software which gives the calorie of the food which the user is going to consume. In order to achieve this, the software will take two images as input from the user, the top view and the side view. The image will have a probe object as well which is a coin whose volume will be known. The food item in the image will be detected with the help of Faster R-CNN algorithm. We are using Faster R-CNN algorithm since it is the fastest among all the object detection algorithms. In Faster R-CNN we don't use selective search algorithm on the feature map to identify the region proposals, we use a separate network is used to predict the region proposals. This makes it the fastest algorithm and therefore it can be used for real-time object detection. In the next step, we will segment the image using the grab cut algorithm. It is needed for foreground extraction with minimal user interaction. After segmentation of images, the volume of the food item is calculated using the known volume of the probe object. After the calculation of volume, the mass of the food item is calculated with the help of formulas and then the calories of the food item will be calculated using the relation between mass and calories.

Keywords— *Machine learning, Calorie estimation, Image processing*

1. INTRODUCTION

Health is one of the most important aspects of an individual's life. It takes some amount of effort from a person to stay in shape and maintain a healthy diet. 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 calories intake with the help of charts or timetable. Or they used to maintain a strict diet where the food item which has to be consumed was fixed along with its quantity. These methods are a bit tedious and difficult for the user to follow judiciously. We have come up with a project to help the user track the number of calories which it takes in with the help of simple images of the food item. There are already many different apps and products available to do the same. In these apps, the user inputs the ingredients and their amount which they are consuming. The apps then search them in their database and calculate the calories present in them with values present in the database. Computer vision is also used to estimate the amount of calories present. The accuracy of these projects is determined by two factors, the accuracy of the object detection algorithm and the method to calculate the volume. There is a various algorithm which is available for object detection for eg. R-CNN, Fast R-CNN, Faster R-CNN and YOLO. Each method has its own advantages and disadvantages. In our project, we are using Faster R-CNN algorithm and the reason for this is that it is the fastest among all the other algorithms. It does not use selective search and hence it is quicker in detecting the objects. We use a very extensive dataset of thousands of images of different types of fruit for our project. This gives our model an edge over other such projects. The accuracy of our model is also very high. Most of the other projects either take the ingredients along with the food item. We calculate the calories with the help of segmented image using formulas and probe object whose dimensions are already known.

2. METHODOLOGY

Our System architecture is depicted in figure 1. For our system to work we take an image as input from the user. The input image contains two views the top and side view. The image is then passed through faster R-CNN block which is used for image

detection. This is explained in section 2.2. The Faster R-CNN gives coordinates, confidence and category. After image detection, we perform image segmentation. For image segmentation, we use grab cut algorithm. The grab cut algorithm is explained in section 2.3. Since segmentation of the image is done we perform calorie estimation on the segmented image. For calorie estimation, we calculate the volume of the food item. The volume is then used to calculate the calories present in the food item. We explain this in section 2.4. All the above modules are then integrated to make the software for calorie estimation using images of the food item.

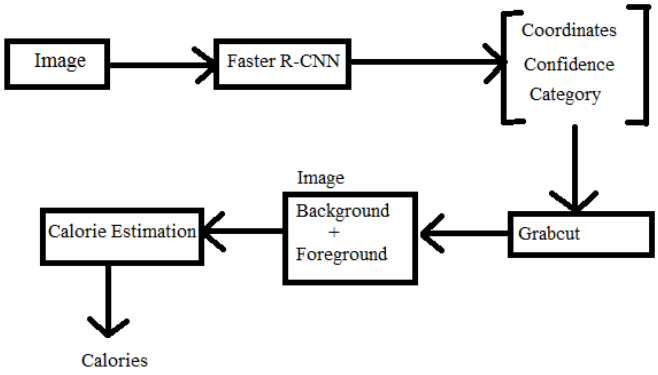

Fig. 1: System architecture

2.1. Dataset

Our dataset is made from manually clicking pictures of a food item by the members of this project. We have also used the dataset from GitHub. The dataset contains images of food items along with the XML file of that image. The XML file is used for obtaining the coordinates of the food item in the particular image. The dataset contains images of ten different food items. These include apple, banana, bread, donut, mango, lichi, lemon, kiwi, orange and egg. The dataset is used for training the images. The dataset has thousands of images of each food item. In our dataset, we have images with one food item as well as multiple food items.

2.2 Image Detection

For image detection, we use Faster R-CNN algorithm. In this algorithm, we provide the image to the convolution network as an input and we get a convolution feature map as output. The algorithm uses a separate network to predict the region proposals. The ROI pooling layer is used for reshaping the predicted region proposals. The offset values for the bounding boxes are predicted with their help and the images are also classified within the proposed region. Faster R-CNN is used for image detection because it is the fastest among all the other detection algorithms such as R-CNN and Fast R-CNN. Faster R-CNN architecture is shown below.

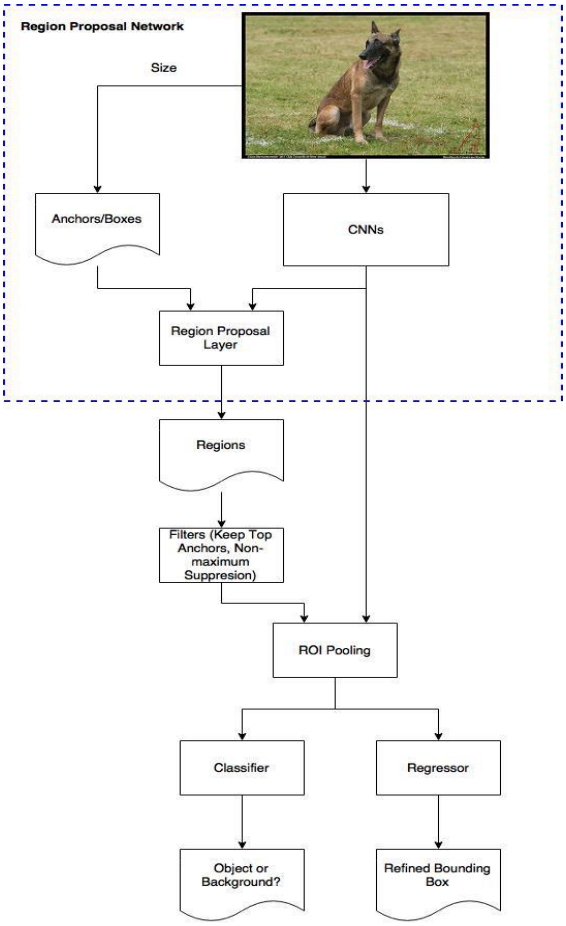

Fig. 2: Faster R-CNN architecture

Anchors are used in Faster R-CNN. An anchor is basically a box. A position of an image consists of 9 anchors in the default configuration of Faster R-CNN. There is a chance that if a good set of anchors is used it can help in improving the speed as well as accuracy. The region proposal network (RPN) gives the output of a bunch of boxes/proposals. The classifier and regressor will examine the output to check for the occurrence of objects.

2.3 Image Segmentation

Image Segmentation using the grab cut algorithm. In the grab cut algorithm, we mark the object with the help of a box. The box is used to distinguish the object from the background. The outer part of the rectangle is the background and the inner part of the algorithm is the combination of some background and the object. After that, we use an iterative method which assigns each pixel as a background or foreground pixel. The initial labelling is done by the computer according to the data provided by the user. Then we use the Gaussian mixture model to model the foreground and the background. Then we mark the probable foreground or probable background pixels to unknown pixels. It depends on their relationship with other hard-labelled pixels in terms of colour statistics. This is very similar to clustering. Every node in the graph is a pixel additional nodes such as source and sink nodes are added. The source node is connected to the foreground pixel and sink node are connected background pixel. The probability of a pixel whether it is foreground/background determines the weight of the edges which are connecting the pixels to the source node or end node. The information on the edge or the similarity between pixels is used to determine the weights between the pixels. Low weight will be assigned to the edge between the pixel if the difference in pixel colour is very large. Then the Segmentation of the graph is done using mincut. The graph is cut into two separating source node and sink node with a minimum cost function. The sum of all weights of the edges that are cut is known as the cost function. After the cut, all the pixels connected to the Source node become foreground and those connected to Sink node become background. The process is continued until the classification converges.

2.4. Volume and Calorie Estimation:

For the calculation of calories, we take the help of a probe object whose dimensions are already known to us. We use a 1 rupee coin as a probe object. The diameter of the used coin is 25 cm, we calculate the side views scale factor with the help of equation 1.

$$\boldsymbol{\alpha = 2.5/(W_s + H_s)/2} \tag{1}$$

Here, the width of the bounding box is represented by *Ws*, the height of the bounding box is represented by *Hs*.

After this, we calculate the volume of the food item based on its shape. There are three equations for the same. The three equation is shown in Equation 3.

$$\begin{aligned} \boldsymbol{v = \beta \times \frac{\pi}{4} \times \sum_{K=1}^{H} \left(L^k\right)^2 \times \alpha^3} \quad & \textbf{if the shape is ellipsoid} \\ \boldsymbol{v = \beta \times s \times \alpha^3 \times H} \quad & \textbf{if the shape is other} \end{aligned} \tag{2}$$

Here, the height of the side view is represented by HS, the number of foreground pixels in row k (k ∈ 1,2,…,HS) is represented by PS and LkS. The max number of foreground pixel in PS is recorded by LMAX = max(Lk ,…,Lk). ß is a compensation factor whose default value = 1.0.

Calculation of calories is done after the calculation of volume. We use the relation between volume, calorie and density for the calculation. Equation 4 shows the relation between the density and volume of the food item. Here, the volume is represented by *v (cm³)* and the density is depicted by *ρ (g/cm³)*.

$$\boldsymbol{m = \rho \times v} \tag{3}$$

The calorie is calculated with the help of Equation 5.

$$\boldsymbol{C = c \times m} \tag{4}$$

The mass of current food is depicted by m (g) and its calories per gram are depicted by c (Kcal/g).

3. RESULTS

In this section, we will discuss the results obtained during the execution of our project. We have three modules in our project and the results of these three modules are discussed separately. All our experiments were run on an Intel(R) Xeon(R) computer with a 2.00 GHz CPU and a 16.0 GB RAM memory and 8.0 GB graphic card.

3.1 Image Detection Result

In this section, we show the results obtained by Faster R-CNN algorithm. For this, we have applied the algorithm on our dataset with a varying number of testing and training images. First, we have applied for 80% training and 20% testing. Then we have applied for 60% training and 40% testing. Then we have applied for 75% training and 25% testing.

Each image was passed through tensorflow library for training the dataset and after training the dataset it was tested and the results were verified by looking at the image. The accuracy of the detection algorithm was found to be above 90% for all the images.

3.2 Image Segmentation Result

In this section, we discuss the results obtained by grab cut algorithm. We have taken our dataset and entered the dataset into grab cut algorithm. The segmented image was shown as output.

3.3 Calories Estimation Result

In this section, we discuss the results obtained by calorie estimation. For analysis of the results, we have calculated we compare it to the actual calorie of the food item available on the internet.

If the error in the calculation is less than ±20% then our project is a success. The below table shows the result of the same.

Food Item	Actual Calories	Obtained Calories	Error in %
Apple	95	92	3.1
Banana	105	99	5.7
Bread	53	57	7.5
Donut	195	203	4.1
Mango	201	210	4.4
Lichi	6	7	16
Lemon	17	18	5.8
Kiwi	42	46	9.5
Orange	87	91	4.5
Egg	78	83	6.4

As we can see the error is less than 10% our project is in good working condition.

4. DISCUSSIONS

In our project, we have kept the scope of the project only for fruits and some uncooked food items like donuts and bread. 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.

5. CONCLUSIONS

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 Faster R-CNN. Then we Segment the image using grab cut 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.

6. ACKNOWLEDGMENT

We would like to thank Riddl for providing the equipment for completing this project and our guide Mrs Sheetal Ignatius Pereira for her timely help.

7. REFERENCES

[1] Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment, Yanchao Lianga, Jianhua Lia. a School of Information Science and Engineering, East China University of Science and Technology, China.
[2] Manal Chokr, Shady Elbassuoni Calories Prediction from Food Images, Department of Computer Science, American University of Beirut, Beirut, Lebanon, mmc35@mail.aub.edu - se58@aub.edu.lb
[3] Interactive Foreground Extraction using GrabCut Algorithm, https://docs.opencv.org/trunk/d8/d83/tutorial_py_grabcut.html
[4] Boykov, Y., and Jolly, M.-P. 2001. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In Proc. IEEE Int. Conf. on Computer Vision.
[5] https://grabcut.weebly.com/background--algorithm.html
[6] Parisa Pouladzadeh1, Pallavi Kuhad1, Sri Vijay Bharat Peddi1, Abdulsalam Yassine1, Shervin Shirmohammadi1,2. 1 Distributed and Collaborative Virtual Environments Research (DISCOVER) Lab University of Ottawa, Ottawa, Canada. Colleges of Engineering and Natural Sciences, Istanbul Şehir University, Turkey Food Calorie Measurement Using Deep Learning Neural Network.
[7] Tatsuya Miyazaki Gamhewage C. de Silva Kiyoharu Aizawa. Interfaculty Initiative in Information Sciences, The University of Tokyo, Tokyo, Japan, Image-based Calorie Content Estimation for Dietary Assessment.

