International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6935
Article Received: 27 March 2023 Revised: 14 May 2023 Accepted: 28 May 2023

Foodopedia: A Convolutional Neural Network Based Food Calorie Estimation

Dr. Pramod Patil[1], Bineeta Bidyut Panja[2], Govind Deepak Mundada[3], Monika Abhiman Nandurkar[4]
[1]Professor, Research and Development Cell
Dr. D.Y. Patil Institute of Technology
Pimpri, Pune
pdpatiljune@gmail.com
[2]Computer Engineering
Dr. D.Y. Patil Institute of Technology
Pimpri, Pune
bineetapanja05@gmail.com
[3]Computer Engineering
Dr. D.Y. Patil Institute of Technology
Pimpri, Pune
govindmundada110@gmail.com
[4]Computer Engineering
Dr. D.Y. Patil Institute of Technology
Pimpri, Pune
nandurkarmonika05@gmail.com

Abstract— The number of calories consumed determines how healthy a body is in the modern world; therefore, it's important to watch your calorie intake to be healthy. People must keep track of their caloric intake in order to become in shape or maintain a healthy weight. The suggested model uses a deep learning algorithm to offer a novel method of calorie measurement. In the medical area, the estimation of dietary calories is crucial. This measurement is derived from the representation of food in various objects, such as fruits and vegetables. The neural network is used to take this measurement. This technique uses a convolutional neural network to determine the calories in food. An image of food is used as the input for this calculated model. The suggested CNN model uses food object identification to calculate the calorie content of the food. Volume error estimation serves as the primary parameter for the outcome, while calorie error estimation serves as the secondary parameter.

I. INTRODUCTION

The key to a human's body is food. Since a poor diet increases the risk of many diseases, individuals are becoming more and more concerned about their nutritional consumption today. To maintain a fit and healthy lifestyle, a diet plan must constantly account for the overall number of calories to be ingested. If your BMI is greater than 30, you are likely obese. Many factors can influence weight. One of these causes is high calorie expenditure. People must keep track of their caloric intake in order to become in shape or maintain a healthy weight. People frequently eat more than they should since they generally avoid difficult and taxing things, which can lead to obesity. Among these investigations, the volume and calorie evaluation approach, as well as object location calculation, are two important factors of the precision modification. We developed a Machine Learning Basis approach to accomplish this. People will actually want to know how many calories are in the meal they are eating. We identify the food in this inquiry, give it a general description, and estimate its volume. Finally, we calculate the food's calories based on the volume that the models predicted. However, we discovered that calculating the calories directly provided us with considerably more accurate results. Unfortunately, the lack of nutritional information, which includes the laborious process of manually recording this information, makes it impossible for people to estimate and measure the amount of food they consume. Because of this, having a system in place for tracking and measuring the calories in meals is quite beneficial.

II. LITERATURE RIVIEW

The process of food recognition has been studied extensively using a variety of methodologies. We have carefully examined a few of the previous approaches for identifying foods that were offered. Color, surface, slope, and filter highlights are collected from food images in the studies by Joutou et al. 2009 and Hoashi et al. 2010, and a partitioned classifier is created for each feature. The authors of a 2018 research report by Subhi et al.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6935
***Article Received**: 27 March 2023 **Revised**: 14 May 2023 **Accepted**: 28 May 2023*

created a model that uses CNN algorithm to recognize food items. They used the FOOD dataset to classify food objects. Furthermore, they asserted that they have developed very complex convolutional networks for the classification of food images. The findings demonstrate the value of network depth in the training of visual representations.

III. PROPOSED METHOD

A. Dataset Collection:

We used over 1000 photos from 7 distinct food classes to create the dataset for food detection. Before training our system, we verified that the correct photos were present in each of the food groups. Our dataset's 7 food classes are as follows:

- Apple, Banana, Carrot, Cucumber
- Onion, Orange, Tomato

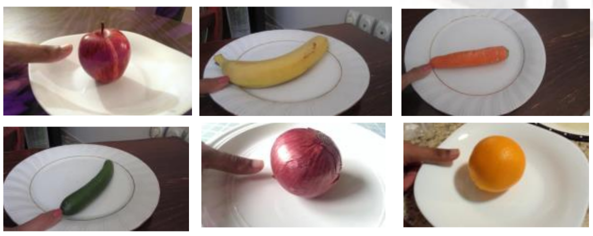
Table 1: Different Food Categories

Category	Camera	Lighting
1	Samsung-S4	Light Environment
2	Samsung-S4	Dark Environment
3	IOS-4	Light Environment
4	IOS-4	Dark Environment
5	CanonSD1400	Light Environment
6	CanonSD1400	Dark Environment

In this project we used 7 food items like apple, banana, carrot, cucumber, onion, orange and tomato

Table 2: Food Nutritional Information

Food	Density (g/cm^3)	Calorie (kcal/g)	Shape
Apple	0.60	0.52	Sphere
Banana	0.94	0.89	Cylinder
Carrot	0.64	0.41	Cylinder
Cucumber	0.64	0.13	Cylinder
Onion	0.51	0.4	Sphere
Orange	0.48	0.47	Sphere
Tomato	0.48	0.18	Sphere



B. Data Preprocessing:

We had to perform the following actions for the dataset we used to recognise several items from an image: loading certain necessary libraries, initialising directories, and resizing the images prior to extracting features. To begin the data preprocessing, we imported the os, tensorflow, numpy as np, glob, matplotlib.pyplot as plt, etc. libraries. Following the import of these libraries, we established the dataset directory so that the model could access the images, begin resizing them, and then begin extracting features. We uploaded our entire dataset to Google Drive, where it was organised in a folder called FooDD. Our photos were then resized to have the following dimensions: 299 pixels in height, 299 pixels in width, and 20 classes. The images were resized before we used our software to extract features.

C. Convolutional Neural Network (CNN) :

CNN also referred to as a convolutional neural network. It primarily employs a pooling technique to decrease the processing time or energy needed to process the data by decreasing the dimension. Moreover, CNN employs a multi-layer system, with the first layer's output serving as an input for the second layer, the second layer's output serving as an input for the third layer, and so on. CNN flattens the image into column vectors after preparing it for multilayer perceptron. These flattened images are sent into a neural network that we refer to as a feed forward network, and we then utilise backtracking and back propagation to analyse and discover mistakes. This process is continued throughout training. Back propagation travels back to the inner layer or hidden layer after computing the mistake to alter the weight in an effort to reduce the error. Until until the algorithm generates the desired output, this procedure is repeatedly repeated through back propagation.

IV. METHOD AND MODEL

The convolutional layer, the pooling layer, and the fully connected layer make up the convolutional neural network's hidden layer.

1. Convolutional layer

It is possible to extract features from input data using the convolutional layer, which several convolution kernels. Each component of the convolution kernel corresponds to a weight coefficient and a bias, just as the neurons in feedforward neural networks. Several neurons close to the layer above are connected to each neuron. In order to produce accurate results, the convolution kernel will periodically scan the input features, multiply and sum the input features by matrix elements, and add the bias

2. The pooling layer:

The pooling layer will carry out feature selection and information filtering of output feature mapping following the convolutional layer's feature extraction. The feature map governed by the pool size, step size, and filling is scanned by

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6935
***Article Received**: 27 March 2023 **Revised**: 14 May 2023 **Accepted**: 28 May 2023*

the convolution kernel concurrently with the choice of the pooling layer.

3. The fully connected layer:
The fully connected layer occupies a comparable place to the hidden layer in a conventional feedforward neural network in the convolutional neural network. Only the other fully connected layer receives signals from the final portion of a convolutional neural network's hidden layer. The feature map converts the spatial topology into a vector and transfers the activation function in the fully linked layer.

4. Output Layer:
The output layer of a convolutional neural network is typically upstream of the fully connected layer, thus its structure and operation are comparable to those of the output layer of a conventional feedforward neural network. The output layer in this project produces the object's centre coordinates, size, and classification.

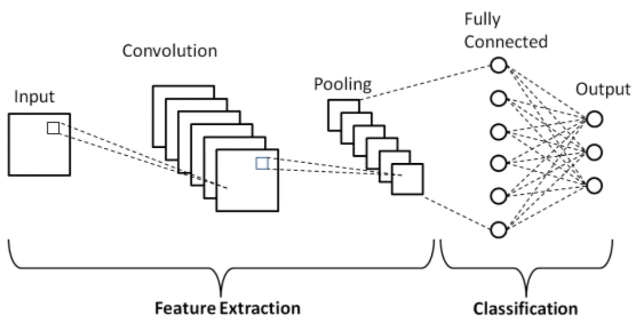


Figure 1: Convolution Neural Network

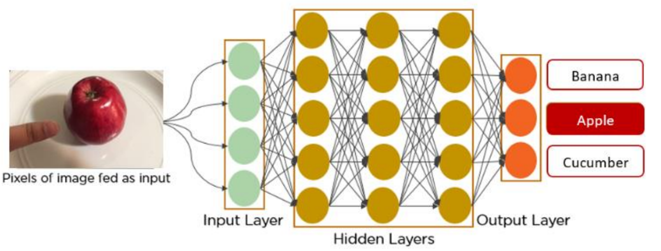
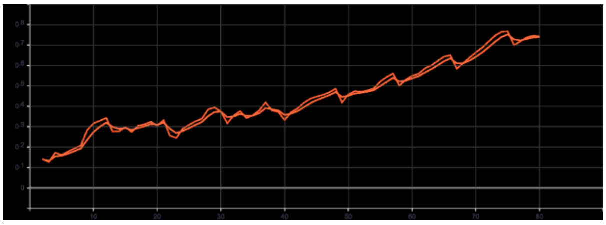
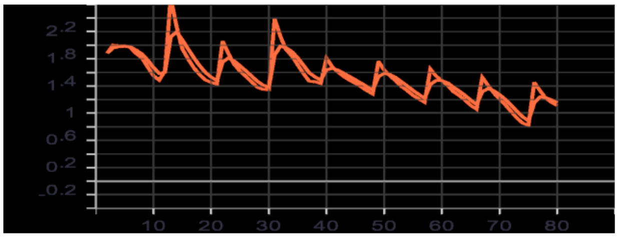
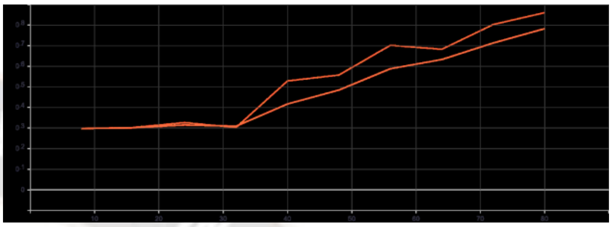


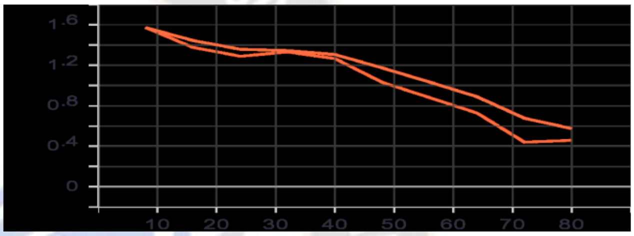
Figure 2: Example of Convolution Neural Network


Accuracy of Training



Loss of Training


Accuracy of Validation



Loss of Validation

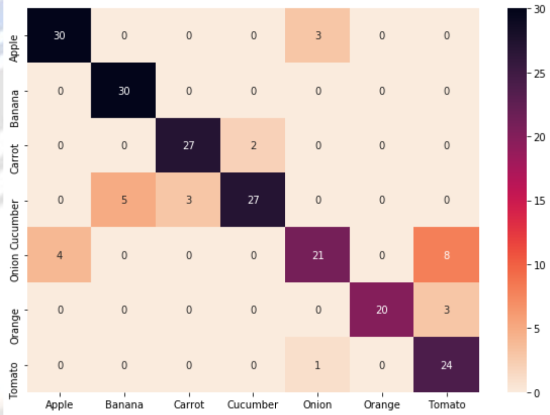
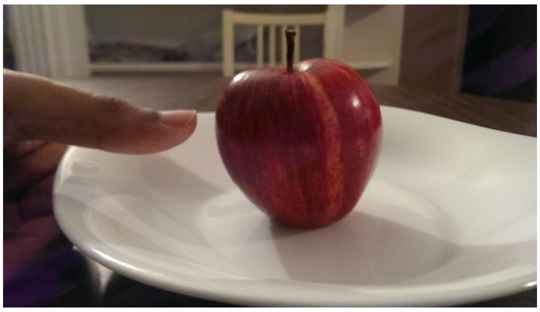


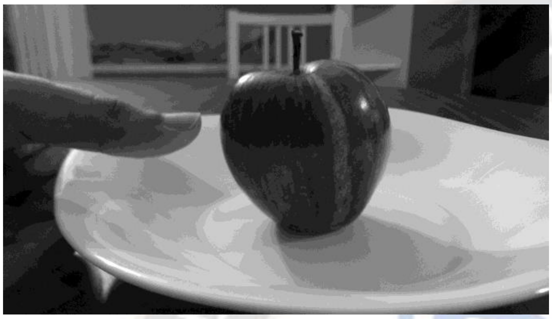
Figure 3: Confusion Matrix

Image Segmentation:
Image Segmentation procedure given below:
We use the thumb finger for calibration purposes. The thumb is placed next to the dish while clicking the photo and this thumb gives us the estimate of the real-life size of the food item and helps estimate volume accurately.

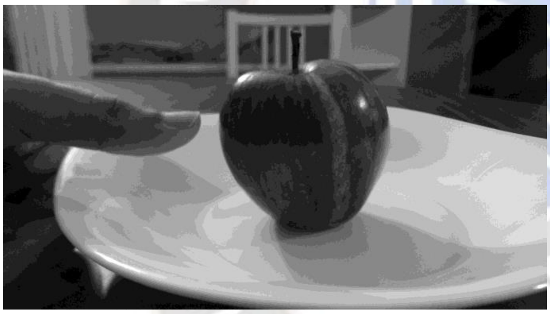
International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6935
***Article Received**: 27 March 2023 **Revised**: 14 May 2023 **Accepted**: 28 May 2023*



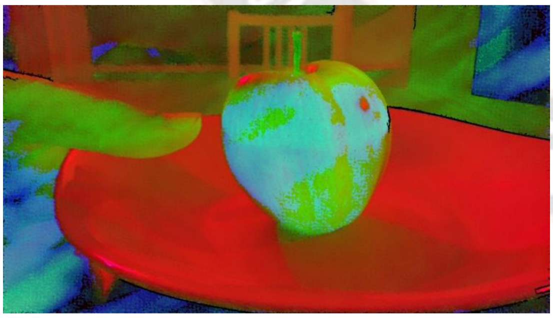
Original Image



Gray Image



Filter Image

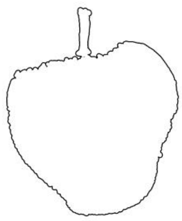


HSV Image

Food Skin Image

Food Structure Image

Final Food Image



Area of Food

rea of food image:

We have 3 factors from image segmentation

1. Foods pixel area
2. Skin pixel area
3. Actual skin area (skin multiplier)

From this factors food estimated area is given below:

$$\textit{Estimated Food Area} = \frac{\textit{Foods Pixel Area} * \textit{Actual Skin Area}}{\textit{Skin Pixel Area}}$$

We have two type of shape of foods
1. Sphere - like apple, orange, tomato, onion
2. Cylinder – like banana, cucumber, carrot

Volume estimation for Sphere:

$$Estimated\ Radius = \sqrt{\frac{Estimated\ Food\ Area}{\pi}}$$

$$Estimated\ Volume = \frac{4}{3} * \pi * Estimated\ Radius^3$$

Volume Estimation for Cylinder:

$$Estimated\ Height = Pixel\ Height * Pix\_To\_Cm\_Multiplier$$

$$Estimated\ Radius = \frac{Estimated\ Food\ Area}{2 * Estimated\ Height}$$

$$Estimated\ Volume = \pi * Estimated\ Height * Estimated\ Radius^2$$

Weight and Calories Estimation of Food:

From table we know food density (g/cm^3) and food calories (kcal/g), using this information we can estimate weight and calories intake in given food.

$$Estimated\ Weight = Actual\ Density * Estimated\ Volume$$

$$Estimated\ Calories = \frac{Estimated\ Weight * Calories\ Per\ 100\ Gr}{100}$$

V. RESULT

Food	Mass	Calories
Apple	43.75	22.75
Banana	147.62	131.38
Carrot	104	42.64
Cucumber	212.56	27.63
Onion	34.94	13.97
Orange	31.38	14.74
Tomato	22.64	4.07

VI. CONCLUSION AND FUTURE SCOPE

The main goal of this study is to identify the calorie of a food from a given image, which will help people overcome diseases caused by obesity, such as diabetes, heart disease, kidney failure, and so on. We believe that educating people about food calories will enable them to live a healthier lifestyle by keeping track of how many calories they consume. As a result, we researched and investigated various methods for food recognition. CNN is best suited for image or object detection tasks. On traditional 2D or 3D image recognition tasks, as well as other object detection tasks, CNN outperforms other neural networks, machine learning, and deep learning algorithms in terms of performance. They also have a high calculating efficiency in terms of object or image classification systems in terms of statistical results. At this point in our research, we are concentrating on classifying different types of foods and determining their nutritional value. We discovered the following limitations in our system: Similar-looking food items pose a problem for multiple food detectors because our system occasionally detects multiple food items from a single image. Finally, the angle at which the image is captured is critical for our system to accurately detect food items. Typically, our system performs poorly for images with a top view.

REFERENCES

[1] Koo, H. C., Joo, N. S., & Kim, H. G. (2020). Deep Learning-Based Calorie Estimation for Food Images. Journal of Food Science, 85(10), 3006-3014.

[2] Qi, Y., Liu, S., & Zhang, X. (2021). Calorie estimation of food using a deep learning approach. International Journal of Food Science & Technology, 56(5), 2365-2372.

[3] Kawa, M., Siemiątkowska, A., Zięba, K., & Szóstak, M. (2021). Calorie Estimation of Meals from Food Images with Convolutional Neural Networks. Nutrients, 13(1), 143.

[4] Hu, Y., Cai, Q., Shi, L., & Zhao, Y. (2021). Calorie Estimation for Mixed Dishes Using a Multitask Deep Learning Model. IEEE Journal of Biomedical and Health Informatics, 25(11), 3738-3746.

[5] Parisa Pouladzadeh, Abdulsalam Yassine, and Shervin Shirmohammadi, "Foodd: An image-based food detection dataset for calorie measurement," in International Conference on Multimedia Assisted Dietary Management, 2015

[6] P.Pouladzadeh, S.Shirmohammadi, and R.Almaghrabi, "Measuring Calorie and Nutrition from Food Image", IEEE Transactions on Instrumentation & Measurement, Vol.63, No.8, p.p. 1947 – 1956, August 2014.

[7] J. Dehais, M. Anthimopoulos, and S. Mougiakakou. Food image segmentation for dietary assessment. In *Proc. of ACMMM Workshop on Multimedia Assisted Dietary Management*, 2016.

[8] T. Ege and K. Yanai. Simultaneous estimation of food categories and calories with multi-task cnn. In *Proc. of IAPR International Conference on Machine Vision Applications (MVA)*, 2017.

[9] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

[10] R. Girshick. Fast R-CNN. In *Proc. of IEEE International Conference on Computer Vision*, 2015.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In

[12] *Proc. of IEEE Computer Vision and Pattern Recognition*, 2014.

[13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg. SSD: Single shot multibox detector. In *Proc. of European Conference on Computer Vision*, 2016.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6935
***Article Received**: 27 March 2023 **Revised**: 14 May 2023 **Accepted**: 28 May 2023*

[14] Y. Matsuda, H. Hajime, and K. Yanai. Recognition of multiple-food images by detecting candidate regions. In *Proc. of IEEE International Conference on Multimedia and Expo*, 2012.
[15] Chalasani , R. D. ., & Radhika Y. (2023). Application of Artificial Neural Networks to Forecast ITK Inhibitor Activity Data. International Journal of Intelligent Systems and Applications in Engineering, 11(1), 70–78. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2445.
[16] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You Only Look Once: Unified, real-time object detection. In *Proc. of IEEE Computer Vision and Pattern Recognition*, 2016.