Proceedings of the 7th International Conference on Computing Methodologies and Communication (ICCMC-2023)
IEEE Xplore Part Number: CFP23K25-ART; ISBN: 978-1-6654-6408-6

Calorie Estimation of Food and Beverages using Deep Learning

1st Dr. Parimala Gandhi A, AP
Department of ECE
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore, India
parimalagandhi@yahoo.com

2nd Ms. Sapna S
Department of ECE
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore, India.
kit.23.19bec097@gmail.com

3rd Mr. Yaswanth K M
Department of ECE
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore, India.
kit.23.19bec125@gmail.com

4th Mr. Praveen Kumar M
Department of ECE
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore, India.
kit.23.19bec079@gmail.com

2023 7th International Conference on Computing Methodologies and Communication (ICCMC) | 978-1-6654-6408-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCMC56507.2023.10083648

Abstract—Obesity, a serious chronic disease, is on the rise as a result of how easily food can be brought to our door steps. People's need for food grew, and at the same time, their anxiety about their nutrition also grew. This study offers an image-based calorie estimation system that asks the user to upload an image of a food item in order to calculate the estimated number of calories in the image. It is a multitasking system that displays weekly information on a user's calorie consumption and the number of calories that must be ingested to prevent obesity-related illnesses like cancer, heart attack, etc. To recognize complex pictures, a collection of food images with 20 classes and 500 images are built in each class. This study has developed a six-layer Convolutional Neural Network (CNN) architecture for the purpose of extracting the traits and classifying the images. The proposed food identification trials had an accuracy of 78.7% during testing and 93.29% throughout training. By using software designed to accurately estimate food calories from still images, users and healthcare experts may be able to more rapidly detect dietary practices and food choices connected to health and health concerns. Calorie calculation has been done by using photographs, however, it is difficult, and there is presently no publicly available program that can conduct both food estimation using images and provide health information about the individual.

Index Terms- Image Segmentation, Convolutional Neural Network (CNN) algorithm, Web application, Datasets.

I. INTRODUCTION

The identification of the food and the presentation of its related calorie content are the main goals of this food calorie detection system. Food is recognized by the program when it is placed in front of the camera, and its name and also nutritional facts are shown. Additionally, by measuring the user's caloric intake, it will offer some dietary advice to keep them healthy.

A number of image processing and classification algorithms are used to categorize the meal, calculate the volume, and determine the nutritional content. A number of methods, including canny edge detection, watershed segmentation, morphological operators, and Otsu's approach were used to segment the food item in order to obtain the contour of the fruit. The segmentation of the food item was done by using Otsu's approach. For calibrating, use the thumb finger. As the photo is being taken, the thumb is placed next to the plate, giving us a sense of the food item's true size and assisting in accurately estimating volume.

The food item is then taken out of the image, and the feature vector is then extracted for use in training and testing. The feature vector is produced by using the area for size, the HSV histogram for color features, Gabor filters for texture features, moments for form, and the HSV histogram for color features. The size of the feature vector is 95 by 1. The Support Vector Machine model was trained on the images using our 95-dimensional feature vector. Our algorithm successfully categorized 94% of the food products.

As people's understanding of the need of a balanced diet has expanded, there is a growing need for autonomous systems that can identify food and beverages. Because they enable both the automated identification of food and drink items as well as the evaluation of their nutritional contents, such systems are particularly useful for dietary assessment and planning. By preventing nutrition-related illnesses, this is true for both healthy people and patients with a variety of dietary limitations.

The issue with drink identification is that images of drink products can only provide a limited amount of information. Examples of this information include the color, brightness, and density of the beverage. The detection and recognition of food and drink images is a particularly tough computer vision problem as a result of all of these challenges. We used deep learning or deep neural networks to overcome this issue. For many computer vision problems, manually defining intricate features takes a lot of time and effort.

Deep learning provides a solution to this issue by allowing computational models with several processing layers to automatically recognize certain qualities and utilize them to represent incoming data. Computer vision is one of the study fields where deep learning models have significantly improved the best outcomes.

We choose deep convolutional neural networks in particular, because they are a subclass of deep neural networks inspired by animal visual cortexes, where individual neurons respond to overlapping parts of the visual field, deep convolutional neural networks are the ones we select in this instance. Convolutional neural networks are particularly well-suited for computer vision since the objective of computer vision systems is the same as the goal of animal vision systems—to interpret input pictures. As information moves through the network layers of a convolutional neural network, 3 numerous operations are carried out on the input image. To do the training, the layer parameters are then frequently modified. The bulk of layers in convolutional neural networks are convolutional, fully-connected, and pooling layers.

Convolutional layer learnable filters are trained to respond to certain features in the input data. We were able to categories 520 distinct food and drink products across a wide variety of food categories using this architecture with an accuracy of 86.72% on a recognition dataset of 225,953 512 512-pixel images and On a detection dataset of 130,517 pictures, 94.47% were detected. On a collection of photos that were both self-acquired and taken, we also conducted a real-world test by people with Parkinson's disease using a smartphone camera. According to the findings, the top-five accuracy for real-world photos was 55%.

II. EXISTING SYSTEM

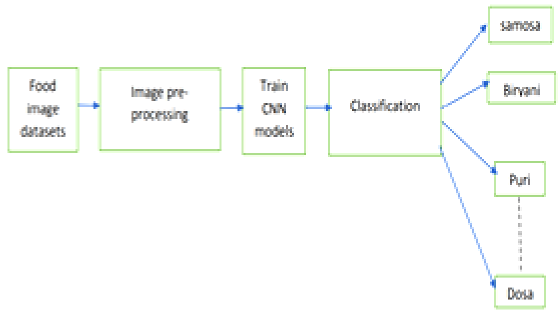


Fig-1: Food image classification

Using Google, the sample in this block was manually downloaded from Kaggle, and some of the train and test pictures included noise, had different color intensities, and had images with the wrong tags. For storage and suitable train and test phases, we additionally downscaled the images to 50x50 pixels.

It becomes clear how tough it is to categories foods because of all the many types that may be found in the actual world. Given the size and variety of the collection, it will be difficult to identify all the different cultures in the sample. The better choice, in many eyes, is to utilize neural networks. The potential of neural networks to absorb patterns that aren't immediately linear-separable is the main reason behind Option 9's success in overcoming scaling issues. It is able to deal with additional environmental factors including noise pictures and much more.

The image-net database may be accessed and utilized by many people. A dataset that may be utilized for picture categorization has undergone extensive CNN training. It now contains a large number of classification categories In order to generalize the system model, the Kaggle dataset was manually downloaded via Google to the CNN model. The model's attributes also include the following: Max Pooling downscale is used to downsize the input photos to 50x50 in each spatial dimension. 0.4 dropout rate with the SoftMax activation function.

Table – 1: Classified food images

Class Number	Class Name
1	Biryani
2	Bisibelebath
3	Butter naan
4	Chaat
5	Chapatti
6	Dhokla
7	Dosa
8	Idly
9	Noodles
10	Upma
11	Poori
12	Samosa

This study presents the findings and analysis of our suggested strategy for the performance evaluation technique for food categorization. We used the Python programming language, an 8 GB RAM system, an Intel i5 CPU, and the Windows 11 operating system to simulate the model repeatedly. Model assessment having the capacity to load and evaluate the models with the greatest accuracy and the least amount of loss. We also acquired the graph accuracy across the 12 classes (As shown in Table-1). There are a total of 12 courses, and there are 709 training photos and 124 testing images in each class. After testing, it will show the output as indicated in Figure 3 of the figure.

Fig-2: Output of existing system

III. PROPOSED SYSTEM

The size, shape, colour, and texture of the food are used in this way to identify it. Based on the input, a system is developed that can recognise food. The technology also helps in calculating how many calories are in the dish. The stages of image processing that produce the competent model, which, utilising the training dataset, can categorise any image, include pre-processing and neural network training. The proposed approaches enable automatic food identification and calorie estimation [31].

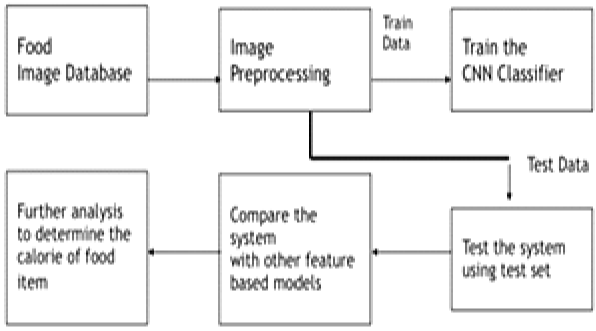


Fig-4: Working process of proposed system

IV. METHODOLOGY

4.1 IMAGE ACQUISITION AND PREPROCESSING:

It uses the 90483 photos of 131 fruits and vegetables from the Fruits 360 collection. In our project, we take into account 15 different kinds of produce. The photos are scaled from their original 100x100 proportions to 224*224. When photos are resized, they are transformed into 4D tensors with the shape (1, 224, 224, 3), which may subsequently be fed into a CNN for learning.

4.2 NEURAL NETWORK TRAINING:

Convolution neural networks (CNN) seem to be the approach utilized, while deep learning works as the model's theoretical foundation. The model is being built entirely from scratch. the shared-weights design and translation invariance features of CNN, sometimes referred to as Convent, a subclass of deep neural networks. Using the numerical technique of fourier is CNN. The linear computation subset covers convolution. Between both the layers of input and output a CNN is built up of numerous hidden layers.

The network-based model now proceeds to process the input array of dimension 4 data. The number of filters was increased since achieving a flawless result with fewer mistakes was the primary objective. We originally started with 16 filters and left padding at its default setting to prevent data loss. Then, in order to reduce the breadth of our data, we employed the Max Pooling layer with pool size=2. After that, 128 filters were added in the following order: In order to categorise and delve further, dropout layers of 16, 32, 64, and 128 were utilised to lessen the chance of overfitting. The CNN layer is linked to the flattering layer via the flattering layer. The concealed layers are then connected using the "relu" triggering mechanism.

The SoftMax function is used in the model's final layer, which contains 15 nodes because there are 15 different kinds of food, to calculate the chance of each type of fruit appearing in the image. The Root Mean Square Propagation (RMS Prop) optimizer, which employs a moving average of squared gradients that normalises itself, is a component of the Sequential CNN model. The category cross entropy loss is also taken into consideration by the model.

4.3 SEGMENTATION OF IMAGE:

To make the depiction of a photograph more understandable and easier to analyse, segmentation is a technique that divides a computer-generated image into a number of segments. An edge detection approach is used to reduce a grayscale picture to just black and white pixels. The numbers between 1 and 254 indicate different tones of grey. A graphic value of 0 typically symbolises white, whereas a value of 255 often represents black. By counting the contours and identifying the greatest contour, a division dependent on contours is accomplished. To convert the photograph the fruit pixels are extracted after using HSV to exclude the plate and plate-related pixels from the picture. After that, we set the two fundamental morphological processes of erosion and dilation into action. Erosion was employed to remove pixels from edges and corners before refraction was applied to add units to the item. The meal's location was then established. After that, we see a picture of the fruit by itself (in this case apple). To acquire the food area, a number of morphological processes must be conducted for various foods several times. To compute area, pixels are converted to square centimetres. To do this, we divide the area by the surface area of the skin and multiply the result by a certain amount.

4.4 CALORIE ESTIMATION:

Food may be photographed at various depths to yield images of varied sizes. In a practical environment, we need a way to determine the amount of food or the number of calories. We need to know the real object sizes once the required food items have been recognised together with their masks because it is hard to do so from a pin-hole camera picture alone. Therefore, by comparing the food-objects to the size of the previously known object, we employ a referencing strategy to extract the true size of the food present in that particular image. In the same manner that the aforementioned experiment employed a coin as a reference object, we propose employing a plate as a standard object for the assessment of the food identified in photographs. Edge detection or training data may be utilised to recognise plates and edibles with a single network, respectively. After the plates have been discovered, the pixels per square inch are calculated using the actual size of the plate in real life. Once a meal's portion has been identified, its volume has been calculated while accounting for all of its different forms. In this instance, it was an apple. You may compute its first radius by multiplying it by pi (3.14). The sphere's volume is then calculated using the formula.

V. DATASET

Table – 2: Dataset table

Fruits	Density	Label	Shape	Calorie
Apple	0.609	1	Sphere	52
Banana	0.94	2	Cylinder	89
Carrot	0.641	3	Cylinder	41
Cucumber	0.641	4	Cylinder	16
Onion	0.513	5	Sphere	40
Orange	0.482	6	Sphere	47
Tomato	0.481	7	Sphere	18

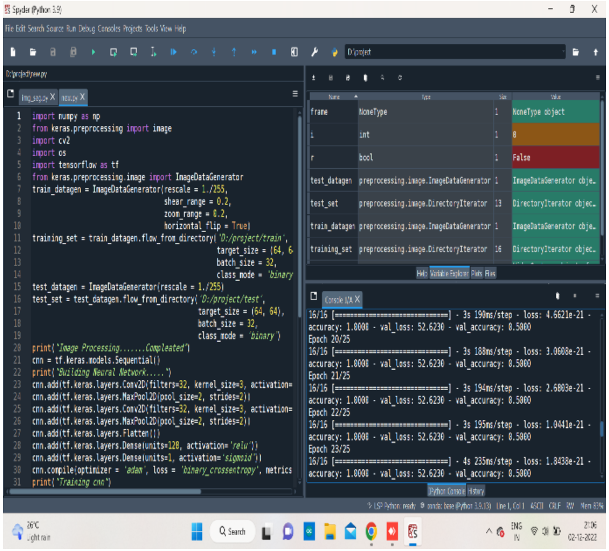
The photos that cell phones took, collectively. For this collection of fruits, four categories have been created. Four fruit types and 2403 real-world photos are offered in this difficult data collection. The images were taken at various angles and from a range of fruit stores. Each of the diverse yet furthermore outwardly and semantically equivalent fruit classes has a total of 565 photos, of which 100 are training images and 465 are test images.

VI. TESTING

The testing is done using tickle software . The first step in testing is by taking the screenshot and extract its shape, color, size etc.. and converted as matrix. By using this it will conform whether it is feasible or not. After that the dataset which we give as a input will compare with the trained dataset.

VII. RESULT AND DISCUSSION

In this study, we trained our programme to test the viability of automatically predicting the kind of takeaway meals based only on a dataset using machine learning techniques. Using labelled training data from an online takeaway food and beverage delivery service, our objective is to develop a model that reliably predicts meal type around three out of four times. Six out of ten food and drink categories had memory rates of more than 71%, and eight out of ten had prediction accuracy of more than 65%. ("excellent"). Multi-fast-food cuisines, as well as South Asian, Southeast Asian, and East Asian cuisines, have the highest rates of prediction accuracy. The lowest recall (44%) and accuracy (54%) were seen at burger establishments. Low recall was created by the most common misclassification of burger outlets as meal categories, while low precision was brought on by the frequent misdiagnosis of restaurants, specifically chicken businesses, as burger outlets. In this case, machine learning was applied since the model beat a crude classification technique that only used keywords.



VIII. CONCLUSION

With the use of this software, we were able to identify and calculate the number of calories in meals after having CNN film it. We produced the fruit photo dataset for this investigation. Additionally, we used this food detector to estimate food calories in order to ascertain the calorie content of a particular fruit. We also produce a dataset of fruit image annotations for calories. As a consequence, we calculated meal calories by photographing the dish. We intend to multitask CNN with food detection and calorie estimates in future work. We anticipate that multi-task learning will increase each task's accuracy. Additionally, we intend to crop higher-resolution photos to match the CNN-estimated bounding boxes and utilize those images to estimate calories more precisely.

IX. ACKNOWLEDGEMENT

We gratefully acknowledge Dr. Parimala Gandhi A, Assistant Professor Department of Electronics and Communication Engineering at [KIT-Kalaignar karunanidhi Institute of Technology, Coimbatore, India] for her contribution to the project in terms of information, expertise, contacts, guidelines and financial assistance. Her ability to motivate, unfailing encouragement, and meticulous oversight throughout the research assist us successfully finish our task.

REFERENCES

[1] G. Eason, B. Nonle, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel function," Phil. Trans. Roy. Soc. London, vol. A247, pp.529-551, April 1955.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp.271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[8] https://www.pantechsolutions.net/food-recognition-using-opencv-andpython/

[9] https://www.ijert.org/fruit-recognition-system-for-calorie-management

[10] https://www.lftechnology.com/blog/ai/image-calorie-estimation-deeplearning/

[11] https://datacarpentry.org/image-processing/aio/index.html

[12] Dr. N. Kumareshan, and et.al,, "M-Chord and Mr-Chord Protocol for High-Throughput Hybrid Wireless Networks" International Journal Of Information And Computing Science Volume 5, Issue 11, November 2018

[13] Dr.N.Kumareshan and et.al,, "Implementing P2p Resource Sharing Applications In Wireless Mesh Networks" Int Jour Of Scientific Engineering And Applied Science (IJSEAS) – Volume-3, Issue-8, August 2017.

[14] Dr.N.Kumareshan, and et.al,, "Performance Analysis Of Various Routing Protocols For Vanet Environments" In Int Jourl Of Innovative Technology And Exploring Engineering (Ijitee) ISSN: 2278-3075, Volume-8 Issue-12, October 2019

[15] Dr. M. Rajalakshmi and et.al" Optimized Energy Aware Routing Based on Suitable Based Antlion Group with Advanced Algorithm (SA-AOA) in Wireless Sensor Network"wireless personal communication volume 113, issue 1, july 2020 (pp:59-77).

[16] N. Prakash and et.al, "Arduino Based traffic congestion control with automatic signal clearance for emergency vehicles and Stolen Vehicle Detection", 2020 Int. Conf. on Couter Comm and Informatics, Coimbatore, 2020, pp. 1-6

[17] Dr. M. Rajalakshmi and et.al" Analysis of QoS for Conveying Authorisation Based on Internet of Things (IoT) in Wireless Sensor Networks (WSN)" 7th International Conference on Smart Structures and Systems (ICSSS), 1-9, 2020

[18] [Dr. M. Rajalakshmi and et.al" Throughput Enhancement in Wireless Sensor Network by Novel Genetic Algorithm-Fuzzy Method (NGA-F)" Recent Trends in Electronics and Communication,273-287

[19] E. Udayakumar and et.al, "GSM-Based Design and Implementation of Women Safety Device Using Internet of Things" Advances in Intelligent Systems and Computing 1167, 169-176

[20] E Udayakumar and et.al, "Design and development of Android based Plant disease detection using Arduino"2020 7th International Conference on Smart Structures and Systems (ICSSS), 1-6

[21] [10] E Udayakumar and et.al, "Implementation of Smart Control of Wheelchair for a Disabled Person"Internet of Things and Its Applications, 137-150

[22] [Prakash. N and et.al, "Light Intensity RTC Based Solar Tracking for Power Conservation in Irrigation System" international journal of engineering research and technology Issue 1 (January - 2014) Pages787-792

[23] K Srihari and et.al, "An IoT- Based Gadget for Visually Impaired People"Integration of Cloud Computing with Internet of Things: Foundations, Analytics, and Applications John Wiley & Sons, Inc. Pages 71-86

[24] N Prakash and et.al, "Implementation of Smart Control of Wheelchair for a Disabled Person", Internet of Things and Its Applications, 137-150

[25] NA Vignesh and et.al, "Implementation of Routing Protocols for Predicting Energy Consumption and route lifetime for Improving network Lifetime in Mobile Adhoc Networks "Solid State Technology 63 (5), 9391-9396

[26] PN Kumareshan.N and et.al, "Improving QoS and Reducing Energy Consumption in Cluster-Based VANET Routing Protocol"International Journal of Recent Technology and Engineering (IJRTE) 8 Issue-4 Pages 4522-4525

[27] https://www.analyticsvidhya.com/blog/2014/12/image-processing-pythonbasics/

[28] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8234200/

[29] https://www.ncbi.nlm.nih.gov/pm

[30] https://www.hindawi.com/journals/cin/2021/1268453/c/articles/PMC5537777/

[31] Attokaren, David J., Ian G. Fernandes, A. Sriram, YV Srinivasa Murthy, and Shashidhar G. Koolagudi. "Food classification from images using convolutional neural networks." In TENCON 2017-2017 IEEE Region 10 Conference, pp. 2801-2806. IEEE, 2017.