IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

INDIAN FOOD RECOGNITION AND CALORIE ESTIMATION USING YOLOV8

[1]V.Gayatri, [2]M.Thanuja

Abstract: Obesity and Overweight is a common and serious disease in most of the adults and children, which refers to abnormal or excessive accumulation of fat that causes serious health problems. A study in 2019 around 38.2 million children under the age of 5 were obese or overweight. There is a rise in the number of overweight children under the age of 5 in Africa by nearly 24% since 2000. Nearly half of the children under the age of 5 in Asia were overweight or obese in 2019. The best way to treat obesity is to follow a reduced-calorie diet. Manually tracking the food and calories in our food is a tedious task. There are Convolutional Neural Network (CNN) models such as InceptionV3 and InceptionV4 to recognize the food items. But the existing CNN-based models are less accurate, computationally expensive and require high computational power. Here we are proposing a Convolutional Neural Network algorithm called YOLOV8, which is mostly used for object detection, to detect the food items. The goal of this project is to develop model which recognizes the food item and estimates the calorie values in it. This system provides a platform where users can estimate the number of calories in a food item by just providing picture or videos of the food item and hence they can track the number of calories in their diet.

***Index Terms* -healthy diet, nutrition, claorie count, computer vision, yolo model, deep learning.**

I. Introduction

Food is essential in our everyday life. The amount of calories in the food has a huge impact on our weight. Since manually tracking the amount of calories in the food is a difficult task there is a need to design a system that estimates the number of calories in the food that we consume. The main goal of this thesis is to design a model that recognizes the food items on the plate more accurately and estimates the number of calories in each food item per gram using Convolutional Neural Networks (CNN). In the previous work, a CNN-based model called InceptionV3 is used to recognize the food items and an API from calorie ninjas is used to estimate the number of calories in each food item. In our work, we have found that the existing model requires high computational power, expensive and slow. Such an understanding could potentially pave the way for a more accurate, fast, cheap, and requires less computational power. We have described a CNN-based model called YOLOV8 which detects the object in a single stage with more accuracy and is faster than the previous model.

II. Literature review

The literature survey on Indian food recognition and calorie estimation highlights the current state of research in the field. Various studies have focused on utilizing machine learning techniques, particularly Convolutional Neural Networks (CNNs), for food recognition and estimation of nutritional values. The aim is to assist individuals in achieving a balanced diet and managing their calorie intake, thereby combating the growing problem of obesity.

Several approaches have been explored in the literature. One study proposed the use of InceptionV3 and InceptionV4 CNN models for food recognition and estimation of nutritional values. They employed linear kernel SVM for classification and utilized an API containing nutritional data to estimate attributes such as vitamins, minerals, and fats. Another study proposed an efficient food image classification method based on InceptionV3 CNN. They trained their model on a dataset comprising real-time South Indian food images with various challenges such as noise, different color intensity, and incorrect labels. They employed techniques such as data augmentation and random crop size to improve efficiency and accuracy.

Other research focused on estimating the nutritional value of food in real-time using deep learning techniques. They employed CNN models like InceptionV3, InceptionV4, VGG-16, VGG-19, and ResNet for food recognition. Attribute estimation was performed using vector space representation and Word2Vec to estimate food attributes from a large dataset. Another study utilized deep learning for the segmentation and weight estimation of food images. They employed Inception networks for weight estimation and YOLO network for segmentation, achieving satisfactory results in calorie and weight estimation.

Overall, these studies emphasize the potential of deep learning techniques, specifically CNNs, for food recognition and calorie estimation. They provide insights into the design and development of mobile applications, utilizing image classification and volume estimation to assist individuals in managing their dietary intake and promoting healthier lifestyles. However, further research is needed to improve accuracy, overcome challenges such as inconsistent labeling and varying image conditions, and develop user-friendly interfaces for seamless food recognition and calorie tracking.

III. PROPOSED METHODOLOGY

In this paper we propose a novel CNN-based model to recognize food items with more accuracy and speed. Using a novel dataset of 5446 food item images of 30 different classes of food we are training and testing our model. The proposed system achieves a highest accuracy of 93.1 %. In this model we have used YOLOV8 to train the food item images and after training we can use this model to detect and recognize food items from new images as well as videos. YOLO has several advantages over other object detection algorithms. It is extremely fast, capable of processing images in real-time, making it suitable for applications that require quick responses. YOLO also has good generalization capabilities and can detect a wide variety of object classes. YOLO algorithm is typically trained using a large dataset with annotated bounding boxes and class labels. During the training process, the model learns to optimize the bounding box predictions and class probabilities based on the provided ground truth annotations. To train the YOLOV8 model we are using a dataset from Roboflow with a large set of images to get better prediction result.

The YOLO (You Only Look Once) algorithm is an object detection algorithm used in computer vision and deep learning. It was first introduced by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi in 2015. YOLO is known for its real-time object detection capabilities and has been widely adopted in various applications, including autonomous vehicles, surveillance systems, and robotics. The main idea behind YOLO is to treat object detection as a regression problem. Instead of using a sliding window or region proposal-based approach, YOLO divides the input image into a grid and predicts bounding boxes and class probabilities directly from the grid cells.

YOLOv8: YOLOv8 is the latest version of YOLO by Ultralytics. As a cutting-edge, state-of-the-art (SOTA) model, YOLOv8 builds on the success of previous versions, introducing new features and improvements for enhanced performance, flexibility, and efficiency. YOLOv8 supports a full range of vision AI tasks, including detection, segmentation, pose estimation, tracking, and classification. This versatility allows users to leverage YOLOv8's capabilities across diverse applications and domains.

Enhancements in YOLOV8:

The YOLOV8 is similar to other versions of YOLO with some special enhancements:

Anchor free detections

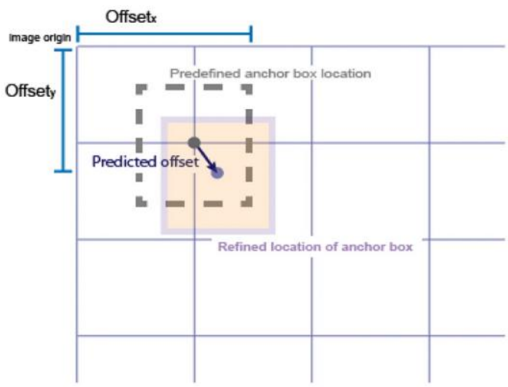

Fig 1. Anchor free detections

Anchor-free detection is when an object detection model directly predicts the center of an object instead of the offset from a known anchor box. Anchor boxes are a pre-defined set of boxes with specific heights and widths, used to detect object classes with the desired scale and aspect ratio. They are chosen based on the size of objects in the training dataset and are tiled across the image during detection. The network outputs probability and attributes like background, IoU, and offsets for each tiled box, which are used to adjust the anchor boxes. Multiple anchor boxes can be defined for different object sizes, serving as fixed starting points for boundary box guesses.

Working of YOLOV8 Algorithm

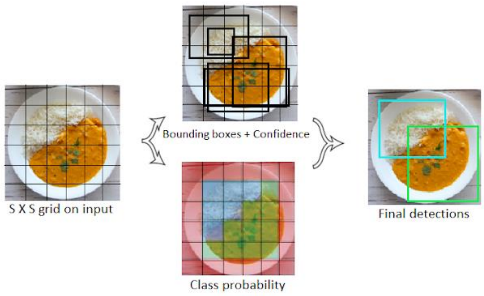

Fig 2. Working of YOLO Algorithm

Input Image: The algorithm takes an input image and resizes it to a fixed size suitable for processing. This size is usually determined based on the network architecture used.
Grid Division: The image is divided into an S x S grid, where each grid cell is responsible for predicting objects present in that cell.
Bounding Box Prediction: Within each grid cell, YOLO predicts multiple bounding boxes. Each bounding box is defined by five attributes: (x, y, w, h, confidence). The (x, y) coordinates represent the center of the bounding box relative to the grid cell, while w and h represent the width and height of the box. The confidence score indicates how confident the algorithm is that the box contains an object.
Class Prediction: Along with the bounding box predictions, YOLO also predicts the probabilities of different classes for each box. The number of class probabilities depends on the dataset being used. These class probabilities represent the likelihood of each class being present in the bounding box.
Confidence Thresholding: YOLO applies a confidence threshold to filter out low-confidence predictions. Bounding boxes with confidence scores below the threshold are discarded as false positives.
Non-Maximum Suppression (NMS): To eliminate duplicate detections and improve accuracy, YOLO applies non-maximum suppression. NMS removes redundant bounding boxes that have significant overlap and keeps only the most confident one. The overlap threshold for suppression is typically determined by a predefined Intersection over Union (IoU) value.
Final Output: The output of the YOLO algorithm is a set of bounding boxes, each associated with a class label and confidence score. These bounding boxes represent the detected objects in the input image.

We tried using yolo model to detect the food items and estimate the number of calories in them, to track the number of calories in our meal in order to avoid Obesity. Here we are using YOLOV8 object detection model to detect the food items. The implementation of the project using YOLOV8 is done as stated below:

Downloading dataset from Roboflow and importing YOLOV8

```
!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="JPYKmiIWJb9BVdXCc7yQ")
project = rf.workspace("indianfoodnet").project("indianfoodnet")
dataset = project.version(1).download("yolov8")
```

Fig 3: Downloading dataset from Roboflow and importing YOLOV8

TrainingYOLOV8model

```
!yolo mode=train model=yolov8m.pt data='/content/drive/MyDrive/food/IndianFoodNet-1/data.yaml' epochs=100
```

Fig 4. Training YOLOV8 model

Obtaining the weights file and using that file in the streamlit web application is the end point of the concept.

Streamlit is an open-source Python library used for building and deploying data-driven web applications. It simplifies the process of creating interactive and user-friendly interfaces for data analysis, machine learning, and visualization. Streamlit allows developers and data scientists to focus on the core functionality of their applications without worrying about the underlying web development aspects. Streamlit is widely used by data scientists, machine learning engineers, and developers to create interactive dashboards, data exploration tools, and machine learning prototypes. Its simplicity, rapid development workflow, and intuitive interface make it a popular choice for building data-driven applications.

The execution of the process will be explained clearly with the help of the continuous screenshots. The whole process includes the uploading the image which is the random image to identify the food item in the image and the number of calories in it.At first the model is trained to recognize the food item and then it will estimate the number of calories from a calorie-chart. It will be trained by using a dataset from Roboflow which contain 5446 images of 30 different classes of food items. The model is trained by taking each image into YOLOV8 Algorithm. After training with all the images in the dataset, it will generate a weights file. Secondly, the image needs to be provided to test, an interface to test. This interface is built using a python library called Streamlit which should be connected with the weights file. This will detect the presence of food items in the image along with the number of calories in each food item per gram.

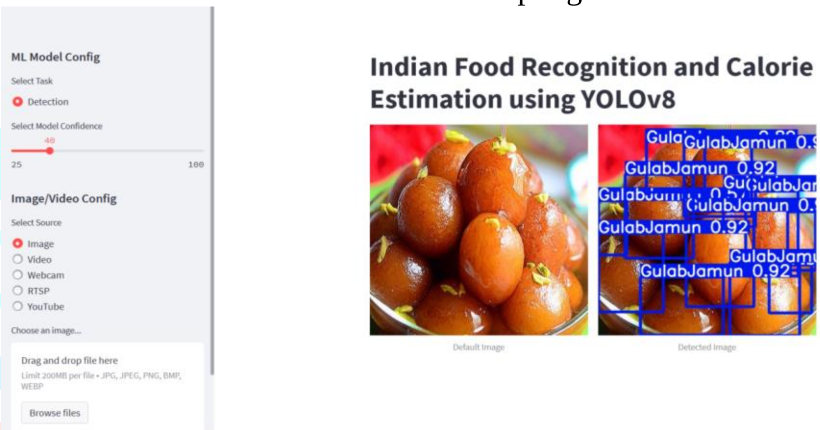

Fig 5. Streamlit webpage

Then we can see the food item in the image that we have uploaded is recognized by our model and the number of calories in the food item per gram is displayed.

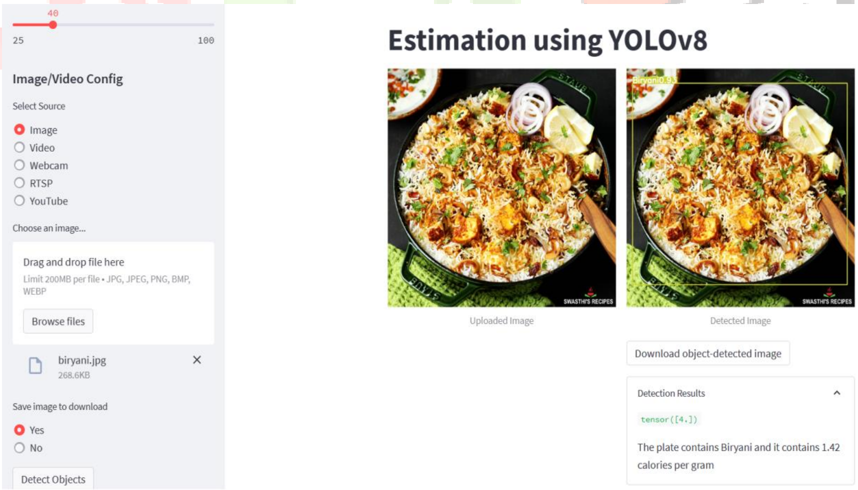

Fig 6. Food item recognition and calorie estimation

IV. CONCLUSION

In this paper, we have used YOLO(You Look Only Once)V8 object detection system to train food images dataset and after training we can use this model to detect the food items in a plate and estimate the number of calories per gram from a calorie-chart. This model has successfully detected 30 classes of food items as mentioned in the dataset and successfully estimated the number of calories in them per gram. Our accuracy is 93.1 when trained through 48 epochs. This model is fast and accurate than other CNN-based models as it detects objects in a single shot rather than two stage detection.

REFERENCES

[1] Raza Yunus, Omar Arif, Hammad Afzal, “To Estimate the Nutritional Value of Food”, December 2018.
[2] Manal Chokr, Shady Elbassouni, “Calories prediction from food images,” in Association for the advancement of artificial intelligence, 2017.
[3] Taichi Joutou and Keiji Yanai, “A food image recognition system with Multiple Kernel Learning,” in 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo,Egypt
[4] Vinay Bettadapura, Edison Thomaz , Aman Parnami, et al,”Leveraging Context to Support Automated Food
[5] Recognition in Restaurants,” in 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA.
[6] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine and S. Shirmohammadi, “ Food calorie measurement using deep learning neural network,” in IEEE International Instrumentation and Measurement Technology Conference Proceedings, Taipei, 2016
[7] Runar Isaksen, Eirik Bø Knudsen, Aline Iyagizeneza Walde “A deep learning segmentation approach to calories estimation of food images,”University of Agder, 2019.
[8] N. Martinel, G. L. Foresti and C. Micheloni, "Wide-Slice Residual Networks for Food Recognition," 2018 IEEE Winter Conference on Applications of Computer Vision(WACV), Lake Tahoe, NV, 2018, pp. 567-576, doi: 10.1109/WACV.2018.00068.
[9] Bossard L., Guillaumin M., Van Gool L. (2014) Food-101 – Mining Discriminative Components with Random Forests. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8694. Springer, Cham. https://doi.org/10.1007/978-3-319-10599-4_29
[10] Viswanath.C.Burkapalli,Priyadarshini.C.Patil,“An Efficient Food Image Classification by InceptionV3 Based nns,” in INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 03, MARCH 2020.
[11] Yash Gosalia, Meet Karnik, Aayush Pandey “ESTIMATION OF NUTRITIONAL VALUES OF FOOD USING INCEPTION v3” IN International Research Journal of Engineering and Technology (IRJET), Volume 08, Issue 03, March 2021.