

FOOD DETECTION AND CALORIE ESTIMATION USING YOLOV3 AND YOLOV5

Project report submitted in partial fulfillment of the requirement for the degree of

BACHELOR OF TECHNOLOGY IN ELECTRONICS AND COMMUNICATION ENGINEERING

By

Srinaga Sai Kartik Vinnakota (191043)

Sumant Kumar Singh (191036)

UNDER THE GUIDANCE OF

Dr. Shweta Pandit



**JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,
WAKNAGHAT**

TABLE OF CONTENTS

Serial No.	Caption	Page No.
I	Declaration	4
II	Acknowledgement	5
III	List Of Abbreviations	6
IV	List Of Figures	7
V	List Of Tables	8
VI	Abstract	9
1.	INTRODUCTION	10
	1.1 Introduction	10
	1.2 Study	11
	1.3 Problem Statement	12
	1.4 Motivation	13
2.	LITERATURE SURVEY	14
3.	PROPOSED SYSTEM MODEL	16
	3.1 Methodology	16
	3.2 Machine Learning Techniques	17
	3.2.1 Intersection Over Union	18
	3.2.2 Yolov3 Network	19
	3.2.3 CNN	20
	3.3 Algorithms Used	20
	3.3.1 Yolov1	21
	3.3.2 Yolov3	22
	3.3.3 Yolov5	23
	3.3.4 Working of Yolov3&v5	23

4.	PROBLEM ANALYSIS	24
	4.1 Product Definition	24
	4.2 Feasibility Analysis	24
	4.3 Technical Feasibility.	24
	4.4 Economic Feasibility	24
5.	SOFTWARE REQUIREMENT ANALYSIS	25
	5.1 Introduction	25
	5.2 Functional Requirement	26
	5.3 Non-Functional Requirements	26
6.	SYSTEM DESIGN	29
	6.1 System Analysis	29
	6.1.1 Flow chart	29
	6.2 Detailed design	30
7.	YOLOv3 IMPLEMENTATION	31
	7.1 Execution	32
	7.2 Result	33
8.	CONCLUSION	35
9.	FUTURE SCOPE & REFERENCES	38
	9.1 References	40

DECLARATION

We hereby declare that the work reported in the B.Tech Project Report entitled “**Food Detection And Calorie Estimation**” submitted at **Jaypee University of Information Technology, Wagnaghat, India** is an authentic record of our work carried out under the supervision of **Dr Shweta Pandit**. We have not submitted this work elsewhere for any other degree or diploma.

Srinaga sai kartik vinnakota, 191043

SumantSingh, 191036

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr Shweta Pandit

Date:

Head of the Department/Project Coordinator

ACKNOWLEDGEMENT

Firstly, we express my gratitude and gratefulness to almighty God for his divine blessings to make it possible to complete the project work successfully.

We are really grateful and wish our profound indebtedness to our supervisor **Dr. Shweta Pandit**, Assistant Professor (SG), department of ECE, Jaypee University of Information Technology, Waknaghat. Deep knowledge and keen interest of our supervisor in the field of Machine Learning to carry out this project . Her endless patience , scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, reading many inferior drafts and correcting them at all stages have made it possible to complete this project successfully.

We would also generously welcome each one of those individuals who have helped us straightforwardly or in a roundabout way in making this project a win. In this unique situation, we might want to thank the various staff individuals, both educating and non-instructing, which have developed their convenient help and facilitated our undertaking.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

LIST OF ABBREVIATIONS

Abbreviation	Word
YOLO	You Only Look Once
ML	Machine Learning
BMI	Body Mass Index
CNN	Convolution Neural Networks
SVM	Support Vector Machine
DGK	Darknet

TABLE OF FIGURES

Figure 3.1	System Architecture
Figure 3.2	Example of computing IOU for various bounding boxes
Figure 3.3	Yolov3 Structure
Figure 3.4	Yolov5 Algorithm
Figure 4.1	Object detection and Boundary Boxes
Figure 5.1	Functional Requirements
Figure 6.1	Flow Chart
Figure 6.2	Detailed Algorithm
Figure 7.1	Feature Selection Technique

LIST OF TABLES

Table 5.1	System List of loaded data-set
Table 5.2	Exploratory Data Analysis

ABSTRACT

As people become more aware of the advantages of a balanced diet, there has been an increase in the demand for automated food recognition systems. Such systems can not only predictably identify food products, but they can also estimate nutritional values, making them especially useful for dietary appraisal and preparation. This can benefit patients with various dietary requirements as well as healthier people by preventing nutrition-related diseases.

With the help of AR technology, we were able to develop our project in such a way that the user doesn't have to manually enter information about the nutrients and calories present in the food item they want to eat. Instead, a 3D image of the food item was provided, and the food item was detected, identified, and its nutrients and calories were stored in a database. This allowed the user to choose a healthy diet without having to manually enter information about the nutrients and calories present.

The main agenda of our model is, to determine the estimate the calories per food item. We used Yolov3 and Yolov5 as our primary algorithm. Our algorithm analyse, the image and create boundary boxes for every image and extract the unwanted image. The Previous dataset is used and according to that, our model analyses the image and produces the estimated calories. Thereby, he/she can understand, and use it for daily purpose.

This project has several advantages as well as disadvantages too, lets talk about the disadvantages first. The main disadvantage we faced is , image quality where our model took several attempts to analyse the picture quality. Secondly, use of large dataset , were our model took several attempts and pre-processing time was slow, after several attempts. model got adapted to the data. There are advantages too, calories will be provided just by a single picture at very less processing time. This project is invented for the people who are suffering from major obesity problems and for the people who are having lack of knowledge to maintain a healthy diet.

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Obesity has grown significantly over the past ten or so years, globally. The use of meal logging and calorie analysis to combat obesity has been the subject of growing research. Numerous food management programs with picture recognition that automatically record meals have emerged as a result of an increase in healthy lifestyles. It's critical for someone to keep track of their daily caloric intake if they want to live a healthy lifestyle. In order to calculate comparatively precise count of calories from photographs of products, the research attempts to combine modern object detection algorithms with image analysis techniques. The method used here uses mathematical calculation of the feature that are taken from the images of the food items using segmentation to finally compute the calorie-count of the food.

The model developed in this research estimates food calories from photos of food items. We are able to identify food from the various data-sets of Indian cuisine and compute the necessary meal calories by utilizing YOLO You Only Look Once for object identification and Image segmentation for calorie calculation.

Digital classification of food, for instance, informs users about their everyday calorie intake and dietary needs. The goal of this effort is to create a smart-phone application that can capture photographs of meals in real time and analyse them for calories so that individuals can modify their eating habits and live healthier lives.

The project employs YOLO algorithm and image segmentation in this study to identify and estimate the caloric content of food products, respectively. A real-time object detection algorithm is called YOLO (You Only Look Once). It determines how the code will identify the items in the image. It takes one quick look at the image before scanning the network to find the object. To calculate the number of calories in the found food item, the image segmentation method from image processing is applied.

1.2 STUDY

Food detection and calorie estimation are areas of active research in the field of computer vision and machine learning. These technologies aim to automate the process of tracking one's diet and caloric intake, which can be helpful for individuals looking to manage their weight or maintain a healthy lifestyle.

A lot of research papers have focused to develop algorithms for Food and Calorie detection using Computer vision techniques. One of the papers have used a data-set of more than ten thousand food item images for training a Deep Neural Network to identify and then classified it into different types of food items. This Deep Neural Network had reached an accuracy level of more than 80% in correct classification of the respective food item and was also capable in estimating the calories content as well with an error percentage of 10% only.

An app for smart-phones was utilized in a different study to gather pictures of food and calculate the meal's caloric content. To identify the foods in the photograph and calculate their caloric content, the app combined image processing and machine learning techniques. According to the study, the software could predict the meal's calories with an accuracy of between 10% and 20% of the real value.

Despite the promising results of these studies, there are still challenges to overcome in developing accurate and reliable food detection and calorie estimation technologies. Factors such as lighting, portion size, and food composition can all affect the accuracy of these systems, and further research is needed to improve their performance in real-world scenarios.

1.3 PROBLEM STATEMENT

- In the fast-paced world, obesity has become a major health issue to the human beings. BMI defines the obesity when it is greater than 30kg/m. Obesity leads to many diseases like high cholesterol, liver failure, knee problems, diabetes and sometimes cancer. When the patient eats healthy food, the obesity can be controlled.
- The obesity problem can be addressed when there is a system that monitors the food consumed by the patient in treatment of obesity. Many of the people find it difficult in monitoring their food intake periodically, due to less knowledge in nutrition.
- In this chapter, identification of food type is made and estimation of calorie is done by using MLP and proposes the results. Single food items types were considered previously, but here mixed food item types are considered.
- Region of interests is used to identify the mixed food item type. The next step includes feature extraction process. The extracted feature Image is fed into MLP classification to classify the image. The volume of food is used to calculate the calories present in the food. The implementation is done by taking 1000 fruit images containing 6 food classes with good accuracy. The automatic dietary control is made available for the diabetic patients.

1.4 MOTIVATION

- According to current data, 650 million people worldwide are obese, and nearly 1.97 billion are overweight. Obesity-related deaths were predicted to number 2.8 million. Because of the consumption of energy-dense food (i.e., bad eating habits), a sedentary lifestyle, a lack of fitness care facilities, and monetary constraints, the developing world faces a huge risk of weight problems and their negative consequences (i.e. diabetes, ischemic coronary heart disorder, etc.). Weight problems affect more than 135 million people in India.
- In India, the occurrence of weight problems varies by age, gender, location (geographic and socioeconomic), and so on.
- The incidence rate of central weight disorders ranges from 11.8% to 31.3% and 16.9% to 36.3%, respectively, according to the 2015 Indian Council of Medical Research. In India, issues with belly obesity pose a significant risk for cardiovascular disease. Women are more likely than men to be obese, according to numerous research. One of the most expensive scientific and financial costs for the government is weight-related disorders. By raising public knowledge of weight issues and how they affect fitness, they can be prevented. Governmental bodies ought to advocate for the benefits of a healthy lifestyle, which includes good diet and exercise practises.
- The purpose of the above words is to explain the superiority of weight problems in various Indian areas and to pique readers' interest in the difficult areas.

CHAPTER 2

LITERATURE SURVEY

A method of estimating calories for dietary assessment using image analysis had been described by Tatsuya Miyazaki, Gamhewage C[1] de Silva, and Kiyoharu Aizawa [2]. They created a dataset with 6000 photos that was made up of food log calorie counts calculated by the nutritionist. Multiple alternatives, including Correlograms, colour Histograms, and SURF options, were used to compare the food image to the food log dataset.

A food calorie and nutrition measurement method that assisted patients and dietitians in controlling daily food consumption was provided by Parisa Pouladzadeh [3], Shervin Shirmohammadi [4], and Rana Al-Maghrabi [5]. According to the report, a smartphone application that measures calorie and nutritional content by taking photos of food before and after consumption would be developed.

YOLO, a unified paradigm for object detection, had been introduced by Joseph Redmon[6], Santosh Divvala[7], Ross Girshick[8], and Ali Farhadi[9]. This model may be built quickly and trained using only full photos. YOLO is trained using a loss function that directly relates to detection performance, in contrast to classifier-based techniques, and as a result the complete model is trained at once.

Manal Chokr [10] and Shady Elbassuoni [11] proposed the easiest approach to estimate the calorie content of the food item by estimating the size of the item by passing the compressed image through a regressor (directed towards eatables like pizzas, doughnuts, chicken, and sandwiches).

YOLOv2 and YOLO9000, real-time detection systems, were introduced by Joseph Redmon[12] and Ali Farhadi[13]. Comparatively speaking, YOLOv2 is quite quick. A real-time framework

called YOLO9000 allows for the classification and detection of over 9000 different item categories.

Sujata Chaudhari [14], Nisha Malkan [15], Ayesha Momin [16], and Mohan Bonde [17] all used YOLO to quickly find things. In this essay, YOLO will be addressed as a modern strategy.

for accurate object detection with a high accuracy. This study explains how accuracy increased during the training procedure. They have discussed the development of CNN in this work, and object detection is carried out with the use of YOLO and CNN.

The major goal of this study by Alexey Bochkovskiy [18], Chien-Yao Wang [19], and Hong-Yuan Mark Liao [20] is to construct an object detector that operates quickly in production systems while maximising for parallel calculations. An effective and potent object detection model has been created by them. Modern methods are improved to be more efficient for training on a single GPU. Numerous features have been confirmed to increase the accuracy of the classifier and detector.

CHAPTER 3

PROPOSED SYSTEM MODEL

3.1 METHODOLOGY

The dataset obtained consists of infinite food images and also the images are belonged to various classes such as Apple , Banana , Orange , Pizza , etc. Our model will be trained in such a way that, it can calibrate the object , the images placed near the webcam.

Image cropping , resizing is the main agenda of our model and ensures that unnecessary parts of the images will be deleted.The graphical image annotation tool i.e. a tool name labelling is used for creating boundary box or rectangular box around the object in the image. The format is pascal voc.

Yolo is a algorithm used for object detection. Yolo's agenda is to used a single neural network to the image, later it classify the images into several regions and predicts that the bounding box applies for each possible region. Therby our model is capable for selecting multiple objects in a single image and able to predict multiple bounding boxes and class probabilities for those boxes.

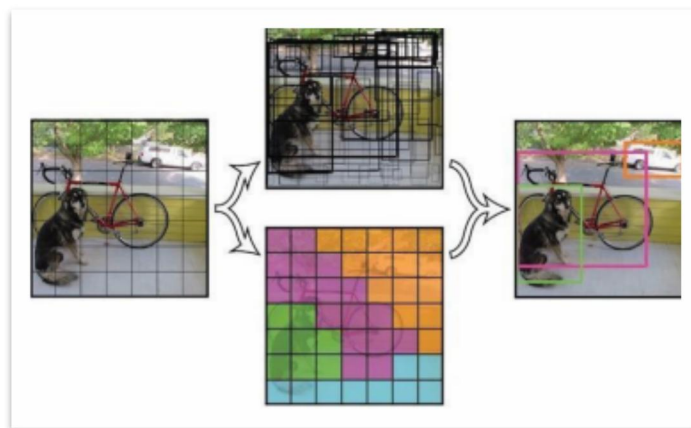


Figure 3.1: System Architecture

3.2 MACHINE LEARNING TECHNIQUES

Convolutional neural networks (CNN) are often used in YOLO set of rules to detect gadgets in real time. To discover gadgets, the set of rules requires only a single forward propagation via a neural network, as the name suggests. This method predicts the entire picture in a single set of rules. The CNN is used to predict a large number of classes.

3.2.1 Intersection Over Union (IOU)

Intersection over union is applicable for overlapping of boxes. The measure to calculate the accuracy of the object detector is to apply IOU , the following parameters are required.

- The ground truth bounding boxes
- The predicted bounding boxes

$$\text{IOU} = \text{Area of overlap} / \text{Area of Union}$$



Figure 3.2: Example of computing IOU for various bounding boxes

3.2.2 YOLO V3 NETWORK

The YOLOv4 network consists of three parts :

a) Backbone

The feature extraction architecture is referred to as the backbone. It is a convolution neural network (CNN) that mixes and creates picture features at various granularities. The backbone distinguishes the many YOLO iterations. CSPDarknet53 is utilized in YOLOv4. Cross-Stage-Partial connections are known as CSP.

b) Neck

The dense prediction block's neck block serves as an additional layer between the backbone and the head. The YOLOv4 employs updated Spatial Pyramid Pooling (SPP), modified Spatial Attention Module (SAM), and modified Pan Aggregation Network (PAN) [10] to add additional information in a layer. In order to transfer image features forward for prediction, a succession of layers must be used.

b) Head

Bounding boxes are located using the head block, which is also used to categorise the objects in the image. It evaluates the neck block's features, finds the bounding box, then sorts it into distinct classifications.

3.2.3 Convolutional Neural Network

An image's objects can be found and classified using the Convolutional Neural Network (CNN), a feed-forward neural network. It also goes by the name ConvNet. An input, an output, and several hidden layers make up a convolutional neural network. Convolutional, pooling, fully connected, and normalisation layers make up a CNN's hidden layers. RELU, which is an activation function, is also one of them.

In CNN, each image is displayed as an array of pixel values. It is a network that can take an image as input, give distinct items in the picture weights, and distinguish between things in the image. Compared to other classification methods, CNN requires substantially less pre-processing. Convolutional neural networks are tasked with converting images into a format that is simpler to comprehend while retaining the elements required for making accurate predictions.

Fully Connected Layer

Bounding boxes are located using the head block, which is also used to categorise the objects in the image. The Fully Connected (FC) Layer is made up of the weights, biases, and neurons that link the neurons of two separate layers together.

The final few layers of the CNN design are these layers, which are positioned before the output layer. The input images from the preceding layers are flattened and supplied to the Fully Connected layer in this layer.

The flattened vector is then put through more Fully Connected layers, which is where mathematical operations are typically performed. This is the stage where the classifying process starts.

3.3 ALGORITHMS USED

3.3.1 Yolo v1

YOLO (You Only Look Once) is the set of guidelines that we've in use in our project , because it is able to anticipate devices in real-time, this set of rules increases the rate of detection superior precision.

YOLO is a predictive strategy that offers precise result with little in the way of background error. acquiring skills Due to the rules' excellent learning capabilities, it is possible for them to analyse the models of devices and use them in item detection.

3.3.2 Yolo v3

It is a supervised learning technique that performs object detection in real time. CNN are classifier based systems that can process input images as structured arrays of data and recognized based systems and recognize patterns between them. It allows model to look at whole image at at time. First, it separate images into a grid, predict some boundary boxes, each boxes has confident score of how accurate it assumes that prediction should be and detects only one per bounding box. The bounding boxes are generated by clustering the dimensions of the ground truth boxes from the original datasets to find the most common shapes and sizes.

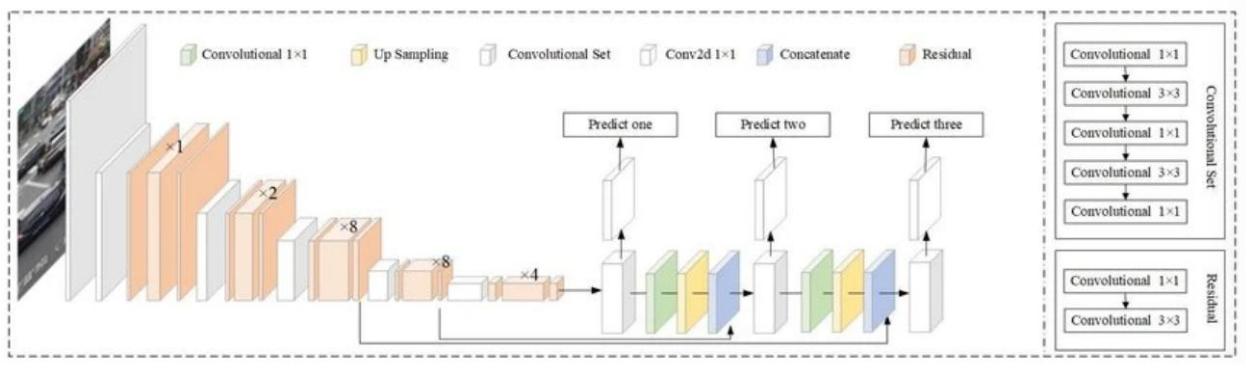


Figure 3.3: Yolo v3 Structure

3.3.3 Yolo V5

The additional network layers YOLOv5 can train more classes in less time than YOLOv3 and produces better outcomes.

This is an advantage of training time over YOLOv3. Transfer Learning is a notion that is used in YOLOv3. Reusing a model that has already been trained on a different problem is known as transfer learning in machine learning.

With transfer learning, a computer can use its understanding of one activity to better generalize about another. Making use of several libraries was made easier by using the Darknet. The primary contribution of YOLOv5 is the conversion of the PyTorch framework for Darknet research.

3.3.4 Working of Yolo v3 and Yolo v5 Algorithms

The detection task is treated by YOLO as a regression drawback, and it is widely used in image processing areas. There have also recently been plans for numerous versions of YOLO, including YOLOv1, YOLOv2, YOLOv3, YOLOv4, and YOLOv5. The following benefits of YOLOv3 over YOLOv1 and YOLOv2: enhanced classification efficiency on challenging datasets. the expanded amount of information in the feature map.

The additional network layers YOLOv5 can train more classes in less time than YOLOv3 and produces better outcomes.

This is an advantage of training time over YOLOv3. Transfer Learning is a notion that is used in YOLOv3. Reusing a model that has already been trained on a different problem is known as transfer learning in machine learning.

With transfer learning, a computer can use its understanding of one activity to better generalize about another. Making use of several libraries was made easier by using the Darknet. The primary contribution of YOLOv5 is the conversion of the PyTorch framework for Darknet research.

The Darknet framework, which is mostly developed in C, gives great control over how the network is operated through coding. Low-level language control is useful for research in many ways, however because one must create specific gradient statistics for each new compilation, it can slow down the submission of data for new research.

No small achievement was accomplished in transferring (and passing) the training processes from Darknet to PyTorch into YOLOv3

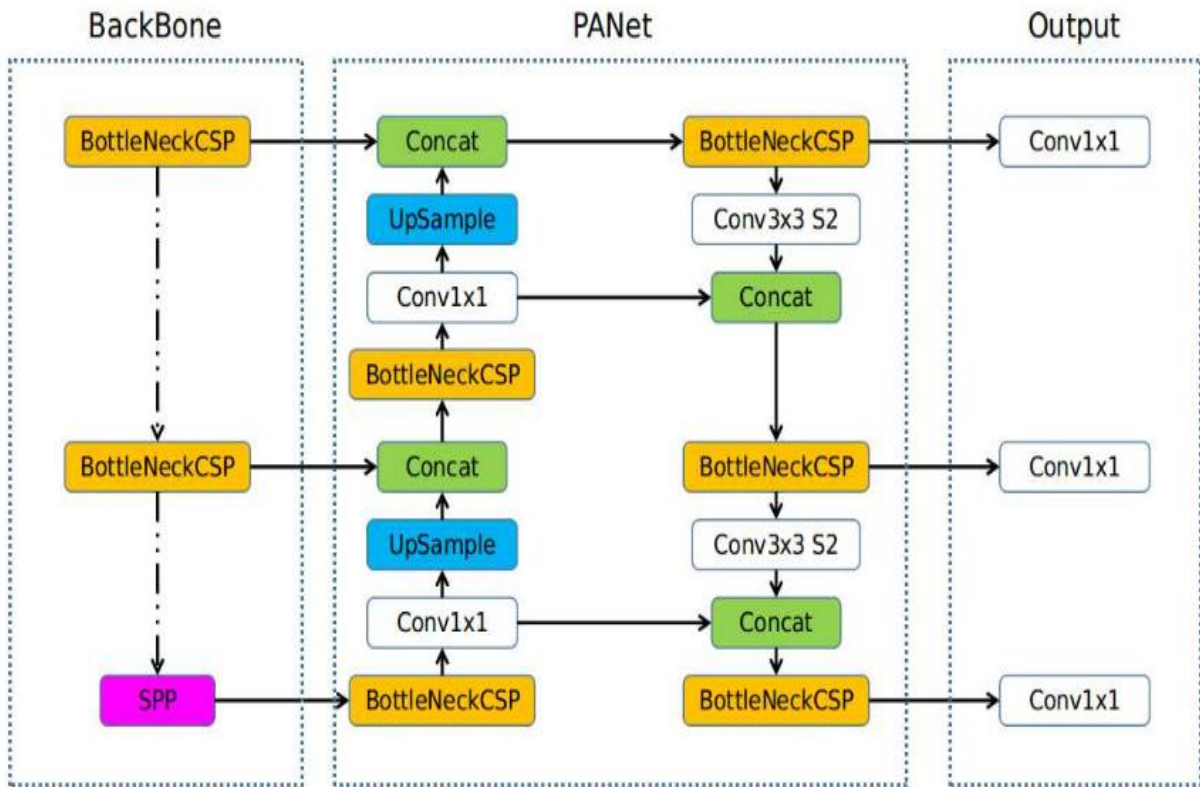


Figure 3.4: YOLO V5 Algorithm

CHAPTER 4

PROBLEM ANALYSIS

4.1 Product Definition

In this modern era, there are different technologies that are emerge to solve various issues regarding our needs. Food classification system field is rapidly changing with rise in new technologies like AI's, ML but our main problem is to show the calories of the food people are in-taking.

Today's world many people might not be able to track their daily calorie intake and because of this they might consume unhealthy amount of calories and this may ultimately lead to obesity .

4.2 Feasibility Analysis

A feasibility study is the evaluation of the practicality of a proposed project .The feasibility study aims to find the strength and limitations of s proposed and the risks associated with the project , the resources required for the completion of the project and the prospects for success.

4.3 Technical Feasibility

Since the project involves deep learning technology, we have worked on Google-collab as well as in Kaggle using python language. In the case of modules, I have used Tensor Flow, Keras, stream lite, NumPy, and pillow. In case of local host, I used Stream-lit. For getting the data-set, we found it from different websites like Kaggle, images online.

4.4 Economic Feasibility

As the whole project is dependent on the deep learning model CNN and python there is no money spent. The data-sets are directly available for free on the internet. The python platform I used for running the model, downloaded directly from net. We trained the model in Google collab.

4.5 Operational Feasibility

With the help of this device the user will be able to use the interface without any difficulty cause as interface is made keeping in mind that ease to use is one of our main priorities and all the required modules are presented in such a way that whatever action user chooses to do, it will always be one click away.

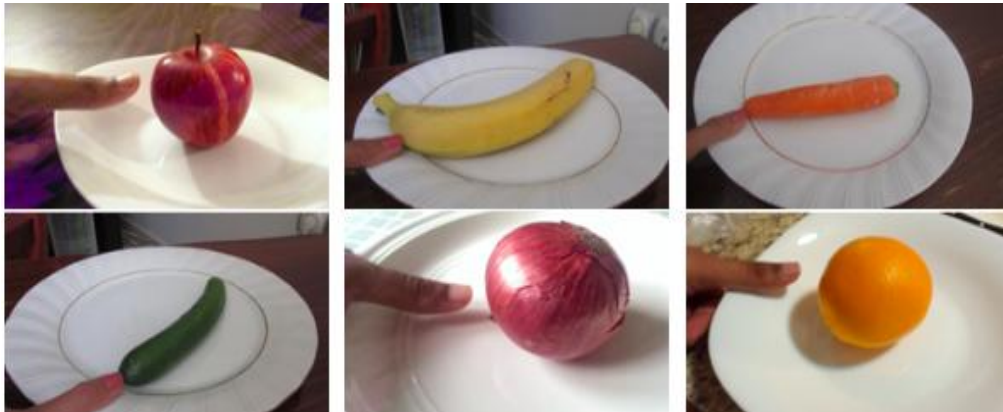


Figure 4.1: Object detection and Boundary boxing

CHAPTER 5

SOFTWARE REQUIREMENT ANALYSIS

5.1 Introduction

This document describes the software requirements and specification for Fruit analyser system. The system is designed to make sure that calories consumed by the user is appropriate or not. In the final output we get to know that, he/she has to follow the particular diet chart given by the system.

PyCharm: PyCharm is an integrated development environment used in computer programming, specifically for the Python programming language. Here PyCharm edition 2021.2.3 edition used for running the model in the local environment. Kaggle: Kaggle is a website where a variety of different datasets of different projects are widely available. Our data set is downloaded from kaggle-"Fruits-360". The images downloaded from Kaggle are-Apple, Banana, and orange. Google colab: Google colab is a software where one can upload and run one programs. It is developed by google. This is used by us to train the model.

5.3.1 Libraries

TensorFlow

TensorFlow is an open-source programming for mathematical calculation. It was initially made to lead ML and deep neural network research

Keras

Keras is an open-source library written in Python, that's used for image pre-processing, modelling, evaluating, and optimization. It can run on the top of TensorFlow library.

It is likewise utilized for making significant level Application Programming Interface as it is taken care of by the back-end. It is intended for making a version with loss and optimizer function, and training manner with fit characteristic. For the back-end, it's miles intended for convolution and low-degree calculation underneath tensors or TensorFlow. Bringing in the python libraries are utilized for pre-processing, optimization, enhancement, and testing.

Stream Lit:

Streamlit is an open-source framework in programming language Python. It allows us create internet apps for technologies like machine learning and data science in a quick time. It is well matched with foremost important Python libraries such as scikit-learn, Keras. So, for creating the local host it is used.

NumPy

NumPy is a library in programming language Python, providing support for massive, multi-dimensional arrays and matrices, together with a huge series of high-stage mathematical function on those arrays. It is used for storing the datasets within the form of arrays.

OpenCV

OpenCV is an extraordinary tool for image processing and to perform computer vision related tasks. It is an open-source library that may be used to carry out tasks like face detection, to track objects and landmark detection.

5.2 Functional Requirements

Local host

Description: To upload the image to be analysed Input: User defined functionality

Processing: When the code execution is started, we get a local host Ip address, where we upload the image and get the output.

Output: Image is uploaded (Image name)

Description: Need to display the name of the underline image uploaded

Input: Image uploaded by the user or from the dataset provided for training

Processing: When an image gets uploaded the food description is visible on the browser screen

Output: Display the food calories

- **Calories of the Food item with respect to the image**

Description: Recognize the food item from the image uploaded.

Input: Image uploaded by the user by clicking a button with embedded functionality.

Processing: Model automatically runs and print the output on the screen.

Output: Display the calories for the food item.

- **Nutrition**

Description: Get the nutrition value of the fod image in calories. Input: Image uploaded by the user for processing.

Processing: To be provided with the nutritional component measure. Output: To be provide with the nutritional component measure.

- **Report bug**

Description: If a bug is found by the user he can report it.

Input: If any error is found in the web Processing: Inbuilt option when using streamlit Output:

Can report a Bug

5.3 Non functional Requirements

The system is web-based as well as mobile-based hence requires a web server and database.

It must be able to perform in adverse conditions with high and low speed of the internet.

Uninterrupted interrupted connections.

The website and application should load within 4 seconds

5.4 Safety Requirements

It should make user that uploaded one should be image of the food or any other dish. Must make sure the image uploaded should be in 255*255-pixel size. If not, the model could not give the correct output

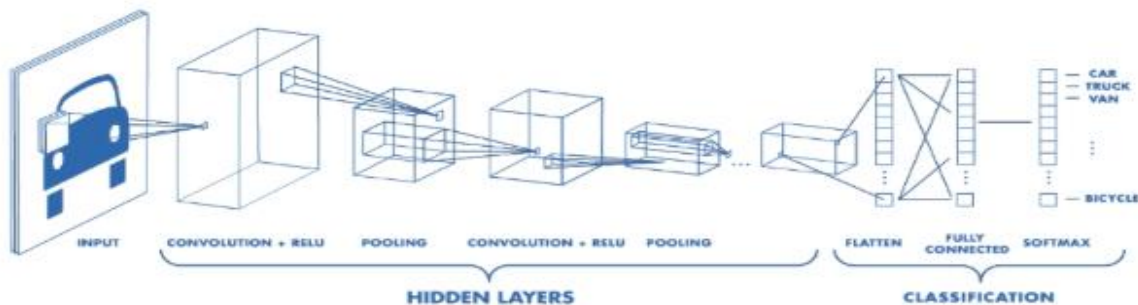


Figure 5.1: Functional Requirements

CHAPTER 6

6.1 System Design

6.1.1 Flow chart

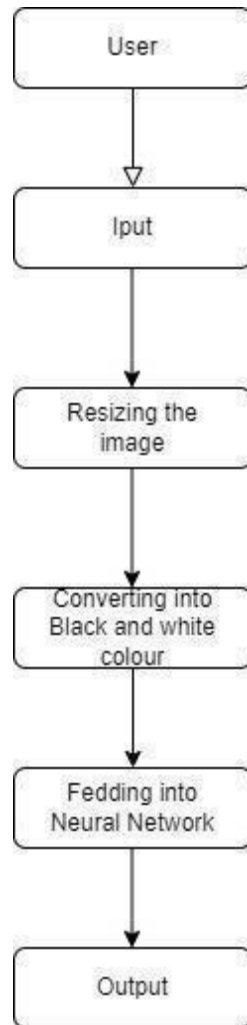


Figure 6.1 Flow chart

6.2 Detailed design

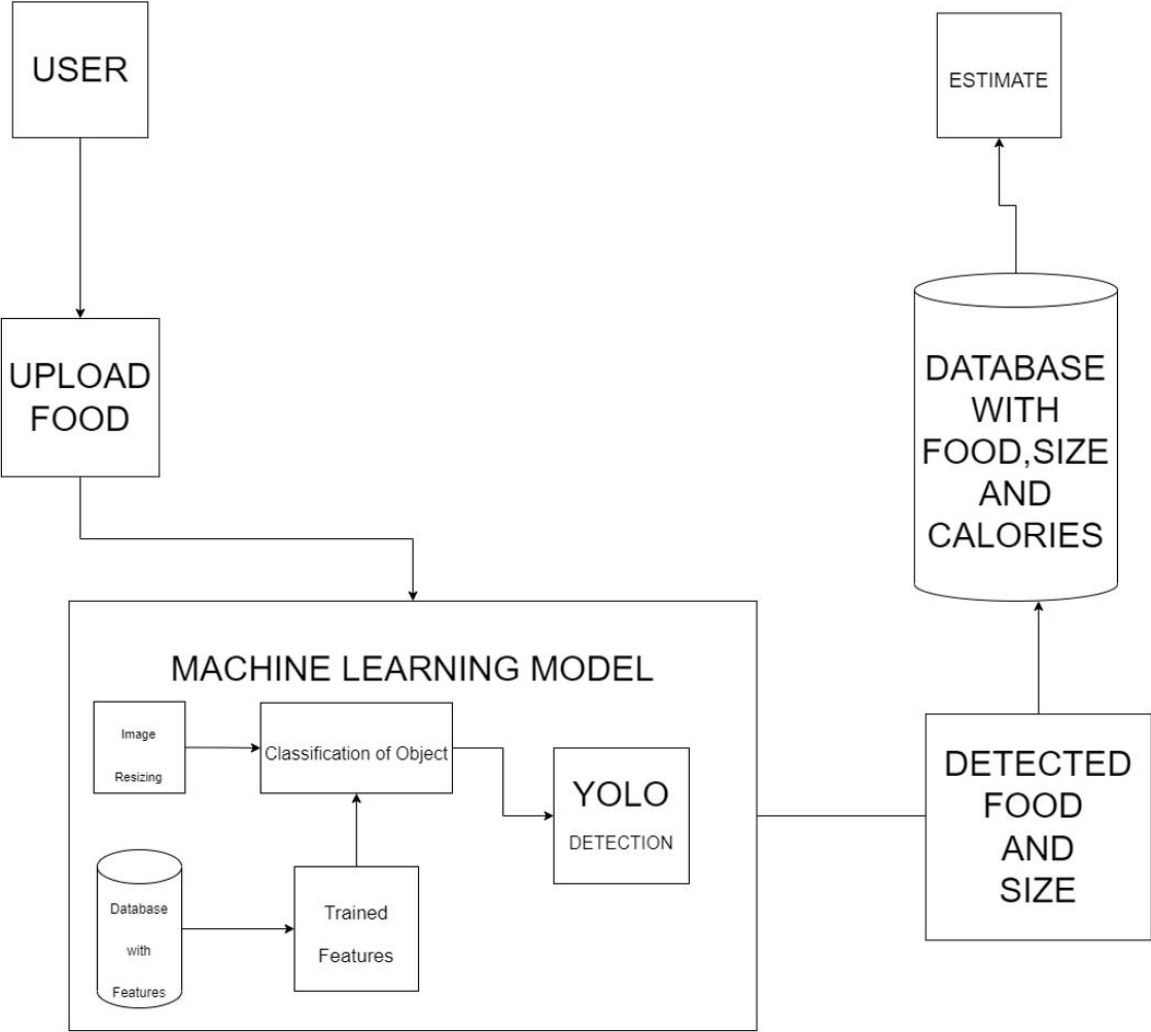


Figure 6.2: Detailed Algorithm

CHAPTER 7

YOLOv3 IMPLEMENTATION

We used our own data-set FINAL 7 for the detection of the food item. For training the model we used an algorithm to randomly divide our data-set into 70% into training and 30% rest data into testing. Then ,the respective data-set is used for training and testing . Also it supports dark-net framework.

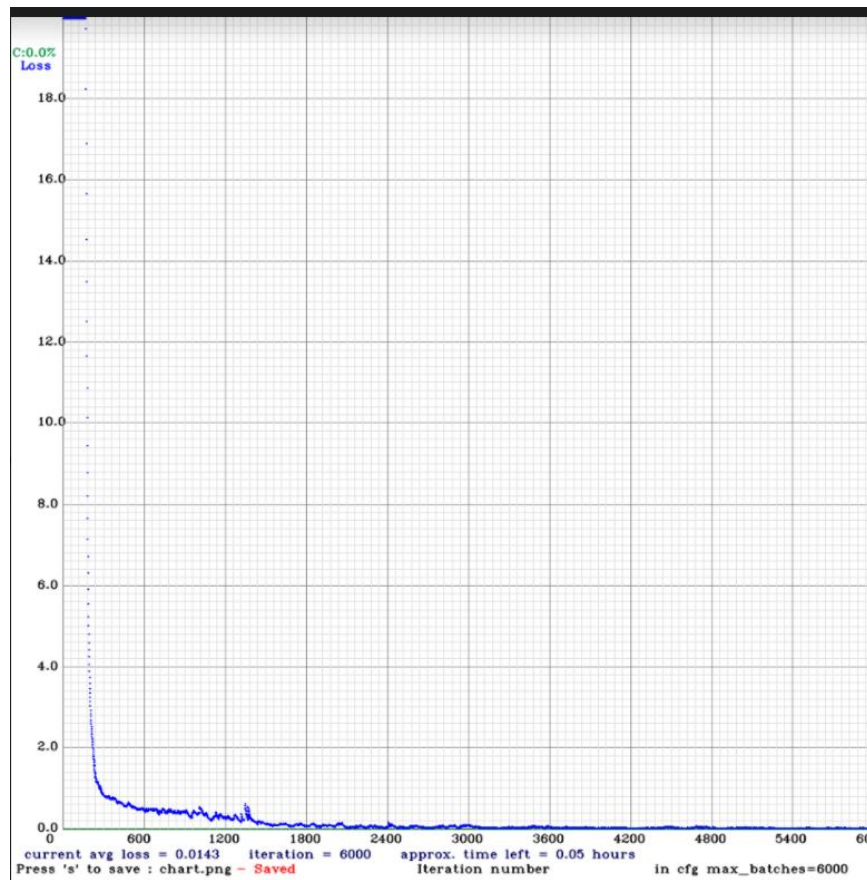
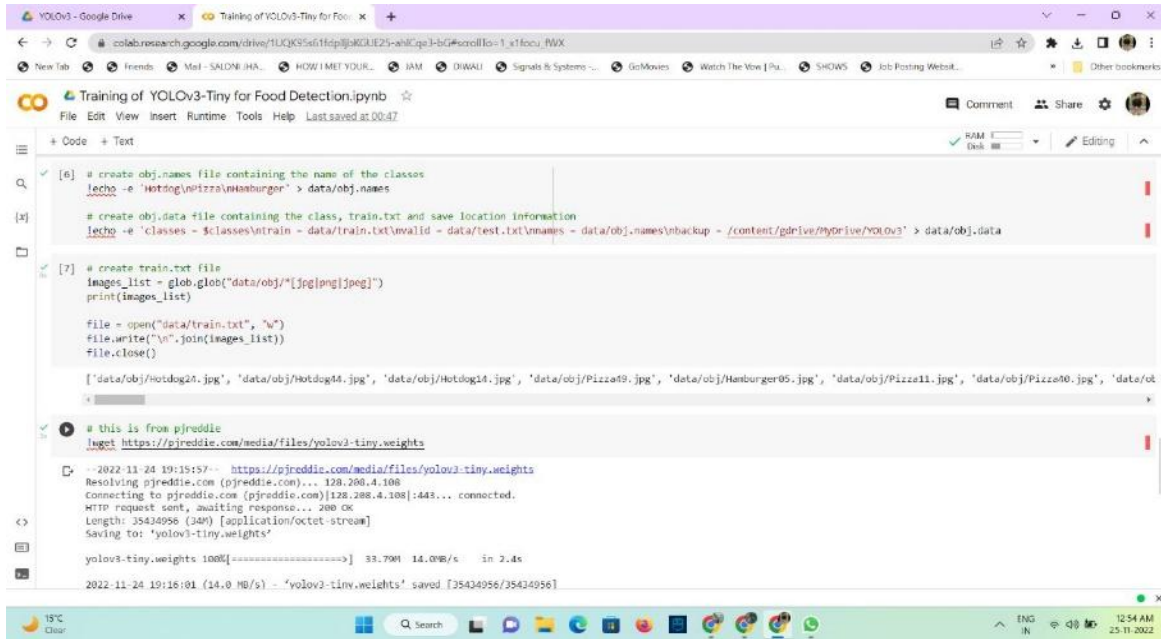


Figure 7.1: Yolov3 Graph

7.1 Execution:



```
[6] # create obj.names file containing the name of the classes
!echo -e 'Hotdog\Pizza\Hamburger' > data/obj.names

# create obj.data file containing the class, train.txt and save location information
!echo -e 'classes = %classes\ntrain = data/train.txt\nvalid = data/test.txt\nnames = data/obj.names\nbackup = /content/gdrive/MyDrive/YOLOv3' > data/obj.data

[7] # create train.txt file
images_list = glob.glob("data/obj/*[jpg|png|jpeg]")
print(images_list)

file = open("data/train.txt", "w")
file.write("\n".join(images_list))
file.close()

['data/obj/Hotdog01.jpg', 'data/obj/Hotdog04.jpg', 'data/obj/Hotdog14.jpg', 'data/obj/Pizza09.jpg', 'data/obj/Hamburger05.jpg', 'data/obj/Pizza11.jpg', 'data/obj/Pizza00.jpg', 'data/obj/Hotdog02.jpg', 'data/obj/Hotdog03.jpg', 'data/obj/Hotdog05.jpg', 'data/obj/Hotdog06.jpg', 'data/obj/Hotdog07.jpg', 'data/obj/Hotdog08.jpg', 'data/obj/Hotdog10.jpg', 'data/obj/Hotdog11.jpg', 'data/obj/Hotdog12.jpg', 'data/obj/Hotdog13.jpg', 'data/obj/Hotdog15.jpg', 'data/obj/Hotdog16.jpg', 'data/obj/Hotdog17.jpg', 'data/obj/Hotdog18.jpg', 'data/obj/Hotdog19.jpg', 'data/obj/Hotdog20.jpg', 'data/obj/Hotdog21.jpg', 'data/obj/Hotdog22.jpg', 'data/obj/Hotdog23.jpg', 'data/obj/Hotdog24.jpg', 'data/obj/Hotdog25.jpg', 'data/obj/Hotdog26.jpg', 'data/obj/Hotdog27.jpg', 'data/obj/Hotdog28.jpg', 'data/obj/Hotdog29.jpg', 'data/obj/Hotdog30.jpg', 'data/obj/Hotdog31.jpg', 'data/obj/Hotdog32.jpg', 'data/obj/Hotdog33.jpg', 'data/obj/Hotdog34.jpg', 'data/obj/Hotdog35.jpg', 'data/obj/Hotdog36.jpg', 'data/obj/Hotdog37.jpg', 'data/obj/Hotdog38.jpg', 'data/obj/Hotdog39.jpg', 'data/obj/Hotdog40.jpg', 'data/obj/Hotdog41.jpg', 'data/obj/Hotdog42.jpg', 'data/obj/Hotdog43.jpg', 'data/obj/Hotdog44.jpg', 'data/obj/Hotdog45.jpg', 'data/obj/Hotdog46.jpg', 'data/obj/Hotdog47.jpg', 'data/obj/Hotdog48.jpg', 'data/obj/Hotdog49.jpg', 'data/obj/Hotdog50.jpg', 'data/obj/Hotdog51.jpg', 'data/obj/Hotdog52.jpg', 'data/obj/Hotdog53.jpg', 'data/obj/Hotdog54.jpg', 'data/obj/Hotdog55.jpg', 'data/obj/Hotdog56.jpg', 'data/obj/Hotdog57.jpg', 'data/obj/Hotdog58.jpg', 'data/obj/Hotdog59.jpg', 'data/obj/Hotdog60.jpg', 'data/obj/Hotdog61.jpg', 'data/obj/Hotdog62.jpg', 'data/obj/Hotdog63.jpg', 'data/obj/Hotdog64.jpg', 'data/obj/Hotdog65.jpg', 'data/obj/Hotdog66.jpg', 'data/obj/Hotdog67.jpg', 'data/obj/Hotdog68.jpg', 'data/obj/Hotdog69.jpg', 'data/obj/Hotdog70.jpg', 'data/obj/Hotdog71.jpg', 'data/obj/Hotdog72.jpg', 'data/obj/Hotdog73.jpg', 'data/obj/Hotdog74.jpg', 'data/obj/Hotdog75.jpg', 'data/obj/Hotdog76.jpg', 'data/obj/Hotdog77.jpg', 'data/obj/Hotdog78.jpg', 'data/obj/Hotdog79.jpg', 'data/obj/Hotdog80.jpg', 'data/obj/Hotdog81.jpg', 'data/obj/Hotdog82.jpg', 'data/obj/Hotdog83.jpg', 'data/obj/Hotdog84.jpg', 'data/obj/Hotdog85.jpg', 'data/obj/Hotdog86.jpg', 'data/obj/Hotdog87.jpg', 'data/obj/Hotdog88.jpg', 'data/obj/Hotdog89.jpg', 'data/obj/Hotdog90.jpg', 'data/obj/Hotdog91.jpg', 'data/obj/Hotdog92.jpg', 'data/obj/Hotdog93.jpg', 'data/obj/Hotdog94.jpg', 'data/obj/Hotdog95.jpg', 'data/obj/Hotdog96.jpg', 'data/obj/Hotdog97.jpg', 'data/obj/Hotdog98.jpg', 'data/obj/Hotdog99.jpg']

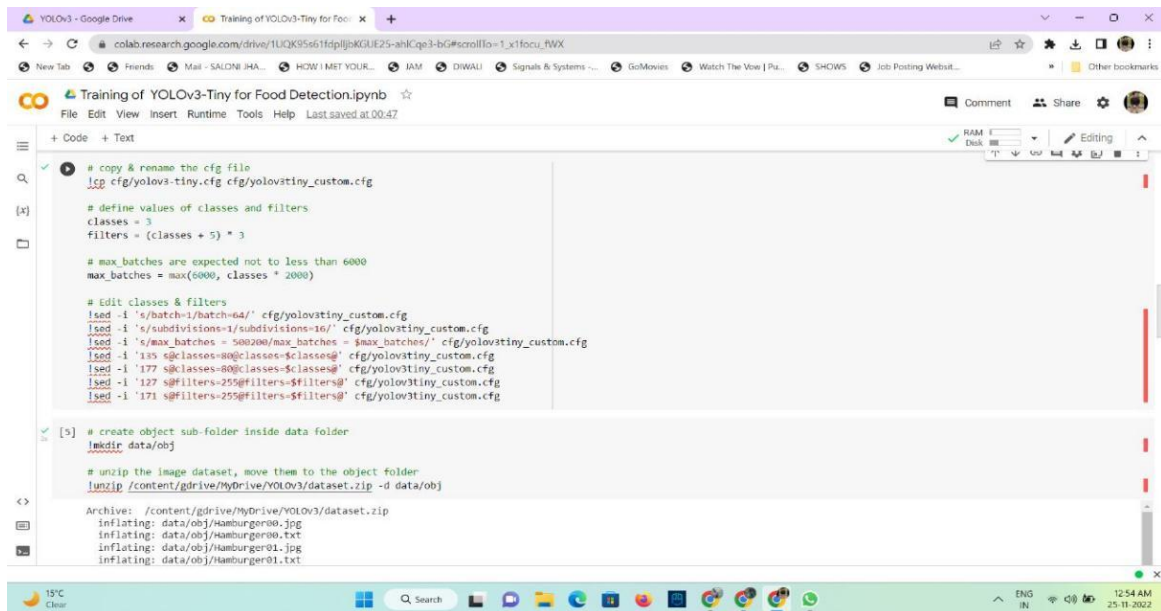
# this is from pjreddie
!wget https://pjreddie.com/media/files/yolov3-tiny.weights

--2022-11-24 19:18:57-- https://pjreddie.com/media/files/yolov3-tiny.weights
Resolving pjreddie.com (pjreddie.com)... 128.208.4.108
Connecting to pjreddie.com (pjreddie.com)[128.208.4.108]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 35424956 (34M) [application/octet-stream]
Saving to: 'yolov3-tiny.weights'

yolov3-tiny.weights 100%[=====] 33.79M 14.0MB/s in 2.4s

2022-11-24 19:18:01 (14.0 MB/s) - 'yolov3-tiny.weights' saved [35424956/35424956]
```

Figure 7.2: Execution Process



```
[5] # copy & rename the cfg file
!cp cfg/yolov3-tiny.cfg cfg/yolov3tiny_custom.cfg

# define values of classes and filters
classes = 3
filters = (classes + 5) * 3

# max_batches are expected not to less than 6000
max_batches = max(6000, classes * 2000)

# Edit classes & filters
!sed -i 's/batch=1/batch=64/' cfg/yolov3tiny_custom.cfg
!sed -i 's/subdivisions=1/subdivisions=16/' cfg/yolov3tiny_custom.cfg
!sed -i 's/max_batches = 50000/max_batches = $max_batches/' cfg/yolov3tiny_custom.cfg
!sed -i 's/135 s@classes=80@classes-$classes/' cfg/yolov3tiny_custom.cfg
!sed -i '177 s@classes=80@classes-$classes/' cfg/yolov3tiny_custom.cfg
!sed -i '127 s@filters=25@filters-$filters/' cfg/yolov3tiny_custom.cfg
!sed -i '171 s@filters=25@filters-$filters/' cfg/yolov3tiny_custom.cfg

# create object sub-folder inside data folder
!mkdir data/obj

# unzip the image dataset, move them to the object folder
!unzip /content/gdrive/MyDrive/YOLOv3/dataset.zip -d data/obj

Archive: /content/gdrive/MyDrive/YOLOv3/dataset.zip
  inflating: data/obj/hamburger00.jpg
  inflating: data/obj/hamburger00.txt
  inflating: data/obj/hamburger01.jpg
  inflating: data/obj/hamburger01.txt
```

Figure 7.3: Execution Continued

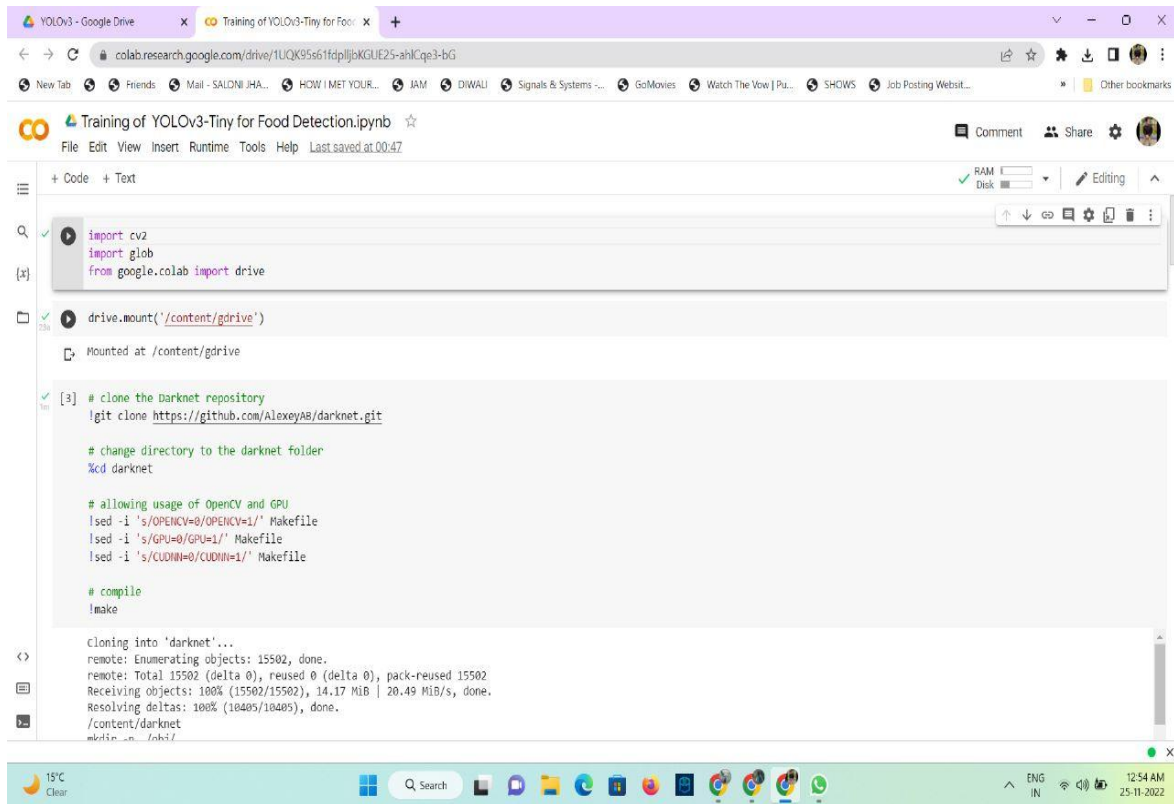


Figure 7.4: Execution

7.2 RESULT:

As we can see, YOLOv3 properly spotted the apple, but the time required by this model was quite high, and the classes trained were very few. Thus, YOLOv3 has more accuracy than its predecessors, but it is very slow while training data. The primary contribution of YOLOv5 is the conversion of the Pytorch framework for Darknet research. The Darknet framework is written mostly in C and allows excellent control over the network's coded functioning. Low-level language control is advantageous for YoloV3 and Yolov5. In one way, it the submission of data and ensure that it requires enough space and rest will be saved for further purpose. The main purpose of this is that one must create specific gradient statistics for each new compilation.

Apple , banana, guava, bread, pomegranate and pizza are the 6 food classes which are considered here. SVM and MLP based classifiers are used here and the results are compared. Based on the precision, recall, F-measure , the efficiency and accuracy of SVM and MLP methods are analyzed and the results are tabulated.

Accuracy represents how close the measurements are to specific value. Whereas, precision represents how close the measurements values are related to each other

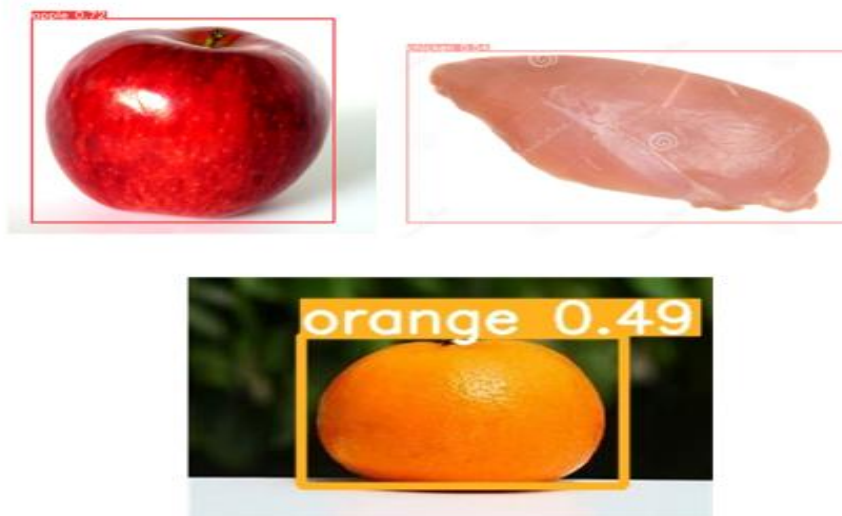


Figure 7.5: Image Size and Re-sizing

CHAPTER 8

CONCLUSION

We trained our data using the YOLOv3 and YOLOv4 versions of YOLO [YOU ONLY LOOK ONCE] algorithm.

In addition to that our model will enhance the displayed image quality as well as the average calorie content for 1gm of each eatable to calculate the calories to a greater extent. A precise calorie amount of the food items is decided using the input weight of the food items with respect to a standard calorie value.

Although YOLOv3 shown respectable accuracy, it took a very long time to train the data for numerous classes. In contrast, YOLOv5 demonstrated respectable accuracy and trained the data for more classes in a shorter length of time than YOLOv3. Additionally, YOLOv3 is compatible with darknet, making it much simpler to apply custom weights with various libraries like open-CV, but YOLOv5 is incompatible because it generates py-Torch files.

When it comes to picture categorization, the conventional machine learning algorithms are not very accurate in practice. The CNN model is used in this research to recognize images. A significant amount of sorting and classifying is done perfectly . CNN does not require clean pixel values because it naturally learns the generalised pattern needed to identify and recognize new images. As a result, the accuracy of the CNN model is significantly higher than that of all other conventional models.

CHALLENGES

Identifying the food item from a single image.

Foods of similar types, such as apples and oranges, have similar shapes that make them challenging to distinguish.

Because food images require a significantly larger dataset and longer training time, we presently use a finite dataset for training.

Making the most use of our own dataset

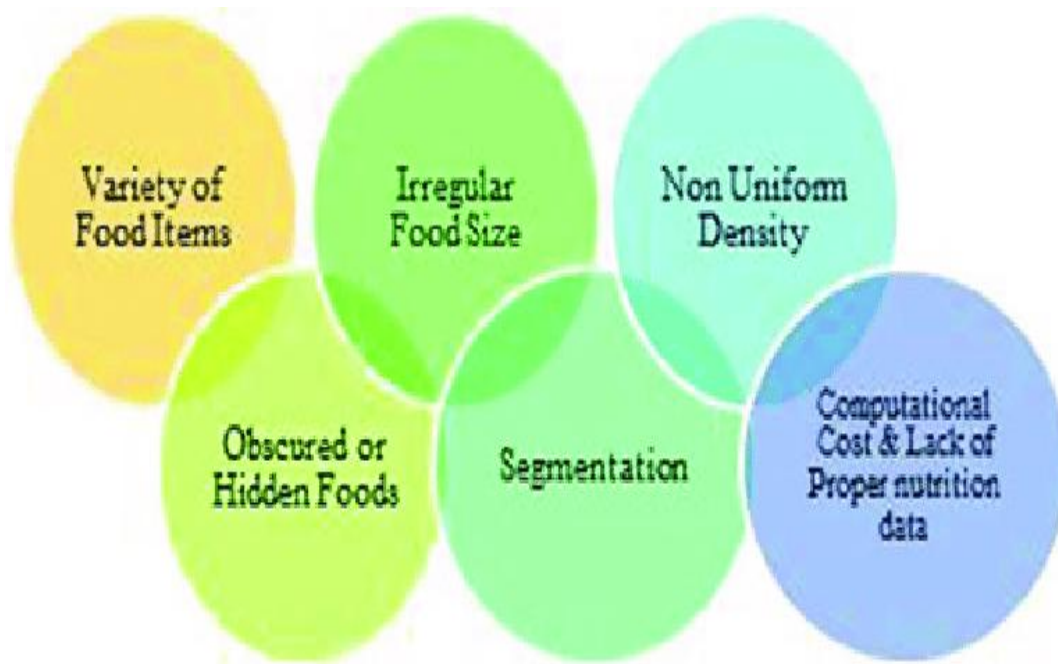


Figure 8.1: Flow Chart Of The Challenges

BENEFITS

- Precise and accurate recognition of food items.
- Helps in keeping track of dietary information.
- The typical goal of calorie restriction is weight loss.
- In humans, it has been discovered that reducing caloric consumption by 20–30% while still meeting nutritional needs can treat age-related disorders such Blood-vascular disease, Rapid increase in sugar levels, rapid increase of thyroid levels. In order to fight among these all diseases, one should perform a regular vigorous physical activities like swimming, running and going into calorie deficient, healthy foods, which will prevent reducing the obesity levels.



Figure 8.2: Benefits of having a good diet

CHAPTER 9

FUTURE SCOPE

The model can be expanded to include more categories, which will enable the classifier to predict the characteristics of a wider range of food products. Recognizes several food items that can be included in a single meal, such as cakes and hamburgers on the same plate, making it easier to determine the meal's calorie content. For more user convenience, the online application can be made into a mobile application. In the future, other food categories could be trained. It is possible to see multiple layers in a single food product. Volume may be used to determine calories.

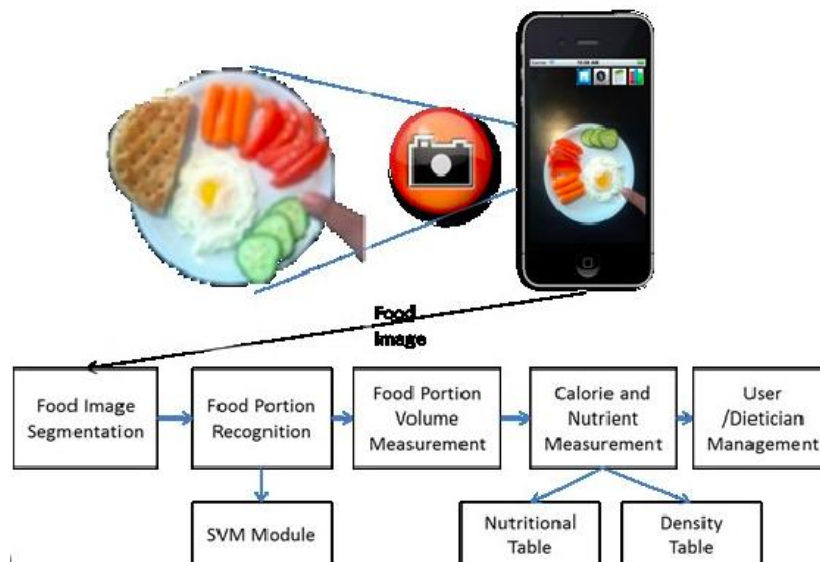


Figure 9.1: Benefits and future scope

References:

- Towards Data Science (2018) R-CNN, Fast R CNN, Faster R-CNN, YOLO Object Detection Algorithm.
- [1]Medium. (2018). (Part 1) Generating Anchor boxes for Yolo-like network for vehicle detection using KITTI dataset.
- [2]Medium. (2018). Real-time Object Detection with YOLO, YOLOV2 and now YOLOV3.
- [3]<https://medium.com/analytics-vidhya/object-detection-using-yolov3-d48100de2ebb>.
- [4] <https://github.com/ultralytics/yolov5>.
- [5] <https://docs.ultralytics.com/>
- [6] <https://.org/abs/1804027671v1>
- [7]<https://papers.readthedocs.io/en/latest/imagedetection/yolov3/>
- [8] S. Gidaris and N. Komodakis. Object detection via a multi-region & semantic segmentation aware CNN mode CORR, abs/ 1505.01749, 2015.7.
- (12) Chandhan G , Ayush Jain Harsh Jain and Mohana Real Time Object Detection and tracking Using Deep Learning And Opencv.
- [13] Joseph Redman, Ali Larhus YOLOV An Incremental improvement University of Washington.